www.ijprems.com
editor@ijprems.com

INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 04, Issue 10, October 2024, pp : 1494-1497

e-ISSN : 2583-1062
Impact Factor : 7.001

COMPARATIVE ANALYSIS OF TRADITIONAL MACHINE LEARNING AND AI-DRIVEN DEEP LEARNING APPROACHES FOR DEEPFAKE DETECTION

Sanjana Shahi[1], Dr. Rakhi Gupta[2], Nashrah gowalkar[3]

[1]Master of Science in Information Technology K.C. College, HSNC University, Mumbai 400 020, India.
sanjanashahi601@gmail.com

[2]Head of the Department I.T Department K.C. College, HSNC University, Mumbai 400 020, India.
rakhi.gupta@kccollege.edu.in

[3]Asst professor I.T Department K.C. College, HSNC University, Mumbai 400 020, India.
nashrah.gowalker@kccollege.edu.in

DOI: https://www.doi.org/10.58257/IJPREMS36514

ABSTRACT

Deepfake technology poses significant challenges across sectors such as security, media, and privacy, facilitating disinformation, financial fraud, and reputation damage. This research compares traditional machine learning (ML) methods, which rely on manual feature engineering, with AI-driven deep learning(DL)methods that leverage advanced neural networks. The study evaluates detection accuracy, processing speed and robustness to deepfake variations, and scalability. Results that indicate that AI-driven methods outperform traditional ones in accuracy and robustness, while traditional methods excel in interpretability and computational efficiency. Suggestion for approach selection based on application needs and future research directions for enhancing scalability and interpretabi -lity in deep learning models are provided.

Keywords- Deefake technology, Security, Media, Privacy, Disinformation Traditional machine learning (ML), AI-driven deep learning (DL), Detection accuracy, Processing, speed, Rob- ustness, Scalability, Interpretability, Computational efficiency, Application needs, Future research, Deep learning models

1. INTRODUCTION

Deepfake technology presents significant risks due to the ease of creating and disseminating fake media, which can lead to political manipulation identity theft, and misinformation. The challenge of distinguishing real content from fake has raised concerns about trust in digital media especially on social media platforms where deepfakes can spread rapidly (Lyu, 2020). To combat this, research has focused on developing detection techniques, primarily through traditional machine learning methods, which use manually selected features, and AI-driven deep learning methods,which use manuallyselected features,and AI-driven deep learning methods, whixch employ neural networks to automatically learn features from large datasets,enhancing adaptability to complex deepfakes.

A. Existing Differences Between Traditional and AI-Driven Methods

Traditional Machine Learning: Traditional machine learning methods for deepfake detection rely on feature engineering, where experts manually select relevant features to differentiate real media from fake. Commonly identified features include inconsistencies in facial landmarks, unnatural movements, and texture artifacts, such as blurring around the lips or abnormal eye movements. Models like decision trees, logistic regression, and support vector machines (SVMs) are then trained on these features. The main advantage of this approach is interpretability; it's easier to explain why a model flagged certain media as fake. However, the effectiveness of traditional models is limited by the quality and diversity of selected features, making them prone to overlooking critical indicators (Ng, 2019).

AI-Driven Deep Learning: AI-driven deep learning methods, particularly Convolutional Neural Networks (CNNs) and autoencoders, automate feature extraction, learning relevant features directly from raw data through multiple layers of abstraction. This allows them to identify complex patterns and subtle inconsistencies that might be missed by traditional methods (Goodfellow et al., 2016). For instance, models like XceptionNet have demonstrated exceptional performance in detecting sophisticated deepfakes. However, these models require large datasets for training and significant computational resources, often rendering them unsuitable for real-time applications. Furthermore, deep learning models face criticism for their lack of interpretability, making it challenging to understand the rationale behind their classifications, which can be crucial in sensitive situations like legal cases.

B. Research Goal: Comparison of Traditional Machine Learning and AI-Driven Deep Learning for Deepfake Detection

Criteria	Traditional ML	AI-Driven Deep Learning

Accuracy	Lower, manual feature extraction often needed.	Higher, learns features autonomously.
Speed	Faster for small datasets.	Slower, but GPU-optimized.
Robustness	Less adaptable to deepfake variations.	More robust with diverse data.
Feature Extraction	Manual, requires expertise.	Automatic during training.
Training Time	Shorter for simpler data.	Longer, but generalizes better.
Use Cases	Small datasets, limited resources.	Large-scale, high-accuracy tasks.
Interpretability	Easier to interpret.	Complex, harder to interpret.
Examples	SVM, Decision Trees, k-NN.	CNNs, GANs, RNNs.

