Iris Presentation Attack: Assessing the Impact of Combining Vanadium Dioxide Films with Artificial Eyes

Darshika Jauhari, Renu Sharma, Cunjian Chen, Nelson Sepulveda, Arun Ross
Michigan State University
{jauharid,sharma90,cunjian,sepulve6,rossarun}@msu.edu

Abstract

Iris recognition systems, operating in the near infrared spectrum (NIR), have demonstrated vulnerability to presentation attacks, where an adversary uses artifacts such as cosmetic contact lenses, artificial eyes or printed iris images in order to circumvent the system. At the same time, a number of effective presentation attack detection (PAD) methods have been developed. These methods have demonstrated success in detecting artificial eyes (e.g., fake Van Dyke eyes) as presentation attacks. In this work, we seek to alter the optical characteristics of artificial eyes by affixing Vanadium Dioxide (VO_2) films on their surface in various spatial configurations. VO_2 films can be used to selectively transmit NIR light and can, therefore, be used to regulate the amount of NIR light from the object that is captured by the iris sensor. We study the impact of such images produced by the sensor on two state-of-the-art iris PA detection methods. We observe that the addition of VO_2 films on the surface of artificial eyes can cause the PA detection methods to misclassify them as bonafide eyes in some cases. This represents a vulnerability that must be systematically analyzed and effectively addressed.

1. Introduction

Iris recognition systems use the texture of the iris in order to recognize individuals [12]. A typical iris recognition system operates in the near-infrared (NIR) spectrum. There are several reasons for using NIR sensors to acquire an image of the iris: (a) NIR illumination is non-invasive and, unlike visible spectrum lighting, does not excite the pupil; and (b) NIR illumination can be used to elicit the texture of even dark-colored irides since it can penetrate the multi-layered iris more effectively than visible spectrum lighting. Despite their success in a number of real-world applications, iris systems are vulnerable to number of attacks [36], including presentation attacks (PAs) [2, 10]. A presentation attack occurs when an adversary presents a fake or altered trait to the sensor in order to obfuscate their own identity, spoof another person's identity or to create a virtual identity. The biometric characteristics or materials used to launch a presentation attack are referred to as Presentation Attack Instruments (PAI). Examples of PAIs in the case of the iris modality include printed iris images [9, 12, 18, 33], plastic, glass, or doll eyes [18, 26], cosmetic contact lenses [4, 22, 34, 44], a video display of an eye image [10, 11, 35], cadaver eyes [10, 11, 30], robotic eye models [24] holographic eye images [32] and synthesized irises [41]. A few examples of iris PAIs are shown in Fig. 1.

Among all the attacks described above, iris pattern printed on a paper is perhaps one of the easiest ones. The efficacy of this type of attack depends on a number of factors including the choice of printer (inkjet or laserjet), paper (matte, glossy, photographic, butter, white, recycled or cardboard), resolution (600 or 1,200 dpi), image type (grayscale or color), configuration (with or without pupil cutout), and sensing device (IrisPass, IrisAccess, or IrisGuard). In the LivDet-Iris 2013 competition [47], a combination of two different printers, two commercial iris sensors and matte paper were used. Later, the dataset was extended in the LivDet-Iris 2015 [48], LivDet-Iris 2017 [46], LivDet-Iris 2020 [11] and LivDet-Iris 2023 [41] competitions by including more variations in the resolution, contrast and texture of the printed irides. In [11], various add-ons were applied to printed paper, including transparent domes and textured as well as clear contact lenses.

The use of cosmetic contact lens as a PAI poses an even greater challenge than the prints, since the former has significantly more manufacturers, brands, and colors [11, 46]. In [47], 22 types of patterned contact lenses were collected, which was later increased to 57 types with different texture patterns [39]. In [48], 20 different varieties of cosmetic contacts were used to generate iris PA samples. This was later extended by adding samples from the Notre Dame subset, which contained five different brands of textured contact lenses, and the IITD-WVU subset, which contained four manufacturers and six colors [46]. In [11], three different brands of cosmetic contacts (Johnson & Johnson, Ciba

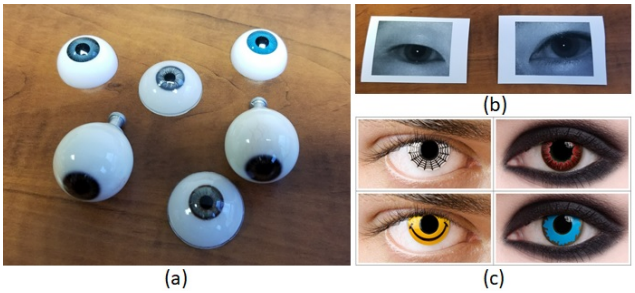


Figure 1. Examples of PAI used to launch presentation attacks (PAs) on the iris modality: (a) plastic eyes, (b) printed images, and (c) cosmetic contacts [20].

Vision, and Bausch & Lomb) were captured using the LG IrisAccess 4000 and IrisGuard AD100 under various illumination setups (two different illuminants in LG4000 and six different illuminants in AD100).

In addition to the printed and cosmetic contact attacks, the replay or display attack also poses a challenge. In this type of attack, a previously captured iris image or video is presented to an iris sensor via a display media. However, most modern computers, laptops and mobile phone screens do not necessarily emit NIR light. So this type of attack has been predominantly tested on iris systems operating in the visible spectrum [10, 35]. However, the display of certain Kindle devices emit NIR light and, consequently, can be more easily imaged using NIR iris sensors [11, 19].

A plastic or prosthetic eye is a highly viable PAI, but has not been as extensively explored in the literature, unlike some of the other PAs. Variants of such artificial eyes can be designed using different materials like Poly Methyl Meta Acrylate [26], glass or plastic. Lee et al. [26] created three different-colored artificial eyes (blue, gray and dark brown). Sun et al. [39] selected 40 different subjects iris images from the UPOL database [29] and printed them on plastic eyeball models. Hoffman et al. [19] collected images of fake eyes using three different plastic/glass eye brands and 10 distinct colors. Das et al. [11] presented two different types of fake eyes: Van Dyke Eyes (which have higher iris quality details) and Scary Eyes (plastic fake eyes with a simple pattern on the iris region). They also presented add-ons for fake eyes, such as textured and clear contacts.

