Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

A Survey on Face Spoof Attack Detection Using Multimodal Machine Learning Approaches

Ahmed Kareem Shahloul Al-hchaaimi[1]*, Ali Abdulkarem Habib Alrammahi[2]
1 Kufa University, Computer Science and Mathematics , Computer Science, 54001, Iraq.
ahmedk.alhchaaimi@student.uokufa.edu.iq
[2] Kufa University, Computer Science and Mathematics , Computer Science, 54001, Iraq.
alia.alramahi@uokufa.edu.iq

https://doi.org/10.46649/fjiece.v4.2.6a.22.9.2025

Abstract. *Face anti-spoofing (FAS) has lately received more attention because of its critical function in protecting face recognition systems against presentation assaults (PAs). With the advent of more realistic face recognition systems of various sorts, early-stage face anti-spoofing approaches based on hand-crafted characteristics have become untrustworthy due to their limited representativeness. With the advent of large-scale academic datasets over the last decade, deep learning-based face anti-spoofing algorithms have outperformed and dominated the area. However, most existing evaluations in this field rely on hand-crafted features, which are now outdated and hinder progress in the face anti-spoofing research community.*

To inspire future research, we give the first complete analysis of recent advances in deep learning-based face anti-spoofing systems. This article addresses numerous innovative and significant components: 1) In addition to binary label supervision (e.g., "0" for truth versus "1" for face recognition systems), we also examine cutting-edge approaches using pixel-level supervision (e.g., pseudo depth map). 2) In addition to traditional evaluation within the dataset, we collect and analyze cutting-edge methods specifically designed for domain generalization and open-set FAS analysis; and 3) In addition to commercial RGB cameras, we summarize deep learning applications under multimodal (e.g., depth and infrared) or specialized (e.g., light field and flash) sensor categories. We end our survey by stressing existing outstanding challenges and future opportunities.

Keywords: *Deep Learning; Face Anti-Spoofing; Machine Learning; and Convolutional neural network*

1. INTRODUCTION

With the rapid advancement of computer science and internet technology, biometric technology is increasingly being used in identity identification. Currently, biometric technology focuses mostly on face and fingerprint identification. Face recognition, in particular, has received substantial attention and is widely employed in a variety of applications. However, current face anti-spoofing systems are vulnerable to attacks including realistic photographs, films, 3D masks, and even cosmetics and cosmetic surgery. As a result, a strong face anti-spoofing solution is critical to ensuring security. Face anti-spoofing entails evaluating whether a facial picture is a genuine face or a fake face assault, with the primary goal of properly determining whether the person in front of the camera is live or an image. Face anti-spoofing is used in financial transactions, access control, and other applications to increase security. There are some differences between actual and synthetic faces, which are notably evident in picture color texture information, motion information, image quality, and depth information [1].

The authors describe a spoofing attack as the act of misleading a biometric sensor by delivering

counterfeit biometric evidence of a legitimate user. This assault is simple; the attacker merely places a replica of a photograph in front of the sensor, with no previous knowledge of the recognition process. Biometric systems are vulnerable to spoofing attacks due to their architecture of recognizing IDs without validating the subject's existence [2].

In recent years, both CNN-based [3] and classical approaches [4] have showed the capacity to discriminate between genuine and forged faces. However, these techniques confront difficulties in determining the nature of spoofing patterns, such as loss of skin features and color distortion [5]. Many writers found that joining handmade traits with deep neural networks greatly boosts model performance. These aid with trait extraction and transfer learning (TL) where models use pre-learned info to make their performance better on new tasks[6].

Handcrafted features can achieve knowledge about the domain which deep learning models may not be able to achieve on their own. Adding these variables into the model helps it perform better and align more closely with the specific needs of the domain. An amalgamation of handcrafted features and a diminutive architecture of CNN, when computation resources are restricted, serves well as an effortless solution that maintains near results [7]. Therefore, this framework will include multilevel features for color space and texture data alongside CNN classification models that will employ TL for detecting face spoofing.

