iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: V Month of publication: May 2025

DOI: https://doi.org/10.22214/ijraset.2025.70664

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

Enhancing Deepfake Detection through Hybrid Mobile Net-LSTM Model with Real-Time Image and Video Analysis

Ewan Ulysses P F[1], Varun Prakash M[2], Dr. Raghavendra V[3]
[1, 2]B Tech CSE, [3]Assistant Professor, SRM Institution of Science & Technology Chennai, India

Abstract: The race to crush information integrity and public trust is being won by one thing: deepfakes, AI manipulated media. The goal is to enhance our Deepfake detection using MobileNet and LSTM network. MobileNet's lightweight CNN architecture is able to coreinduce thevisual features in an image by an image and in avideoand can pick up the slight visual clues of textures and facial structure. A temporal inconsistencies that cannotbeseenby image base methods are then analyzed using an LSTM network. The hybrid model is trained on real and deepfake media datasets, and is thus adaptable to emerging deepfake techniques. This has auserfaceinterfacetoanalyzethereal time and fly mediatogettheanalysisandanalysisscoreand visual feedback of the identified artifacts. Unique to this system is its versatility for images and videos, and its real time capability, making it a suitable choice forpracticaluse in social media,journalism,andlawenforcementcombating the spread of misinformation with a guarantee of digital media authenticity.

Keywords: Media integrity, real-time detection,temporalmodeling,imageanalysis,videoanalysis,mo bilenet lstm, real time detection, deepfake detection.

I. INTRODUCTION

Advancements in artificial intelligence have produced deepfake tech which leverages to create highly convincing manipulated media that might be images or videos. Deepfakes startedoutasseenbysomeasaform of creative expression, but have quickly turned into abigconcerninallfields—frompoliticstonewsmediato social platforms. Deepfake algorithms have enabled voters to create synthetic content where they appear to perform actions or make statements they neverdid,and this medium has now become a powerful purveyor of misinformation, a reputation wrecker andadestroyerof public trust in digital information.

Generative Adversarial Networks (GANs) at the heartof most deepfake generation techniques permit it to generate a person's likeness or actions with extremely high precision. Although deepfake generation has become more accessible (and creates near authentic content with or without specializedtechnicalexpertise), this technology continues to grow in impact,questioning media authenticity and digital integrity.

Most of the existing detection methods are trained to find visual inconsistencies or artifacts like the presence of facial articulation irregularity or some mismatch of lighting from one frame to another. Although very effective against standard manipulations, they do not work well against sophisticated deepfake algorithms aimed at eliminating or camouflaging such cues. In addition, image-based analysis is the main limitationof many systems, restricting them toanalysisofsequential patterns (e.g., video based contents).

In this work we propose a hybrid model for deepfake detection using MobileNet, a lightweight CNN optimized for spatial feature extraction and Long Short Term Memory networks (LSTM) specialized for temporal analysis. Satisfyingly, MobileNet can successfully localize spatial inconsistencies inindividual frames, including texture anomalies and misaligned facial features. This is complemented by LSTM which analyzes sequential dependency among frames to detect temporal inconsistencies characteristic of manipulated videos.

The proposed system addresses the problem using a unified approach using a combination of spatial and temporal features by integrating MobileNet andLSTM. In MobileNet, each video frame are processed independently to extract spatial features, followed by sequential analysis by LSTM on the temporal pattern represented by the extracted spatial features. The system's ability to capture subtle, sophisticated, and high value manipulations is improved due to this dual focus architecture.

On a large and diversified dataset including real and deep fake images and videos created through many methods, the model is trained.

This comprehensive training boosts adaptability tothefutureinthedeepfake generation space. In addition, the lightweight design of the model supports near real time analysis, aprerequisite for real world applications including social media monitoring, journalism, and public safety.

The system provides an easy to use interface enabling uploading of images, videos or URLs for analysis. It provides visual explanations of detected anomalieswhile being accessible to non-technical users, and outputs a detection score so that the process is transparent. Through their application to a range of different domains, this system provides evidence of its ability to fight misinformation and maintain digital media authenticity.

