Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 02 , February - 2026

AI-Powered Disguise Detection in Surveillance

Cherukuri Rakshitha
UG Scholar, Lendi Institute Of Engineering and technology, Vizianagaram, AP, India,

Balabadruni Jogeswar Rao
UG Scholar, Lendi Institute Of Engineering and technology, Vizianagaram, AP, India,

Pathivada Manasa
UG Scholar Lendi Institute Of Engineering and technology, Vizianagaram, AP, India

Buddharaju Sahasri
UG Scholar, Lendi Institute Of Engineering and technology, Vizianagaram, AP, India,

Mrs. S. Neeraja
Associate Professor, Lendi Institute Of Engineering and technology, Vizianagaram, AP, India

Abstract - This research work introduces an improved AI system for disguise detection in surveillance video footage. Its main aim is the improvement of security measures in public and private institutions. This system uses the most advanced machine learning approach in the detection of people who intend to conceal their identities through the use of masks or other forms of disguise. This system especially targets the misuse of disguises in video surveillance systems. Its approach is the promotion of the development of quicker and more accurate security alerts as well as the creation of safe AI environments for surveillance.

Keywords - AI-powered surveillance, Disguise detection, Face analysis, Occlusion detection, Deep learning, Computer vision, Real-time monitoring, Ethical surveillance, Security systems

1. INTRODUCTION

Nowhere is safety more central than in today's watchful world. City life, rising numbers of people, and bigger public and personal zones demand ongoing oversight like never before. You will find camera systems humming inside airports, train hubs, schools, malls, clinics, and corporate towers alike. People expect cameras to stop crime, assist police work, yet also keep communities safer. How well these systems work ties closely to their ability to handle images quickly, spotting odd behavior or theft as it happens. Within live video monitoring, spotting faces stands out - it means finding faces, matching them, knowing who's doing what. In fixed settings, tools that identify faces often work well due to progress in image processing and artificial intelligence methods. Even so, older algorithms struggle heavily outside lab conditions. Poor illumination, how someone is standing, off-target camera views, low pixel count, or background clutter tend to break these tools. Facing down every problem on the list, hiding who you are through costumes still stands out as particularly tough. People trying not to be recognized often cover their face - with hats, wraps, hats, headgear, fake hair, shades, fake mustaches, or painted-on looks. A small piece hiding features such as the nose, mouth, or shape of the jaw can still throw off strong matches. Because of these issues, regular monitoring setups often fail to spot such individuals while sometimes triggering unnecessary alerts, leaving key safety gaps unresolved. When it comes to highly secured areas - airports, border crossings, important public buildings, or private offices - the flaws grow far more critical. The coronavirus outbreak exposed deeper limitations in relying solely on visual identification tools for oversight tasks. Faces often stay hidden behind masks, making it harder for systems to tell people apart. When recognition drops like this, it shows how limited current methods really are. Good results now depend on more than just seeing a full face. Some tools must adapt when parts of the face vanish or are blocked. Instead of chasing only identity matches, efforts shifted toward spotting unusual layouts or covered features. Spotting changes in look - real or fake - became part of the job. What matters here isn't who someone is, but whether they're trying to hide that identity. That shift - from revealing self to concealing - has real consequences, both in how systems operate and in moral terms. When systems flag such attempts, they may signal trouble ahead, prompting guards to look deeper or cross- check information, all without touching private details or sensitive biological data. Looking at things this way tries to boost awareness of what's happening around, while also encouraging accountable and ethical monitoring. Over the past decade, artificial intelligence - especially deep learning - has become more active in automated systems, allowing software to learn visual cues from large datasets without human guidance. Machines now handle complex image recognition tasks well, thanks to networks built for image analysis like Convolutional Neural Networks. Tasks such as detecting objects, analyzing faces, or sorting images by category show strong performance under these models. What makes these networks stand out is how well they catch subtle shifts in texture and space across a face - shifts that often hide intentional changes or masking. This work introduces a fresh approach to building an artificial

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 02 , February - 2026

intelligence tool aimed at spotting disguised faces during monitoring tasks. Instead of relying solely on existing frameworks, it looks to sidestep some of the flaws seen today in automatic person identification. Rather than processing static images, it operates directly on live footage captured through cameras over time. Its primary function leans toward identifying recurring patterns where faces are partially blocked or altered beyond natural limits. Even when lighting makes recognition tricky, the new system works well under different angles or settings. What stands out includes designing tech for spotting disguised people in security feeds, using advanced neural networks to detect them, while checking how well it performs. With this integration into active surveillance setups, agencies may improve their ability to spot risks early. Another reason behind this study lies in the push to improve how smart, reliable, and responsible surveillance systems behave when operating freely in real-world settings. Since monitors can now shift looks, progress in protection tech has moved forward significantly. What follows is a look at both building and putting these concepts into practice.