As stated above, attacks using artificial eyes made of glass or plastic have not been heavily studied [7, 18, 26, 49]. The goal of this work is to leverage recent developments in material science to test the robustness and measure the susceptibility of iris systems to such type of spoof attacks. Specifically, we are interested to produce spoofs which are fabricated by affixing chemically modified films on these artificial eyes. [1] The iris is generally imaged in the NIR spectrum; accordingly, we have attempted to use NIR-sensitive Vanadium dioxide (VO_2) films to generate these spoofs. VO_2 is a typical thermochromic material that has been widely studied as smart coatings for buildings fenestrations [3, 14–16, 50]. The synthesis of VO_2 films has been reported briefly in the literature, and its manufacturing is easy and cost effective. It is an advantage to use VO_2 for our work as it is deposited on a glass substrate and its handling is smooth. A further advantage is its low toxicity and high stability at room temperature conditions for such a short period of usage. These films show transmittance drop, close to a temperature of 68°C, in the NIR region [1, 27, 28, 31, 43]. This implies that at temperatures below 68°C, the film allows maximum light to pass through, but as the temperature increases above 68°C, the film behaves in a completely different manner, only allowing a portion of light to pass (Fig. 2). This change in behavior of the film allows us to image the fake eyes in 2 different arrangements. Thus, in order to generate an effective spoof, we used the VO_2 coated and uncoated (blank) films in varied configurations on the fake glass eye. This is a unique kind of presentation attack, which combines multiple attack modes and that has never been attempted before.

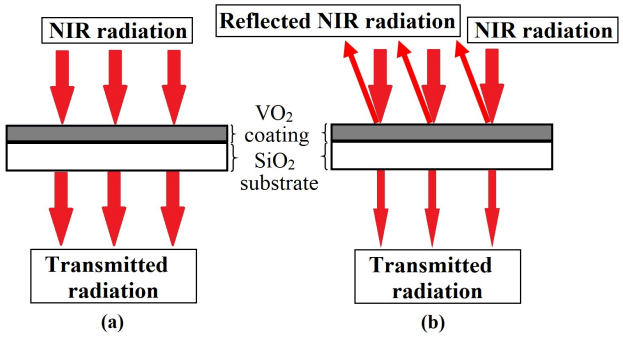


Figure 2. Schematic of thermochromic behavior of VO_2 films (a) below and (b) above 68°C (critical temperature).

2. Experiment and Setup

2.1. Fabrication

Vanadium dioxide (VO_2) thin films were deposited by pulsed laser deposition over fused silica (SiO_2) substrates (2" in diameter, 250 μm thick), following a process similar to what has been described in the past [8, 13]. The substrate was heated to 600° C in a vaccum chamber with a background pressure of 1x10^{-7} Torr. After reaching this setpoint, oxygen and argon gas was introduced to the chamber, while the chamber pressure was controlled through a butterfly valve to be at 35 mTorr. At this point, laser pulses from an excimer laser (wavelength λ=248 nm, 20 ns pulse duration, and $\sim$ 4 J/cm^2 fluence) ablated a metallic vanadium target, and the pressure was controlled to 35 mTorr. A total

[1] In principle, it can be used on other types of PA artifacts.

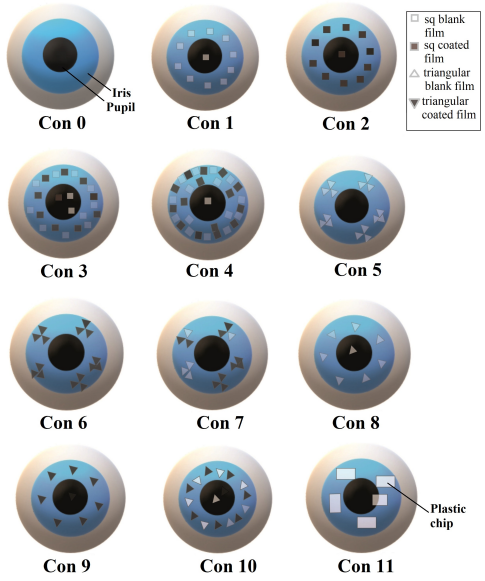


Figure 3. Diagrammatic representation of the various patterns in which the blank films and VO_2 coated films were arranged on the fake eyes.

of 320,000 pusles resulted in a 280 nm thick VO_2 thin film over the SiO_2 substrate. After VO_2 deposition, the 2" sample was diced into squares and triangles (2 mm $\times$ 2 mm). Another blank identical SiO_2 substrate (i.e., with no VO_2 thin film deposited) was also diced with the same dimensions. The resulting samples were multiple bare SiO_2 and VO_2-coated 2 mm $\times$ 2 mm "pixels".

Our aim in this work is to fabricate fake eyes (Van Dyke eyes, made of soft glass) with different patterns of films on it. This patterning was based on different factors such as shape, type and orientation of the films (Fig. 3). To achieve this, VO_2 coated films and blank films were fixed on the fake eyes in 11 different geometrical configurations as described below. For the first set of images (Con 0), the naked Van Dyke eyes were imaged in different angular and lighting conditions (Fig. 5 (a-j)). To achieve this, the Van Dyke eyes were first attached to fake Halloween glasses using double-sided tape. The user then mounted these glasses and approached the iCAM 7100S iris sensor for imaging. This triggered the activation of the sensor, as indicated by the appearance of an orange dot on the mirror. Now, at the correct distance, once the orange dot is aligned over the bridge of the nose, it turns green, and both the irides are acquired. This process was subsequently repeated by using the tilt up/down button on the sensor unit. Multiple other images (Fig. 5 i (f-j)) were also captured by focusing some extra light (120 V, GE-IR table lamp) on the fake eyes (mounted on the user). For Con 1, a few blank square films were removed from the whole blank diced lot using a pair of tweezers. These films were then carefully stuck on the fake iris in a circular pattern (Fig. 5 ii (a-j)), with a couple of them on the pupil portion of the fake eyes. This patterned eye was imaged using the same process as stated above. One additional change was the *in situ* heating of the films using the IR lamp. The film was heated for 2.5 min to reach a temperature of 80°C, and a picture was acquired immediately. This was done to appreciate the difference in image and PA scores with and without heating (Fig. 5 ii (j)).

A similar procedure was adopted for Con 2, where VO_2 coated square films were used instead of blank films (Fig. 5 (a-j)). Again, for Con 3, VO_2 coated and blank film were arranged alternately on the iris and pupil of the fake eyes (Fig. 5 iv (a-j)). Con 4 was designed by closely placing the coated and blank films in 2 rings on the iris, with one blank film on the pupil (Fig. 5 v (a-j)). Con 5 was fabricated by choosing triangular blank films. These triangular films were placed in a group of 3's to form a flower-like pattern (Fig. 5 vi (a-j)). Similarly, VO_2 coated films were arranged in triangles of 3, forming flower-like pattern for Con 6 (Fig. 5 vii (a-j)). The Con 7 was designed using both coated and uncoated triangular films in grouping of 3 on the fake eyes (Fig. 5 viii (a-j)). Con 8, 9, and 10 were fabricated by placing triangular-shaped blank; triangular-shaped coated; and triangular-shaped blank and coated on the fake iris, respectively. (Fig. 5 ix-xi (a-j)). For the last configuration (Con 11), transparent plastic chips were stuck on the Van Dyke eyes (Fig. 5 (a-j)).

