



The Interplay of AI and Biometrics: Challenges and Opportunities

Christian Berghoff, Matthias Neu, and Arndt von Twickel, Federal Office for Information Security

Current biometric applications exhibit numerous connections to artificial intelligence (AI). They use AI to boost their performance, can fall prey to attacks targeting AI components or based on AI, and can use AI to ward off attacks.

Biometric applications serve to identify individuals based on their biological characteristics or behaviors. Currently, fingerprints and faces are the predominantly used biometric modalities,

but systems based on hand veins, irises, or voices are also available. Biometric applications can use single modalities or combine multiple ones, and they can process static data (for example, for facial recognition) as well as data sequences (for example, for video identification or speaker verification). Data can be acquired using single or multiple sensors of different types (for example, optical or acoustic).

Nowadays, biometrics are being increasingly used in many applications in different sectors. Such applications range from automatic border control and physical access control in some contexts to a plethora of use cases in which biometrics are used for authenticating individuals (for example, for authorizing user actions on mobile

devices in the consumer sector). This widespread use is mostly based on substantial performance improvements due to the use of connectionist artificial intelligence (AI) methods, in particular, deep neural networks (DNNs). One striking example of their superiority over traditional systems is their ability to match facial images taken from different angles with high probability.¹



Since biometric applications implement a security functionality, they constitute valuable targets for attackers. Despite the impressive results shown by current biometric AI systems, these systems exhibit a range of vulnerabilities—many of which can themselves be exploited with the use of AI.² Some of these vulnerabilities are specific to the AI systems whereas others also apply to traditional biometric systems. Depending on the application-specific ambient conditions and the security level required, these vulnerabilities must be mitigated, and, again, AI methods play an important role in this.² Figure 1

provides an overview of the interplay of biometrics and AI based on the example of authentication using facial biometrics.

VULNERABILITIES OF BIOMETRIC SYSTEMS

Attacks on AI

It is well known that applications that rely on connectionist AI methods (for example, DNNs) face novel attacks.³ Also, the prevalent ambient conditions of biometric applications do not usually detract from these new threats.

Connectionist AI methods are trained using a large amount of data.

A robust performance of an AI system depends on the availability of training and of test data of a sufficient quantity and quality, that is, the data should be representative, unbiased, and correctly labeled. Attackers may target this training process by inserting maliciously crafted samples into the training data set. Such poisoning attacks can aim to degrade a system's overall performance or, more concerningly, they can aim to insert backdoors that an attacker can use when interacting with the deployed system later on. In a backdoor attack on a biometric system, malicious samples include a