To this end, it can be summarized that the presented modelofthisprojectisarobust,efficientand user-friendly hybrid model that involves spatial and temporal analysis. Its real time capabilities, and the appeal it carries in real timeapplication to safeguard digitized information from manipulation and misuse makes it a nice to have tool.

II. LITERATURE SURVEY

1) J.Choy,S.W.Lee,andC.P.M.ChanSee(2021)

A deepfake detection model wasproposedin[1],which integrates deep learning and biometric analysis, basedon iris patterns as detection metrics. Their model took advantage oftheseirisreplicationchallengestoimprove upon detection accuracy by real fine-grained details. Nevertheless, the model's addressing of high resolution media is restricted in cases of low video quality. Furthermore, since temporal inconsistencies in video sequences are crucial for a robust detection, this approach does not address them.

2) W.Wang,J.Dong,Y.Tang,Y.Zhang,andH.Liu (2021)

The Inception-Res-net architecture coupled with attention mechanisms was proposed by Dong et al. [2] for the implementation of a novel deep fake detection system. Facial regions susceptible to manipulation are selectively isolated by this integration so themodelcan discern fine artifacts in high fidelity deepfake videos. Attention mechanisms lead to reliability in detection of deepfakes, yet the system is unable toanalyzetemporal dependencies and thus limits effectiveness in video based deepfake detection.

3) A. Kurniawan,M.Pamungkas,andM.N.Husen (2020)

In Husen et al. [3],anInception-ResNet-v2architecture mated withLSTMlayerswaspresentedtomodelspatial and temporal features for deepfake detection. CNN components worked well to extract frame specific features and the LSTM layers helped to capture the inconsistencies acrosstheframes.Weshowedthatusing the spatialaswellastemporalanalysisisimportant,still the computational cost is rather highwhichhindersreal time deployment.

4) J.Cheng,A.Jaiswal,Y.Wu,L.Nataraj,M.Sabir and S.Chandrasekaran (2020)

In a sequential inconsistency detection of deepfake videos, Sabir et al. [4] proposed a recurrent convolutional model. The model captured motion and facial expression inconsistencies that are not evident in single frame evaluation by combining CNNs with recurrent layers. However, this method significantly enhanced video analysis by relying on plentiful high quality datasets but itwasnotefficientinauthenticating low resolution deepfakes.

5) S.MarcelandP.Korshunov(2020)

He compared human versus machine perfomance in detecting deepfakes, according to Korshunov and Marcel [5]. However, their findings showed that machine learning models are betteratidentifyingsubtle manipulations that are not seen by humans. But the study stressed that automated systems must learn to generalize well to new types ofdeepfakes,especiallyin real world applications.

6) V.Agarwal,R.Chugh,S.Subramanian,andK.R. Ramakrishnan (2021)

In Chugh et al. [6], a hybrid deepfake detection model consisting of CNNs and Transformer architectures is proposed. Xception addressed limitations of traditional CNNs where the extra temporal context cannot be captured, by utilizing Transformers to leverage long range dependencies. However, the detection accuracy significantly improved for long video sequences at the price of high computational cost.

7) A.Verma,R.Jain,andS.Mittal(2020)

Mittal et al. [7] find that transfer learning for deepfake detection can both reduce training time and increase detection accuracy by using pre-trained models. Their results showed the capability of transfer learning on scenarios with little training data. But adaptation to the new manipulation techniques uses a pre-trained model, thus reducing its ability and bias by the sources dataset in the pre-trained model.

8) M. Llewellyn, T. Hartley, T. S. Ip, M.AhmedandL. Stroebel (2023)

Stroebel et al. [12] performed a systematic review of deepfake detection techniques and distinguished between them as biological, spatial, and temporal approaches. Most generalization and adaptational problemsarehighlightedinthestudyandhybridmodels are needed to make generalization and adaptability in effective ways.

9) Y.Zhang,H.LiuandW.Wang(2020)

In [8], Wang et al. surveyed deeply a comprehensive survey of deepfake detection methods, classifying approaches directly based on CNNs,RNNs,GANs,and hybrid architectures. Finally, they analyze the strengths and weaknesses of these methods, especially to deal with real time applications and various data formats.

