

A Review on Defake Detection Technique for Human Faces

1st Dr. Jayashri .P. Shinde

Department of Computer Engineering,
Marathwada Mitra Mandal's Institute of
Technology (MMIT), Pune
Jayashri.shinde@mmit.edu.in

2nd Himanshu Patil

Department of Computer Engineering,
Marathwada Mitra Mandal's Institute
of Technology (MMIT), Pune
Himanshu.patil22@mmit.edu.in

3rd Shruti Patil

Department of Computer Engineering,
Marathwada Mitra Mandal's Institute
of Technology (MMIT), Pune
Shruti.patil22@mmit.edu.in

4th Atharva Phuge

Department of Computer Engineering,
Marathwada Mitra Mandal's Institute
of Technology (MMIT), Pune
Atharva.Phuge22@mmit.edu.in

5th Rohan Mate

Department of Computer Engineering,
Marathwada Mitra Mandal's Institute
of Technology (MMIT), Pune
Rohan.mate22@mmit.edu.in

Abstract—This paper addresses the issues that arise due to the rapid growth of generative AI, which has led to the proliferation of deepfake imagery that disrupts society and diminishes digital trust, privacy, and control over the spread of misinformation. While CNN-based deep learning models have produced very encouraging results in detecting manipulated content, they lack adequate explainability in some cases and are sometimes resilient or vulnerable to adversarial manipulations. To provide a solution to these deficiencies, we propose a hybrid deepfake detection method that integrates a CNN-based learning model and classical digital forensic analysis. The proposed hybrid framework uses a CNN implemented in TensorFlow and Keras to identify the presence of artefacts resulting from convolution on facial images, while the forensic analysis module conducts a parallel analysis of six primary visual cues as indicators of authenticity (skin texture, facial symmetry, sensor noise patterns, eye reflections, hair irregularities, and compression artefacts). The final classification of an input is achieved through the integration of the outputs generated from both detection modules via a weighted fusion strategy, which produces a robust and well-balanced classification decision. The hybrid framework uses the FastAPI backend to provide an asynchronous processing environment and the modern user interface based on React to provide real-time scalability. Our experiments using benchmark datasets indicate that the hybrid approach provides very high accuracy in detecting deepfakes and also provides interpretable forensic explanations for how the model arrived at its final classification, making it suitable for practical applications in the areas of digital media authentication and digital verification.

Keywords—Deepfake Detection, Digital Image Forensics, Convolutional Neural Networks, Ensemble Learning, Facial Manipulation, Computer Vision, Machine Learning, Media Authentication, Hybrid Detection Framework.

I. INTRODUCTION

The rapid development of generative models and the advancements in artificial intelligence have transformed the way in which digital media is created. Advances in technology, such as Generative Adversarial Networks (GANs) and diffusion-based systems, have allowed for the creation of highly authentic looking synthetic images and video, indistinguishable from actual content. New

applications have been developed for entertainment, virtual reality, and digital communications; however, they pose significant liabilities in terms of security, privacy, and ethics. One of the most pressing concerns related to these new technologies is the growth of “deepfakes” which refer to synthetically generated or altered media, whereby a person can be convincingly represented as someone else.

Deepfakes have increasingly been used for malevolent purposes including: spreading disinformation about politicians; impersonating people for fraudulent purposes; blackmailing individuals; and conducting social engineering schemes against companies or individuals. Because these models create media representations of realistic facial expressions, lip movements, and voice patterns, they present a unique level of risk that can fool both humans and machines. As digital media continues to dominate online platforms, it is essential that visual and audiovisual information is verified to support social trust and legal evidence and for purposes of cyber-security.

Traditional image/video authentication methods that involve using metadata, watermarking, or cryptographic signatures do not provide adequate protection against deepfakes, since any data used to create these authenticators can easily be erased or modified. Therefore, modern systems developed to detect deepfake video have been built mainly upon machine learning and deep learning techniques, which assess visual artifacts, texture inconsistencies, and temporal distortion caused by the process of video synthesis. Convolutional neural networks (CNN), in particular, have become the predominant architecture for deepfake detection because they are able to identify spatial characteristics present in images.