2. RELATED WORKS

This study links to several earlier works that struggle with identifying fake images since they display events that could have taken place at different times.

S. Priya et al., 2019 [8], introduced different methods for feature extraction in face spoofing which are "Histogram of Oriented Gradients (HOG)", "Local Binary Pattern (LBP)", and "Scale-Invariant Feature Transform (SIFT)" along with DL approaches called" VGG16, Shallow CNN, and Inception-ResnetV2". A comparison among these techniques was done in their work using three different classifiers: decision trees," Support Vector Machines (SVM)", and" Artificial Neural Networks (ANN)". For this experiment, the YALE dataset was utilized. It was observed that deep feature transfer from the Inception-ResnetV2 model gave classification accuracy close to 96% when classifying a spoof image versus a real image.

Polash Kumar Das and others, in 2019 [9], used LBP to get hand-crafted features and a VGG-16 convolutional neural network for deep features. This study uses data collected from the SSIJRI Dataset that has both real and fake images and videos of faces. From the results, it is seen that combining both types of traits improves sensitivity and makes detecting a fake face from an real one much easier, with accuracies reaching 96.97% on the SSIJRI dataset.

In Yi Lu et al., 2020 [10], a face-morphing attack was detected by distinguishing HOG and a multi-feature CNN with micro-texture and color features. This morphing attack detection method uses the HSV color space transformation where LBP and HOG serve for image preprocessing and then boost morphing attack traces in terms of micro-texture as well as color. These two feature maps plus the original image are then fed into a finely-tuned multi-feature CNN. Experimental results offered subjectivity levels up to 94.65%.

Neenu Daniel and A. Anitha, 2021 [11], a method for detecting face faking using entropy-based texture and quality characteristics was described and validated with the Replay Attack Database. The procedure starts with recognizing and resizing their faces using the Viola-Jones technique. The photos are then transformed to the hue, saturation, value, or HSV color space, and entropy-based texture and fractal dimensions are extracted. In addition, chromatic moments and blurriness are assessed to determine image quality. These attributes are then put out end to end, and binaries are identified using a K-Nearest Neighbors (KNN) classifier. The strategy achieved an accuracy of 98.2%, which is much superior to standard problem-solving methodologies.

Vinutha H1 and Dr. Thippeswamy G, 2022 [12], proposed a method for feature extraction based on

a Stockwell transform combined with an elegant pairing method that uses the Szudzik index to represent the feature vector of the Stockwell transform of face samples, supporting an SVM classifier for the detection of 2D face presentation attacks. Such studies were conducted on the NUAA Photograph Impostor dataset, with an average accuracy of 91.1%.

Leyla G. Muradkhanli and Parviz A. Namazli, 2023 [13], proposed a technique for detecting fraudulent facial alteration using CNNs. The approach entails training a CNN with many layers of convolution and pooling to discover distinguishing characteristics. The Replay Attack Dataset, which is publically available, was utilized for training and assessment. The experiment findings obtained an accuracy of 89%.

Muhammad Amir Malik et al., 2023 [14], a hybrid model was used by which combined machine learning and computer vision for face spoof detection. Features extracted included texture patterns, color distortions, and geometric distortions. Inconsistencies observed between different color spaces (HSV, YCbCr) were utilized thereafter by CNN-based classifiers. The same model was trained and tested on the MSU MFSD, IDIAP Replay Attack, and SiW datasets. The results revealed that the proposed CNN-based model achieved maximum accuracy near 87.5%.

ZHAOPENG XU et al., 2024 [15], created This system called AFace . It has two main parts: an iso-depth model which helps to map the structures of faces using echoes, and a range-adaptive or RA algorithm which adapts dynamically based on how far away the user is from the smartphone. Testing with 40 participants revealed that A Face had an average accuracy of 96.9% in rejecting image/video-based spoofing attempts.

A comprehensive review of relevant literature is conducted to better understand the existing advancements in this domain. The table below summarizes key approaches, methodologies, and findings from related work.