10) J.Lv,H.Song,Z.Yu,G.Yang,X.Luo(2020)

Luo et al. [9] proposed a CNN approach on combined facial and environmental stream context in deepfake videos. Dual focus improved detection accuracy by capturing inconsistencies in the background and facial features. Although effective, the approach is computationally expensive, which restricts its real time applicability.

11) Singh, S.; Kaushal, A.; Negi, S; and Chhaukar, S. (2022)

In their work, Kaushal et al. [13] compare CNN and Transformer based, and hybrid deepfake detection models, while revealing that the latter outperforms the previous in detecting more advanced manipulations. In addition, theydeterminedthatattentionmechanismscan improve detection precision.

12) M.Ma,C.Liu,Z.Lei,Y.Wen,andY.Yang(2022)

GMM-MobileNet architecture, combining Gaussian Mixture Models for preposing and MobileNet forfeature extraction was introduced by Wen et al.[14]. With the focus on inconsistencies in codec based applications this lightweight design has delivered significant efficiency in real time applications, butdoes restrict versatility.

13) Y.Li,H.JiangandX.Wang(2019)

CNNs are used by Wang et al. [10] to target compression related artifacts in low quality deepfake videos. For low bandwidth environments, their model worked well at detecting manipulations in compressed media. However, confidence in artifact based detection can suffer, particularly in high quality content.

14) N.Rosetti(2020)

In this thesis, I explore the use of Transformers and LSTMs in deepfake detection, where a frame based inconsistency and temporal misalignments are investigated by Rosetti [11]. Although computational complexity is a concern, the study showed that combining these architectures reduces detection accuracy and greatly increases accuracy in multi frame scenarios.

15) K.S.Charan,M.Abhineswari, and B. N. Shrikarti (2024)

In transfer learning for deepfake detection,theprevious work Abhineswari et al. [15] had noticed good results for fine grained artifacts with models like EfficientNet and ResNet. However, the use of pre-trained architectures revealed the demands of fine tuning to handle specific challenges in the dataset.

16) Singh,G;Guleria,K;Sharma,S(2024)

Singh et al. [16] alsosuggestafinetunedMobileNetV3 model for detecting subtle distortions in manipulated media. Since the model is lightweight, the model is fit fortheedgedevices,butitismainlyapplicabletoimage based analysis.

17) Raza,A.,Munir,K.andAlmutairi,M.(2022)
Toincreasethedeepfakedetectionreliability,Razaetal. used Ensemble learning techniques. While the approach showed higher computational overhead upon implementation, the accuracy was generally improved over a range of datasets.

18) H.S.Chen,S.Hu,S.You,andC.C.J.Kuo(2022)
Defakehop++ is a lightweight detection framework emphasizing computational efficiency which is presented by Chen et al.[18].Althoughitissuitablefor large scale deployment, its simplicity limits its effectivenessatidentifyingsophisticatedmanipulations.

III. PROPOSED METHODOLOGY

In this work, a complete deepfake detection modelbased MobileNet and Long Short-Term Memory (LSTM) networks is proposed. A hybrid approach coupling MobileNet's onboard spatialpatternextraction to LSTM's temporal dependency analysis in video sequences, affords a robust manipulation identification system based on deepfake.The followingstepsoutline the methodology in detail:

A. Data Collection and Preprocessing

A huge and diverse set of real and deepfakemediawas collected in order to create a robust, versatile model.The dataset encompasses different methodsofdeepfake generation and it generalizes them over various manipulation techniques.

1) Data Augmentation: Transformation of data set such as rotation,flipping,additionofnoise, cropping for enhancing dataset diversity to create a real world dataset and make them robust.
2) Frame Extraction and Normalization: Resizing frames of videos to have consistent resolution and fix pixel bit values to avoid skewing the data.
3) Balanced Dataset:Preventing bias by ensuring repeated distribution of real versus fake data over multiple deepfake techniques.

Preprocessing with these techniques provides the same consistent, well structured input dataset that is critical for training models and detecting.

