Biosignal Authentication Considered Harmful Today

Veena Krish, *Stony Brook University;* Nicola Paoletti and Milad Kazemi, *King's College London;* Scott Smolka and Amir Rahmati, *Stony Brook University*

https://www.usenix.org/conference/usenixsecurity24/presentation/krish

This paper is included in the Proceedings of the
33rd USENIX Security Symposium.

August 14–16, 2024 • Philadelphia, PA, USA

978-1-939133-44-1

Open access to the Proceedings of the
33rd USENIX Security Symposium
is sponsored by USENIX.

Biosignal Authentication Considered Harmful Today

Veena Krish
Stony Brook University

Nicola Paoletti
King's College London

Milad Kazemi
King's College London

Scott Smolka
Stony Brook University

Amir Rahmati
Stony Brook University

Abstract

User authentication systems based on cardiovascular biosignals have gained prominence in recent years, as these signals are presumed to be difficult to forge. We challenge this assumption by showing that an observer who has access to one type of cardiac data – such as a user's pulse waveform, readily obtainable from video and commercial smartwatches – can design a spoofing attack strong enough to fool authentication systems based on other cardiovascular biosignals. We present *BioForge*, an approach that leverages a cycle-consistent generative adversarial network to synthesize realistic physiological signals for a given user without relying on simultaneously collected supervision data. We evaluate BioForge on multiple open-access datasets and an array of verification systems, many of which can be fooled over 50% of the time in 10 or fewer attempts. Notably, we are able to fool systems that rely not just on heart rate and peak locations but also on the morphology of the waveforms. We additionally showcase how BioForge can be used to spoof authentication systems from biosignal data extracted from video clips of a target user. Our work demonstrates that authentication systems should not rely on the secrecy of cardiovascular biosignals.

1 Introduction

Recent advances in biosensing and health monitoring have spurred interest in using physiological data as the basis for user authentication systems. In particular, commercial interest in cardiovascular monitoring has stimulated research in authentication systems based on cardiovascular biosignals. Many commercially available wearables (*e.g.,* smartwatches, fitness trackers) advertise the ability to monitor a user's cardiovascular data: small, inexpensive, and high-quality sensors allow for the collection of electrocardiograms (ECGs), photoplethysmograms (pulse waveforms or PPGs), seismocardiograms (SCGs), and ballistocardiograms (BCGs) for cardiovascular health monitoring.

Leveraging these sensors to design authentication systems has gained traction for three main reasons. First, the known disadvantages of current authentication methods have motivated the study of other, more seamless input modalities [9, 21, 59]. Secondly, state-of-the-art machine learning methods have demonstrated remarkable effectiveness in extracting information from physiological time-series data. Traditional authentication systems based on cardiovascular biosignals tend to rely heavily on manual features (*e.g.,* heartbeat segmentation, peak-to-peak distances, spectral analysis); recent deep-learning methods not only perform better but also reduce or eliminate the need for manual feature selection [30, 37, 47]. These advancements have enabled the development of "end-to-end" authentication systems that extract relevant features for user identity prediction along with training the system.

Third, biosignals are assumed to be confidential: an observer cannot easily obtain a user's physiological data and masquerade as that user [30, 37]. Previous work [15] has shown that biometric devices can be compromised if an adversary has access to similar recordings of a target user's biosignal: the BrokenHearted attack demonstrated this vulnerability in 2017 against the Nymi band, a commercially available ECG authentication device in trials at the time with MasterCard [15]. However, system designers assume that physiological signals are difficult to obtain since measurements typically require physical contact with the user. Recent work in video-based, remote estimation of PPG data has chipped away at this assumption [36, 38]. This, however, has still not impeded new research on biosignal-based authentication, with hundreds of mentions of ECG and PPG authentication have been published each year over the past decade with no signs of slowing down; see Figure 2.

In this paper, we show that the key assumption of confidentiality for biosignals is false and that the uniquely identifying features found in a user's cardiovascular signal, such as the ECG, can be consistently leaked from other types of cardiovascular signals from that user. We design black-box spoofing attacks on a wide array of biosignal-based authentication systems to demonstrate that an observer who has some physiological information (obtainable through compromised devices, a leaked database, or even videos of a user) can synthesize a target biosignal to masquerade as the given user. We refer to

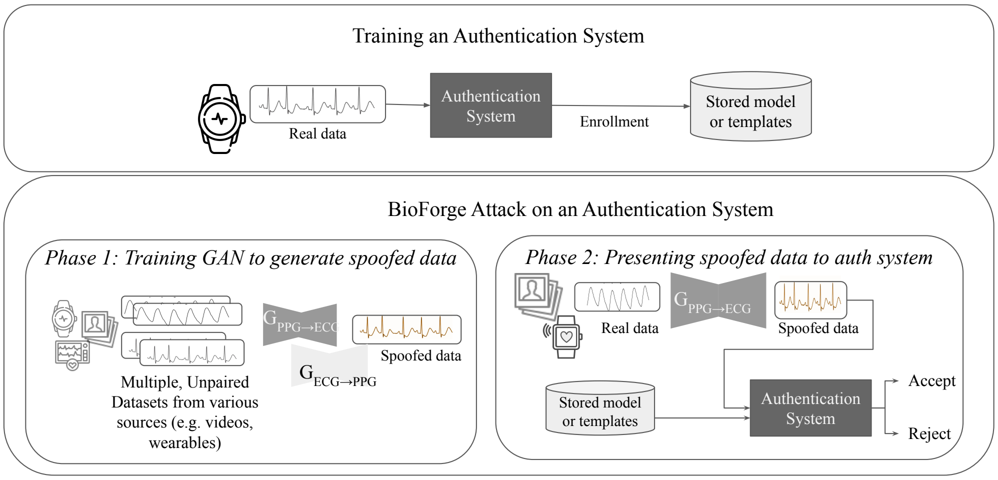
Training an Authentication System
Real data
Authentication System
Enrollment
Stored model or templates
BioForge Attack on an Authentication System
Phase 1: Training GAN to generate spoofed data
$G_{PPG→ECG}$
$G_{ECG→PPG}$
Spoofed data
Multiple, Unpaired Datasets from various sources (e.g. videos, wearables)
Phase 2: Presenting spoofed data to auth system
Real data
$G_{PPG→ECG}$
Spoofed data
Stored model or templates
Authentication System
Accept
Reject

Figure 1: Overview of our BioForge approach. Top: We assume that a user authentication system has been trained on measured biosignal data (*e.g.,* ECGs). Reference templates (or the entire trained model) are stored for retrieval the next time users wish to authenticate themselves. Bottom: Our attack methodology comprises two stages. First, a biosignal generator is trained for a given pair of modalities (*e.g.,* source: PPG, target: ECG). This data is obtained from public datasets or leaked data, such as from health wearables or video clips, and does not need to contain examples from the intended victim. Then, the adversary acquires some true source data from the victim user (again from videos or a data leak), generates spoofed target data using the trained model, and presents the spoofed data to the trained authentication system. The stored templates or model are retrieved for the system to compare against the spoofed data to either accept or reject the attempt.

our overall approach as *BioForge*.

Our approach is comprised of three parts: First, we develop a generative model to synthesize fake, but realistic, biosignal data from another biosignal modality. We call these *target* and *source* signals, respectively. This approach is *system-agnostic*: we do not require details of the target authentication system, and it uses *unpaired* data, *i.e.,* we do not require simultaneously-recorded source-target biosignals. In particular, our model builds on the CycleGAN architecture for unsupervised image-to-image translation [53, 64], which we extend with a contrastive loss term to favor the learning of similar neural representations for biosignals coming from the same user.

Secondly, we implement a wide array of authentication systems published over the past five years and evaluate the success of the forged biosignals created using our generative model in fooling the state-of-the-art authentication techniques. The (offline) user enrollment and the two phases of the attack are illustrated in Figure 1. We primarily focus our evaluation on ECG-based authentication systems because of their popularity due to both the richness of the ECG signal and the wide availability of single-lead ECG sensors in commercial wearables. We also consider systems based on PPGs: pulse waveforms obtained from sensors in many wearable devices with heart monitoring features. To demonstrate the universality of our approach, we additionally evaluate systems based on SCG and BCG signals obtained from accelerometers. We find that spoofed data is able to fool most of these systems within 10 or fewer attempts at authentication.

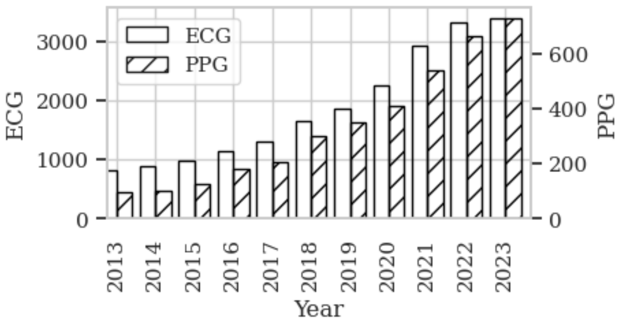

Figure 2: New papers published each year on Google Scholar that mention "authentication" along with either "ECG" (left bars) or "PPG" (right bars)

Finally, we showcase how BioForge can leverage video data to break biosignal authentication systems. We extract remote PPG (rPPG) traces from videos of target users and train BioForge to generate spoofed ECGs from the video-derived traces. We show how these spoofed signals can successfully compromise various authentication systems. Our evaluations illustrate the inherent weakness of biosignal-based authentication systems to spoofing attacks and serve as a warning against the use of these systems in real-world applications.

In summary, our contributions are as follows:

1. We develop a generative model that can synthesize realistic target biosignals from a different cardiovascular biosignal modality in an unsupervised manner.
2. We demonstrate that synthesized ECG traces from PPGs can fool an array of state-of-the-art ECG-based authentication systems.
3. We show the potential for this architecture to generate realistic PPG, BCG, and SCG traces that can fool the corresponding authentication systems.
4. We demonstrate an end-to-end attack that, starting solely from video clips of target users, successfully breaks an array of ECG-based authentication systems.