We captured 10 images of an eye for each configuration, resulting in a database of 120 samples. These images were taken with the help of six different subjects. More than one subject was used to eliminate any subject-specific errors during data collection. These images were then assessed using two state-of-the-art PA detectors: D-NetPAD and IrisTL-PAD (Fig. 4). Both PA detection methods produce a single-valued PA score. These PA scores range from 0 to 1, where 1 indicates a PA sample and 0 indicates a bonafide or live iris.

2.2. Iris Presentation Attack Detection Methods

The two iris PA detection algorithms that are utilized to assess the vulnerability of adhering VO_2 films on artificial eyes are described below. They both are based on deep neural architectures.

D-NetPAD: D-NetPAD [38][2] is based on a densely connected convolutional neural network where each layer connects to every other layer in a feed-forward fashion. Its base architecture is DenseNet-121 [21], which consists of 121 convolutional layers in a series of four Dense Blocks and three Transition Layers. A detailed description of the architecture is provided in [21]. To detect iris PA, the iris region is first cropped from the ocular image and resized to 224 $\times$

[2]https://github.com/iPRoBe-lab/D-NetPAD

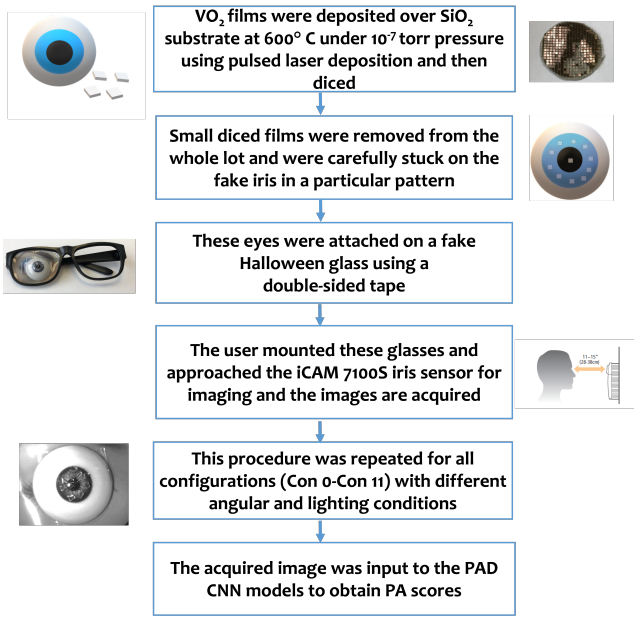


Figure 4. Step-wise procedure for fabrication of VO_2 modified fake eye starting from its deposition to PA score procurement [23].

224. The cropped and resized iris region is then input to the D-NetPAD, which produces a presentation attack (PA) score between 0 and 1. A flowchart of the D-NetPAD is shown in Fig. 7. It utilizes a pre-trained ImageNet model to initialize weights and fine-tune them with iris PA samples. Fine-tuning has been performed with a proprietary dataset and the NDCLD2015 [40] dataset. The proprietary dataset consists of 6,610 bonafide irides and 3,839 PA samples. PA samples include 130 Kindle replay attacks, 1,651 printed eyes, 1,537 plastic eyes, and 521 cosmetic contact lenses. From NDCLD2015 [40], 2,236 cosmetic contact lens images are used for training.

IrisTL-PAD: IrisTL-PAD [4–6] operates on the cropped iris regions and offers a simple and fast solution for PA detection. It also utilizes the pre-trained ImageNet model to initialize the weights and then performs transfer learning. First, an off-line trained iris detector was used to obtain a rectangular region encompassing the outer boundary of the iris. Then, the iris region was automatically cropped based on the estimated rectangular coordinates. Finally, the cropped iris region was input to a CNN (ResNet50) to train the iris PA detection model (Fig. 6). The training was fine-tuned on an existing ImageNet model, by leveraging extensive data augmentation schemes. The IrisTL-PAD model was trained on 9,072 bonafide images and 7,352 PA images as summarized in Table 1.

Both PAD algorithms are state-of-the-art methods that resulted in the best performance another proprietary dataset. The data were collected using the iCAM7000 NIR sensor from 1,315 subjects. A total of 3,315 iris images were acquired, out of which 2,963 were bonafide irides and 352 were PA samples. PAs in the dataset include two types of VanDyke eyes and 10 different types of cosmetic contact lenses. The D-NetPAD and IrisTL-PAD methods resulted in a True Detect Rate (TDR) of 98.58% and 92.61%, respectively, at a False Detect Rate (FDR) of 0.2%.[3] The TDR denotes the fraction of PA samples that were correctly classified, while the FDR denotes the fraction of bonafide samples that were incorrectly classified as PA samples. **In addition, both PAD algorithms were the best performing algorithms in the LivDet-Iris 2020 competition [11].**

3. Evaluation and Results

To determine whether the designed configurations of the Vanadium dioxide films on artificial eyes can be used to attack the system or not, we compared their PA scores to that of bonafide, i.e., live human eyes. The PA score for a live human eye ranges from 0.0-0.5 for IrisTL-PAD and 0.0-0.4 for D-NetPAD. As depicted in Table 2, Cons 0, 1 and 2 showed PA scores more than the threshold value (0.5 for IrisTL-PAD, 0.4 for D-NetPAD) for both the algorithms. This indicates that these configurations were detected as spoofs by both the algorithms. However, as we move onto Con 3, the PA scores dip below the threshold for all 10 images for IrisTL-PAD and 6 images for D-NetPAD. **This is a successful configuration that fools the PA detection systems and passes as a live or bonafide eye** (Fig. 9). Con 4, which has VO_2 coated and blank films in 2 concentric circles in the iris region, has an attack success rate of 40% for IrisTL-PAD and 10% for D-NetPAD (Fig. 9). Attack success rate was calculated as the percentage of attacks below the given threshold value. A configuration having success rate of 50% or more was considered to be a successful attack. Con 5 which has triangular blank films arranged in a group of 3, shows attack success rate of 50% for IrisTL-PAD and 0% for D-NetPAD. Con 6 images have slightly lower chances of working as a spoof (rates: 40% IrisTL-PAD and 10% D-NetPAD). Con 7 on the other hand has higher chances of passing as a live eye, with attack success percentage at 90% for IrisTL-PAD and 50% for D-NetPAD. Con 8 has 30% success for IrisTL-PAD, and 10% for D-NetPAD. Con 9 too has a lower chance to deceive the system (only 20% success for IrisTL-PAD and none for D-NetPAD). Cons 10 (10% success rate for D-NetPAD) and 11 also do not pose a threat to the two PA detection methods. One point to be mentioned is that the heating of VO_2 films upto a temperature of 80°C does not bring a significant change in the PA scores. Cons 1(j), 5(i), 9(i), 9(j) and

[3]ISO/IEC 30107-3:2023 specifies Attack Presentation Classification Error Rate (APCER) and Bonafide Presentation Classification Error Rate (BPCER) as evaluation metrics for PAD. TDR is 1−APCER, and FDR is the same as BPCER.