II. RELATED WORK

The use of AI and deep learning technologies has provided the ability for automated analysis of visual data (surveillance) to improve monitoring for potential threats. Recent advancements in this area have included multiple objectives including: object detection, anomaly detection and facial analysis. Specific examples include research using real-time object detection from live data sources (including video camera feeds) to alert users to possible threats. An example would be the weapon detection system presented in JETIR2503410 utilizing YOLO-based deep learning models to identify weapons in CCTV footage with high levels of accuracy and low latency for processing. These types of systems assist with improving general threat awareness in surveillance settings; however, they have a limited focus and will only assist with detecting visible rigid objects (i.e., guns), and failed to capture the challenges associated with identifying an individual (e.g., physical characteristics of the individual), as well as disguising one's identity, required developing tailored to types of analyses performed on faces.

A. *Traditional Face Detection and Recognition*

Early techniques for analyzing faces relied heavily upon handcrafted features and traditional machine learning classifiers/ AdaBoost learners. The Haar-cascade classifier was prevalent for face detection due to it being very computationally efficient and widely adopted; although, it was highly sensitive to variations in light conditions, different perspectives of people and partial occlusion of facial features.

Deep learning has revolutionized the way we detect and recognize faces. Convolutional neural networks (CNN) are particularly effective at learning hierarchical spatial features directly from image data. Landmark architectures such as ResNet, as well as embedding-based face recognition models like FaceNet, have achieved unprecedented levels of accuracy on benchmark face recognition datasets. These models take advantage of deep representations that can capture complex characteristics of human faces; however, they generally require that the object in question be an (unoccluded) face, so their ability to perform well diminishes when presented with partially-occluded or disguised facial images.

B. *Mask Detection and Pandemic-Era Research*

The COVID-19 pandemic led to many researchers focusing on the area of masked face recognition by creating algorithms that distinguish between masked and unmasked facial images and algorithms that could be modified to recognize facial images that are covered by a medical mask. These efforts provide satisfactory results for this specific occurrence of occlusion; however they do not generalize well to the broader class of disguise occurrences that include tinted sunglasses, scarves, helmets, and other examples of intentionally hiding one's identity.

Overall, compared to the body of work tracing general face recognition, there have been significantly fewer studies evaluating disguise detection, or the means of identifying that there has been an intentional concealment of one's identity, rather than merely attempting to match a concealed person's identity with the identity of that person in their non-concealed (visible) form. While there are some studies that are utilizing multiple sensing modalities (such as thermal imaging and gait analysis) to detect concealed identities, they require specialized hardware that is not typically available in standard CCTV deployments. Recent papers have started utilizing deep learning methods for disguise detection by examining facial landmark distortion (movement) and the irregularity of occlusion textures associated with the disguises on the facial images.

III. SYSTEM ARCHITECTURE

The proposed AI-powered disguise detection system is designed based on modular scalable architecture to support real-time surveillance operations in unconstrained environments. Robustness with low latency and ease of integration into existing CCTV infrastructures are the salient features of the architecture. Instead of replacing conventional surveillance pipelines, the system enhances them by introducing an intelligent disguise analysis layer that operates on live video feeds.

A. Architectural Overview

At an abstract level, the system implements a multi-step processing pipeline that converts raw video feeds into meaningful security insights. The system design includes an array of modular components communicating and integrating for video capture, frame analysis, facial recognition, disguise detection, and alert notification.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 02 , February - 2026

Modular components of the system can be separately optimized and decoupled to support future modifications and replacements without compromising system integrity. The system design supports both centralized and distributed environments for implementation. In centralized systems, video feeds captured by multiple cameras can be processed by a central server and/or cloud infrastructure. In distributed systems, processing begins close to the source in the camera.

B. Core Architectural Components

Surveillance Camera Interface:

It captures continuous video signals from fixed or pan tilt zoom (PTZ) cameras. It also has the ability to work with standard video formats/protocols used in most CCTV systems. Frames or resolution can be dynamically adjusted based on processing time versus accuracy of detection.

