Enhanced Deepfake Detection Using Resnet50 And Facial Landmark Analysis

S.Pavithra,[1], R.Sridevi[2], Mahalakshmi .N[3],Devika R[4]
[1]Assistant Professor, Dept of IT
[2]Dept of IT
[1, 2] K S R Institute for Engineering and Technology, Tiruchengode - 637215

Abstract- *This research focuses on accuracy enhancement in the detection of deepfakes using the ResNet50 algorithm designed through deep learning. It analyzes anomalies in artificial facial images. Materials and Methods: The two implemented deep learning models include MobileNetV2 (Group 1) and ResNet50 (Group 2), each trained and tested with 40 image samples, comprising 20 real images and 20 deepfake images. Here, a facial irregularity detector based on ResNet50 was trained against one whose model was created through MobileNetV2. Result: ResNet50 was shown to have a detection accuracy of 91.81 % to 97.87 % for distinguishing between real and fake photographs. Its effectiveness for real-time applications is demonstrated. Statistical study revealed a significant improvement in detection accuracy than the MobileNetV2 Model (p-value < 0.05). Conclusion: According to the study's results, the ResNet50 algorithm is very good at identifying deepfake photos and real photos with a low mistake rate and high accuracy. Due to its efficiency in processing synthetic and genuine images, it can be a dependable tool for handling the problems created by deepfake media.*

Keywords: Deepfake Detection, ResNet50, Deep Learning, MobileNetV2, Detection Accuracy, Fake image, Dataset, Face.

I. INTRODUCTION

This research focuses on identifying deepfakes using Residual Neural Networks (ResNet50), leveraging their proven success in image recognition [1]. The study examines the detection performance of ResNet50 in distinguishing deepfake images by extracting subtle irregularities and inconsistencies introduced during their generation [2]. Deep learning advancements have improved methods of detecting deepfakes, and the project here is to improve this method using the robust and reliable architecture of ResNet50. Training ResNet50 on a large dataset of real and manipulated images/videos would enable the recognition of anomalies that distinguish deepfakes from authentic content [3]. The objective of this research is to use a deep learning-based ResNet50 architecture to develop a trustworthy deepfake detection system. The ResNet50 model will be trained on a large dataset of actual and changed images or videos, allowing the system to identify subtle irregularities and inconsistencies introduced throughout the deepfake manufacturing process. This research is important since deepfakes seriously jeopardize the legitimacy of digital information and affect journalism, media, and law enforcement [4]. The strategy will increase public trust in digital media and assist stop the spread of false information [5 - 8].

II. RELATED WORKS

The total number of articles published on this topic in the last five years, there have been more than 1954 publications on the topic in IEEE Xplore and 836 in Google Scholar. In contrast to conventional techniques, the combination of preprocessing steps, feature-based, residual connection, and batch normalization improves the detection accuracy of deepfake videos in the presence of facemasks. According to the study's findings, face-mask-deepfakes can be detected with 94.81 % accuracy when compared to the conventional InceptionResNet50V2 and VGG19, which have accuracy rates of 77.48 % respectively. Future research should assess the precision of creating a follow-up experimental study for improved deepfake detection using facemasks.

The user interface further displays a confidence score for each prediction to provide users an idea of the model's dependability. With training and testing accuracy rates over 95 %, this study provides a comprehensive technique to differentiate between real and fake faces [9]. The ability of deep learning to produce and identify deepfakes is still developing. deepfake detection models are created using older datasets, may become obsolete over time, and constantly need new detection methods. With an accuracy rate of over 90 %, the research findings are encouraging, but there is still room for improvement [10][11]. The CT outperforms the state-of-the-art in the deepfake detection task, demonstrating its better discriminative capacity and resilience to various forms of attacks. With an overall classification accuracy of over 97 % despite accounting for deepfakes from 10 different GAN architectures that are not only involved in face photos, the CT demonstrates its dependability and independence from image semantics. Finally, tests employing deepfakes created by

FACEAPP demonstrated the effectiveness of the proposed strategy in a real-world scenario, achieving 93 % accuracy in the fake detection task [12]. Additionally, in order to assess the resilience of Arabic-AD, Arabic recordings from non-Arabic speakers were gathered, taking into account the accent.

