RESEARCH ARTICLE OPEN ACCESS

Single Input Multi Factor User Authentication Protocol for Smartphone

Prof. Prajakta R. Mali[1], Prof. Vaishali Londhe[2]
[1](Lecturer in Computer Engineering Department
Government Polytechnic, Thane,
[2] (HOD Computer Engineering Department
YadavraoTasgaonkar Institute of Engineering and Technology

Abstract:

Smart phones are increasingly becoming a tool for ubiquitous access to a number of services including but not limited to e-commerce and home banking, and are more and more used for sensitive data storage. If on the one hand this makes the smartphone a powerful tool in our private and professional life, on the other it has brought about a series of new challenging security and privacy threats and raised the need to protect users and their data through new secure authentication protocols. In this article, we illustrate how the security level of a human authentication system increases from traditional systems based on the use of passwords or badges to modern systems based on biometrics. We have moved a step forward by conceiving an authentication protocol based on the combined recognition of human face, voice and smartphone fingerprint.

Thanks to image processing techniques, both the distinctive characteristics of the face, voice and of the device that captured the face image can be extracted from a single video frame and used for a triple check of user identity. The fast-technological development of smartphones, allows performing sophisticated operations on the device itself. In the edge computing perspective, the burden of biometric recognition and source camera identification can be moved on the end user side.

Keywords **– Smartphone Authentication, Multi factor, Authentication protocol**

I. INTRODUCTION

Smartphones are by definition devices able to perform many of the functions of a computer. Their technology has a rapid development that is quickly overcoming the initial limits related to insufficient memory or low computational power. This has widened their use in daily life such as for email checking, messaging, and personal data storage (including private photos and passwords), but also in security-critical tasks, namely home banking operations, use of credit cards or other payment methods for online shopping, and remote access to workstations. The scenario described above has led to two consequences that are addressed in this article:

- The user and their smartphone are inseparable.
- Sensitive data and access to remote services must be protected.

The number of smartphone users worldwide is forecast to reach 2.1 billion in 2020 (from Statista - The portal for statistics, 2017). It is reported that in 2015 about eight-in-ten Americans used to shop online, 51% using a cell phone (source: Survey “Online Shopping and E-Commerce”, by Pew Research Centre). In 2016, Kaspersky Security Network (KSN), estimated mobile banking attacks increase of 1.6 times, compared to 2015. Pew Research Centre also reports (January 2016) that 28% of smartphone owners do not use a screen lock or other security features in order to access their phone or protect sensitive data stored on it. Finally, Acuity Market Intelligence has published its latest “Biometric Smartphone Update”, which reveals that the number of smartphones incorporating biometrics has doubled since January 2016. These data define the scenario that has given rise to the proposed authentication protocol. On one hand, there is the ever-increasing need of secure

authentication procedures and on the other, biometrics are spreading through smartphone applications. One important aspect addressed by the proposed protocol, is the ease of use. For convincing smartphone owners, including the ones that do not use any kind of security feature, it is important to design authentication protocols easy-to-use and as transparent as possible to the user. In the following sections, it is illustrated how the security and ease of use requirements are achieved by the proposed solution.

As mentioned before, the initial smartphone limits related to insufficient memory or low computational power are being overcome by fast-technological development. This allows performing sophisticated operations, including image and video processing on the device itself without requiring more demanding computation to be processed on the server side. In the edge computing perspective [1], the burden of biometric recognition and source camera identification can be moved on the end user side.

Kaspersky Lab Resource Center lists the "Top 7 Mobile Security Threats" on their website. Data leakage, unsecured Wi-Fi, and phishing attacks are part of them. These threats can be faced by a wise user behavior, such as avoid sharing personal information, always check the source, and use unique passwords. The threat addressed here, is the attempt to access the smartphone itself or a remote service by fooling the authentication system. Given the predominant role that the smartphone has assumed in our daily lives it is very unlikely to lend it to someone for a long period or do not immediately realize to have forgotten or lost it. Thus, it is a much more suitable object for user authentication compared to badges or keys, something that the user always brings with them. Existing techniques for authentication on smartphones includes personal identification number (PIN), numeric password, pattern, and biometrics. The latter have been increasingly adopted on smartphones in recent years, but also often been fooled. The famous hacking of the Apple Touch ID fingerprint scanner and then the Samsung Galaxy S8 iris scanner bypassed less than a month after it was shipped to public, have demonstrated the need of new and robust protocols for user authentication.