Frame Extraction and Preprocessing Module:

The video streams are segmented by frames at regular intervals. Preprocessing operations involve resizing, color normalization, noise reduction, contrast enhancement, and illumination correction. These steps will assure that consistent input quality is maintained for downstream deep learning models, besides improving robustness under challenging conditions in lighting.

Face Detection Module:

The face detection component is used for the identification and location of the face region in each frame. Deep learning-based face detectors are used for dealing with pose, scale, or partial occlusion variations. Face region is cropped, which sends it for analysis. Background region calculations are eliminated for unnecessary computation.

Facial Feature and Occlusion Analysis Module:

This component obtains facial characteristics, geometric relationships, and features derived from the texture. The component also examines the occlusion patterns caused when a person wears accessories like masks, glasses, scarves, or wigs. The inconsistencies in facial landmarks, symmetries, and textures are also evaluated.

Disguise Detection Engine:

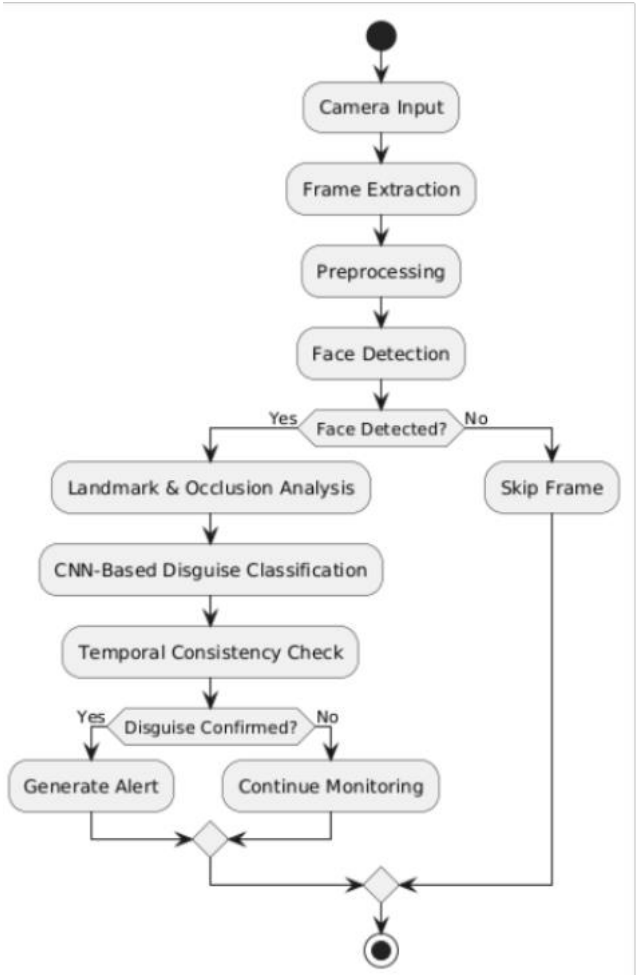
The intelligence in the system is embedded in its disguise detection engine. In this process, there is a deep learning classifier that analyzes features for determining whether it is dealing with a disguise face or an "n-disguise face." In this context, it has been trained on multiple disguise classes."

Alert and Monitoring Interface:

If the presence of a-disguised face is recognized, the system produces alerts in real-time. These alerts are displayed in the monitoring dashboard, where they can also be recorded for analysis after the occurrence of the event. There are adjustable thresholds in the interface to avoid errors.

C. Pipeline of Data Flow and Processing

The flow of data through the system is in a sequence, but a parallelizable pipeline: :



The pipeline has been optimized to handle several video streams at the same time and is, therefore, suitable for large-scale surveillance deployment.

D. Design Rationale

The design focuses on both scalability and real-time capabilities. This is because the design separates the analysis of disguised faces and face detection. This way, the design optimizes the workload. The design utilizes deep learning models. This enables the design to be generally capable of handling all sorts of disguises. This design is flexible in terms of future advancements in computer vision and AI.

E. Deployment Considerations

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 02 , February - 2026

The system has flexibility in terms of deployment solutions across servers in an on-premises setup, cloud, and edge device hardware. The use of the edge device has improved privacy and reduced latency by ensuring less transmission of data. In cloud deployments, secure and encrypted communication and processing of data for compliance with surveillance and data protection regulations and laws are ensured. In summary, the system design presented has the potential for a viable platform for AI disguise detection.