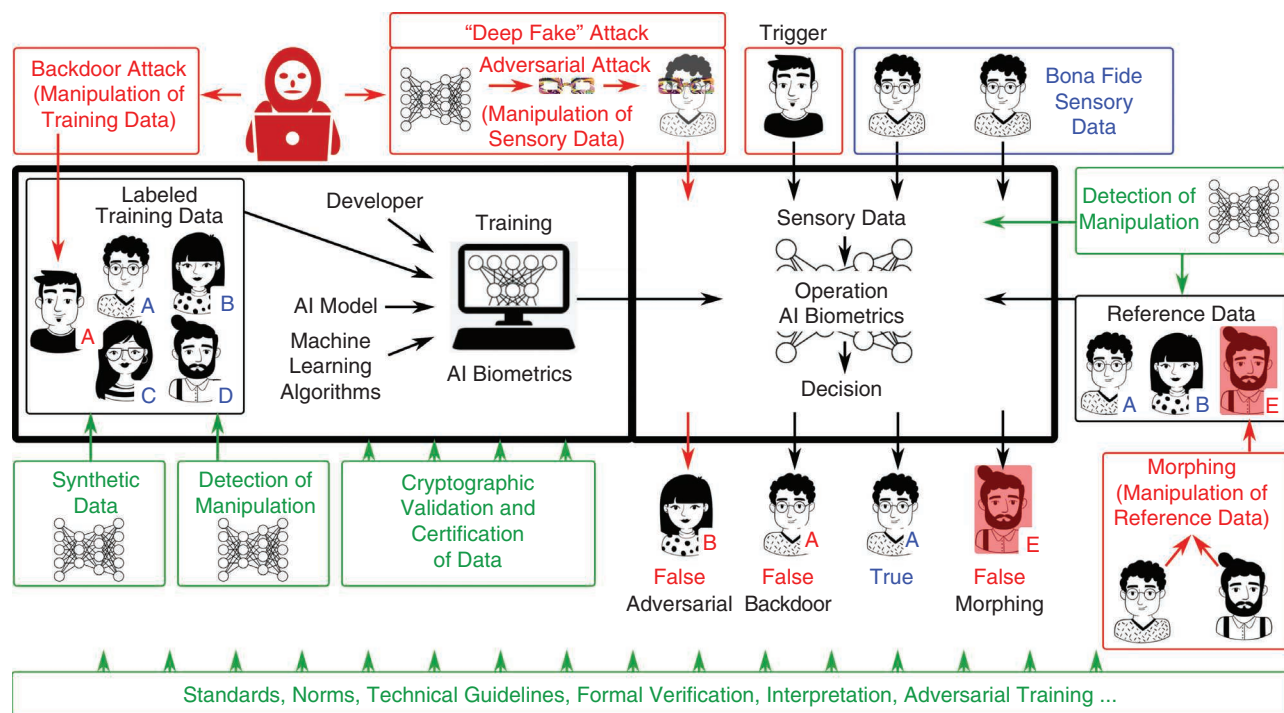


FIGURE 1. A facial-biometrics system based on AI is trained using labeled facial images for the purpose of identifying individuals (left part). During operation, individual users can be enrolled to the system, which stores reference images (right). Later on, new images of enrolled users, which are acquired through sensors, can be matched with the stored references to identify the users (right part). The system can fall prey to backdoor and adversarial attacks directly targeting the AI components (upper part) and to morphing attacks manipulating the reference data (lower right). Mitigations to these attacks (green parts) can use AI methods themselves but can also rely on cryptographic techniques to prevent manipulations by protecting the integrity of the data, which allows for certifying their quality. (© BSI; used with permission.)

trigger pattern strongly associated with a particular identity chosen by the attacker. If an AI model is trained using these samples, presenting the trigger pattern to it will cause targeted misclassifications, allowing an attacker to impersonate another individual (see Figure 1).⁴

Another AI-specific vulnerability is exploited by evasion attacks (also known as *adversarial attacks*), which use specially crafted inputs (called *adversarial examples*) to cause a deployed AI model to make unintended decisions. In facial recognition, such inputs can, for instance, be special patterns attached to a glasses frame

Attacks Using AI

Biometric systems may be fooled by fake inputs, which are created by manipulating original biometric data using AI methods. Currently, a particular target of these attacks is biometric systems processing data sequences. For instance, speaker-verification systems can be deceived by means of AI-based voice-generation and modification methods, such as voice conversion or text-to-speech methods. Using voice-conversion methods and given arbitrary words and sentences spoken by an individual, it is possible to convert audio data in such a way that the speaker characteristics are changed to those of a selected speaker

Due to the large space for possible inputs and the lack of interpretability of current connectionist AI models, it is very hard to reliably prevent and detect such attacks.

or a cap, which allow an attacker to impersonate some other individual (see Figure 1).^{5,6} When considering speaker verification, adversarial examples are special noise patterns, which are added to an audio signal to change the output of the system under attack while being inaudible to the human ear.^{7,8} It is highly unlikely for adversarial examples to occur naturally. However, they can be efficiently precomputed if the attacker has access to the model internals. The attacker can sidestep this requirement by training a surrogate model.

Due to the large space for possible inputs and the lack of interpretability of current connectionist AI models, it is very hard to reliably prevent and detect such attacks. In some major jurisdictions, where biometric data are specially protected (for example, by the European Union General Data Protection Regulation) and, thus, not easily available in large quantities, the threat of poisoning attacks is aggravated since AI developers may be tempted to use whatever data they can legally attain.

while the linguistic content of the audio remains unaltered.⁹ Video-based biometrics can fall prey to AI-based manipulation techniques.^{10,11} Depending on the use case, attackers may employ, for example, face-swapping techniques, which use videos of an attacker and a victim to create fake videos by seamlessly replacing the attacker's face with the victim's.

Creating high-quality audio and video fakes requires training an AI model with a sufficient amount of data of the victim of an attack. Hence, targeted attacks may be challenging unless such data are publicly available, as is the case for persons of public interest. However, it should be noted that recent research has made significant progress in reducing the amount of training data required of a victim. In general, due to the increasing availability of free tools, public data sets, and computing power, creating realistic fakes has become much easier, and even real-time manipulations seem within reach. These techniques can be used not only to attack biometric

authentication systems but also, for instance, to influence public opinion via sophisticated fake news.

Another type of attack in which AI techniques are increasingly being used is called the *morphing attack*.¹² It is well-documented in facial recognition, where the facial images of multiple individuals can be fused together to create a new facial image that has a great similarity to all of the original faces. Since facial-recognition systems are designed to be robust against natural variances in human faces, this similarity is usually sufficient to achieve high verification results. In this setting, the attack consists of replacing the reference images enrolled to the system that are used for identifying an individual (see Figure 1). An attacker may, for instance, aim to insert the morphed image into a passport, allowing more than one individual to use it for passing border control.

