

Toward a Privacy-Preserving Face Recognition System: A Survey of Leakages and Solutions

LAMYANBA LAISHRAM, Kyungpook National University, Daegu, South Korea
MUHAMMAD SHAHERYAR, Kyungpook National University, Daegu, South Korea
JONG TAEK LEE, Kyungpook National University, Daegu, South Korea
SOON KI JUNG, Kyungpook National University, Daegu, South Korea

Abstract Recent advancements in face recognition (FR) technology in surveillance systems make it possible to monitor a person as they move around. FR gathers a lot of information depending on the quantity and data sources. The most severe privacy concern with FR technology is its use to identify people in real-time public monitoring applications or via an aggregation of datasets without their consent. Due to the importance of private data leakage in the FR environment, academia and business have given it a lot of attention, leading to the creation of several research initiatives meant to solve the corresponding challenges. As a result, this study aims to look at privacy-preserving face recognition (PPFR) methods. We propose a detailed and systematic study of the PPFR based on our suggested six-level framework. Along with all the levels, more emphasis is given to the processing of face images as it is more crucial for FR technology. We explore the privacy leakage issues and offer an up-to-date and thorough summary of current research trends in the FR system from six perspectives. We also encourage additional research initiatives in this promising area for further investigation.

CCS Concepts: • **Computing methodologies** → **Computer vision**; **Biometrics**; • **Security and privacy** → **Privacy-preserving protocols;**

Additional Key Words and Phrases: Privacy preservation, face recognition, surveillance system, face obfuscation, de-identification, face encryption, face privacy, deep learning

ACM Reference Format:
Lamyanba Laishram, Muhammad Shaheryar, Jong Taek Lee, and Soon Ki Jung. 2025. Toward a Privacy-Preserving Face Recognition System: A Survey of Leakages and Solutions. *ACM Comput. Surv.* 57, 6, Article 147 (February 2025), 38 pages. https://doi.org/10.1145/3673224

L. Laishram and M. Shaheryar contributed equally to this research.

This work was supported by the Innovative Human Resource Development for Local Intellectualization program through the Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korean government (MSIT) (IITP-2024-RS-2022-00156389) and also by the Digital Innovation Hub project supervised by the Daegu Digital Innovation Promotion Agency(DIP) grant funded by the Korea government (MSIT and Daegu Metropolitan City) in 2024 (No. DBSD1-04, Smart management system for preventing to lonely deaths of elderly people living alone based on automatic meter reading information and CCTV access information). ※ MSIT: Ministry of Science and ICT.

Authors' Contact Information: Lamyanba Laishram, School of Computer Science and Engineering, Kyungpook National University, Daegu, South Korea; e-mail: yanbalaishram@knu.ac.kr; Muhammad Shaheryar, School of Computer Science and Engineering, Kyungpook National University, Daegu, South Korea; e-mail: shaheryar@knu.ac.kr; Jong Taek Lee, School of Computer Science and Engineering, Kyungpook National University, Daegu, South Korea; e-mail: jongtaeklee@knu.ac.kr; Soon Ki Jung, School of Computer Science and Engineering, Kyungpook National University, Daegu, South Korea; e-mail: skjung@knu.ac.kr.



This work is licensed under a Creative Commons Attribution International 4.0 License.

© 2025 Copyright held by the owner/author(s).
ACM 0360-0300/2025/02-ART147
https://doi.org/10.1145/3673224

1 Introduction

1.1 Brief Introduction to Face Recognition (FR)

FR is a biometric authentication method that identifies people by analyzing and matching facial features that are captured from images or video frames with the stored template. The FR system processes face images or videos to obtain higher-level information such as human identification, emotions, demographic data, and so on. The input data for general FR can vary and may comprise images and video acquired from numerous sources, including portable devices or live camera feeds. In addition, FR can be divided into two categories: face identification and face verification [60]. Face verification uses 1:1 matching to determine whether two images are of the same person by comparing them, while face identification uses 1 matching to select the best match of a person in an image from among all potential outputs. Moreover, FR is a prevalent issue in **machine learning** (**ML**) research, especially in biometrics and computer vision, with safety and security being key benefits of FR systems.

Furthermore, FR technology is extensively used in various organizations, such as law enforcement and airports, and in applications like security, access control, and entertainment. One of the most significant examples of its use is in surveillance systems, which have grown in popularity with technological advancements. These systems are used in residential societies, businesses, and public spaces to monitor activities for safety and security [184]. CCTV technology is employed to identify people or monitor activities continuously in many areas, enabling authorities to employ recordings to identify suspects after a crime [65].

In China, FR technology for surveillance has rapidly expanded [44], prioritized for public security, AI development support, and COVID-19 prevention research. The CCTV market is expected to grow by approximately 46.52 billion USD by 2030 [61]. China leads in public surveillance with 200 million CCTV cameras, averaging 14.36 cameras per 100 people [200]. The United States ranks second with 50 million cameras. For instance, Atlanta has 15.56 public CCTV cameras per 1,000 people, and Chicago has 13.06 cameras per 1,000 people with 35,000 cameras in total [174]. Moreover, CCTV-based surveillance systems reduce security costs by preventing crime and discerning suspicious behavior. However, increased FR technology use raises additional privacy concerns.

1.2 Motivation for Privacy Protection

Surveillance ideally should not allow personal information to be derived from footage, but constant observation makes this impossible. The growing use of FR technology with CCTV systems has heightened privacy concerns, such as the ability to identify individuals from an image or video stream discreetly. The advancements in FR technology pose privacy risks for people unintentionally recorded in public places. Serious concerns exist over recording faces for recognition tasks due to potential security flaws or leaks in FR systems, potentially exposing biometric information without consent.

1.2.1 CCTV Incidents. In a CCTV incident in August 2022, a vast Chinese database containing millions of faces was left open on the internet for months before disappearing [203]. Xinai Electronics collected millions of license-plate numbers and facial images through a network of cameras, claiming the data were "secretly kept." However, the data were not securely stored, allowing anyone with the correct information to access it via a web browser, raising legitimate concerns about privacy breaches.

In daycares or nursery schools, CCTV allows parents to access live feeds, giving them confidence in the staff and understanding of the preschool system [128]. However, constant monitoring invades the privacy of children, caretakers, and institutions [51], and raises the risk of adversaries monitoring victims' lives.

The ability of FR technology to monetize and exploit personal biometric data without consent has increased public awareness of privacy invasion issues. Clearview AI collected facial images from social media to develop an FR system that could match images from databases [90]. Law enforcement agencies in Australia, the U.K., and the U.S. have used Clearview AI to scan social media images and identify suspects [1].

1.2.2 Data in Social Media. Millions of images and videos are published monthly on social media sites like Facebook, Instagram, X, and YouTube, resulting in extensive collections that can be processed and analyzed using advanced computer vision and FR technology. These collections pose significant privacy concerns due to the ease of deriving sensitive information, profiling chosen individuals, or identifying specific individuals and comparing them to existing databases without their consent [4].

Advancements in computer vision can create powerful, potentially intrusive systems capable of revealing a person's identity, including full name, address, and hobbies, from a single image [77]. Additionally, sensitive biometric data such as gender, age, or ethnicity can be directly extracted from facial characteristics.

1.2.3 Government Policies. Many individuals want more privacy protection while taking advantage of technological advancements. Governments worldwide are taking necessary actions to control biometric data use, collection, and processing due to privacy concerns in biometric recognition systems.

According to the current data privacy regulations, many countries have procedures for managing personally identifiable information. In the U.S., legislation has been enacted to protect FR data privacy. For example, Oregon and New Hampshire have banned the use of FR technology in police body cameras [114]. Major software providers like IBM, Amazon, and Microsoft have also considered suspending their FR initiatives [75].

Privacy and security are growing concerns linked to the need for trustworthy AI, especially for FR systems. Trustworthy AI considers legal and ethical aspects in AI system development. The **European Commission** (**EC**) has formed an expert committee to define ethical standards for AI, promoting a human-centric approach and establishing seven key requirements: Human agency and oversight, Technical robustness and safety, Privacy and data governance, Transparency, Diversity, non-discrimination and fairness, Societal and environmental well-being, and Accountability [34].

Significant legal progress in FR technology has been made in the last 10 years in Australia, the U.K., and the U.S. For example, Smith et al. [175] explored the ethical and legal implications of biometric recognition technology. In 2019, the Biometrics Institute established extensive privacy regulations for biometrics [15, 195], addressing issues of prejudice or harm caused by biometric systems. In 2017, Japan implemented the Act on the Protection of Personal Information [135]. One of the most crucial regulations for privacy protection is the **General Data Protection Regulation** (**GDPR**) of the **European Union** (**EU**) [195]. Additionally, the U.K. has enacted regulations concerning CCTV systems [170].

On December 1, 2021, the Italian National Assembly, for the first time in an EU member state, prohibited video-surveillance systems that employ FR technology [25]. This regulation prohibits private entities from applying these systems in public spaces, recognizing the threats to people's rights and freedoms. The moratorium will last until December 31, 2023, unless new biometric surveillance laws are introduced before then. Shops, athletic facilities, and public transportation will not be permitted to utilize FR technology in video surveillance systems; if they do, they risk administrative fines similar to those imposed by the GDPR.

Along with the GDPR, stringent privacy rules with extraterritorial reach have increased, impacting multiple economies globally. 120 nations have implemented privacy laws to ensure transparent

Table 1. International Privacy Laws and Their Key Characteristics

Law Name	Jurisdiction	Main Purpose	Effective Date
Personal Information Protection and Electronic Documents Act (PIPEDA) [144]	Canada	To regulate the use of personal or subjective information, recorded or not, about an identifiable individual.	April 2000
Personal Data Protection Act (PDPA) [69]	Argentina	To govern the lawful processing of personal data for individuals and entities in Argentina, including collection, storing, and sharing.	October 2000
Act on the Protection of Personal Information (APPI) [150]	Japan	To protect individual rights and the utility of personal information in the context of an advanced information society.	April 2005
Personal Information Protection Act (PIPA) [68]	South Korea	To safeguard personal information by regulating its collection, use, disclosure, and preventing abuse to protect citizens' rights.	September 2011
Protection of Freedoms Act (POFA) [170]	United Kingdom	To control the use and security requirements for surveillance and biometric data.	May 2012
General Data Protection Regulation (GDPR) [195]	European Union	To protect the fundamental rights of persons, particularly regarding their personal data privacy.	May 2018
Privacy Act 2020 [102]	New Zealand	To provide a framework for protecting an individual's right to privacy of personal information and access to personal information.	December 2020

data management, with further local regulations enforcing the responsibilities of data controllers and processors [169]. Table 1 provides an overview of significant privacy rules in various jurisdictions, emphasizing their primary objectives and dates of implementation, showcasing diverse global strategies for data protection.

Several **International Organizations for Standardization** (**ISO**) standards have been proposed to protect privacy regarding biometric technology [179]. These standards focus on system design and specifications to ensure privacy. Moreover, ISO/IEC 24745 [87] identified key security requirements for face-verification systems in the client-server model:

— **Irreversibility:** Biometric data should be irreversibly transformed before storage to prevent misuse.
— **Unlinkability:** Biometric data should not be linkable across applications to avoid unauthorized imitation.
— **Confidentiality:** Biometric data must be confidential to prevent unauthorized access and privacy risks.
— **Performance Preservation:** The protection scheme must not significantly impair biometric performance.

1.2.4 Academic Research. Extensive research interest in **privacy-preserving** (**PP**) techniques has been observed concerning the current privacy issues. To observe the growth in this interest, we conducted a keyword search on "**privacy-preserving face recognition**" (**PPFR**) tasks on the Web of Science platform [206] from 2009 to August 2022. Research interest in PPFR is clear, with 225 publications, and Figure 1 depicts the growth of this interest in the past few years.

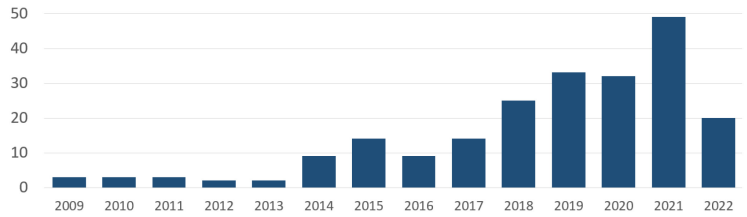

Fig. 1. Research trend for the keyword "Privacy-Preserving Face Recognition" until August 2022 [206].

Table 2. Comparison of PP Surveys

Ref	Year	Camera Level		Image Level		Face Level		Model Level		Database Level		Communication Level	
		A/L	PP method	A/L	PP method	A/L	PP method	A/L	PP method	A/L	PP method	A/L	PP method
VSN [205]	2014	✓	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
OSN [111]	2020	✗	✗	✓	✓	✗	✗	✗	✗	✗	✗	✗	✗
B–PET [125]	2021	✗	✗	✓	✓	✓	✓	✗	✗	✗	✗	✗	✗
CAFD [101]	2021	✗	✗	✗	✗	✓	✓	✗	✗	✗	✗	✗	✗
MIA [80]	2021	✗	✗	✗	✗	✗	✗	✓	✓	✗	✗	✗	✗
UDC [218]	2022	✗	✗	✗	✗	✗	✗	✗	✗	✗	✓	✗	✗
VP-DL [215]	2022	✗	✗	✓	✓	✗	✗	✓	✓	✗	✗	✗	✗
PPFR (Ours)	2022	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Symbol explanation: A/L: attack or leakage, PP method: privacy-preserving methods.

1.3 Contributions

Due to the importance of data privacy protection and the European **AI Act** (**AIA**) for trustworthy AI, FR models that preserve privacy are of interest to both academia and industry. Furthermore, CCTV-based surveillance-system privacy encompasses various topics in network security, computer vision, and deep learning. While this area has been gaining increased attention, no surveys have been conducted on FR privacy in the context of surveillance systems to date.

1.3.1 Existing Surveys. Existing surveys have primarily concentrated on one area of the FR privacy research field. This study collects and summarizes some existing surveys that are partially related to our work.

Winkler et al. [205] classified security features of **visual-sensor network** (**VSN**) applications, addressing security risks and potential attacks. Furthermore, Hu et al. [80] conducted a comprehensive survey on **membership inference attacks** (**MIAs**) and defenses, providing taxonomies for both. Moreover, Liu et al. [111] evaluated privacy requirements for image sharing in **online social networks** (**OSNs**), focusing on privacy demands over the lifespan of OSN image sharing. In addition, Zhang et al. [215] examined privacy concerns in visual data and deep learning, highlighting how deep learning can be used both for privacy attacks and preservation. Ying et al. [218] investigated differential privacy for unstructured data, summarizing privacy options for protecting such data before sharing with untrustworthy parties. Meden et al. [125] reviewed **Biometric Privacy-Enhancing Techniques** (**B-PETs**) for FR, assessing their efficiency. Laishram et al. [101] studied face deepfake challenges, analyzing manipulation techniques, datasets, and evaluation matrices. Table 2 compares these surveys based on attack surfaces, including camera, image, face, model, database, and communication.

Two crucial elements distinguish this survey apart from the mentioned ground-breaking efforts. First, earlier surveys were conducted using a problem-driven approach, which involved defining various privacy issues and seeking certain defenses. In contrast, we apply a life-cycle-based privacy-analysis paradigm to create a more comprehensive overview of this subject from a

user-centric standpoint. This survey offers a unique perspective in the field of CCTV-based surveillance systems by comprehensively evaluating challenges and solutions across every operational level'camera, image, face, model, database, and communication, resulting in a more thorough understanding of security in CCTV-based FR systems. Second, whereas other studies have generally focused on the privacy of an image or a model in isolation, we investigate how all modules interact while considering privacy concerns for a dynamic CCTV-based FR system.

1.3.2 Our Contribution. This article offers the first study that outlines PPFR systems, particularly privacy-leakage issues at each level of an FR system, and their PP solutions for complete FR models that can harm privacy or cause data leaks. This survey study provides a more in-depth examination of the concepts, theories, techniques, taxonomy-based categorizations, and future research prospects. The primary concern is the feasibility of creating a trustworthy FR system that complies with the PP criteria of most regulatory authorities.

In this regard, we proposed our contributions in five-folds:

- We identify the various privacy aspects of a CCTV-based FR system. We conduct a comprehensive analysis encompassing all dimensions rather than focusing on specific privacy areas or a single type of technology connected to FR.
- We provide a thorough and multidimensional view of PPFR by proposing a six-level-based taxonomy to carefully examine potential privacy-leakage problems in FR systems. The suggested six-level-based taxonomy includes the six essential components of a typical CCTV-based FR system.
- Analyze several incidents and potential data leaks for FR systems using the proposed analysis frameworks.
- We conduct a detailed review of several significant and representative PP methods for each of the six levels of FR systems proposed in the literature.
- We discuss insights and ideas regarding the current problems and research directions in this field.

The remainder of this article is organized as follows: The preliminaries and essential definitions are presented along with the overall architecture for PPFR in Section 2. We also introduce potential data leaks and attacks with respect to our defined levels. We analyze level-wise potential leaks by level in depth in Section 3. Section 4 presents level-wise privacy-preservation techniques, comparisons, and classifications of published works. Then, Section 5 identifies open issues and recommends future research. Finally, Section 6 concludes the article.

2 PPFR: An Overview

2.1 Preliminaries and Definitions

Based on biometric template protection, FR is intended to meet the major requirements of irreversibility and unlinkability defined in ISO/IEC IS 24745. Numerous explicit and implied references to secure facial identifications can be found in the AIA and ISO. Table 3 presents the selected few keywords from the complete list of definitions in the AIA, Article 3, and the ISO document. These terms primarily concern classifying biometrics and various attacks and their countermeasures.

We provide a brief introduction of the relevant terminology in the literature.