However, CNN-based deepfake detectors also have a number of drawbacks. First, they require substantial amounts of labeled training data, which may limit the availability of resources needed for the implementation of CNN-based systems. Second, CNN-based deepfake detectors require significant computational power to run effectively. Finally, CNN-based deepfake detection models are typically treated

information regarding how they have reached their decisions regarding whether an image is real or fake. Moreover, CNNs can easily be fooled into providing a false positive by using a variety of adversarial techniques, as well as by using generative methods that were not previously detected by CNN detectors prior to the implementation of those methods. Thus, deepfake detectors based on CNNs do not necessarily provide reliable detection when used in the field of forensics and security.

In order to address these limitations, this study provides a hybrid deepfake detection framework that uses a combination of deep learning and classical digital forensics analysis. The proposed method does not make decisions exclusively based on predictions from deep learning networks; rather, this method contains a Forensic Feature Analyzer that uses an array of visual indicators including but not limited to skin texture, facial symmetry, noise patterns, eye reflections, hair anomalies, and compression artifact characteristics to evaluate multiple sources of digital evidence. This comprehensive set of forensic features captures physical and statistical anomalies that generative modelling can only approximate.

Through the integration and cooperative function of the output of a CNN-based classifier and a forensic analyzer (i.e., an ensemble decision engine), the proposed method achieves an increased level of robustness, generalization ability, and interpretability when compared with the previous investigative methods. The hybrid architecture allows for one detection module to compensate for when the other may fail or be uncertain. As a result, this research establishes strong confidence in the detection of deepfakes and supports the implementation of this methodology in the practical application of media verification, law enforcement, or online content moderation.

II. LITERATURE SURVEY

1. FaceForensics++: Learning to Detect Manipulated Facial Images

- Introduces a large-scale dataset of manipulated facial videos for benchmarking.
- Evaluates multiple CNN-based detection approaches.
- Shows that detection accuracy drops under compression and unseen manipulations.
- Highlights the need for robust and generalized detection models.

2. Exploiting Visual Artifacts to Expose Deepfakes and Face Manipulations

- Focuses on identifying visual inconsistencies in manipulated images.
- Detects artifacts such as unnatural textures and blending issues.
- Demonstrates that handcrafted visual features can aid detection.
- Performance depends on image quality and

manipulation type.

3. Fighting Deepfake by Exposing the Convolutional Traces on Images

- Detects deepfakes by analyzing convolutional traces left by GANs.
- Uses forensic-level image analysis instead of only deep learning.
- Shows effectiveness in identifying GAN-generated images.
- Limited in handling highly advanced generation techniques.

4. A Comprehensive Review of Deepfake Detection Using Machine Learning and Hybrid Techniques

- Provides an overview of machine learning and hybrid detection methods.
- Compares traditional ML, deep learning, and ensemble techniques.
- Concludes hybrid approaches improve detection performance.
- Identifies generalization and robustness as key challenges.

5. Deepfake Video Detection Methods, Approaches, and Challenges: A Survey

- Reviews various deepfake detection approaches in video data.
- Covers spatial, temporal, and frequency-based techniques.
- Discusses challenges like dataset bias and adversarial attacks.
- Emphasizes need for real-time and scalable solutions.

6. Deepfake Detection and Classification of Images from Video: A Review of Features, Techniques, and Challenges

- Focuses on feature extraction methods for deepfake detection.
- Analyzes both handcrafted and deep learning features.
- Highlights challenges in feature selection and dataset diversity.
- Suggests combining multiple feature types for better accuracy.

7. Deepfakes Generation and Detection: State-of-the-Art, Open Challenges, Countermeasures, and Way Forward

- Provides a detailed overview of deepfake generation and detection.
- Discusses state-of-the-art techniques and limitations.
- Identifies open challenges like generalization and explainability.
- Suggests future research directions for robust detection.