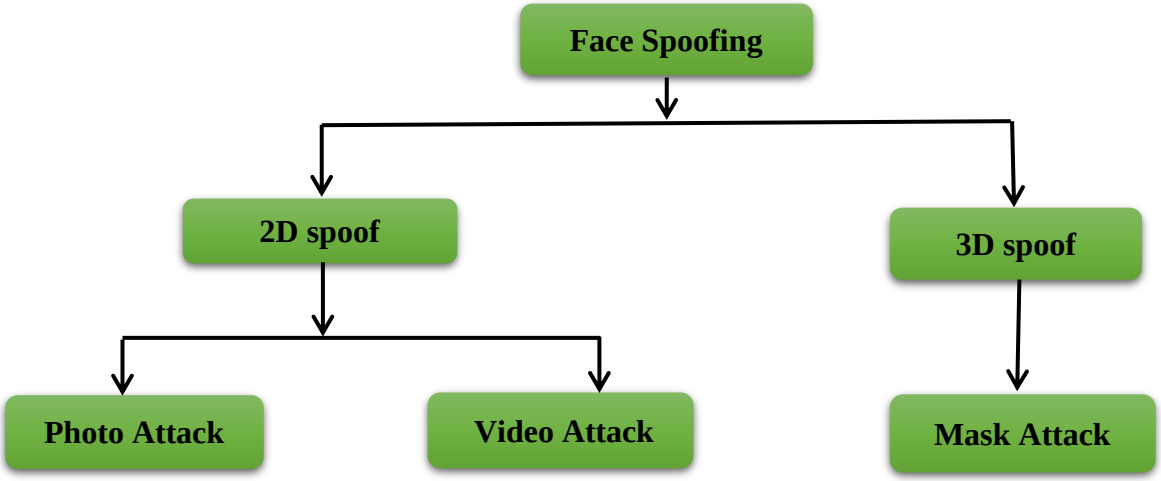
Table 1. Summary of Related Works

References, Year	Dataset	Feature Extraction	Classifier	Accuracy	limitations
[8] ,2019	YALE face dataset	LBP, Shallow CNN, SIFT, HOG, VGG16 and Inception -Res- Netv2	Decision Tree, ANN , SVM	96.23%	1. Computational Complexity & Resource Requirements 2. Generalization Issues 3. Feature Redundancy & Fusion Challenges 4. Real-time Implementation Constraints 5. Dataset Bias & Ethical Concerns
[9],2019	SSIJRI, Replay-Attack, Mobile-Replay, 3DMAD	LBP	VGG16	92.05, 75.25%, 90.52%, and 96.97%	1- Sensitivity to Noise 2- Limited Discriminative Power 3- Feature Extraction Overhead 4- Ineffectiveness Against AI-Generated Spoofs
[**10**], 2020	FEI	LBP, HOG	CNN	94.65%	1- Limited Texture Representation 2- Limited Scalability 3- False Positives and False Negatives

[11], 2021	Replay – Attack	facial image quality and texture features	KNN	98.2%	1- Vulnerability to High-Quality Spoofing Attacks 2- Limited Generalization Across Datasets 3- Dependence on Handcrafted Features 4- False Positives and False Negatives
[12], 2022	NUAA	Stockwell Trans- form	SVM with the radial basis kemel	91.1%	1- Computational Complexity 2- Overfitting Risk 3- Limited Robustness Against Advanced Attacks 4- Hardware Dependency
[13], 2023	replay at- tacks	CNN	CNN	89%	1- High Computational Cost 2- Overfitting to Training Data 3- Difficulty Handling Real-Time Detection
[14], 2023	MSU MFSD, ID- IAP Re- play At- tack, and SiW	texture patterns, color distortions, and geometric dis- tortions	CNN	87.5%	1- Computational Complexity 2- Difficulty in Capturing Live Movement 3- Dataset Bias
[15], 2024	40 volun- teers Users	Distance, Energy	Deep Neu- ral Network	96.9%	1- Limited Robustness Against Advanced Attacks 2- Feature Sensitivity 3- Computational Complexity 4- False Positives/Negatives 5- Difficulty Implementation

