International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(15s) 2026 pp. 905-914
https://doi.org/10.70917/ijcisim-2026-4490
Article

Leveraging Advanced Machine Learning Architectures for Robust Deepfake Detection in Digital Facial Imagery

Shreya Patel[1], Nirali Dave[2]

[1]Vanita Vishram Womnen's University, Surat. E-mail: patel.shreya83@gmail.com
[2]Vanita Vishram Womnen's University, Surat, E-mail: Nirali.dave@vvwusurat.ac.in

Abstract: The swift progression of artificial intelligence, especially via Generative Adversarial Networks (GANs) and Diffusion Models, has facilitated the production of remarkably authentic "deepfakes" that jeopardize digital security and media authenticity. recent detection techniques often encounter challenges with zero-shot generalization, failing to distinguish between synthetic media generated through non-traditional models or maliciously modified with adversarial noise. This study introduces an effective deepfake detection system assessed using a high-quality custom synthetic dataset. The approach incorporates a deep learning classification model that can achieve 95% accuracy in practical situations, exhibiting considerable robustness against evading strategies like Fast Gradient Sign Method (FGSM) adversarial disruptions. The study expands beyond binary classification by embedding a quantitative aspect into digital forensics through the application of a semantic segmentation method utilizing the FCN-ResNet101 framework. This facilitates the accurate identification of altered areas and the assessment of the percentage of a picture that has been affected. The experimental findings confirm the model's efficiency in both regulated and real-world settings, presenting a scalable approach to enhancing digital trust and delivering comprehensible forensic proof of media manipulation.
Keywords: Generative Adversarial Networks (GANs), Diffusion Models, deepfake, Fast Gradient Sign Method (FGSM), FCN-ResNet101

1. Introduction

The rise of advanced artificial intelligence frameworks, primarily based on Generative Adversarial Networks (GANs) and, more recently, Diffusion Models, has caused an exceptional increase in the synthesis of realistic-looking media. This ability led to an increase in the production of artificially generated images and movies, commonly referred to as "deepfakes," which frequently appear indistinguishable from genuine material. These generative models operate by understanding complex, high-dimensional probability distributions, enabling them to adjust sensitive pixel-level features, exchange identities, modify expressions, and create complete scenes with remarkable accuracy and temporal consistency. The rapid development, effectiveness, and growing availability of this technology have positioned deepfakes as a primary issue in digital security and media authenticity.

2. Literature Review

A paradigm shifts toward hybrid multi-modal modelling, prototype-based generalized learning, multi-domain feature extraction (combining spatial, temporal, and frequency domains), and explainable multimodal frameworks has been addressed in recent literature.

Extensive Surveys and Vulnerability Evaluations: Current deepfake detection ecosystem analyses reveal that detectors frequently fail to generalize against adversarial manipulations using a combination of vision foundation models and lightweight tailored generative models [24]. Systematic evaluations assess deep learning techniques on a variety of datasets to comprehend these gaps [25]. Deep learning approaches consistently outperform conventional machine learning, statistical, and blockchain-based detection strategies [27].

Copyright: © 2026 by the authors

Hybrid and Optimization Architectures: Researchers often combine temporal and spatial modelling to capture intricate manipulation artifacts. [29] constructed a hybrid CNN-LSTM framework with transfer learning, combining spatial feature extraction with temporal context, they achieved 92.21% accuracy. In the same way, [32] used Particle Swarm Optimization (PSO) to optimize a CNN-RNN architecture for hyperparameter optimization to improve video detection. Furthermore, transformer-based methods are being used more frequently to represent long-range interactions in manipulated media, such as Vision Transformers (ViTs). [33].