II. LITERATURE REVIEW

In this section, we review the recent literature emphasizing on the types of authentication mechanisms and the ways on which they are developed, and analyse them from security and usability point of view. More specifically, we present the assessment of commonly used user-authentication-mechanisms on smartphones, focusing on the security and usability issues.

Ways of Authentication

The usability of authentication mechanisms is one of the dominant attribute that influence users' acceptance of a particular authentication scheme. The ISO standard: 13407 define usability as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction, in a specified context of use". Further, the study, suggests that the usability can be done on the basis of three criteria: task performance, user satisfaction, and user cost.

Conventional authentication schemes, i.e., PIN, passwords, graphical patterns, are no more considered secure and convenient, because they are not able to distinguish between the users, rather they authorize everyone (regardless of whether that person is the legitimate owner of the device or not) who enter the correct credentials. Physiological biometric-based solutions are considered more secure because it is assumed human body traits cannot be shared, copied, lost or stolen. Moreover, they genuinely authenticate their users by forcing them to present themselves physically to the system. However, they are less preferable on smartphones due to their inherent usability issues. As such, security experts are focusing on developing the usable authentication systems because they believe that behavioural biometrics will restructure the authentication landscape in the next 5-8 years.

In each subsections, we have included tables presenting the synopsis of each authentication ways being used as different authentication types along with the references that either indicating usability pros and cons, or reporting security solutions and concerns.

1) Something You Know: As per the web-report, aver-age smartphone users get themselves engaged in 76 separate phone sessions, while heavy users (the top 10%) peaked to 132 sessions per day. PIN/passwords, and graphical patterns, require users to memorize their text, they had set earlier, to unlock their devices, every time they need to initiate the session (76 times a day). The capacity of the human brain to process the information varies from person to person. Zhang et al. found that users faced problems in remembering their passwords and more especially, to memorize and correctly recall numerous passwords. This encouraged users going for an easy or simple password which is quick to remember but this opens plenty of opportunities for attackers to guess or crack their passwords, easily. When the system enforces stringent password policies, users due to memorability issues, allow their browsers or password managers to save their username/password information to make future logins easier. However, users trusting their browsers or password managers are more likely to be a victim of a wide variety of attacks. Overall, 82% of end users are frustrated with managing passwords. Clearly, this indicates the lack of usability, and a result, nearly; 75 million smartphones users in the US do not use any of PIN, pattern, or passwords, because they consider them annoying and an obstacle in quick access to their smartphones

From security perspective, PINs, passwords, are vulnerable to various attacks, e.g., guessing, because users choose date of births, easier digits (1111, 2222, etc.), to set up their PIN. Alternatively, Android users (40% of them) pre-fer graphical patterns for device unlocking. But this approach too, requires users to remember them, hence users choose simple and less secure patterns, i.e., if a user connects at least four dots without repeating any of them in their patterns, the maximum number of combinations are 389,112 which could be easily cracked by brute-force. Ye at al. managed to crack 95% of 120 unique patterns collected from 215 independent users within just five attempts by recording their smartphone screen, remotely, while they were unlocking their devices. In addition, these schemes are more vulnerable to shoulder surfing than textual passwords.