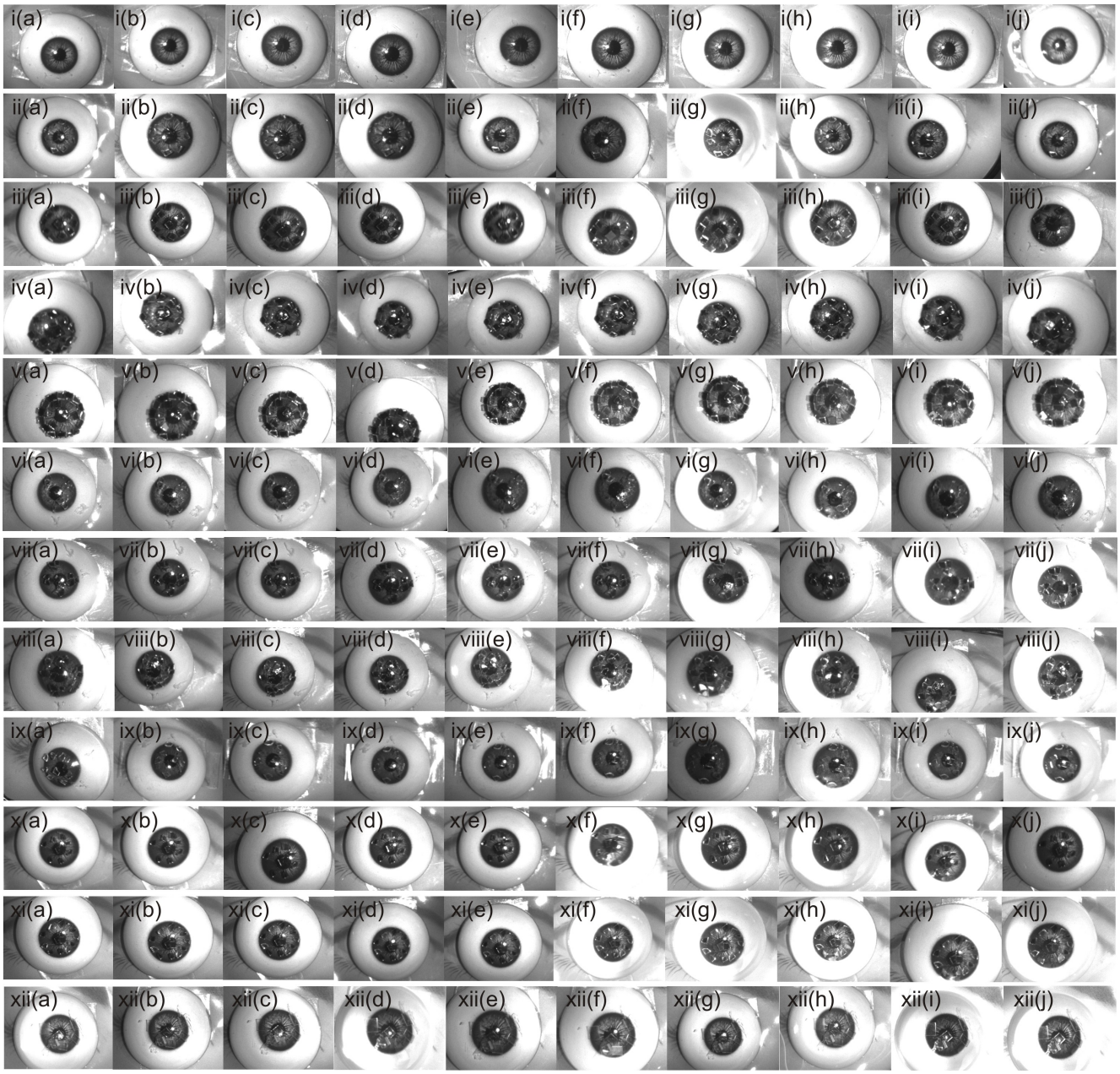


Figure 5. Images with various configurations of films as captured by iCAM 7100S. Images (a)-(j) represent a particular configuration taken in different angular, lighting and temperature conditions, sequentially (for a detailed label refer to Table 2). Images (i)-(xii) represent configurations Con 0 to Con 11.

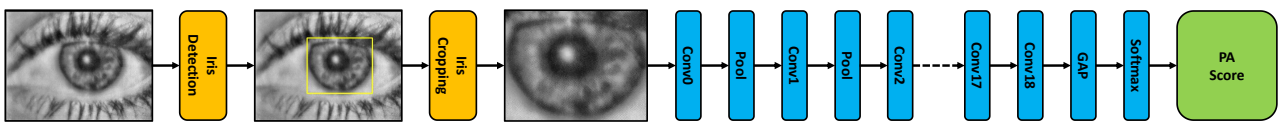


Figure 6. Flowchart depicting the IrisTL-PAD method.

10(j) show high PA scores even when the films are above the critical temperature, and reflecting most of the light. This is

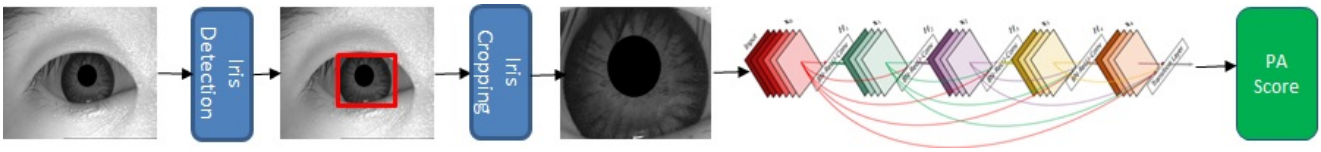


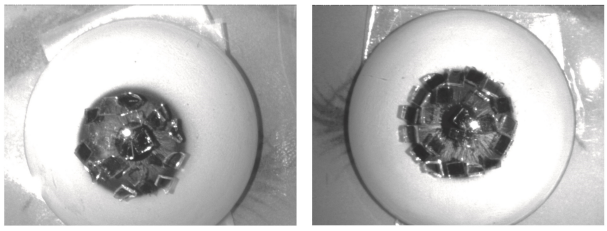
Figure 7. Flowchart of the D-NetPAD algorithm. Iris region (red box) is detected and cropped from the ocular image and input to the D-NetPAD architecture. The base architecture used in D-NetPAD is DenseNet-121 [21]. It produces a single PA score which determines whether an input image is a bonafide or a PA.

Table 1. A summary of datasets used to train IrisTL-PAD.