B. . Feature ExtractionwithMobileNet

MobileNet, a lightweight CNN architecture, During individual frames, we use a lightweight CNN architecture, MobileNet, to efficiently extract spatial features. It then makes a prediction on whether the common cues (e.g. texture, facial alignment, lighting variation) imply manipulation. Key improvements include:

1) Pre-Trained Weights:Enhanced feature extractionusingtransferlearningusingweights pre trained on ImageNet dataset.
2) Batch Normalization:However, by usingbatch normalization layers, we can stabilize training and accelerate convergence.
3) Real-Time Capability:Enabling real time applicationswithsamedetectionprecisionwith MobileNet's fast processing speed.

These extracted spatial features constitute the input for subsequent temporal analysis used to detect inconsistencies in video sequences.

C. Temporal Analysis with LSTM

An LSTM network is used based on the MobileNet spatial features extracted from video frames to analyze the sequential feature relationships. It is shown that LSTM is able to capture temporal patterns such as motion and expressioninconsistencythatataredifficultto replicate in the deepfakes videos. The enhancements include:

1) Hyperparameter Tuning:Improving the temporal dependent analysis via optimization of the number of LSTM units, dropouts rates and activation functions.
2) Temporal Loss Function:Motivation for introducing a loss function for emphasizing sequential consistency andimprovingtemporal feature learning.

This integration allows addressing the manipulation detection over time, which is typically ignored with single shot analysis.

D. HybridModelTraining

Optimization abilities towards spatial and temporal patternofdetectionareusedtotraina MobileNet-LSTM hybrid model, over the preprocessed dataset.

Traininginvolves:

1) Learning Rate Scheduling:Convergence efficiently using ReduceLROnPlateau for learning rate adjustment dynamically.

2) Adversarial Training:To improve model robustness against sophisticated deepfake techniques, we exposethemodeltoadversarial examples.
3) Combined Loss Functions: We use a combination of binary cross entropy loss for classification and temporal loss for sequential analysis.

Our hybrid model is intended to be elastic to multiple deepfake approaches, both robust and versatile in detecting.

E. UserInterfaceDevelopment

For practical applications a user friendly interface is developed which permits a user to upload images or video or files to be analyzed.Key features include:

1) Real-Time Detection:Real-time deepfake analysis is supported for the interface displaying a detection score as well as visual explanations of identified anomalies.
2) Transparency and Usability:Media verification by identifying and highlighting manipulated regions within media targeting user trust and providing actionable insightsfor the media verifier.

As for the interface, it is designed to be accessible for non technical users which makes it apt at application areas, journalism, social media monitoring, law enforcement, etc.

F. ModelEvaluationandOptimization

Metrics like accuracy, precision, recall andF1scoreare usedtoevaluatetrainedmodelssothatit'sreliable.Key steps include:

1) Performance Validation:Adaptability and robustness on unseen deepfake techniques.
2) Hyperparameter Fine-Tuning:Change of MobileNet and LSTM parameters to achieve high performance.
3) Processing Speed Optimization:Architectural adjustment to achieve real time requirement for detection speed.

The optimizations made sure that the model remains stable and efficient in different operational situations.

G. DeploymentandReal-WorldTesting

We optimize the model for deployment on platforms from social media to news agencies and content verification services.Real-world testing involves:

1) Diverse Media Sources: Running the modelon media from different sources to validate its accuracy and speed under varied conditions.
2) User Feedback Integration: Tinkering the model and interface to get better usability and performance through feedback.

New deepfake detection methods are incorporated regularly, taking as an underlying assumption that the model must be regularly updated for the model to stay effective with evolving manipulation techniques.

H. ContinuousImprovementandUpdates

The model continues to be retrained periodically with the latest deepfake data in order to maintain effectiveness. Key strategies include:

1) Online Training: We propose to incorporate online learning to adapt to novel manipulation techniques without full retraining.
2) Multimodal Analysis: Improving the model's ability to analyze combined audio and visual data, inadditiontoexpandingitscapabilitiesto detect.
3) These improvements guarantee that the MobileNet-LSTM solution continues to be a powerful weapon against deepfakemediainan ever evolving digital space.