3. FACE SPOOFING

In the digital era, facial recognition technology has become a cornerstone of many security systems. However, the broad usage of this technology has resulted in the development of advanced kinds of face spoofing. Almost all biometric systems involve facial recognition. With such systems missing liveness detection mechanisms for face spoofing, it has become necessary to prioritize security [16]. Spoofing human faces is quite simple since it is one of the few biometric features that can only be deceived by a valid user data (i.e., photographs or videos) [17]. Spoofing attacks are among the most significant dangers to biometric-based authentication systems, and researchers have expressed worry about them. Spoofing attacks occur when an unauthorized user attempts to obtain access to a system using forged fingerprints. Today's mobile technologies are mobility devices with embedded sensors such as cameras that are extensively utilized in societies all over the globe. They are used not only for day-to-day communications, but also as a tool for conducting personal business and altering data using biometrics. Face recognition system attackers have advanced beyond typical picture prints and 3D images to include photo screening, video reply attacks, silicon 3D masks, and printed paper masks, as shown in Figure (1). [17].



Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

Fig. 1. Different Methods for Face Spoofing Attack

A flat printed picture is arguably the least expensive and most likely to occur; most individuals have their face photos on the Internet (such as social media) or might be captured by impostors without their knowledge or consent, as seen in Figure (2) [18].

Fig. 2. depicts several types of face spoofing attacks: (a) real user; (b) printed photo assault; (c) eye-cut picture; (d) warped photo; (e) video reply attack; (f) silicon 3D mask; and (g) printed paper mask [18].

Face spoofing detection is the technique of distinguishing between the genuine face and the phony face, which is done by several methods, including:

***A. Texture-Based Approaches*:** These methods identify photo assaults by evaluating texture patterns in pictures collected by sensors. They assume that real and fake ages have distinct texture characteristics. This method is popular for its simplicity [19].

***B. Motion-Based Approaches*:** These approaches compare the motion patterns of real users with collected photos. They presume that 2D faces move differently from actual ones. Optical flow in video sequences is used to analyze techniques like lip movement, head rotation, and eye blinking. High-quality pictures are required for proper analysis [19].

***C. Image Quality-Based Approaches*:** This method focuses on identifying quality variations between actual and simulated faces. It implies that phony faces are of inferior quality. Image quality is assessed by measuring features such as chromatic moments, blurriness, and specular reflection [19].

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

D. Frequency-Based Approaches: These methods use noise signals in captured videos to differentiate between real and spoofed faces. They use frequency analysis, assuming there are frequency fluctuations in recovered movies [19].

E. Other Approaches: Recent approaches include deep learning, person-specific algorithms, and others include 3D depth, infrared, and vein flow identification. Deep learning uses convolutional neural networks (CNNs) to identify face faking, whereas person-specific techniques rely on enrollment samples [19].

4. FACE DETECTION

Face detection is a computer system that detects the existence, size, and location of human faces in digital images. It concentrates solely on recognizing face characteristics, ignoring other factors like trees, buildings, and bodies. This procedure is an example of object-class detection, which involves identifying the positions and sizes of all objects in a picture that belong to a specified class. Face detection is a larger notion than face localization, which focuses on finding the positions and sizes of a set number of faces, usually only one [20]. Face detection provides an approximate estimation of the face's location and size, whereas face landmarking precisely recognizes facial characteristics such as the eyes, nose, mouth, and facial shape [21].

This may be performed with a land marking module or a face alignment module. ML approaches such as neural networks or (SVM) are commonly used to categorize regions of an image as face or non-facial based on training data [22].

Face recognition is one of the most compelling issues in computer vision, prompting the development of numerous software libraries for face detection. Face detection may be done using HOG, as seen in Figure (3). HOG is a lightweight technique that works well in conjunction with frontal [23].