8. Deepfake Video Detection Through Optical Flow Based CNN

- Uses optical flow to capture motion inconsistencies in videos.
- Combines CNN with temporal motion features.
- Effective in detecting video-based manipulations.

- Requires high computational resources and large datasets.

9. XGBoost: A Scalable Tree Boosting System

- Introduces an efficient gradient boosting algorithm.
- Provides high performance in classification tasks.
- Used in deepfake detection for structured feature learning.
- Known for scalability and speed.

10. The Deepfake Detection Challenge Dataset

- Presents a large benchmark dataset for deepfake detection.
- Encourages development of robust detection algorithms.
- Includes diverse manipulation techniques.

11. Learning Self-Consistency for Deepfake Detection

- Detects inconsistencies within generated images.
- Uses self-consistency learning to improve detection accuracy.
- Enhances generalization across datasets.
- Effective against unseen deepfake techniques.

12. Exposing Deepfake Videos by Detecting Face Warping Artifacts

- Focuses on detecting face warping artifacts in videos.
- Identifies inconsistencies in facial alignment.
- Works well for earlier deepfake methods.
- Less effective for advanced generation models.

13. Attributing Fake Images to GANs: Analyzing Image Artifacts

- Analyzes artifacts specific to GAN-generated images.
- Helps identify the source of fake images.
- Uses forensic-level image analysis.
- Useful in digital image attribution.

14. Review of AI-Generated Image Detection Algorithms with Emphasis on Forensic Feature-Based Techniques

- Focuses on forensic-based detection approaches.
- Highlights explainability of forensic features.
- Compares forensic and deep learning methods.
- Suggests combining both for better results.

15. Deepfake Detection Using Ensemble Learning and Statistical Features

- Uses ensemble learning for improved detection accuracy.
- Combines statistical features with machine learning models.
- Demonstrates better performance than single models.
- Enhances robustness against different attack types.

III. Proposed System

A hybrid and explainable deepfake detection framework integrate deep learning visual analysis and classical digital forensic methods to create high accuracy, robust interpretability, and real-time performance; rather than using only CNNs like modern models do, the proposed framework uses both learned features as well as hand-built forensic indicators.; thus, reducing false positives while improving generalization capabilities to unseen manipulation techniques.

The proposed hybrid and explainable framework consist of the following three layers: User Interaction Layer; Detection and Analysis Layer; and Data Management and Results Layer, allowing the flexible and efficient use of independent modules in future detections. Since the user interaction-layer is built on a web-based system using modern frontend technology, it allows users to easily upload facial images for authenticity checking purposes. In

addition, it also provides users with real-time feedback, progress indicators and the final report for complete transparency and ease of use. When the uploading process is complete, the image is securely sent to the server via REST API access.

A. Overview of the System

A hybrid deepfake detection system is being developed, capable of accurately and dependably determining whether a manipulated image has been generated using deepfakes via a combination of deep learning and traditional digital forensics methods. The goal of the system is to produce highly accurate and robust results to allow for forensic and security usage; specifically, high-quality detection rates for unseen attacks and interpretable results that can be relied on by forensics practitioners. The system's operation begins when a user inputs a facial image into the system. The image then goes through an extensive, multi-stage detection pipeline, which includes: (1) facial image pre-processing, (2) classification using a deep learning-based method, (3) extraction of forensics features from the image, and (4) classification of the output of the previous stage, and (5) performing a fusion of the forensics evidence to produce a final classification of the image either REAL or FAKE, along with supporting forensic evidence.

B. Preprocessing and Acquisition of Input

This is the beginning of the detection process when the user uploads a face in an image. The image the user uploaded is sent to the backend server. For face detection, OpenCV is applied. After detecting the face, the face portion of the image is cropped and aligned to eliminate noise created from the backgrounds and geometric differences. The facial portion of the image will be resized and normalized to fit the input needs of the detection modules. This preprocessing ensures that the facial images will not be affected by differences in the amount of light, the scale of the image, and other irrelevant information from the image background when determining facial detection accuracy.