2) Something You Have: Smartphones are being utilized for authentication purposes in several sensitive operations by the means of OTP via SMS, offline OTP using Apps, or pairing the wearable devices, e.g., smart-watches, smart-glasses, smartcards, etc. However, this idea of enhancing security with multi-factor authentication, i.e., topping knowledge based authentication with token based authentication (one-time-passcode), eventually perishes too due to side-channel attacks, e.g., MITM (Man-in-the-Middle), and MITPC/Phone (Man-in-the-PC/Phone). Software-based OTP solutions also do not guarantee the confidentiality of the generated passwords or the seeds as the mobile OS could be compromised, at the same time, could also suffer from denial-of-service attacks on the account of mobile OS crashes.

3) Something You Have -Insertable Biometrics: Insertable biometrics (see Table III) including implantable medical devices (IMDs) and emerging technologies such as Bespoke devices Neodymium Magnets NFC or RFID chips smart-piercings [smart-tattoos are the newer addition to biometrics that potentially can be used to provide increased usability over the existing solutions. Researches are exploring the further possibilities of insertable biometrics as go-to solution for improving digital security and usability in smartphones.

4) Something You Are - Physiological Biometrics: Mobile device manufacturers have started embedding biometric sensors in their flagship smartphones for reliable and convenient user authentication with the intuition that biometric approaches are better than their conventional authentication schemes. For example, Apple, Huawei, Lenovo (Motorola), Microsoft (Nokia), Samsung, and many other leading manufacturers have integrated fingerprint sensors, iris scanners, and face recognition algorithms, in some of their high-end devices. These advancements are akin to replacing a hay castle with a glass house to ward off attacks from sophisticated cyber-pirates.

Physiological biometrics e.g., face, fingerprint, iris, eyes etc., are commonly used as a one-shot, or multi-factor/multi-model (combining

with other modalities) authentication schemes for smartphones (see Table IV). Unexpectedly, biometric systems have shown to be exposed to different types of attacks, e.g., impersonation, replay, spoofing, hill climbing, etc., exposing their security loop-holes. These schemes suffer from their data leakage, i.e., a user's face can be easily found on social media websites, or her fingerprints can be extracted from the photos from their gestures, to mount a presentation attack against them. Additionally, these solutions also suffer from lack of secrecy and vulnerability to various spoofing attacks.

Recent research has shown that these schemes can be hacked very easily with almost negligible investment and efforts. For example, iPhone X face ID was hacked with 3D-printed mask costing just $150 approximately, while Samsung S8 facial recognition technology was simply fooled with a photo of the owner. Similarly, German Chaos Computer Club cracked the Samsung Galaxy S8 iris scanner with a dummy eye made from pictures of the iris, taken by a digital camera in a night mode, and covered it with a contact lens to match the curvature of the eye, within a month of S8 launch. The same club earlier cracked the iPhone 5S fingerprint sensor protection within two days after the device went on sale worldwide. Their hacking team photographed the glass surface containing the fingerprint of a user and created a "fake fingerprint" using a thin film to unlock the phone. Japan's National Institute of Informatics (NII) researcher Isao Echizen demonstrated that fingerprints can easily be recreated from photos, taken just from three meters distance, without the use of any sophisticated process and warned casually making a peace sign in front of a camera, which could lead to fingerprint theft.

III. PROPOSED SYSTEM

Existing system

Several surveys deal with the security needs of mobile phone users. Experimental results towards various authentication methods had been reviewed to illustrate the advantages and disadvantages of different approaches. The existing authentication methods can be divided into 3 groups as illustrated in Figure 1.

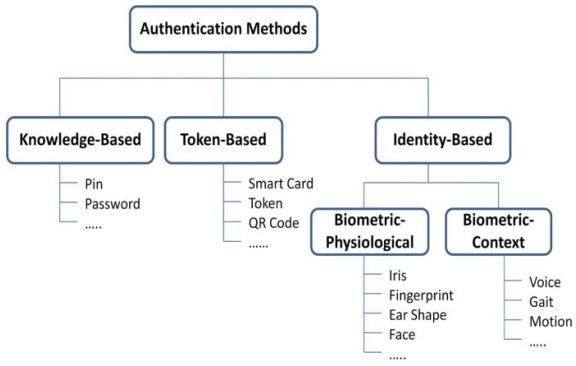

Figure 1 Classification of authentication methods.