To evaluate the suggested approach and contrast it with established standards in the literature, three comprehensive experiments were carried out. With the lowest EER rate (0.027 %) and the highest detection accuracy (97 %) while avoiding the requirement for lengthy training, Arabic-AD thereby surpassed other state-of-the-art techniques [13]. The study's YOLO models all exhibited almost flawless accuracy in recognizing X-ray images that revealed osteoarthritis in the knee. Despite the generally excellent performance of the YoloV8 models, the YoloV5 models yielded the best and poorest results in lung CT scan images.The YoloV5su model had the highest recall (0.997), while the YoloV5nu model had the lowest (0.91). Additionally, YoloV5su, the best model, performs 60 % better than YoloV8x, the second-best model. The results demonstrate YoloV5su's speed and accuracy in detecting medical deepfakes [14][15]. The neural network will be trained extensively using these photos and their labels in order to accurately predict an image's legitimacy. The model has been trained on precisely 7104 photos, and based on the outcomes, the model's predictions are reasonably accurate. The accuracy and precision of the results are 88.81 % and 87.93 %, respectively. The suggested technique classifies the input video frames into genuine and deepfake by extracting features using a CNN-MLP model that is based on deep learning. The suggested approach detected deepfake movies with a high accuracy of 81.25 % using Celeb-df, a publicly available deepfake dataset. Comparative tests using a different algorithm show how accurate our model is, confirming its effectiveness in identifying modified content [16]. The feature extraction stage, which is the next and most important step, is when particular features relating to face orientation are taken out of every frame. A histogram of oriented gradients, scale-invariant feature transforms, and local binary patterns are the features that are extracted [17]. Normalization of features takes place in other neural network pipelines later in the design. The following neural network will then categorize the videos as either deepfake or not after combining the features. This architecture has a validation accuracy of 95.89 % [18].
Previous studies have shown that using MobileNetV2 methods to detect deepfake content might be challenging. An important factor in the rapidly changing realm of cybersecurity is effective deepfake detection. In contrast to traditional detection methods, the goal of this study is to improve detection accuracy by employing a unique deep learning-based ResNet50 algorithm.

III. MATERIALS AND METHODS

The study measures the performance of the ResNet50 model compared with that of MobileNetV2 in deepfake image detection based on predictive accuracy. The models are built and tested on a prepared dataset to verify real vs fake image classification [19]. The experiments were conducted in the IT Lab at KSR Institute for Engineering and Technology, where high-performance computing systems and necessary computational support are available for deep learning-based research. The dataset used for training and testing was obtained from Kaggle.com, specifically the Deepfake Detection Dataset. Statistical methods such as a 95 % confidence interval and a 0.05 % level of significance at a G-power value of 80 % were used to validate the results.
In this study, Group 1 is the MobileNetV2 model, which was trained and tested on 2178 deepfake and real images. The model, however, failed to predict images well, resulting in inconsistent performance in detection and obtained an accuracy level of 89 % [20]. Group 2 refers to the ResNet50 deep learning model, trained and tested on 2193 deepfake and real images. The aim is to improve detection accuracy through deeper feature extraction and enhanced classification capabilities. The ResNet50 model was found to have improved the detection of deepfake images by about 5 % to 7 % when compared to MobileNetV2.