Dataset	Total	Live	Print	Contact Lenses	Artificial Eye
LivDet-Iris 2017-IIT-WVU [46]	1,750	750	-	1000	-
LivDet-Iris 2017-NotreDame [46]	1,200	600	-	600	-
LivDet-Iris 2017-Warsaw [46]	4,513	1,844	2,669	-	-
BERC-Iris-Fake [25]	4,598	2,778	1,600	140	80
CASIA-Iris-Interval [17]	740	-	-	740	-
Private Dataset	3,623	3,100	6	334	183
Combined	**16,424**	**9,072**	**7,352**		

an indication that the thermochromic behaviour of the film does not play a big role in deceiving the system as far as our experimental protocol is concerned.

In summary, our preliminary observations indicate that Cons 3, 5 and 7 have high presentation attack success rates. The high presentation attack success rate for Cons 3, 5 and 7 could be due to the kind of geometrical arrangement of films on them. Note that Cons 3 and 4 have a similar type of arrangement for both the films (coated and uncoated), but Con 3 has more space between the films (Fig. 8). This causes a change in the captured iris pattern and impacts the PA detection methods.



Con 3 (iv (g)) **Con 4 (v(e))**

Figure 8. Comparison of the geometrical arrangement of Cons 3 and 4. Con 3 has VO_2 coated and blank films arranged in an alternate manner, but in no particular geometrical pattern all over the artificial eye. Con 4 on the other hand, has these films arranged in 2 concentric circles inside the iris region and 2 blank films on the pupil of the fake eyes.

The result was further visually analyzed by generating "heatmaps" using Gradient-weighted Class Activation Mapping (Grad-CAM) [37]. Grad-CAM produces a coarse localization map highlighting the salient regions in an image that were used by the network in order to generate its inference. Fig. 10 presents the "heatmaps" for configurations that were unsuccessful (Con 0) as well as those that were successful (Con 3 and Con 7) in defeating the D-NetPAD algorithm. The red regions indicate high activation, whereas the blue regions represent low activation when inferring the final decision (i.e., bonafide or PA). The first row of Fig. 10 shows the heatmaps of Con 0 images, where the high activation region is at the pupillary zone of the printed iris pattern

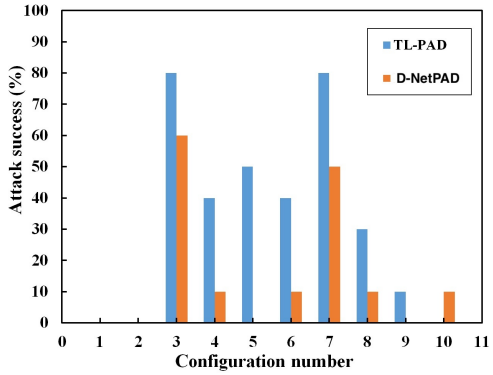


Figure 9. Presentation attack success rate across all configurations. Cons 3, 5 and 7 have a higher success rate (against the IrisTL-PAD method) compared to other configurations. This suggests that the films may have to be strategically placed on the fake iris pattern in order to defeat an iris PA detection system.

Table 2. Detailed table of PA scores for each image captured across all 12 configurations. Red colored cells represent PA scores for IrisTL-PAD which are less than or equal to its threshold value (0.5). Yellow colored cells represent PA scores for D-NetPAD which are less than or equal to its threshold value (0.4). Orange fonted numbers represent images taken with extra lightning conditions. Blue fonted numbers represent PA scores of images taken after heating the films.

Configuration		PA scores									
		a	b	c	d	e	f	g	h	i	j
Con 0 (Van Dyke eyes)	IrisTL-PAD	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
	D-NetPAD	1.00	0.93	0.95	0.97	0.87	0.93	0.92	0.89	0.87	0.66
Con 1 (Blank sq all over)	IrisTL-PAD	0.99	0.98	0.99	0.99	1.00	1.00	0.91	0.97	0.98	**1.00**
	D-NetPAD	0.66	0.57	0.58	0.60	0.57	0.64	0.57	0.59	0.56	**0.65**
Con 2 (VO_2 sq all over)	IrisTL-PAD	0.76	0.70	0.90	0.96	0.94	0.94	0.99	0.92	0.91	0.99
	D-NetPAD	0.43	0.55	0.54	0.56	0.47	0.47	0.51	0.49	0.53	0.61
Con 3 (VO_2 and Blank sq alternate)	IrisTL-PAD	0.40	0.05	0.10	0.42	0.47	0.10	0.01	0.10	**0.23**	**0.09**
	D-NetPAD	0.35	0.48	0.37	0.38	0.35	0.43	0.38	0.39	**0.41**	**0.41**
Con 4 (VO_2 and Blank sq alternate ring)	IrisTL-PAD	0.16	0.92	0.35	0.89	0.70	0.89	0.25	0.73	**0.13**	**0.60**
	D-NetPAD	0.42	0.43	0.43	0.43	0.42	0.42	0.44	0.43	**0.43**	**0.40**
Con 5 (Blank triangle in flower)	IrisTL-PAD	0.26	0.87	0.51	0.11	0.07	0.93	0.52	0.11	**0.95**	**0.13**
	D-NetPAD	0.51	0.54	0.53	0.50	0.52	0.52	0.53	0.62	**0.52**	**0.46**
Con 6 (VO_2 triangle in flower)	IrisTL-PAD	0.88	0.89	0.19	0.02	0.80	0.99	0.55	0.02	**0.72**	**0.06**
	D-NetPAD	0.53	0.47	0.49	0.48	0.46	0.40	0.44	0.46	**0.54**	**0.48**
Con 7 (Blank and VO_2 triangle in flower)	IrisTL-PAD	0.08	0.07	0.55	0.23	0.01	0.03	0.26	0.44	0.14	**0.23**
	D-NetPAD	0.31	0.35	0.43	0.37	0.47	0.44	0.39	0.43	0.38	**0.37**
Con 8 (Blank triangle all over)	IrisTL-PAD	0.24	0.92	0.90	0.70	0.73	1.00	0.60	0.99	0.09	0.01
	D-NetPAD	0.53	0.63	0.40	0.47	0.59	0.54	0.58	0.63	0.61	0.50
Con 9 (VO_2 triangle all over)	IrisTL-PAD	0.98	0.86	0.46	0.04	0.93	0.99	0.94	0.99	**0.76**	**0.70**
	D-NetPAD	0.67	0.68	0.64	0.61	0.60	0.75	0.62	0.61	**0.65**	**0.65**
Con 10 (VO_2 and Blank triangle all over)	IrisTL-PAD	0.59	0.89	0.79	0.97	1.00	0.96	0.94	0.93	0.67	**0.83**
	D-NetPAD	0.48	0.44	0.44	0.40	0.42	0.50	0.44	0.57	0.47	**0.47**
Con 11 (Plastic chips sq all over)	IrisTL-PAD	0.99	0.98	0.96	0.77	0.98	0.87	0.99	0.99	0.99	0.99
	D-NetPAD	0.50	0.51	0.58	0.51	0.53	0.51	0.49	0.44	0.55	0.53