Are passwords dead? Not entirely—but as the sole means to log in, protect sensitive information, and link important accounts, usernames and passwords alone are no longer enough. Brute force attacks, phishing scams, data breaches, and SQL injection attacks have become so common that usernames and passwords can be easily cracked, captured, and leaked. Pile on top of that the use of weak passwords, same passwords across multiple accounts, and the use of unsecure wifi networks, and many people are in jeopardy of getting hacked.

To overcome these dis-advantages of single verification, multi modal and TWO FACTOR (2FA) Authentication protocols are introduced. But still run risk of a man in the middle (MITM) or man in the browser (MITB) attack, as sessions in a browser that aren't closed can be compromised Phishing and social engineering are always factors. 2FA isn't foolproof; a hacker can call a phone company and impersonate you, activating a new SIM card and intercepting your SMS tokens.

Proposed system

The proposed authentication protocol combines the recognition of the user's smartphone with the recognition of the user based on their face and voice. The user is only required to record a short video clip of their face with oral password phrase for voice recognition. From that single clip, both faces, voice and device recognition is performed. The system work flow is illustrated in Fig. 2. Besides ensuring a higher level of security

than using biometric recognition only, as detailed in the following section, the proposed system has several advantages:

- Although the system consists of a triple recognition, the acquisition process is very easy and fast.
- The system is more robust to attacks since both the face, voice features and the smartphone signature must be replicated to fool the system.

Given the pervasiveness of technology in our lives, just think of the so-called Internet of things (IoT), this approach can be further applied for fast and secure authentication for any kind of smart object [2]. This article is an extended version of the paper "Secure User Authentication on Smartphones via Sensor and Face Recognition on Short Video Clips" [3], previously presented at the 12th International Conference on Green, Pervasive and Cloud Computing (GPC 2017).

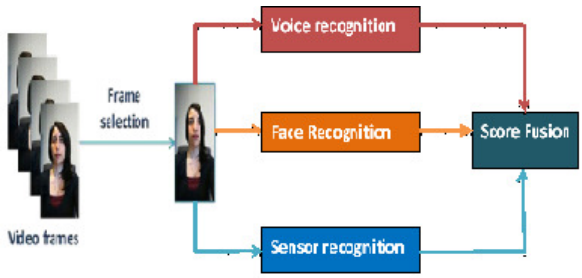

Figure 2 Proposed authentication system work flow.

User Authentication

Authentication can be performed based on one or a combination of the following items [4]:

- Something the user knows (e.g., password, personal identification number (PIN), secret answer, pattern).
- Something the user has (e.g., smart card, ID card, security token, software token, smartphone).
- Something the user is or does (e.g. fingerprint, face, iris, gait).

The last are known as biometrics. As a premise, it is worth considering that passwords can be forgotten or snatched by malicious people, physical objects such as badges and ID documents can be lost or stolen. Biometrics can hardly be stolen and the process of falsification is much more complicated (e.g. plastic surgery). The most recent biometric recognition systems also embed mechanisms to recognize live biometrics (liveness detection) and fakes (presentation attack detection). If we consider all possible combinations of the three factors of authentication, we obtain the ranking, from lowest to highest security, illustrated in Figure 3 [4].

- Something The User Knows;
- Something The User Has;
- Something The User Knows + Something The User Has;
- Something The User Is Or Does;
- Something The User Has + Something The User Is Or Does;
- Something The User Knows + Something The User Is Or Does;
- Something The User Knows + Something The User Has + Something The User Is Or Does

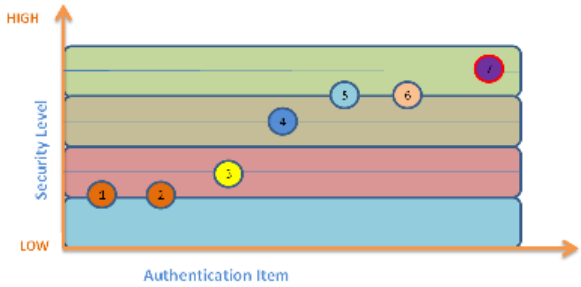

Figure 3 Authentication systems security levels:

Figure 3 plots the relative degrees of security. As mentioned before, the proposed system is of the type "Something the user knows (oral password phrase) + Something the user has (smartphone) + something the user is (face, voice)", and assures a higher security level compared to the use of biometrics only.