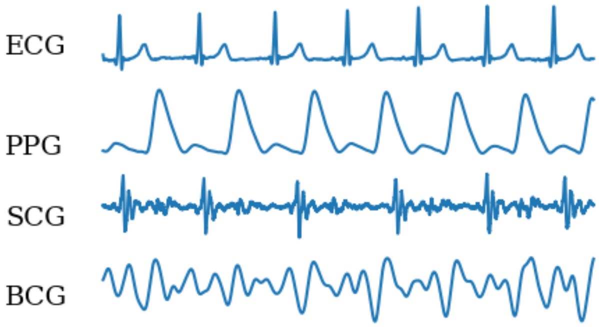

Figure 3: Example traces of each biosignal modality

2 Background

In this section, we provide an overview of the cardiovascular signals of interest and existing techniques for leveraging their statistical properties to develop authentication systems.

2.1 Cardiovascular Biosignals

Cardiovascular biosignals are physiological time-series measurements of cardiac activity. These measurements can be taken from a variety of sensors, on or near the body, and can relay information about heart rate, blood pressure, pulse volume, oxygen saturation, and the general cardiovascular health of the subject. The most recognizable biosignals are electrocardiograms (ECGs internationally; EKGs in the U.S.) and photoplethysmograms (pulse waves, or PPGs). Latif *et al.* [35] and Rathore *et al.* [50] outline the relationships among at least eight types of measurements. For our study, we narrow our focus to the most commonly studied waveforms that have a well-understood morphology, ease of collection, and generally accepted definition: ECG, PPG, SCG (seismocardiogram), and BCG (ballistocardiogram). ECGs are measured via electrodes, PPGs via light reflection or absorbance, and SCGs and BCGs via accelerometers. These four kinds of waveforms are illustrated in Figure 3.

ECG (Electrocardiogram) is a measurement of the heart's electrical activity via recording action potentials from electrodes placed on the skin. The ECG is sufficiently rich to capture the full cardiac cycle, composed of muscle depolarization and repolarization and, as such, it is considered a gold standard for diagnosing cardiovascular diseases in clinical settings. While medical ECG devices consist of 12 leads, an increasing number of wearable devices feature 1-lead ECG sensors. These require at least two points of contact to complete the electrical circuit (e.g., for wrist-worn devices, typically the user needs to touch a sensor with the other hand).

PPG (Photoplethsmogram) is a recording of the changes in blood volume and content, typically obtained by an LED and photodiode pair used to measure the reflection or absorption of light by various components in peripheral blood. The PPG signal is most often used to determine oxygen saturation and heart rate. Conventional sensors (aka pulse oximeters) obtain this measurement at the fingertip, but other common sensor placements include the wrist (for smartwatches), ear, and foot. When PPG is derived from video, we refer to the biosignal as "remote" PPG, or *rPPG*. An rPPG signal is an approximation of PPG derived by tracking the slight fluctuation of red channel values in segments of an individual's face from full-color video recordings.

SCG (Seismocardiogram) is a recording of the vibrations of the chest induced by a heartbeat, typically obtained via an accelerometer placed on the chest.

BSC (Ballistocardiogram) also uses an accelerometer to sense hearbeat-induced vibrations. In contrast to the SCG, the BCG is intended to capture the response at the body's center of mass. Accelerometers for BCG are integrated into chairs, beds, or flat plate scales intended for a user to stand on, and a typical BCG trace is aggregated from multiple sensors.

2.2 Biosignal-based Authentication

Biosignal-based authentication systems leverage the unique statistical properties of physiological waveforms to identify or verify users. A wealth of literature over the past two decades has explored techniques for extracting features from cardiac biosignals [30, 47]. Proposed systems typically consist of the following components in a pipeline: sensing, filtering, segmentation, feature extraction, and template matching. In short, signals are first filtered and then segmented to isolate information about individual cardiac cycles. Features relevant to authentication are subsequently extracted from the segments. During the so-called *enrollment* phase (illustrated in the top half of Figure 1), the extracted features are aggregated to form a user template, which is then stored in a database. When the user later attempts to authenticate themselves, the same process for sensing, filtering, segmentation, and feature extraction is used to build a "test template" to compare against the reference template. A matching algorithm is then used to determine whether the test and reference templates are similar enough to authenticate the user.

Surveys by Li *et al.* [37], Odinaka *et al.* [47] , Ingale *et al.* [30], and Melzi *et al.* [44] have detailed the state-of-art in sensing and signal-processing techniques for designing authentication systems based on PPGs and ECGs. Of particular note is the evidence in support of using single-lead biosignals, which facilitates systems built around wearable devices. Additionally, these surveys enumerate the various segmentation and feature extraction techniques from cardiovascular data, most of which focus on determining significant (*fiducial*) points of the cardiac cycle. Recent end-to-end deep learning approaches learn unique features from the entire morphology of the waveform, rather than relying on manual annotations.

Identification vs. Verification Some systems are primarily designed to perform closed-set user *identification*: a known set of users is fixed, and the system is responsible for predicting the identity of the test subject. These models perform multi-class prediction: given a new test sample, the model is expected to output the known user index most closely associated with the presented sample. A different approach, termed *verification*, seeks to determine whether the test sample belongs to the intended subject. As such, verification systems output a binary prediction rather than the user index. Verification can be realized through recognition, *i.e.,* checking whether or not the test sample and a stored reference sample belong to the same subject, or through one-vs-all identification, where we use a closed-set identification model and check whether the predicted identity corresponds to the intended one. Our work is suited to both types of systems.

3 BioForge

We design a spoofing attack on biosignal-based biometric authentication systems, assuming access to a signal of a different modality. Figure 1 presents the workflow of the attack, which occurs in two stages: training and presentation. First, a generative model is trained on unpaired segments of source and target biosignals using a Contrastive CycleGAN architecture. Then, at attack time, a user's synthetic target biosignal is generated and presented to the authentication system. For concision, Figure 1 only illustrates spoofing ECG (target modality) from PPG (source modality). Our evaluation, however, explores spoofing multiple cardiovascular signal modalities (various target and source types). We detail our threat model in Section 3.1 and the generative model in Section 3.2.

3.1 Threat Model

We assume that a user authentication system has been trained on biosignal data, and we seek to synthesize data that can fool the system. This process is shown in Figure 1. We make the following assumptions about the adversary:

1. The adversary has black-box access to the authentication system, with solely the ability to present synthesized traces from an arbitrary waveform generator as input to the system. Note that in our evaluation, we reimplemented the authentication systems only to evaluate them; no information about the systems is used to generate the spoofed samples.
2. The adversary has access to datasets of source and target cardiovascular samples to train the CycleGAN models. For better generalization, these datasets should contain samples from a variety of subjects (with subject identifier labels[1]), but the biosignals do not need to be simultaneously recorded. Also, source and target data do not need to originate from the same sets of users nor from recorded signals. As we discuss in Section 4.1, such datasets are widely available online.
3. The adversary has access to some type of cardiovascular biosignal from the target user (a signal of a different modality than what is used by the authentication system). In this study, we evaluate scenarios where the adversary has access to either PPG, ECG, SCG, or BCG as source data. We believe that this assumption is realistic. Data leakages in the healthcare sector are alarmingly common and on the rise. Between 2009-2023, healthcare organizations suffered more than 5,000 data breaches, affecting more than 400 million medical records [6]. Recently reported leaks of biosignal data from devices by Fitbit and Apple [16, 42] demonstrate a significant likelihood that an attacker obtains access to some type of cardiovascular data. The situation is exacerbated by a rapidly growing market for commercial health devices, including many products with less rigorous security than those from major brands. Moreover, as discussed in Section 2, the attacker can acquire PPG signals even remotely from video data, and such signals can be directly used by our method.

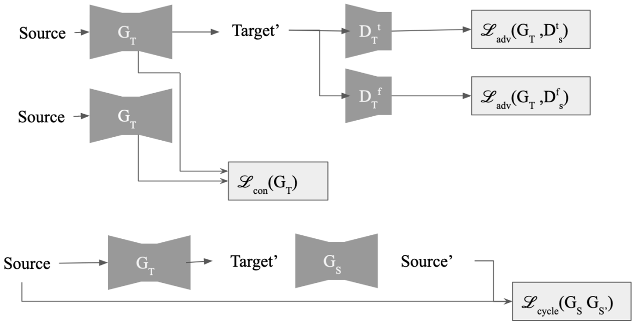

Figure 4: Overview of one half of the signal generation model. The components shown involve translating a Source signal to a Target form using a Generator network G_T. The discriminator networks D_T^t and D_T^f are trained at the same time to distinguish between real and synthetic data. We simultaneously train the inverse mapping: G_S, D_S^t and D_S^f for translating from Target to Source modalities. The loss functions associating with the inverse are $\mathcal{L}_{adv}(G_S,D_T^t)$, $\mathcal{L}_{adv}(G_S,D_T^t)$, $\mathcal{L}_{cycle}(G_T,G_{T'})$, and $\mathcal{L}_{con}(G_S)$.

3.2 Biosignal Generation Model

The proposed biosignal generation model is based on the CardioGAN model [53] with an additional contrastive loss component [34]. CardioGAN leverages the CycleGAN architecture [64] to translate signals between different modalities in an unsupervised manner (*i.e.,* with unpaired data), which in turn builds on the popular GAN model [20].