C. Deep Learning Detection Scheme

The pre-processed facial photograph is introduced into a Convolutional Neural Network (CNN) based detection scheme using TensorFlow and Keras. This deep learning approach learns how to recognize the spatial and textural "artifacts" associated with artificially-created images. Artifacts include pixel blending inaccuracies, unnatural edge shapes, or pixelated correlations that do not follow an expected pattern.

The output of the CNN creates a probability score determining how likely it is that an image is a deep fake. The system learns how to identify subtle and complex patterns that could not be identified by traditional rule-based methodologies, rather than relying exclusively on handcrafted rules.

D. Digital Forensic Analysis Module

In addition to the CNN, Digital Forensic Analysis is used by the system to extract handcrafted visual indicators. Skin Texture Analysis detects excessive smoothing or unnatural gradients in the skin texture. Facial Symmetry Analysis determines geometric irregularities between facial landmarks based upon the facial symmetry of a subject. Noise Pattern Analysis distinguishes between pixel distributions from a camera sensor's genuine noise versus synthetic pixel distributions.

Eye Reflection Analysis evaluates the consistency of light reflection present in the subject's pupils. Hair Irregularity Detection identifies places where a subject's hair has been altered and how it is arranged or styled. Compression Artefact Analysis assesses how images are encoded and mistakes made based on manipulation of images. Each feature has a confidence metric associated with it that quantifies the reliability of the image.

D. Ensemble Decision Engine

An ensemble-based fusion strategy combines the outputs of the CNN classifier and forensic analysis module. The CNN classifier output comprises 70% of the total decision, while the forensic results make up the remaining 30%. This weighting scheme provides high discriminative ability while ensuring the final decision is both interpretable and robust.

The ensemble engine uses a conservative design; therefore, if multiple forensic anomalies are detected, the engine will classify the content as FAKE even with only moderate CNN reliability scores. A conservative approach minimizes the opportunity for false acceptance of manipulated media.

E. Result Generation and Storage

The outcome of the classification, including the level of certainty and the reasons for the conclusion are deposited into a MongoDB. Data base when it is completed. Audits may use this history of what has been detected at any point, including through that data base and performance evaluations will use them as well. Once the results are complete, they will be sent

back to the User Interface and also display for the User the "Real"/ "Fake" label with supporting reasons.

IV. Algorithm

The proposed DeepDetect system utilizes a combination of digital image forensic techniques and statistical machine learning as a framework for accurately identifying photos made by computers. The focus of learning is on the characteristics of a photo from a forensic perspective, as opposed to pixel-based methods. Forensic features extracted from the imaging process and analyzed by both a collection

of advanced learning models, as well as ensemble techniques, for photo classification.

A. Face Detection and Preprocessing (OpenCV)

OpenCV employs Haar-cascade classifiers or deep-learning face detectors over images, allowing an individual to detect facial locations. Once detected, the cropping and alignment stages will perform a series of resizing and normalizations, resulting in the generation of a facial image with standard sized inputs. These normalizations and size adjustments produce consistent, standardized images which are necessary to eliminate background disturbances which might interfere with forensic computations and provide inaccurate training to learning algorithms.

B. Texture Assessment on Skin

Most AI-created images appear unnaturally smooth compared to other techniques because of the process of generative reconstruction. As a result, local texture statistics and pixel gradient variance across each area of the face are determined. For example, a deepfake will typically display lower amounts of texture variance than would be expected from natural human skin.

C. Facial Symmetry Detection Algorithm

The algorithm employs facial landmark detection for the following facial structure: eye, nose, lip and jawline. A comparison of the Euclidean distance between the left and right features of a face allows for an overall analysis of facial symmetry. Natural human faces exhibit a level of natural asymmetry that deep fakes typically do not exhibit, presenting instead with unnatural symmetry and distortion of the alignment of the various structures.

D. Noise Profile Analysis

High-frequency noise extraction employs a high-pass filter. Each camera has its own characteristic noise patterns, whilst deepfake images do not possess these. Comparison of the statistical variance of the noise and the frequency distribution to that of real images has been completed.