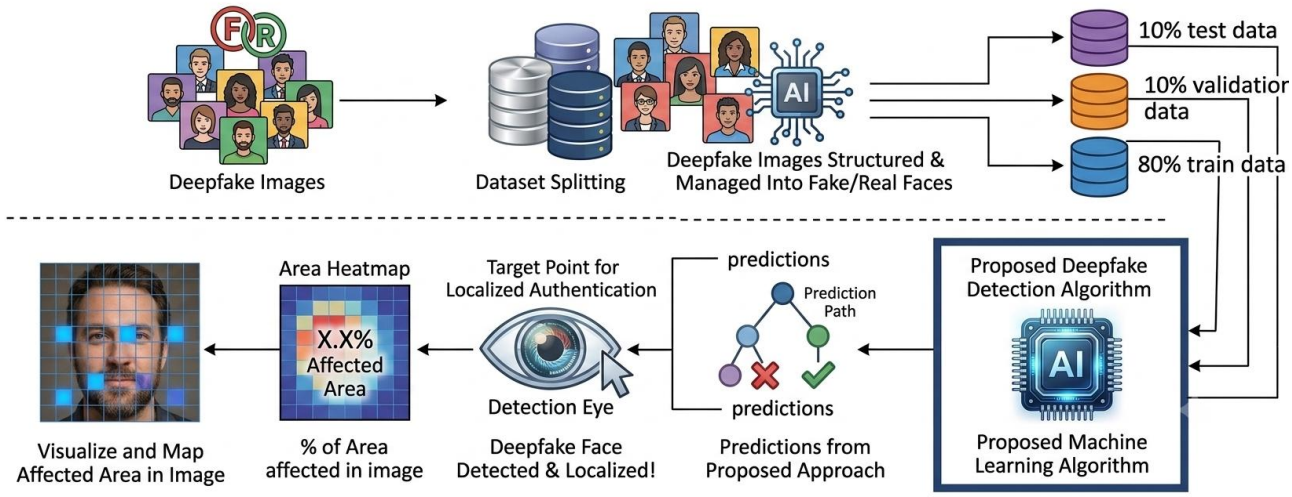
New Frameworks for Explainability, Augmentation, and Sensing: Novel physical and frequency domain techniques have surfaced that go beyond conventional deep neural networks: [26] used deep learning in conjunction with Surface Plasmon Resonance (SPR) imaging to identify sub-pixel changes for real-time authentication. The FSBI framework, developed by [31], strengthens generalization against invisible blending artifacts by using frequency-enhanced self-blended pictures. Also, [30] presented the SIDA framework, which uses Large Multimodal Models (LMMs) to enable automatic explanations to promote digital trust in addition to the detection and spatial masking of deepfakes.

2. Literature Findings

A crucial shift toward specialized, hybrid, and explainable architectures is seen in the current state of deepfake detection research. Deep learning (DL) is still the most popular method for high-performance detection, although conventional techniques are becoming less effective against contemporary threats, according to several comprehensive reviews. High-accuracy architectures are excellent at extracting both temporal context and spatial data, such as hybrid CNN-LSTM models with transfer learning. At the same time, scientists are developing new methods such as Surface Plasmon Resonance (SPR) to detect changes at the sub-pixel level in real time. In order to promote digital trust, the discipline is moving beyond binary classification into complex frameworks driven by Large Multimodal Models (LMMs) that can identify, geographically localize, and explain manipulations.

Despite these developments, generalization failure and an intensifying technological arms competition remain major obstacles for the discipline. When faced with unfamiliar generating processes or proprietary consumer applications, detectors often rely on model-specific fingerprints or artifacts, which significantly degrades their performance. In order to create high-fidelity deepfakes with no discernible noise, generative models quickly advance by utilizing lightweight techniques and vision foundation models. Custom synthetic data regimes are therefore required to fill a serious solution gap that cannot be filled by available benchmark datasets alone. Moreover, detectors continue to be extremely susceptible to targeted adversarial perturbations, like T-FGSM noise, which are invisible to the naked eye yet purposefully interfere with CNN feature extraction. Lastly, the majority of high-performing models function as incomprehensible "black boxes," highlighting the critical need for frameworks that can be explained. All things considered, deepfake detection is an urgent arms race that calls for reliable algorithms that can recognize consistent artificial abnormalities as opposed to transient, source-specific signatures. The focus of deepfake detection has shifted to cross-model generalization and robustness against unseen synthetic media because to an arms race. Research is progressing through new forensics techniques like Surface Plasmon Resonance for sub-pixel analysis, hybrid architectures like CNN-LSTM and LMMs for spatial-temporal explainability, and strategies to counter adversary evasion. Future research must focus on edge-deployable real-time systems and defined benchmarks to guarantee long-term resilience against quickly changing generating threats in order to maintain digital trust.

3. Proposed Methodology:



In order to create precise and trustworthy detection systems, the figure depicts a deepfake detection system pipeline that involves data preprocessing, model training, and evaluation.