A GAN is comprised of a generator and a discriminator network, which are trained simultaneously in a competitive scheme. The generator aims to improve its ability to generate realistic data to fool the discriminator, and the discriminator aims to improve its ability to distinguish between real and generated data. The two models are updated in an alternating fashion. After training, only the generator is used to synthesize data. A CycleGAN network consists of two (conditional) GANs that are trained in tandem to enforce a condition of cyclical consistency (in addition to the usual GAN discriminator loss). These generators are $G_T : S \rightarrow T$, the generator that translates source data into target data, and $G_S : T \rightarrow S$, the generator that translates target data into source data. Cycle consistency imposes that a source signal s should be close to $G_S(G_T(s))$, i.e., when s is translated into the target modality

[1] This is required by our contrastive loss update.

through G_T, and then back to the source modality with G_S, the resulting signal should be close to the original signal s. The same condition is enforced for a target signal t and its cycle-translation $G_T(G_S(t))$. This two-step process allows training generators without requiring pairs of supervised training data, which may not be available (as in our case). We show one half of the model (relative to the S-to-T translation) in Figure 4.

Each generator is based on the CardioGAN design [53], which combines elements of a U-Net with self-attention mechanisms [48] to encourage the network to focus on critical sections of the waveforms. Each generator follows an encoder-decoder architecture consisting of multiple convolutional downsampling, attention, and upsampling layers. CardioGAN introduces two discriminators (instead of one) to preserve time- and frequency-domain information of the generated signals. In particular, the generator G_T (G_S) is paired with two discriminators D^t_T and D^f_T (D^t_S and D^f_S): the former works in the time domain and is trained to distinguish between true signals t (s) and synthetic ones $G_T(s)$ ($G_S(t)$); the latter operates in the frequency domain and discriminates between $DFT(t)$ ($DFT(s)$) and $DFT(G_T(s))$ ($DFT(G_S(t))$), where $DFT(\cdot)$ is the discrete Fourier transform.

Each discriminator's architecture consists of the same downsampling convolutional layers used for the encoder portion of the generators. The discriminator's output represents the probability that the given input comes from real data. Generators and discriminators are updated using three loss functions, defined below only for the half involving G_T, D^t_T and D^f_T. These are the GAN loss relative to the time domain, $\mathcal{L}_{adv}(G_T, D^t_s)$, the GAN loss relative to the frequency domain, $\mathcal{L}_{adv}(G_T, D^f_s)$, and the cycle-consistency loss (involving only the time domain), $\mathcal{L}_{cyclic}(G_T, G_S)$:

$$\begin{aligned}\mathcal{L}_{adv}(G_T, D^t_T) = \mathbb{E}_t[\log(D^t_T(t))]&\\ +\mathbb{E}_s[\log(1-D^t_T(G_T(s)))]&\quad (1)\\ \mathcal{L}_{adv}(G_T, D^f_T) = \mathbb{E}_t[\log(D^f_T(t))]&\\ +\mathbb{E}_s[\log(1-D^f_T(G_T(s)))]&\quad (2)\\ \mathcal{L}_{cyclic}(G_T, G_S) = \mathbb{E}_s[\|G_S(G_T(s))-s\|_1]&\quad (3)\end{aligned}$$

where $\|\cdot\|_1$ is the L1-norm. Loss functions for the inverse mapping are defined analogously. We denote with $\mathcal{L}_{adv}(G_T, D^t_T, \mathcal{S}, \mathcal{T})$, $\mathcal{L}_{adv}(G_T, D^f_T, \mathcal{S}, \mathcal{T})$, and $\mathcal{L}_{cyclic}(G_T, G_S, \mathcal{S})$ the empirical approximations of the above expectations via unpaired batches of source signals $\mathcal{S}$ and target signals $\mathcal{T}$.

On top of the CardioGAN model, we additionally adopt a supervised contrastive loss component. Its purpose is to ensure that the latent features of the generators (*i.e.,* the bottlenecks) are similar for signals coming from the same patient. In doing so, we condition the network to extract a representation of the subject that remains (close to) invariant despite signal variability. This invariant representation can be seen as a proxy for a subject's identifying features; *i.e.,* if two signals have similar latent features, then they likely belong to the same subject. The same principle is applied in popular biometrics and face recognition architectures such as Hadsell *et al.* [22] and Schroff *et al.* [55].

In particular, we consider the SupCon loss from [34] (using the implementation by [45]) for its ability to handle an arbitrary number of positive and negative pairs in a batch of signals. Given a batch of signals $\mathcal{D}$ and a signal $x \in \mathcal{D}$, we denote with $[x] \subseteq \mathcal{D} \setminus \{x\}$ the set of signals other than x but that belong to the same subject as x. For a signal x and generator G, we denote with $z_G(x)$ the latent representation associated with $G(x)$. Then, the contrastive loss relative to G and $\mathcal{D}$ is given by

$$\mathcal{L}_{con}(G, \mathcal{D}) = -\sum_{x\in\mathcal{D}} \frac{1}{[x]} \sum_{p\in[x]} \log\left[\frac{\exp\frac{z_G(x)\cdot z_G(p)}{\tau}}{\sum_{x'\in\mathcal{D}\setminus\{x\}} \exp\frac{z_G(x)\cdot z_G(x')}{\tau}}\right] \quad (4)$$

where $z_G(x) \cdot z_G(p)$ denotes the dot product between vectors and $z_G(x)$ and $z_G(p)$ and is used as a similarity measure. Hence, the softmax argument of the log function represents a smooth approximation of whether p has the highest similarity to x across all signals in the batch. The temperature $\tau > 0$ controls such smoothness (larger τ, less smooth). We set $\tau = 0.1$ in our evaluation. It is clear to see that minimizing (4) corresponds, for each $x \in \mathcal{D}$, to maximizing the (dot-product) similarity between the representations of x and all signals in $[x]$, consequently minimizing the similarity with signals not in $[x]$ (thanks to the softmax term).

Finally, we update the networks' weights in two separate steps to avoid interference between different loss functions: first, in the direction of the GAN and cyclic losses to improve the quality of the generated signals (see Equation 5 below), then, in that of the contrastive loss to capture key identifying information (see Equation 6). For batches of source and target signals $\mathcal{S}$ and $\mathcal{T}$, the final losses become:

$$\begin{aligned}\mathcal{L}_{GAN} = \alpha(\mathcal{L}_{adv}(G_T, D^t_S, \mathcal{S}, \mathcal{T}) + \mathcal{L}_{adv}(G_S, D^t_S, \mathcal{T}, \mathcal{S}))&\\ +\beta(\mathcal{L}_{adv}(G_T, D^f_T, \mathcal{S}, \mathcal{T}) + \mathcal{L}_{adv}(G_S, D^f_S, \mathcal{T}, \mathcal{S}))&\\ +\gamma(\mathcal{L}_{cyclic}(G_T, G_S, \mathcal{S}) + \mathcal{L}_{cyclic}(G_S, G_T, \mathcal{T}))&\quad (5)\end{aligned}$$

and

$$\mathcal{L}_{con} = \mathcal{L}_{con}(G_S, \mathcal{S}) + \mathcal{L}_{con}(G_T, \mathcal{T}) \quad (6)$$

where $\alpha, \beta, \gamma > 0$ are hyper-parameters for combining the individual loss components in the first update.

4 Evaluation

In this section, we present the results of our evaluation. We first describe the open-access datasets and authentication systems chosen to demonstrate our approach in Section 4.1 and 4.2, along with performance on true data. Section 4.3 presents details on the training and performance of the CycleGAN model. Next, we present our comprehensive evaluation of the BioForge approach in Section 4.4-4.6. Finally, Section 4.7 presents our findings from a study using PPG traces extracted from video clips.

4.1 Datasets

We performed our primary evaluation on four publicly available datasets that contain simultaneous measurements of multiple cardiovascular signals. These datasets represent collections from a variety of device types (clinical vs wrist-worn wearables), sampling frequencies (64Hz to 5000Hz), and participants (varied in age and gender). We note that we do not need paired (simultaneously recorded) data to train our generative model and run the attack, but we do need paired recordings to understand and evaluate the performance of our approach. We use data from one modality (*e.g.,* ECG) to train the authentication systems and data from a second modality (*e.g.,* PPG) as a source to generate forged data to present to the trained system. We include four datasets for ECG/PPG translation, one dataset for ECG/SCG translation, and one dataset for ECG/PPG/BCG translation. We additionally evaluated the ability for BioForge to succeed using a biosignal extracted solely from video clips of a target user. For this purpose, we use the MANHOB-HCITagging dataset [57], which includes video/ECG pairings.

***BIDMC* [19, 49]** contains clinical PPG and ECG readings from 53 ICU patients. Each recording is 8 minutes long, and both PPG and ECG are sampled at 125 Hz. BIDMC is a subset of the larger MIMIC II matched waveform database, which contains thousands of vital sign recordings collected bedside in adult and neonatal ICUs.

***CAPNO* [33]** contains clinical PPG and ECG readings from 42 hospitalized children and adults. Each recording is 8 minutes long and all traces are sampled at 300Hz. This is a subset of the larger CapnoBASE benchmark dataset, used to develop algorithms for monitoring CO2 breath saturation and respiratory rate.

***DALIA* [51]** contains PPG and ECG signals from 15 people wearing portable devices. ECGs are obtained from a chest-worn device (RespiBAN[2]) and PPGs were obtained from a wrist-worn device (Empatica E4[3]). Each recording is about 2 hours. PPG was recorded at 64Hz and ECG was recorded at 700Hz.

***WESAD* [54]** contains PPG and ECG signals from 15 people wearing the same ECG and PPG devices used for the DALIA dataset. Each recording is about 1 hour long.

***CEBSDB* [17, 19]** contains SCG and ECG signals for 20 participants in a lab environment recorded using the Biopac MP36 system[4]. Each recording is about 5 minutes. The SCG data is an aggregation of measurements obtained from a triaxial accelerometer and filtered between 0.5Hz and 100Hz. All data was recorded at 5000Hz.