IV. RESULTS AND DISCUSSION

The performance of the proposed MobileNet-LSTM hybrid model for deepfake detection is compared to baseline models such as standalone MobileNet, standalone LSTM, and a CNN baseline. The proposed approach was evaluated based on accuracy, precision, recall and F1 score for robustness and reliability purposes. The results are presented quantitatively, graphicallyanalyzed,andfullydiscussedinthissection.

A. Quantitative Results

Table 1 gives a comparative performance of the MobileNet-lstm hybrid model against baseline models. Results show the hybrid approach achieves more accuracy, precision, recall, F1-score, compared to the existing policies..

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Standalone MobileNet	84.5	82.3	85.0	83.6
Standalone LSTM	80.2	78.9	80.0	79.4
CNNBaseline	83.1	81.5	82.0	81.7
MobileNet-LSTM(Proposed)	91.8	89.4	90.5	89.9

Table1:ComparativePerformanceMetricsofProposed Model and Baseline Models

B. *InterpretationofQuantitativeResults*

ResultsinTable1verifythattheinclusionofspatialand temporal analysis significantly improves deepfake detection performance.

1) High Accuracy:All baselinemodelsareoutperformed by the MobileNet-LSTM hybrid model with anaccuracy of 91.8. Thus, we combined spatial feature extraction with temporalsequenceanalysis,andshowed that this is effective for detecting deepfake manipulations.
2) Improved Precision and Recall:The hybrid model captures the dilemma between low false positives and false negatives by achieving identification of deepfake content with a precision of 89.4% and recall of 90.5%. Themodel'sreliability is validated by these metrics on practical problems such as media verification and security.
3) Balanced F1-Score:The F1 score of 89.9% provides evidence the model achieves an appropriate trade off between precision and recall to reliablydetecteventsin diverse test scenarios.

C. *GraphicalPerformanceAnalysis*

The further illustration of comparison performance can beseeninFig.2,inwhichtheaccuracy,precision,recall andF1 scorearedisplayedgraphicallyforMobileNet-LSTM and benchmark approaches.

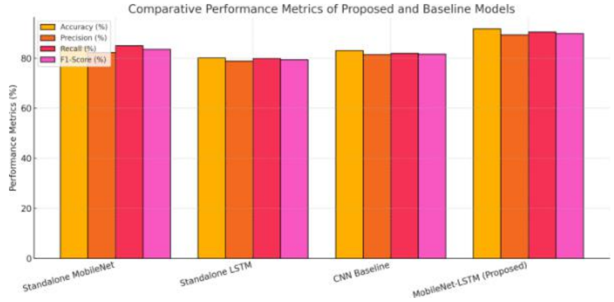


Fig. 2: Comparative Performance Metrics of Proposed and Baseline Models

We show that the MobileNetLSTM hybrid modelclearly outperforms all other models in all evaluation metrics. Through spatial and temporal analysis integration, it is evident that the enhancements are real as recall and F1 score are significantly improved.

D. Discussion

Experimental results demonstrate the significant improvements in deepfake detection achieved by the proposed MobileNet-LSTM hybrid model:

1) High Accuracy and Robustness:We achieved an accuracy of 91.8% which is a power improvementover the standalone MobileNet and LSTM model. Since we demonstraterobustnessofthecombined spatial-temporal analysis to a wide range of deepfake techniques and scenarios,wearehighlightingtheability of this approach to unearth such scenarios through verification.
2) Enhanced Precision and Recall:The model's ability to detect manipulations has precision and recall values of 89.4% and 90.5%, respectively. Thanks to higher recallreducingtheriskofmissingdeepfakecontent,and higher precision making fewer false alarms, which is important for public safety use casesandmediacontent authentication.
3) Generalized Performance with Balanced F1-Score: The good F1 score of 89.9% allows for a precisionand recall balanced model for data from different datasets and under different test conditions. The balanced performance of the model guaranteesitsapplicabilityto the real world scenarios.
4) Real-Time Applicability:MobileNet allows for real timedetectioncapabilitiesandcoupledwiththeeffortof the LSTM guarantees lightweight architecture, and hence, the system is feasible for applicationslikesocial media monitoring, journalism, or law enforcement.
5) Significance of Spatial-Temporal Integration:By fusing MobileNet for spatial feature extraction and LSTM for sequence level analysis, the model can find these frame level and sequence level inconsistencies.A limitation of the previous approach is that they only consider spatial or temporal features. This integration addresses those limitations.