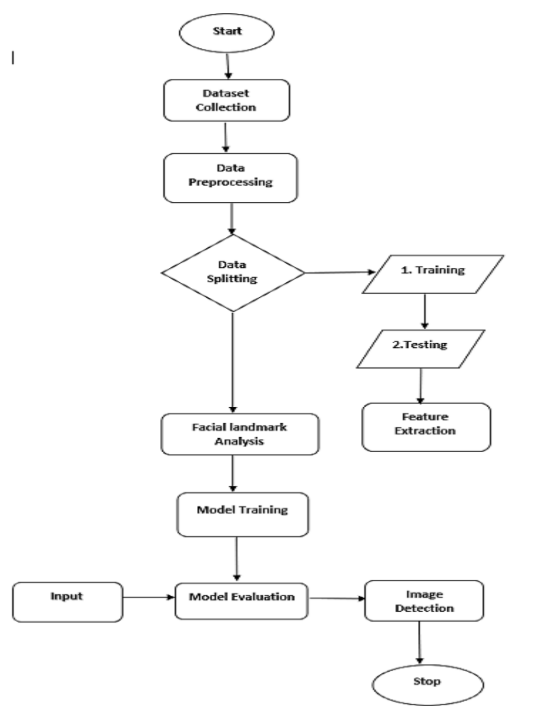


Fig. 1 The workflow of the deepfake detection process using images

The Deepfake detection system begins with the collection of the dataset from sources such as Kaggle, and preprocessing steps like resizing, normalization, and augmentation are performed on it. Then, the data is split into training and testing. The ResNet50 model learns deepfake-specific patterns during the training phase. Feature extraction and facial landmark analysis using dlib enhance classification accuracy. Metrics like accuracy, precision, recall, and F1-score are used to evaluate the trained model. During detection, the final phase, this model classifies images as being real or fake based on some confidence scores which ensure an efficient deepfake-detecting model.

IV. STATISTICAL ANALYSIS

The presentation of the deepfake recognition framework is analyzed by using SPSS adaptation 26 for a quantitative study. The review centers on two essential measurements: discovery precision and levels of certainty [21]. A one-tail free t-test was conducted to analyze the results. The ResNet model showedhigh dependability, accomplishing more than 97 % exactness for ensured pictures and 94 % accuracy for counterfeit ones. These outcomes highlight the design's capacity to see genuine and controlled pictures with surprising accuracy, making it a strong reaction for combating deepfakes.

V. RESULT

The results show that ResNet50 is more accurate, with accuracy ranging from 91.81 % to 97.87 %, while for the MobileNetV2 variants it was between 83.64 % and 89.56 %. It follows that the error for ResNet50 lies in the range of 2.13 % to 8.19 % compared to a much higher range for MobileNetV2, which is from 10.44 % to 13.71 %. Further statistical measures such as standard deviation, variance, and t-tests emphasize the significance of the differences. Graphical analysis further emphasizes the better nature of ResNet50 about it's always good performance.

The T-test comparison between the ResNet50 and MobileNetV2 models shows a high difference in the accuracy, with $p < 0.05$. ResNet50 has higher mean accuracy as 95.155 with low standard deviation 0.86375, and MobileNetV2 has mean accuracy as 87.557 with a high standard deviation of 1.01267. This indicates that although ResNet50 is more accurate, its consistency is slightly higher due to lower standard deviation. The graph of accuracy distribution also shows the comparative advantage of ResNet50: multiple iterations indicate higher accuracy in the case of the ResNet50 model. This again shows that the model proposed here is better and more consistent than MobileNetV2 in terms of deepfake detection.

TABLE 1. The model of mobileNetV2 realizes an accuracy of 86.37 % TO 89.56 %, AN ERROR RATE OF 10.87 TO 13.63.

ITERATION	ACCURACY	ERRORRATE
1	86.37	13.63
2	87.29	12.71
3	86.72	13.28
4	88.02	11.98
5	88.64	11.36
6	89.56	10.44
7	86.42	13.58
8	89.13	10.87
9	87.81	12.19
10	86.38	13.62
11	86.44	13.56
12	87.31	12.69
13	88.14	11.86
14	86.69	13.31
15	87.55	12.45

TABLE 2. The model of resnet50 realizes an accuracy of 91.81 % to 97.87 %, AN ERROR RATE OF 2.13 TO 7.65.

ITERATION	ACCURACY	ERROR RATE
1	94.24	5.76
2	95.41	4.59
3	96.05	3.95
4	97.12	2.88
5	96.80	3.20
6	97.35	2.65
7	95.85	4.15
8	97.87	2.13
9	96.30	3.70
10	94.75	5.25
11	96.51	3.49
12	91.81	8.19
13	92.48	7.52
14	94.03	5.97
15	92.35	7.65
16	94.92	5.08

17	93.49	6.51
18	94.95	5.05
19	94.63	5.37
20	95.46	6.54

TABLE 3. The mean accuracy of ResNet50 is 95.155 and MobileNetV2 has 87.557. Even though ResNet50 has a significantly lower variability or Std.Deviation of 0.86375 while MobileNetV2 has Std.Deviation of 1.01267, means only slightly high stability.