IV. METHODOLOGY

Smartphone Recognition

The smartphone is identified based on the distinctive pattern, also called camera fingerprint or camera signature, left by its digital camera on the

captured photos. That is why this technique is also referred as source digital camera identification. Each imaging sensor has a noise pattern originated from imperfections during the manufacturing process and different sensitivity of pixels to light due to the inhomogeneity of silicon wafers of which the sensor is composed [5]. Even sensors of the same model can be distinguished by analyzing the sensor pattern noise (hereinafter SPN). The technique to extract and compare the SPN from an image has been first presented by Lukas et al. in [6] and further improved by Li in [5].

An image can be represented by its frequencies in the so-called frequency domain. Low frequencies correspond to homogeneous image regions while high frequencies describe the image details including edges but also the sensor noise. The SPN of a sensor is obtained by applying a de-noising filter in the wavelet domain to isolate the frequencies associated with the sensor noise. However, since both noise and scene details are located in high frequencies, it is observed that the SPN can be affected by the image content [5]. Li's approach, namely the enhanced sensor pattern noise (ESPN), is based on the idea that strong SPN components are more likely to be originated from the scene details and thus must be suppressed, while weak components should be enhanced. Figure 4 illustrates how the SPN is still contaminated by picture details (i.e. edges) while the ESPN is less affected.

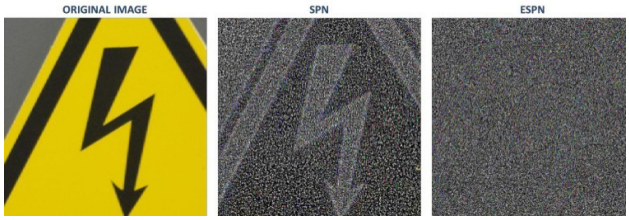

Figure 4 Comparisons between the SPN (middle) and the ESPN (right) extracted from a picture (left).

Sensor Pattern Noise from Videos

It is known that videos are strongly compressed. The SPN comparison achieves optimal performances on still images but source digital camera identification from videos is much more challenging. The sensor noise pattern is strongly impacted by video compression, and it is demonstrated that SPN performance drastically drops [7]. The identification rate can be improved by selecting from the video only the I-frames for SPN estimation. A compressed video is made up of three kinds of frames: the intra-coded picture (I-frame), the predicted picture (P-frame), and the bidirectional predictive picture (B-frame). I-frames are the least compressible. An I-frame is a complete image, like a JPG image file. P and B frames hold only part of the image information (the part that changes between frames, e.g. moving objects). Thus, part of the SPN is lost. P□frames hold only the changes in the image from the previous frame. For example, in a scene where a person moves across a stationary background, only the person's movements are encoded. B□ frames only store differences between the current frame and both the preceding and following frames. Therefore, the SPN is best preserved in I-frames. In the H.264/MPEG-4 AVC standard, the granularity is brought down to the "slice level." A slice is a spatially distinct region of a frame that is encoded separately from any other region in the same frame. I-slices, P-slices, and B-slices take the place of I, P, and B frames [8]. In [9] it is shown how performance improves by selecting only I-frames, or a weighted combination of I-frames and P-frames. In [10], Chen et al. propose a technique for determining whether two video clips came from the same camcorder by mean of the Maximum Likelihood Estimator for estimating the SPN, and of normalized cross-correlation for SPN comparison. Other factors can affect the SPN, for example video stabilization and the additional video compression operated by some website when uploading a video [11]. The latter is a major issue since videos with criminal content are often posted on line on social networks or web platforms for video sharing and the additional compression steps make more difficult to associate the video to the source digital camera.