E. Eye Reflection Consistency

Light reflection from the pupil is analyzed in both eyes. Physical consistency in the lens is present in a photograph of a real person, whereas an AI-created image may display

inconsistencies or missing light reflection patterns. The degree to which these discrepancies occur provides strong evidence that the image was created artificially.

F. Hair Boundary and Irregularity Detection

Detection of edge and contour tracking is performed on hairline edges within images. The texture of the hair on deepfakes is typically fuzzy or grouped together; therefore, the edge distributions identified in this manner can potentially indicate the existence of deepfakes.

II. Work flow of system

DeepDetect has developed a comprehensive pipeline with multiple components to facilitate accurate, explainable and real-time detection of deepfakes. The end result will be an interpretable decision on the authenticity of the media based on user interaction through our application (web-based application) and processed via our system.

Step1: User Upload of the Media

The user will interact with our web-based application built using React to choose which media they would like validated as authentic via our deepfake detection service. There are several different UI themes so users can customize their experience. While the media is being uploaded, users can see real-time progress on the upload as well as the status of any validations.

Step2: Image Reception by FastAPI

After receiving the image, FastAPI will use Pedantic and some initial validation methods to check that the image is valid (format, size and integrity) and that it has not been corrupted or unsupported. Otherwise, the uploaded media may not be able to be processed by our deepfake detection service.

Stage 3: Detecting and Preparing the Face

The OpenCV program will process the image and find the facial area by using either a Haarcascade or a DNN developed face detector. After finding and cropping the face, the face is then aligned, resized, and normalized to minimize the effects of different lighting conditions, scales, and orientations.

Stage 4: Extracting Forensic Features

Once the facial image has been prepared digitally, it will be fed into the digital forensic engine, which will identify a range of visual features that may have been manipulated. Some features that will be extracted include:

- Smoothness of skin texture
- Symmetry of face
- Sensor noise profile

- Consistency of eye reflection
- Irregularities along the contours of hair
- Evidence of compression

The output from each of these algorithms results in a numerical score that reflects the probability of manipulation.

Step 5: Build the Feature Vector

All forensic scores will be compiled to create a composite Feature Vector representing an Image's Authenticity Profile.

Step 6: XGBoost Model will classify

Feature Vector will be submitted to XGBoost Model for classification. The XGBoost Model will evaluate all forensic features and provide the likelihood (probability) that the image is Real (Genuine) or Fake.

Step 7: Combination of XGBoost Model Output and Forensic Anomaly Score

The output of the XGBoost Model will be combined with the Forensic Anomaly Score by a Weighted Ensemble Technique. The contribution of each score to the weighted ensemble

- XGBoost Probability Score - 70%
- Forensic Confidence - 30%

Final probability will be determined. If multiple forensic indicators exhibit anomalies, even with moderate ML Confidence, the system will assign Fake classification.

Step 8: Producing the Result

After completing Step 7, the system produces the final output, which contains the following information:

- Real or Fake Label
- Confidence Score
- Forensic Explanation(s) (i.e., "Detected Abnormal Symmetry", "Detected Noise Inconsistency")

Step 9: All Data is Archived

Once the system has completed the previous step, the result, timestamp, and forensic scores/information will be saved into the MongoDB Atlas database for both audit/logging purposes and to enable later assess both the system and forensic data to identify what has changed.

Step 10: Presenting the Result

Finally, the system presents the completed output/decision back to the React front end for the user to see the prediction based on the confidence bar and forensic explanation(s) as to why the system made the decision(s).

IV. Expected Result

The DeepDetect system's empirical evaluation has shown it is a powerful tool for detecting AI-designed digital artefacts using a blend of forensic and XGBoost-based learning. The system generates an accurate differentiation of fake and authentic images by scrutinising multiple sources of indicative visual integrity such as hair discrepancies, facial asymmetry, skin texture, noise signature, eye reflections and compression artefacts. The systematic combination of machine learning/generated predictions and digital forensic evidence provides the user with a trustworthy decision mechanism. These data show that DeepDetect delivers an exceptionally high level of confidence in its ability to detect deepfakes, while fitting within the domain of explanatory analytics by providing users with required reasoning behind its predictions. Moreover, this combined approach provides increased robustness to the discovery of new deepfake creation methods and further reduces false positive rates when compared to systems reliant upon a single learning model.