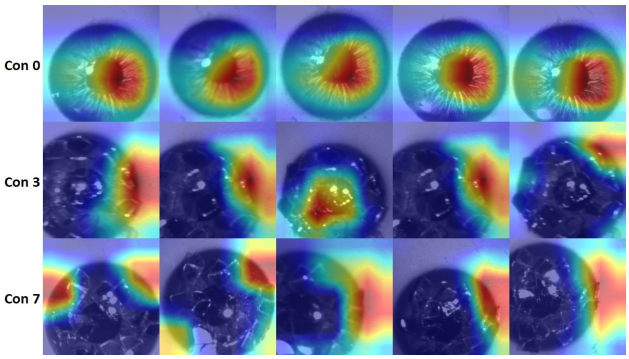


Figure 10. Grad-CAM [37] heatmaps of images corresponding to Con 0, Con 3 and Con 7 configurations. Con 0 has low PA success rate, whereas Con 3 and Con 7 have a high PA success rate. Red-colored regions represent highly focused region by the D-NetPAD. The blue region represents low priority regions. These regions help in making the final decision about being a bonafide or a PA.

of the fake eye. The other two rows of Fig. 10 correspond to Cons 3 and 7 (high presentation attack success rate). The high activation regions in these two rows of images are distributed throughout the iris pattern.

We hypothesize that the combination of VO_2 and blank films when placed on the Van Dyke eyes interfered with the iris pattern inscribed on the fake eye. This presumably resulted in a pattern that was never seen by the algorithm during training. As a result, the focus shifted away from the iris pattern (see the last two rows of Fig. 10), resulting in PA scores that were in the vicinity of the threshold (0.40). The chances of mis-classification seems to have increased with an increase in the density of the VO_2 and blank films. The VO_2 films appear to obscure the underlying pattern due to their special optical property with NIR illumination, whereas the blank films distort the pattern. Thus, Cons 3 and 7, which have a high concentration of VO_2 and blank films (Fig. 5), show high misclassification rates.

After Grad-CAM visualization, which utilizes back-propagation, we also visualized the features fed into the architecture in the forward direction for the final decision. The features were extracted from the penultimate layer (just before the fully connected layer) of D-NetPAD and reduced to two dimensions using t-Distributed Stochastic Neighbor Embedding (t-SNE) [42]. t-SNE plots are shown in Fig. 11, where the green and blue data points represent bonafide and

fake eye images from the proprietary dataset, respectively. The red data points (Fig. 11) represent configurations with a high PA success rate (Cons 3, 4, 6, 7, 8), whereas the pink data points represent configurations with a low PA success rate (Cons 0, 1, 2, 5, 9, 10, 11). Fig. 11 shows the distribution of the configurations departing from that of the fake eyes, as well as being spaced out. This divergence further substantiates the effectiveness of using VO_2 films in performing iris presentation attacks.

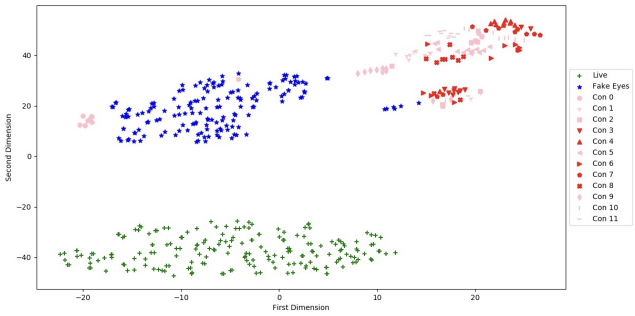


Figure 11. t-SNE plot of the D-NetPAD algorithm on bonafide (green) and fake eye (blue) images on the proprietary dataset. It also shows the t-SNE of all the configurations considered in this work. Red-colored data points represent configurations with a high PA attack success rate (Con 3, 4, 6, 7, 8), whereas pink represents configurations with a low PA attack success rate (Con 0, 1, 2, 5, 9, 10, 11). The distribution of the configurations is observed to be substantially different from that of the fake eyes, suggesting the novel nature of the attack.

4. Role of Coated and Blank Films

It is clear from the experiments conducted in this work that introduction of the VO_2 films along with blank films made a difference in the optical properties of the fake eyes. This change triggered a shift of focus of the PAD algorithms away from the iris portion, labelling them as bonafide. To check the function of VO_2 films, we carried out some preliminary experiments using metal coated films. These films when used alternately with blank films on fake eyes lowered the PA scores. This clearly shows that the new films (metal), just like VO_2 ones, caused changes in the optical properties of the fake eyes. These films worked as an attack only when used with blank films but not just by themselves. This suggests that a patch-based configuration is able to fool the PAD algorithm and pass as a genuine eye. But extensive experiments have to be carried out with the new films which can help us strengthen this hypothesis.

5. Mitigation Measures

We performed another experiment to find a potential solution for the misclassification of VO_2 and blank films coated fake eyes as bonafides. We utilized all samples (=10) from a subset of configurations defined in this paper to perform incremental training of the D-NetPAD algorithm. The configurations used for the training were 1, 2, 5, 9 and 11 as they were correctly classified as PAs by the D-NetPAD. For incremental training, we fine-tuned the D-NetPAD model with the selected samples. Next, we recomputed the PA scores of all samples, including those pertaining to Cons 3, 4, 6, 7, 8, 10 which were not used for training. We observed that all the samples were now correctly classified as PAs. The experiment shows that incremental training with only a few samples can extend the discriminative power of the model in detecting such new attacks.

6. Discussion and Future Work

In this work, we assessed the possibility of combining Vanadium dioxide films with artificial glass eyes in order to create PAs that can potentially evade presentation attack detection. VO_2 films can be used to selectively regulate NIR transmission thereby causing such artificial eyes to be misclassified as bonafide samples. Our experimental results suggest that the placement of these films in specific configurations can indeed confound a PA detection system.

Having said that, there are some ways to detect these types of attacks: (a) Patch-based PAD: The Vanadium dioxide films used to create the PAs, modified the iris texture in configurations that can be described by local patches (see Fig. 5). Both IrisTL-PAD and D-NetPAD solutions extract global features from the cropped iris images. By using local regions for PA determination, it is likely that patches which are not modified with VO_2 films would produce high PA scores. Hence, averaging the PA scores across individual patches can increase the robustness of these PAD solutions to the proposed attack [19]. (b) One-class Classification: It is difficult to model the distribution of every unknown or unseen PAs. To tackle such PAs, a one-class classifier concept can be leveraged where only bonafide distribution is required to create the PAD model [45].