Homomorphic Encryption (HE): This is a form of encryption [162]. where an encryption algorithm E is called homomorphic over an operation $\odot$ if $E(m1) \odot E(m2) = \mathrm{E}\ (\mathrm{m1} \odot \mathrm{m2})$, $\forall$ m1, m2 $\in$ M, where M is the set of all possible messages. HE schemes can be classified into three types depending on the number [3] and the supported $\odot$ operations are: **partially homomorphic encryption (PHE)** [147], **somewhat homomorphic encryption (SWHE)** [20], and **fully homomorphic encryption (FHE)** [64]. PHE schemes enable a single operation type without

Table 3. Definitions Pertinent to Face Identification and Attacks in the EC's Proposed AIA [35]

Key term	Description
Biometric data	Personal information are obtained by processing a natural person's physical, physiological, or behavioral traits that enable or verify the individual identity of that person, such as facial photographs or dactyloscopy information.
Adversary	One who compromises an enrollment database may gain access to the generative biometric data of the individuals enrolled therein.
Biometric template protection	Protection of biometrics must ensure secrecy, irreversibility, and renewability during storage and transfer.
Privacy compromise	Event is when an adversary discovers part of the generative biometric data of an individual enrolled in the database of a biometric verification or identification system.
Privacy leakage	Amount of information about an individual's generative biometric data that an adversary can learn from the stored reference data.

usage limits and are the oldest HE schemes, supporting addition or multiplication. FHE schemes support unlimited operations and provide more robust encryption but are computationally costly. SWHE schemes allow multiple operation types but only a limited number of times. Due to FHE's performance issues, SWHE versions of FHE schemes have been proposed.

Secret Sharing (SS): In this cryptographic scheme, a secret key consisting of n shares can be reconstructed only if a sufficient number of shares are combined [171].

Secure Multiparty Cryptography (SMC): This cryptographic method enables distributed participants to calculate an objective function collaboratively without revealing their data [211].

Visual Cryptography (VC): This cryptographic method encrypts visual information so that the decrypted information appears as a visual image [131].

Transformation Techniques: Image transformation is a many-to-one function intended to purposefully alter a biometric image in a feature or signal space context, acting as an agent in template security without needing to be kept secret since it doesn't directly interact with raw biometrics.

Cancelable Biometrics (CB): Soutar et al. [178] introduced CB, which deliberately and consistently alters biometric signals through transformations. These transformations enhance privacy protection and template security in existing biometric systems and enable a comparison of biometric templates in the transformed domain.

Differential Privacy (DP): The DP privacy model provides the highest level of privacy by minimizing the chance of individual-record identification [91]. It limits the personal information revealed to third parties or adversaries, controlled by a "privacy budget" denoted as ϵ. There are two DP approaches: **global differential privacy** (**GDP**) [27] and **local differential privacy** (**LDP**) [91], with Laplace and Gaussian mechanisms being the most common noise-generation methods.

Synthetic Faces: Generating synthetic faces involves creating an independent dataset unrelated to the original dataset and is a separate method for face anonymization. Recently, attention has shifted to synthetic data production as a critical component of releasing privacy-protected data [105]. The main difficulties in creating synthetic faces include making fake faces to replace the real ones, preserving the original attributes, and assessing the value and effectiveness of the newly created image for anonymity.

Network Models: Neural networks are advanced techniques for creating synthetic data, capable of synthesizing unstructured data like images and videos and handling richer data distributions than conventional algorithms. They are also used in various FR applications, providing perfect classification accuracy. Some of these networks are discussed below:

— **Encoder-Decoder (ED) Network:** An ED architecture is an unsupervised method for generating synthetic data through a two-step transformation involving an encoder and a decoder. It maps a dataset's distribution to a latent space and then back. Symmetrical encoders and decoders form an **Autoencoder** (**AE**), aiming to reconstruct input images. A **Variational Autoencoder** (**VAE**) produces more disentangled latent spaces, generating synthetic face images from input vectors similar to the source but not identical.
— **Convolutional Neural Network (CNN):** The CNN is a neural network with convolutional layers that recognize and process image patterns. Convolutional layers use kernel filters to create feature maps. Pooling layers reduce image dimensions while up-sampling layers increase them. An ED network can be built with a CNN for image processing.
— **Generative Adversarial Network (GAN):** A GAN comprises two neural networks, the generator, and discriminator, which iteratively train the model [36]. The generator creates fake images from random sample data to fool the discriminator while the discriminator learns to distinguish between original and fake images. GANs are commonly used to generate realistic images and have diverse applications in synthetic data. **Conditional GAN** (**CGAN**) is a variant that applies a condition to the data during the generator and discriminator's training.

Federated Learning (FL): FL is an ML paradigm enabling collaborative model training across distributed devices while keeping data decentralized and ensuring user privacy. McMahan et al. [121] present FL as a solution for PP and efficient learning from distributed data sources. Local models are trained on respective devices, and only model updates are shared and aggregated to create a global model, providing PP training without exposing sensitive data to a centralized server.

2.2 Proposed Framework

The suggested framework uses FR technology and CCTV-based surveillance to improve FR capabilities in public areas. Moreover, FR technology and CCTV-based surveillance are inextricably linked because FR relies on CCTV cameras to gather and provide video footage for analysis and identification in public FR systems. The FR system collects facial traits from the recorded CCTV video and compares them to a database of known persons, allowing for the real-time identification of specific individuals. This connection considerably improves surveillance capabilities by allowing for the efficient tracking and identification of targets.

CCTV is video surveillance that employs a large number of cameras. It is linked to several monitoring systems, including schools, banks, airports, hospitals, and supermarkets. Initially, this technology could only view live videos, but it swiftly evolved into a recording system that could save data. A CCTV system typically consists of cameras and recorders. The cameras continuously capture images that are transferred wirelessly or through a cable to a display and recording device. Numerous studies have examined CCTV footage to understand better the contents of CCTV monitoring and the need to process them to yield distinct facts.

The fundamental goal of FR technology is to attempt to identify and confirm a person's identity by comparing various images with each other or to images in a database repository. While the major goal of the FR system is automated face recognition, the integration of CCTV-based surveillance within this framework serves the objective of security monitoring and aiding accurate and automated FR in public settings. The FR system uses the current CCTV infrastructure to perform face analysis on collected video data, enabling proactive identification and surveillance of individuals of interest.

This section proposes a CCTV-surveillance FR system derived from the literature, presents a full-flow system diagram, and highlights potential privacy breaches or attack types categorized by the attack surfaces. In a CCTV surveillance FR system, an attack surface refers to the specific level

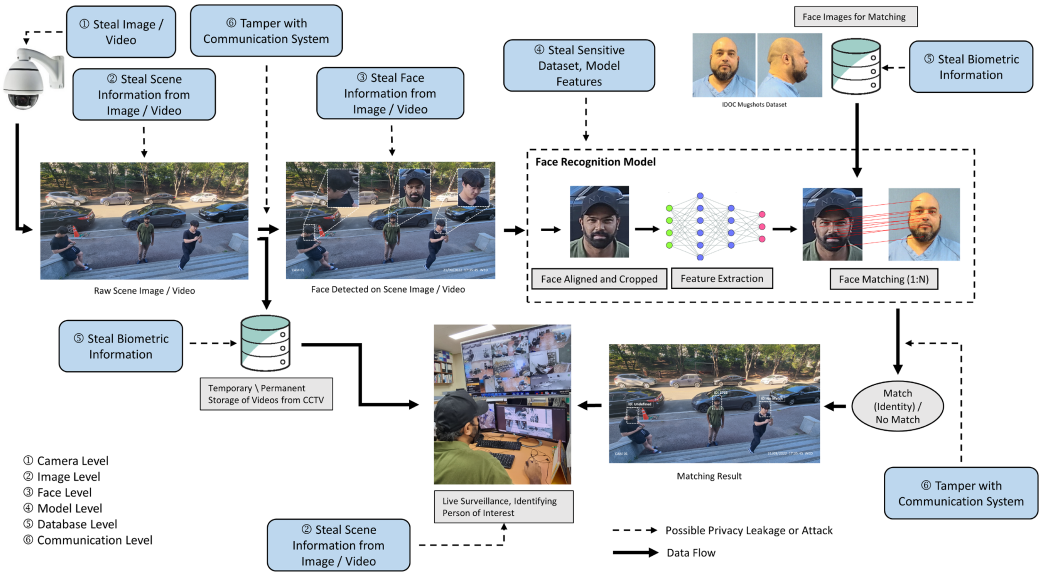

Fig. 2. A general CCTV surveillance FR system with proposed levels for privacy leakages.

or point within the system where potential vulnerabilities exist, such as the camera, image, face, model, database, or communication surfaces. Categorizing attacks by surface in these systems enables comprehensive analysis focused on security measures and effective resource allocation. It facilitates clear communication and ensures compliance with privacy and security standards, enhancing the overall impact of the research. This approach allows us to identify potential vulnerabilities and privacy breaches at specific points in the system, facilitating a comprehensive analysis and targeted mitigation strategies. We propose a privacy-analysis framework to aid in discovering various leaks and investigating privacy-preservation solutions. Using this framework, we conduct a fine-grained study and comparison of the data leaks and the PP solutions proposed in the literature in Sections 3 and 4, respectively.

The overall framework is shown in Figure 2. The system supports live video streaming from CCTV cameras and storage in a temporary database. Generally, storage for 30 days is recommended. In specific incidents that require investigation or reporting to authorities, at least 14 days of footage should be kept to provide sufficient time for review. The video from the camera is input into a monitoring system, and the entire footage can be viewed in real-time. Simultaneously, the video feed from the CCTV camera is fed into an FR system, which performs a series of processes to identify the face. First, it preprocesses the face (i.e., detects and aligns it). Facial features are extracted and matched with features in the database for the individual of interest. Once the matching operation is finished, the matching result is sent to the monitoring system to indicate whether a person of interest is in the surveillance area.

We classified potential leaks or attacks on the privacy of the FR system based on the attack surfaces. This method addresses the specific points in the system where privacy breaches are possible. These attack surfaces represent distinct levels of vulnerability within the architecture.

(1) **Camera Level:** A camera-level leak or attack occurs at the signal level during data capture, compromising the integrity and confidentiality of the transmitted signal. At this surface, adversaries employ various techniques to extract sensor data, signal, or streaming data directly from camera equipment. Their tactics may involve physical tampering with the camera or

intercepting data transmission. The primary objective in such instances is to gain unauthorized access to the raw footage captured by the camera.

(2) **Image Level:** An image-level leak or attack happens when the adversary attempts to steal information from an entire video scene or CCTV image. At this surface, leaks or attacks mainly involve obtaining or manipulating digitalized image data, including sensitive information beyond facial features. Attackers can exploit vulnerabilities at the image level to access all the information contained within the images captured by CCTV, posing significant privacy and data security risks.

(3) **Face Level:** A face-level leak or attack happens when the adversary attempts to steal the face image from a monitoring device, either from a handheld device or the monitoring system. At this surface, adversaries employ techniques to extract facial information from images containing detected faces. The primary objective is to acquire the captured face images, which may reveal sensitive facial characteristics and identity-related data. These attacks pose significant privacy risks and can be exploited for unauthorized identification or surveillance.

(4) **Model Level:** A model-level leak or attack happens when the adversary tries to obtain information about the model. The adversary can steal sensitive data or model features. On this surface, adversaries try to acquire confidential information about the FR model. This can be done using black-box or white-box techniques to access the model's internal workings, parameters, or learned features. By doing so, attackers may gain insights into the model's functioning or obtain data used to train it, potentially leading to privacy breaches.

(5) **Database Level:** A database-level leak or attack happens when an attacker steals biometric information from the database storing the FR-related information. Adversaries may attempt to gain unauthorized access to the database, manipulate, modify, delete, or copy the stored data, including personal information, face templates, or other relevant records.

(6) **Communication Level:** A communication-level leak or attack happens when an adversary tries to tamper with the communication system or eavesdrops on the channel to steal information or modify responses. Attackers at this surface can intercept or manipulate data exchanged through the channel, potentially compromising the transmitted information's confidentiality, integrity, or authenticity.

3 Privacy Leakages in FR Systems

The public and government may become very concerned about privacy due to data breaches. For instance, at the annual Black Hat Hacking Conference, hackers cracked Apple's iPhone Face ID in about 120 seconds, highlighting vulnerabilities and increasing the risk of face ID theft among major crimes [204].

In Section 2.2, we provided an overview of the CCTV-surveillance FR system and potential privacy leaks at various stages. To ensure a secure FR system, we identified stages where privacy could be compromised and discussed both accidental and intentional data leaks. We selected examples from all six stages where leaks can occur and elaborated on them in the following subsections.

3.1 Camera Level

Hacking is a serious risk with CCTV cameras. For example, if a camera with FR is tampered with in some fashion, a hacker may obtain identifying information about anyone in that camera's system. One of the most widely used methods for hiding face areas in surveillance recordings is post-processing [138]. In this method, videos with distinct face areas are initially recorded. Subsequently, specific face parts of the videos are either manually or automatically concealed. The difficulty with

Table 4. Attacks on CCTV Camera

Attack Event	Year	Location	Description
Brute-force attack [139]	2018	South Korea	Nearly 2,000 cameras with open IP addresses on the Internet were identified by Kaspersky.
Privacy hack [137]	2018	Japan	Hackers have officially disabled more than 60 Canon security cameras.
Command-and-control attack [8]	2019	India	CCTV cameras suffered backdoor accesses and remote hacking.
Privacy attack [83]	2021	USA	A breach of 150,000 security cameras, including banks, prisons, schools, the Tesla factory, and other places, was made.

this method is that videos with distinct facial regions are captured, transmitted, and processed. As a result, these face areas are susceptible to privacy leakage or hacker activities before they are preserved, as the privacy enhancement has not yet been applied [153, 216].

New exploits and vulnerabilities are continuously being used by attackers to target these systems [193]. Someone else could potentially see what a user sees when they utilize an Internet-connected camera from a computer. A "hacker" or an application to which we inadvertently or knowingly granted access permission can be that person. For instance, when a straightforward search term like "webcamXP" is entered on the Shodan search engine [172], one can view random video footage of urban areas, retail stores, and home spaces using an **Internet-of-Things** (**IoT**) search engine. Large-scale assaults on the IoT ecosystem have become common owing to resource limitations, obsolete firmware, inadequately protected IoT devices, and built-in vulnerabilities.

Numerous factors could lead to an attack on a CCTV system [185]. Examples of real-world attacks against CCTV surveillance systems are listed in Table 4. The most frequent causes for Internet-connected CCTV systems to be vulnerable include devices responding to ping requests, using default IP socket ports (such as port 80 for HTTP, port 21 for FTP, and port 23 for Telnet), weak or default passwords, unrestricted access, and flaws in the **network video recorder/digital video recorder** (**NVR/DVR**) video-management software or firmware.

3.2 Image Level

When people appear in images they are unaware of, privacy concerns can arise. These difficulties are exacerbated by the rising use of photo-capture devices in public places, such as smartphones, surveillance cameras, and wearable cameras. Humans can easily examine the rest of the image or video despite the faces being impossible to identify. Because each individual in a video/image must permit the use of his or her picture data, compiling high-quality datasets containing humans is particularly difficult for computer-vision researchers. The Duke **multi-target, multi-camera** (**MTMC**) dataset [161], a widely used resource for individual re-identification, was recently taken down for privacy concerns.

Image-level leaks in FR systems pose a privacy threat by unintentionally disclosing sensitive information. When visual data is not properly sanitized, attackers can extract not only facial identities but also additional details like location, time, the height of a person, clothes, and background from the images, compromising individuals' privacy. One potential vulnerability arises when visual content is shared with the public, as in the case of Google Street View. Insufficiently obscuring individuals' identities before image publication can enable malicious actors to draw undesirable inferences or engage in activities like voyeurism or stalking. This ease of access allows

attackers to leverage unprotected images to extract sensitive information, violating individuals' privacy [38].

3.3 Face Level

Face detection does not identify the faces it finds in images. It simply determines whether a face is present. It is a necessary component of the FR system, and if data is compromised there, it can be used for various things.

According to a Microsoft poll [127], 70% of recruiters in the U.S. have turned down applicants because of details, including images, they discovered online. The majority of British citizens supported this technology in 2019, according to the Ada Lovelace Institute [95], when they could see a value to society, such as in criminal investigations, unlocking cell phones, or checking passports. However, 29% of the people were uneasy about the police using this technology because they lacked confidence in their ability to use it morally.

Face images have a considerable potential for fraud. Medical records can be identified using facial photos. Companies that provide FR software have built enormous datasets to power their algorithms. The controversial Clearview AI, for instance, has 3 billion photographs it can search against (scraped from Google, Facebook, YouTube, LinkedIn, and Venmo). However, these technologies pose a significant security risk.

In the past, hackers have gained access to databases [62], including facial scans used by banks, police forces, and defense companies. This information can be used by criminals to stalk, harass, or perform identity fraud against victims. After Clearview experienced a data breach in 2020 [176], U.S. Senator Edward Markey stated that face-image data is not something one would want to have fall into the wrong hands. "If your password gets breached, you can change it. If your credit card number gets breached, you can cancel your card. But you can't change biometric information like your facial characteristics."

Although using our faces to unlock our iPhones or computers may appear harmless, this technology is increasingly being used to capture face images of children. FR technology is being used by Chinese gaming company Tencent to prevent kids from playing games between 10 p.m. and 8 a.m. [136]. Players must provide a facial scan to verify their age during this period.

Surveillance technology was tested in a Texas school district as part of a pilot program. There were 164,000 facial detections over the course of seven days, including 1,100 detections of one student [119]. Attackers may attempt to force an authentication if they cannot spoof a system. For instance, they may use a face image to unlock a phone [134].

3.4 Model Level

A model can be compromised when an adversary learns the features in the model structure or learns to identify a member in the database. Model leaks or attacks can be categorized into model-inversion attacks and MIAs.

Model Inversion Attacks: These attacks refer to removing private information or features from a publicly available model by running prediction queries against it. White-box attacks were launched by Fredrikson et al. against facial-recognition APIs based on neural networks by taking advantage of the confidence values that are frequently provided with predictions [57]. The authors showed that an attacker could reconstruct the image of a person using nothing more than their name or a unique identification number. The authors implemented it as an optimization problem where they attempted to find the maximum confidence value of the target class.

MIAs: Given a model and a record, an adversary can determine whether the record is included in the model training dataset. If the attack is successful, it suggests that individuals who provide training data may experience privacy leaks. For instance, a patient's medical record used to train

Table 5. Database Leaks and Attacks

Target	Year	Record Type	Record Size	Method
Apple [67]	2013	Tech, retail	12.3 million	Accidentally published
Marriott International [59]	2014	Personal information	500 thousand	Hacked
Facebook [143]	2019	User personal data	530 million	Hacked
Alibaba [207]	2019	User IDs, mobile numbers, and comments	1.1 billion	Hacked
Sina Weibo [33]	2020	User information	538 million	Hacked
Clearview AI [90]	2020	Face images	3 billion	Hacked
LinkedIn [55]	2021	User information	700 million	Hacked
Cognyte [155]	2021	User information	5 billion	Weak system
Accenture [140]	2021	Corporate and customer data	10 thousand	Stolen and leaked
Xinai Electronics [203]	2022	Faces and vehicle license plates	800 million	Leaked
NtechLab [14]	2022	Face-biometrics licenses	1.1 thousand	Leaked

a disease model can reveal that the patient has such a disease. Shokri et al [173] observed that training data respond differently to known ML models compared to those they have never "met" before. A shadow-training approach (black-box attack) was performed by the author to recognize those behavioral differences.

3.5 Database Level

One of the barriers authorities must overcome is that technology is evolving incredibly quickly, and transparency regarding what happens to saved face images, how they are processed, and for how long is insufficient. This situation is especially true for closed-box technology systems used by police forces. Table 5 shows a list of database attacks and leaks suffered by large technology companies.

3.6 Communication Level

An eavesdropping attack occurs when an adversary tries to intercept, delete, or modify data that are transmitted through an FR system. Image data transmitted across an open network, from the camera device to the recognition system and toward the database, can be obtained by an adversary. This results in privacy concerns about the face images. Adversaries can attack the channels where the FR result is transmitted. They can also override responses, such as changing a match to a non-match (or vice versa) or a person's identity information.

4 PP Methods

In recent years, many researchers have implemented numerous PP techniques in different applications. Here, we outline a number of PP techniques that can be used at each level of an FR system. We provide a brief overview of the advantages and disadvantages associated with existing defenses at each level, aiding researchers in identifying areas for further improvement and innovation.

4.1 Camera Level

Understanding and sensing persons and their faces is a key objective of computer vision and computational-imaging research. Researchers developed approaches to safeguard privacy by employing specifically engineered optical components (optical-level approach) or image sensors (sensor-level approach). The data generated by the image sensors are already secured before they are delivered to the host board within the camera. Thus, the susceptible period experienced by the post-processing-based solution is eliminated. Novel optics [115] and image sensors are offered as **integrated circuits** (**IC**) and optics-design advances, providing additional dimensions for privacy protection.

Privacy is also needed during image formation (i.e., during read-out, amplification, and digitization). This implies that privacy is ensured when signal measurements are taken, and there is no direct picture capture of sensitive facial information. At the CCTV/camera level, privacy is achieved either by direct managing from the device's sensors or by altering the raw retrieved data before they reach the device's general storage, i.e., processing and filtering.

Sensor Level: To protect privacy, a sensor-level technique employs specifically developed image sensors. Pittaluga et al. proposed a PP exposure-bracketing technique [153], which is useful for sensor-level thermal-image processing and offers the possibility of creating computational images. As incoming photons are transformed into electrical signals and then processed into picture measurements, privacy is enforced at the sensor level. A PP picture-generation process is offered by altering the sensor operations of gain, digitization, exposure duration, and bias voltage; no raw facial data are directly collected or stored. The ability to identify a person based on biometrics derived from photos, such as silhouette or gait, is a structural limitation. The method can remove facial information from the image.

Optical Level: Ordinary image sensors are typically used in optical-level approaches. However, specifically constructed optical components are used in front of the image sensors. For example, a collection of precisely crafted optical elements, made available by Pittaluga and Koppal [152], can carry out activities, such as monitoring and counting people, motion analysis, and so on, without recording any sensitive information. A technique for anonymous optical-level image sensing was introduced to improve privacy [116]. The device can optically conceal the facial regions, making it impossible to physically record any facial parts. The authors proposed a neural network that combines an AE for modeling typical behaviors with a 3D CNN for feature extraction. Because the prototype was large, its normal use was difficult. It is easier to execute optical-level solutions because they do not require specially created image sensors.

Other researchers proposed a scalable, affordable, and real-time de-identification camera as a solution to privacy-protection challenges [132]. The key contribution is a privacy-protection architecture that alters the video at the camera level itself. The system's implementation on a Texas Instrument OMAP4-based embedded platform enabled real-time video de-identification. Transforming the video on the camera itself also offers protection from various attacks.

Advantages: Image protection starts right at the moment of capture, adding an extra layer of guard against potential threats. Authenticity and integrity of images and video frames can also be established by embedding specific information in the image metadata to identify the original source. Protection in the early stage provides security in all the stages that follow.

Disadvantages: The problem with these approaches is that when the image sensor is constructed, the privacy-enhancing technique used by the image-processing circuit is fixed and typically difficult to change. IC design and manufacturing are both very complex and expensive. Specialized hardware or additional components lead to increased costs and performance for the system, resulting in slower image capture, processing, or transmission, which are undesirable in real-time applications. Providing privacy enhancement results in the increase of computational cost for the whole system.

4.2 Image Level

Online image sources may be used as data points in an intense and possibly invasive method of revealing a person's identity. Using a single image, this method could obtain a person's full name, address, and hobbies. Obfuscation is "the generation of noise modeled on an existing signal to make a data collection more ambiguous, confusing, harder to exploit, more difficult to act on, and hence, less useful." This process hides information from view, a promising method for protecting privacy. While using privacy-aware obfuscations, considering how they affect the aesthetics or visual appeal of the image is crucial.

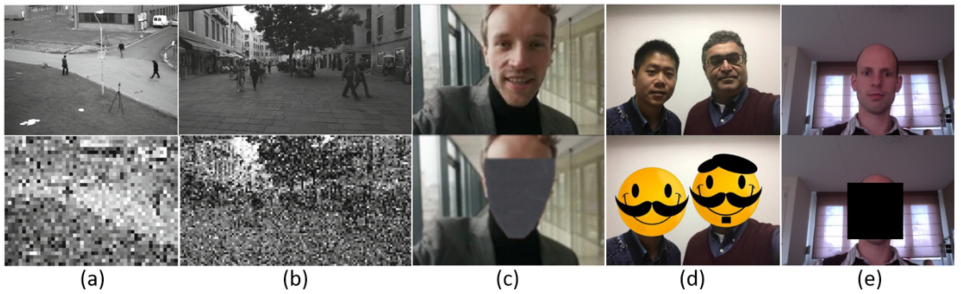

Fig. 3. Examples of PP methods in Image and Face levels. Image obfuscated method [52] on two datasets: (a) *PETS* dataset [151] with 768 × 576 resolution, and (b) *Venice* dataset [192] with 1920 × 1080 resolution. (c) anonymous face masking with a prototype camera [216], (d) visually obfuscated JPEG image [214], and (e) faces obfuscated using the IoT app [41]. The first row is the raw image, and the second is the corresponding obfuscated images. We used different examples for (c) and (e) because the authors did not provide the input faces.

Researchers have already proposed several obfuscation techniques to safeguard picture privacy, including pixelation, blurring, and masking; these protect privacy by obscuring important information in the images [23, 47, 74, 84, 108, 160, 194]. With blurring and pixelation techniques, the block size (blur radius) significantly influences the privacy-protection level. An excessively small value provides insufficient protection, whereas an excessively large value reduces the usability [92]. Furthermore, even with an optimal balance of the two, the provided anonymity level is still lower than some of the more current solutions [76, 122].

Masking methods typically use a mask square to cover the sensitive part of an image. This method provides good privacy protection but renders the image largely ineffective. Thus, to ensure maximum privacy, one can use these techniques. Blurring, pixelation, and redaction have long been employed to protect image privacy. However, such strategies are irreversible and ineffectual, particularly when compared to automatic recognition methods [76, 122, 138, 217].

Fan et al. [52] proposed a masking method by pixelization that protects sensitive information and represents the image with different pixel grids. They achieve this by performing differential privacy pixelization and experimenting on different grid cells, pixel sizes, and image utility with different privacy parameters. The results of the differentially private pixelization approach on two different datasets are shown in Figure 3 (a, b).

Bertalmio et al. described an inpainting method as a type of identity obfuscation [13]. A GAN creates an extremely realistic and organic image with the same data distribution as the original image. Although this technique partially satisfies the usability criteria, it does not completely safeguard the private data present in the original image.

Advantages: Image transformation and filtering methods conceal the image information as a whole. The method can hide important information relative to the person or the surrounding environment, like the present time of the incident, incident location, surrounding locations, weather conditions, the height of the individual, whom the individual is with, clothes they are wearing, and so on. The process can also be used to authenticate the raw images or videos from alteration by adversaries.

Disadvantages: Once the visual quality is reduced, it is difficult to identify and retrieve the original visual information. It can interfere with image analysis tasks. Many privacy enhancement techniques are irreversible, and accurately restoring the original details can be challenging. There should be a balance between the privacy protection and usability of the scene image or a way to consider the reversibility of the enhancement process.

4.3 Face Level

An innovative, effective PP face representation must safeguard the privacy of the face data while they are being stored and used in real-world scenarios. The core of these problems is maintaining privacy while allowing for viewing and using the image. Maintaining privacy while yet being able to view and use the image is at the heart of these problems.

Privacy techniques at the face level are categorized into face obfuscation, face de-identification, and face-swapping methods. Face obfuscation degrades or alters a face image so that it cannot be recognized with a face-biometric identification system. Face de-identification methods replace the original face with a synthetic face. Face swapping exchanges one person's face for another person's face.

The objective of privacy at the face level is to protect the privacy of individuals' faces and reduce the chance of information leaks. One of the earliest and most basic privacy-preservation methods concentrates largely on the use of common image-processing techniques, such as blurring [58], warping [177], and pixelation [168].

Face Obfuscation: Face obfuscation techniques aim at making the original face image unrecognizable by lowering its quality. The technique reduces or alters the face biometric data, which helps individuals avoid FR systems and enables the faces of the subjects to be used while preserving their privacy. Face obfuscation uses masking, filtering, and transformation methods. Figure 3 (c-e) shows some face obfuscation examples.

Face masking techniques, often considered the baseline for preserving face privacy, were demonstrated by Chinomi et al. [10]. They implemented the PriSurv surveillance system, which swaps people or their faces with masks or background pixels to protect privacy in video footage. Video subjects are identified via background subtraction, and faces are obscured using an adaptive masking technique following a predetermined privacy policy. This policy dictates the level of usefulness to be preserved and the type of visual abstraction, such as opaque or semi-transparent masks, to hide individuals in the monitoring data.

An anonymous camera, as used by Zhang et al. [216], employs privacy-enhancement masking techniques. The camera, equipped with a cold mirror, a liquid crystal on a silicon chip, and built-in charge-coupled device (RGB) and infrared image sensors, performs real-time facial masking based on thermal images recorded by the IR section. Heat signatures, which appear as warmer items in the image, are used to detect faces for masking. Schiff et al. [167] implemented a similar concept using visual color markers to identify sensitive face regions.

Wang et al. [197] demonstrated a real-world application for enhancing privacy using face masking and blurring methods with the RTFace system. This system consists of a face trainer for face identification and user registration and a privacy mediator determining whether a face should be visible or masked. Das et al. [41] also applied the RTFace concept in IoT-based applications, allowing consumers to choose to have their faces masked in the video.

Yuan and Ebrahimi introduced a reversible technique for face masking [214], known as JPEG transmorphing, which masks and alters visual regions of faces. Initially, sensitive image portions are masked to hide information, creating a binary mask of these regions. The unmasked areas are encrypted with symmetric encryption, and a secret key, and the transmorphing process overlays the mask and encrypted image on the obscured image, with decryption retrieving the original image.

Face filtering is a face-obfuscation process where linear or nonlinear filtering operations are applied to the face or specific regions in images or video frames, typically reducing image quality through blurring, averaging, and gradient techniques. Erdélyi et al. [48] introduced an adaptive cartoon-based filtering technique that abstracts actual photographs into cartoonish representations. These cartoons are superimposed on the original images using edge-intensity-based

weighting, edge detection, mean-shift filtering, and blurring. The simple computational processes allow for implementation in embedded intelligent cameras, and the cartooning approach enhances privacy by reducing biometric utility while still allowing behavioral inference.

Fradi et al. [56] address privacy concerns in crowd-surveillance applications using crowd-density data to determine the level of privacy augmentation needed for different input components while maintaining usability. A crowd-density estimation mechanism was developed using optical flow and local characteristics. Faces are detected with a histogram of oriented gradients-based detector. Privacy is enhanced through blurring and pixelation, providing greater protection in low-populated and less crowded areas.

Although most research on filtering-based privacy focuses on video and decreases in biometric utility [98], other studies have begun exploring filtering techniques that can preserve attribute information. Ye et al. [213] examined privacy-enhancement approaches like blurring that protect privacy while allowing accurate age assessment. Letournel et al. [104] proposed a technique preserving facial expressions using variational adaptive filtering while hiding identity. Fan et al. [53] implemented differentially private methods of pixelization and Gaussian blur, using the m-neighborhood concept to protect sensitive image information up to m pixels.

Privacy concerns have increased with the deployment of cameras on **micro aerial vehicles** (**MAVs**). Bonetto et al. [21] examined the efficiency of various filters for privacy protection in MAV-based video surveillance systems, while Sarwar et al. [166] proposed another privacy-filtering method for MAVs. These techniques are effective only at high image resolutions. To address reconstruction attacks on privacy filters, the author extended the work using adaptive Gaussian blurring [165].

Face transformation involves hiding facial images through various image transformations to decrease their utility by altering visual characteristics. Dadkhah et al. [39] applied several half-toning techniques, converting grey-level pixels into black and white dots, which are comprehensible for humans but not for ML approaches. The study examines the effects of techniques such as Stucki dithering, Bayern half-toning, Jarvis half-toning for grayscale photos, and Floyd–Steinberg dithering for RGB images.

Kobayashi et al. [96] introduced a privacy-enhancing surveillance camera system combining reversible watermarking with reversible mosaicking. Faces are divided into 8×8 pixel blocks, and quantized **discrete cosine transform** (**DCT**) coefficients are calculated. Low-frequency pixel values are encrypted using the advanced encryption standard and embedded in high-frequency pixels, with the encrypted data embedded in the DCT domain of the mosaic image via reversible watermarking. Hoshino et al. [78] later improved this design. Additionally, Fan et al. [54] proposed a metric privacy-based image transformation for **regions of interest** (**ROIs**), transforming the ROI into a feature vector and then performing an inverse transform, resulting in an obfuscated ROI image that maintains good utility and indistinguishability.

Reilly et al. [157] evaluated face obfuscation methods, synthesizing differentially private techniques including differentially private pixelization [52], pixel-sampling [196], singular value decomposition [54], and snow [89]. They assessed image utility, privacy risk, and computational time with varying parameter values and levels.

Face de-identification: Earlier approaches failed to prevent identification and compromised the integrity of the entire image, necessitating improvements in reliable FR software. These methods do not sufficiently conceal identity information and cannot provide complete anonymity. Additionally, reverse-transformation methods, such as de-pixelation and de-blurring, can recreate the original face from the output face. Instead of techniques that degrade face-image quality, privacy can be preserved through face de-identification by concealing identifying information or swapping the face with a synthetic one. This prevents the misuse of face data. Synthetic faces, created

from an independent dataset, are an increasingly important method for face anonymization and privacy protection [105].

The *k*-Same [183] family of methods laid the foundation for face de-identification algorithms by comparing each element of a set of de-identified faces to at least *k* elements of a set of person-specific faces without discrimination. The vanilla *k*-Same approach [138] produces a fresh face by averaging it with the *k* closest in a collection, often resulting in ghosting artifacts. To preserve original face characteristics, the *k*-Same-Select approach [66] divides a gallery into mutually exclusive subgroups and was later upgraded with the active appearance model method [43], successfully removing ghostly aspects on synthetic faces.

However, the *k*-Same methods lack generalization, requiring a sizable and diverse group of faces with each subject appearing only once. Moreover, these techniques cannot use all possible classifiers for facial attributes. To address these issues, Guo et al. [71] introduced k-Dive-Net, a deep-learning-based model that improves on the *k*-Same-Net model by integrating diversity, ensuring that de-identified faces are distinct in expression and identity, moving beyond the limitations of the *k*-Same family techniques.

Generative deep models like GANs have become popular for preserving the privacy of face images. Chen et al. [29] presented a GAN-based privacy preservation technique using synthetic faces to replace real ones. Their method includes an attribute-transformation module that preserves key facial attributes. They evaluated the utility of the images by analyzing how closely the attributes of the synthetic and real photos matched, using this similarity to gauge the degree of privacy. Sun et al. [181] introduced GAN-based methods using **3D morphable models** (**3DMM**) [18] for face de-identification. This approach involves two stages: first, creating a synthetic face with 3DMM to hide identity information while preserving utility, and second, rendering the entire head with a GAN-based inpainting method to integrate it with the rest of the original image.

Wu et al. [208] proposed PP-GAN, an image-synthesis technique that generates de-identified faces from a single face image using a U-Net-based ED network with a shallow convolutional discriminator. Hao et al. [73] introduced the utility-preserving GAN, which creates surrogate faces for de-identification while maintaining biometric characteristics such as age, gender, skin tone, stance, and expression without revealing identity.

A generator-based approach [117] separates facial identity from other image elements, anonymizing specific faces in images and videos by modifying identity vectors in the generator's identity representation space. This method produces several anonymous faces from a single original face while maintaining high similarity. Cosine similarity is used to gauge image usefulness. Meden et al. [123] proposed another generative network for deidentification, using the centroid of the mapped cluster and the input test image, though it does not preserve facial attributes. Cho et al. [31] introduced CLEANIR, a VAE that learns a disentangled representation, separating identity from other properties. This decoupled latent space allows for editing identification information separately, creating images similar to the originals in all but identity, thus concealing identification information.

A temporally consistent video de-identification technique based on GAN inpainting was proposed in [11]. The authors created a diverse facial video dataset to develop and test their GAN-based method, which inpaints faces into anonymous images and videos without relying on landmarks. They assess privacy utility by measuring the average changes to the identity of generated faces across successive frames. Additionally, [129] presented a photorealistic de-identification approach that replaces distinct facial features like the nose, chin, and eyes with those from different individuals.

A CGAN-based approach for face de-identification was proposed in [82], successfully removing 100% of sensitive information from the original faces. They introduced the **Flickr Diverse**

Faces (**FDF**) dataset, containing 1.47 million human faces with bounding boxes and keypoint annotations, achieving higher accuracy in face detection than traditional methods despite increased training time. Maximov et al.'s **conditional identity-anonymization GAN model** (**CIAGAN**) uses an ED network to swap faces by masking and inpainting facial regions with synthetic surrogate identities. A StyleGAN-based approach in [130] maps latent space to produce high-quality (1,024 × 1,024) anonymized images, evaluated using Euclidean distance and FID score.

Another alternative approach is the ED architecture for face de-identification [93], which proposed a **privacy-preserving protector network** (**PPAPNET**). This method creates realistic synthetic images from input images, protecting them from model-inversion attacks. The PPAPNet framework uses adversarial-mode training of the protector and attacker to optimize image security. Privacy gain, measured between the original and reconstructed images, indicates privacy performance, with a higher gain reflecting better privacy protection.

Another ED method with DP [219] enforces privacy in high-semantic latent vectors for face images, producing high-quality images with adjustable privacy protection and achieving a balanced tradeoff between visual quality, privacy, and utility. Li and Lyu [107] developed a synthesis-based ED network for privacy improvement through facial-attribute transfers, transferring only non-identity-related attributes from the source to the target face. Different decoders are used for different identities, while the same encoder ensures each decoder maintains identity-specific properties for mapping onto target faces.

Croft et al. [37] introduced a GAN framework using differential privacy for face de-identification, which does not require subject-specific image collections. Wen et al. [202] proposed a minimal-altering de-identification approach that allows individuals to control their degree of anonymity using differential privacy. The network generates disentangled identity and attribute representations, then produces anonymous faces through Laplace mechanism perturbation, requiring few resources and having lower computational complexity.

An approach similar to those mentioned above was proposed in [124], but it was ineffective in practice as the subject's identity could still be visually determined. A combination of VAE and GAN (PPRL-VGAN) using a conditional identity vector for privacy preservation was highlighted in [28], allowing for controlled de-identification. However, the control vector is a one-shot encoded vector, limiting the de-identification range in the training set. [82] proposed preserving facial position using a facial-landmark detector to generate a random face with the same pose, but GAN-based limitations mean the new face may not be distinct from the input face.

Another method [106] defined identity as a combination of facial characteristics, employing 40 classifiers to predict facial traits and selecting which attributes to maintain based on privacy requirements. A new face was then created using the recorded data. However, the objective of preserving useful information about the original face is compromised by the high processing capacity required to run 40 classifiers.

A reversible de-identification method called UU-Net [154] trains two U-Nets in succession: one for de-identification and the other for reversing the process. The first U-Net enhances video for data analysis and sharing, while the second is used privately to restore original data when required by authorities. Another method by Wen et al. [201] introduces a protection module to safeguard identity information and a recovery module to restore the original identity using an identity-disentanglement network guided by a flow module. The visual output of these reversible methods is shown in Figure 4.

Face Swapping: Face swapping replaces a face image with another real face image. The facial exchange method based on morphable models [17] was a great innovation in face-swapping, utilizing 3DMM [18]. However, these methods require numerous images of the target and source



Fig. 4. A reversible de-identification method with identity protection and recovery on the VoxCeleb dataset [201].

subjects to construct a 3D model, limiting their effectiveness. Estimating 3D geometries using 3DMM under various lighting conditions remains challenging.

The success of deep learning has led to numerous deep-learning-based methods. The authors of [99] developed a novel face-switching technique for the style-transfer problem, viewing identity and facial characteristics as style aspects. **Region-separative GAN** (**RS-GAN**) [133] swaps faces in the latent space by detaching the latent representations. **Face-swapping GAN** (**FSGAN**) is the first technique to simultaneously manipulate expression and identity while generating high-quality, temporally coherent outcomes without requiring person- or pair-specific training [141]. It employs a novel reenactment-based method for continuous interpolation between different views of the same face, producing realistic, temporally consistent results with an end-to-end trainable model.

Table 6 summarizes the surveyed techniques for the face level. The table compares the techniques in terms of (i) the type of input data, (ii) the testing datasets for these techniques, (iii) the mechanism used to model such techniques, (iv) the type of techniques for preserving privacy at the face level, and (v) to check whether the technique fulfills privacy requirements. The privacy requirements are the ISO standards discussed in Section 1.2.3 and are marked as "✓" if they fulfill the requirement and "✗" if not.

Advantages: By blurring or pixelating faces, it becomes more difficult for others to identify or recognize individuals, thus reducing the risk of potential harm, such as unauthorized surveillance or stalking. Face privacy enhancement enables greater flexibility in sharing and using images without infringing privacy rights. Face deidentification techniques preserve the anonymity of individuals and protect their privacy while still being able to utilize facial attributes for analysis or training ML models.

Disadvantages: Face obfuscation techniques can completely prevent facial recognition. While this may be desirable in certain scenarios for privacy reasons, it can also limit the ability to accurately identify individuals, which can have practical implications in areas such as security, law enforcement, or access control systems. The method introduces additional computational complexity and resource requirements, leading to performance and scalability in real-time scenarios. Face deidentification can hinder or even prevent accurate identity change, and it mainly depends on whose face has been replaced, which leads to another privacy concern for the replaced face. It can also impede facial properties like facial emotions and attributes. One may attempt to reverse-engineer or deobfuscate the obscured faces using sophisticated algorithms.

4.4 Model Level

Obtaining a fully functional PPFR system is challenging, but some methods have been successfully designed. If an FR system is trained with PP techniques, model-level attacks, like leaks in model features, are inconsequential as augmented face images lack identity information. We categorized

Table 6. PP Techniques for the Face Level

Ref.	Year	Applied to		Database	Mechanism	Privacy Method	Privacy Requirements			
		Images	Videos				IR	UL	CF	PP
[10]	2008	✓	✓	IHD	Adaptive masks	Face obfuscation	✓	✗	✗	M
[98]	2012	✓	✓	IHD	Various filters	Face obfuscation	✓	✗	✗	M
[56]	2013	✓	✗	PETS2009	Blurring and pixelation	Face obfuscation	✓	✓	✓	M
[96]	2014	✓	✓	IHD	Reversible mosaicking	Face obfuscation	✗	✓	✓	M
[48]	2014	✓	✓	Pevid	Cartooning	Face obfuscation	✓	✓	✓	M
[104]	2015	✓	✓	LFW	Adaptive filters	Face obfuscation	✓	✓	✓	M
[21]	2015	✓	✓	IHD	Various filters	Face obfuscation	✓	✓	✓	M
[165]	2016	✓	✓	Dataset from [79]	Adaptive blurring	Face obfuscation	✓	✓	✓	M
[167]	2014	✓	✓	IHD	Opaque masks	Face obfuscation	✓	✓	✓	M
[214]	2017	✓	✓	PIPA	Arbitrary masks	Face obfuscation	✗	✓	✓	M
[197]	2017	✓	✗	LFW	Denaturing masks	Face obfuscation	✓	✓	✓	M
[41]	2017	✓	✓	IHD	Opaque masks	Face obfuscation	✓	✓	✓	M
[213]	2018	✓	✗	Adience	Various filters	Face obfuscation	✓	✓	✓	M
[180]	2018	✓	✗	PIPA	AE + GAN	Face obfuscation	✓	✓	✓	H
[78]	2018	✓	✓	IHD	Reversible mosaicking	Face obfuscation	✗	✓	✓	M
[39]	2018	✓	✓	IHD	Half toning	Face obfuscation	✓	✓	✓	M
[71]	2018	✓	✗	BU-3DFE	GNN	Face de-identification	✓	✓	✓	H
[208]	2018	✓	✗	FERG, MUG	GAN	Face de-identification	✓	✓	✓	H
[123]	2018	✓	✗	XM2VTS	GNN, k-Anonymity	Face de-identification	✓	✓	✓	M
[28]	2018	✓	✓	FERG, MUG	GAN, VAE, Synthetic face	Face de-identification	✓	✓	✓	H
[148]	2019	✓	✗	RAFD, CelebA	GAN	Face de-identification	✓	✓	✓	H
[107]	2019	✓	✗	LFW, PIPA	Attribute transfer	Face de-identification	✓	✓	✓	H
[106]	2019	✓	✗	CelebA	GAN	Face de-identification	✓	✓	✓	H
[82]	2019	✓	✓	WIDER-Face, IHD (FDF)	GAN	Face de-identification	✓	✓	✓	H
[82]	2019	✓	✗	WIDER-Face	CGAN	Face de-identification	✓	✓	✓	H
[141]	2019	✓	✓	Forensics++, CelebA	RNN, GAN	Face swapping	✓	✓	✓	H
[93]	2019	✓	✗	CelebA	ED	Face de-identification	✓	✓	✓	H
[73]	2019	✓	✗	FaceScrub	GAN	Face obfuscation	✓	✓	✓	H
[37]	2019	✓	✗	RAFD, KDEF	GNN	Face obfuscation	✓	✓	✓	H
[166]	2019	✓	✓	LFW, IHD	Adaptive blurring	Face obfuscation	✓	✓	✓	M
[120]	2020	✓	✓	CelebA, MOTS, LFW	AE, GAN	Face de-identification	✓	✓	✓	H
[31]	2020	✓	✓	LFW, JAFFE, MUG	AE	Face de-identification	✓	✓	✓	H
[154]	2020	✓	✓	IHD (BIODI), MARS, P-DESTRE	AE	Face de-identification	✓	✓	✓	H
[29]	2021	✓	✗	CelebA	StarGAN, STGAN, Synthetic face	Face de-identification	✓	✓	✓	H
[130]	2021	✓	✗	CelebA	StyleGAN + Synthetic face	Face de-identification	✓	✓	✓	H
[117]	2021	✓	✓	LFW, FaceForensics++	GAN + ED	Face de-identification	✓	✓	✓	H
[11]	2021	✓	✓	FDF	GAN + U-Net	Face de-identification	✓	✓	✓	H
[202]	2022	✓	✗	CelebA-HQ	CGAN, Differential privacy	Face de-identification	✓	✓	✓	H
[219]	2022	✓	✗	CelebA, Flickr-Faces-HQ	GAN + ED	Face de-identification	✓	✓	✓	H
[201]	2022	✓	✓	CelebA-HQ, VoxCeleb	Ukey, AE	Face de-identification	✗	✓	✓	H

Symbol explanation: AE: autoencoder, CF: confidentiality, CGAN: conditional generative adversarial network, ED: encoder-decoder, FDF: Flickr Diverse Faces, GAN: generative adversarial network, GNN: generative neural network, H: high preservation IR: irreversibility, M; medium preservation, PP: performance preservation, RNN: recurrent neural network, UL: unlinkability, VAE: variational autoencoder.

all existing PPFR techniques into hand-crafted and deep learning approaches applicable to methods like face identification and face verification.

4.4.1 Hand-crafted Approaches. Numerous approaches use hand-crafted methods for PPFR, categorized by face image representation and techniques used. These include HE [50], index-based feature extraction [145], sparse representation [88], differential privacy [26], VC [158], and randomness techniques [72].

HE: Erkin et al. [50] introduced PPFR, a secure two-party protocol between the camera and database server, operating in a semi-honest model where participants are **honest but curious** (**HBC**).

Using a standard Eigenface [188, 189] recognition system, the method processes encrypted images through **principal component analysis** (**PCA**) and employs **additive homomorphic encryption** (**AHE**) with the Paillier [147] and DGK [40] cryptosystems. The matching process involves projecting the image into a feature vector, calculating distances between image and database features, and selecting the minimum distance, all in the encrypted domain, without degrading the reliability compared to standard Eigenface recognition.

Sadeghi et al. [164] improved Erkin's method [50] by enhancing communication and computational efficiency through a hybrid cryptographic scheme combining HE and **garbled circuits** (**GCs**) [110, 212]. Their method uses HE for feature projection and distance calculation and GCs for the minimum-distance phase. Additionally, they implemented a protocol entirely using GCs. Boddeti et al. [19] introduced a FHE-based framework for securing face template databases, enabling face verification through template matching in the encryption domain.

Bringer et al. [24] analyzed PP face-biometric identification using various distance metrics and oblivious transfer [97, 126]. Upmanyu et [190] proposed a blind and secure biometric authentication system utilizing public-key cryptography for secure communication and a generic linear classifier, such as a support vector machine with multiplicative HE for privacy. Testing on an encrypted face dataset showed no impact on accuracy.

Previous PPFR algorithms face scalability issues due to high computational and communication costs. Bentafat et al. [12] proposed a PP video surveillance architecture for identification that maintains an encrypted suspect database for video cameras. Their camera and server protocols reduce communication to a single round, expedite computation time, and instantly reveal suspect identification results through extensive precomputations. The authors used OpenFace [7], ensuring that the server receives trivial information insufficient for identifying individuals.

Index-based Feature Extraction: Osadchy et al. [145] proposed an index-based secure computation face-identification method, utilizing an index-based feature-extraction technique instead of the eigenface method in an "honest but curious" manner. Full faces are represented in appearance and spatial components, with face representation performed on the client end and secure computation on the server end using AHE with oblivious transfer. This method performs well on faces with varying illumination and occlusions, making it applicable to real-life scenarios.

Differential Privacy (DP): A PPFR protocol called **privacy using eigenface perturbation** (**PEEP**) [26] applies local DP in three steps: acquiring original faces, producing eigenfaces, and randomizing images with Laplacian noise. A multilayer perceptron classifier is trained using only the randomized face image. Eigenfaces hide some features due to dimensional reduction, and adding noise with PEEP using ϵ enhances privacy. PEEP achieves acceptable accuracy when the privacy budget is between 4 and 8 ($0 < \epsilon \leq 9$). Yan et al. [210] proposed another DP-based PPFR scheme, DP-Face, using a Siamese network, balancing utility and privacy through the privacy budget.

VC: A PP biometric-recognition system based on VC was proposed by Ren et al. [158]. They implemented an extended VC scheme to maintain the size of the recovered images, using block encryption instead of pixel-by-pixel before and after encryption. They also utilized transfer learning to expedite and increase the learning efficiency of the model. The significant generalization capability of transfer learning also eliminates noise interference for images recovered from EVCS. The results of the method show high accuracy in different experiments for feature recognition while maintaining privacy.

Sparse Representation: Jin et al. [88] proposed a new approach to PP face identification using sparse representation and cryptographic tools. A dictionary representation is learned from a third-party face database, with sparse vectors representing input and cloud-stored faces. Matching is performed in an encrypted space using Paillier encryption, ensuring that only the client knows the matching results while the server remains unaware.

Wang et al. [199] designed a FR framework for edge and cloud networks to coordinate sparse coding. They developed a low-complexity encryption technique based on random unitary transformations, ensuring privacy without affecting FR accuracy. Utilizing a distributed-learning framework, the system leverages multi-device diversity and minimizes data transmission between edge and cloud, making it suitable for general and large-scale settings. Sun et al. [182] proposed a sparse-matrix transformation method for encryption/decryption that significantly reduces the computational load on the local client due to the simplicity and sparsity of these matrices. Designed under a dishonest-and-curious single-server model, the large matrix sizes ensure security. The algorithm's strictly one-way privacy allows the client to detect dishonest server behaviors with an optimal probability of one.

Randomness Technique: Guo et al. [72] proposed a PPFR algorithm using randomness techniques involving various affine transformations such as shift, diffusion, and permutation. A **pseudorandom generator** (**PRG**) with a secret key performs the transformations. First, permutation transformations by the PRG scramble the face images. Second, a random matrix is used for diffusion encryption, and finally, a shift operator modifies the properties of the previous transformations. The final encrypted face is used for the FR process on the server. This encryption is applied to the original face, random face, mean face, and eigenface.

Xue et al. [209] introduced PPFR for IoT systems using Bloom filters to generate a system-representation protocol. This implementation meets the resource constraints of IoT devices by converting input faces or face features into a Bloom filter space in a trusted environment, allowing all utility operations to be performed in an untrusted environment. Another PP face-identification method was introduced, using a product-quantization-based hash lookup table to index and retrieve protected face templates [146]. FHE techniques ensure the privacy protection of enrolled subjects' face templates. Experimental assessments in closed-set and open-set environments demonstrated the viability of this method for use in large-scale FR systems.

Ergun [49] proposed a PPFR framework that combines image-processing operations and cryptographic techniques. Faces are recognized in an encrypted domain using a fast and robust cryptographic method based on the **truly random number generator** (**TRNG**) [45] and one-time-pad encryption. Cryptographic keys, generated using a TRNG based on a continuous chaotic function, encrypt faces via a bitwise XOR operation on each image pixel. The encrypted faces are then used for PCA-based FR.

Liu et al. proposed a face-encryption technique using a chaotic method [113]. The face-feature extraction was performed using a ridgelet DCT, and the face images were encrypted using the Tent–Henon double chaos model. A neural network based on PCA was designed to recognize the encrypted face images.

CB: The methods discussed above are CB that use facial features and perform a transformation to randomly project the feature projection, bypassing the FR model. The methods can be further classified based on the type of face-biometric projection: (i) Geometric Transforms [94], (ii) Random Projections [149], (iii) Filters based [156], (iv) Bloom Filters [209], (v) Hashing [146], and (vi) Hybrid Methods [221]. The cancelable biometric template can also be obtained by combining two or more methods which are represented as the hybrid method.

4.4.2 Deep Learning Approaches. With the interest in PPFR technology, deep learning models are increasingly used in these applications. Some techniques are deep learning approaches with HE [118], and other uses with synthetic faces [159] or masked faces [198].

HE: Ma et al. [118] proposed a face-verification method using deep learning. They used CNN for facial feature extraction and Paillier encryption for privacy, storing features in a database. The model operates as a three-party system with a client, data server, and verification server. The client

collects and sends information to the servers; the data server encrypts and saves the information, while the verification server handles verification operations. Experimental results confirmed data security but a slight reduction in accuracy.

Huang et al. [81] proposed an efficient PP face-verification method using FHE and GCs. The system uses a three-party model, where the client captures the user's face image, and the server stores the encrypted face image and provides a verification service. Finally, a cryptographic service provider maintains public and private keys. A pretrained FaceNet network is used for feature vector generation from a face image. The face verification protocol is fast and takes less than a second.

Synthetic Faces: Ren et al. [159] proposed a principle approach for training a video face anonymizer. They performed adversarial training where a video anonymizer modifies the original video by removing the privacy-sensitive information while a discriminator tries to extract privacy-sensitive information from the anonymized videos. The method produced an anonymized video while maintaining the detection performance.

FaceMAE is a cutting-edge framework based on face masking that uses an instance-relation matching module to minimize the distribution gap between original and reconstructed faces, balancing privacy protection and recognition performance [198]. The framework involves randomly masking the original face and training a deep-learning model for reconstruction. The authors achieved a 50% error rate reduction on several public datasets and decreased privacy leakage to 20%. Boutros et al. [22] investigated the viability of trained FR models using only synthetic face datasets. They proposed an accurate FR method called SFace, using a class-CGAN to generate class-labeled synthetic face images. They trained an FR system using three alternative learning strategies: multi-class classification, label-free knowledge transfer, and a combination of both, as the first step towards generating a privacy-friendly dataset.

FL: FedFace [5] stands out as the pioneering work in addressing FR model training within the federated setup. it improves the performance of the CosFace FR system by leveraging distributed face data in a federated setup, preserving data privacy. It utilizes a mean feature initialization scheme and a spreadout regularizer to enhance the separation of class embeddings. In contrast, [112] takes a different approach by prioritizing privacy protection and scalability. They avoid transmitting identity prototypes to prevent potential leaks of private identity information from clients. Furthermore, their proposed framework, FedFR, is designed to be scalable across various scenarios where each client may possess multiple identities. Through joint optimization, FedFR improves the overall face representation of the global model and enhances the personalized user experience.

Table 7 summarizes all surveyed techniques for the model level. Similar to the face level, the model level table compares the techniques in terms of the (i) input data type, (ii) testing datasets, (iii) design mechanism, (iv) technique types, and (v) whether the technique fulfills privacy requirements. The privacy requirements are the ISO standards discussed in Section 1.2.3 and are marked as "✓" if they fulfill the requirement and "✗" if not.

Advantages: The FR accuracy of cryptography-based schemes is always similar to the original FR schemes, which ensure security while providing accurate measures. Realistic faces with different identities can be achieved while preserving privacy and maintaining the attributes of the original face. Moreover, different FR models can be designed in relation to the modified input data, in which adversaries will find difficulty in accessing the model.

Disadvantages: During the face-verification protocol, the server must not learn any information about the face image, intermediate results, or verification results, or the client could learn more than the verification results. Due to the complex cryptographic techniques, such as HE algorithms, encrypting faces and performing FR operations is computationally inefficient and time-consuming.

Table 7. PP Techniques for the Model Level

Ref.	Year	Applied to		Database	Mechanism	Privacy Method	Privacy Requirement			
		Images	Videos				IR	UL	CF	PP
[50]	2009	✓	✗	ORL	S2PC, AHE	Homomorphic encryption	✗	✗	✗	M
[164]	2009	✓	✗	ORL	S2PC, AHE + GC	Homomorphic encryption	✗	✗	✗	M
[145]	2010	✓	✓	CMU-PIE, FERET	S2PC, Index, AHE + OT	Homomorphic encryption	✓	✓	✓	H
[190]	2010	✓	✗	YaleB	SMC, MHE, SVM	Homomorphic encryption	✓	✓	✓	H
[186]	2013	✓	✗	FERET, XM2VTS, LFW	FHE + GC, SVM	Homomorphic encryption	✓	✓	✓	H
[49]	2014	✓	✗	ORL, CMU	TRNG, Chaotic system	Randomness technique	✓	✓	✓	H
[24]	2014	✓	✗	ORL	AHE + OT	Homomorphic encryption	✗	✗	✗	M
[88]	2015	✓	✗	Own	AHE + OT	Sparse representation	✓	✓	✓	H
[118]	2017	✓	✗	LFW, Face94	CNN, CSP, AHE	Homomorphic encryption	✓	✓	✓	M
[159]	2018	✓	✓	CASIA-WebFace, LFW	GAN	Deep learning	✓	✓	✓	H
[19]	2018	✓	✓	LFW, IJB-A, IJB-B, CASIA	FHE	Homomorphic encryption	✓	✓	✓	H
[42]	2019	✓	✗	FERET	CKKS and BFV, FHE	Homomorphic encryption	✓	✓	✓	H
[199]	2020	✓	✓	YaleB	Bloom filter, SVM	Sparse representation	✓	✓	✓	H
[26]	2020	✓	✗	CelebA, LFW	Laplacian noise	Differential privacy	✓	✓	✓	M
[103]	2020	✓	✓	CASME II, Honda/UCSD	LBP-TP, SVM	Sparse representation	✓	✓	✓	H
[100]	2021	✓	✓	LFW	BLS short signature, FaceNet	Randomness technique	✓	✓	✓	H
[12]	2021	✓	✓	Own	S2PC, AHE	Homomorphic encryption	✓	✓	✓	H
[113]	2022	✓	✗	ORL	Ridgelet transform, THM, Chaotic system	Randomness technique	✓	✓	✓	H
[182]	2022	✓	✗	AR face	SRC based	Sparse representation	✓	✓	✓	H
[158]	2022	✓	✗	IMM Face, LFW	Secret sharing	Visual cryptography	✓	✓	✓	H
[81]	2022	✓	✗	LFW	Facenet, CSP, FHE + GC	Homomorphic encryption	✓	✓	✓	H
[198]	2022	✓	✗	LFW, CFP-FP, AgeDB	Masked AE	Deep learning	✓	✓	✓	H
[22]	2022	✓	✗	LFW, CASIA-WebFace	GAN, synthetic faces	Deep learning	✓	✓	✓	H
[5]	2022	✓	✗	CASIA-WebFace, LFW, IJB-A, IJB-C	Federated learning	Deep learning	✓	✓	✓	H
[112]	2022	✓	✗	MS-Celeb-1M, IJB-C	Federated learning	Deep learning	✓	✓	✓	H

Symbol explanation: AE: autoencoder, AHE: additive homomorphic encryption, BFV: Brakerski–Fan–Vercauteren, CF: confidentiality, CKKS: Cheon–Kim–Kim–Song, CNN: convolutional neural network, CSP: cryptographic service provider, FHE: fully homomorphic encryption, GAN: generative adversarial network, GC: garbled circuit, GNN: generative neural network, H: high privacy protection, IR: irreversibility, LBP-TP: local binary-pattern temporal planes, M: medium privacy protection, MHE: multiplicative homomorphic encryption, OT: oblivious transfer, PP: performance preservation, S2PC: secure two-party protocol, SMC: secure multiparty computation, SVM: support vector machines, THM: Tent–Henon map, TRNG: truly random number generator, UL: unlinkability, VAE: variational autoencoder.

This limitation renders it unsuitable for tasks that need to operate on high-dimensional vectors. VC is primarily designed for simple images with binary pixels, and it struggles to handle complex images with multiple colors or grayscale variations. Additionally, designing new FR models that are specific to an input type can be challenging.

4.5 Database Level

A considerable number of data related to face images or videos and sophisticated FR techniques have been implemented and are running on cloud servers, facilitating large-scale video surveillance applications, such as FR and suspect searches. However, this approach exposes the privacy of the visual data. Handling and storing facial data in an untrustworthy setting with third parties is complex because privacy standards classify facial data as sensitive. This section describes secure and PP solutions for distributed FR.

The FR system's most important asset is arguably the valuable information contained in its database, and it needs to be preserved. To safeguard database-management systems from harmful cyberattacks and unauthorized use, a wide variety of strategies are deployed. Database security systems are designed to protect not only the data contained within the database but also the entire data-management system, as well as any applications that use it, from exploitation, damage, and infiltration.

Traditional methods of guaranteeing database security include database partitioning, cards, and encryption. Encryption is a standard technique for protecting the confidentiality of sensitive private data stored in the cloud. Multiple encryption techniques can be used on different granularities of relational data [6].

A defensive matrix of best practices, paired with regular internal controls, is required to provide effective protection for an organization's databases. In place of that, Rubens presented the finest database-security practices to keep databases secure from data leakage and intruders [163]. According to Rubens, the best possible solution is to allow as few persons as feasible to have access to the database. Administrators should have only the privileges necessary to execute their job, and they should only have access when they need it.

The other alternative solution includes periodically monitoring logins (and attempted logins) to the database and checking logs to detect unusual activity. Alerts can be set up to warn the appropriate team members when potentially malicious behavior is detected. Encrypting stored data is a common practice in many applications and organizations. The five best practices for database security include managing user access rights, training employees in risk-mitigation strategies, examining database flaws and classifying sensitive data, real-time monitoring of database activities to detect data leaks and unauthorized SQL commands, and concealing sensitive information by archiving external data, encrypting databases, and masking database fields [46]. Some methods store data as plaintext in the database [16, 50, 70, 164], whereas others store data as ciphertext [19, 30, 32, 42, 85, 86, 187, 191, 220, 222], providing more privacy than those stored in plaintext.

Advantages: Protected databases provide a secure method for storing sensitive information, reducing the risk of data breaches and leaks. In the event of a data breach or unauthorized access, privacy-enhanced databases significantly reduce the potential impact. This data can be transmitted or shared with confidence. A data lifecycle can also be provided for more secure database storage.

Disadvantages: Protected data loses its original structure, making it challenging to perform efficient search and query operations on the encrypted database. Privacy protection algorithms and protocols must be properly implemented, and any changes or updates to the protection scheme may require significant coordination and migration efforts. The complexity and challenges with indexing, key management, backup, and recovery processes can also increase.

4.6 Communication Level

In the communication channel, PPFR can be divided into secure two-party or three-party computation. In a two-party client-server arrangement, the client has a private key, and the server offers FR services. In the three-party model, the server that offers the FR service is assisted by a cryptographic service provider with a private key.

Secure Two-Party Computation (S2PC): Several PP face-identification technologies operate on a client-server architecture to provide a secure recognition procedure. The client keeps the private key, and the server handles the recognition. Between the camera and database server, a secure two-party protocol offers security.

Erkin et al. [50] introduced the first face-identification protocol that protects privacy based on AHE. Erkin's protocol was enhanced by Sadeghi et al. [164] by fusing AHE with a GC. Using AHE and GC, Blanton et al. [16] proposed biometric identification techniques that protect user privacy. The client–server protocol is also used for secure channel recognition in PP face verification. Upmanyu et al. [191] proposed a PP biometric verification method based on multiplicative HE. A face-verification approach based on AHE was introduced by Im et al. [86]. A face-verification mechanism based on a corrupted circuit was proposed by Sedenka et al. [222]. FHE and message-authentication code-based biometric verification techniques were proposed by Cheon et al. [30] to protect user privacy. A PP biometric-verification system based on quadratic HE algorithm was

proposed by Im et al. [85]. A threshold-predicate encryption-based biometric-verification mechanism was proposed by Zhou et al. [220]. Troncoso-Pastoriza et al. [187] and Boddeti et al. [19] also proposed PP face-verification systems based on FHE.

The face-template database must be encrypted and kept on the server to use any of the aforementioned protocols in the client–server architecture. These procedures have certain drawbacks. First, by keeping a private key on the client side, there is a risk of key leakage. Second, either the server could learn the verification result or the client could learn additional information than the verification result. In either case, these protocols will reveal more information than necessary.

Cryptographic Service-Provider Model: The three-party model consists of a client that collects the image, a server that stores the data and processes biometric matching, and a cryptographic service provider that maintains the encryption methods and manages the public and private keys. Because the CSP is typically deployed in a more secure domain, this strategy has an advantage over the client–server architecture in that it minimizes the possibility of private-key leaking.

The three-party model is used in some face-identification applications. Using AHE and GC, Chun et al. [32] proposed an identification protocol that protects privacy. The server knows the final identification outcome, although it remains unaware of the client access pattern. An FHE method was proposed by Drozdowski et al. [42]. It is a three-party model face-identification technique that protects the privacy of the individual. However, because it has access to the intermediate ciphertexts, the server that possesses the private key can learn more about the users' facial images than necessary. Three-party models have also been used in PP face-verification techniques. An AHE-based face-verification protocol for the three-party scenario was proposed by Abidin et al. [2]. Another one based on AHE and GC was introduced by Lin et al. [109]. One server outputs the verification result using these two protocols; thus, it will learn more than is necessary. A PP face-verification methodology in the three-party paradigm was proposed by Gunasinghe et al. [70]. However, this protocol does not safeguard the anonymity of the face templates because they are kept in public view on a server known as an **identity provider** (**IDP**). Ma et al. [118] used AHE for their face-verification system. Because the verification server with the private key has direct access to the ciphertexts, it gains more knowledge than necessary about the users' facial photographs. In a three-party paradigm based on GC, Gasti et al. [63] suggested a face-verification technique that protects privacy. It differs from the conventional three-party model in that it considers the client to own the secret key rather than the servers. As a result, the client agrees to assume the risk of key leaking. The server also knows the outcome of the verification; thus, it gains access to more data than is necessary. Huang et al. [81] introduced an effective PP face-verification system based on FHE and GC. The scheme ensures that the server learns nothing from the face images, and the client learns nothing but the results.

Table 8 summarizes all techniques surveyed for the communication level. Similar to the face and model levels, the communication level table compares the techniques in terms of the (i) input data type, (ii) testing datasets, (iii) design mechanism, (iv) privacy techniques, and (v) whether the technique fulfills privacy requirements. The privacy requirements are the ISO standards discussed in Section 1.2.3 and are marked as "✓" if they fulfill the requirement and "✗" if not.

Advantages: Privacy protection of the biometric provides a secure recognition procedure, and several PP face identification technologies operate on a client–server architecture. The client keeps a private key, which is not shared and can only be used by the client. Even if the communication is compromised, the protected data will not be of use without the appropriate keys. The process fosters trust between communicating parties while transmitting sensitive information over Wi-Fi networks, cellular networks, or other public connections.

Disadvantages: During PPFR with cryptography, the communication costs of these schemes are also unsatisfactory because the interaction protocols are needed to handle nonlinear

Table 8. PP Techniques for the Communication Level

Ref.	Year	Applied to		Database	Mechanism	Privacy Method	Privacy Requirements			
		Images	Videos				IR	UL	CF	PP
[50]	2009	✓	✗	ORL	AHE	Secure two-party protocol	✗	✗	✗	M
[164]	2009	✓	✗	ORL	AHE + GC	Secure two-party protocol	✗	✗	✗	M
[191]	2010	✓	✗	YALE	MHE	Secure two-party protocol	✓	✓	✓	M
[16]	2011	✓	✗	Own	Transformation	Secure two-party protocol	✗	✗	✗	M
[187]	2012	✓	✗	LFW	FHE	Secure two-party protocol	✓	✓	✓	H
[2]	2014	✓	✗	Own	AHE	Cryptographic service provider	✓	✓	✓	H
[32]	2014	✓	✗	Own	AHE + GC	Cryptographic service provider	✓	✓	✓	M
[222]	2015	✓	✗	LTU, Frank	GC	Secure two-party protocol	✓	✓	✓	H
[30]	2016	✓	✗	CASIA-iris	FHE	Secure two-party protocol	✓	✓	✓	H
[85]	2016	✓	✗	Own	AHE	Secure two-party protocol	✓	✓	✓	M
[63]	2016	✓	✗	Own	GC	Cryptographic service provider	✓	✓	✓	H
[70]	2017	✓	✗	ORL	ZKPK	Cryptographic service provider	✗	✗	✗	M
[118]	2017	✓	✗	LFW, Face94	AHE	Cryptographic service provider	✓	✓	✓	M
[19]	2018	✓	✓	LFW, IJB-A, IJB-B, CASIA	FHE	Secure two-party protocol	✓	✓	✓	H
[220]	2018	✓	✗	Own	TPE	Secure two-party protocol	✓	✓	✓	M
[109]	2018	✓	✗	Own	AHE + GC	Cryptographic service provider	✓	✓	✓	H
[42]	2019	✓	✗	FERET	FHE	Cryptographic service provider	✓	✓	✓	H
[86]	2020	✓	✗	ORL	QHE	Secure two-party protocol	✓	✓	✓	H
[12]	2021	✓	✓	Own	AHE	Secure two-party protocol	✓	✓	✓	M
[81]	2021	✓	✗	LFW	FHE + GC	Cryptographic service provider	✓	✓	✓	H

Symbol explanation: AHE: additive homomorphic encryption, CF: confidentiality, FHE: fully homomorphic encryption, GC: garbled circuit, H: high privacy protection, IR: irreversibility, M: medium privacy protection, MHE: multiplicative homomorphic encryption, PP: performance preservation, QHE: quadratic homomorphic encryption, TPE: threshold-predicate encryption, UL: unlinkability, ZKPK: zero-knowledge proof of knowledge.

operations. HE requires low communication complexity but huge round and computation complexity, whereas GC has low online complexity (rounds, communication, and computation) but large offline communication complexity. The computation and communication of the protocol that must be performed for every face captured by the camera increases the overall cost. Also, keeping a private key on the client side, there is a risk of key leakage.

5 Open Issues and Future Directions

5.1 Open Issues

FR technology can potentially assist people and society; however, it has some serious privacy issues. In a typical FR system, the users' raw face images must be sent to remote servers with **graphics-processing units** (**GPUs**), raising user concerns about exploiting face images and potential privacy breaches during the transmission process. Although there is no question that these worries are legitimate, we must be cautious when making broad generalizations. For instance, the extent of privacy concern FR poses depends on the precise FR technology utilized, how it is implemented, and what computing method it uses.

Institutions worldwide have enacted strong policies to protect people's biometric data, reflecting these concerns. The GDPR in Europe exemplifies this. It specifies that citizens have a right to privacy and that violating it will result in penalties.

Hacking security cameras is a severe risk. For example, if an FR camera is hacked, the hacker may obtain identifying information about anyone in that camera's system. An additional problem is what happens to the faces that the camera records. The most vital component in all fields is security. Thus, this technology must advance alongside other innovations in computer vision and ML. It

is not necessary to completely remove identity traits to protect privacy. Balancing the danger of exposure against the value of the data is one of the most crucial privacy-protection tenets. Over the last eight years (2010–2018), the failure rate in FR has dropped from 5% to 0.2%, and in some cases, FR has proven to be more successful than manual attempts to recognize a human face successfully [142].

The processing of personal data, including special data categories, is inextricably related to using FR systems. Additionally, processing data affects many fundamental rights established in the EU Charter of Fundamental Rights, directly or indirectly. As a result, all applications of FR technology must strictly adhere to the legal framework. This research is meant to be considered while evaluating potential legislation and using current methods in specific situations involving FR systems. Privacy issues related to FR are only anticipated to grow in the next years, owing to continued advancements in FR technology and its implementation in an ever-increasing number of application sectors. Despite the advantages of distributed computing, privacy and security issues are increasing. These issues are directly linked to the need for PPFR.

The subject of explainability-mechanism designs is underexplored. A significant barrier is a lack of ground truth to compare and quantify explainable findings across models. Conversely, GAN-based synthetic-face generation approaches aim to safeguard the privacy and are expanding quickly; however, new adversarial attack types are continuously emerging and threaten system privacy. FL is a new field that aims to cooperatively train an AI model using parameters from trained models while keeping the raw training data local (e.g., on local PCs, mobile, or other connected devices). It is gradually becoming a critical privacy-protection strategy, particularly in contexts where data confidentiality is severely enforced, such as healthcare. The potential of deep learning for PPFR, conversely, is mostly untapped. In a recent open letter to the EC, 51 organizations demanded the outlawing of any FR software used for mass surveillance [9].

After reviewing more than 100 studies, we can confidently affirm that there is a tradeoff between the utility of face data and privacy preservation. While poor PP strategies lead to the recognition of private information, high PP techniques fail to preserve utility for various tasks. PP techniques provide stringent privacy assurances for each level; however, extending them to a broader context, such as applying noise to images and training an FR model, is complicated and inaccurate.

5.2 Future Directions

Multiple technologies are employed to preserve privacy in FR systems: secure cameras, secure communication channels, encrypted databases, GANs (for generating synthetic datasets), and DP for PP. Using one type of technology for numerous tasks is another issue that must be addressed.

Adversarial examples are created to counter deep learning models to protect face data privacy. Deep learning has the potential to either compromise or uphold data privacy. Technologies based on deep learning will be applied more in the future to counter data privacy threats and create PP solutions.

Moreover, the system employs a preservation mechanism to preserve privacy, which can be broken by a skilled attacker, leading to a never-ending cycle of attack and defense. Breaking the loop necessitates new research on attack methods or PP solutions, which is still challenging. Creating new deep learning frameworks and models that consider PP requirements is also a constant challenge.

The tradeoff between model privacy and utility should be carefully considered using DP. Model performance is also affected when a certain amount of differential private noise is added. As more noise is added, model performance significantly declines. The problem of providing a model that preserves privacy while maintaining model utility must still be solved. A general metric framework should be developed to protect the privacy of sensitive data. The framework could be applied to any

protected dataset and indicate the protection level through a score, considering which attributes are protected.

This study aims at guiding those interested in PPFR and in knowing how human rights and data-protection principles (e.g., lawfulness, necessity, proportionality, or other principles) relating to the processing of personal data are interpreted (including policymakers, stakeholders, scholars, and citizens). Comparing deep learning-based PP techniques and the future directions covered in this literature survey can help researchers reach this objective.

6 Conclusion

Focusing on the critical privacy requirements of modern FR systems, this review addresses PPFR, a critical research area in compliance with the GDPR. We expanded on this subject by first discussing the necessity and importance of FR and general privacy problems concerning international laws. Moreover, we investigated numerous privacy problems, solutions, and challenges in depth in this interdisciplinary subject and proposed a six-level leakage and PP framework based on the complete life cycle of a CCTV-based FR system. In addition, we presented a thorough and methodical analysis of the most recent privacy leaks and PP advancements using the framework. We also outlined the methods, benefits, and drawbacks of the intelligent solutions assessed at each level and enumerated their common design principles. In addition, we mapped these solutions concerning linkability, irreversibility, confidentiality, and privacy level. The data leakage and privacy-intelligence solutions investigated in this study are adequate to construct a PPFR system capable of preserving the privacy of all parties involved. This study is expected to aid modern privacy management and clarify the ever-increasing PPFR systems.

References

[1] Bogle A. 2020. Australian federal police officers trialled Clearview AI, ABC News. Retrieved July 2022 from https://www.abc.net.au/news/science/2020-04-14/clearview-ai-facial-recognition-tech-australian-federal-police/12146894

[2] Aysajan Abidin, Kanta Matsuura, and Aikaterini Mitrokotsa. 2014. Security of a privacy-preserving biometric authentication protocol revisited. In *International Conference on Cryptology and Network Security*. Springer, 290–304.

[3] Abbas Acar, Hidayet Aksu, A. Selcuk Uluagac, and Mauro Conti. 2018. A survey on homomorphic encryption schemes: Theory and implementation. *ACM Computing Surveys* 51, 4 (2018), 1–35.

[4] Alessandro Acquisti, Ralph Gross, and Frederic D. Stutzman. 2014. Face recognition and privacy in the age of augmented reality. *Journal of Privacy and Confidentiality* 6, 2 (2014), 1.

[5] Divyansh Aggarwal, Jiayu Zhou, and Anil K. Jain. 2021. Fedface: Collaborative learning of face recognition model. In *International Joint Conference on Biometrics*. IEEE, 1–8.

[6] Divyakant Agrawal, Amr El Abbadi, and Shiyuan Wang. 2013. Secure and privacy-preserving database services in the cloud. In *International Conference on Data Engineering*. IEEE, 1268–1271.

[7] Brandon Amos, Bartosz Ludwiczuk, and Mahadev Satyanarayanan. 2016. Openface: A general-purpose face recognition library with mobile applications. *CMU School of Computer Science* 6, 2 (2016), 20.

[8] Abhishek Anand. 2020. Installation of 1.4 lakh Chinese CCTV cameras by Delhi Govt Sparks Row. Retrieved July 2022 from https://www.indiatoday.in/mail-today/story/installation-of-1-4-lakh-chinese-cctv-cameras-by-delhi-govt-sparks-row-1696032-2020-07-02

[9] Article19. 2021. Ban on biometric surveillance. Retrieved July 2022 from https://www.article19.org/resources/european-commission-must-ban-biometric-mass-surveillance/

[10] Noboru Babaguchi, Takashi Koshimizu, Ichiro Umata, and Tomoji Toriyama. 2009. Psychological study for designing privacy protected video surveillance system: PriSurv. In *Protecting Privacy in Video Surveillance*. Springer, 147–164.

[11] Thangapavithraa Balaji, Patrick Blies, Georg Göri, Raphael Mitsch, Marcel Wasserer, and Torsten Schön. 2021. Temporally coherent video anonymization through GAN inpainting. *arXiv preprint arXiv:2106.02328*.

[12] Elmahdi Bentafat, M. Mazhar Rathore, and Spiridon Bakiras. 2021. Towards real-time privacy-preserving video surveillance. *Computer Communications* 180 (2021), 97–108.

[13] Marcelo Bertalmio, Guillermo Sapiro, Vincent Caselles, and Coloma Ballester. 2000. Image inpainting. In *Annual Conference on Computer Graphics and Interactive Techniques*. 417–424.

[14] Biometric. 2022. Leak reveals Western companies with face biometrics licenses from Russia's NtechLab. Retrieved July 2022 from https://www.biometricupdate.com/202208/leak-reveals-western-companies-with-face-biometrics-licenses-from-russias-ntechlab
[15] Biometrics-Institute. 2019. First Universal Privacy Guidelines for Biometrics. Retrieved August 2022 from https://identityweek.net/biometrics-institute-launches-first-universal-privacy-guidelines/
[16] Marina Blanton and Paolo Gasti. 2011. Secure and efficient protocols for iris and fingerprint identification. In *European Conference on Research in Computer Security*. Springer, 190–209.
[17] Volker Blanz, Kristina Scherbaum, Thomas Vetter, and Hans-Peter Seidel. 2004. Exchanging faces in images. In *Computer Graphics Forum*. Wiley Online Library, 669–676.
[18] Volker Blanz and Thomas Vetter. 2003. Face recognition based on fitting a 3D morphable model. *Transactions on Pattern Analysis and Machine Intelligence* 25, 9 (2003), 1063–1074.
[19] Vishnu Naresh Boddeti. 2018. Secure face matching using fully homomorphic encryption. In *International Conference on Biometrics Theory, Applications, and Systems*. IEEE, 1–10.
[20] Dan Boneh, Eu-Jin Goh, and Kobbi Nissim. 2005. Evaluating 2-DNF formulas on ciphertexts. In *Theory of Cryptography Conference*. Springer, 325–341.
[21] Margherita Bonetto, Pavel Korshunov, Giovanni Ramponi, and Touradj Ebrahimi. 2015. Privacy in mini-drone based video surveillance. In *International Conference and Workshops on Automatic Face and Gesture Recognition*. IEEE, 1–6.
[22] Fadi Boutros, Marco Huber, Patrick Siebke, Tim Rieber, and Naser Damer. 2022. SFace: Privacy-friendly and accurate face recognition using synthetic data. arXiv:2206.10520. Retrieved from https://arxiv.org/abs/2206.10520
[23] Michael Boyle, Christopher Edwards, and Saul Greenberg. 2000. The effects of filtered video on awareness and privacy. In *ACM Conference on Computer Supported Cooperative Work*. 1–10.
[24] Julien Bringer, Herve Chabanne, Melanie Favre, Alain Patey, Thomas Schneider, and Michael Zohner. 2014. GSHADE: Faster privacy-preserving distance computation and biometric identification. In *Workshop on Information Hiding and Multimedia Security*. 187–198.
[25] Laura Carrer. 2021. Italy introduces a moratorium on video surveillance systems that use facial recognition. Retrieved July 2022 from https://edri.org/our-work/italy-introduces-a-moratorium-on-video-surveillance-systems-that-use-facial-recognition/
[26] Mahawaga Arachchige Pathum Chamikara, Peter Bertok, Ibrahim Khalil, Dongxi Liu, and Seyit Camtepe. 2020. Privacy preserving face recognition utilizing differential privacy. *Computers and Security* 97 (2020), 101951.
[27] T-H Hubert Chan, Mingfei Li, Elaine Shi, and Wenchang Xu. 2012. Differentially private continual monitoring of heavy hitters from distributed streams. In *International Symposium on Privacy Enhancing Technologies Symposium*. Springer, 140–159.
[28] Jiawei Chen, Janusz Konrad, and Prakash Ishwar. 2018. Vgan-based image representation learning for privacy-preserving facial expression recognition. In *CVPRW*. IEEE, 1570–1579.
[29] Zhenfei Chen, Tianqing Zhu, Ping Xiong, Chenguang Wang, and Wei Ren. 2021. Privacy preservation for image data: A gan-based method. *International Journal of Intelligent Systems* 36, 4 (2021), 1668–1685.
[30] Jung Hee Cheon, HeeWon Chung, Myungsun Kim, and Kang-Won Lee. 2016. Ghostshell: Secure biometric authentication using integrity-based homomorphic evaluations. *Cryptology ePrint Archive* (2016). Retrieved August 2022 from https://eprint.iacr.org/2016/484
[31] Durkhyun Cho, Jin Han Lee, and Il Hong Suh. 2020. CLEANIR: Controllable attribute-preserving natural identity remover. *Applied Sciences* 10, 3 (2020), 1120.
[32] Hu Chun, Yousef Elmehdwi, Feng Li, Prabir Bhattacharya, and Wei Jiang. 2014. Outsourceable two-party privacy-preserving biometric authentication. In *ACM Symposium on Information, Computer, and Communications Security*. 401–412.
[33] Catalin Cimpanu. 2020. Selling data of Weibo users. Retrieved July 2022 from https://www.zdnet.com/article/hacker-selling-data-of-538-million-weibo-users/
[34] European Commission. 2018. Trustworthy AI. Retrieved July 2022 from https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines/1.html
[35] EUROPEAN COMMISSION. 2021. Regulation of the European Parliament and of the Council. Retrieved July 2022 from https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX%3A52021PC0206&from=EN
[36] Antonia Creswell, Tom White, Vincent Dumoulin, Kai Arulkumaran, Biswa Sengupta, and Anil A Bharath. 2018. Generative adversarial networks: An overview. *Signal Processing Magazine* 35, 1 (2018), 53–65.
[37] William L. Croft, Jörg-Rüdiger Sack, and Wei Shi. 2019. Differentially private obfuscation of facial images. In *International Cross-Domain Conference for Machine Learning and Knowledge Extraction*. Springer, 229–249.
[38] William L. Croft, Jörg-Rüdiger Sack, and Wei Shi. 2021. Obfuscation of images via differential privacy: from facial images to general images. *Peer-to-Peer Networking and Applications* 14 (2021), 1705–1733.

[39] Sajjad Dadkhah, Mario Koeppen, Somayeh Sadeghi, and Kaori Yoshida. 2018. Bad Ai: Investigating the effect of halftoning techniques on unwanted face detection systems. In *International Conference on New Technologies, Mobility and Security*. IEEE, 1–5.
[40] Ivan Damgård, Martin Geisler, and Mikkel Krøigaard. 2007. Efficient and secure comparison for on-line auctions. In *Australasian Conference on Information Security and Privacy*. Springer, 416–430.
[41] Anupam Das, Martin Degeling, Xiaoyou Wang, Junjue Wang, Norman Sadeh, and Mahadev Satyanarayanan. 2017. Assisting users in a world full of cameras: A privacy-aware infrastructure for computer vision applications. In *CVPRW*. IEEE, 1387–1396.
[42] Pawel Drozdowski, Nicolas Buchmann, Christian Rathgeb, Marian Margraf, and Christoph Busch. 2019. On the application of homomorphic encryption to face identification. In *International Conference of the Biometrics Special Interest Group*. IEEE, 1–5.
[43] Liang Du, Meng Yi, Erik Blasch, and Haibin Ling. 2014. GARP-face: Balancing privacy protection and utility preservation in face de-identification. In *International Joint Conference on Biometrics*. IEEE, 1–8.
[44] Lauren Dudley. 2020. China's ubiquitous facial recognition tech sparks privacy backlash. *The Diplomat* (2020). Retrieved August 2022 from https://thediplomat.com/2020/03/chinas-ubiquitous-facial-recognition-tech-sparks-privacy-backlash/
[45] Frederic Dufaux. 2011. Video scrambling for privacy protection in video surveillance: Recent results and validation framework. In *Mobile Multimedia/Image Processing, Security, and Applications*. SPIE, 11–24.
[46] Geekflare Editorial. 2022. Database threats and prevention. Retrieved July 2022 from https://geekflare.com/database-threats-and-prevention-tools/
[47] Passant Elagroudy, Mohamed Khamis, Florian Mathis, Diana Irmscher, Andreas Bulling, and Albrecht Schmidt. 2019. Can privacy-aware lifelogs alter our memories?. In *Conference on Human Factors in Computing Systems*. 1–6.
[48] Adám Erdélyi, Tibor Barát, Patrick Valet, Thomas Winkler, and Bernhard Rinner. 2014. Adaptive cartooning for privacy protection in camera networks. In *International Conference on Advanced Video and Signal Based Surveillance*. IEEE, 44–49.
[49] Ovgu Ozturk Ergun. 2014. Privacy preserving face recognition in encrypted domain. In *APCCAS*. IEEE, 643–646.
[50] Zekeriya Erkin, Martin Franz, Jorge Guajardo, Stefan Katzenbeisser, Inald Lagendijk, and Tomas Toft. 2009. Privacy-preserving face recognition. In *International Symposium on Privacy Enhancing Technologies Symposium*. Springer, 235–253.
[51] The Indian Express. 2019. CCTV in daycare. Retrieved August 2022 from https://indianexpress.com/article/parenting/blog/cctv-camera-daycare-place-trust-in-child-staff-5677027/
[52] Liyue Fan. 2018. Image pixelization with differential privacy. In *Data and Applications Security and Privacy XXXII*. Springer, 148–162.
[53] Liyue Fan. 2019. Differential privacy for image publication. In *Theory and Practice of Differential Privacy Workshop*, Vol. 1. 6.
[54] Liyue Fan. 2019. Practical image obfuscation with provable privacy. In *ICME*. IEEE, 784–789.
[55] Fortune. 2021. LinkedIn. Retrieved August 2022 from https://fortune.com/2021/06/30/linkedin-data-theft-700-million-users-personal-information-cybersecurity/
[56] Hajer Fradi, Volker Eiselein, Ivo Keller, Jean-Luc Dugelay, and Thomas Sikora. 2013. Crowd context-dependent privacy protection filters. In *International Conference on Digital Signal Processing*. IEEE, 1–6.
[57] Matt Fredrikson, Somesh Jha, and Thomas Ristenpart. 2015. Model inversion attacks that exploit confidence information and basic countermeasures. In *ACM SIGSAC Conference on Computer and Communications Security*. 1322–1333.
[58] Andrea Frome, German Cheung, Ahmad Abdulkader, Marco Zennaro, Bo Wu, Alessandro Bissacco, Hartwig Adam, Hartmut Neven, and Luc Vincent. 2009. Large-scale privacy protection in Google Street View. In *ICCV*. IEEE, 2373–2380.
[59] Josh Fruhlinger. 2020. Marriott data breach. Retrieved August 2022 from https://www.csoonline.com/article/3441220/marriott-data-breach-faq-how-did-it-happen-and-what-was-the-impact.html
[60] Md Tahmid Hasan Fuad, Awal Ahmed Fime, Delowar Sikder, Md Akil Raihan Iftee, Jakaria Rabbi, Mabrook S Al-Rakhami, Abdu Gumaei, Ovishake Sen, Mohtasim Fuad, and Md Nazrul Islam. 2021. Recent advances in deep learning techniques for face recognition. *IEEE Access* 9 (2021), 99112–99142.
[61] Market Research Future. 2022. CCTV market. Retrieved August 2022 from https://www.marketresearchfuture.com/sample_request/677
[62] David Gargaro. 2022. Face Recognition Technology. Retrieved August 2022 from https://www.itpro.co.uk/security/privacy/356882/the-pros-and-cons-of-facial-recognition-technology
[63] Paolo Gasti, Jaroslav Šeděnka, Qing Yang, Gang Zhou, and Kiran S. Balagani. 2016. Secure, fast, and energy-efficient outsourced authentication for smartphones. *IEEE Transactions on Information Forensics and Security* 11, 11 (2016), 2556–2571.

[64] Craig Gentry. 2009. *A Fully Homomorphic Encryption Scheme*. Stanford University.
[65] IFSEC Global. 2021. Role of CCTV Cameras: Public, Privacy and Protection. Retrieved August 2022 from https://www.ifsecglobal.com/video-surveillance/role-cctv-cameras-public-privacy-protection/
[66] Ralph Gross, Edoardo Airoldi, Bradley Malin, and Latanya Sweeney. 2005. Integrating utility into face de-identification. In *International Workshop on Privacy Enhancing Technologies*. Springer, 227–242.
[67] The Guardian. 2013. Apple Developer site hack. Retrieved August 2022 from https://www.theguardian.com/technology/2013/jul/22/apple-developer-site-hacked
[68] Data Guidance. 2011. Data Protection, South Korea. Retrieved April 2024 from https://www.dataguidance.com/notes/south-korea-data-protection-overview
[69] Data Guidance. 2023. Data Protection, Argentina. Retrieved April 2024 from https://www.dataguidance.com/notes/argentina-data-protection-overview
[70] Hasini Gunasinghe and Elisa Bertino. 2017. PrivBioMTAuth: Privacy preserving biometrics-based and user centric protocol for user authentication from mobile phones. *Transactions on Information Forensics and Security* 13, 4 (2017), 1042–1057.
[71] Siyu Guo, Songhe Feng, Yidong Li, Shushan An, and Hairong Dong. 2018. Integrating diversity into neural-network-based face deidentification. In *Chinese Control Conference*. IEEE, 9356–9361.
[72] Shangwei Guo, Tao Xiang, and Xiaoguo Li. 2019. Towards efficient privacy-preserving face recognition in the cloud. *Signal Processing* 164 (2019), 320–328.
[73] Hanxiang Hao, David Güera, Amy R. Reibman, and Edward J. Delp. 2019. A utility-preserving GAN for face obscuration. *arXiv preprint* arXiv:1906.11979.
[74] Rakibul Hasan, Patrick Shaffer, David Crandall, and Eman T. Apu Kapadia. 2017. Cartooning for enhanced privacy in lifelogging and streaming videos. In *Conference on Computer Vision and Pattern Recognition Workshops*. IEEE, 29–38.
[75] Rebecca Heilweil. 2020. Big tech companies back away from selling facial recognition to police. That's progress. *Vox* (2020). Retrieved August 2022 from https://www.vox.com/recode/2020/6/10/21287194/amazon-microsoft-ibm-facial-recognition-moratorium-police
[76] Steven Hill, Zhimin Zhou, Lawrence Saul, and Hovav Shacham. 2016. On the (in) effectiveness of mosaicing and blurring as tools for document redaction. *Proceedings on Privacy Enhancing Technologies* 2016, 4 (2016), 403–417.
[77] Benjamin Hodges. 2020. The varying laws governing facial recognition technology. Retrieved July 2022 from https://www.ipwatchdog.com/2020/01/28/varying-laws-governing-facial-recognition-technology/id=118240/
[78] Kota Hoshino, Keiichi Iwamura, and Kitahiro Kaneda. 2018. Improvement of privacy protection surveillance camera system and its applications. In *International Conference on Mobile and Secure Services*. IEEE, 1–5.
[79] Hwai-Jung Hsu and Kuan-Ta Chen. 2015. Face recognition on drones: Issues and limitations. In *Workshop on Micro Aerial Vehicle Networks, Systems, and Applications for Civilian Use*. 39–44.
[80] Hongsheng Hu, Zoran Salcic, Lichao Sun, Gillian Dobbie, Philip S Yu, and Xuyun Zhang. 2021. Membership inference attacks on machine learning: A survey. *ACM Computing Surveys* 54, 80 (2021), 1–37.
[81] Hai Huang and Luyao Wang. 2021. Efficient privacy-preserving face verification scheme. *Journal of Information Security and Applications* 63 (2021), 103055.
[82] Håkon Hukkelås, Rudolf Mester, and Frank Lindseth. 2019. Deepprivacy: A generative adversarial network for face anonymization. In *International Symposium on Visual Computing*. Springer, 565–578.
[83] IBM. 2020. Securing mobile devices. Retrieved July 2022 from https://www.slideshare.net/ibmsoftwareindia/securing-mobile-devices-in-the-business-environment-36342389
[84] Panagiotis Ilia, Iasonas Polakis, Elias Athanasopoulos, Federico Maggi, and Sotiris Ioannidis. 2015. Face/off: Preventing privacy leakage from photos in social networks. In *ACM SIGSAC Conference on Computer and Communications Security*. 781–792.
[85] Jong-Hyuk Im, JinChun Choi, DaeHun Nyang, and Mun-Kyu Lee. 2016. Privacy-preserving palm print authentication using homomorphic encryption. In *International Conference on Dependable, Autonomic, and Secure Computing*. IEEE, 878–881.
[86] Jong-Hyuk Im, Seong-Yun Jeon, and Mun-Kyu Lee. 2020. Practical privacy-preserving face authentication for smartphones secure against malicious clients. *Transactions on Information Forensics and Security* 15 (2020), 2386–2401.
[87] ISO. 2011. IEC 24745: Biometric information protection. *International Organization for Standardization* (2011). Retrieved August 2022 from https://www.iso.org/standard/75302.html
[88] Xin Jin, Yan Liu, Xiaodong Li, Geng Zhao, Yingya Chen, and Kui Guo. 2015. Privacy preserving face identification in the cloud through sparse representation. In *Chinese Conference on Biometric Recognition*. Springer, 160–167.
[89] Brendan John, Ao Liu, Lirong Xia, Sanjeev Koppal, and Eakta Jain. 2020. Let it snow: Adding pixel noise to protect the users identity. In *ACM Symposium on Eye Tracking Research and Applications*. 1–3.

[90] Hill K. 2020. The Secretive Company. Retrieved July 2022 from https://www.nytimes.com/2020/01/18/technology/clearview-privacy-facial-recognition.html
[91] Peter Kairouz, Sewoong Oh, and Pramod Viswanath. 2014. Extremal mechanisms for local differential privacy. *Advances in Neural Information Processing Systems* 27 (2014), 2879–2887.
[92] Mohamed Khamis, Habiba Farzand, Marija Mumm, and Karola Marky. 2022. DeepFakes for privacy: Investigating the effectiveness of state-of-the-art privacy-enhancing face obfuscation methods. In *International Conference on Advanced Visual Interfaces*. 1–5.
[93] Taehoon Kim and Jihoon Yang. 2019. Latent-space-level image anonymization with adversarial protector networks. *IEEE Access* 7 (2019), 84992–84999.
[94] Youngsung Kim and Kar-Ann Toh. 2008. Sparse random projection for efficient cancelable face feature extraction. In *Conference on Industrial Electronics and Applications*. IEEE, 2139–2144.
[95] Carly Kind and Jenny Brennan. 2019. Beyond face value: Public attitudes to facial recognition technology. Retrieved July 2022 from https://www.adalovelaceinstitute.org/report/beyond-face-value-public-attitudes-to-facial-recognition-technology/
[96] Kento Kobayashi, Keiichi Iwamura, Kitahiro Kaneda, and Isao Echizen. 2014. Surveillance camera system to achieve privacy protection and crime prevention. In *International Conference on Intelligent Information Hiding and Multimedia Signal Processing*. IEEE, 463–466.
[97] Vladimir Kolesnikov and Ranjit Kumaresan. 2013. Improved OT extension for transferring short secrets. In *Annual Cryptology Conference*. Springer, 54–70.
[98] Pavel Korshunov, Claudia Araimo, Francesca De Simone, Carmelo Velardo, J-L Dugelay, and Touradj Ebrahimi. 2012. Subjective study of privacy filters in video surveillance. In *International Workshop on Multimedia Signal Processing*. IEEE, 378–382.
[99] Iryna Korshunova, Wenzhe Shi, Joni Dambre, and Lucas Theis. 2017. Fast face-swap using convolutional neural networks. In *International Conference on Computer Vision*. IEEE, 3677–3685.
[100] Xiaoyu Kou, Ziling Zhang, Yuelei Zhang, and Linlin Li. 2021. Efficient and privacy-preserving distributed face recognition scheme via FaceNet. In *ACM Turing Award Celebration Conference*. 110–115.
[101] Lamyanba Laishram, Md Maklachur Rahman, and Soon Ki Jung. 2021. Challenges and applications of face deepfake. In *International Workshop on Frontiers of Computer Vision*. Springer, 131–156.
[102] New Zeland legislation. 2020. Privacy Act. Retrieved April 2024 from https://www.legislation.govt.nz/act/public/2020/0031/latest/LMS23227.html
[103] Shu-Min Leong, Raphaël C-W Phan, Vishnu Monn Baskaran, and Chee-Pun Ooi. 2020. Privacy-preserving facial recognition based on temporal features. *Applied Soft Computing* 96 (2020), 106662.
[104] Geoffrey Letournel, Aurélie Bugeau, V-T Ta, and J-P Domenger. 2015. Face de-identification with expressions preservation. In *ICIP*. IEEE, 4366–4370.
[105] Haoran Li, Li Xiong, Lifan Zhang, and Xiaoqian Jiang. 2014. DPSynthesizer: Differentially private data synthesizer for privacy preserving data sharing. In *PVLDB Endowment International Conference on Very Large Data Bases*. NIH Public Access, 1677.
[106] Tao Li and Lei Lin. 2019. Anonymousnet: Natural face de-identification with measurable privacy. In *CVPRW*. IEEE, 56–65.
[107] Yuezun Li and Siwei Lyu. 2019. De-identification without losing faces. In *ACM Workshop on Information Hiding and Multimedia Security*. 83–88.
[108] Yifang Li, Nishant Vishwamitra, Bart P Knijnenburg, Hongxin Hu, and Kelly Caine. 2017. Effectiveness and users' experience of obfuscation as a privacy-enhancing technology for sharing photos. *ACM on Human-Computer Interaction* 1, CSCW (2017), 1–24.
[109] Dan Lin, Nicholas Hilbert, Christian Storer, Wei Jiang, and Jianping Fan. 2018. UFace: Your universal password that no one can see. *Computers and Security* 77 (2018), 627–641.
[110] Y. Lindell and B. Pinkas. 2004. A proof of Yao's protocol for secure two-party computation. ECCC report TR04-063. In *Electronic Colloquium on Computational Complexity*.
[111] Chi Liu, Tianqing Zhu, Jun Zhang, and Wanlei Zhou. 2020. Privacy intelligence: A survey on image privacy in online social networks. *ACM Computing Surveys* 55, 8 (2020), 1–35.
[112] Chih-Ting Liu, Chien-Yi Wang, Shao-Yi Chien, and Shang-Hong Lai. 2022. FedFR: Joint optimization federated framework for generic and personalized face recognition. In *AAAI Conference on Artificial Intelligence*. 1656–1664.
[113] Zilong Liu, Jingbing Li, and Jing Liu. 2022. Encrypted face recognition algorithm based on ridgelet-DCT transform and THM chaos. *Mathematical Biosciences and Engineering* 19, 2 (2022), 1373–1387.
[114] Jens-Martin Loebel. 2012. Is privacy dead?–an inquiry into GPS-based geolocation and facial recognition systems. In *IFIP International Conference on Human Choice and Computers*. Springer, 338–348.
[115] Huimin Lu and Seiichi Serikawa. 2013. Design of freely configurable safety light curtain using hemispherical mirrors. *Transactions on Electrical and Electronic Engineering* 8, S1 (2013), S110–S111.

[116] Chao Ma, Atsushi Shimada, Hideaki Uchiyama, Hajime Nagahara, and Rin-ichiro Taniguchi. 2019. Fall detection using optical level anonymous image sensing system. *Optics and Laser Technology* 110 (2019), 44–61.
[117] Tianxiang Ma, Dongze Li, Wei Wang, and Jing Dong. 2021. CFA-Net: Controllable face anonymization network with identity representation manipulation. *arXiv preprint* arXiv:2105.11137.
[118] Yukun Ma, Lifang Wu, Xiaofeng Gu, Jiaoyu He, and Zhou Yang. 2017. A secure face-verification scheme based on homomorphic encryption and deep neural networks. *IEEE Access* 5 (2017), 16532–16538.
[119] The Markup. 2021. Popular Facial Recognition Tool Shows Just How Much the Software Tracks People. Retrieved July 2022 from https://themarkup.org/privacy/2021/07/06/this-manual-for-a-popular-facial-recognition-tool-shows-just-how-much-the-software-tracks-people
[120] Maxim Maximov, Ismail Elezi, and Laura Leal-Taixé. 2020. Ciagan: Conditional identity anonymization generative adversarial networks. In *CVPR*. IEEE, 5447–5456.
[121] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas. 2017. Communication-efficient learning of deep networks from decentralized data. In *Artificial Intelligence and Statistics*. PMLR, 1273–1282.
[122] Richard McPherson, Reza Shokri, and Vitaly Shmatikov. 2016. Defeating image obfuscation with deep learning. *arXiv preprint* arXiv:1609.00408.
[123] Blaž Meden, Žiga Emeršič, Vitomir Štruc, and Peter Peer. 2018. k-same-net: k-anonymity with generative deep neural networks for face deidentification. *Entropy* 20, 1 (2018), 60.
[124] Blaž Meden, Refik Can Mallı, Sebastjan Fabijan, Hazım Kemal Ekenel, Vitomir Štruc, and Peter Peer. 2017. Face deidentification with generative deep neural networks. *IET Signal Processing* 11, 9 (2017), 1046–1054.
[125] Blaž Meden, Peter Rot, Philipp Terhörst, Naser Damer, Arjan Kuijper, Walter J. Scheirer, Arun Ross, Peter Peer, and Vitomir Štruc. 2021. Privacy–enhancing face biometrics: A comprehensive survey. *Transactions on Information Forensics and Security* 16 (2021), 4147–4183.
[126] O. Rabin Michael and M. O. Rabin. 1981. *How to Exchange Secrets by Oblivious Transfer*. Technical Report. Technical report, Aiken Computation Laboratory. Harvard University.
[127] Microsoft. 2009. Microsoft - Online Reputation in a Connected World. Retrieved July 2022 from https://download.microsoft.com/download/C/D/2/CD233E13-A600-482F-9C97-545BB4AE93B1/DPD_Online%20Reputation%20Research_overview.doc
[128] Cambridge Montessori. 2018. CCTV in Preschool. Retrieved July 2022 from https://cambridgemontessoriglobal.org/importance-of-cctv-in-preschool/
[129] Saleh Mosaddegh, Loic Simon, and Frédéric Jurie. 2015. Photorealistic face de-identification by aggregating donors' face components. In *ACCV*. Springer, 159–174.
[130] Taichi Nakamura, Yuiko Sakuma, and Hiroaki Nishi. 2021. Face-image anonymization as an application of multidimensional data k-anonymizer. *International Journal of Networking and Computing* 11, 1 (2021), 102–119.
[131] Moni Naor and Adi Shamir. 1994. Visual cryptography. In *Workshop on Cryptographic Techniques*. Springer, 1–12.
[132] M. Mrityunjay and P. Narayanan. 2011. The de-identification camera. In *National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics*. IEEE, 192–195.
[133] Ryota Natsume, Tatsuya Yatagawa, and Shigeo Morishima. 2018. RSGAN: Face swapping and editing using face and hair representation in latent spaces. In *ACM SIGGRAPH Posters (SIGGRAPH'18)*. Association for Computing Machinery, New York, NY, USA, Article 69, 1–2.
[134] Francis Navarro. 2019. These phones can be unlocked with nothing more than a photograph. Retrieved July 2022 from https://www.komando.com/security-privacy/these-phones-can-be-unlocked-with-nothing-more-than-a-photograph/529870/
[135] NEC. 2018. Privacy Measures of Biometrics Businesses. Retrieved July 2022 from https://www.nec.com/en/global/techrep/journal/g18/n02/180205.html
[136] BBC News. 2021. China's 'Midnight patrol' cracks down on Young Gamers. Retrieved July 2022 from https://www.bbc.co.uk/news/technology-57752782
[137] Kyodo News. 2018. Dozens of canon security cameras hacked in Japan. Retrieved July 2022 from https://english.kyodonews.net/news/2018/05/91ec861ae24d-dozens-of-security-cameras-hacked-in-japan.html?phrase=ham+fighters&words=
[138] Elaine M Newton, Latanya Sweeney, and Bradley Malin. 2005. Preserving privacy by de-identifying face images. *Transactions on Knowledge and Data Engineering* 17, 2 (2005), 232–243.
[139] Alfred Ng. 2018. Spy with smart cameras. Retrieved July 2022 from https://www.cnet.com/home/smart-home/your-smart-camera-may-have-been-spying-on-you/
[140] Shaun Nichols. 2021. Accenture sheds more light on August data breach. Retrieved July 2022 from https://www.techtarget.com/searchsecurity/news/252508243/Accenture-sheds-more-light-on-August-data-breach
[141] Yuval Nirkin, Yosi Keller, and Tal Hassner. 2019. Fsgan: Subject agnostic face swapping and reenactment. In *CVPR*. IEEE, 7184–7193.

[142] NIST. 2018. Advance in Face Recognition Software's Capabilities. Retrieved July 2022 from https://www.nist.gov/news-events/news/2018/11/nist-evaluation-shows-advance-face-recognition-softwares-capabilities
[143] npr. 2021. Facebook Data Breach. Retrieved July 2022 from https://www.npr.org/2021/04/09/986005820/after-data-breach-exposes-530-million-facebook-says-it-will-not-notify-users
[144] Office of the Privacy Commissioner of Canada. 2021. The Personal Information Protection and Electronic Documents Act. Retrieved April 2024 from https://www.priv.gc.ca/en/privacy-topics/privacy-laws-in-canada/the-personal-information-protection-and-electronic-documents-act-pipeda/
[145] Margarita Osadchy, Benny Pinkas, Ayman Jarrous, and Boaz Moskovich. 2010. Scifi-a system for secure face identification. In *Symposium on Security and Privacy*. IEEE, 239–254.
[146] Dailé Osorio-Roig, Christian Rathgeb, Pawel Drozdowski, and Christoph Busch. 2021. Stable hash generation for efficient privacy-preserving face identification. *Transactions on Biometrics, Behavior, and Identity Science* 4, 3 (2021), 1–1.
[147] Pascal Paillier. 1999. Public-key cryptosystems based on composite degree residuosity classes. In *International Conference on the Theory and Applications of Cryptographic Techniques*. Springer, 223–238.
[148] Yi-Lun Pan, Min-Jhih Haung, Kuo-Teng Ding, Ja-Ling Wu, and Jyh-Shing Jang. 2019. K-same-siamese-GAN: K-same algorithm with generative adversarial network for facial image de-identification with hyperparameter tuning and mixed precision training. In *International Conference on Advanced Video and Signal Based Surveillance*. IEEE, 1–8.
[149] Vishal M. Patel, Rama Chellappa, and Massimo Tistarelli. 2010. Sparse representations and random projections for robust and cancelable biometrics. In *International Conference on Control Automation Robotics and Vision*. IEEE, 1–6.
[150] Japan Personal Information Protection Commission. 2003. Personal Protection. Retrieved April 2024 from https://www.cas.go.jp/jp/seisaku/hourei/data/APPI.pdf
[151] PETS09. 2009. MOT benchmark. Retrieved July 2022 from https://motchallenge.net/vis/PETS09-S2L1
[152] Francesco Pittaluga and Sanjeev Jagannatha Koppal. 2016. Pre-capture privacy for small vision sensors. *Transactions on Pattern Analysis and Machine Intelligence* 39, 11 (2016), 2215–2226.
[153] Francesco Pittaluga, Aleksandar Zivkovic, and Sanjeev J Koppal. 2016. Sensor-level privacy for thermal cameras. In *International Conference on Computational Photography*. IEEE, 1–12.
[154] Hugo Proença. 2021. The uu-net: Reversible face de-identification for visual surveillance video footage. *Transactions on Circuits and Systems for Video Technology* 32, 2 (2021), 496–509.
[155] Habiba Rashid. 2021. Cognyte data breach. Retrieved July 2022 from https://www.hackread.com/cybersecurity-firm-expose-data-breach-records/
[156] Christian Rathgeb, Marta Gomez-Barrero, Christoph Busch, Javier Galbally, and Julian Fierrez. 2015. Towards cancelable multi-biometrics based on bloom filters: A case study on feature level fusion of face and iris. In *IWBF*. IEEE, 1–6.
[157] Dominick Reilly and Liyue Fan. 2021. A comparative evaluation of differentially private image obfuscation. In *International Conference on Trust, Privacy and Security in Intelligent Systems and Applications*. IEEE, 80–89.
[158] Lijing Ren and Denghui Zhang. 2022. A privacy-preserving biometric recognition system with visual cryptography. *Advances in Multimedia* 2022, 1 (2022), 1057114.
[159] Zhongzheng Ren, Yong Jae Lee, and Michael S. Ryoo. 2018. Learning to anonymize faces for privacy preserving action detection. In *ECCV*. 620–636.
[160] Chi-Hyoung Rhee and C. Lee. 2013. Cartoon-like avatar generation using facial component matching. *International Journal of Multimedia and Ubiquitous Engineering* 8, 4 (2013), 69–78.
[161] Ergys Ristani, Francesco Solera, Roger Zou, Rita Cucchiara, and Carlo Tomasi. 2016. Performance measures and a data set for multi-target, multi-camera tracking. In *ECCV*. Springer, 17–35.
[162] Ronald L Rivest, Len Adleman, and Michael L. Dertouzos. 1978. On data banks and privacy homomorphisms. *Foundations of Secure Computation* 4, 11 (1978), 169–180.
[163] Paul Rubens. 2022. Database security: Best practices. Retrieved July 2022 from https://www.esecurityplanet.com/networks/database-security-best-practices/
[164] Ahmad-Reza Sadeghi, Thomas Schneider, and Immo Wehrenberg. 2009. Efficient privacy-preserving face recognition. In *International Conference on Information Security and Cryptology*. Springer, 229–244.
[165] Omair Sarwar, Bernhard Rinner, and Andrea Cavallaro. 2016. Design space exploration for adaptive privacy protection in airborne images. In *International Conference on Advanced Video and Signal Based Surveillance*. IEEE, 159–165.
[166] Omair Sarwar, Bernhard Rinner, and Andrea Cavallaro. 2019. A privacy-preserving filter for oblique face images based on adaptive hopping Gaussian mixtures. *IEEE Access* 7 (2019), 142623–142639.
[167] Jeremy Schiff, Marci Meingast, Deirdre K. Mulligan, Shankar Sastry, and Ken Goldberg. 2009. Respectful cameras: Detecting visual markers in real-time to address privacy concerns. In *Protecting Privacy in Video Surveillance*. Springer, 65–89.
[168] Laura E. Schulz and Elizabeth Baraff Bonawitz. 2007. Serious fun: Preschoolers engage in more exploratory play when evidence is confounded. *Developmental Psychology* 43, 4 (2007), 1045.

[169] Securiti. 2024. Data Privacy Laws and Regulations Around the World. Retrieved August 2022 from https://securiti.ai/data-privacy-laws/
[170] Calder Security. 2022. Surveillance Camera Laws. Retrieved April 2024 from https://www.caldersecurity.co.uk/surveillance-camera-laws-for-business/
[171] A. Shamir. 1979. How to share a secret. *Commun. ACM* 22, 11 (1979), 612–613.
[172] Shodan. 2022. Search Engine for the Internet of Everything. Retrieved August 2022 from https://www.shodan.io/
[173] Reza Shokri, Marco Stronati, Congzheng Song, and Vitaly Shmatikov. 2017. Membership inference attacks against machine learning models. In *Symposium on Security and Privacy*. IEEE, 3–18.
[174] SmartCites. 2019. 6 US cities top list of world's most surveilled. Retrieved August 2022 from https://bit.ly/3x4zo5X
[175] Marcus Smith and Seumas Miller. 2022. The ethical application of biometric facial recognition technology. *AI and Society* 37, 1 (2022), 167–175.
[176] Mike Snider. 2020. Clearview AI, which has facial recognition database of 3 billion images, faces data theft. Retrieved August 2022 from https://www.usatoday.com/story/tech/2020/02/26/clearview-ai-data-theft-stokes-privacy-concerns-facial-recognition/4883352002/
[177] Hosik Sohn, Wesley De Neve, and Yong Man Ro. 2011. Privacy protection in video surveillance systems: Analysis of subband-adaptive scrambling in JPEG XR. *Transactions on Circuits and Systems for Video Technology* 21, 2 (2011), 170–177.
[178] Colin Soutar, Danny Roberge, Alex Stoianov, Rene Gilroy, and Bhagavatula Vijaya Kumar. 1998. Biometric encryption using image processing. In *Optical Security and Counterfeit Deterrence Techniques II*. SPIE, 178–188.
[179] ISO standards. 2022. STANDARDS. Retrieved August 2022 from https://www.iso.org/standards.html
[180] Qianru Sun, Liqian Ma, Seong Joon Oh, Luc Van Gool, Bernt Schiele, and Mario Fritz. 2018. Natural and effective obfuscation by head inpainting. In *Conference on Computer Vision and Pattern Recognition*. IEEE, 5050–5059.
[181] Qianru Sun, Ayush Tewari, Weipeng Xu, Mario Fritz, Christian Theobalt, and Bernt Schiele. 2018. A hybrid model for identity obfuscation by face replacement. In *ECCV*. 553–569.
[182] Xin Sun, Chengliang Tian, Changhui Hu, Weizhong Tian, Hanlin Zhang, and Jia Yu. 2022. Privacy-preserving and verifiable SRC-based face recognition with cloud/edge server assistance. *Computers and Security* 118 (2022), 102740.
[183] Latanya Sweeney. 2002. k-anonymity: A model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 10, 05 (2002), 557–570.
[184] Collins T. 2019. Facial recognition: Do you really control how your face is being used? USA Today. Retrieved August 2022 from https://www.usatoday.com/story/tech/2019/11/19/police-technology-and-surveillance-politics-of-facial-recognition/4203720002/
[185] Taylored. 2022. CCTV Security. Retrieved August 2022 from https://www.taylored.com/blog/cctv-security-how-to-detect-and-eliminate-botnet-attacks/
[186] Juan Ramón Troncoso-Pastoriza, Daniel González-Jiménez, and Fernando Pérez-González. 2013. Fully private non-interactive face verification. *Transactions on Information Forensics and Security* 8, 7 (2013), 1101–1114.
[187] Juan Ramón Troncoso-Pastoriza and Fernando Pérez-González. 2012. Fully homomorphic faces. In *ICIP*. IEEE, 2657–2660.
[188] Matthew Turk and Alex Pentland. 1991. Eigenfaces for recognition. *Journal of Cognitive Neuroscience* 3, 1 (1991), 71–86.
[189] Matthew A. Turk and Alex P. Pentland. 1991. Face recognition using eigenfaces. In *Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE Computer Society, 586–587.
[190] Maneesh Upmanyu, Anoop M. Namboodiri, Kannan Srinathan, and C. V. Jawahar. 2010. Blind authentication: A secure crypto-biometric verification protocol. *Transactions on Information Forensics and Security* 5, 2 (2010), 255–268.
[191] Maneesh Upmanyu, Anoop M. Namboodiri, Kannan Srinathan, and C. V. Jawahar. 2010. Blind authentication: A secure crypto-biometric verification protocol. *Transactions on Information Forensics and Security* 5, 2 (2010), 255–268.
[192] Venice-2. 2022. MOT benchmark. Retrieved July 2022 from https://motchallenge.net/vis/Venice-2
[193] Preethi Vennam, Pramod TC, Thippeswamy BM, Yong-Guk Kim, and Pavan Kumar BN. 2021. Attacks and preventive measures on video surveillance systems: A review. *Applied Sciences* 11, 12 (2021), 5571.
[194] Nishant Vishwamitra, Bart Knijnenburg, Hongxin Hu, and Yifang P. Kelly Caine. 2017. Blur vs. block: Investigating the effectiveness of privacy-enhancing obfuscation for images. In *IEEE Conference on Computer Vision and Pattern Recognition Workshops*. 39–47.
[195] Paul Voigt and Axel Von dem Bussche. 2017. The EU general data protection regulation (GDPR). *A Practical Guide, 1st Ed., Cham: Springer International Publishing* 10, 3152676 (2017), 10–5555.
[196] Han Wang, Shangyu Xie, and Yuan Hong. 2020. VideoDP: A flexible platform for video analytics with differential privacy. *Proceedings on Privacy Enhancing Technologies* 2020, 4 (2020), 277–296.

[197] Junjue Wang, Brandon Amos, Anupam Das, Padmanabhan Pillai, Norman Sadeh, and Mahadev Satyanarayanan. 2017. A scalable and privacy-aware IoT service for live video analytics. In *ACM Multimedia Systems Conference*. 38–49.
[198] Kai Wang, Bo Zhao, Xiangyu Peng, Zheng Zhu, Jiankang Deng, Xinchao Wang, Hakan Bilen, and Yang You. 2022. FaceMAE: Privacy-preserving face recognition via masked autoencoders. *arXiv preprint* arXiv:2205.11090.
[199] Yitu Wang and Takayuki Nakachi. 2020. A privacy-preserving learning framework for face recognition in edge and cloud networks. *IEEE Access* 8 (2020), 136056–136070.
[200] Earth Web. 2022. Do Cameras Deter Crime? Retrieved July 2022 from https://earthweb.com/surveillance-camera-statistics/
[201] Yunqian Wen, Bo Liu, Jingyi Cao, Rong Xie, Li Song, and Zhu Li. 2022. IdentityMask: Deep motion flow guided reversible face video de-identification. *Transactions on Circuits and Systems for Video Technology* 32, 12 (2022), 8353–8367.
[202] Yunqian Wen, Bo Liu, Ming Ding, Rong Xie, and Li Song. 2022. Identitydp: Differential private identification protection for face images. *Neurocomputing* 501 (2022), 197–211.
[203] Zack Whittaker. 2022. Chinese database leak. Retrieved July 2022 from https://techcrunch.com/2022/08/30/china-database-face-recognition/
[204] Davey Winder. 2022. Forbes: Apple's iPhone FaceID hacked in less than 120 seconds. Retrieved July 2022 from https://www.forbes.com/sites/daveywinder/2019/08/10/apples-iphone-faceid-hacked-in-less-than-120-seconds/?sh=5ae4094c21bc
[205] Thomas Winkler and Bernhard Rinner. 2014. Security and privacy protection in visual sensor networks: A survey. *ACM Computing Surveys* 47, 1 (2014), 1–42.
[206] WOS. 2022. WEB OF SCIENCE. Retrieved July 2022 from https://www.webofscience.com/
[207] WSJ. 2021. Alibaba. Retrieved July 2022 from https://www.wsj.com/articles/alibaba-falls-victim-to-chinese-web-crawler-in-large-data-leak-11623774850
[208] Yifan Wu, Fan Yang, Yong Xu, and Haibin Ling. 2019. Privacy-protective-GAN for privacy preserving face de-identification. *Journal of Computer Science and Technology* 34, 1 (2019), 47–60.
[209] Wanli Xue, Wen Hu, Praveen Gauranvaram, Aruna Seneviratne, and Sanjay Jha. 2020. An efficient privacy-preserving iot system for face recognition. In *Workshop on Emerging Technologies for Security in IoT*. IEEE, 7–11.
[210] Nazhao Yan, Hang Cheng, Meiqing Wang, Qinjian Huang, and Fei Chen. 2021. DP-Face: Privacy-preserving face recognition using siamese network. In *International Symposium on Distributed Computing and Applications*. IEEE, 92–95.
[211] Andrew C. Yao. 1982. Protocols for secure computations. In *Annual Symposium on Foundations of Computer Science*. IEEE, 160–164.
[212] Andrew Chi-Chih Yao. 1986. How to generate and exchange secrets. In *FOCS*. IEEE, 162–167.
[213] Linwei Ye, Binglin Li, Noman Mohammed, Yang Wang, and Jie Liang. 2018. Privacy-preserving age estimation for content rating. In *International Workshop on Multimedia Signal Processing*. IEEE, 1–6.
[214] Lin Yuan and Touradj Ebrahimi. 2017. Image privacy protection with secure JPEG transmorphing. *IET Signal Processing* 11, 9 (2017), 1031–1038.
[215] Guangsheng Zhang, Bo Liu, Tianqing Zhu, Andi Zhou, and Wanlei Zhou. 2022. Visual privacy attacks and defenses in deep learning: A survey. *Artificial Intelligence Review* 55, 6 (2022), 4347–4401.
[216] Yupeng Zhang, Yuheng Lu, Hajime Nagahara, and Rin-ichiro Taniguchi. 2014. Anonymous camera for privacy protection. In *ICPR*. IEEE, 4170–4175.
[217] Qiang Alex Zhao and John T Stasko. 1998. *The Awareness-privacy Tradeoff in Video Supported Informal Awareness: A Study of Image-filtering Based Techniques*. Technical Report. Georgia Institute of Technology.
[218] Ying Zhao and Jinjun Chen. 2022. A survey on differential privacy for unstructured data content. *ACM Computing Surveys* 54, 10 (2022), 1–28.
[219] Yuan Zhao, Bo Liu, Tianqing Zhu, Ming Ding, and Wanlei Zhou. 2022. Private-encoder: Enforcing privacy in latent space for human face images. *Concurrency and Computation: Practice and Experience* 34, 3 (2022), e6548.
[220] Kai Zhou and Jian Ren. 2018. PassBio: Privacy-preserving user-centric biometric authentication. *Transactions on Information Forensics and Security* 13, 12 (2018), 3050–3063.
[221] Hua-Hong Zhu, Qian-Hua He, and Yan-Xiong Li. 2012. A two-step hybrid approach for voiceprint-biometric template protection. In *International Conference on Machine Learning and Cybernetics*. IEEE, 560–565.
[222] Jaroslav Šeděnka, Sathya Govindarajan, Paolo Gasti, and Kiran S. Balagani. 2015. Secure outsourced biometric authentication with performance evaluation on smartphones. *Transactions on Information Forensics and Security* 10, 2 (2015), 384–396.

Received 14 October 2022; revised 3 June 2024; accepted 6 June 2024