***BedBased* [23]** contains simultaneous BCG, ECG, and PPG recordings from 40 participants in a lab environment. BCGs were collected from a custom bed comprised of four electromechanical films and four load cells. We average the four load cells to obtain a single BCG signal. Each recording is about 7 minutes long and sampled at 1000Hz.

***MAHNOB-HCITagging* [57]** contains data originally collected for studying emotion and affect recognition and includes numerous sessions of video and ECG recordings from 27 participants. ECGs were monitored at 256Hz, and videos were recorded at 60Hz with a resolution of 780x580 pixels. We used all available ECG traces and only videos recorded in color.

4.2 Authentication Systems

We evaluated the ability of our generated biosignal traces to spoof a range of published authentication systems. We looked for systems proposed within the past five years with comparatively high citation counts as representative of state-of-the-art techniques. We selected more ECG-based systems than other modalities due to the abundance of research on ECG systems and the richness of the signal. We also gave preference to publications that made their code publicly available. For training the authentication systems, we followed, where possible, the individual systems' procedures for processing data and hyperparameter selection. We additionally resampled, detrended, standardized, and rescaled data as needed.

4.2.1 Performance Metrics

We first report each system's EER (Equal Error Rate) on true (non-spoofed) data. EER is a preferred evaluation metric for authentication systems, as accuracy can be misleading if the number of samples from each user varies. If we see an authentication system as a function that outputs the likelihood that the given signal is from the intended user, then the EER is obtained by finding the decision threshold for the said likelihood that yields equal rates of false acceptances (FAR) and false rejections (FRR). This threshold is found using a held-out set of calibration signals.

There is a natural tradeoff between FAR and FRR, and the EER is the position on this tradeoff curve where the two are equal (or closest) as the decision threshold is changed. An EER of 0 indicates perfect performance. The FAR and FRR rates can be derived from the empirical counts of False Positives (FP), False Negatives (FN), True Positives (TP), and True Negatives (TN) as follows: $\text{FAR} = \frac{\text{FP}}{\text{FP+TN}}$ and $\text{FRR} = \frac{\text{FN}}{\text{FN+TP}}$.

We note that the EERs obtained in our evaluation do not necessary align with those reported in the respective original studies. This is because we take a different approach to splitting data for training and testing. Most studies shuffle all available data segments and randomly assign them to training and test sets. We instead split our data over the time dimension to mimic a realistic scenario: the systems are trained on the earliest section of available data (by time), tested on the next section to obtain EERs and associated EER thresholds, and evaluation of spoofed data is performed on the last section in time. Also, we remark that most systems present approaches highly tuned to their collected data, and thus we should ex-

[2] https://support.pluxbiosignals.com/knowledge-base/cardioban-getting-started/

[3] https://www.empatica.com/research/e4/

[4] https://www.biopac.com/product-category/research/systems/mp36r-4-channel-systems/

Table 1: Performance of ECG-based authentication systems. Equal Error Rates (%) are presented for true (non-spoofed) ECG traces.

System Name	System Description	BIDMC	CAPNO	DALIA	WESAD
KeyToHeart [52]	Support Vector Machine on 25 principal components from ECG segments	35.35%	35.36%	31.19%	29.13%
Deep-ECG [14]	CNN on 8 consecutive QRS complexes	8.00%	8.00%	17.0%	11.80%
EDITH [28]	CNN feature extractor and FFNN siamese network on segmented QRS complexes	2.85%	3.51%	10.59%	1.00%
ECGXtractor [44]	Autoencoder feature extractor and CNN siamese network from a consensus of segmented QRS complexes	7.68%	10.62%	15.95%	14.6%

Table 2: Performance of PPG-based authentication systems. Equal Error Rates (%) presented for true PPG traces.

System Name	System Description	BIDMC	CAPNO	DALIA	WESAD
Hwang2020 [27]	CNN and LSTM on single-pulse segments transformed via FFT, dynamic time warping, and wavelet transforms	7.12%	15.35%	43.45%	42.03%
CorNet [7]	CNN with dual heads to predict both heart rate and identity	3.75%	1.49%	20.25%	29.18%

Table 3: Performance of SCG-based authentication systems on CEBSDB dataset. Equal Error Rates (%) presented for true SCG traces.

System Name	System Description	EER
WaveletTransform [26]	Template matching from wavelet-transformed segments	23.48%
MotionResilient [25]	Support Vector Machine classification on segments	13.29%

Table 4: Performance of BCG-based authentication systems on BEDBASED dataset. Equal Error Rates (%) presented for true BCG traces.

System Name	System Description	EER
HebertCNN [23]	CNN on 3-sec rolling windows	23.48%
ZhangRNN [62]	RNN on segments centered around corresponding ECG R-peaks	13.29%
ZhangRNN+ECG [62]	RNN on concatenated BCG and ECG segments centered around corresponding ECG R-peaks	13.29%

pect performance to deviate when these systems are applied to different datasets (as we do).

4.2.2 ECG Authentication

The performance of the following four ECG-based authentication systems is reported in Table 1.

***KeyToYourHeart* [52]** is a proof-of-concept system for user identification and verification using a commercially-available "pocket" ECG sensor: the KardiaMobile device by AliveCor. [5] The system extracts the top 25 principal components of ECG segments to obtain template features. Classification is performed on these templates using a support vector machine with a radial basis function in a binary one-vs-all scheme. We based our implementation on a codebase provided by the authors.

***Deep-ECG* [14]** by Labati *et al.* is one of the first end-to-end approaches that leverage deep convolutional networks for ECG-based identification and verification. They use a deep CNN trained on QRS complexes (the portion of the ECG cycle closely surrounding the highest peak) and evaluate closed-set identification, identity verification, and periodic re-authentication across time.

***EDITH* [28]** by Ibtehaz *et al.* present a sophisticated deep learning system with two components: a convolutional neural network (with multi-resolution blocks and spatial pooling layers) to learn a feature embedding from segmented QRS complexes, and a Siamese network to train the identification/verification model. The authors released their implementation on GitHub.[6]

***ECGXtractor* [44]** by Melzi *et al.* surveys existing literature to identify the best combination of signal processing methods for single- and multi-session user authentication. They implement a two-step approach for feature extraction and identity recognition. They train an autoencoder with multiple convolutional layers to learn a feature embedding from segmented QRS complexes, and the bottleneck of the autoencoder is then used as inputs for a Siamese model that performs user verification. The authors of ECGXtractor released their implementation and some pre-trained models on GitHub.[7] We use their pre-trained autoencoder and train separate Siamese networks on our individual datasets.

We report that ECG systems based on deep learning models attain low EERs, no larger than 17%, while we see a consistent performance degradation in the only “non-deep” system (KeyToYourHeart).

4.2.3 PPG Authentication

The performance of the following PPG-based systems is reported in Table 2. We acknowledge the high error rates of the DALIA and WESAD datasets, which we explain by the different nature of the datasets. The Hwang2020 system [27] was developed on clinical PPG waveforms that have clear dicrotic notches and follow diastolic peaks. Their proposed data transforms take advantage of the clear peaks and troughs of the signals. The data in DALIA and WESAD were collected from wrist-worn bands (rather than clinical pulse oximeters) and their PPG traces generally do not have pronounced dicrotic notches. Nonetheless, we decided to include Hwang2020 as their work is among the most highly referenced recent efforts on PPG authentication, and the authors made their code avail-

[5] https://store.kardia.com/products/kardiamobile
[6] https://github.com/nibtehaz/EDITH
[7] https://github.com/BiDAlab/ECGXtractor

able for further research.

***Hwang2020* [27]** by Hwang *et al.* stacks the results of multiple data transforms (in both time and frequency domains) on segmented pulse waves to achieve a standard signal length before feeding the data into a CNN+LSTM for user identification and authentication. The authors released an example of their approach on GitHub. [8]

***CorNET* [7]** by Biswas *et al.* presents a deep learning approach to perform PPG-based identification by leveraging two loss functions: one for closed-set user identification (cross-entropy) and another (mean squared error) for heart rate estimation. The network consists of 2 CNN layers and 2 LSTM layers, with two final dense layers for predicting identity and HR, independently.

4.2.4 SCG Authentication

The performance of the following SCG-based authentication systems is reported in Table 3.

***WaveletTransform* [26]** explores a variety of wavelet-transformed signals, feature extraction, and matching techniques to perform authentication on SCGs. While many of their models achieve similar performance, we selected the Morse Wavelet and L2-norm distance for matching.

***MotionResilient* [25]** The approach consists of averaging five one-second segments of SCG data centered around the ECG's R-peaks for motion artifact removal and then using a Support Vector Machine classifier to predict a user's identity. We implement their approach as a one-vs-all scheme for each user.

4.2.5 BCG Authentication

The performance of the following BCG-based authentication systems is reported in Table 4.

***ZhangRNN* [62]** explored BCG traces of various lengths and a few types of recurrent neural networks. Of their proposals, we adopted the system that concatenated 15 segmented heartbeats to feed as input to a single-layer LSTM network. However, their heartbeat segmentation relies on having an ECG source for R-peak detection. Because of the assumption that ECG is available, they also explored a multimodal approach: they concatenated ECG and BCG traces for increased identification accuracy.

***HebertCNN* [23]** explored BCG-based authentication from a head-mounted wearable device (the Google Glass). They trained a convolutional network on BCG data from independent sensor streams split into 3-second segments for each user, using a one-vs-all identification scheme.

4.3 BioForge Signal Generation Performance

In this subsection, we detail the training process of the BioForge generation model and describe its performance in synthesizing realistic biosignals.