Future work will explore the application of Vanadium dioxide films to generate PA artifacts that can be used for training and increase the robustness of existing PAD methods. Further, their thermochromic behavior can possibly be used to design new iris hardware for PA detection. We will also study the efficacy of this attack on other PAD techniques. In this work, the attack has been studied in the context of PAD only; future work will involve analyzing the impact on the iris recognition method also.

***Ethical Implications**: The goal of this paper was to alert researchers and practitioners to potential attacks and provide a preliminary solution to detect them. However, it should not be misused to launch an attack against iris recognition systems.*

References

[1] A. S. Barker, H. W. Verleur, and H. J. Guggenheim. Infrared optical properties of vanadium dioxide above and below the transition temperature. *Physical Review Letters (PRL)*, 17:1286–1289, 1966. 2

[2] Aidan Boyd, Jeremy Speth, Lucas Parzianello, Kevin Bowyer, and Adam Czajka. State of the art in open-set iris presentation attack detection. *arXiv*, 2208.10564, 2022. 1

[3] Tianci Chang, Xun Cao, Liv R. Dedon, Shiwei Long, Aibin Huang, Zewei Shao, Ning Li, Hongjie Luo, and Ping Jin. Optical design and stability study for ultrahigh-performance and long-lived vanadium dioxide-based thermochromic coatings. *Nano Energy*, 44:256 – 264, 2018. 2

[4] Cunjian Chen and Arun Ross. Exploring the use of IrisCodes for presentation attack detection. *Biometrics: Theory, Applications and Systems (BTAS)*, pages 1–9, 2018. 1, 4

[5] Cunjian Chen and Arun Ross. A multi-task convolutional neural network for joint iris detection and presentation attack detection. *IEEE Winter Applications of Computer Vision (WACV) Workshops*, 00:44–51, 2018. 4

[6] Cunjian Chen and Arun Ross. An explainable attention-guided iris presentation attack detector. *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV) Workshops*, pages 97–106, January 2021. 4

[7] Rui Chen, Xirong Lin, and Tianhuai Ding. Liveness detection for iris recognition using multispectral images. *Pattern Recognition Letters (PRL)*, 33(12):1513–1519, 2012. 2

[8] Horacio Coy, Rafmag Cabrera, Nelson Sepúlveda, and Félix E. Fernández. Optoelectronic and all-optical multiple memory states in vanadium dioxide. *Journal of Applied Physics (JAP)*, 108(11):113–115, 2010. 2

[9] A. Czajka. Database of iris printouts and its application: Development of liveness detection method for iris recognition. In *International Conference on Methods Models in Automation Robotics (MMAR)*, pages 28–33, 2013. 1

[10] Adam Czajka and Kevin W. Bowyer. Presentation attack detection for iris recognition: An assessment of the state-of-the-art. *ACM Computing Surveys (CSUR)*, 51(4):86:1–86:35, 2018. 1, 2

[11] P. Das, J. McGrath, Z. Fang, A. Boyd, G. Jang, A. Mohammadi, S. Purnapatra, D. Yambay, S. Marcel, M. Trokielwicz, P. Maciejewicz, K. Bowyer, A. Czajka, S. Schuckers, J. T. Farias, S. Gonzalez, M. Fang, N. Damer, F. Boutros, A. Kuijper, R. Sharma, C. Chen, and A. Ross. Iris liveness detection competition (livdet-iris) – the 2020 edition. *International Joint Conference on Biometrics (IJCB)*, 2020. 1, 2, 4

[12] John Daugman. Countermeasures against subterfuge. *Biometrics: Personal Identification in Networked Society*, pages 103–121, 1999. 1

[13] José Figueroa, Yunqi Cao, Henry Dsouza, Juan Pastrana, and Nelson Sepúlveda. A simplified approach for obtaining optical properties of VO_2 thin films, and demonstration of infrared shape-shifting devices. *Advanced Materials Technologies*, 4(4):1800599, 2019. 2

[14] Yanfeng Gao, Hongjie Luo, Zongtao Zhang, Litao Kang, Zhang Chen, Jing Du, Minoru Kanehira, and Chuanxiang Cao. Nanoceramic VO_2 thermochromic smart glass: A review on progress in solution processing. *Nano Energy*, 1:221–246, 2012. 2

[15] Yanfeng Gao, Shaobo Wang, Litao Kang, Zhang Chen, Jing Du, Xinling Liu, Hongjie Luo, and Minoru Kanehira. VO_2–Sb:SnO_2 composite thermochromic smart glass foil. *Energy Environmental Science*, 5:8234–8237, 2012. 2

[16] Yanfeng Gao, Shaobo Wang, Hongjie Luo, Lei Dai, Chuanxiang Cao, Yiliao Liu, Zhang Chen, and Minoru Kanehira. Enhanced chemical stability of VO_2 nanoparticles by the formation of SiO_2/VO_2 core/shell structures and the application to transparent and flexible VO_2-based composite foils with excellent thermochromic properties for solar heat control. *Energy Environmental Science*, 5:6104–6110, 2012. 2

[17] Lingxiao He, Haiqing Li, Fei Liu, Nianfeng Liu, Zhenan Sun, and Zhaofeng He. Multi-patch convolution neural network for iris liveness detection. In *IEEE International Conference on Biometrics: Theory Applications and Systems (BTAS)*, 2016. 6

[18] Steven Hoffman, Renu Sharma, and Arun Ross. Convolutional neural networks for iris presentation attack detection: Toward cross-dataset and cross-sensor generalization. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pages 1701–17018, 2018. 1, 2

[19] Steven Hoffman, Renu Sharma, and Arun Ross. Iris + ocular: Generalized iris presentation attack detection using multiple convolutional neural networks. *International Conference on Biometrics (ICB)*, 2019. 2, 8

[20] Hoover Vision Center. Halloween Hazard: The Dangers of Cosmetic Contact Lenses, http://hoovervisioncenter.com/2015/10/21/halloween-hazard-the-dangers-of-cosmetic-contact-lenses/, visited: 2018-01-03. 2

[21] G. Huang, Z. Liu, L. v. d. Maaten, and K. Q. Weinberger. Densely connected convolutional networks. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2261–2269, 2017. 3, 6

[22] K. Hughes and K. W. Bowyer. Detection of contact-lens-based iris biometric spoofs using stereo imaging. *Hawaii International Conference on System Sciences (HICSS)*, pages 1763–1772, 2013. 1

[23] Iris ID system Inc. iCAM7 series: User interface, https://www.irisid.com/productssolutions/hardwareproducts/icam7-series/. 4

[24] O. V. Komogortsev, A. Karpov, and C. D. Holland. Attack of mechanical replicas: Liveness detection with eye movements. *IEEE Transactions on Information Forensics and Security (TIFS)*, 10(4):716–725, 2015. 1