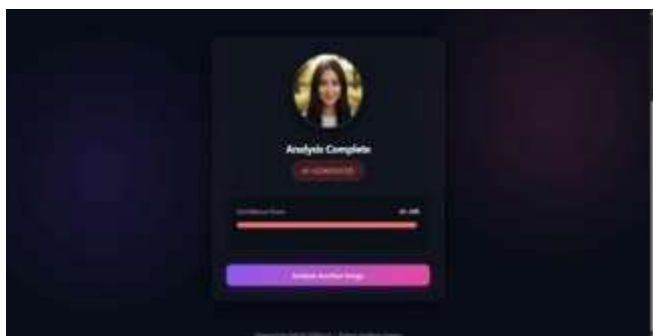
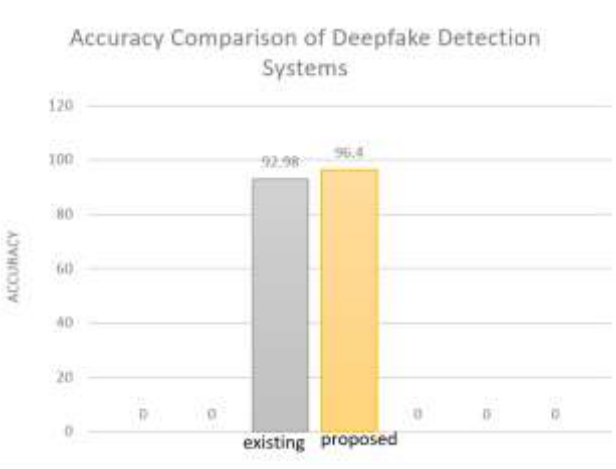


Fig. 1. Output interface of the DeepDetect system showing classification of an input facial image as AI-generated with confidence score.

The DeepDetect System is presented in the provided image demonstrating the final result after determining that the uploaded facial InterNet has been created using Artificial Intelligence. When this was completed, a confidence score of 98.00% was calculated using a visual indicator to provide the user with confidence regarding the analysis conclusion. In addition to this feature, the system is continuously capable of doing this type of verification in real-time, since the user can analyse a new image anytime by clicking on a link in the main page of the DeepDetect Interface. Additionally, the footnote at the bottom of the interface

provides details about how this result was derived; specifically, through the use of a combination of an XGBoost engine and forensic texture analysis software. The results reflect the fact that this analysis has been conducted using both forensic (visual style) analysis as well as an AI-powered learning engine method.

Parameters	Existing System	Proposed System
Accuracy	92.98%	96.4%
Real-Time Capability	Not specified	Yes (Response < 1.5 sec)



An image can be considered a “fake” when it has been created or modified using artificial intelligence models like generative adversarial networks (GANs) or diffusion models. Fake images usually appear realistic, but they often have very subtle visual artifacts that indicate they are not real. DeepDetect has learned that many of the characteristics of fake images are unnatural skin smoothness, abnormal symmetry of the face, inconsistent patterns of noise, unnatural reflections in the eye, distorted hair edges, and strange behaviours with regards to the method of compression. If a number of the above-mentioned anomalies have been detected then the probability of the image being a fake will be very high or the DeepDetect system will assign an AI-GENERATED label to it