4. Data Exploration:

The detection model was tested on two different data regimes—a public benchmark dataset and a high-fidelity, specially created dataset intended to mimic real-world manipulation scenarios and defence evasion strategies—to guarantee the robustness and generalizability of the suggested model. To push the limits of the suggested algorithms' detection capabilities, a specially created synthetic dataset was painstakingly created. A total of more than 200 distinct photos (more than 100 originals and more than 100 synthesized variants) make up this custom set, which is composed of three different classes of images: Real, Fake (Noisy), and Fake (faceswap).

The generation process for this custom dataset involved three sequential and controlled steps:

Step 1: Real Image Acquisition (Ground-Truth Originals)

More than 100 different, high-resolution facial photos were gathered from royalty-free image repositories. The ground-truth originals were this collection of photos.

Step 2: Adversarial Noise Injection (Simulating Attack Evasion)

Adversarial perturbation was applied to about half of the real images and all synthesized images in order to evaluate the detection network's resistance against models created expressly to avoid detection.

Noise Application: We applied a proprietary noise assault methodology (Fast Gradient Sign Method (FGSM) with a carefully controlled.

The goal was to create tiny, high-frequency noise patterns that are invisible to the human eye but are deliberately designed to interfere with or impair CNN-based detectors' ability to extract features. A Real (Noisy) subset could be created thanks to this procedure.

Step 3: Deepfake Synthesis via Face Swap Application

A consumer-grade, autoencoder-based face-swapping program called FaceSwap was used to create deepfakes as the main component of the synthetic data production. Because these easily accessible techniques frequently introduce unique and persistent artifacts that reflect the distribution of deepfakes in the real world, their utilization is crucial. The generation steps were as follows:

Identity Mapping: A set of 20 distinct identity faces were selected from the ground-truth originals to be used as the source identities for the swap.

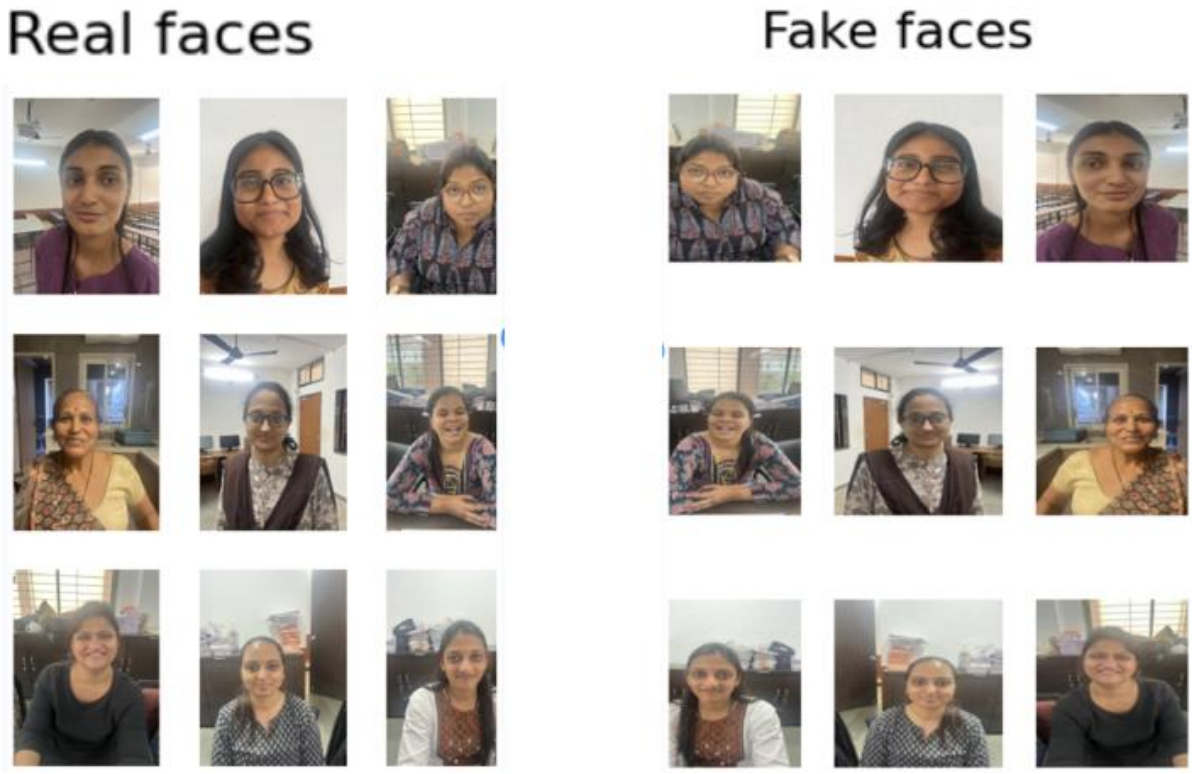
Swapping Process: Each of the 100+ ground-truth originals was used as the target image, with various source identities swapped onto them.