Face Recognition

Face recognition, and biometric recognition in general, consists in compactly representing the features of the face. This representation is also known as biometric template. The method we adopted is based on the histogram of oriented

gradients (HOG) [12]. The idea behind this technique is that object appearance and shape can be represented by the distribution of local intensity gradients (i.e. a directional change in the intensity or colour) or edge directions, even without precise knowledge of the corresponding gradient or edge positions. The resulting HOG descriptors are then used as input of a conventional support vector machine (SVM) based classifier.

V. EXPECTED RESULTS

The main expectations of smartphone users are –

- The smartphone user needs highest security for his personal, private and financial data present in their smartphone.
- The smartphone user needs such authentication protocols which cannot be easily fooled by hackers.
- Also smartphone user doesn't want the burden of authentication process such as typing complicated password every time while unlocking.
- The smartphone user likes to use such security protocol which does not consume more memory of the device and more processing power of the device.
- Also user doesn't want the burden of remembering the password and patterns to unlock the device.

The proposed protocol, of which the workflow is illustrated in Fig. 1, requires in input a short video clip depicting the user face. A single I-frame is selected, it can be chosen according to many criteria such us image quality in terms of focusing. The frame is processed by three modules that can work independently, namely the voice recognition module, the face recognition module and the source digital camera identifier. Each module provides a score indicating how likely is that the input image comes from the authorized user.

The proposed authentication protocol satisfies all the user expectations because it provides highest security level i.e. Level 7 security as shown in figure 2. The proposed system is more robust to attacks since voice, face and the smartphone signature must be replicated to fool the system. Although the system consists of triple recognition, the acquisition process is very easy and fast. According to proposed system the smartphone user need not to remember or typing complicated password every time while unlocking.

On one hand, there is the ever-increasing need of secure authentication procedures and on the other, biometrics are spreading through smartphone applications. One important aspect addressed by the proposed protocol, is the ease of use. For convincing smartphone owners, including the ones that do not use any kind of security feature, it is important to design authentication protocols easy-to-use and as transparent as possible to the user. The user is only required to record a short video clip of their face with oral password phrase for voice recognition. From that single clip, both faces, voice and device recognition is performed. In this way the proposed system illustrates how the security and ease of use requirements are achieved.

As mentioned before, the initial smartphone limits related to insufficient memory or low computational power are being overcome by fast-technological development. This allows performing sophisticated operations, including image and video processing on the device itself without requiring more demanding computation to be processed on the server side. In the edge computing perspective [1], the burden of biometric recognition and source camera identification can be moved on the end user side.

In this way the proposed system will satisfy all the user expectations regarding smartphone authentication protocol.

VI. CONCLUSION

The need of protecting smartphone users' sensitive data and access to remote services led us to conceive an innovative authentication protocol. To the best of our knowledge, this is the first work proposing the combination of source camera identification, voice recognition and face recognition for real-time user authentication from videos. The authors have previously presented a system combining iris and sensor recognition from still images in [14]. Here, the use of videos as input data presents a considerable challenge since the

SPN is significantly affected by strong video compression. However, by simply selecting I-frames from a set of short video clips, a rate of correct classification equal to 77% is obtained by the source camera recognition module and a rate of 97% is achieved by the combination with the face module.

When dealing with biometric recognition, a question arises about privacy protection. How to protect sensitive data, such as the face picture, used for authentication? The solution mostly adopted is to never store the original picture/biometric sample, but only circulate its compact representation, namely the template. In addition, the template must never be externally visible decrypted.

The proposed protocol assures a more secure authentication by combining different authentication items, namely the user's face and their smartphone. Also, the acquisition process is simple and fast.