V. CONCLUSION

This work introduces ahybridMobileNet-LSTMmodel resulting in a robust and effective solutionfordeepfake detection, leveraging the joint usage of spatial and temporal analysis. By virtue of the efficient extractionof fine grained spatial features by the MobileNet component and temporal patterns on sequential frames by the LSTM network, manipulations in the images as well as video are detected. The experimental results verify the superiority of the proposed model compared to standalone architectures and traditional CNN based methods that result in superior accuracy, precision, recall and F1-score.

Finally, we show the accuracy of theMobileNet-LSTM model is 91.8%, which is robust and adaptable to different deep fake techniques. The model achieves a precision of 89.4% and recall of 90.5%, with high reliability in termsofdetectingmanipulatedcontentand low rate of false positives and negatives.Thisresultsin balanced F1-score of 89.9% which demonstrates thatit generalizes well across all sorts of real world testing scenarios rendering it well suited to be used as part of numerous applications.

The proposed model is lightweight and enables realtime detection, supporting real time deepfake media detection requiring timely and accurate detection. This capability makes social media monitoring, digital forensics, journalism and public safety practical applications of this capability. The user-friendly interface also improves usability by providing non technical users an intuitive way to get detection results and visual explanations of detected anomalies helping build trust and transparency for business users.

Overall, the proposed hybrid MobileNet-LSTM model makes the contributions of removing the spatial and temporal limitations of deepfake detection systems by using spatial analysisandtemporalanalysis.Andowing to its real time capability, high performance and user friendly design, it is an appropriate tool to combat the persistent problem of deepfake media.Laterworksmay be aimed at incorporating multimodal inputs,e.g.audio and textual analysis, to further enhance detection accuracy and flexibility in dynamic digitalenvironments.

VI. FUTURE WORK

Thereby future work can improve the scalability, adaptability and robustness of mobile deepfakedetection model MobileNet LSTM on mobile model either on real world test cases. Pruning techniques and lightweight architectures would make large scale real time deployment in particular on the resource constrained environment of mobile devices. The future of deepfakes is, if we want to catch up with the technology, to include online training or somehow transfer learning to avoid a full restart every time the manipulation technique gets new. To increase detection accuracy even further, we can extend the model by adding multimodal detection by joint audio and visual analysis for consistency across channels, particularly when dealing with video content.

Apart from user trust and adoption, the model should ideally become more transparent and interpretable, through visualizationtechniquesfocusedonregionsthat have been manipulated or at least on temporal inconsistencies developed. But we cannot ignore thatwe can also make the model more robust to subtle manipulation attempts by also dealing with adversarial modelrobustnessthroughadversarialtraining.Such expanded testing of the model on diverse scenarios including livemediastreamswouldvalidatethemodel’s performance in beyond deriving test datasets, whichalso align with the claims of ethical and privacy concerns raised by journalism and law enforcement. ThesemodelswillcontinuetoadvanceMobileNet-LSTM will as a powerful tool for deepfake detection, for an ever evolving media landscape.

REFERENCES

[1] Chan, C. P. M., Lee, S. W., & See, J. (2021). DeepFakeVideoDetectionUsingDeepLearningand IrisDetection.2021IEEEInternational Conference on Image Processing (ICIP), Anchorage,AK, USA, pp. 1239–1243. doi: 10.1109/ICIP42928.2021.9506664.

[2] Dong,J., Wang, W.,Tang,Y.,Zhang,Y.,&Liu,H.(2021).Deepfake Video Detection Using Inception-ResNet and AttentionMechanism. 2021 IEEE International Conference onMultimediaandExpo (ICME), Shenzhen, China, pp. 1–6. doi:10.1109/ICME51207.2021.9428434.