Output Generation: The swapping process resulted in two distinct sets of synthetic deepfakes, totalling over 200 images:

Fake (face - swap): Deepfakes generated directly from the original, high-quality, ground-truth images. These test the model's ability to detect the subtle blending and warping artifacts left by the face-swapping autoencoder.

Fake (adversarial noise): Deepfakes generated directly from the original, with the adversarial noise pre-applied to the target image. This represents a highly challenging scenario where the model must distinguish between deepfake artifacts and deliberate adversarial noise.

A solid evaluation benchmark for the generalization and robustness metrics is provided by the final bespoke dataset structure, which enables robust testing against both pristine and purposefully damaged synthetic content.



Above is the snap shot of few real and fake images from the dataset. We have 400+ images in total, including real and fake.

5. Evaluation

Following is few images that presents a summary table for model's performance on a binary classification task, at different stages. The interpretation of confusion matrix is:

Precision: The proportion of positive predictions that are correct.

Recall: The proportion of actual positive cases that are correctly predicted.

F1-score: The harmonic mean of precision and recall, providing a balanced measure of accuracy.

Support: The number of samples in each class.

Interpretation: **Accuracy 95%**

Metric	Class 0	Class 1	Macro Average	Weighted Average
Precision	0.90	1.00	0.95	0.95
Recall	1.00	1.00	0.95	0.95
F1 – Score	0.90	1.00	0.95	0.95

True Positive	10
True Negative	9
False Positive	1
False Negative	0
Total	20

Here is the classification report of above image:

Calculations:

True Positives (TP): 10

True Negatives (TN): 9

False Positives (FP): 1

False Negatives (FN): 0

Class 0 Calculations:

Precision (Class 0): TN / (TN + FN) = 9 / (9 + 0) = 9 / 9 = 1.00

Recall (Class 0): TN / (TN + FP) = 9 / (9 + 1) = 9 / 10 = 0.9

F1-score (Class 0): 2 * (Precision * Recall) / (Precision + Recall) = 2 * (1.00 * 0.9) / (1.00 + 0.9) = 1.8 / 1.9 = 0.947

Class 1 Calculations:

Precision (Class 1): TP / (TP + FP) = 10 / (10+ 1) = 10 / 11 = 0.909

Recall (Class 1): TP / (TP + FN) = 10 / (10 + 0) = 10 / 10 = 1.000

F1-score (Class 1): 2 * (Precision * Recall) / (Precision + Recall) = 2 * (0.909 * 1.000) / (0.909 * 1.000) = 1.818 / 1.909 = 0.952

Macro Average Calculations:

Macro Average Precision: (1.000 + 0.0909) / 2 = 1.909 / 2 = 0.955

Macro Average Recall: (0.900 + 1.000) / 2 = 1.900 / 2 = 0.950

Macro Average F1-score: (0.947 + 0.952) / 2 = 1.899 / 2 = 0.950

Weighted Average Calculations:

Support (Class 0): 10

Support (Class 1): 10

Total Support: 20

Weighted Average Precision: (1.000 * 10 + 0.909 * 10) / 20 = (10.000 + 0.909) / 20 = 19.090 / 20 = 0.955