REFERENCES

1. *Chiara Galdi, Michele Nappi, Jean-Luc Dugelay, and Yong Yu, "Exploring New Authentication Protocols for Sensitive Data Protection on Smartphones" in IEEE Communications Magazine · January 2018, DOI: 10.1109/MCOM.2017.1700342*
2. *B. P. Rimal, D. P. Van, and M. Maier, "Mobile Edge Computing Empowered Fiber-Wireless Access Networks in the 5G Era," in IEEE Communications Magazine, vol. 55, no. 2, February 2017, pp. 192-200. doi: 10.1109/MCOM.2017.1600156CM*
3. *M. Shahzad and M. P. Singh, "Continuous Authentication and Authorization for the Internet of Things," in IEEE Internet Computing, vol. 21, no. 2, Mar.-Apr. 2017, pp. 86-90. doi: 10.1109/MIC.2017.33*
4. *C. Galdi, M. Nappi, and J.-L. Dugelay, "Secure User Authentication on Smartphones via Sensor and Face Recognition on Short Video Clips," in International Conference on Green, Pervasive, and Cloud Computing, May 2017, pp. 15-22.*
5. *L. O'Gorman, "Comparing passwords, tokens, and biometrics for user authentication," in Proceedings of the IEEE, vol. 91, no. 12, Dec 2003, pp. 2021-2040. doi: 10.1109/JPROC.2003.819611*
6. *C. T. Li, "Source camera identification using enhanced sensor pattern noise," in IEEE Trans-actions on Information Forensics and Security 5(2), 2010, pp. 280-287.*
7. *J. Lukas, J. Fridrich, and M. Goljan, "Digital camera identification from sensor pattern noise," in IEEE Trans. Inf. Forensics Security, vol. 1, no. 2, June 2006, pp. 205-214.*
8. *C. Galdi, F. Hartung, and J.-L. Dugelay, "Videos versus still images: Asymmetric sensor pattern noise comparison on mobile phones," in Electronic Imaging, Media Watermarking, Security, and Forensics 2017, 2017, pp. 100-103(4).*
9. *G. J. Sullivan and T. Wiegand, "Video Compression - From Concepts to the H.264/AVC Standard," in Proceedings of the IEEE, vol. 93, no. 1, Jan. 2005, pp. 18-31.*
10. *W. H. Chuang, H. Su, and M. Wu, "Exploring compression effects for improved source camera identification using strongly compressed video," in 18th IEEE International Conference on Image Processing, Brussels, 2011, pp. 1953-1956. doi: 10.1109/ICIP.2011.6115855*
11. *M. Chen, J. J. Fridrich, M. Goljan, and J. Lukas, "Source digital camcorder identification using sensor photo response non-uniformity," in Electronic Imaging 2007, SPIE Proceedings Vol. 6505, Security, Steganography, and Watermarking of Multimedia Contents IX, 2007, pp. 65051G-65051G.*
12. *W. Van Houten and Z. Geradts, "Using sensor noise to identify low resolution compressed videos from YouTube," in International Workshop on Computational Forensics, 2009, pp. 104-115.*
13. *N. Dala and B. Triggs, "Histograms of Oriented Gradients for Human Detection," in IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 1, June 2005, pp. 886-893.*
14. *A. Jain, K. Nandakumar, and A. Ross, "Score normalization in multimodal biometric systems," in Pattern Recognition, Volume 38, Issue 12, December 2005, pp. 2270-2285. doi: 10.1016/j.patcog.2005.01.012*
15. *C. Galdi, M. Nappi, and J.-L. Dugelay, "Multimodal authentication on smartphones: Combining iris and sensor recognition for a double check of user identity," in Pattern Recognition Letters, Volume 82, Part 2, 15 October 2016, pp. 144-153. doi: 10.1016/j.patrec.2015.09.009*
16. *M. De Marsico, C. Galdi, M. Nappi, and D. Riccio, "Firme: Face and iris recognition for mobile engagement," in Image and Vision Computing, 32(12), 2014, pp. 1161-1172.*