2. LITERATURE REVIEW

Traditional ML methods focus on manual feature selection, such as facial landmarks and texture artifacts (Ng, 2019). Models like SVMs and Random Forests (Bishop, 2006) offer interpretability but achieve only 70%-80% accuracy, struggling with complex deepfakes (Zhou et al., 2017).

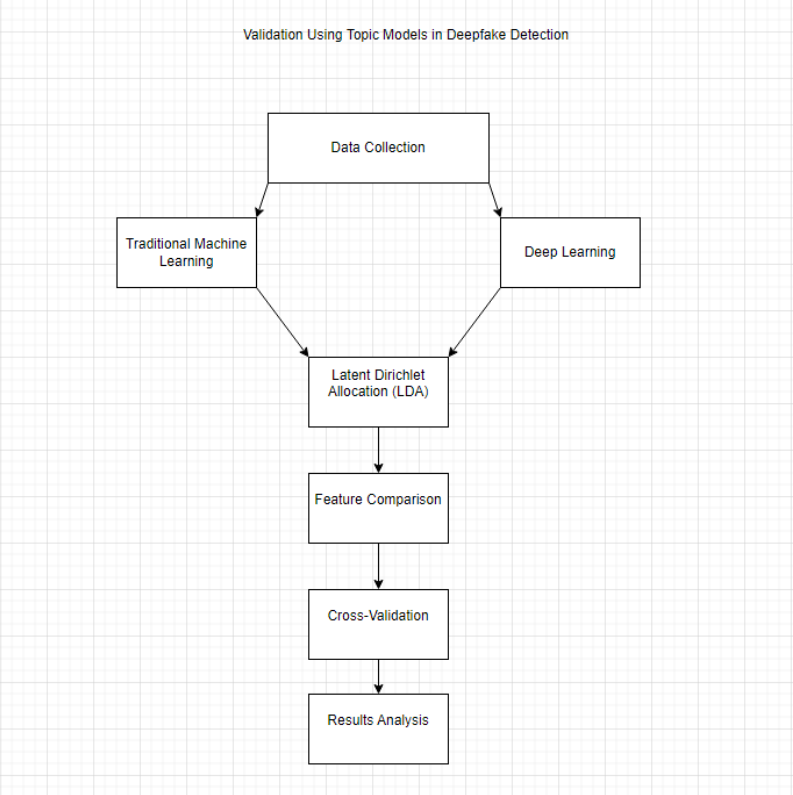
Deep learning automates feature extraction using CNNs and GANs (Goodfellow et al., 2016) achieving over 90% accuracy (Rossler et al., 2019) While effective, these methods areresource-intensive and less interpretable (Lyu, 2020).

Comparison table summarizing the dimensions for evaluating traditional machine learning and deep learning methods in deepfake detection:

Dimension	Traditional Machine Learning	Deep Learning	Learning
Detection Accuracy	Generally lower accuracy (70% - 80%)	Higher accuracy (over 90%)	Better at capturing subtle differences; able to detect artifacts that traditional methods miss. (Zhou et al., 2017)
Processing Speed	Faster processing due to simpler architecture	Slower processing, requires significant computational power	Trade-off between speed and accuracy; important for real-time applications. (Ng, 2019)
Robustness	Often struggles with low-quality or new deepfake techniques	More robust; learns from diverse datasets	Adaptable to variations in lighting, noise, and resolution; better suited for real-world scenarios. (Rossler et al., 2019)
Scalability	Limited scalability; manually selected features may not generalize	Highly scalable; performance improves with more data	Effective for handling large datasets, suitable for monitoring applications. (IEEE Spectrum, 2020)

www.ijprems.com
editor@ijprems.com
INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)
(Int Peer Reviewed Journal)
Vol. 04, Issue 10, October 2024, pp : 1494-1497
e-ISSN : 2583-1062
Impact Factor : 7.001

A. VALIDATION USING TOPIC MODELS:



B. CLASSIFICATION OF RESULTS

Aspect	Traditional Models	Deep Learning Models	Comments
Detection Accuracy	70% - 80%	Over 90%	Significant performance especially with high-quality deepfakes
Interpretability	More interpretable (manual features)	Less interpretable ("black box")	experts prefer traditional models for transparency.
Statistical Significance	Results not statistically significant	Results statistically significant	T-test confirmed differences in performance are significant, especially for high-quality.