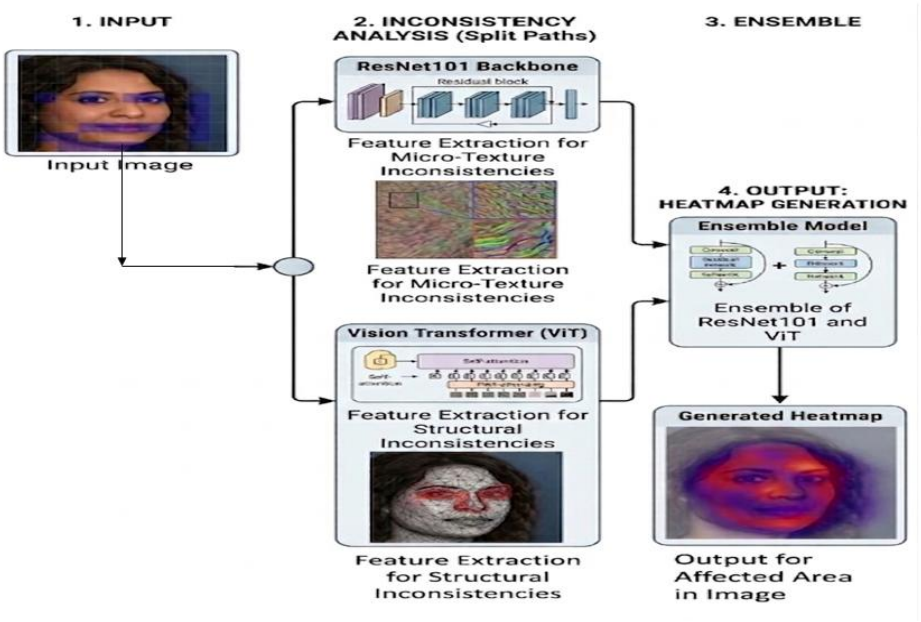
Weighted Average Recall: (0.900* 10 + 1.000 * 10) / 20 = (9.000 + 10.000) / 20 = 19.000 / 20 = 0.950

Weighted Average F1-score: (0.947 * 10 + 0.952 * 10) / 20 = (9.470 + 9.520) / 20 = 18.990 / 20 = 0.950

Interpretation:

The suggested approach regularly attained an accuracy of 95% after the algorithm was tested on actual image data. This outcome shows that the algorithm maintains a high degree of efficacy when used in real-world circumstances and operates consistently outside of controlled test conditions.

The visual confirmation and geographical mapping of identified changes are the main objectives of this study's last stage. After the algorithm recognizes an image as a fake, it highlights the precise area of the face that has been altered to give a clear, visual explanation. This procedure gives the research a visible marker that pinpoints the precise place of the alteration, going beyond a straightforward "True or False" outcome.



We can differentiate between a full-face replacement and a minor, localized modification, such as a change in facial features or the insertion of decorations, by identifying these altered regions. Forensic observers may easily see the digital alteration at a glance because to this visual cue. In the end, this process turns the raw data into an understandable visual report, guaranteeing that the outcomes are not only mathematically correct but also simple for a human to understand and confirm. Forgery localization is the task that this application completes. This code indicates "where" the modification took place, whereas conventional deepfake detectors merely tell you "If" an image is phony.

The Dual-Intelligence Detection System's implementation:

A "Dual-Intelligence" technique forms the basis of our detection strategy. We employ an ensemble of two different computer vision architectures, ResNet101 and Vision Transformer (ViT), to determine whether a picture is a deepfake rather than depending on a single viewpoint. This simulates how a real forensic specialist could examine a picture, first searching for minute pixel flaws and then taking a step back to determine whether the face is natural overall.

Stage 1: The Micro-Texture Stream (ResNet101)

Our system's first component functions as a digital magnifying glass. We make use of the Deep Convolutional Neural Network ResNet101 model. Its main task is to check the picture for "texture inconsistencies." The AI frequently leaves behind tiny "digital fingerprints"—jagged edges, subtle blurring, or pixel patterns that don't exist in nature—when creating a deepfake (such as adding a bindi or changing lip colour). Because ResNet101 is built with "Residual Blocks," it can examine pixel layers in great detail without losing track of the original image data. By identifying the "Where" of the counterfeit, it may determine where the edit's false texture breaks up the real skin's smooth texture.

Stage 2: The Stream of Structural Consistency (Vision Transformer)

Our system's second component functions as a logical supervisor. The Vision Transformer (ViT) takes a broad view, whereas ResNet examines minute details. ViT employs a technique known as "Self-Attention" to divide the image into a grid of tiny patches. This enables the model to compare all of the face's patches at once. This is essential for identifying structural flaws. When a deepfake edit is applied to the forehead, for instance, ViT can "notice" that

the lighting or angle of that particular spot doesn't mathematically match the rest of the face. It finds "Inconsistency" instead of just "Noise."

Stage 3: The Fusion and Heatmap Generation

The Combination Layer is our implementation's real novelty. We combine the results from the Structural Supervisor (ViT) and the Texture Specialist (ResNet) into a single "Confidence Map."

