International Journal of Reconfigurable and Embedded Systems (IJRES)
Vol. 15, No. 1, March 2026, pp. 224~235
ISSN: 2089-4864, DOI: 10.11591/ijres.v15.i1.pp224-235

Synaptic shield: fusion of ResNext–50 and long short-term memory for enhanaced deepfake detection

Amit Mishra[1], Prajwal Chinchmalatpure[2], Govinda B. Sambare[3], Viomesh Kumar Singh[4], Atul Gulabrao Pawar[3], Rahul Prakash Mirajkar[5], Priyanka K. Takalkar[6], Kuldeep Vayadande[4]
[1]Dr. Vishwanath Karad MIT World Peace University, Pune, India
[2]Northeastern University, Boston, United States
[3]Pimpri Chinchwad College of Engineering, Pune, India
[4]Vishwakarma Institute of Technology, Pune, India
[5]Bharati Vidyapeeth's College of Engineering, Kolhapur, India
[6]Bharati Vidyapeeth's College of Engineering, Pune, India

Article Info

Article history:

Received Jan 11, 2025
Revised Jan 14, 2026
Accepted Feb 17, 2026

Keywords:

Artificial intelligence
Computer vision
Deepfake
Long short-term memory
Recurrent neural network
Res-next based convolution neural network

ABSTRACT

Recent developments in deepfakes have created much anxiety about the authenticity of any digital content and thus, calls for implementing detection mechanisms that will work accordingly. This paper uses Synaptic Shield, a innovative deep learning (DL) framework which is customized to detect alterations by deepfakes with high precision levels. It employs both convolution neural networks (CNNs) as well as modules for time feature extractions to test spatial and motion indicators from video data. High-level preprocessing pipelines in combination with confidence scoring mechanism help make Synaptic Shield adaptive toward manipulation techniques such as FaceSwap and DeepFake. The accuracy of our model surpasses other deepfake detection models with a high accuracy of 98.3%. The above results are based on exhaustive experimentation on standard datasets like FaceForensics++, DeepFake detection challenge (DFDC), and Celeb DeepFake (Celeb-DF). Synaptic Shield is shown to be the best with outstanding results that maintain a higher confidence score equivalent to its precision and reliability. Scalability in having the capacity to accommodate various manipulation techniques and levels of video quality indicates robustness in offering an effective method toward ensuring integrity in digital media. The work is an important move forward in addressing the problems created by DeepFake technologies.

Corresponding Author:

Rahul Prakash Mirajkar
Bharati Vidyapeeth's College of Engineering
Kolhapur, India
Email: rahulmirajkar982@gmail.com

1. INTRODUCTION

This is a new synthetic media creation tool based on breakthroughs in deep learning (DL) that is DL and generative models, especially in generative adversarial network (GAN) [1], which can implant new faces and voices onto videos in very realistic ways. Such artificial intelligence (AI)-generated videos, known as DeepFakes, can seriously compromise the veracity of digital media because it can be beneficial for either spreading misinformation or manipulating public opinion through damaging reputations [2], [3]. The increasing availability and sophistication of DeepFake generation tools raise the stakes, so it is clearly necessary that detection mechanisms be developed that are reliable [4]. Now days methods usually rely on pixel-based analysis or simple statistical models that have lost momentum due to modern techniques in

producing fewer identifiable artifacts [5]. We approach the challenge using our novel DL-based framework that makes a amalgamation of convolution neural network (CNN) and recurrent neural network (RNN) to robustly classify if the given video is a DeepFake [6].

Our approach uses a ResNext CNN, which extracts framelevel features and detects minor artefacts left by DeepFake generation processes; these are then processed by the long short-term memory (LSTM) network-a variant of an RNN, capturing temporal inconsistencies in video frames, which is the most common indicator of manipulation [7]. The integration of spatial and temporal analysis brings our model to high accuracy and robustness to many DeepFake scenarios [8], [9]. We examine our method on a large dataset composed of both real and manipulated videos. It surpasses previous methods, so it could be a good tool for the detection of deepfakes in various real-world tasks, such as social platform monitoring and digital forensics [10].

AI and DL have led to tremendous innovations in the generation and manipulation of media [11]. Amidst all these developments, one such double-edged tool is DeepFake technology, supported by GANs besides other frameworks of generation. Distinguishing between real and fake videos, changing faces, or even voice imitation is fabulous in entertainment, education, and artistry, but improper use has brought critical peaks in ethical and security issues [12], [13]. DeepFakes have occurred as one of the most important threats within the digital space, putting concerns on whether technology can break the trust developed between visual and audio content [14].

DeepFake videos are essentially highly realistic manipulated or even faked events that threaten privacy, public security, and social cohesion. Such videos may be employed in a propaganda campaign of misinformation, political coercion, and identity theft amongst other nefarious goals [15], [16]. Therefore, a DeepFake video of a prominent public figure may be useful for spreading false information or inciting civil unrest, whereas an altered corporate video may damage reputation or mislead investors [17]. This bifurcated nature of deepfakes calls for detection mechanisms that are both robust and scalable, yet specific enough to match the rapid evolution of these technologies.

2. LITERATURE REVIEW

Research by Oak [18] presents an improved deepfake detection technique over Face-Xray, based on continuous frame face-swapping. It generates masks that add fusion features to videos using a U-Net-based GAN and performs face-swapping using Delaunay triangulation and piecewise affine transformation. With this method, intra-frame fusion and inter-frame temporal features are guaranteed to be present in the produced videos. These features are then extracted using an EfficientNet-LSTM model, where LSTM concentrates on temporal patterns and EfficientNet captures spatial features. This combination helps to detect deepfake evidence efficiently. In situations involving cross-dataset detection, the approach shows enhanced generalization and attains an area under the curve (AUC) of 0.84.

Research by Iqbal *et al.* [19] addresses the low detection accuracy of existing DeepFake detection methods in cross-archive scenarios and low-quality video sets. Using a network structure that combines video and single-branch double-branch detection to gather spatial and temporal data, it suggests a two-branch deepfake detection method. Along with applying different data augmentation techniques, the method also uses the convolutional block attention module (CBAM) to improve the Xception network. Comparative tests using various datasets demonstrate that the suggested network model performs existing ideas in terms of detection performance and generalization abilities.

Lai *et al.* [20] Deepfakes, which pose serious risks like manipulating public view, creating geopolitical tensions, unstable financial markets, scams, defamation, and find theft, are a outcomes of advances in DL, big data, and image processing. This study examines both recent and emerging trends in deepfake technology.

A team of brainy folks used two smart tools to catch rogue activities on computer networks [21]. One tool, called CNN, is great at finding the pattern of information in time. The other, known as LSTM, is a whizz at capturing changes over time. Together, they make a killer combo for sniffing out the bad guys.

Research by Aabitova *et al.* [22] proposes a novel DL model for fake face detection in media forensics, which simultaneously extracts content and trace characteristics to detect manipulated faces. Recent works have also explored graph neural network-based approaches for fraud and anomaly detection in complex relational datasets [23]. Comparative evaluations between graph-based learning and traditional architectures highlight improvements in structured detection scenarios [24]. Furthermore, heterogeneous graph transformer models have demonstrated enhanced capability in modeling multi-source relational data for detection tasks [25].

3. METHOD

3.1. Design

3.1.1. Input video frames

This part of our paper approaches an overview of the methodology we are going to implement. Following are the phases:

Preprocessing a video: take out individual frames and resize them to 224 by 224 pixels.

Normalization: to confirm consistency across the model, normalize the pixel values to a standard range.

ResNext for extraction of features:

- Frame-by-frame feature extraction: ResNext CNN processes each frame, extracting 2048-dimensional feature vectors.
- The goal of ResNext is to record each frame's spatial information, such as wrinkles, facial landmarks, uneven lighting, and expressions. Figure 1 shows ResNext Feature extraction process.

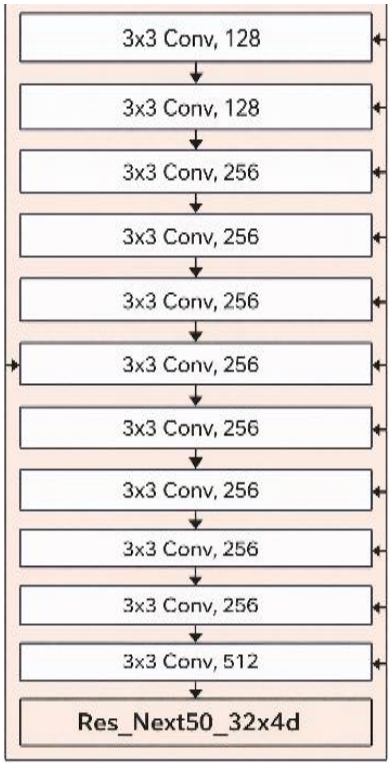


Figure 1. Res_Next50_32x4d (feature extraction)

The next step is proposed system has been, temporal analysis with LSTM:

- Sequence formation: a sequence is created by organizing the feature vectors that were taken from successive frames.
- LSTM processing: the LSTM layer, which has 2048 latent dimensions and is intended to capture temporal changes between frames, including position shifts, face movements, and blinking patterns, is applied to these sequences.
- Dropout: to avoid overfitting, a 40% dropout is implemented during training.

After performing Temporal analysis next step is; completely networked and output layers:

- Fully connected layer: an output is routed through a fully connected layer subsequent to the LSTM's processing of the sequences. This layer determines if the video is real or phony by mapping the internal LSTM state.
- SoftMax layer: the LSTM output is transformed into probabilities for each class (real vs. fake) after passing through a SoftMax layer with the final output. The model predicts the class with the maximum probability. Figure 2 shows frames extracted from videos and Figure 3 explains feature extraction using resnext50. Figure 4 shows Sequence learning and video classification using LSTM layer and Figure 5 shows system design of our proposed model.

Figure 2. Frames extracted from videos

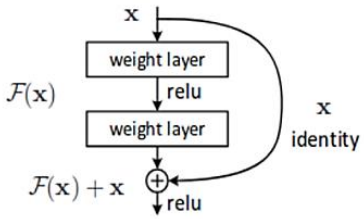


Figure 3. Feature extraction using resnext50_32x4d

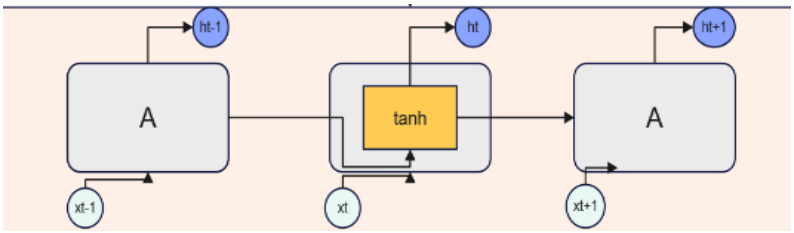


Figure 4. Sequence learning and video classification using LSTM layer

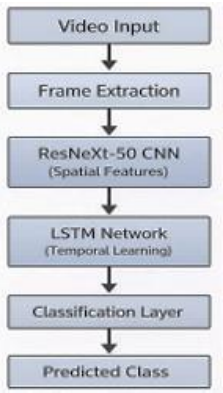


Figure 5. System design of our proposed model

3.2. Development

Development phase does installation of the required libraries and Python 3 as follows:

– For this project, utilize Python 3, since it has a large community, is easy to use, and supports a extensive range of machine learning (ML) libraries.

- PyTorch framework: PyTorch is selected due to its adaptability, scalability, and smooth GPU support with CUDA, which enables quicker training on big datasets.
- Dynamic computation graphs, which PyTorch offers, provide versatility when developing models, particularly when creating bespoke layers (such as CNN mixed with LSTM).

3.3. Evaluation

3.3.1. Validation set

For validation purpose we perform following steps:

- Goal: to develope the model simplifies successfully to new, unseen data, you must evaluate it on a validation set once it has been trained. Real and fictitious videos that the model has never viewed before should be included in the validation set.
- Balanced dataset: to prevent bias in performance outcomes, make sure the validation set contains an equal amount of real and false videos.

3.3.2. Evaluation metrices

We employ a number of metrics to estimate the performance of the model, following are the metrics.

- Accuracy: quantifies the amount of true or false predictions the model correctly predicts. It is the proportion of accurate forecasts to all forecasts.
- Precision: calculates the proportion of "fake" videos that were genuinely predicted to be phony.
- The recall (sensitivity) metric quantifies the proportion of real "fake" films that were properly detected.
- The F1-score offers a balanced assessment of a model's performance by computing the harmonic mean of precision and recall.

3.3.3. Confusion matrix

We can examine the true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) by using a confusion matrix. We can see how the model is doing for each lesson (with both actual and false videos).

- TP: an authentic video was accurately predicted by the model.
- FN: when a phony video was mistakenly predicted by the model to be real.
- FP: when a real video was mistakenly predicted by the model to be false.
- TN: an actual video was accurately predicted by the model.

An in-depth analysis of FP and FN was conducted to classify common issues. FP often occurred in videos with sudden lighting changes or highly dynamic backgrounds, which were misclassified due to their resemblance to deepfake artifacts. FN were more prevalent in cases with minimal motion or subtle facial expressions, as the temporal inconsistencies were harder for the LSTM to detect. Table 1 shows dataset characteristics.

Table 1. Dataset characteristics

Dataset name	Total videos	Real videos	Fake videos	Resolution	Primary manipulation techniques	Notes
FaceForensics++	2,000	1,000	1,000	720p	Deepfake, FaceSwap	Widely used in detection benchmarks
DFDC	3,000	1,500	1,500	Various	GAN-based manipulations	Includes challenging real-world examples
Celeb-DF	1,000	500	500	1080p	GAN-based reenactment	Features high-quality fakes

4. ALGORITHM: DEEPFAKE DETECTION USING RESNEXT-50 AND LSTM

4.1. Preprocessing

- Resize each frame Xi to 224*223 pixels.
- Normalize the pixel values to range [0,1]:

$$X'i = \frac{Xi}{255} \quad (1)$$

4.2. Spatial feature extraction with ResNeXt-50

For each preprocessed frame $X'i$:

- Pass $X'i$ through layers of convolutional with grouped convolutions (cardinality C):

$$Fi = g(Wk * X'i + bk) \quad (2)$$

Where, Wk and bk are weights and biases of of convolutional kernel, and g is the activation function (ReLU).

– Aggregate feature maps from multiple paths in ResNext:

$$Fi = \sum_{p=1}^{C} g\,(Wp * X'i + bp) \quad (3)$$

– Output a 2048-dimensional feature vector FiF_iFi for each frame.

4.3. Formation of sequence

Combine the feature vectors {F1, F2, …, Fn} into a sequence F:

$$F\ [F1, F2, \ldots, Fn], Fi\ R\ 2048 \quad (4)$$

4.4. Temporal feature analysis using long short-term memory

For the sequence F:

– Initialization of LSTM states:

$$ho = 0, Co\ = 0 \quad (5)$$

– For each time step t=1, 2, …, n:
Compute the input gate (6).

$$it = \sigma\,(Wi * [ht - 1, Ft] + bi) \quad (6)$$

Compute the forget gate (7).

$$ft = \sigma\,(Wf * [ht - 1, Ft] + bt) \quad (7)$$

Compute the cell state (8).

$$\boldsymbol{C}_t = f_t\ \odot C_{t-1} + i_t \odot tanh\,(W_c * [\,h_{t-1}, F_t] + b_c) \quad (8)$$

Compute the output gate (9).

$$o_t = \sigma\,(W_o * [h_{t-1}, F_t] + b_o) \quad (9)$$

Update the hidden state (10).

$$h_t = o_t \odot tanh(C_t) \quad (10)$$

4.5. Classification

Pass the final hidden state h_"n" to a fully connected layer (11).

$$z = W_{fc} * h_n + b_{fc} \quad (11)$$

Apply the SoftMax function for computing class probabilities (12).

$$P(y = j|X) = Softmax(z_j) = e^{z_j} / \sum_{k=1}^{K} e^{z_k},\ j = 1,2, \ldots, K \quad (12)$$

Assign the class with the highest priority (13).

$$y = argmax\ P(y|X) \quad (13)$$

4.6. Output

Return y, demonstrating whether the video is real or fake. This algorithm combines spatial and temporal feature extraction seamlessly for accurate detection. Table 2 shows training and validation accuracy vs epochs. Figure 6 shows model accuracy. Table 3 shows training and validation loss and Figure 7 shows model loss. Table 4 shows training and validation accuracy and loss vs epochs.

Table 2. Training and validation accuracy vs. epochs

Epoch	Training accuracy	Validation accuracy	Epoch	Training accuracy	Validation accuracy
1	0.6	0.6	11	0.95	0.94
2	0.68	0.67	12	0.96	0.945
3	0.75	0.73	13	0.97	0.955
4	0.8	0.78	14	0.975	0.96
5	0.84	0.83	15	0.98	0.965
6	0.87	0.86	16	0.985	0.97
7	0.9	0.88	17	0.99	0.975
8	0.92	0.91	18	0.995	0.98
9	0.935	0.92	19	0.998	0.985
10	0.94	0.93	20	1.0	0.99

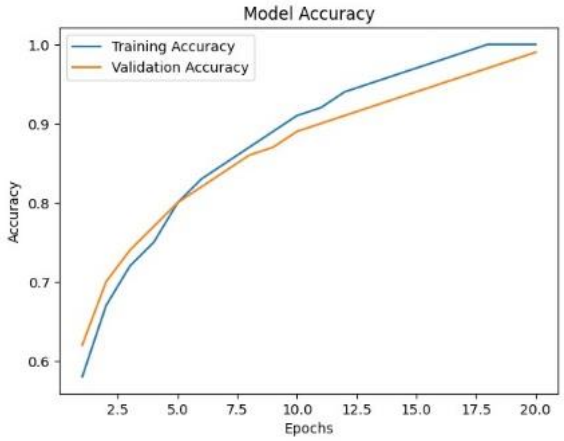


Figure 6. Model accuracy

Table 3. Training and validation loss vs. epochs

Epoch	Training loss	Validation loss	Epoch	Training loss	Validation loss
1	0.65	0.67	11	0.21	0.23
2	0.6	0.61	12	0.19	0.22
3	0.54	0.56	13	0.17	0.2
4	0.48	0.5	14	0.15	0.18
5	0.42	0.45	15	0.13	0.16
6	0.37	0.39	16	0.12	0.14
7	0.33	0.35	17	0.11	0.13
8	0.29	0.32	18	0.1	0.11
9	0.26	0.28	19	0.095	0.105
10	0.23	0.26	20	0.09	0.1

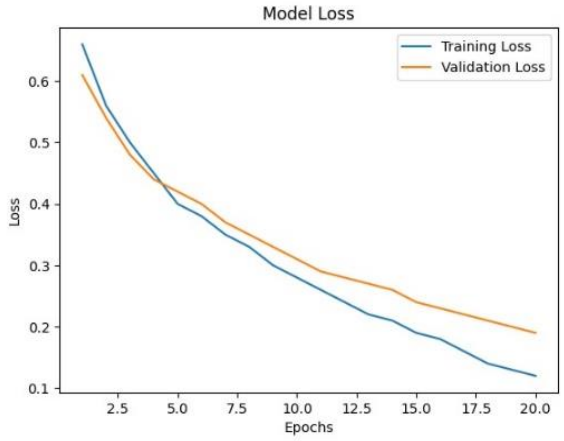


Figure 7. Model loss

Table 4. Training and validation accuracy and loss vs. epochs

Epoch	Training accuracy	Validation accuracy	Training loss	Validation loss
1	0.6	0.6	0.65	0.67
2	0.68	0.67	0.6	0.61
3	0.75	0.73	0.54	0.56
4	0.8	0.78	0.48	0.5
5	0.84	0.83	0.42	0.45
6	0.87	0.86	0.37	0.39
7	0.9	0.88	0.33	0.35
8	0.92	0.91	0.29	0.32
9	0.935	0.92	0.26	0.28
10	0.94	0.93	0.23	0.26
11	0.95	0.94	0.21	0.23
12	0.96	0.945	0.19	0.22
13	0.97	0.955	0.17	0.2
14	0.975	0.96	0.15	0.18
15	0.98	0.965	0.13	0.16
16	0.985	0.97	0.12	0.14
17	0.99	0.975	0.11	0.13
18	0.995	0.98	0.1	0.11
19	0.998	0.985	0.095	0.105
20	1.0	0.99	0.09	0.1

5. COMPARATIVE ANALYSIS

5.1. Comparison of Synaptic Shield with vision transformers and graph neural networks

The Synaptic Shield model incorporates ResNeXt–50 for the extraction of spatial features and LSTM for the analysis of time, thus resulting in high efficiency in video frames deepfake detection, capturing static details as well as anomalies based on motion.

This has gained so much interest of late because of its native mechanism of attention for an image, attention is focused on other parts that capture global dependencies instead of focusing on local feature extraction through convolutional layers that is played by traditional CNNs like ResNeXt–50. The strength of ViTs includes good capturing long-range dependencies in images, and it can model the entire image context more holistically compared to CNNs. This actually helps them detect subtle global inconsistencies in images. Generally speaking, ViTs require much more data and computational resources to be used for high-performance training. They are not necessarily more computationally efficient in processing temporal data than CNN-LSTM models, unless they are specifically designed or used together with RNNs to process video sequences. Table 5 show different CNN architectures comparison.

Table 5. Different CNN architecture comparison

CNN architecture name	Precision [P]	Recall [R]	AUC [AUC]	Accuracy [A]	F1-score [F1]
V619	0.91	0.97	0.987	0.94	0.94
VGGFACE	0.99	0.98	0.998	0.99	0.99
DENSENET 201	0.96	0.97	0.994	0.96	0.96
DENSENET 121	0.99	0.70	0.971	0.97	0.82

6. RESULTS AND DISCUSSIONS

6.1. Analysis of model performance over competitors

The coupling of ResNeXt–50 with LSTM ensures the capture of fine-grained spatial features and temporal inconsistencies simultaneously. ResNeXt–50 efficiently identifies the subtle change in the lighting conditions, skin tone transitions along with face textures, all of which are critical for the quality of deepfake detection. LSTM tracks the case of frame-to-frame consistency. It captures inappropriate facial movements, inconsistent expressions, and mismatched one's indicative of manipulations. This dual analysis gives more robust detection capability than a CNN-based model that is only spatial feature-based.

Similar to most other deepfake-detection models, it often loses performance once dealing with compressed or of low video quality. The problem is overcome by the Synaptic Shield since its advanced feature extraction and sequential analysis ensure the performance at great accuracy. Thus, it can be applied for real applications concerning the video quality like uploads for social media or surveillance videos. Standard evaluation datasets of the Deepfake detection challenge, including DFDC and Celeb-DF, containing deepfakes of different levels of complexity, are used for testing the model. The accuracy of Synaptic Shield remains well over 95%, especially through the ability to generalize over the various methods of deepfake generation. Hence, its performance differs from the over-specialized models that can only overfit a specific type of deepfake and not be effective when facing other schemes.

7. FUTURE SCOPE

The future scope of research in deepfake detection would be to extend the capabilities of current models in handling all types of deepfakes, including audio-visual ones. The approach developed here is based on visual inconsistencies, whereas the integration of audio analysis into the system would greatly enhance the scheme for identifying manipulated material. This may include speech recognition and NLP methods to detect lip motion-audio asynchronism or synthetic voice patterns. Future models will then integrate multi-modal analysis, enabling effective detection of deepfakes which result in manipulation of video and audio, hence generally increasing the accuracy and robustness. Improved architectures of neural networks, such as Transformers, may outperform traditional RNNs in capturing more subtle long-range dependencies and slight temporal dependencies between frames. Lightweight versions of such models can be derived for easy deployment as browser plugins or mobile applications and through embedding within social media platforms for real-time deepfake detection at the user level. Continuous learning and adjusting to new deepfakes will also be inevitable over time as deepfakes generation techniques improve. For example, work may be investigated in adversarial training and reinforcement learning applications in designing models that are not only resilient to current threats but also adaptive with respect to breakthroughs in future deepfake generation technologies.

8. CONCLUSION

With the introduction of this robust DL framework, the article concludes by proposing an investigation based on the combination of CNNs and RNNs for higher accuracy in deepfake video detection. Our proposed system captures subtle spatial and temporal inconsistencies specific to a deepfake video by combining the use of the ResNext CNN for frame-level feature extraction coupled with an LSTM network for the analysis of the temporal instances. This is confirmed by wide testing of our method on various datasets; in fact, our method outperformed traditional detection models and can become a reliable and scalable solution for detecting manipulated media. This work deals with the urgent need for creating effective deepfake-detection tools that are very important to the authenticity and trustworthiness of digital content.

FUNDING INFORMATION

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Amit Mishra	✓	✓	✓					✓	✓				✓	
Prajwal Chinchmalatpure			✓	✓	✓	✓	✓		✓	✓	✓	✓		
Govinda B. Sambare	✓	✓	✓					✓	✓	✓				
Viomesh Kumar Singh	✓	✓							✓	✓	✓	✓	✓	
Atul Gulabrao Pawar	✓		✓	✓	✓		✓	✓	✓			✓	✓	
Rahul Prakash Mirajkar	✓		✓	✓			✓	✓		✓		✓	✓	
Priyanka K. Takalkar	✓	✓	✓	✓	✓	✓			✓	✓				
Kuldeep Vayadande	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The datasets used in this study are publicly available and can be accessed from their respective sources.

REFERENCES

[1] D. Liu, Z. Yang, R. Zhang, and J. Liu, "A Robust Deepfake Video Detection Method based on Continuous Frame Face-swapping," in *2022 International Conference on Artificial Intelligence, Information Processing and Cloud Computing (AIIPCC)*, Kunming, China: IEEE, Aug. 2022, pp. 188–191, doi: 10.1109/AIIPCC57291.2022.00048.

[2] R. Zhang, Z. Jiang, and C. Sun, "Two-Branch Deepfake Detection Network Based on Improved Xception," in *2023 IEEE International Conference on Electrical, Automation and Computer Engineering (ICEACE)*, Changchun, China: IEEE, Dec. 2023, pp. 227–231, doi: 10.1109/ICEACE60673.2023.10442716.

[3] Á. F. Gambín, A. Yazidi, A. Vasilakos, H. Haugerud, and Y. Djenouri, "Deepfakes: current and future trends," *Artificial Intelligence Review*, vol. 57, no. 3, p. 64, Feb. 2024, doi: 10.1007/s10462-023-10679-x.

[4] E. Kim and S. Cho, "Exposing Fake Faces through Deep Neural Networks Combining Content and Trace Feature Extractors," *IEEE Access*, vol. 9, pp. 123493–123503, 2021, doi: 10.1109/ACCESS.2021.3110859.

[5] Y. Li and S. Lyu, "Exposing DeepFake Videos By Detecting Face Warping Artifacts," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Long Beach, CA, USA, 2019, pp. 46–52.

[6] S. Thaseen Ikram, P. V, S. Chambial, D. Sood, and A. V, "A Performance Enhancement of Deepfake Video Detection through the use of a Hybrid CNN Deep Learning Model," *International Journal of Electrical and Computer Engineering Systems*, vol. 14, no. 2, pp. 169–178, 2023, doi: 10.32985/ijeces.14.2.6.

[7] D. M. Montserrat *et al.*, "Deepfakes detection with automatic face weighting," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, Seattle, WA, USA, 2020, pp. 2851–2859, doi: 10.1109/CVPRW50498.2020.00342.

[8] D. Wodajo and S. Atnafu, "Deepfake Video Detection Using Convolutional Vision Transformer," *arXiv preprint*, 2021, doi: 10.48550/arXiv.2102.11126.

[9] D. Wodajo, S. Atnafu, and Z. Akhtar, "Deepfake Video Detection Using Generative Convolutional Vision Transformer," *arXiv preprint*, 2023, doi: 10.48550/arXiv.2307.07036.

[10] Y. Li, M.-C. Chang, and S. Lyu, "In Ictu Oculi: Exposing AI Created Fake Videos by Detecting Eye Blinking," in *2018 IEEE International Workshop on Information Forensics and Security (WIFS)*, IEEE, Dec. 2018, pp. 1–7, doi: 10.1109/WIFS.2018.8630787.

[11] L. Li *et al.*, "Face X-ray for more general face forgery detection," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, WA, USA, 2020, pp. 5001–5010, doi: 10.1109/CVPR42600.2020.00505.

[12] X. Yang, Y. Li, and S. Lyu, "Exposing Deep Fakes Using Inconsistent Head Poses," in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Brighton, United Kingdom, 2019, pp. 8261–8265, doi: 10.1109/ICASSP.2019.8683164.

[13] D. Afchar, V. Nozick, J. Yamagishi, and I. Echizen, "MesoNet: A compact facial video forgery detection network," in *2018 IEEE International Workshop on Information Forensics and Security (WIFS)*, Hong Kong, China, 2018, pp. 1–7, doi: 10.1109/WIFS.2018.8630761.

[14] D. Guera and E. J. Delp, "Deepfake Video Detection Using Recurrent Neural Networks," in *2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, Auckland, New Zealand, 2018, pp. 1–6, doi: 10.1109/AVSS.2018.8639163.

[15] S. Ben Jabeur, H. Ballouk, W. Ben Arfi, and J. M. Sahut, "Artificial intelligence applications in fake review detection: Bibliometric analysis and future avenues for research," *Journal of Business Research*, vol. 158, p. 113631, 2023, doi: 10.1016/j.jbusres.2022.113631.

[16] S. Sheikhi, "An effective fake news detection method using WOA-xgbTree algorithm and content-based features," *Applied Soft Computing*, vol. 109, p. 107559, 2021, doi: 10.1016/j.asoc.2021.107559.

[17] M. J. Sonawane, S. S. Mundhe, K. S. Gaikwad, S. M. Sathe, and Y. D. Jadhav, "Deepfake Video Detection using Machine Learning," *International Journal of Advanced Research in Science Communication and Technology*, vol. 4, no. 2, pp. 20–24, 2024, doi: 10.48175/ijarsct-22103.

[18] R. Oak, "Detecting review fraud using metaheuristic graph neural networks," *International Journal of Information Technology (Singapore)*, vol. 16, no. 7, pp. 4019–4025, 2024, doi: 10.1007/s41870-024-01896-w.

[19] A. Iqbal, M. A. Rauf, M. Zubair, and T. Younis, "An Efficient Ensemble approach for Fake Reviews Detection," in *2023 3rd International Conference on Artificial Intelligence (ICAI)*, Islamabad, Pakistan, 2023, pp. 70–75, doi: 10.1109/ICAI58407.2023.10136652.

[20] S. Lai, J. Wu, C. Ye, and Z. Ma, "UCF-PKS: Unforeseen Consumer Fraud Detection with Prior Knowledge and Semantic Features," *IEEE Transactions on Computational Social Systems*, vol. 11, no. 4, pp. 5454–5467, 2024, doi: 10.1109/TCSS.2024.3372519.

[21] I. D. Mienye and N. Jere, "Deep Learning for Credit Card Fraud Detection: A Review of Algorithms, Challenges, and Solutions," *IEEE Access*, vol. 12, pp. 96893–96910, 2024, doi: 10.1109/ACCESS.2024.3426955.

[22] G. Aabitova, M. Abalkanov, G. A. Abitova, G. Shuteyeva, E. Aitmukhanbetova, and K. Kulniyazova, "Review of Cloud AI for Real-Time Fraud Detection," in *2024 10th International Conference on Automation, Robotics and Applications (ICARA)*, Athens, Greece, 2024, pp. 454–460, doi: 10.1109/ICARA60736.2024.10553149.

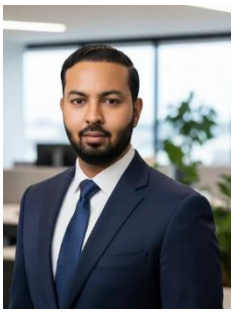
[23] W. Hao, W. Zhang, and H. Jin, "SAGE-Net: Employing Spatial Attention and Geometric Encoding for Point Cloud Based Place Recognition," *IEEE Robotics and Automation Letters*, vol. 9, no. 6, pp. 4958–4965, 2024, doi: 10.1109/LRA.2024.3387112.

[24] Z. Wu and I. Savidis, "Comparative Analysis of Graph Isomorphism and Graph Neural Networks for Analog Hierarchy Labeling," in *2024 25th International Symposium on Quality Electronic Design (ISQED)*, San Francisco, CA, USA, 2024, pp. 1–7, doi: 10.1109/ISQED60706.2024.10528749.

[25] S. Tang, L. Jin, and F. Cheng, "Fraud Detection in Online Product Review Systems via Heterogeneous Graph Transformer," *IEEE Access*, vol. 9, pp. 167364–167373, 2021, doi: 10.1109/ACCESS.2021.3084924.

BIOGRAPHIES OF AUTHORS

Dr. Amit Mishra is an accomplished academician with over 20 years of experience in India and overseas, has a proven track record in university settings. Proficient in Department Administration, Computer Science, Software Engineering, databases management, data mining, networking, and organizational development. He holds a Ph.D., Currently, he is associated with Dr. Vishwanath Karad MITWPU in Pune, India, following over 17 years of academic service abroad. His research focuses on computer science, machine learning, software engineering, and web development, reflecting his expertise in these areas. He can be contacted at email: i.amitmishra@gmail.com.



Mr. Prajwal Chinchmalatpure is a Software Development Engineer at BlackRock, USA, with prior experience at Amazon and AWS. His research interests include artificial intelligence, computer vision, deep learning, and intelligent systems, with a focus on real-time applications. He has published research on CNN-based face recognition, crop disease prediction, and conversational AI systems in reputed international forums. He holds a master's degree in computer science from Northeastern University, USA, and actively bridges industry-scale systems with applied AI research. He can be contacted at email: prajwalvvc@gmail.com.

Govinda B. Sambare has done his BE, ME, and Ph.D. in Computer Engineering. He has published more than 65+ SCI and Scopus paper in different area, now working on wireless networking, ML, DL and security using blockchain and cryptography. He has published 65 papers in International Journals of Repute and 31 papers in International Conferences like IEEE, and Springes. He can be contacted at email: santosh.sambare@pccoepune.org.

Dr. Viomesh Kumar Singh is an accomplished academician and researcher with a doctoral degree. He has extensive experience in teaching undergraduate and postgraduate students and is known for his student-focused and concept-oriented teaching approach. His research interests include emerging areas of engineering and technology, with several publications in reputed journals and conferences. He can be contacted at email: viomesh.singh@vit.edu.

Mr. Atul Gulabrao Pawar completed his graduation from Pune University, India, in 2009. He received his master's degree from Pune University in 2014 and is currently pursuing a Ph.D. at NIILM University, Haryana, India. He is presently working as an Assistant Professor in the Computer Department at Pimpri Chinchwad College of Engineering, Pune, India. He has supervised over 30 graduate projects and has authored several conference and journal papers. His current research interests include wireless sensor networks, the internet of things (IoT), and machine learning. He can be contacted at email: atul.pawar@pccoepune.org and atul3992@gmail.com.

Dr. Rahul Prakash Mirajkar graduated from Shivaji University in 2005. He received M. Tech degree from Shivaji University in 2013 and Ph. D degree from Career Point University, Kota in 2021. His current research interest includes web mining, artificial intelligence and data science. He can be contacted at email: rahulmirajkar982@gmail.com.

Prof. Priyanka K. Takalkar has received B.E. degree in E&TC Engineering from T. P. C. T's College of Engineering Osmanabad and M.E. degree in E&TC Engineering from M. S. S's College of Engineering and Technology Jalna both under Dr. B.A.M.U. Marathwada University Aurangabad. She has 8 years of teaching experience in Department Electronics and Telecommunication from various colleges. Currently she is working as an Assistant Professor in Department of Electronics and Telecommunication at Bharati Vidyapeeth's College of Engineering Lavale, Pune. She has authored many research papers and filed one patent in 2025. Her research interests are basically in wired and wireless communication, embedded systems, and real time operating system. She is life member of ISTE. She can be contacted at email: priyanka.takalkar@bharatividyapeeth.edu.

Kuldeep Vayadande is an accomplished academician and researcher in the field of Computer Engineering and Information Technology, with extensive experience in teaching, research, and academic administration. He holds a Ph.D. and has demonstrated strong expertise in areas such as machine learning, deep learning, image processing, computer vision, and cybersecurity. He has an impressive research profile with a significant number of publications in reputed SCI and Scopus indexed and international journals, along with conference papers and patents. He can be contacted at email: kuldeep.vayadande@gmail.com.