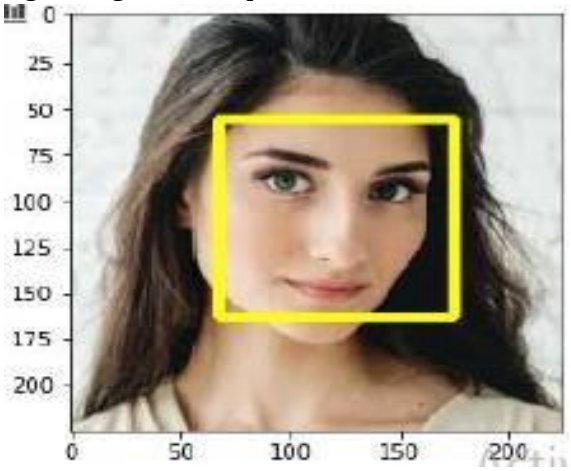


Fig. 3. HOG for Frontal Face Detection result [24].

4.1. Face Cropping & Resizing

Image cropping is a typical method for improving the visual quality of images. Captured photographs were enhanced using digital editing. Cropping is one of the most prevalent types of such adjustments, which removes sections of a picture that are not inside a specified rectangular zone [24]. The primary goal of picture resizing is to keep the essential contents of the photos while resizing the less-important areas, ensuring that image characteristics stay in the contents without any alterations. Some resizing procedures are applied to the original image, and the results are displayed [25].

1- Cropping-Based Resizing: This approach selects only one section of the image.
2. Resizing using Scaling Method: The image has changed from its original form.
3- Resizing using Scene's Content-Aware approach: This approach preserves the image's original form while removing sections with lower energy. [26].

Face spoofing detection systems are continually improving to address the issues posed by the growing use of phony photos and videos in fraudulent assaults. These approaches depend on several new algorithms developed to enhance spoofing detection. Significant gains have been made in integrating different algorithms and methodologies for face spoof detection, with temporal analysis being one of them that can significantly enhance system robustness and accuracy. Temporal analysis refers to analyzing temporal trends in video sequences for plagiarism detection, i.e., performing frame-by-frame processing and heartbeat detection using rPPG remote photoplethysmography. Temporal analysis can be fused with either motion or texture detection-based algorithms.

5. FEATURE EXTRACTION FOR SPOOF DETECTION

Three relevant face spoof detection methods are mostly built on feature extraction, which transforms the raw picture data into a feature vector composed of significant face points, etc., by using disparate techniques. This is because color, texture, geometric aspects, etc. are all essential features in face spoofing detection. After face posting, feature extraction enables us to differentiate between real and altered images [27].

5.1. Color Features

Color features are commonly used to extract color information from images, represent visual content, and distinguish between images [27]. Color spaces for images include grayscale, RGB, HSV, Ycbcr, and CMYK.

5.2. Texture Features

To begin, texture is a pattern that appears on surfaces and is distinguished by varying colors and intensities that are not always consistent. This feature is found on almost all surfaces and gives critical information on their structure and interactions with their surroundings [28]. The features of texture are represented by image texture features in feature space. The current dimensional im- age texture is calculated based on a set of characteristics (which are related to operators purposefully intended to extract textural attributes). LBP and HOG are well-known approaches for effective texture manipulations [29].

6. DATASETS FOR FACE ANTI-SPOOFING

- **CASIA-FASD**: A dataset for face anti-spoofing, containing both real and spoofed images.
- **Replay-Attack**: Includes video sequences of both genuine and spoofed faces for detecting attacks like photo and video replay.
- **Spoof in the Wild (SiW)**: A large-scale dataset with various real and spoofed images for evaluating anti-spoofing methods.

7. Future Works

Analogously to the traditional authentication methods based on color, image texture, image motion, and image quality; involve significant contributions to face spoofing detection, hybrid methods are now proving to be more efficient in combating complex and multi-modal attacks. These methods integrate a bulk of complementary detection methods in order to enhance accuracy and generalization in terms of datasets and attack types.