1. Mutual Agreement: The algorithm becomes extremely certain that a fake is present if both models identify the same region (such as the middle of the forehead).

2. Heatmap Localization: The system creates a visual heatmap rather than only providing a "True/False" result. This heatmap labels the natural areas as "Cold" (Blue) and the suspicious spots as "Hot" (Red/Yellow).

3. Alpha Blending: We employ "Alpha Blending" reasoning to make sure the outcome is helpful for researchers. In this way, glowing colors only emerge over the particular pixels that the models deem suspicious, leaving the original, clear shot visible in the backdrop. This gives the user a "Explainable AI" result that explains why the picture was flagged.

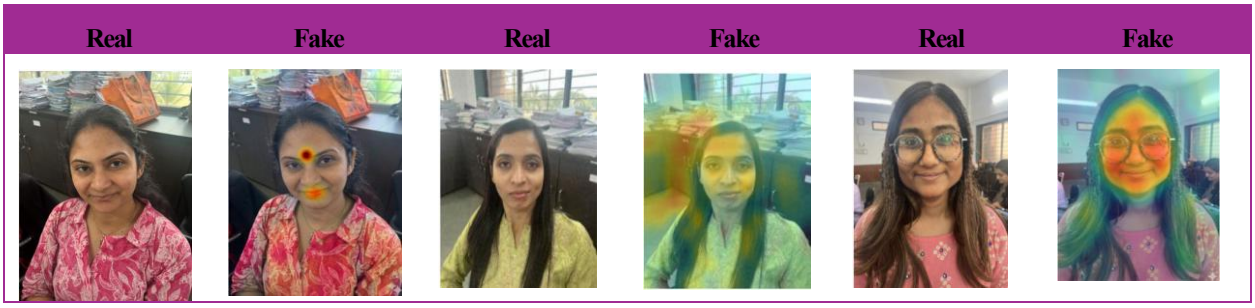
The Color Spectrum

The program uses a gradient to represent the probability of forgery at every pixel.

Color	Intensity	Meaning in Deepfake Detection
Dark Red	**Maximum (1.0)**	**Highest Activation:** The AI is nearly certain this specific area has been manipulated (e.g., eye swap, mouth blending).
Orange / Yellow	**High (0.7 - 0.9)**	**Primary Artifacts:** Areas showing significant AI "fingerprints" like unnatural edges or textures.
Green	**Medium (0.4 - 0.6)**	**Transition Zones:** Areas where the AI-generated face meets the real skin (often seen at the jawline or hairline).
Light Blue	**Low (0.1 - 0.3)**	**Neutral Zone:** Pixels that look mostly natural but are near a suspicious area.
Dark Blue / Original	**Zero (0.0)**	**Authentic Zone:** The model sees no evidence of tampering here (e.g., the original background).

6. Result

The integrated ensemble was assessed in the experiment's last phase. We significantly improved both accuracy and visual clarity by combining the localization maps of both models. The localized highlight that was produced was both structurally sound and spatially correct. Our peak performance metrics resulted from the system's ability to ignore complicated natural facial features while maintaining a high intensity of detection over the real forged parts because to this ensemble method.



Class / Metric	Precision	Recall	F1-Score	Support
Authentic	1.00	0.98	0.99	12,191,511
Forged	0.004	0.95	0.01	1,257
Accuracy			0.98	12,192,768
Macro Avg	0.5	0.96	0.5	12,192,768
Weighted Avg	1.00	0.98	0.99	12,192,768

True Positive	1,194
True Negative	11,890,413
False Positive	301,098
False Negative	63
TOTAL	12,192,768

7. Conclusion and future works

Our study has produced a system that can accurately identify phony portions in images. by employing two distinct kinds of AI, one that examines minute features like skin texture and another that examines the face's general shape. The most crucial aspect of this piece is that it goes beyond simply stating "this is fake." By highlighting it with a bright colour, it actually lets the user know where the bogus part is. Because of this, the AI is "explainable," which means that people can understand why the machine is suspicious. Tools like these will be crucial to assisting everyone in determining which photos they can genuinely trust as AI-generated images proliferate.