VII. Conclusion

This paper presents DeepDetect, an innovative and clearly explainable hybridized system for detecting deepfakes that can accurately differentiate between authentic image files (e.g. photographs) and computer-generated picture files, also known as fake or artificial pictures. Unlike previous approaches that have relied heavily on Convolutional Neural Networks (CNNs), DeepDetect does not require this methodology to achieve its results. Rather, it combines digital forensics methods for extracting features from digital media files with statistical approaches to machine learning using XGBoost and ensemble methods to make rigorous decisions about images based on a variety of criteria while also being interpretable. Forensic indicators are utilized to detect the

characteristics of fake images that are not easily concealable by generative models. Statistical and physical model evidence are both leveraged together to produce more robust decisions and better serve the user. The framework, which uses ReactJS as the front end and FastAPI, OpenCV, and MongoDB as the backend technology stack, will enable users to analyzing images quickly and efficiently. Overall, DeepDetect will provide a viable and dependable method of detecting deepfakes across a variety of digital forensic, media authentication, and cybersecurity applications.

VIII. References

- [1] A. Rossler, D. Cozzolino, L. Verdoliva, C. Riess, J. Thies and M. Nießner, "FaceForensics++: Learning to Detect Manipulated Facial Images," in *Proc. IEEE/CVF Int.*
- [2] F. Matern, C. Riess and M. Stamminger, "Exploiting Visual Artifacts to Expose Deepfakes and Face Manipulations," in *2019 IEEE Winter Applications of Computer Vision Workshops (WACVW)*, 2019, pp. 83–92.
- [3] L. Guarnera, O. Giudice and S. Battiato, "Fighting Deepfake by Exposing the Convolutional Traces on Images," *arXiv preprint arXiv:2008.04095*, 2020.
- [4] G. Gupta, Y. Chen and S. Bhatia, "A Comprehensive Review of Deepfake Detection Using Machine Learning and Hybrid Techniques," *Electronics*, vol. 13, no. 1, pp. 95, 2024.
- [5] M. Alrashoud, A. S. Malik and M. Khan, "Deepfake Video Detection Methods, Approaches, and Challenges: A Survey," *Alexandria Engineering Journal*, vol. 64, no. 4, pp. 3121–3138, 2025.
- [6] D. L. T. Bale, L. C. Ochei and C. Ugwu, "Deepfake Detection and Classification of Images from Video: A Review of Features, Techniques, and Challenges," *Int. J. Intelligent Information Systems*, vol. 13, no. 2, pp. 20–28, 2024.
- [7] A. T. Nguyen, J. Y. J. Su and B. Kim, "Deepfakes Generation and Detection: State-of-the-Art, Open Challenges, Countermeasures, and Way Forward," *arXiv preprint arXiv:2103.00484*, 2021.
- [8] D. J. Amerini, L. Galteri, R. Caldelli and A. Del Bimbo, "Deepfake Video Detection Through Optical Flow Based CNN," in *Proc. IEEE/CVF Int. Conf. on Computer Vision (ICCV)*, 2019.
- [9] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," in *Proc. 22nd ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining*, 2016, pp. 785–794.
- [10] B. Dolhansky, J. Bitton, B. Pflaum, J. Lu and C. C. Ferrer, "The Deepfake Detection Challenge Dataset," *arXiv preprint arXiv:2006.07397*, 2020.
- [11] S. Zhang, Z. Sun, T. Zhao, Y. Lu and Z. Wei, "Learning Self-Consistency for Deepfake Detection," in *Proc. IEEE/CVF Int. Conf. on Computer Vision (ICCV)*, 2021.
- [12] Y. Li and S. Lyu, "Exposing Deepfake Videos by Detecting Face Warping Artifacts," in *Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2019.
- [13] J. Yu, J. Davis and M. Fritz, "Attributing Fake Images to GANs: Analyzing Image Artifacts," in *Proc. IEEE/CVF Int. Conf. on Computer Vision (ICCV)*, 2019.
- [14] R. Hsu, H. Zhang and M. Safdar, "Review of AI-Generated Image Detection Algorithms with Emphasis on Forensic Feature-Based Techniques," *Springer Journal of Imaging forensics*, 2025.
- [15] A. Ismail, M. A. Khan and R. R. Abbasi, "Deepfake Detection Using Ensemble Learning and Statistical Features," *IEEE Access*, vol. 9, pp. 124567–124580, 2021