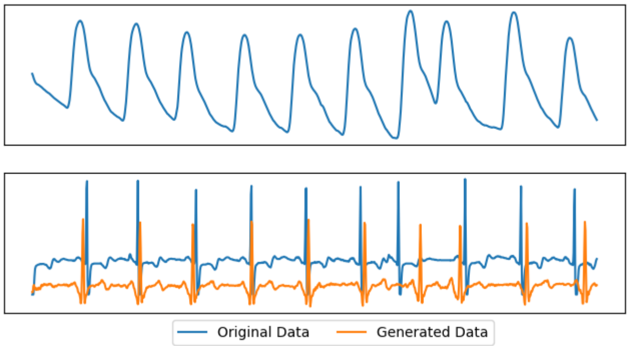

Figure 5: Example original and generated traces of True PPG (blue, top), True ECG (blue, bottom), and BioForge generated ECG (orange, bottom). The generated ECG appears to capture the inter-peak interval morphology of the true ECG well, in addition to reproducing R-peak locations. We observe that the generated data carries over inaccuracies from the source: the true PPG contains an inaccurate extra waveform (the 8th PPG peak), and that extra peak is reflected in the generated ECG.

Data Preparation. A portion of data from the BIDMC, CAPNO, DALIA, and WESAD datasets were used to train the BioForge model for translating between ECG and PPG traces. For translating between ECG and SCG, only the CEBSDB dataset was used. For translating among ECG, PPG, and BCG, only the BedBased dataset was used. We split traces in time and used the first half of each trace to train and test the BioForge model, and the second for training and evaluating the authentication systems. After combining the datasets and segmenting traces into overlapping segments, all segments were shuffled to break relationships among users and datasets. In this way, we prevent the model from memorizing mappings between signals of different modalities for the same user, and between signals of different modalities for a given dataset. We implemented the same signal processing pipeline of CardioGAN [53], which includes resampling all data to 128Hz (or 512Hz for SCG), filtering, standardizing, and shuffling segments to break pairs between simultaneously recorded data.

Training settings. The BioForge models were trained for 25 epochs for the ECG/PPG targets, and 50 epochs for the SCG/BCG targets. We used a total of four Adam optimizers (one for both generators and the GAN loss, one for both generators and the contrastive loss, and one for each discriminator). For all Adam optimizers, we set $\beta_1 = 0.5$ and $\beta_2 = 0.999$ (i.e., the momentum parameters for the gradient and its square) and a learning rate of 10^{-4}, which kept constant for 10 epochs and then decays to 10^{-6} linearly (as done in the CardioGAN paper). We set the GAN loss weights to $\alpha = 3, \beta = 1, \gamma = 100$, as recommended in [53].

Performance. The CardioGAN authors report the accuracy of generated traces in terms of the difference in estimated heart rate (in beats per minute, or BPM) between the generated ECG and ground truth. While we are interested in more than just accurate heart rate estimation, this metric provides an intuitive interpretation of the performance of the models. We observe that

[8] https://github.com/eoduself/PPG-Verification-System

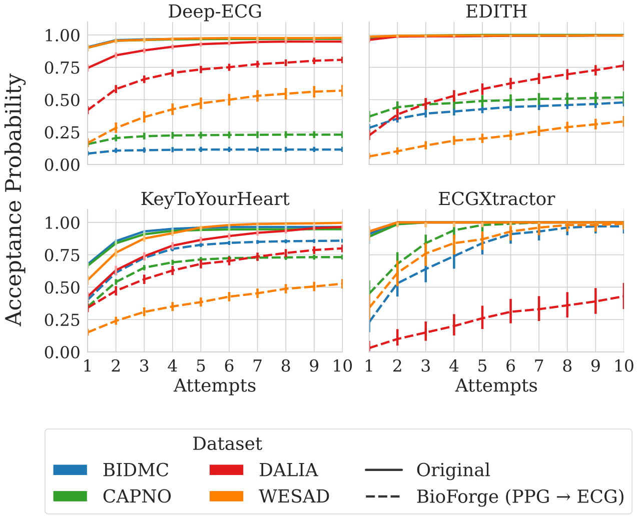

Figure 6: Accumulating probability of accepting an ECG trace, up to 10 attempts. The solid lines show the true probability of acceptance on original ECG traces from the intended user (True Positive Rate, TPR, on true data over time). Corresponding dashed lines indicate the acceptance of spoofed data for the intended user (FPR over time from synthesized data). We observe that for most systems and datasets, the spoofed data was accepted over 50% of the time after 10 attempts.

the mean absolute error (MAE) in heart rate for samples generated by our model is comparable with CardioGAN results. MAE between true and generated ECG ranges from 0.80 BPM to 8.39 BPM across datasets for 60-second traces from a test set.

4.4 Spoofing Success over Attack Attempts

We evaluate the False Acceptance Rate (FAR) of spoofed data in comparison to the True Positive Rates of each system. We report our findings as an accumulating rate of success over 10 attempts at authentication. Typically, a system will allow for multiple authentication attempts before limiting or blocking the user. We chose 10 as the maximum number of tries because many systems, including Apple iOS devices, lock users out permanently after 10 attempts [56]. Results are reported with 95% confidence intervals.

We first focus our attention on ECG-based systems. Figure 6 presents the acceptance rate of true ECG data (solid lines) and spoofed data synthesized from PPG for the four datasets used to train the generation model. We additionally demonstrate the performance of spoofed PPG (from ECG) over the selected PPG-based authentication systems, in Figure 7. Lastly, the performances of spoofed BCG and SCG data are displayed in Figure 8 and Figure 9.

Across all datasets and modalities, we observe that over multiple attempts, the likelihood of a spoofed trace being falsely accepted often surpasses 50%, approaching the authentication performance of true data in multiple settings. Moreover, while some systems (*e.g.,* Deep-ECG) seem to be robust to spoofed data from some datasets, the chance of falsely accepting spoofed data over 10 attempts still often exceeds the reported

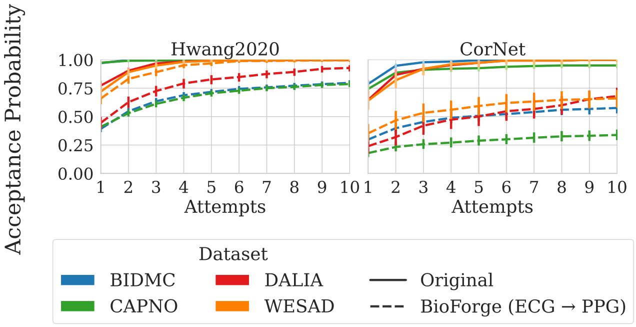

Figure 7: Accumulating probability of accepting a PPG trace, up to 10 attempts. The solid lines show the true probability of acceptance on true PPG (true positive rate), and the dashed lines show the probability of acceptance on spoofed PPG (false positive rate from spoofed data). Similar to ECG results shown in Figure 6, we observe that the spoofed FAR is often higher than expected from reported EERs.

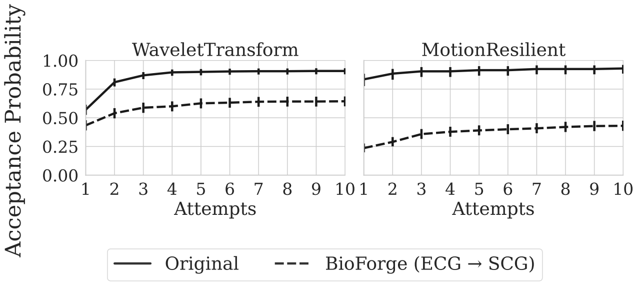

Figure 8: Performance of original and spoofed data from the CEBSDB dataset on SCG based authentication systems. Solid lines represent the True Positive rates from true data, and dashed lines represent the acceptance rates of spoofed data.

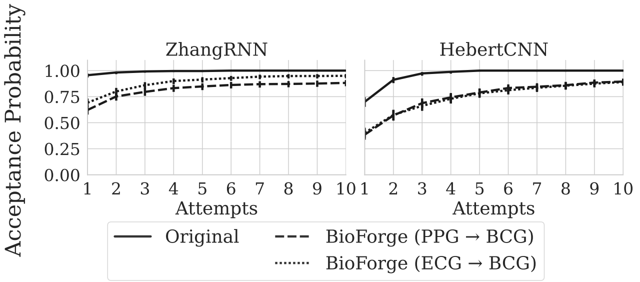

Figure 9: Performance of original and spoofed data on BCG based authentication systems. Solid lines represent the True Positive rates from true data, and dashed lines represent the acceptance rates of spoofed data. The dataset from which BCG is sourced (BEDBASED) contains simulatneous ECG and PPG recordings as well, so we are able to spoof BCG from either source.

test EER. Moreover, a low EER doesn't necessarily provide any information about robustness against spoofed data: both Deep-ECG and ECGXtractor reported similarly low EERs on our datasets, yet spoofed attempt successes differed greatly.

4.5 Multi-modal Setting (ECG+BCG)

A common defense against potential spoofing attacks is to incorporate multiple modalities of physiological data to temper the effects of one tampered data source. However, with our BioForge approach, if an attacker has access to one of these

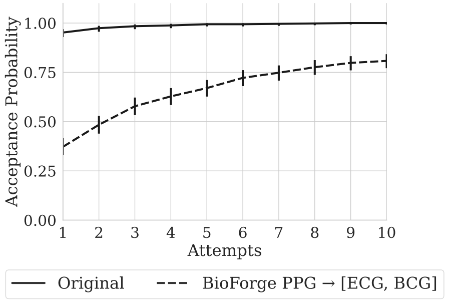

Figure 10: Case study of a multimodal authentication system. The ZhangRNN system considers concatenated ECG and BCG traces. We show that BioForge can spoof both ECG and BCG data sufficiently well from PPG.