[3] Husen, M. N., Kurniawan, A., & Pamungkas, M. (2020).Deepfake Detection on Video Sequences Using Inception-ResNet-v2and LSTM. 2020 3rd International Conference on IntelligentAutonomous Systems (ICoIAS), Singapore, pp. 134–138. doi:10.1109/ICoIAS49312.2020.9081914.

[4] Sabir, M., Cheng, J., Jaiswal, A., Wu, Y., Nataraj, L., &Chandrasekaran, S. (2020). Recurrent Convolutional Strategies forFace Manipulation Detection in Videos. 2020 IEEE/CVF Conferenceon Computer Vision and Pattern Recognition Workshops (CVPRW),Seattle, WA, USA, pp. 1643–1652. doi:10.1109/CVPRW50498.2020.00203.

[5] Korshunov, P., & Marcel, S. (2020). Human vs. Machine:Benchmarking Humans Against Deepfake Detection Systems. 20208th International Conference on Biometrics Theory,ApplicationsandSystems (BTAS), Washington, DC, USA, pp. 1–6. doi:10.1109/BTAS48898.2020.9528289.

[6] Chugh, R., Agarwal, V., Subramanian, S., &Ramakrishnan, K. R. (2021). Not Made ForEachOther— CombiningCNN and Transformers for Deepfake Detection. 2021 IEEE/CVFInternational Conference on Computer Vision (ICCV),Montreal,QC,Canada, pp. 15024–15033. doi: 10.1109/ICCV48922.2021.01514.

[7] Mittal, S., Verma, A.,&Jain,R.(2020).TransferLearningfor Deepfake Detection. 2020 6th InternationalConferenceonSignalProcessing and Communication (ICSC), Noida,India,pp.69–74.doi:10.1109/ICSC48311.2020.9182767.

[8] Wang, W., Zhang, Y., & Liu, H. (2020). Deep NeuralNetworks forDeepfakeDetection:ASurvey.2020IEEEInternationalConference on Artificial Intelligence and Knowledge Engineering(AIKE), Phuket, Thailand, pp. 290–296. doi:10.1109/AIKE48582.2020.00048.

[9] Luo,X.,Lv,J.,Song,H.,Yu,Z.,&Yang,G.(2020). Dual-Stream CNNs for Forgery Detection in DeepFake Videos.2020IEEE International Conference on Image Processing (ICIP), AbuDhabi, UAE, pp. 2556–2560. doi: 10.1109/ICIP40778.2020.9191035.

[10] Wang,X.,Li,Y.,&Jiang,H.(2019).FakeVideoDetectionwith Convolutional Neural Networks. 2019 10th InternationalConference on AdvancedComputationalIntelligence(ICACI),Guilin,China, pp. 182–186. doi: 10.1109/ICACI.2019.8778503.