A hybrid approach usually combines several modalities e.g.

- Thermal Imaging: Thermal images are registered infrared radiation that contains unique thermal patterns from living tissue. Spoofing instances such as printed photos or silicone masks do not possess genuine thermal characteristics, which make thermal imaging a potent anti-spoofing modality. It has been successful in distinguishing live human skin from spoof materials.
- **Remote Photoplethysmography (rPPG)**: rPPG enables pulse rate estimation of a subject remotely, based on subtle changes in skin color that occur in time with blood flow. This thermophysiological feature is not shown in the process of a spoof attack, such as image or video input, thus making rPPG a good feature for liveness detection. Recent works combine rPPG and CNN-based models to improve the temporal feature extraction of CVFs.

- **Vision Transformers (ViTs):** Vision Transformers (ViTs) have shown great success in achieving state-of-the-art results on computer vision tasks, thanks to their attention mechanisms which allow for a modeling of long-range dependencies across their input image. In the context of spoof detection, ViTs have the ability to capture global features of spoofing artifacts and patterns neglected by traditional CNNs.

CONCLUSIONS

Researchers are continually improving algorithms to identify face faking with greater accuracy and speed. The most promising approaches are deep learning and multimodal solutions, which, when combined with behavior-based cues and sophisticated imaging capabilities (such as depth and infrared), form a solid foundation for current anti-spoofing systems.

Using many approaches for face spoof detection can considerably improve the system's capacity to identify various forms of assaults. A multimodal strategy that combines texture, motion, depth, and deep learning approaches is frequently the most effective. In addition, using ensemble learning, transfer learning, and explainable AI can improve the system's robustness and interpretability.

In addition in this study we conclude that the LBP method use for convert image to the binary image and get the texture features and based on previous works the higher accuracy was 96.97%, CNN use to extract the deep features from images and the accuracy was 96.23%, and ViT concert on physiological signals extracted from video sequences and this method will use in future work.

Modern algorithms for detecting face spoofing are critical components of today's digital security solutions. As spoofing tactics grow, there is an urgent need to create more advanced and effective methods to assure the safety and reliability of face recognition systems.

REFERENCES

[1] Y. Li, Y. Wang, and Z. Zhao, "Face anti-spoofing methods based on physical technology and deep learning," in Proc. Int. Conf. Computer Vision, Application, and Design (CVAD), vol. 12155, SPIE, pp. 173–184, Dec. 2021.
[2] I. Chingovska, A. Anjos, and S. Marcel, "On the effectiveness of local binary patterns in face anti-spoofing," in Proc. BIOSIG – Int. Conf. Biometrics Special Interest Group, pp. 1–7, Sept. 2012.
[3] Y. Liu et al., "Aurora guard: Real-time face anti-spoofing via light reflection," arXiv preprint arXiv:1902.10311, 2019.
[4] T. de Freitas Pereira, A. Anjos, J. M. De Martino, and S. Marcel, "LBP−TOP based countermeasure against face spoofing attacks," in Computer Vision – ACCV 2012 Workshops, Springer, pp. 121–132, 2013.
[5] Z. Wang et al., "Deep spatial gradient and temporal depth learning for face anti-spoofing," in Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition (CVPR), pp. 5042–5051, 2020.
[6] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?," in Adv. Neural Inf. Process. Syst. (NIPS), vol. 27, 2014.
[7] S. M. Sarwar, "FedBiometric: Image Features Based Biometric Presentation Attack Detection Using Hybrid CNNs-SVM in Federated Learning," Ph.D. dissertation, Univ. Texas Rio Grande Valley, 2023.
[8] S. Priya, S. Pawar, and A. Joshi, "Evaluation of local descriptors and deep CNN features for face anti-spoofing," Int. J. Recent Technol. Eng., vol. 8, no. 2(Special Issue 8), pp. 1644–1648, 2019.
[9] P. K. Das et al., "A new approach for face anti-spoofing using handcrafted and deep network features," in Proc. IEEE Int. Conf. Service Operations and Logistics, and Informatics (SOLI), pp. 33–38, Nov. 2019.

[10] Y. Lu, K. Xu, T. Sun, K. Qi, and L. Yao, "Face morphing detection with convolutional neural network based on multi-features," in Proc. Int. Conf. Aviation Safety and Information Technology, pp. 611–616, Oct. 2020.
[11] N. Daniel and A. Anitha, "Texture and quality analysis for face spoofing detection," Comput. Electr. Eng., vol. 94, p. 107293, 2021.
[12] H. Vinutha and G. Thippeswamy, "Spoof detection using Elegant Pairing method on Stockwell Transform in Face Biometric System," in Proc. Int. Conf., 2022.
[13] L. G. Muradkhanli and P. A. Namazli, "Face spoof detection using convolutional neural network," Problems Inf. Soc., pp. 40–46, 2023.
[14] M. A. Malik et al., "A novel deep learning-based method for real-time face spoof detection," in Proc. Int. Conf., 2023.
[15] Z. Xu et al., "AFace: Range-flexible anti-spoofing face authentication via smartphone acoustic sensing," Proc. ACM Interact. Mobile Wearable Ubiquitous Technol., vol. 8, no. 1, pp. 1–33, 2024.
[16] S. D. Thepade et al., "The comprehensive review of face anti-spoofing techniques," Int. J. Adv. Sci. Technol., vol. 29, pp. 8196–8205, 2020.
[17] M. K. Rusia and D. K. Singh, "A color-texture-based deep neural network technique to detect face spoofing attacks," Cybern. Inf. Technol., vol. 22, no. 3, pp. 127–145, 2022.
[18] L. Souza, L. Oliveira, M. Pamplona, and J. Papa, "How far did we get in face spoofing detection?," Eng. Appl. Artif. Intell., vol. 72, pp. 368–381, 2018.
[19] N. Daniel and A. Anitha, "A study on recent trends in face spoofing detection techniques," in Proc. Int. Conf. Inventive Comput. Technol. (ICICT), pp. 583–586, Nov. 2018.
[20] A. Kumar, A. Kaur, and M. Kumar, "Face detection techniques: A review," Artif. Intell. Rev., vol. 52, pp. 927–944, 2019.
[21] Y. H. Ali and I. S. A. Aljabar, "Real-time face recognition in video using linear discriminate analysis and local binary patterns," Eng. Technol. J., vol. 33, no. 4B, 2015.
[22] E. Hjelmås and B. K. Low, "Face detection: A survey," Comput. Vis. Image Underst., vol. 83, no. 3, pp. 236–274, 2001.
[23] A. Jadhav, S. Lone, S. Matey, T. Madamwar, and S. Jakhete, "Survey on face detection algorithms," Int. J. Innov. Sci. Res. Technol., vol. 6, no. 2, pp. 291–297, 2021.
[24] D. Vaquero et al., "A survey of image retargeting techniques," in Proc. SPIE Appl. Digit. Image Process. XXXIII, vol. 7798, pp. 328–342, Sep. 2010.
[25] A. Mumuni and F. Mumuni, "Data augmentation: A comprehensive survey of modern approaches," Array, vol. 16, p. 100258, 2022.
[26] K. A. Hassan, H. S. Abdullah, and A. E. Ali, "Proposed image similarity metric with multi-block histogram used in video tracking," Eng. Technol. J., vol. 34, no. 4B, pp. 578–584, 2016.
[27] N. J. Ibrahim and E. K. J. Alfatly, "Video image for security system by using chaotic oscillator for segmentation," Eng. Technol. J., vol. 27, 2009.
[28] M. Sonka, V. Hlavac, and R. Boyle, "Image processing, analysis and machine vision," Springer, 2013.