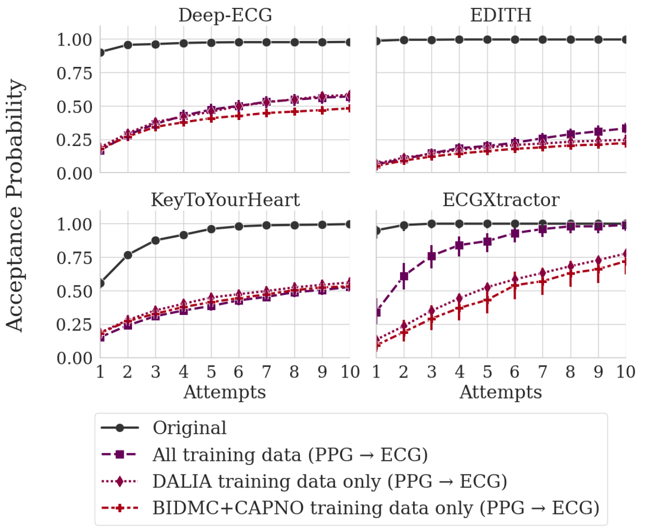

Figure 11: BioForge performance on WESAD dataset for ECG-based system, using various sources of data for training the CycleGAN. The solid lines refer to the original performance of true ECG data on each system. The remaining dashed lines represent various trained models, relying on data from all datasets (■), from just a similar dataset (⧫), and from different datasets (+).

multiple modalities, they can generate signals for the other ones, which renders this particular defense helpless. We evaluate the ability of spoofed data to fool systems that are multimodal by design: we choose to investigate the ZhangCNN BCG system, as the authors propose a version that explicitly operates on a combination of ECG and BCG traces, generating spoofed data from a different source modality (PPG). As shown in Figure 10, the performance of this attack achieves a significant success rate (80%) after 10 attempts. Note that we do not consider here multi-modal systems that simply aggregate predictions from separately trained uni-modal systems, as the performance of these can be trivially obtained by combining the results from the individual models.

4.6 Cross-dataset Transfer

In all previous experiments, our BioForge generative model was trained on an aggregation of data across multiple datasets. Here, we explore the potential of our model trained solely on one dataset to produce realistic traces for a different dataset.

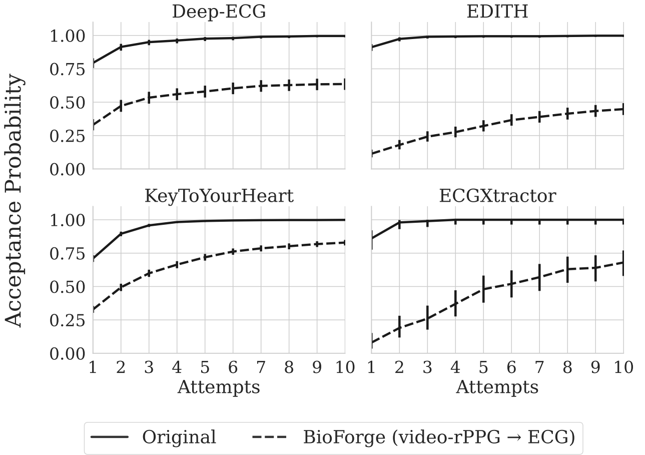

Figure 12: ECG Spoofing results for HCITagging dataset using video as starting point. rPPG traces were extracted from video clips, and the BioForge model was trained on pairs of rPPG and ECG to generate synthetic, spoofed ECG.

Figure 11 shows False Acceptance Rates for various systems trained on WESAD data. The Combined Training Set lines (square markers) correspond to the results shown earlier, where all datasets are used to train the CycleGAN. Next, we examine the attack performance when the generative model is trained only on DALIA data. DALIA and WESAD datasets were collected using the same devices (chest- and wrist-worn wearables). Lastly, we show how a model trained on purposefully different data – BIDMC and CAPNO, which are clinical datasets – can generate data that can still fool the authentication systems. Spoofed data generated using just DALIA performed similarly to that generated from BIDMC and CAPNO, which indicates that there is sufficient similarity among biosignals obtained from different devices to facilitate transfer attacks.

4.7 Video-derived source biosignal

In Section 3.1, we noted that emerging work on remote physiological data sensing would facilitate access to leaked biosignal data. In this section, we demonstrate the feasibility of an end-to-end attack that starts solely with video data rather than a leaked biosignal. We showcase this performance using the MAHNOB-HCITagging dataset, which contains video and ECG recordings from participants. From these videos, we leverage recent advances in remote physiological sensing to extract an *rPPG* pulse trace. For this purpose, we use the CHROM method [13] from the rPPG-Toolbox [41]. [9] An advantage of CHROM is that, unlike other PPG extraction methods, it does not require examples of true PPGs or any other physiological information to train the extraction model. After extracting rPPG from videos, we train our BioForge generative model on shuffled pairs of rPPG and ECG. We compare the performance of the array of ECG-based authentication systems using the true ECG and the spoofed data, shown in Figure 12.

We do not directly calculate the extent to which the video-

[9] https://github.com/ubicomplab/rPPG-Toolbox

derived PPG degrades the performance of BioForge, because the HCITagging dataset does not contain true PPG data. To the best of our knowledge, no large public research dataset exists that contains video, PPG, and ECG recordings from the same participants.[10] Even though the attack employs a video-derived approximation of the true PPG, we observe that most systems accept such spoofed data over 50% of the time after 10 attempts, which is comparable with the results on the other datasets starting with the true PPG.

5 Discussion and Countermeasures

Our evaluations illustrate that biosignal authentication systems cannot solely rely on a presumption of biosignal secrecy: synthesized data can compromise an array of authentication systems that operate on various cardiovascular biosignals. Nonetheless, this work prompts further lines of research regarding safeguards and potential countermeasures:

5.1 Countermeasures

Multi-modal authentication systems These systems can combine information from multiple sources to build a more comprehensive profile of someone's identity. Our experiments illustrate that practitioners should be careful to avoid using solely cardiac biosignal data for all modalities – one compromised modality can be used to synthesize others and serve as a single point of failure. Multi-modal systems that combine data from clearly independent sources might offer a path forward.

Obfuscating physiological information in video While preventing large-scale data leaks is a pressing area of concern for healthcare security in general, our video-based attacks in Section 4.7 demonstrate that keeping physiological data private remains a challenge. We used a method for rPPG extraction which requires nothing but videos of the victim. As methods for remote physiological sensing advance (see Section 6.1 for more details), we can expect the quality of rPPG and subsequent other cardiovascular data only to improve. However, recent research shows that it is possible to remove physiological information from videos to preserve privacy without altering the videos in noticeable ways [10, 12, 58].

Adversarial training to improve authentication systems Our work prompts a line of defensive strategies that actively uses spoofed data to improve the robustness of the authentication systems at training time. One could include synthetic data as negative examples to encourage models to differentiate between true and fake data. Alternatively, these systems could include a liveness test or an initial signal filter to identify anomalous or synthesized traces before allowing an authentication attempt. Nonetheless, we argue that BioForge produces synthetic data that is remarkably similar to true data (along time and frequency domains), and detection of synthetic data is likely a difficult challenge. Moreover, even if a system can detect spoofed data generated using BioForge, future generations of models for synthesizing biosignals might be able to evade detection.

5.2 Ethical Considerations

To the best of our knowledge, no commercial devices that use biosignal authentication are available off-the-shelf to consumers. The lone device that uses ECG for authentication (*i.e.,* Nymi Band [29]) does so in conjunction with fingerprint authentication for the purpose of continuous authentication. While our work reveals a fundamental limitation in the design of biosignal authentication systems in general, we do not develop an exploit to compromise the Nymi Band. However, we have initiated contact with Nymi to make them aware of this potential vulnerability. We hope our paper serves as a forewarning to future system designers to seriously consider the threat model associated with spoofed data when using such systems. All our experiments were conducted on publicly available datasets and open-source biosignal authentication systems.

6 Related Work

Our approach involves synthesizing biosignals for spoofing authentication systems, and builds on related work in both biosignal synthesis and spoofing attacks.

6.1 Synthesis of Cardiovascular Biosignals

There exists a wealth of literature on synthesizing physiological data, with the goal of strengthening medical diagnosing. Much of this work has focused on synthesizing ECGs, as clinical interpretation of high-quality ECG recordings is considered the gold standard for diagnosing heart failure. An early popular approach generated ECG signals as the sum of Gaussian curves parameterized with time-domain features manually extracted from the patient's ECG [43]. Current state-of-the-art approaches avoid relying on manual feature selection and mapping, as they instead leverage deep generative models. Various frameworks have been proposed to synthesize ECG, such as the Variational Autoencoder [4] and diffusion models [2]. Notably, Generative Adversarial Nets (GANs) [20] have emerged as a popular framework for the synthesis of ECGs to augment or improve the quality of existing data [1, 5, 61, 61]. For example, multiple works have used GANs to reconstruct a clinical 12-lead ECG recording from a single lead signal for improved arrhythmia diagnosis [3, 31, 63]. Other related approaches consider generating ECG data given another cardiovascular biosignal. The CardioGAN approach by Sarkar *et al.* [53] leverages a GAN framework to generate ECGs from PPG waveforms, and they evaluate the performance of the model in terms of estimated heart rate across a variety of datasets. We base our model on the CardioGAN approach and describe CardioGANs further in Section 3.2. Rather than assessing model performance in terms of heart rate estimation, we evaluate performance in terms of authentication security.

Considerable attention has also been placed on synthesizing PPG data, mainly from remotely-collected data. Traditional

[10] The OBF database [39] did include all three; however, it is no longer available due to GDPR concerns.

PPG measurement requires physical contact with the user, but recent research has shown that is is possible to estimate a person's PPG waveform from videos [11, 18, 40, 60]. While this is an area of active research, typical approaches involve the detection of regions of interest from videos that correspond to informative areas of the skin, followed by the detection of changes in color intensity within these regions that are correlated with the cardiac cycle. Boccignone *et al.* [8] and Liu *et al.* [40] provide open-source Python toolkits for studying methods of pulse-rate estimation and PPG reconstruction using remote methods.