Presentation Attacks

Yet another, non AI-specific threat is posed by presentation attacks (PAs), which use artifacts that counterfeit human characteristics to fool a biometric system. While some PAs require bespoke equipment, such as authentic facial masks, rather basic PAs, for example, those using simple 2D facial images, are remarkably successful in many cases. PAs exploit shortcomings in the sensory information available to and its processing by the biometric system. For example, facial-recognition systems using only 2D optical information can often be easily fooled by printed images or images shown on screens, and fingerprint sensors often fail to distinguish real fingers made of flesh and blood from fake ones made of materials like silicone, wood glue, or latex.^{13,14}

MITIGATIONS

Mitigating Attacks on AI

Several approaches exist for mitigating AI-specific attacks.³ Most of these approaches are not application-specific and, hence, can directly be applied

to AI systems used in biometrics (see the green parts of Figure 1).

Although many mitigations to evasion and poisoning attacks have been proposed and explored, an arms race is taking place in this area, and, so far, there is no knowledge of a mitigation that can reliably thwart an adaptive attacker adjusting his or her attack to take the defenses in place into account. Nevertheless, deploying defenses can help raise the bar for attackers, thus decreasing the likelihood of successful security breaks.

Common countermeasures to evasion attacks include robustifying the training procedure by using regularization methods or adversarial training, which integrates relevant adversarial examples into the training procedure, or preprocessing inputs to hamper the effectiveness of adversarial examples or to detect them.¹⁵ A general approach consists of restricting access to the model and the information on it. This can hinder direct query attacks as well as model-extraction attacks used to create substitute models for crafting attacks. Whereas this general approach may not prevent black-box attacks altogether, it raises the bar for attackers.

Similarly, several generic mitigations to poisoning attacks have been proposed, in particular, methods for detecting poisoned samples in the training data. Another mitigation transcending the level of the AI system is to guarantee the quality of the training data by carefully selecting them from reliable sources and protecting their integrity along the whole data supply chain, thus preventing an attacker from tampering with them. Research on creating synthetic biometric data for training AI models may also alleviate the problem of the limited availability of biometric data in some jurisdictions and the increased susceptibility to poisoning attacks resulting from it.¹⁶

Other fields of research in AI can likewise be profitable for dealing with attacks. On the one hand, explainable AI (XAI) methods can improve the

interpretability of AI models, which may help to identify vulnerabilities in them and to more easily detect attacks. On the other hand, the formal verification of AI models can be used to guarantee their resistance against certain types of attacks. So far, the verification of AI is restricted to very limited boundary conditions.

Mitigations Using AI

AI doesn't just introduce new vulnerabilities: It is a double-edged sword in IT security that can also be wielded for defense. Attackers may be able to use AI models for creating fake inputs using voice conversion or face swapping, but defenders can train AI models to

can also be used for PAD, targeting the root cause of PAs and, thus, considerably raising the bar for attackers.^{13,14}

A general organizational measure consists of changing the ambient conditions when biometrics are used. Partial human supervision may be used to deter attackers and expose them more easily, for example, in border control.

STANDARDS AND REGULATION

Although biometrics are widely used in many applications in the public, industrial, and consumer sectors, the number of security standards is quite limited so far, and AI-related attacks and defenses are not their focus. Generally speaking,

These techniques can be used not only to attack biometric authentication systems but also, for instance, to influence public opinion via sophisticated fake news.


spot these attacks.^{17,18} In addition, AI models can help to tackle the notoriously difficult problem of PA detection (PAD) by efficiently extracting and assessing all of the available and possibly multimodal sensory information.¹⁹ However, it should be noted that such AI-based defense methods can themselves be the targets of AI-specific attacks, like adversarial attacks.

FURTHER MITIGATIONS

Other technical and organizational measures can also mitigate attacks. When dealing with fake inputs, challenge-response methods can make attacks more difficult, preventing attackers from precomputing the inputs and, instead, forcing them to compute them online. For instance, in video identification, having a user move his or her finger in front of his or her face or having the user use reflective objects can help to bring forth inconsistencies in fake inputs. Additional sensors of different types (for example, for detecting the material or 3D structure of an object)

the security levels that different standards enforce vary considerably.