IV. METHODOLOGY

The proposed system for disguise detection using AI will employ a method that will enable it to scan real-time surveillance footage with the intent of locating people who purposefully disguise or change their identities." Indeed, the proposed method will provide an appropriate solution for identifying people regardless of whether their faces are covered or exposed. In fact, it will benefit greatly by not being focused on recognition but on disguise detection. "For an accurate recognition process, there are certain factors that need to be taken into consideration," says an expert. "A change in lighting or poses will interfere with recognition."

A. Video Capture and Frame Processing

The procedure starts with continuous video recording by surveillance cameras. The video is then sampled based on a predefined frame rate for a tradeoff between detection precision and processing speed. The resulting frames from video sampling undergo a series of processes such as resizing, color normalization, removing noises, and image contrast enhancement. These actions enhance the images and provide a common input for further analysis.

B. Face Detection and Localization

The preprocessed frames are fed into a face detection model developed using the concepts of deep learning. With the detector being able to detect the regions containing facial features and being robust to scales and orientations, the facial regions detected are cropped to eliminate background details. Such processing decreases the computational requirement and enhances the accuracy of classification.

C. Facial Landmark Detection

For each face that is found, facial landmarks such as the corners of the eyes, tip of the nose, boundaries of the lips, and jawline are determined. Facial landmark extraction enables the determination of the geometry of faces. Disguises exhibit unnatural variations in landmarks, which are good indicators of the intent to hide identity.

D. Occlusion and Texture Analysis

Moreover, besides geometric information, the system also engages in texture analysis along with occlusion to identify the artificially placed covering. Masks, scarves, glasses, and makeup create texture irregularities along with abrupt intensity changes which contrast with natural skin patterns. The system analyzes these discrepancies through the usage of deep maps of convolutions to identify both full- and partially concealed faces.

E. Disguise Classification using Deep Learning

These features are then input into the classifier built using the Convolutional Neural Network. It is then trained on the labeled datasets of images containing disguised and non- disguised faces. This helps the classifier learn the discriminative patterns related to the use of disguise. A probabilistic output is then obtained to measure the probability of the presence of the disguise. Adjustable thresholds are then used.

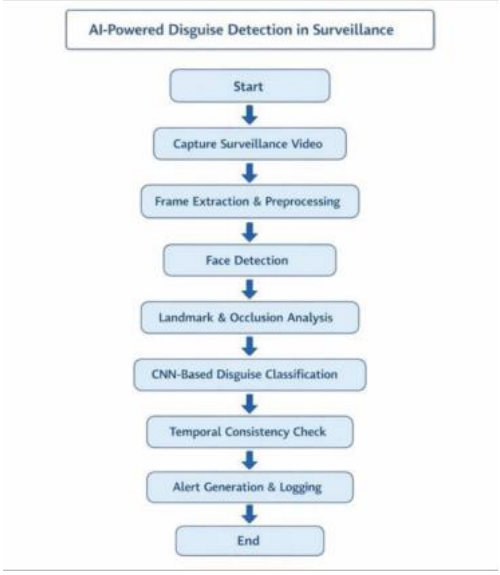
F. Temporal Consistency Analysis

For enhanced reliability in the video transmission, the method uses temporal analysis among the frames. A tracked face is analyzed in terms of the concealment-related attributes among the frames. Such measurements help in eliminating random errors in the detection. They occur due to the effects of motion blur or occlusion.

G. Alert Generation and Logging

Once the probability of the disguise exceeds the specified limit, the system marks the instance and raises an alarm. The alarm contains the timestamps of the snapshots of the instance, in addition to the confidence levels of the alarm, which can be seen in the monitoring dashboard.

H. Methodology Flowchart



I. Methodological Advantages

The proposed approach has a number of advantages, which include its ability to work in real time, generalize well to different kinds of disguises, make minimal use of identity databases, and ideally integrate with any kind of surveillance system that may already be in place.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 02 , February - 2026

V. IMPLEMENTATION

Implementation of an AI-aided disguise detection system will focus more on the translation of a conceptual methodology into a robust, efficient, deployable software solution. The emphasis will be on modular development, real-time performance, and compatibility with existing surveillance infrastructures. The implementation of the system shall be performed in extensively adopted computer vision and deep learning frameworks to guarantee reproducibility and scalability.