[25] Sung Lee, Kang Park, Youn Lee, Kwanghyuk Bae, and Jai Kim. Multifeature-based fake iris detection method. *Optical Engineering*, 46(12), 2007. 6

[26] S. J. Lee, K. R. Park, and J. Kim. Robust fake iris detection based on variation of the reflectance ratio between the iris and the sclera. *Biometrics Symposium: Special Session on Research at the Biometric Consortium Conference*, pages 1–6, 2006. 1, 2

[27] S. Lysenko, A. Rúa, V. Vikhnin, F. Fernández, and H. Liu. Insulator-to-metal phase transition and recovery processes in

VO_2 thin films after femtosecond laser excitation. *Physical Review B (PRB)*, 76:035104, 2007. 2

[28] S. Lysenko, V. Vikhnin, F. Fernandez, A. Rua, and H. Liu. Photoinduced insulator-to-metal phase transition in VO_2 crystalline films and model of dielectric susceptibility. *Physical Review. B (PRB)*, 75:075109, 2007. 2

[29] M. Dobeš and L. Machala. UPOL Iris Database. http://phoenix.inf.upol.cz/iris/. 2

[30] Sébastien Marcel, Mark S. Nixon, Julian Fiérrez, and Nicholas W. D. Evans, editors. *Handbook of Biometric Anti-Spoofing - Presentation Attack Detection, Second Edition*. Advances in Computer Vision and Pattern Recognition. Springer, 2019. 1

[31] F. J. Morin. Oxides which show a metal-to-insulator transition at the neel temperature. *Physical Review Letters (PRL)*, 3:34–36, 1959. 2

[32] Andrzej Pacut and Adam Czajka. Aliveness detection for iris biometrics. *IEEE International Carnahan Conferences Security Technology (ICCST)*, pages 122 – 129, 2006. 1

[33] R. Raghavendra and Christoph Busch. Robust Scheme for Iris Presentation Attack Detection using Multiscale Binarized Statistical Image Features. *IEEE Transactions on Information Forensics and Security (TIFS)*, 10(4):703–715, 2015. 1

[34] R. Raghavendra, K. B. Raja, and C. Busch. Contlensnet: Robust iris contact lens detection using deep convolutional neural networks. *IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 1160–1167, 2017. 1

[35] K. B. Raja, R. Raghavendra, and C. Busch. Video presentation attack detection in visible spectrum iris recognition using magnified phase information. *IEEE Transactions on Information Forensics and Security (TIFS)*, 10(10):2048–2056, 2015. 1, 2

[36] Nalini Ratha, Jonathan Connell, and Ruud Bolle. Enhancing security and privacy in biometrics-based authentication systems. *IBM Systems Journal*, 40:614–634, 01 2001. 1

[37] Ramprasaath R. Selvaraju, Michael Cogswell, Abhishek Das, Ramakrishna Vedantam, Devi Parikh, and Dhruv Batra. Grad-CAM: visual explanations from deep networks via gradient-based localization. *The IEEE International Conference on Computer Vision (ICCV)*, 2017. 6, 7

[38] Renu Sharma and Arun Ross. D-NetPAD: An explainable and interpretable iris presentation attack detector. *International Joint Conference on Biometrics (IJCB)*, 2020. 3

[39] Z. Sun, H. Zhang, T. Tan, and J. Wang. Iris image classification based on hierarchical visual codebook. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 36(6):1120–1133, 2014. 1, 2

[40] The Notre Dame Contact Lenses Dataset 2015. https://cvrl.nd.edu/projects/data/#the-notre-dame-contact-lense-dataset-2015ndcld15. 4

[41] Patrick Tinsley, Sandip Purnapatra, Mahsa Mitcheff, Aidan Boyd, Colton Crum, Kevin Bowyer, Patrick Flynn, Stephanie Schuckers, Adam Czajka, Meiling Fang, Naser Damer, Xingyu Liu, Caiyong Wang, Xianyun Sun, Zhaohua Chang, Xinyue Li, Guangzhe Zhao, Juan Tapia, Christoph Busch, Carlos Aravena, and Daniel Schulz. Iris liveness detection competition (livdet-iris) – the 2023 edition. *arXiv*, 2310.04541, 2023. 1

[42] L.J.P. van der Maaten and G.E. Hinton. Visualizing high-dimensional data using t-sne. *Journal of Machine Learning Research*, page 2579–2605, 2008. 7

[43] H. W. Verleur, A. S. Barker, and C. N. Berglund. Optical properties of VO_2 between 0.25 and 5 ev. *Physical Review*, 172:788–798, 1968. 2

[44] D. Yadav, N. Kohli, J. S. Doyle, R. Singh, M. Vatsa, and K. W. Bowyer. Unraveling the effect of textured contact lenses on iris recognition. *IEEE Transactions on Information Forensics and Security (TIFS)*, 9(5):851–862, 2014. 1

[45] Shivangi Yadav, Cunjian Chen, and Arun Ross. Relativistic discriminator: A one-class classifier for generalized iris presentation attack detection. *IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2020. 8

[46] D. Yambay, B. Becker, N. Kohli, D. Yadav, A. Czajka, K. W. Bowyer, S. Schuckers, R. Singh, M. Vatsa, A. Noore, D. Gragnaniello, C. Sansone, L. Verdoliva, L. He, Y. Ru, H. Li, N. Liu, Z. Sun, and T. Tan. LivDet iris 2017 — iris liveness detection competition 2017. *IEEE International Joint Conference on Biometrics (IJCB)*, pages 733–741, 2017. 1, 6

[47] D. Yambay, J. S. Doyle, K. W. Bowyer, A. Czajka, and S. Schuckers. LivDet-iris 2013– iris liveness detection competition 2013. *IEEE International Joint Conference on Biometrics (ICB)*, pages 1–8, 2014. 1

[48] David Yambay, Brian Walczak, Stephanie Schuckers, and Adam Czajka. LivDet-Iris 2015 — iris liveness detection competition 2015. In *IEEE International Conference on Identity, Security, and Behavior Analysis (ISBA)*, pages 1–6, 2017. 1

[49] Hui Bin Zhang, Zhenan Sun, Tieniu Tan, and Jianyu Wang. Learning hierarchical visual codebook for iris liveness detection. *International Joint Conference on Biometrics (IJCB)*, 2011. 2

[50] Zongtao Zhang, Yanfeng Gao, Hongjie Luo, Litao Kang, Zhang Chen, Jing Du, Minoru Kanehira, Yuzhi Zhang, and Zhong Wang. Solution-based fabrication of vanadium dioxide on F:SnO_2 substrates with largely enhanced thermochromism and low-emissivity for energy-saving applications. *Energy Environmental Science*, 4:4290–4297, 2011. 2