	Method	N	Mean	Std.deviation	Std.error mean
Accuracy	MobileNetV2	20	87.557	1.01267	0.22644
Accuracy	Resnet50	20	95.155	0.86375	0.15378

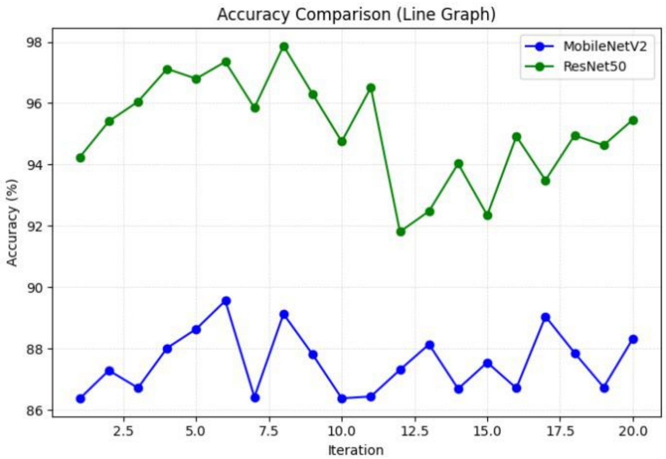
TABLE 4. From SPSS Independent samples test. T-test comparison of the gain in MobileNetV2 and ResNet50 ($p<0.05$)

		Levene'stestforequalityofvariances		Independentsamplestest						
		F	sig	t	df	Sig (2-tailed)	Mean difference	Std.errordifference	95% confidenceintervalofthedifference	
									lower	Upper
Accuracy	equal variance assumed	4.381	0.043	-16.761	38	0.000	-7.59800	0.45332	-8.51570	-6.68030
Accuracy	equal variances not assumed			-16.761	30.377	0.000	-7.59800	0.45332	-8.52332	-6.67268

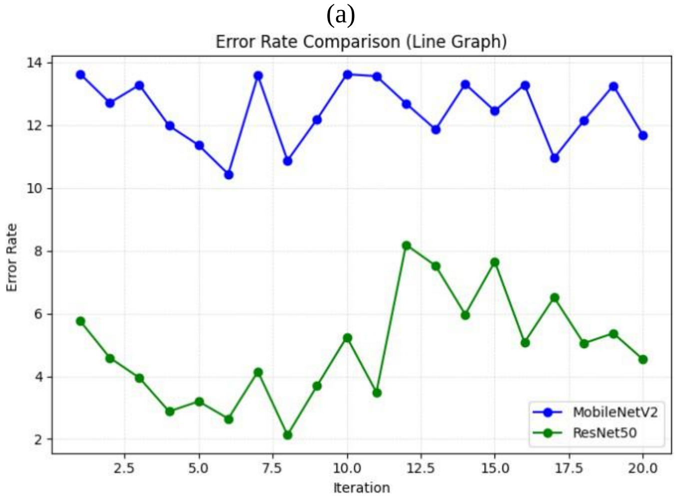
(a)

(b)

Fig. 2. The outcome with the Accuracy is displayed along with the result of the uploaded image, and it determines that the image is real with high Accuracy as illustrated in (a) and identifies the image as fake with strong Accuracy in the classification as shown in (b). The interface clearly displays the outcome of the classification to ensure an efficient process of deepfake detection.

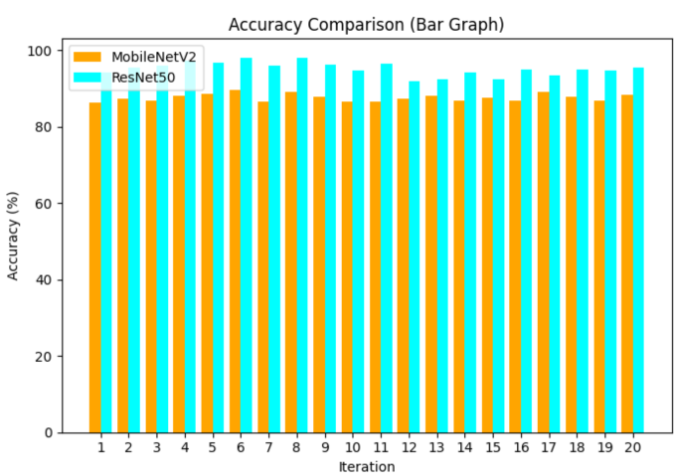


(a)