In the consumer sector, the Fast Identity Online standards play an important role for facial- and fingerprint-recognition systems on mobile devices.²⁰ They serve as a first step by creating a baseline security level, but they do not consider even slightly more advanced attacks. Several standards address mitigations to specific attacks, like morphing²¹ and PAs.²² In the public sector, more stringent regulation is generally in place, particularly for high-security applications, such as automatic border control. New regulation also targets sensitive applications in the consumer sector,²³ where biometrics-based authentication has become immensely popular on mobile devices for unlocking them, accessing their functionalities, and authorizing money transfers via fingerprint or facial recognition,²⁴ for example, in widespread payment solutions provided by companies like Apple, Google, Samsung, and Alibaba.

The relations between AI and biometrics are numerous. AI has been of paramount importance for increasing the performance of biometric systems to levels unseen with previous technology. Yet, AI can also bring forth new attacks, both as a new target and as a new tool for attackers. There are mitigations of varying effectiveness to these attacks, and many of these are based on AI themselves. However, in the arms race between attackers and defenders that is so typical of security-related areas, it is important that defenders keep pace with their counterparts. As much as ever, the concrete defensive measures to be applied depend strongly on the applications and their specific ambient conditions. The consumer sector needs special attention in this respect. Taking into account that biometrics, even when using AI methods, are not the silver bullet for secure authentication, existing standards should be updated, tightened, and enforced with a focus on sensitive consumer applications. Germany's Federal Office for Information Security (BSI) is working toward this direction by developing new mitigation techniques and test criteria for certifying biometric systems on higher security levels.²³ 

ACKNOWLEDGMENTS

We would like to thank our colleague Ralph Breithaupt for carefully proofreading the manuscript and providing valuable suggestions for improvement. We would also like to thank the reviewers for their helpful suggestions.

REFERENCES

1. P. J. Grother, M. L. Ngan, and K. K. Hanaoka, "Ongoing Face Recognition Vendor Test (FRVT) part 2: Identification," National Institute of Standards and Technology, Gaithersburg, MD, Tech. Rep. 8238, 2019. [Online]. Available: <https://www.nist.gov/publications/ongoing-face-recognition-vendor-test-frvt-part-2-identification>
2. B. Biggio, G. Fumera, P. Russu, L. Didaci, and F. Roli, "Adversarial biometric recognition: A review on biometric system security from the adversarial machine-learning perspective," *IEEE Signal Process. Mag.*, vol. 32, no. 5, pp. 31–41, 2015. doi: 10.1109/MSP.2015.2426728.
3. C. Berghoff, M. Neu, and A. von Twickel, "Vulnerabilities of connectionist AI applications: Evaluation and defense," *Front. Big Data*, vol. 3, p. 23, July 2020. doi: 10.3389/fdata.2020.00023.
4. B. Biggio, G. Fumera, F. Roli, and L. Didaci, "Poisoning adaptive biometric systems," in *Proc. Joint IAPR Int. Workshops on Statist. Techn. Pattern Recognit. (SPR) and Structural Syntactic Pattern Recognit. (SSPR)*, 2012, pp. 417–425. doi: 10.1007/978-3-642-34166-3_46.
5. M. Sharif, S. Bhagavatula, L. Bauer, and M. K. Reiter, "Accessorize to a crime: Real and stealthy attacks on state-of-the-art face recognition," in *Proc. 2016 ACM SIGSAC Conf. Comput. Commun. Security*, 2016, pp. 1528–1540. doi: 10.1145/2976749.2978392.
6. F. Vakhshiteh, R. Ramachandra, and A. Nickabadi, "Threat of adversarial attacks on face recognition: A comprehensive survey," 2020, arXiv:2007.11709.
7. G. Chen, S. Chen, L. Fan, X. Du, Z. Zhao, F. Song, and Y. Liu, "Who is real bob? Adversarial attacks on speaker recognition systems," 2020, arXiv:1911.01840.
8. Z Li, C Shi, Y Xie, J Liu, B Yuan and Y Chen, "Practical adversarial attacks against speaker recognition systems," in *Proc. 21st Int. Workshop on Mobile Comput. Syst. Appl.*, 2020, pp. 9–14.
9. B. Sisman, J. Yamagishi, S. King, and H. Li, "An overview of voice conversion and its challenges: From statistical modeling to deep learning," *IEEE Trans. Audio Speech Language Process.*, vol. 29, pp. 132–157, 2021. doi: 10.1109/TASLP.2020.3038524.
10. A. Rössler, D. Cozzolino, L. Verdoliva, C. Riess, J. Thies, and M. Nießner, "FaceForensics++: Learning to detect manipulated facial images," in *Proc. 2019 Int. Conf. Comput. Vision*, pp. 1–11.
11. A. Tewari et al., "State of the art on neural rendering," *Comput. Graph. Forum*, vol. 39, no. 2, pp. 701–727, 2020. doi: 10.1111/cgf.14022.
12. U. Scherhag, C. Rathgeb, J. Merkle, R. Breithaupt, and C. Busch, "Face recognition systems under morphing attacks: A survey," *IEEE Access*, vol. 7, pp. 23,012–23,026, Feb. 2019. doi: 10.1109/ACCESS.2019.2899367.
13. S. Marcel, M. S. Nixon, J. Fierrez, and N. W. D. Evans, *Handbook of Biometric Anti-Spoofing: Trusted Biometrics under Spoofing Attacks*. Cham: Springer-Verlag, 2019.
14. D. Yambay, L. Ghiani, G. L. Marcialis, F. Roli, and S. Schuckers, "Review of fingerprint presentation attack detection competitions," in *Handbook of Biometric Anti-Spoofing: Trusted Biometrics under Spoofing Attacks*, S. Marcel, M. S. Nixon, J. Fierrez, and N. W. D. Evans, Eds. Cham: Springer-Verlag, 2019, pp. 109–131.
15. R. Theagarajan and B. Bhanu, "Defending black box facial recognition classifiers against adversarial attacks," in *Proc. IEEE/CVF Conf. Comput. Vision Pattern Recognit. Workshops*, 2020, pp. 812–813.
16. A. Kortylewski, B. Egger, A. Schneider, T. Gerig, A. Morel-Forster, and T. Vetter, "Analyzing and reducing the damage of dataset bias to face recognition with synthetic data," in *Proc. IEEE Conf. Comput. Vision Pattern Recognit. Workshops*, 2019, pp. 2261–2268.
17. M. Todisco et al., "ASVspoof 2019: Future horizons in spoofed and fake audio detection," in *Proc. Interspeech*, 2019, pp. 1008–1012. doi: 10.21437/Interspeech.2019-2249.
18. R. Tolosana, R. Vera-Rodríguez, J. Fierrez, A. Morales, and J. Ortega-Garcia, "DeepFakes and beyond: A survey of face manipulation and fake detection," *Inf. Fusion*, vol. 64, pp. 131–148, Dec. 2020. doi: 10.1016/j.inffus.2020.06.014.
19. L. Spinoulas et al., "Multispectral biometrics system framework:

- Application to presentation attack detection," *IEEE Sensors J.*, vol. 21, no. 13, pp. 15,022–15,041, July 1, 2021. doi: 10.1109/JSEN.2021.3074406.
20. "Biometric component certification," FIDO Alliance, 2020. <https://fidoalliance.org/certification/biometric-component-certification/>
 21. "Face Recognition Vendor Test (FRVT) Part 4: MORPH- Performance of automated face morph detection," National Institute of Standards and Technology, Gaithersburg, MD, Tech. Rep. 8292, 2020. [Online]. Available: <https://www.nist.gov/publications/face-recognition-vendor-test-frvt-part-4-morph-performance-automated-face-morph>
 22. *Information Technology — Biometric Presentation Attack Detection — Part 2: Data Formats*, ISO/IEC 30107, 2016.
 23. "Technical guideline for biometric authentication components in devices for authentication," Bundesamt für Sicherheit in der Informationstechnik (BSI), Bonn, Germany, Tech. Rep. TR-03166, 2021. [Online]. Available: <https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/Publications/TechGuidelines/TR03166/BSI-TR-03166.pdf>
 24. W. R. Malatji, R. van Eck, and T. Zuva, "Acceptance of biometric authentication security technology on mobile devices," in *Proc. 2nd Int. Multidisciplinary Inform. Technol. Eng. Conf. (IM-ITEC)*, 2020, pp. 1–5. doi: 10.1109/IMITEC50163.2020.9334082.

CHRISTIAN BERGHOFF is a technical officer at the Federal Office for Information Security, Bonn, 53175, Germany. Contact him at christian.berghoff@bsi.bund.de.

MATTHIAS NEU is a technical officer at the Federal Office for Information Security, Bonn, 53175, Germany. Contact him at matthias.neu@bsi.bund.de.

ARNDT VON TWICKEL is a technical officer at the Federal Office for Information Security, Bonn, 53175, Germany. Contact him at arndt.twickel@bsi.bund.de.



IEEE Computer Graphics and Applications bridges the theory and practice of computer graphics. Subscribe to *CG&A* and

- stay current on the latest tools and applications and gain invaluable practical and research knowledge,
- discover cutting-edge applications and learn more about the latest techniques, and
- benefit from *CG&A*'s active and connected editorial board.



Digital Object Identifier 10.1109/MC.2021.3102347