While less attention has been focused on SCG and BCG biosignals, recent work has considered the synthesis of whole-body BCG signals for data augmentation and from chest SCG measurements [24, 46].

6.2 Spoofing Authentication Systems

Three related lines of research have addressed the threat of spoofing attacks against biosignal authentication. In 2017, [15] demonstrated the first practical spoofing attack against an ECG authentication system. The authors targeted the Nymi band, a commercially available ECG wristband that advertised the ability to facilitate user authentication. They developed models that could translate ECG recordings collected from a portable ECG monitor ("sufficiently close" to the Nymi band) to the form collected by the band. They demonstrated how an adversary could compromise the wristband and present the spoofed data, masquerading as a victim user.

Karimian *et al.* [32] added to this threat by demonstrating various techniques to synthesize a target user's ECG trace, including using another subject's (i.e., an adversary's) ECG from the same device, or by converting a small segment of the victim user's ECG into longer traces. They moreover simulated several types of spoofing attacks against ECG authentication systems that generate victim traces which could pass verification. While promising, these attacks rely on access to supervised training data that include samples of the target user's ECG. This assumption limits its practical threat to authentication systems: if an adversary has access to a sample of the subject's ECG data to train the model, they might be able to replay that sample directly.

Other closely related work focuses on spoofing PPG-based authentication systems using other forms of PPG or remote PPG (rPPG). Karimian et al. developed a method to convert a PPG signal obtained from an arbitrary sensor location to one that looks like it was measured by a target device used for authentication; the resulting signal was then able to spoof the system [32]. Again, this requires pairs of PPG recordings from a target subject for training the translation model. More recently, attention has moved to leveraging remote PPG measurements; rPPG techniques transform a video of a subject (focused on the face or a small section of a finger) into reliable PPG waveforms. While these techniques are still an area of active research, Li *et al.* showed that rPPG data obtained from recordings of participants' faces can be used to spoof an authentication system [36, 38] trained on their regular PPGs. Li's studies were the first to show that PPG-based systems are vulnerable to remote attacks, but so far the attack has only been demonstrated for one dataset and against one example system that relies heavily on fiducial point detection (for heartbeat segmentation). Their work contributes to discrediting the security of PPG-based authentication; however, the extent of this threat to various physiological signals and forms of authentication systems has yet to be shown.

7 Conclusion

Biosignal-based authentication systems are extensively studied as a potential form of biometric authentication. In this paper, we challenge the prevailing assumption that cardiovascular biosignals are confidential and illustrate that spoofed biosignals can fool existing state-of-the-art authentication systems. Our BioForge method shows that spoofed biosignals can be generated both from video of the target and from other biosignal modalities that are plausible for an adversary to obtain from data leaks. Moreover, we show that generating fake biosignals does not require access to simultaneously recorded supervision data. We also illustrate how multimodal authentication, a commonly proposed defense, is similarly vulnerable to spoofing attacks if these multiple modalities include only cardiovascular signals. We anticipate a growing risk landscape on this front as database leaks become more likely, wearable monitoring devices become more widespread, and the performance of generative models continues to advance in the coming years. Our results highlight the inherent weakness of biosignal-based authentication systems to spoofing attacks and the need for strengthened security analysis in the design of future biometric authentication systems.

Availability

We have made our system and evaluation code available at `https://github.com/Ethos-lab/biosignal-auth-harmful` to foster future research.

Acknowledgements

We thank Hannah Lydon and Pratik Vaishnavi for their valuable help and feedback. This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) under Award EP/W014785/2, the Army Research Office under ARO W911NF-24-1-0082, and the National Science Foundation under awards ITE-2134840, CCF-1918225, and CPS-1446832. Any opinions, findings, or conclusions expressed in this material are those of the authors and do not necessarily reflect the views of the sponsors.

References

[1] Edmond Adib, Fatemeh Afghah, and John J Prevost. Synthetic ECG signal generation using generative neural networks. *arXiv preprint arXiv:2112.03268*, 2021.

[2] Edmonmd Adib, Amanda S Fernandez, Fatemeh Afghah, and John J Prevost. Synthetic ECG signal generation using probabilistic diffusion models. *IEEE Access*, 2023.

[3] Max Bagga, Hyunbae Jeon, and Alex Issokson. ECGNet: A generative adversarial network (GAN) approach to the synthesis of 12-lead ECG signals from single lead inputs. *arXiv preprint arXiv:2310.03753*, 2023.

[4] Marcel Beetz, Abhirup Banerjee, Yuling Sang, and Vicente Grau. Combined generation of electrocardiogram and cardiac anatomy models using multi-modal variational autoencoders. In *2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI)*, pages 1–4. IEEE, 2022.

[5] Laurenz Berger, Max Haberbusch, and Francesco Moscato. Generative adversarial networks in electrocardiogram synthesis: Recent developments and challenges. *Artificial Intelligence in Medicine*, page 102632, 2023. Publisher: Elsevier.

[6] Paul Bischoff. Medical breaches accounted for 422.7 million leaked records from 5,478 data breaches. `https://www.comparitech.com/blog/vpn-privacy/medical-data-breaches/`, August 2022.

[7] Dwaipayan Biswas, Luke Everson, Muqing Liu, Madhuri Panwar, Bram-Ernst Verhoef, Shrishail Patki, Chris H. Kim, Amit Acharyya, Chris Van Hoof, Mario Konijnenburg, and Nick Van Helleputte. CorNET: Deep Learning Framework for PPG-Based Heart Rate Estimation and Biometric Identification in Ambulant Environment. *IEEE Transactions on Biomedical Circuits and Systems*, 13(2):282–291, 2019.

[8] Giuseppe Boccignone, Donatello Conte, Vittorio Cuculo, Alessandro D'Amelio, Giuliano Grossi, and Raffaella Lanzarotti. An Open Framework for Remote-PPG Methods and Their Assessment. *IEEE Access*, 8:216083–216103, 2020.

[9] Joseph Bonneau, Cormac Herley, Paul C Van Oorschot, and Frank Stajano. The quest to replace passwords: A framework for comparative evaluation of web authentication schemes. In *2012 IEEE symposium on security and privacy*, pages 553–567. IEEE, 2012.

[10] Mingliang Chen, Xin Liao, and Min Wu. Pulseedit: Editing physiological signals in facial videos for privacy protection. *IEEE Transactions on Information Forensics and Security*, 17:457–471, 2022.

[11] Weixuan Chen and Daniel McDuff. Deepphys: Video-based physiological measurement using convolutional attention networks. In *Proceedings of the european conference on computer vision (ECCV)*, pages 349–365, 2018.

[12] Weixuan Chen and Rosalind W. Picard. Eliminating physiological information from facial videos. In *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, page 48–55. IEEE Press, 2017.

[13] Gerard De Haan and Vincent Jeanne. Robust pulse rate from chrominance-based rPPG. *IEEE transactions on biomedical engineering*, 60(10):2878–2886, 2013.

[14] Ruggero Donida Labati, Enrique Muñoz, Vincenzo Piuri, Roberto Sassi, and Fabio Scotti. Deep-ECG: Convolutional Neural Networks for ECG biometric recognition. *Pattern Recognition Letters*, 126:78–85, September 2019.

[15] Simon Eberz, Nicola Paoletti, Marc Roeschlin, Marta Kwiatkowska, Ivan Martinovic, and Andrea Patané. Broken hearted: How to attack ECG biometrics. In *Network and Distributed System Security Symposium 2017*. Internet Society, 2017.

[16] Matt Evans. Where is all your health data going? The Google and Fitbit scandal explained. *Tech Radar*, 1, September 2023.

[17] Miguel A García-González, Ariadna Argelagós-Palau, Mireya Fernández-Chimeno, and Juan Ramos-Castro. A comparison of heartbeat detectors for the seismocardiogram. In *Computing in Cardiology 2013*, pages 461–464. IEEE, 2013.

[18] John Gideon and Simon Stent. The way to my heart is through contrastive learning: Remote photoplethysmography from unlabelled video. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 3995–4004, 2021.

[19] Ary L Goldberger, Luis AN Amaral, Leon Glass, Jeffrey M Hausdorff, Plamen Ch Ivanov, Roger G Mark, Joseph E Mietus, George B Moody, Chung-Kang Peng, and H Eugene Stanley. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *circulation*, 101(23):e215–e220, 2000. Publisher: Am Heart Assoc.

[20] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. *Advances in neural information processing systems*, 27, 2014.

[21] Abdenour Hadid. Face biometrics under spoofing attacks: Vulnerabilities, countermeasures, open issues, and research directions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 113–118, 2014.

[22] Raia Hadsell, Sumit Chopra, and Yann LeCun. Dimensionality reduction by learning an invariant mapping. In *2006 IEEE computer society conference on computer vision and pattern recognition (CVPR'06)*, volume 2, pages 1735–1742. IEEE, 2006.

[23] Josh Hebert. *Ballistocardiography-based Authentication using Convolutional Neural Networks*. PhD Thesis, Worcester Polytechnic Institute, 2018.

[24] Sinan Hersek, Beren Semiz, Md Mobashir Hasan Shandhi, Lara Orlandic, and Omer T Inan. A globalized model for mapping wearable seismocardiogram signals to whole-body ballistocardiogram signals based on deep learning. *IEEE journal of biomedical and health informatics*, 24(5):1296–1309, 2019. Publisher: IEEE.

[25] Po-Ya Hsu, Po-Han Hsu, Tsung-Han Lee, and Hsin-Li Liu. Motion artifact resilient SCG-based biometric authentication using machine learning. In *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pages 144–147. IEEE, 2021.