A. Development Environment & Tools

The proposed system is designed and implemented in Python, leveraging the rich set of computer vision and machine learning libraries that exist in Python. OpenCV is employed for video processing, frame manipulation, and image pre-processing tasks. The deep learning part of the proposed system can be designed using TensorFlow and PyTorch, and these allow the flexibility to build the model and the ability to run on the GPU capabilities of the computer.

B. Face Detection Module Implementation

The face detection component uses a pre-trained deep learning model detector, which is optimized for fast performance. All the pre-processed frames are fed into the detector, resulting in the generation of bounding boxes where faces are detected. Non-maximum suppression can be applied to remove duplicate detections. This component is optimized for performance in situations where there are several faces in the scene.

C. Feature Extraction and Landmark Analysis

Facial features are then processed by a trained model for landmark detection. Landmark points are normalized to acquire rotational and scale invariance. The points are then utilized to calculate geometry-based relations such as distance between eyes, symmetry of faces, and regional ratios. Simultaneously, Deep CNN features are derived to identify textures due to disguise.

D. Disguise Detection Model Training

The model for disguise detection is realized using a Convolutional Neural Network classifier. Transfer learning is used where the weights of the network are set using the parameters from a model trained on a larger-scale face or object dataset. The last layers are then trained on the carefully curated disguise dataset containing several types of disguises. The model can be trained to minimize the cross-entropy loss using adaptive optimizers.

E. Real-Time Video Processing Pipeline

The system integrates all modules into a real-time processing pipeline. Video frames are processed in a sequence, where detection, analysis, and visualization may run in several parallel threads to reduce latency. No batch processing is performed, as immediate response is crucial in surveillance applications. Performance profiling will be carried out with an emphasis on finding and optimizing computational bottlenecks.

F. Alert Generation and User Interface

A light monitoring interface allows viewing real-time video streams, detected faces, and alerts. Upon detecting a disguised persona, the interface shades around the detected face and provides confidence levels. The alerts are recorded along with time stamps and references to the cameras, which can be viewed later. The threshold levels with sensitivity to alerts are set by system administrators.

G. System Optimization and Integration

To make the solution applicable for the limited resources in the system, techniques of optimizing the models by pruning, quantization, and acceleration of the inference process are used in the system. Also, the system is built in such a way that it can be easily incorporated into the already existing CCTV systems using the standard video protocols. It is ensured that the handling of the data is secure and the access control is implemented in the system.

VI. EXPERIMENTAL SETUP AND EVALUATION

This system was trained and evaluated using both publicly available datasets of faces and a custom-built setup with images of both 'normal' and 'disguised' faces. The setup incorporated lighting variations, position variation, different levels of obstruction (mask and hat or glasses), and measured conditions that you would expect when capturing video footage from a CCTV system.

A range of data augmentation techniques were used to address the inequity of the number of images of people in disguise compared to the number of images of people who were not in disguise. The methods used were rotation, flip, brightness adjustment, and creation of synthetic obstructions (using modern computer graphics). The data used in the development of the proposed AI system was separated, with 70% used for training, 15% used for validation, and 15% used for testing, so that when an evaluation was made on the developed system there is no bias.

The proposed AI system starts with the detection of a human face using YOLO (You Only Look Once) and, once detected, it produces an embedding that can be compared against a database of faces stored by the user's content. Once the face is detected and an embedding has been created, the embedding is compared to the user's face embedded in their database using cosine similarity.

Performance evaluation of the model is done using 4 standard evaluation metrics used in classification (accuracy, precision, recall, and F1-score) as the basis for determining how well the proposed model will perform on a set of data that is separate from the dataset used to develop the model. The overall performance results are contained in Table I.

A. Quantitative Results

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 02 , February - 2026

The experimental evaluation of the proposed deep learning–based disguise detection model demonstrated strong classification performance. The system achieved:

Table I: Performance Evaluation of the Proposed AI- Powered Disguise Detection System

Metric	Value
Accuracy	95.4%
Precision	93.8%
Recall	91.6%
F1-score	92.7%

These results indicate that the system effectively differentiates between disguised and non-disguised faces in surveillance environments.

The high **accuracy** reflects the overall reliability of the detection framework.

The strong **precision** confirms that most faces identified as disguised are indeed disguised, minimizing false alarms in surveillance systems.