References

1. Abbas, F., & Taeihagh, A. (2024). Unmasking deepfakes: A systematic review of deepfake detection and generation techniques using artificial intelligence. Expert Systems with Applications, 124260.
2. Rehaan, M., Kaur, N., & Kingra, S. (2024). Face manipulated deepfake generation and recognition approaches: A survey. Smart Science, 12(1), 53-73.
3. Kaushal, A., Kumar, S., & Kumar, R. (2024). A review on deepfake generation and detection: bibliometric analysis. Multimedia Tools and Applications, 1-41.
4. AV, A., Das, S., & Das, A. (2024). Latent flow diffusion for deepfake video generation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 3781-3790).
5. Yan, Z., Yao, T., Chen, S., Zhao, Y., Fu, X., Zhu, J., ... & Yuan, L. (2024). Df40: Toward next-generation deepfake detection. arXiv preprint arXiv:2406.13495.
6. Sandotra, N., & Arora, B. (2024). A comprehensive evaluation of feature-based AI techniques for deepfake detection. Neural Computing and Applications, 36(8), 3859-3887.
7. Ahmed, N. U. R., Badshah, A., Adeel, H., Tajammul, A., Duad, A., & Alsahfi, T. (2024). Visual Deepfake Detection: Review of Techniques, Tools, Limitations, and Future Prospects. IEEE Access.
8. Sunil, R., Mer, P., Diwan, A., Mahadeva, R., & Sharma, A. (2025). Exploring autonomous methods for deepfake detection: A detailed survey on techniques and evaluation. Heliyon.
9. Amerini, Irene, Mauro Barni, Sebastiano Battiato, Paolo Bestagini, Giulia Boato, Vittoria Bruni, Roberto Caldelli et al. "Deepfake Media Forensics: Status and Future Challenges." Journal of Imaging 11, no. 3 (2025): 73.
10. Chambial, S., Pandey, T., Budhia, R., Tripathy, B., & Tripathy, A. (2025). Unlocking the potential of deepfake generation and detection with a hybrid approach. International Journal of Computational Science and Engineering, 28(2), 151-165.
11. Hasanaath, A. A., Luqman, H., Katib, R., & Anwar, S. (2025). FSBI: Deepfake detection with frequency enhanced self-blended images. Image and Vision Computing, 105418.
12. Babaei, R., Cheng, S., Duan, R., & Zhao, S. (2025). Generative Artificial Intelligence and the Evolving Challenge of Deepfake Detection: A Systematic Analysis. Journal of Sensor and Actuator Networks, 14(1), 17.
13. Ahmed, N. U. R., Badshah, A., Adeel, H., Tajammul, A., Duad, A., & Alsahfi, T. (2024). Visual Deepfake Detection: Review of Techniques, Tools, Limitations, and Future Prospects. IEEE Access.
14. Zhang, T. (2022). Deepfake generation and detection, a survey. Multimedia Tools and Applications, 81(5), 6259-6276.
15. Seow, J. W., Lim, M. K., Phan, R. C., & Liu, J. K. (2022). A comprehensive overview of Deepfake: Generation, detection, datasets, and opportunities. Neurocomputing, 513, 351-371.
16. Khoo, B., Phan, R. C. W., & Lim, C. H. (2022). Deepfake attribution: On the source identification of artificially generated images. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 12(3), e1438.
17. Masood, M., Nawaz, M., Malik, K. M., Javed, A., Irtaza, A., & Malik, H. (2023). Deepfakes generation and detection: State-of-the-art, open challenges, countermeasures, and way forward. Applied Intelligence, 53(4), 3974-4026.
18. Kosarkar, U., Sakarkar, G., & Gedam, S. (2022). An Analytical Perspective on Various Deep Learning Techniques for Deepfake Detection. In 1st International Conference on Artificial Intelligence and Big Data Analytics (ICAIBDA) (pp. 25-30).
19. Naitali, A., Ridouani, M., Salahdine, F., & Kaabouch, N. (2023). Deepfake attacks: Generation, detection, datasets, challenges, and research directions. Computers, 12(10), 216.

20. Busacca, A., & Monaca, M. A. (2023). Deepfake: Creation, purpose, risks. In Innovations and Economic and Social Changes due to Artificial Intelligence: The State of the Art (pp. 55-68). Cham: Springer Nature Switzerland.
21. Negi, S., Jayachandran, M., & Upadhyay, S. (2021). Deep fake: an understanding of fake images and videos. International Journal of Scientific Research in Computer Science Engineering and Information Technology, 7(3), 183-189.
22. Tolosana, Ruben, et al. "Deepfakes and beyond: A survey of face manipulation and fake detection." Information Fusion 64 (2020): 131-148.
23. T. T. Nguyen, C. M. Nguyen, D. T. Nguyen, D. T. Nguyen and S. Nahavandi, "Deep learning for deepfakes creation and detection", 2019.
24. Abdullah, S. M., Cheruvu, A., Kanchi, S., Chung, T., Gao, P., Jadliwala, M., & Viswanath, B. (2024, May). An analysis of recent advances in deepfake image detection in an evolving threat landscape. In 2024 IEEE Symposium on Security and Privacy (SP) (pp. 91-109). IEEE.
25. Heidari, A., Jafari Navimipour, N., Dag, H., & Unal, M. (2024). Deepfake detection using deep learning methods: A systematic and comprehensive review. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 14(2), e1520.
26. Maheshwari, R. U., Paulchamy, B., Pandey, B. K., & Pandey, D. (2025). Enhancing sensing and imaging capabilities through surface plasmon resonance for deepfake image detection. Plasmonics, 20(5), 2945-2964.
27. Sharma, V. K., Garg, R., & Caudron, Q. (2025). A systematic literature review on deepfake detection techniques. Multimedia Tools and Applications, 84(20), 22187-22229.
28. Al-Dulaimi, O. A. H. H., & Kurnaz, S. (2024). A hybrid CNN-LSTM approach for precision deepfake image detection based on transfer learning. Electronics, 13(9), 1662.
29. Ashani, Z. N., Ilias, I. S. C., Ng, K. Y., Ariffin, M. R. K., Jarno, A. D., & Zamri, N. Z. (2024). Comparative analysis of deepfake image detection method using vgg16 vgg19 and resnet50. Journal of Advanced Research in Applied Sciences and Engineering Technology, 47(1), 16-28.
30. Huang, Z., Hu, J., Li, X., He, Y., Zhao, X., Peng, B., ... & Cheng, G. (2025). Sida: Social media image deepfake detection, localization and explanation with large multimodal model. In Proceedings of the Computer Vision and Pattern Recognition Conference (pp. 28831-28841).
31. Hasanaath, A. A., Luqman, H., Katib, R., & Anwar, S. (2025). FSBI: Deepfake detection with frequency enhanced self-blended images. Image and Vision Computing, 154, 105418.
32. Al-Adwan, A., Alazzam, H., Al-Anbaki, N., & Alduweib, E. (2024). Detection of deepfake media using a hybrid CNN–RNN model and particle swarm optimization (PSO) algorithm. Computers, 13(4), 99.
33. Ghita, B., Kuzminykh, I., Usama, A., Bakhshi, T., & Marchang, J. (2024, June). Deepfake image detection using vision transformer models. In 2024 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom) (pp. 332-335). IEEE.
34. Pellicer, A. L., Li, Y., & Angelov, P. (2024). Pudd: Towards robust multi-modal prototype-based deepfake detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 3809-3817).

Dataset Reference:

https://www.kaggle.com/datasets/greatgamedota/dfdc-part-29/code

https://www.kaggle.com/datasets/ericth/deep-fake-detection-on-faces

https://www.kaggle.com/datasets/vijaydevane/deepfake-detection-challenge-dataset-face-images

https://paperswithcode.com/dataset/wilddeepfake

https://www.kaggle.com/datasets/haingn/bigger-dataset-for-image-deepfake-detection

https://www.kaggle.com/datasets/yihaopuah/deep-fake-images

Web Reference:

https://paperswithcode.com/task/deepfake-detection

https://github.com/paperswithcode

https://www.techtarget.com/searchenterpriseai/definition/convolutional-neural-network

https://www.geeksforgeeks.org/convolutional-neural-network-cnn-in-machine-learning/

Video References:

https://www.youtube.com/watch?v=2uIq1FvVfp8&t=462s

https://www.youtube.com/watch?v=w8Qx40tHeEM&t=29s

https://www.youtube.com/watch?v=kYeLBZMTLjk&pp=ygUeYXJ0aWZhY3QgYW5hbHlzaXMgaW4gZGVlcGZha2Vz

https://www.youtube.com/watch?v=7iV00Pu3ARg&pp=ygUeYXJ0aWZhY3QgYW5hbHlzaXMgaW4gZGVlcGZha2Vz