[26] Po-Ya Hsu, Po-Han Hsu, and Hsin-Li Liu. Exploring seismocardiogram biometrics with wavelet transform. In *2020 25th International Conference on Pattern Recognition (ICPR)*, pages 4450–4457. IEEE, 2021.

[27] Dae Yon Hwang, Bilal Taha, Da Saem Lee, and Dimitrios Hatzinakos. Evaluation of the time stability and uniqueness in PPG-based biometric system. *IEEE Transactions on Information Forensics and Security*, 16:116–130, 2020. Publisher: IEEE.

[28] Nabil Ibtehaz, Muhammad E. H. Chowdhury, Amith Khandakar, Serkan Kiranyaz, M. Sohel Rahman, Anas Tahir, Yazan Qiblawey, and Tawsifur Rahman. EDITH: ECG biometrics aided by Deep learning for reliable Individual auTHentication. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 6(4):928–940, August 2022.

[29] Nymi Inc. Wearable identity purpose-built for the workplace. https://www.nymi.com/nymi-band.

[30] Mohit Ingale, Renato Cordeiro, Siddartha Thentu, Younghee Park, and Nima Karimian. ECG Biometric Authentication: A Comparative Analysis. *IEEE Access*, 8:117853–117866, 2020.

[31] Jinho Joo, Gihun Joo, Yeji Kim, Moo-Nyun Jin, Junbeom Park, and Hyeonseung Im. Twelve-Lead ECG Reconstruction from Single-Lead Signals Using Generative Adversarial Networks. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 184–194. Springer, 2023.

[32] Nima Karimian. How to attack PPG biometric using adversarial machine learning. In *Autonomous Systems: Sensors, Processing, and Security for Vehicles and Infrastructure 2019*, volume 11009, pages 31–37. SPIE, 2019.

[33] Walter Karlen, Srinivas Raman, J Mark Ansermino, and Guy A Dumont. Multiparameter respiratory rate estimation from the photoplethysmogram. *IEEE Transactions on Biomedical Engineering*, 60(7):1946–1953, 2013. Publisher: IEEE.

[34] Prannay Khosla, Piotr Teterwak, Chen Wang, Aaron Sarna, Yonglong Tian, Phillip Isola, Aaron Maschinot, Ce Liu, and Dilip Krishnan. Supervised contrastive learning. *Advances in neural information processing systems*, 33:18661–18673, 2020.

[35] Tahmid Latif, James Dieffenderfer, Rafael Luiz Da Silva, Edgar Lobaton, and Alper Bozkurt. Wearable Cyber-physical Systems for Biomedicine. In *Encyclopedia of Sensors and Biosensors*, pages 63–85. Elsevier, 2023.

[36] Lin Li, Chao Chen, Lei Pan, Yonghang Tai, Jun Zhang, and Yang Xiang. Hiding Your Signals: A Security Analysis of PPG-based Biometric Authentication, 2022. _eprint: 2207.04434.

[37] Lin Li, Chao Chen, Lei Pan, Jun Zhang, and Yang Xiang. SoK: an overview of PPG's application in authentication. *arXiv preprint arXiv:2201.11291*, 2022.

[38] Lin Li, Chao Chen, Lei Pan, Jun Zhang, and Yang Xiang. Video is all you need: Attacking PPG-based biometric authentication. In *Proceedings of the 15th ACM Workshop on Artificial Intelligence and Security*, pages 57–66, 2022.

[39] Xiaobai Li, Iman Alikhani, Jingang Shi, Tapio Seppanen, Juhani Junttila, Kirsi Majamaa-Voltti, Mikko Tulppo, and Guoying Zhao. The obf database: A large face video database for remote physiological signal measurement and atrial fibrillation detection. In *2018 13th IEEE international conference on automatic face & gesture recognition (FG 2018)*, pages 242–249. IEEE, 2018.

[40] Si-Qi Liu and Pong C. Yuen. A General Remote Photoplethysmography Estimator with Spatiotemporal Convolutional Network. In *2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020)*, pages 481–488, 2020.

[41] Xin Liu, Girish Narayanswamy, Akshay Paruchuri, Xiaoyu Zhang, Jiankai Tang, Yuzhe Zhang, Roni Sengupta, Shwetak Patel, Yuntao Wang, and Daniel McDuff. rppg-toolbox: Deep remote ppg toolbox. *Advances in Neural Information Processing Systems*, 36, 2024.

[42] J McKeon. 61M Fitbit, Apple users had data exposed in wearable device data breach. *Health IT Security*, 1, 2021.

[43] Patrick E McSharry, Gari D Clifford, Lionel Tarassenko, and Leonard A Smith. A dynamical model for generating synthetic electrocardiogram signals. *IEEE transactions on biomedical engineering*, 50(3):289–294, 2003. Publisher: IEEE.

[44] Pietro Melzi, Ruben Tolosana, and Ruben Vera-Rodriguez. ECG Biometric Recognition: Review, System Proposal, and Benchmark Evaluation. *IEEE Access*, 11:15555–15566, 2023.

[45] Kevin Musgrave, Serge J. Belongie, and Ser-Nam Lim. PyTorch Metric Learning. *ArXiv*, abs/2008.09164, 2020.

[46] Mohammad Nikbakht, Asim H Gazi, Jonathan Zia, Sungtae An, David J Lin, Omer T Inan, and Rishikesan Kamaleswaran. Synthetic seismocardiogram generation using a transformer-based neural network. *Journal of the American Medical Informatics Association*, 30(7):1266–1273, 2023.

[47] Ikenna Odinaka, Po-Hsiang Lai, Alan D. Kaplan, Joseph A. O'Sullivan, Erik J. Sirevaag, and John W. Rohrbaugh. ECG Biometric Recognition: A Comparative Analysis. *IEEE Transactions on Information Forensics and Security*, 7(6):1812–1824, 2012.

[48] Ozan Oktay, Jo Schlemper, Loic Le Folgoc, Matthew Lee, Mattias Heinrich, Kazunari Misawa, Kensaku Mori, Steven McDonagh, Nils Y Hammerla, Bernhard Kainz, Ben Glocker, and Daniel Rueckert. Attention U-Net: Learning where to look for the pancreas. In *Medical Imaging with Deep Learning*, 2018.

[49] Marco A. F. Pimentel, Alistair E. W. Johnson, Peter H. Charlton, Drew Birrenkott, Peter J. Watkinson, Lionel Tarassenko, and David A. Clifton. Toward a Robust Estimation of Respiratory Rate From Pulse Oximeters. *IEEE Transactions on Biomedical Engineering*, 64(8):1914–1923, 2017.

[50] Aditya Singh Rathore, Zhengxiong Li, Weijin Zhu, Zhanpeng Jin, and Wenyao Xu. A survey on heart biometrics. *ACM Computing Surveys (CSUR)*, 53(6):1–38, 2020. Publisher: ACM New York, NY, USA.

[51] Attila Reiss, Ina Indlekofer, Philip Schmidt, and Kristof Van Laerhoven. Deep PPG: Large-scale heart rate estimation with convolutional neural networks. *Sensors*, 19(14):3079, 2019. Publisher: MDPI.

[52] Nikita Samarin and Donald Sannella. A Key to Your Heart: Biometric Authentication Based on ECG Signals, June 2019. arXiv:1906.09181 [cs, eess].

[53] Pritam Sarkar and Ali Etemad. Cardiogan: Attentive generative adversarial network with dual discriminators for synthesis of ECG from PPG. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 488–496, 2021.

[54] Philip Schmidt, Attila Reiss, Robert Duerichen, Claus Marberger, and Kristof Van Laerhoven. Introducing wesad, a multimodal dataset for wearable stress and affect detection. In *Proceedings of the 20th ACM international conference on multimodal interaction*, pages 400–408, 2018.

[55] Florian Schroff, Dmitry Kalenichenko, and James Philbin. Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 815–823, 2015.

[56] Apple Platform Security. Passcodes and passwords. `https://support.apple.com/guide/security/passcodes-and-passwords-sec20230a10d/web`, February 2021.

[57] Mohammad Soleymani, Jeroen Lichtenauer, Thierry Pun, and Maja Pantic. A multimodal database for affect recognition and implicit tagging. *IEEE transactions on affective computing*, 3(1):42–55, 2011.

[58] Zhaodong Sun and Xiaobai Li. Privacy-phys: Facial video-based physiological modification for privacy protection. *IEEE Signal Processing Letters*, 29:1507–1511, 2022.

[59] Wencheng Yang, Song Wang, Jiankun Hu, Guanglou Zheng, and Craig Valli. Security and accuracy of fingerprint-based biometrics: A review. *Symmetry*, 11(2):141, 2019.

[60] Zitong Yu, Xiaobai Li, and Guoying Zhao. Remote photoplethysmograph signal measurement from facial videos using spatio-temporal networks. In *30th British Machine Visison Conference: BMVC 2019. 9th-12th September 2019, Cardiff, UK*. The British Machine Vision Conference (BMVC), 2019.

[61] Peng Zhang, Mingfeng Jiang, Yang Li, Ling Xia, Zhefeng Wang, Yongquan Wu, Yaming Wang, and Huaxiong Zhang. An efficient ECG denoising method by fusing ECA-Net and CycleGAN. *Mathematical Biosciences and Engineering*, 20(7):13415–13433, 2023.

[62] Xianwen Zhang, Yandong Zhang, Liyan Zhang, Heng Wang, and Jintian Tang. Ballistocardiogram based person identification and authentication using recurrent neural networks. In *2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, pages 1–5. IEEE, 2018.

[63] Yu-He Zhang and Saeed Babaeizadeh. Synthesis of standard 12-lead electrocardiograms using two-dimensional generative adversarial networks. *Journal of Electrocardiology*, 69:6–14, 2021. Publisher: Elsevier.

[64] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pages 2223–2232, 2017.