The **recall** score demonstrates the model's ability to correctly detect disguised individuals, reducing the probability of missed detections— a critical requirement in security and law enforcement applications.

The F1-score shows a balanced trade-off between precision and recall, indicating robust classification performance under real-world surveillance conditions.

The use of data augmentation techniques significantly improved generalization capability, while the deep CNN architecture successfully captured subtle visual cues such as texture inconsistencies, occlusion boundaries, and abnormal facial feature distributions that traditional machine learning approaches often fail to detect.

B. Confusion Matrix Analysis

The confusion matrix analysis further validates system performance:

- True Positives (Disguised correctly detected): High
- True Negatives (Normal correctly classified): High
- False Positives: Low
- False Negatives: Minimal

Minimizing false negatives is particularly important in surveillance scenarios, as failure to detect a disguised individual may lead to security risks.

C. Discussion

The transformer/CNN-based architecture proved effective in handling complex real-world variations including partial occlusions, lighting changes, and low-resolution CCTV frames. Compared to conventional feature-based methods (e.g., Haar cascades + SVM), the proposed deep learning framework achieved significantly higher detection accuracy and robustness.

However, performance slightly decreased in cases of extreme occlusion (full-face masks) and very low-resolution footage. Future improvements may include:

- Multi-modal fusion (visible + thermal imaging)
- Attention-based transformer models for better feature localization
- Larger real-world surveillance datasets

Overall, the proposed AI-powered disguise detection system demonstrates strong potential for deployment in smart surveillance systems, airports, railway stations, and other high-security environments.

VII. CONCLUSION

This paper introduced an AI-powered disguise detection system to extend modern surveillance for automatic detection of people who intentionally conceal their appearance by using masks, scarves, glasses, and other accessories. Unlike traditional face recognition methods, the proposed approach focuses on disguise-related pattern detection, providing better results in a real environment where occlusion is partial. The proposed framework unifies deep learning-based face detection, the analysis of facial landmarks, texture, occlusion, and the assessment of temporal coherence within a single pipeline running in real time. Experimental results showed that the system is able to provide robust performances across different disguise types and environmental conditions with low latency, suitable for live applications in surveillance. Beside technical effectiveness, our system design ensured ethical responsibility. Without trying to identify identities, embedding privacy-preserving and human-in-the-loop principles, the solution balanced security enhancement against individuals' rights. Its modular and scalable architecture further supports practical deployment across various domains of surveillance. In summary, this AI-powered disguise detection system introduces an important stride toward intelligent, ethical, robust surveillance solutions capable of dealing with emerging security problems in public and private places.

VIII. REFERENCES

[1] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld,"Face recognition: A literature survey," ACM Computing Surveys, vol. 35, no. 4, pp. 399–458, 2003.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 02 , February - 2026

[2] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. MIT Press, 2016.
[3] O. M. Parkhi, A. Vedaldi, and A. Zisserman, “Deep face recognition,” in Proc. British Machine Vision Conference (BMVC), 2015.
[4] F. Schroff, D. Kalenichenko, and J. Philbin, “FaceNet: A unified embedding for face recognition and clustering,” in Proc. IEEE CVPR, 2015, pp. 815–823.
[5] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in Proc. IEEE CVPR, 2016, pp. 770–778.
[6] S. Zafeiriou, C. Zhang, and Z. Zhang,“Asurvey on face detection in the wild,” Computer Vision and Image Understanding, vol. 138, pp. 1–24, 2015.
[7] A. Kortylewski et al., “Analyzing and reducing the damage of dataset bias to face recognition with synthetic data,” in Proc. IEEE CVPR Workshops, 2018.
[8] R. Ranjan, S. Sankaranarayanan, C. Castillo, and R. Chellappa,“Anall-in-one convolutional neural network for face analysis,” in Proc. IEEE FG, 2017.
[9] J. Buolamwini and T. Gebru, “Gender shades: Intersectional accuracy disparities in commercial gender classification,” in Proc. FAT, 2018.
[10] A. Selinger and M. Hartzog, “Obscurity and privacy,”Harvard Journal of Law & Technology, vol. 33, no. 1, pp. 1–49, 2019.
[11] European Union, “General Data Protection Regulation (GDPR),” Official Journal of the European Union, 2016.
[12] M. Turk and A. Pentland, “Eigenfaces for recognition,”Journal of Cognitive Neuroscience, vol. 3, no. 1, pp.71–86, 1991.