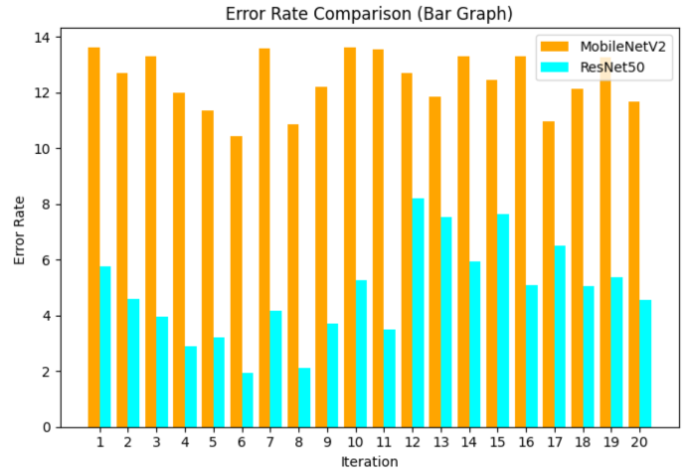


(b)

Fig. 3. The accuracy comparison between MobileNetV2 and ResNet50 is shown in Fig (a), while the error rate is depicted in Fig (b). The performance of ResNet50 was better compared to both the metrics, that is, with greater accuracy and maintaining a lesser error rate compared to MobileNetV2 throughout all iterations.



(a)



(b)

Fig. 4. Figure (a) contrasts the accuracy of MobileNetV2 and ResNet50 with a bar graph, and Figure (b) uses another bar graph to depict their error rates. In all iterations, ResNet50 outperforms MobileNetV2 by having greater accuracy and lower error rates, which makes them highly effective in detecting deepfakes.

VI. DISCUSSION

The present study showed how the ResNet50 algorithm provides a remarkable degree of improvement in detecting deepfake images compared to the MobileNetV2 model [22]. With accuracies between 91.81 and 97.87 % achieved with ResNet50, the model took precedence over MobileNetV2, showing a significance in difference with statistical confidence (p-value < 0.05) [23][24]. This emphasizes the potential of ResNet50 for accurately distinguishing between real and deepfake images, especially in real-time applications where accuracy becomes critical. With skip connections incorporated into a deep architecture, ResNet50 is capable of extracting rather complex features from facial images that sensitively detect small anomalies which might become possible red flags in deepfakes but invisible to standard detection methods [25]. The model can also learn hierarchical representations of features across many layers that further augment its power to identify variations in textures and contours and minor details that remain undetectable for standard economic functions. For example, slight differences in skin texture, eye movements, or maybe facial expressions are manipulations that happen in deep-fake content [26][27]. Slightest differences are what make ResNet50 bestcapable of doing deep-fake detection and, therefore, good at unveiling minute manipulations [28][29].

Deepfake generation methods, including facial swapping, expression manipulation, or synthesis-based facial animation, are all variable methods that will require a contrasting range of capabilities for accurate identification and distinguishable traits incorporated within them. ResNet50 scores high in feature extraction, which allows it to withstand variation in the described attributes [30]. Due to its great generalizability across varied datasets, it can distinguish deepfakes coming from different sources and formats, which is an exciting feature for real-world deployment. The extremely high accuracy shows that deep learning-based methods have prospects like ResNet50 [31]. Future work will extend to increase diversity of the dataset, enhance pre-processing, and boost the accuracy with ensemble learning. Attention mechanisms and adversarial training will improve robustness. Optimizations will minimize further computational overhead. Extracting both image and audio features for generating the model for detection shall increase this performance.

VII. CONCLUSION

The deepfake detection system was designed to train, and test for the performance in the differentiating between real and deepfake images utilizing the ResNet50 algorithm. The results showed that the ResNet50 model displayed significantly better than the performance of the MobileNetV2 model and demonstrated between 91.81 % to97.87 % detection accuracy, while MobileNetV2 achieved 86.37 % to 89.56 % accuracy. This signifies the wide margin for the improvement of the determines ResNet50 model's capability of detecting minute anomalies and manipulations in facial images and essential for high-accuracy applications and considering stability with the MobileNetV2 showing a standard deviation of 1.01267, indicating some variability in the performance across the different datasets. On the other hand, the lower standard deviation of 0.86375 for the ResNet50 model indicates the model had consistently good performance across varying datasets.

REFERENCES

[1] D. Zhu, C. Li, Y. Ao, Y. Zhang, and J. Xu, “Position detection of elements in off-axis three-mirror space optical system based on ResNet50 and LSTM,” Opt Express, vol. 33, no. 1, pp. 592–603, Jan. 2025.

[2] C. Yang, S. Ding, and G. Zhou, “Wind turbine blade damage detection based on acoustic signals,” Sci Rep, vol. 15, no. 1, p. 3930, Jan. 2025.

[3] S. Xue and C. Abhayaratne, “Region-of-Interest Aware 3D ResNet for Classification of COVID-19 Chest Computerised Tomography Scans.” Accessed: Feb. 03, 2025. [Online].Available: https://doi.org/10.1109/ACCESS.2023.3260632

[4] S. Song, J. C. K. Lam, Y. Han, and V. O. K. Li, “ResNet-LSTM for Real-Time PM2.5 and PM_{10} Estimation Using Sequential Smartphone Images.” Accessed: Feb. 03, 2025. [Online].Available: https://doi.org/10.1109/ACCESS.2020.3042278

[5] A. Qadir, R. Mahum, M. A. El-Meligy, A. E. Ragab, A. AlSalman, and M. Awais, “An efficient deepfake video detection using robust deep learning,” Heliyon, vol. 10, no. 5,p. e25757, Mar. 2024.

[6] R. Yang, K. You, C. Pang, X. Luo, and R. Lan, “CSTAN: A Deepfake Detection Network with CST Attention for Superior Generalization,” Sensors (Basel), vol. 24, no. 22, Nov. 2024, doi: 10.3390/s24227101.

[7] S. M. Hassan and A. K. Maji, “Pest Identification Based on Fusion of Self-Attention With ResNet.” Accessed: Feb. 03,2025.[Online].Available:https://doi.org/10.1109/ACCESS.202 4.3351003

[8] M. Saratha et al, “Research and Application of Boundary Optimization Algorithm of Forest Resource Vector Data Based on Convolutional Neural Network.” Accessed: Feb. 03, 2025. [Online].Available:https://doi.org/10.1109/SmartTechCon575 26.2023.10391589

[9] C. Wang, C. Shi, S. Wang, Z. Xia, and B. Ma, “Dual-Task Mutual Learning With QPHFM Watermarking for Deepfake Detection.” Accessed: Feb. 03, 2025. [Online]. Available: https://doi.org/10.1109/LSP.2024.3438101

[10] N. M. Alnaim, Z. M. Almutairi, M. S. Alsuwat, H. H. Alalawi,A. Alshobaili, and F. S. Alenezi, “DFFMD: A Deepfake Face Mask Dataset for Infectious Disease Era With Deepfake Detection Algorithms.” Accessed: Feb. 03, 2025. [Online].Available: https://doi.org/10.1109/ACCESS.2023.3246661

[11] S. Karthikeyan et al, “An attempt to enhance the time of reply for web service composition with QoS,” International Journal of Enterprise Network Management, Dec. 2020, Accessed: Feb.03,2025.[Online].Available:https://www.inderscienceonli ne.com/doi/10.1504/IJENM.2020.111750

[12] Y. Xu, P. Terhörst, M. Pedersen, and K. Raja, “Analyzing Fairness in Deepfake Detection With Massively Annotated Databases.” Accessed: Feb. 03, 2025. [Online]. Available: https://doi.org/10.1109/TTS.2024.3365421

[13] L. Zhou, C. Ma, Z. Wang, Y. Zhang, X. Shi, and L. Wu, “Robust Frame-Level Detection for Deepfake Videos With Lightweight Bayesian Inference Weighting.” Accessed: Feb. 03,2025.[Online].Available:https://doi.org/10.1109/JIOT.202 3.3337128

[14] T. Qiao, S. Xie, Y. Chen, F. Retraint, and X. Luo, “Fully Unsupervised Deepfake Video Detection Via Enhanced

ContrastiveLearning."Accessed:Feb.03,2025.[Online].Av ailab le:https://doi.org/10.1109/TPAMI.2024.3356814

[15] P. Jahnavi et al, "IOT based Innovative Irrigation using Adaptive Cuckoo Search Algorithm comparison with the State of Art Drip Irrigation to attain Efficient Irrigation." Accessed: Feb.03,2025.[Online].Available:https://doi.org/10.1109/ICTA CS56270.2022.9987981

[16] N. Saravanan et al, "Revolutionizing Air Quality Prognostication: Fusion of Deep Learning and Density-Based Spatial Clustering of Applications with Noise for Enhanced Pollution Prediction." Accessed: Feb. 03, 2025. [Online].Available: https://doi.org/10.1109/ICPCSN62568.2024.00029

[17] I. N. K. Wardana, "Design of mobile robot navigation controller using neuro-fuzzy logic system," Computers and Electrical Engineering, vol. 101, p. 108044, Jul. 2022.

[18] M. Venkatesan et al, "A New Data Hiding Scheme with Quality Control for Binary Images Using Block Parity." Accessed:Feb.03,2025.[Online].Available:https://doi.org/10.1 109/IAS.2007.26

[19] L. Pham, P. Lam, T. Nguyen, H. Nguyen, and A. Schindler, "Deepfake Audio Detection Using Spectrogram-based Feature and Ensemble of Deep Learning Models." Accessed: Feb. 03, 2025.[Online].Available:https://doi.org/10.1109/IS262782.20 24.10704095

[20] E. Şafak and N. Barışçı, "Detection of fake face images using lightweight convolutional neural networks with stacking ensemble learning method," PeerJ Comput Sci, vol. 10, p. e2103, Jun. 2024.

[21] K. Stehlik-Barry and A. J. Babinec, Data Analysis with IBM SPSS Statistics. Packt Publishing Ltd, 2017.

[22] L. Si et al., "A Novel Coal-Gangue Recognition Method for Top Coal Caving Face Based on IALO-VMD and Improved MobileNetV2 Network." Accessed: Feb. 03, 2025. [Online].Available: https://doi.org/10.1109/TIM.2023.3316250

[23] S. Khairnar, S. Gite, K. Mahajan, B. Pradhan, A. Alamri, andS. D. Thepade, "Advanced Techniques for Biometric Authentication: Leveraging Deep Learning and Explainable AI." Accessed: Feb. 03, 2025. [Online]. Available: https://doi.org/10.1109/ACCESS.2024.3474690

[24] P. Raghavendra Reddy et al, "Novel Detection of Forest Fire using Temperature and Carbon Dioxide Sensors with ImprovedAccuracy in Comparison between two Different Zones." Accessed: Feb. 03, 2025. [Online]. Available: https://doi.org/10.1109/ICIEM54221.2022.9853107

[25] Y. Mao, Y. Lv, G. Zhang, and X. Gui, "Exploring Transformer for Face Mask Detection." Accessed: Feb. 03, 2025. [Online].Available: https://doi.org/10.1109/ACCESS.2024.3449802

[26] M. C. Gursesli, S. Lombardi, M. Duradoni, L. Bocchi, A. Guazzini, and A. Lanata, "Facial Emotion Recognition (FER) Through Custom Lightweight CNN Model: Performance Evaluation in Public Datasets." Accessed: Feb. 03, 2025. [Online].Available:https://doi.org/10.1109/ACCESS.2024.338 0847

[27] A. V. Santhosh Babu et al, "Performance analysis on cluster-based intrusion detection techniques for energy efficient and secured data communication in MANET," International Journal of Information Systems and Change Management, Aug. 2019, Accessed: Feb. 03, 2025. [Online]. Available: https://www.inderscienceonline.com/doi/10.1504/IJISCM.201 9.101649

[28] M. Venkateswarlu and V. R. R. Ch, "DrowsyDetectNet: Driver Drowsiness Detection Using Lightweight CNN With Limited Training Data." Accessed: Feb. 03, 2025. [Online]. Available: https://doi.org/10.1109/ACCESS.2024.3440585

[29] N. Saravanan et al, "Accurate Prediction and Detection of Suicidal Risk using Random Forest Algorithm." Accessed: Feb. 03, 2025. [Online]. Available: https://doi.org/10.1109/ICPCSN62568.2024.00053

[30] H. Chen, G. Hu, Z. Lei, Y. Chen, N. M. Robertson, and S. Z. Li, "Attention-Based Two-Stream Convolutional Networks for Face Spoofing Detection." Accessed: Feb. 03, 2025. [Online].Available: https://doi.org/10.1109/TIFS.2019.2922241

[31] N. Gautam and D. K. Vishwakarma, "Obscenity Detection in Videos Through a Sequential ConvNet Pipeline Classifier." Accessed: Feb. 03, 2025. [Online]. Available: https://doi.org/10.1109/TCDS.2022.3158613