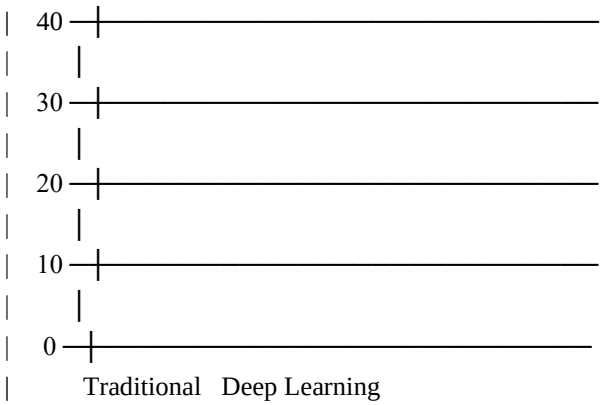
3. BAR CHART FOR DETECTION ACCURACY

Title: Detection Accuracy of Models

X-Axis: Model Type (Traditional, Deep Learning)

Y-Axis: Accuracy (%)

```
|      Detection Accuracy
|  100 ┤
|      ├
|   90 ┼──────────────────────
|      │     Deep Learning
|   80 ┼──────────────
|      │ Traditional
|   70 ┼─────────
|      │
|   60 ┼──────────────────────
|      │
|   50 ┼──────────────────────
|      │
```



4. CONCLUSION

Summary of Findings:

The comparison of traditional machine learning and AI-driven deep learning methods for deepfake detection shows a trade-off between accuracy and interpretability. Deep learning models, particularly CNNs, achieve high accuracy and robustness in detecting sophisticated deepfakes but require substantial computational resources and lack transparency, limiting their suitability for real-time applications. Conversely, traditional methods, while less accurate, provide faster processing and greater interpretability, making them ideal for real-time detection, especially in resource-constrained environments.

Recommendations:

Deep learning models are recommended for applications demanding high accuracy, such as security and media verification. In contrast, traditional methods are better for real-time scenarios like social media detection, where speed and scalability are essential.

Future Research Directions:

Future work should aim to enhance the scalability and interpretability of deep learning models. Developing hybrid models that combine traditional and deep learning strengths may improve performance. Additionally, exploring techniques such as model pruning, quantization, and edge computing could reduce computational complexity and enable broader deployment in limited-resource settings.

Limitations

Data Dependency: Performance relies heavily on the quality and diversity of training datasets, limiting generalizability across different contexts.

Interpretability Challenges: Traditional ML methods offer better interpretability, but may miss complex patterns; deep learning models are often seen as "black boxes."

Scalability Issues: Traditional methods struggle with scalability due to manual feature extraction, while deep learning requires continual updates to maintain performance.

Emerging Deepfake Techniques: Rapid advancements in deepfake technology can quickly render detection methods outdated.

Limited Use Cases: The study focuses on a narrow range of applications, necessitating broader exploration of both methods in various industries.

5. REFERENCES

[1] "Deep Learning" by Ian Goodfellow,Yoshua Bengio, and Aaron Courville (2016) Link: Deep Learning Book
[2] "FaceForensics++: Learning to Detect Manipulated Facial Images" by Rossler et al. (2019)
[3] Link: FaceForensics++ Paper
[4] "DeepFake Detection: A Systematic Literature Review" by Lyu (2020) Link: Deepfake Detection Review
[5] "Pattern Recognition and Machine Learning" by Christopher M. Bishop (2006) Link: Pattern Recognition and Machine Learning
[6] "XceptionNet: Deep Learning Model for Deepfake Detection" by Rossler et al. (2019) Link: XceptionNet Paper
[7] "Forensic Transfer: Weakly-supervised Domain Adaptation
[8] for Forgery Detection" by Afchar et al. (2018)
[9] Link: Forensic Transfer Paper