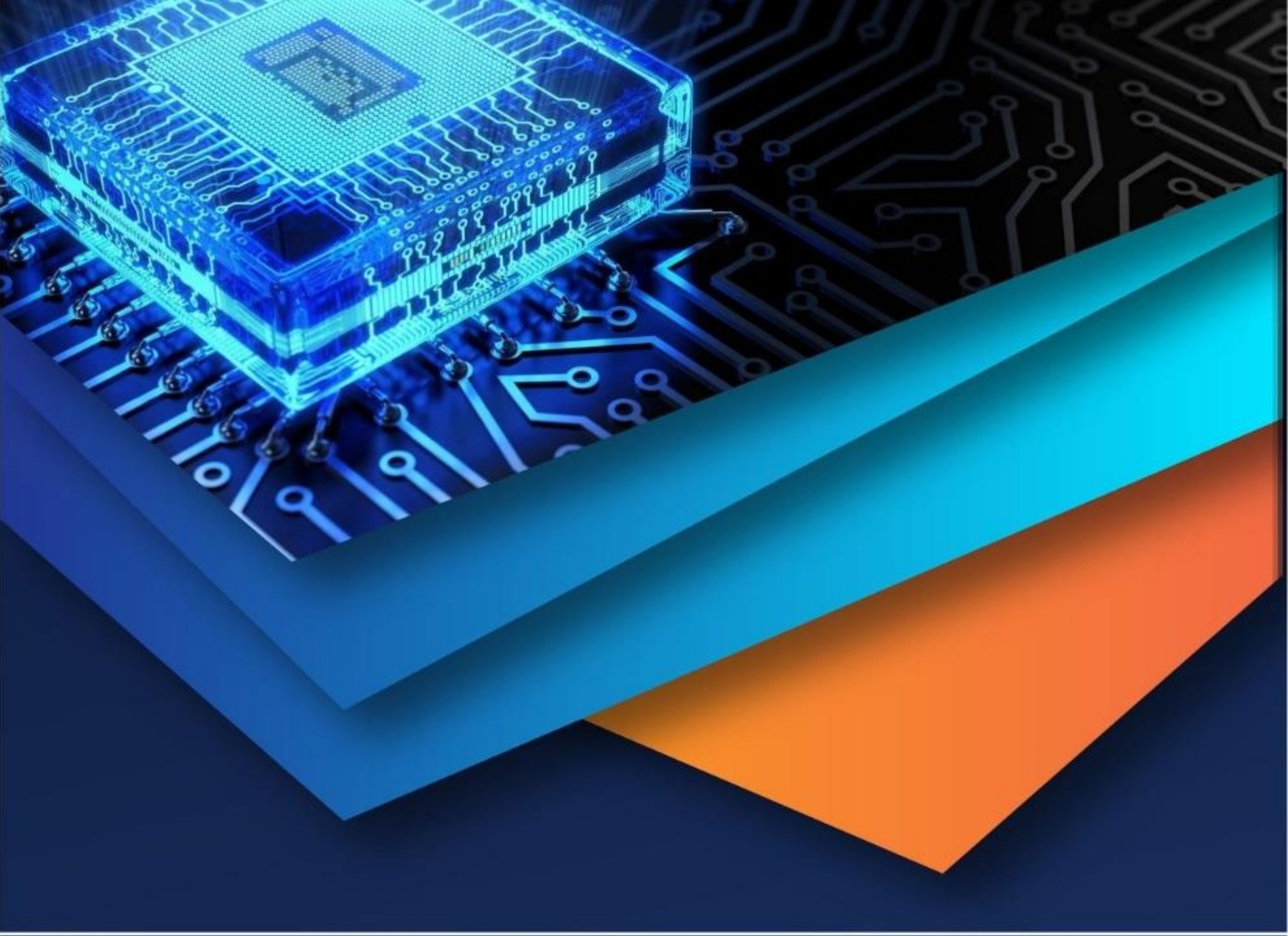
[11] Rosetti, N. (2020). Deepfake Detection Using LSTMs,Transformers and Video-Level Artifacts.

[12] Stroebel, L., Llewellyn, M., Hartley, T., Ip, T. S., &Ahmed, M. (2023). A Systematic Literature Review on theEffectiveness of Deepfake Detection Techniques. Journal of CyberSecurity Technology, 7(2), 83–113.

[13] Kaushal, A., Singh, S., Negi,S.,&Chhaukar,S.(2022).AComparative Study on DeepFake Detection Algorithms. 2022 4thInternational Conference onAdvancesinComputing,CommunicationControl and Networking (ICAC3N), pp. 854–860. IEEE.

[14] Wen,Y.,Lei,Z.,Yang,Y.,Liu,C.,&Ma,M.(2022). Multi-PathGMM-MobileNetBasedonAttackAlgorithmsandCodecsfor Synthetic Speech and Deepfake Detection. INTERSPEECH, pp.4795– 4799.

[15] Abhineswari, M., Charan, K.S.,&Shrikarti,B.N.(2024).Deep Fake Detection Using Transfer Learning:AComparativeStudyof Multiple Neural Networks. 2024 International Conference onSignal Processing, Computation, Electronics, Power andTelecommunication (IConSCEPT), pp. 1–6. IEEE.



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY