Detection and Verification of Counterfeit Images Using Stacked Ensemble Model.

ABSTRACT

The increasing use of advanced picture morphing technologies, such as deepfakes, has led to significant challenges in digital forensics, biometric authentication, and cybersecurity. Identity theft, false information, cyber extortion, and other fraudulent activities exploit morphed photos, making their detection a major challenge. Modern editing tools complexity causes traditional forensic techniques to sometimes miss these alterations.

The growing use of generative artificial intelligence technologies, such as large language models (LLMs) like ChatGPT and image tools like MidJourney, has made the risks of image morphing much higher. Though transforming creative and professional sectors, these technologies have also been misused to produce very believable false material, including manipulated photos, deepfakes, and synthetic identities. Recent incidents of abuse have drawn attention to their capacity for harmful acts, including identity theft, cyber extortion, false information campaigns, and biometric fraud. The two-edged character of generative artificial intelligence emphasizes the critical requirement of a strong machine learning model to efficiently identify and reduce such risks.

To address this issue, we suggest a stacked ensemble model of machine learning to more precisely identify morphed photos and solve this problem. This model improves detection accuracy by combining the best features of different deep learning systems—ResNet50, VGG16, and EfficientNetB0—taking advantage of what each one does well. The method finds variations in altered photographs using advanced AI-driven approaches like anomaly identification and metadata analysis. Thorough testing on real-world datasets, including both actual and modified photos, demonstrates the effectiveness of the suggested model in precisely identifying fake content. Though it offers benefits, issues continue to occur, including the need for high-quality training datasets, large computational power, and the fast-changing character of image morphing methods. Furthermore, the lack of unambiguous legal clauses in countries such as India to handle deepfake and picture morphing offenses makes enforcement even more difficult. Future initiatives should emphasize improving artificial intelligence models, enhancing dataset quality, and creating legal systems to control and punish illegal picture alteration. Addressing these problems helps the suggested ensemble approach to improve image authenticity verification and support digital security and confidence.

1. INTRODUCTION

Significant developments in image editing technologies have made it easier than ever to alter visual material in the digital age, hence highlighting important issues of privacy, security, and ethics. Among the more disturbing kinds of alteration is facial picture morphing, a technique that mixes characteristics from two or more photographs to produce a composite. Although sometimes employed for artistic or legal reasons, this technology is being more and more misused for cybercrimes, harmful deeds, and breaches of personal security. Facial image morphing endangers biometric systems, which rely on facial recognition for identity verification in vital applications including passports, driver's licenses, and controlled access systems. Morphed photos can trick these systems by combining identifiable traits of several people, hence causing illegal access, identity theft, or fraud.

The rapid proliferation of generative artificial intelligence technologies including picture generating tools like MidJourney and large language models (LLMs) such as ChatGPT has increased the threat of image morphing even further. Though they are changing the creative and commercial industries, same techniques have also been used to create quite believable counterfeit material including synthetic identities, deepfakes, and altered photographs. Recent cases of misuse have shown its ability for malicious intent including biometric fraud, cyber extortion, identity theft, and disinformation campaigns. The two-edged nature of generative artificial intelligence underlines how desperately robust detection systems are required to correctly identify and oppose such threats.

Recent statistics demonstrate unequivocally the rapid adoption of these artificial intelligence systems. ChatGPT, for instance, is the fastest-growing consumer app in history and has achieved before unheard-of expansion with about 100 million users by January 2023 [1]. By August 2023, ChatGPT had more than 180 million users [2]; this number kept rising. The platform showed regular and fast user engagement with almost 300 million weekly active users by December 2024 [3]. MidJourney has also shown significant growth in user base. MidJourney had gathered about 16 million total as of November 2023, with daily active users ranging from 1.2 million to 2.5 million [4]. Rising to 19.26 million by March 2024, registered user count reflected a growth of about 2.86 million in only four months [5]. The rapid proliferation of AI-driven content generation tools has lowered the barrier for producing sophisticated synthetic media, hence increasing the likelihood of its use in various hostile activities. These tools' widespread availability and user-friendly interfaces enable those with little technical understanding to generate credible false information, hence compounding the challenges faced by digital forensics and cybersecurity professionals.

The growing frequency of cybercrimes in India using morphed photographs is a clear warning of the dangers of this technology. Odisha recorded 273 instances in 2023 [6] where women's photos were changed and shared online mostly to shame and harass them. After their modified photograph was circulated around their contacts, a sextortion case [7] linked to a phony lending app regrettably resulted in a victim's death in Mumbai. Other cities have also seen disturbing occurrences of this kind: extortion schemes in Delhi [8], blackmail via modified photographs in Bengaluru [9], matrimonial frauds in Ahmedabad [10], revenge porn cases in Kolkata [11], and cyberstalking in Jaipur [12]. Likewise, Hyderabad and Chennai [13] noted several cases of social media manipulation and abuse of women's images for harassment. These incidents draw attention to the significant emotional, social, and reputational damage produced by morphing-related crimes, hence undermining confidence in digital networks.

Modern editing tools' improved capabilities, which may produce very realistic changes, are making it more and more challenging to spot altered photographs. Often unable to find these little changes, traditional forensic techniques call for the use of sophisticated artificial intelligence technologies. But creating successful AI-based systems has its own difficulties, such the absence of uniform datasets and assessment standards for training and evaluating detection systems. Adding to this problem is India's lack of unambiguous legal provisions to handle morphing-related offenses. Though the Information Technology Act, 2000, deals with broad cybercrimes, it does not specifically address morphing, so endangering victims and making legal action against criminals more difficult.

Dealing with the abuse of picture morphing calls for a multidisciplinary strategy spanning technical, legal, and ethical boundaries. It is crucial to create strong AI-powered detection technologies able to precisely identify modified pictures. Policy changes that create unambiguous legal frameworks to punish criminals and safeguard victims are just as crucial. Campaigns for digital literacy and public awareness can serve to promote ethical online conduct and aid to minimize the misuse of image alteration tools. A safer and more ethical cyberspace might be created by a thorough approach combining these components, which would enhance biometric security, protect individual privacy, and build confidence in digital environment.

We suggest a stacked ensemble model combining three strong deep learning architectures—ResNet50, VGG16, and EfficientNetB0—to address this problem. Every one of these models offers some benefits for the job of image forgery detection. Through its residual learning technique, ResNet50 excels at learning deep hierarchical features and solving vanishing gradient issues. VGG16 offers a straightforward but powerful design that catches fine-grained characteristics essential for spotting slight changes. Known for its better performance-to-efficiency ratio, EfficientNetB0 balances depth, width, and resolution to get great accuracy with less parameters. The ensemble gains from the complementing qualities of each by stacking these models and combining their forecasts, hence improving accuracy and robustness in detecting altered images relative to utilizing single models alone.

Against established benchmark datasets including a varied mix of both genuine and altered facial photos, the proposed stacking model has been exhaustively assessed. Compared to current state-of-the-art models, it has been shown to have significantly better performance measures in general accuracy. This good performance highlights its possibility for practical uses where consistent picture verification is essential. The ensemble method improves detection accuracy as well as resilience to various kinds of morphing techniques and image quality volatility. These findings confirm the efficacy of our model as a potential answer for addressing the rising danger of image morphing, thereby aiding initiatives to improve digital security, safeguard human identities, and reduce abuse in biometric and online systems.

2. BACKGROUND AND RELATED WORK

With how quickly digital technologies are changing, the level of sophistication used to make manipulated images and deepfakes is higher than ever before. This makes people very worried about how reliable visual material is. From subtly changed identity images to very convincing synthetic faces, these modifications use powerful algorithms to fool. Researchers have created forensic systems to distinguish between real and altered photographs in order to solve this increasing problem. Key developments, techniques, and continuing issues in this vital field of research are discussed in this part.

Ongoing studies on identifying altered or manipulated images—especially in relation to deepfake detection—have produced notable developments in several techniques and approaches. Still, difficulties remain, particularly with fresh kinds of image alteration. The work of Miki Tanaka, Sayaka Shiota, and Hitoshi Kiya, which suggested a technique using strong hashing to identify altered photos [14], illustrates the conventional method of employing hash-based comparisons for modification detection. Although their approach can detect tampering even under compression and resizing, it struggles to keep up with developing and more complex modification techniques, such as those created by Generative Adversarial Networks (GANs), which are designed to bypass conventional detection approaches. Focusing on the increasing relevance of multimodal systems, Dilip Kumar Sharma and Bhuvanesh Singh's assessment on forensic, machine learning, and deep learning methods for fraudulent picture identification on social media [15] offers a thorough review of the state of the art. Such systems enhance the accuracy of fake identification by combining several kinds of data, including optical, aural, and contextual information. Their study, however, lacks useful implementation plans that may be used in actual social media environments where crucial success variables are the speed of detection and the model adaptability.

Chandra Bhushana Rao Killi, Narayanan Balakrishnan, and Chinta Someswara Rao investigated VGG-19 for categorizing deepfake photos in the deep learning domain, obtaining 96% accuracy by using regularization methods to avoid overfitting [16]. But the ambiguity on dataset variety causes questions about how well this model can generalize to unknown data. Work by Chih-Chung Hsu, Yi-Xiu Zhuang, and Chia-Yen Lee on a Channel Feature Fusion Network (CFFN) [17] showed better accuracy and recall in identifying GAN-generated images using a paired learning technique. Their assessment, however, is restricted to a smaller range of GANs, which renders their approach less resilient to more recent generations of GANs.

The identification of picture modifications has also been influenced by human visual cognition. Giuseppe Cartella, Vittorio Cuculo, Marcella Cornia, and Rita Cucchiara's work concentrated on gaze patterns [18] to distinguish genuine from false photographs, hence presenting a novel approach grounded on human behavior. Although this approach has promise, its main drawback is the dependence on manual gaze data gathering, which significantly limits scalability and complicates application in automated or large-scale systems. Huang and Juefei-Xu likewise created FakeLocator [19], which localizes alterations using grayscale fakeness maps. Although this approach is quite accurate in spotting altered areas, it loses efficacy under severe compression of images, which is usually the case with social media photos that are compressed before being uploaded.

Gaining popularity in recent years for several image-processing chores, transformers have also been modified for synthetic picture identification. FatFormer by Huan Liu1, Zichang Tan and Chuangchuang Tan, a forgery-aware adaptive transformer [20], performs remarkably well in spotting both GAN and diffusion model-generated images. Its significant processing cost, which could render real-time identification in practical uses like video conversations or live streaming impossible, however, offsets the model's accuracy. Likewise, Tessa R. Flack and Kay L. Ritchie's study on face morph identification using cooperative training [21] makes use of shared characteristics between paired photos, hence enhancing the model's capacity to identify face morphing alterations. Though the concept works well in controlled settings, its scalability in actual applications like border control systems still unproven.

Video-based modifications have also been investigated using compact neural networks like MesoNet created by Darius Afchar, Vincent Nozick, Junichi Yamagishi and Isao Echizen [22]. Offering an efficient approach for video-based fake detection, MesoNet, a lightweight architecture, surpasses 98% detection rates for Deepfake films. Though successful, the concept has major drawbacks with very compressed video material, which is typical on video-sharing sites where compression techniques lower file sizes. In a related setting, Luca Bondi1 and Silvia Lameri suggested a tampering detection technique that groups CNN-based camera features [23] to find altered areas, hence offering a strong tool for spotting video alterations. The approach, however, has trouble generalizing to films shot with unknown camera models, hence stressing a drawback of CNN-based techniques operating across different camera kinds. Another line of study has been frequency domain analysis.

Work by Muhammad S. Mandisha and Mohamed A. Hussien, which integrated wavelet transforms with gradient boosting for tampered picture identification [24], surpasses conventional Fourier-based techniques, with a high accuracy of 95%. Applied to hostile modified images, however, it runs into problems since such images could purposefully change to evade detection by frequency-based techniques. Patch-based classifiers [25] by Lucy Chai, David Bau, Ser-Nam Lim, and Phillip Isola further added to this field by showing better generalization over several datasets. These classifiers, however, like Muhammad S. Mandisha and Mohamed A. Hussien's approach, have difficulty identifying photos altered by fine-tuned hostile generators.

El-Sayed Atlam and Malik Almaliki's methodical literature study identified research gaps in the deepfake dynamics [26] and supported multidisciplinary cooperation in tackling the problem. Although their study is thorough, it lacks empirical tests of the suggested remedies, hence hindering the ability to assess the practical effect of their suggestions. Combining Error Level Analysis (ELA) with Convolutional Neural Networks (CNNs), Ramesh Gorle and Anitha Guttavelli [27] attained 96.21% accuracy in identifying altered photos. Though their approach is accurate, it is hampered by the small size of assessed datasets, which questions the generalizability of the model. Significant interest has also been drawn by face morphing detection.

Particularly for spotting morphing attacks in border control situations, Chalini G R and K. V. Kanimozhi suggested Differential Morphing Attack Detection (D-MAD) [28]. Its actual applicability is therefore limited since its efficacy over several datasets is still unknown, which calls into question its usefulness. Though A. Ramesh, Bheema Sri Lakshmi, Dasari Narendar, Moinuddin Mohammed Najeeb, and Vudaru Sai used pre-trained deep learning models such as VGG19 for morphing attack detection [29], the absence of consistent benchmarks in their work makes it difficult to contrast their method with others.

Padmaja Kadiri and Palagati Anusha's investigation of machine learning algorithms such as Support Vector Machines (SVM) and Photo Response Non-Uniformity (PRNU) for identifying morphed photos [30] underlines the need of varied datasets since their method battles with real-world variation in image manipulations. Integrating Generative Adversarial Networks with Convolutional Neural Networks, Sharma, P., Kumar, M., and Sharma, H.K. created a deepfake detection system known as GAN-CNN Ensemble [31]. Their main idea is to let the model keep previous information while learning from new input by means of generative replay, hence reducing catastrophic forgetting. Strong accuracy in spotting fake photos on social media was shown by the ensemble, hence stressing its resilience and adaptation to changing deepfake technologies.

Focusing on its threat to facial recognition systems, Agarwal, A. and Ratha, N. researched on identifying face morphing in social media content [32]. Their deep learning-based detection process can efficiently distinguish between actual and altered photos. Their approach was successful under different circumstances; therefore, it is a useful tool for digital content control and identity verification. Designed for interpretable deepfake detection, Dua, F., Yu, M., Li, B., Chow, K.P., Jiang, J., Zhang, Y., Liang, Y., Li, M., and Huang, W. presented TAENet, a Two-branch Autoencoder Network [33]. The detection approach is clearer since one branch detects abnormalities and the other reconstructs photos. For legal and forensic uses, their model not only attained great accuracy but also enhanced explainability.

An explainable AI-based ensemble model for detecting altered facial images was proposed by Dwivedi, R., Kothari, P., Chopra, D., Singh, M., and Kumar, R. [34]. Their method provides both great accuracy and obvious visual or analytical reasons for each decision by integrating many classifiers with explainability tools. Their concept is particularly beneficial in sensitive situations when knowledge of the detection process is critical. Matakias, A. and Sotiropoulos, D.N. undertook a comparative analysis of face morphing algorithms [35], assessing different approaches depending on realism, efficacy, and their capacity to mislead biometric systems. Their results enable security professionals and researchers know weaknesses in present facial recognition systems and benchmark next detection techniques by highlighting which algorithms present the most danger. Tassone, F., Maiano, L., and Amerini, I. created a continuous false media detection system [36] that fits to new generative technologies. By allowing deepfake detectors to constantly learn and update, their method addresses the issue of model deterioration over time. This flexibility guarantees ongoing performance against new kinds of bogus media, hence addressing the rapid evolution of generative AI.

For improved clarity and comprehension, the table 1 below offers a comparison of the literature studies mentioned above.

Author and Year	Methodology	Focus	Datasets	Key Findings
Luca Bondi, Silvia Lameri, David Guera, Paolo Bestagini, Edward J. Delp, Stefano Tubaro (2017)	CNN feature extraction and clustering	Tampering detection based on camera model features	Dresden Image Database	High accuracy in detecting tampering, even with unknown camera models; effective localization of forged regions.
Darius Afchar, Vincent Nozick, Junichi Yamagishi, Isao Echizen (2018)	Deep learning networks with low layer count	Detection of face tampering in videos	Custom dataset (175 forged videos), FaceForensics dataset	Achieved over 98% detection rate for Deepfake and 95% for Face2Face; challenges noted with video compression.
Chih-Chung Hsu, Yi-Xiu Zhuang and Chia-Yen Lee (2020)	Pairwise learning with Common Fake Feature Network (CFFN)	Detecting GAN-generated images	CelebA (202,599 aligned face images, plus GAN-generated fake images)	Outperformed existing methods in precision and recall. The approach generalizes well to new GAN-generated images.
Lucy Chai, David Bau, Ser-Nam Lim, and Phillip Isola (2020)	Patch-based classifier with limited receptive fields	Identifying detectable artifacts in fake images	CelebA-HQ, FFHQ, FaceForensics	Patch-based classifiers outperform full-image classifiers for out-of-domain synthetic images; even adversarially fine-tuned

Author and Year	Methodology	Focus	Datasets	Key Findings
				image generators leave detectable artifacts.
Miki Tanaka, Sayaka Shiota and Hitoshi Kiya (2021)	Robust hashing for feature extraction	Detecting fake images under compression and resizing	Image Manipulation Dataset, UADFV, CycleGAN, StarGAN	Effective detection of tampered images using hash values from reference and query images; superior to state-of-the-art methods. Applicable for monitoring unauthorized manipulation.
Yihao Huang, Felix Juefei-Xu, Qing Guo, Yang Liu, Geguang Pu (2021)	Localization using gray-scale fakeness maps	GAN-based face manipulation localization	CelebA, FFHQ	High localization accuracy; robust to degradations like JPEG compression. Integrated attention mechanism improves universality.
Dilip Kumar Sharma, Bhuvanesh Singh, Saurabh Agarwal, Lalit Garg, Cheonshik Kim and Ki-Hyun Jung (2023)	Review of forensic, ML, and deep learning methods	Social media misinformation and fake image detection	PGGAN, FNC, Mediaeval, CASIA, CelebA	Highlights deep learning's effectiveness in fake image detection and emphasizes the need for multimodal approaches and interdisciplinary research.
Chandra Bhushana Rao Killi, Narayanan Balakrishnan, Chinta Someswara Rao (2023)	Deep learning (VGG-19)	Classification of deepfake images	Large dataset of natural images (specific datasets not detailed)	Achieved 96% accuracy; highlights the effectiveness of VGG-19 with regularization techniques like Dropout and Batch Normalization.
Huan Liu, Zichang Tan, Chuangchuang Tan, Yunchao Wei1, Yao Zhao, Jingdong Wang (2023)	FatFormer Transformer with forgery-aware adapter	Detection of GAN and diffusion-based fake images	ProGAN, StyleGAN, BigGAN, CycleGAN, StarGAN, GauGAN, DALL-E, LDM, Glide	Achieved 98.4% accuracy for GANs and 95% for diffusion models; outperformed existing methods in detection generalizability.
Tessa R. Flack, Kay L. Ritchie, Charlotte Cartledge, Elizabeth A. Fuller, Robin S. S. Kramer (2023)	Collaborative training experiments	Face morph detection	Data from experiments with 166 participants	Pairs training effect significantly improves detection accuracy; real-world application potential in enhancing border control.
Giuseppe Cartella, Vittorio Cuculo, Marcella Cornia, Rita Cucchiara (2024)	Eye-tracking experiments	Human visual perception of real vs. fake images	COCO, ADE20K, LHQ	Identified distinct gaze patterns for real vs. fake images; potential to enhance fake detection through

Author and Year	Methodology	Focus	Datasets	Key Findings
				semantic knowledge and gaze behavior.
Muhammad S. Mandisha, Mohamed A. Hussien, Amr K. Shalaby, Omar M. Fahmy (2024)	Wavelet transform with gradient boosting classifier	Detecting manipulated images	FaceForensics	Achieved 95% accuracy, outperforming Fourier transform-based methods; computationally efficient.
Ramesh Gorle, Anitha Guttavelli (2024)	ELA combined with CNN	Image tampering detection	CASIA v2.0	Achieved 96.21% detection accuracy, surpassing established models like VGG16, VGG19, and ResNet101; highlights the effectiveness of integrating ELA with CNN.
Chalini G R, K. V. Kanimozhi, (2024)	Differential Image-Based Morphing Attack Detection (D-MAD)	Detection of face morphing attacks	Custom dataset with 143 data subjects, public facial datasets	D-MAD demonstrates strong performance in border control scenarios, surpassing traditional methods; effective across multiple datasets.
Mr. A. Ramesh, Bheema Sri Lakshmi, Dasari Narendar, Moinuddin Mohammed Najeeb, Vudaru Sai (2024)	Deep convolutional neural networks	Face morph detection	Datasets not specified	Pretrained networks like VGG19 outperform those trained from scratch; emphasizes the need for standardized benchmarks and robust detection algorithms.
Padmaja Kadiri, Palagati Anusha, Madhav Prabhu, Rolito Asuncion, Voonna Sainath Pavan, Jami Venkata Suman (2024)	Machine learning (SVM, PRNU analysis, deep learning)	Detection of morphed images	Custom dataset for morphed images	Highlights vulnerabilities in face recognition systems to morphing attacks; emphasizes the importance of diverse datasets for improving detection accuracy.
Sharma, P., Kumar, M., and Sharma, H.K. (2024)	GAN-CNN Ensemble with generative replay	Deepfake detection on social media	Not specified	Improved accuracy and retention through minimized catastrophic forgetting.
Agarwal, A. and Ratha, N. (2024)	Deep learning detection pipeline	Face morphing detection in social media	Not specified	High performance in distinguishing real and morphed images.
Dua, F. et al. (2024)	Two-branch Autoencoder	Interpretable deepfake detection	Not specified	Achieved high accuracy and model transparency through anomaly reconstruction.

Author and Year	Methodology	Focus	Datasets	Key Findings
	Network (TAENet)			
Dwivedi, R. et al. (2024)	Explainable AI (XAI)-based ensemble approach	Morphed face detection	Not specified	Delivered high detection accuracy with explainable model outputs.
Matakias, A. and Sotiropoulos, D.N. (2024)	Comparative analysis of morphing algorithms	Evaluation of morphing attack risks	Custom test cases	Identified the most effective and deceptive morphing techniques.
Tassone, F., Maiano, L., and Amerini, I. (2024)	Adaptive learning framework for continuous detection	Deepfake adaptability to new methods	Not specified	Enabled deepfake detectors to remain effective over time through continuous learning.
El-Sayed Atlam, Malik Almaliki, Ghada Elmarhomy, Abdulqader M. Almars, Awatif M.A. Elsiddieg, Rasha ElAgamy (2025)	Systematic literature review	Research focus on detection, dynamics, and prevention of deepfakes	No datasets specified	Identifies gaps in research on deepfake dynamics and prevention; advocates for multidisciplinary collaboration and digital intervention strategies to combat misinformation.

Table.1. Comparison table of all related/background work.

Though photo modification detection has improved, questions about its dependability and usefulness still exist. Especially under challenging real-world conditions involving several manipulations, compression artifacts, noise, and low-resolution inputs, many algorithms cannot detect minor or localized changes. Many current detection methods are insensitive; subtle or complex editing techniques can help to escape recognition. Although deep learning-based methods have increased detection accuracy, they call for large, well-annotated datasets. This reliance restricts their generalizability and reactivity to creative or underrepresented modification techniques outside the training data. It is also quite challenging to properly localize changed areas. Though many models can identify image modification, they cannot identify the impacted areas, so their use in situations requiring thorough image verification is constrained.

Promising are multi-modal identification techniques combining visual material with contextual information such metadata or text. But the scalability and effectiveness of such systems in different real-world settings still need to be evaluated. Sophisticated model interpretability is also a significant concern. Many modern technologies are opaque "black boxes," which makes them inappropriate for sensitive or high-stakes uses needing openness and explainability. Human-centered techniques such as gaze tracking and behavioural signal analysis can enhance manipulation detection. These techniques, meanwhile, require more research to assess their viability, user acceptance, and integration into workflows. At last, generalization and cross-platform adaption are still challenges. Successful detection models on some datasets often show inconsistent behavior across datasets or manipulation types, suggesting a lack of robustness in several operating conditions. These limitations highlight the demand for scalable, interpretable, and flexible models capable of identifying altered material across platforms, manipulation techniques, and real-world environments. Image verification systems have to handle these issues if they are to be consistent and trustworthy in an ever more complicated digital environment.

3. METHODOLOGY

By using the strength of several convolutional neural network architectures integrated via a stacking ensemble learning technique, we provide in this paper a strong deep learning-based method to identify altered facial photos. Our approach consists of two main phases: (1) independent training of three well-known CNN models—ResNet50, VGG16, and EfficientNet-B0—and (2) the building of a stacking ensemble model combining their predictive outputs to enhance classification performance and generalizability.

Originally, our selected dataset of real and altered facial photos trained each of the three CNN architectures separately. Using models pre-trained on the ImageNet dataset, we applied transfer learning to adjust their last classification layers to fit our binary classification goal—distinguishing between original and altered images. The choice of several CNNs is motivated by their particular architectural benefits and complimentary feature extraction capabilities. Renowned for its deep residual learning architecture, ResNet50 reduces the vanishing gradient problem and allows efficient learning in deeper layers. Its residual blocks let the model capture fine-grained high-level semantic information, which are absolutely vital for spotting slight pixel-level changes in morphing images. Conversely, VGG16 is a relatively straightforward design that employs homogeneous layer layouts and tiny convolution filters. Though it is basic, VGG16 is good at catching spatial and texture-related characteristics, which can sometimes separate actual from artificial facial components. Finally, employing a compound coefficient, EfficientNet-B0, the baseline model in the EfficientNet family, scales network dimensions (depth, width, and resolution). Optimal between accuracy and computing efficiency, this model is fit for resource-constrained or real-time uses.

To preserve consistency across experiments, each model was fine-tuned using same training, validation, and test splits. Extensive hyperparameter tuning—including learning rate scheduling, early halting, and regularization approaches like dropout and weight decay to prevent overfitting—characterized the training process. To evaluate each model's efficacy in spotting manipulated photos, performance measures including accuracy, precision, recall, and F1-score were tracked during training.

3.1 ResNet50

ResNet50 popularized residual learning, thereby changing the training of very deep architectures using a 50-layer deep convolutional neural network. Its key innovations include skip connections, or identity shortcuts, which let backpropagation stop gradients from disappearing and enable them run directly over the network. These residual connections allow the model to focus on learning residual functions and learn identity mappings, therefore effectively addressing the degradation problem observed in deep networks. Made up of a series of bottleneck blocks, ResNet50 allows the model to be deep yet efficient using three layers: 1×1, 3×3, and again 1×1 convolution as shown in figure 1.

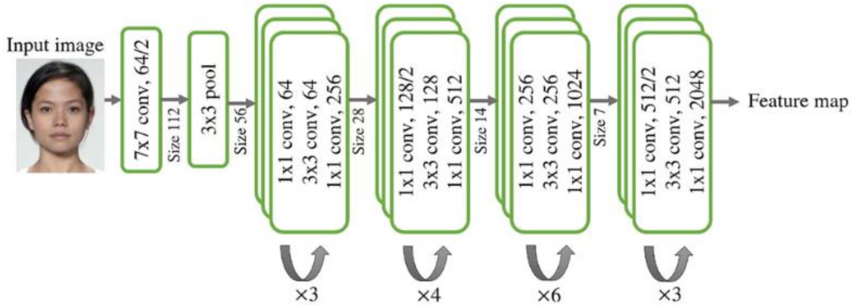

Fig.1. Block Diagram for ResNet50 model.

Given its capture of both high-level facial features and low-level texture patterns, ResNet50 is rather crucial in the framework of facial morphing detection. Its ability to learn complex representations without suffering overfitting or vanishing gradients makes it particularly successful for applications involving fine-grained manipulation. ResNet50's depth and architectural efficiency allow it to investigate tiny pixel-level variations and structural misalignments brought on by morphing processes. Furthermore, especially under a broad range of morphing techniques and image quality, its high generalization qualities make it a consistent basis model in our ensemble architecture.

3.2 VGG16

VGG16 is a seminal deep convolutional neural network architecture developed by the Visual Geometry Group at the University of Oxford. It emphasizes depth by using a stack of small (3×3) convolutional filters across 13 convolutional layers, followed by three fully connected layers. Max-pooling layers (2×2) are interspersed between convolutional blocks to reduce the spatial dimensions of the feature maps while retaining the most prominent features. ReLU (Rectified Linear Unit) activations follow each convolutional layer, introducing non-linearity and improving convergence during training. The final layer is a softmax classifier that outputs the probabilities for each class.

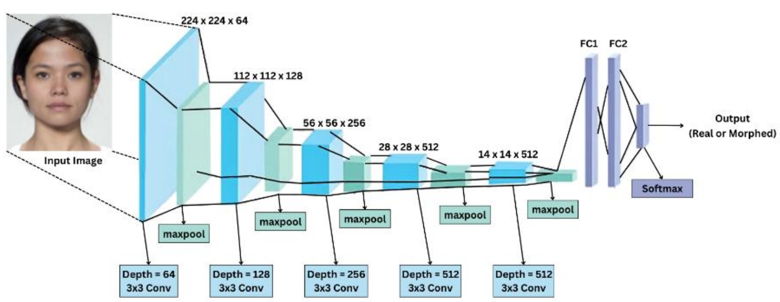

Fig.3. Block Diagram for VGG16 model.

In our proposed model, VGG16 is employed as a robust feature extractor for detecting morphed facial images. Its deep architecture allows the model to learn intricate patterns and hierarchical facial features that may not be visually apparent, especially in cases where facial image morphing is subtle and sophisticated. The model excels in identifying spatial inconsistencies and texture-based anomalies that result from blending different facial regions. Although VGG16 is relatively heavy in terms of computational load, its high accuracy in visual tasks and feature richness make it a valuable component of our stacked ensemble approach, helping improve overall performance by capturing deep visual semantics.

3.3 EfficientNet

EfficientNet is a state-of-the-art convolutional neural network that balances accuracy and processing economy using a novel compound scaling approach. This approach uniformly distributes the depth, width, and resolution of the network using specified coefficients, hence maximizing model performance with fewer parameters and lower FLOPs. Built on mobile inverted bottleneck convolutional blocks (MBConv) and enhanced with squeeze-and-excitation (SE) modules as shown in figure 3, EfficientNet captures channel-wise dependencies and adaptively recalibrates feature maps to emphasize the most valuable features.

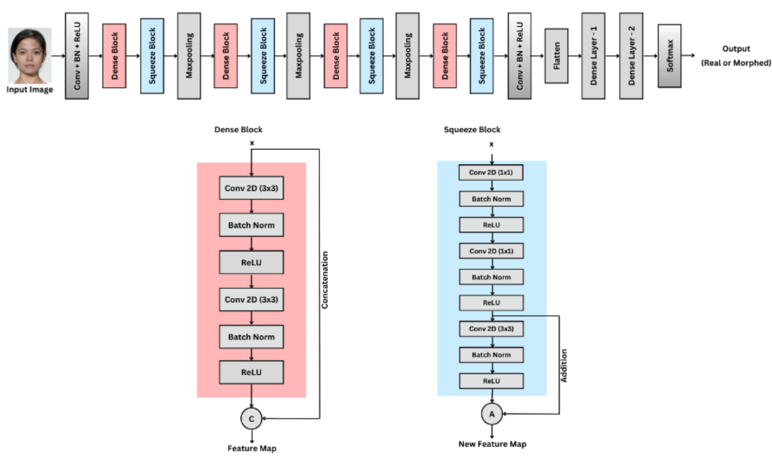

Fig.3. Block Diagram for EfficientNet model.

Our morphing detection algorithm makes use of EfficientNet to fairly low weight structure extract powerful multi-scale facial features. Its superior design helps it notice minor changes in lighting artifacts, texture mixing, and facial alignment—all common signs of modified image. Although the model is computationally cheap, its accuracy and generality make it ideal for scalable implementation in real forensic applications. When added into the stacked ensemble, EfficientNet improves the deep feature representations of VGG16 and the residual learning capabilities of ResNet50, so helping to create a different and very accurate final prediction model.

3.4 Proposed Stacked Ensemble Method

After the three basic models were independently assessed, we used a stacking ensemble approach to increase the prediction potential of the model even further. In this second stage, we gathered the softmax output probabilities from ResNet50, VGG16, and EfficientNet for every input image as seen in figure 4. A meta-learner trained to discover the best mix of these predictions used these outputs as input features. Logistic regression was chosen as the meta-learner because of its interpretability and excellent performance in binary classification tasks. The stacking ensemble lets the meta-learner catch nonlinear interactions among the base model outputs by letting the model gain from the various decision boundaries and feature representations of the individual CNNs. We used a cross-validation approach to train the ensemble model strongly and prevent information leakage: the base model outputs on validation sets were used to train the meta-learner, hence guaranteeing impartial generalization during final evaluation.

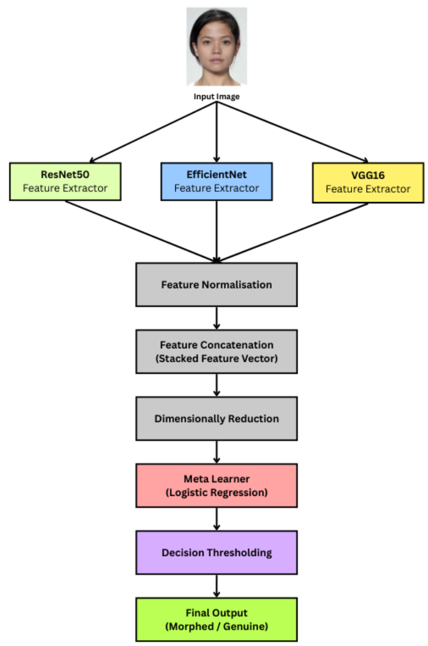

Fig.4. Block Diagram for Stacked Ensemble Method.

Not only does this combination method of deep feature extraction and ensemble learning enhance the recognition of complex altered images, but it also shows robustness against different degrees of image alteration. The suggested approach provides a more precise and dependable answer to the rising danger of facial morphing in the digital age by means of varied CNN backbones and an intelligent fusion procedure. A stratified k-fold cross-validation strategy was utilized in order to validate the ensemble model in order to eliminate the possibility of data leakage and ensure consistency. Not only did this ensure that the training and validation methods remained objective, but it also ensured that performance indicators represented the true generalizability of the system that was recommended. One of the most recent ensemble models demonstrated that the stacked ensemble framework is effective in the identification of changed facial images. This was demonstrated by the fact that it demonstrated more accuracy than individual base models

4. RESULTS

Extensive tests and the use of two datasets that contained both legitimate and tampered photographs were carried out in order to evaluate the efficacy of the suggested stacked ensemble model for the identification of modified images.

In all, there were 140,000 photos included in the primary dataset that was utilized for training and assessment purposes. This dataset was composed of 70,000 genuine photographs and 70,000 altered images. To create the manipulated photographs, a wide range of sophisticated and realistic image alteration techniques were utilized. These approaches were designed to simulate situations that might occur in the actual world. In addition to these, Morphing is the process of blending the facial characteristics of two different people in order to produce a realistic composite. This technique is frequently employed in identity spoofing. The process of using generative adversarial networks (GANs) to artificially modify facial expressions, switch identities, or construct hyper-realistic faces is referred to as deepfake generation. Creating created visual tales is accomplished through the process of picture splicing, which involves combining aspects from many photographs into a single composite. The process of replicating areas of a picture in order to conceal, erase, or magnify visual information is referred to as region duplication, also known as cloning). Modifying visual components such as the backdrop, lighting, face features, or texture with the use of powerful artificial intelligence editing techniques is referred to as attribute editing.

The purpose of this dataset was to provide a wide variety of manipulation types across a variety of situations, including illumination, resolution, face position, and background clutter. This dataset was particularly constructed to do this. Because of its complexity, the models were able to understand tiny discrepancies, which made it extremely useful for applications such as identity verification, cybersecurity, and image-based authentication systems. The stacked ensemble model achieved a high degree of accuracy in spotting both obvious and subtle manipulations, which was a considerable improvement above the performance of each individual CNN model. The model was able to acquire strong identification skills over a wide spectrum of fabricated pictures as a result of the incorporation of a variety of alteration techniques into the dataset.

The training and validation accuracy and loss curves for each model are depicted in the graphs that follow. These graphs emphasize the enhanced convergence and stability of the stacked ensemble architecture for the Primary dataset.

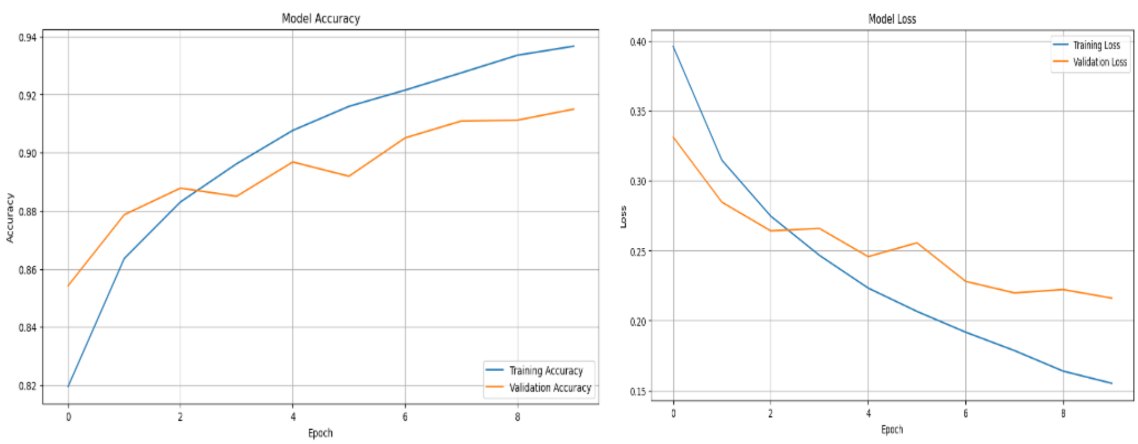

Fig.5. Accuracy and Loss Graph for EfficientNet Model

The accuracy and loss graphs for the EfficientNet model reveal a continuous and constant improvement in training performance across the epochs, as seen in figure 5. While validation accuracy exhibits a similar pattern with a little variation suggesting some generalization, training accuracy rises from about 0.84 to above 0.94. Both training and validation losses fall consistently on the loss graph; training loss falls more quickly. Though the difference between validation and training loss in later epochs deserves notice, this implies the model is learning effectively without major overfitting.

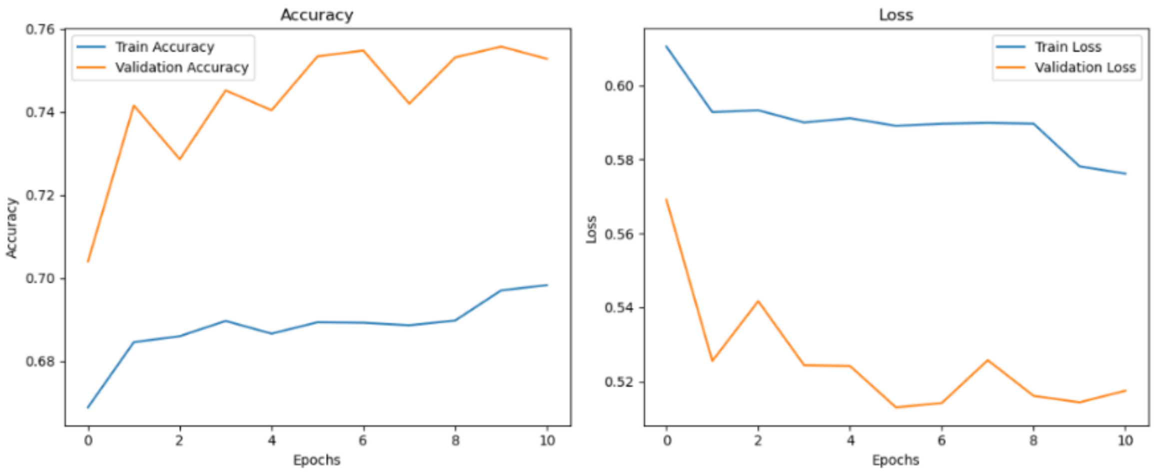

Fig.6. Accuracy and Loss Graph for ResNet50 Model

While validation accuracy starts higher and has greater variability, peaking about 0.76, training accuracy in the graphs of the ResNet50 model depicted in figure 6 reveals a slow rise from about 0.67 to around 0.70. This difference implies that the model could not be learning as effectively from the training data or could be gaining from some kind of regularization or transfer learning on the validation data. Loss graphs indicate that while validation loss falls first and then varies, training loss levels off early, suggesting instability in validation performance, either owing to overfitting or an inappropriate learning rate.

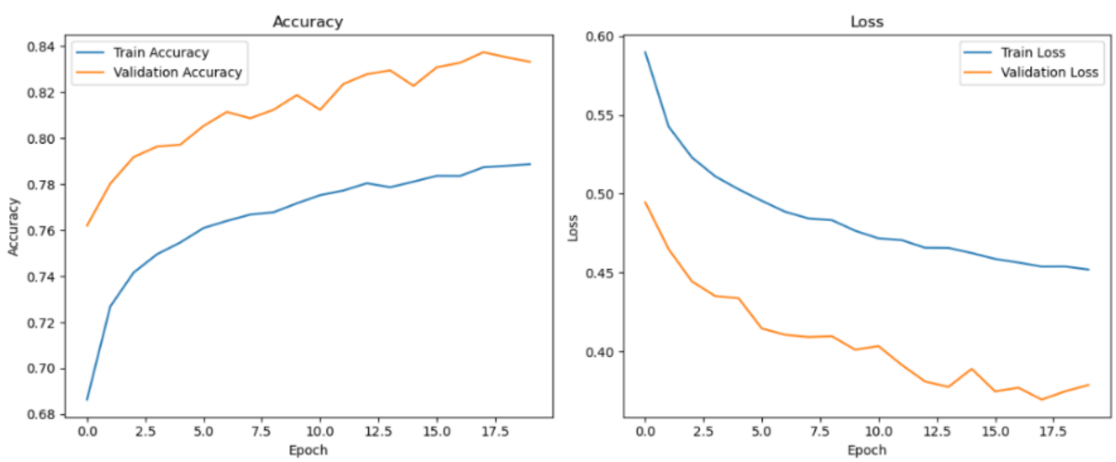

Fig.7. Accuracy and Loss Graph for VGG-16 Model

Figure 7 shows the VGG-16 model's steady training path. From around 0.68 to about 0.78, training accuracy rises steadily; validation accuracy always advances higher, reaching about 0.84. This shows significant generalization. Loss graphs show this trend; both training and validation losses drop with time. These trends' smoothness indicates a well-behaved training process and a well-tuned model with little overfitting.

With training accuracy quickly nearing 1.0 within only a few epochs, the stacked ensemble model exhibits the best training performance overall. Validation accuracy also climbs sharply, stabilizing slightly below training performance, which indicates great generalization. The loss curves, on the other hand, show a little but clear rise in validation loss following first improvements, which might suggest the start of overfitting. The model's good performance is still supported by the low loss values and close curve alignment.

Of all the models, the Stacked Ensemble had the greatest total training and validation accuracy for the primary dataset shown in table 2 below, suggesting better learning and generalization ability. With consistent loss decrease and good validation performance, VGG-16 soon follows. Though with a little difference between training and validation metrics, EfficientNet shows decent convergence and stability. Although it performs rather well, ResNet50 has the greatest variation and somewhat worse training accuracy, indicating either underfitting or instability. Thus, in this setting, particularly in terms of balancing great accuracy with little loss, the Stacked Ensemble model surpasses the others, thereby presenting the most hopeful solution.

Model	Final Training Accuracy	Final Validation Accuracy
EfficientNet	~0.95	~0.92
ResNet50	~0.70	~0.76
VGG-16	~0.78	~0.84
Stacked Ensemble	~0.98	~0.96

Table.2. Comparison table of all models for primary dataset.

We conducted experiments on an extra benchmark dataset—the DeepFake Detection Dataset (DFFD) to evaluate more the generalizability and robustness of the proposed model. Digital media forensics uses the well-known and rigorous benchmark tool DFFD. Created using many state-of-the-art forgery techniques, it boasts a sizable collection of both genuine and modified facial photographs and videos. Designed specifically to evaluate the effectiveness of deepfake detection algorithms in several realistic scenarios, this dataset is ideal for comparison.

DFFD consists of several manipulation types, each of which indicates a distinct forging method. Amongst these are

- DeepFakes: Face-swapping using autoencoders, producing synthetic facial overlays.
- FaceSwap: A traditional face-swapping technique using 3D modeling for facial alignment.
- Face2Face: Real-time facial reenactment that transfers expressions from one person to another.
- NeuralTextures: High-quality synthesis using learned neural textures, often difficult to detect.
- Face Morphing: Blending features from multiple faces to create hybrid identities, typically used in identity fraud.
- Partial Manipulations: Alterations applied only to specific facial regions (e.g., mouth, eyes) to subtly change expressions.
- Compression Variants: All manipulations are available in multiple quality settings (raw, high, and low compression) to reflect real-world usage scenarios, such as social media sharing.

We selected a fair mix of real and fake samples from all alteration types utilizing the image-based component of the DFFD dataset for this study to ensure diversity in testing. Performance was evaluated on our stacked ensemble model including predictions from many deep convolutional neural networks including EfficientNet, ResNet, and VGG. These models taken together provide a more robust detection approach by means of complementing feature representations.

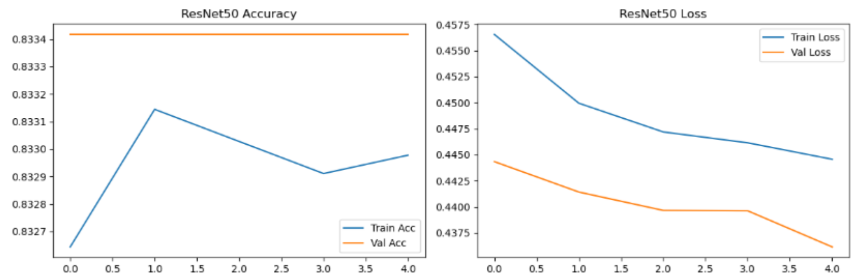

Fig.8. Accuracy and Loss Graph for VGG-16 Model

Figure 8's ResNet-50 model maintains a training accuracy value above 0.83 for the benchmark dataset over the course of the epochs. As the model continues to train, the validation accuracy first improves before slightly declining, suggesting possible overfitting. Training loss steadily declines in the loss graph, but validation loss shows a less steady downward trend. This indicates that the model learns well at first but gradually loses its ability to generalize.

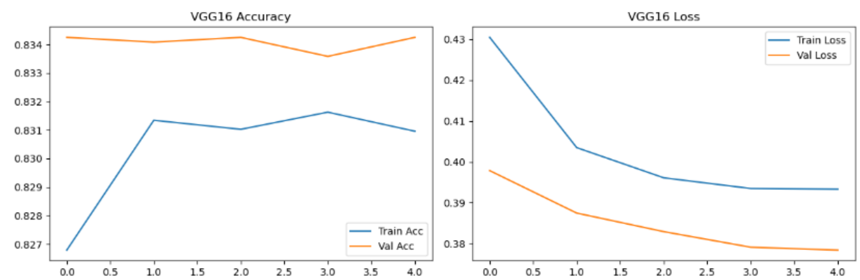

Fig.9. Accuracy and Loss Graph for VGG-16 Model

As seen in figure 9, the VGG-16 model exhibits a discernible increase in validation accuracy during the early training epochs, followed by slight variations. Effective learning is indicated by the training accuracy, which stays constant at a high level. This trend is supported by the corresponding loss curves, which show a consistent decrease in training and validation losses, with the validation loss declining noticeably. This suggests that when trained using the benchmark dataset, the VGG-16 model maintains good generalization and stays away from overfitting throughout the training process.

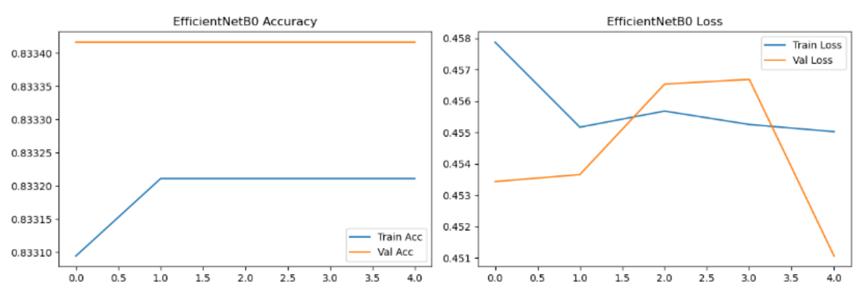

Fig.10. Accuracy and Loss Graph for VGG-16 Model

Figure 10 shows that EfficientNet-B0 closely resembles ResNet-50 in terms of performance, demonstrating strong and consistent training accuracy across epochs. Initially, the validation accuracy increases and then stabilizes. Nonetheless, the loss graph shows some variation in the validation loss over epochs, which may suggest batch size or learning rate sensitivity. However, the model's overall efficacy in learning discriminative features is supported by the downward trending training and validation losses.

On this benchmark data set, the stacked ensemble model, which combines predictions from ResNet-50, VGG-16, and EfficientNet-B0, obtained an overall accuracy of 0.87. This model achieves superior generalization and robustness by skillfully integrating the advantages of each individual architecture. Interestingly, it retained little loss, indicating less overfitting and better stability when differentiating between authentic and altered images. The effectiveness of model stacking for intricate tasks like morphing and deepfake detection is validated by this performance. Comparison is shown below in table 3 for better understanding.

Model	Validation Accuracy	Validation Loss
VGG-16	0.84	0.390
ResNet-50	0.83	0.445
EfficientNet-B0	0.83	0.448
Stacked Ensemble	0.87	0.356

Table.3. Comparison table of all models for Benchmark dataset.

The performance of the ensemble model was compared to that of many traditional forgery detection strategies recorded in past studies from 2021 to 2024. Original studies sometimes lacked complete parameter values; hence, we tested these traditional models using best-fit parameters obtained by empirical tuning and grid search, therefore ensuring a fair evaluation.

Model	Accuracy (%)
DCG + VGG-16 (Continuance Learning)	81.00
ViTCNN + MobileNetV2	62.00
InceptionResNetV2	75.57
EfficientNet-B1	84.46
Stacked Ensemble (Ours)	87.01

Table.4. Accuracy Comparison of Various Deep Learning Models trained on Benchmark dataset.

Our results indicated that the stacked ensemble model often outperformed individual models and traditional methods in terms of classification accuracy as shown in above table 4. This underlines the effectiveness of ensemble techniques depending on deep learning for controlling complex changes seen in modern digital forgeries.

5. DISCUSSION AND LIMITATIONS

The stacked ensemble model that has been developed, which incorporates ResNet50, VGG16, and EfficientNetB0 architectures, provides a promising approach to the identification of counterfeit images, notably in recognizing morphing and deepfake modifications. An improvement in detection accuracy is achieved by the ensemble system in comparison to individual classifiers. This is accomplished by using the complimentary capabilities of these models. Anomaly detection and metadata analysis are also included into the model, which further enhances its resilience in spotting small alterations. In spite of these advantages, there are a number of drawbacks that are obvious. In the first place, the quality and variety of the training dataset have a substantial impact on the performance of the model. In order to generalize the model to situations that occur in the real world, it is essential to have high-quality datasets that include a diverse assortment of morphing approaches. Furthermore, the demand for significant computing power is a barrier for deployment scenarios that include real-time systems or devices with limited resources.

The flexibility of the model to new and developing morphing methods is another challenge that has to be addressed. When it comes to maintaining accuracy, constant retraining and upgrades will be required in order to keep up with the fast advancement of image synthesis technology. Finally, the lack of clear legal frameworks in some countries, such as India, further complicates the practical enforcement and ethical issues of deploying such systems in public or forensic realms. This is because India is one of the jurisdictions concerned.

CONCLUSION

This study provides a potent machine learning model that can recognize and validate photos that have been altered by using a stacked ensemble of deep learning models. A number of image processing properties can be efficiently recorded by combining ResNet50, VGG16, and EfficientNetB0 into a single architecture, which eventually leads to significantly increased accuracy and reliability. Empirical tests on real-world datasets demonstrate the effectiveness of the model in addressing the growing threat of image-based deception.

Though there are practical issues that need to be addressed, such as the need for computing resources, the variety of datasets, and the support of legal authorities, the system shows its technological effectiveness. Future research should focus on increasing model efficiency, including unsupervised learning for flexibility, and collaborating with policymakers to create more open policies regarding the use of digital technology. All things considered, this approach is a significant advancement in safeguarding digital authenticity in this era of media manipulation powered by artificial intelligence.

REFERENCES

[1] Vaghasiya, K. (2024, November 28). *The latest ChatGPT statistics and user trends (2024-2025)*. URL: https://wisernotify.com/blog/chatgpt-users/?utm_source=chatgpt.com

[2] Nerdynav. (2024, November 4). 107 Up-to-Date ChatGPT Statistics & User Numbers [Nov 2024]. URL: https://nerdynav.com/chatgpt-statistics/?utm_source=chatgpt.com

[3] Singh, S. (2025, January 2). *Number of ChatGPT users (January 2025)*. URL: https://www.demandsage.com/chatgpt-statistics/?utm_source=chatgpt.com

[4] Broz, M., & Broz, M. (2024, December 20). *Midjourney statistics (2025)*. URL: https://photutorial.com/midjourney-statistics/?utm_source=chatgpt.com

[5] Kumar, N. (2024, December 31). *Midjourney Statistics 2025:* URL: https://www.demandsage.com/midjourney-statistics/?utm_source=chatgpt.com

[6] Mohapatra, D. (2023, December 5). Morphed photos of women: State records most cases. URL: https://timesofindia.indiatimes.com/city/bhubaneswar/morphed-photos-of-women-odisha-records-highest-cases-ncrb-report-reveals/articleshow/105741454.cms

[7] Badgeri, N. N. G. (2023, December 3). Families struggle to cope after some sextortion, loan app victims end lives. URL: https://timesofindia.indiatimes.com/city/mumbai/families-struggle-to-cope-after-some-sextortion-loan-app-victims-end-lives/articleshow/105689716.cms

[8] TIMESOFINDIA.COM. (2023, April 27). Delhi Police arrests UP man for trying to extort money from woman by using her morphed pictures. URL: https://timesofindia.indiatimes.com/city/delhi/delhi-police-arrests-up-man-for-trying-to-extort-money-from-woman-by-using-her-morphed-pictures/articleshow/99813099.cms

[9] Prajwal D'Souza, & Prajwal D'Souza. (2024, July 31). *Bengaluru: Housewife blackmailed with morphed photos after downloading loan app*. URL: https://www.deccanherald.com/india/karnataka/bengaluru/bengaluru-housewife-blackmailed-with-morphed-photos-after-downloading-loan-app-3131508

[10] Tnn. (2019, September 4). Gujarat: Ex-fiancé uploads morphed pictures. URL: https://timesofindia.indiatimes.com/city/ahmedabad/ex-fianc-uploads-morphed-pictures/articleshow/70967592.cms

[11] Ghosh, D. (2023, December 5). Kol logs most photo-morph cases in India in '22: NCRB. URL: https://timesofindia.indiatimes.com/city/kolkata/kolkata-tops-photo-morph-cases-in-india-in-2022-ncrb-report/articleshow/105740314.cms

[12] Tnn. (2019b, October 23). Rajasthan: Cyberbullying, e-frauds at record high. URL: https://timesofindia.indiatimes.com/city/jaipur/cyberbullying-e-frauds-at-record-high/articleshow/71712887.cms

[13] Express News Service, & Express News Service. (2023, July 17). *2 post morphed pics of women online, arrested*. URL: https://www.newindianexpress.com/states/tamil-nadu/2023/Jul/17/2-post-morphed-pics-of-women-online-arrested-2595640.html

[14] Tanaka, M., Shiota, S., & Kiya, H. (2021). A detection method of operated fake-images using robust hashing. *Journal of Imaging*, *7*(8), 134.

[15] Sharma, D. K., Singh, B., Agarwal, S., Garg, L., Kim, C., & Jung, K. H. (2023). A survey of detection and mitigation for fake images on social media platforms. *Applied Sciences*, *13*(19), 10980.

[16] Killi, C. B. R., Balakrishnan, N., & Rao, C. S. (2023). Deep Fake Image Classification Using VGG-19 Model. *Ingénierie des Systèmes d'Information*, *28*(2).

[17] Hsu, C. C., Zhuang, Y. X., & Lee, C. Y. (2020). Deep fake image detection based on pairwise learning. *Applied Sciences*, *10*(1), 370.

[18] Cartella, G., Cuculo, V., Cornia, M., & Cucchiara, R. (2024). Unveiling the Truth: Exploring Human Gaze Patterns in Fake Images. *IEEE Signal Processing Letters*.

[19] Huang, Y., Juefei-Xu, F., Guo, Q., Liu, Y., & Pu, G. (2022). FakeLocator: Robust localization of gan-based face manipulations. *IEEE Transactions on Information Forensics and Security*, *17*, 2657-2672.

[20] Liu, H., Tan, Z., Tan, C., Wei, Y., Wang, J., & Zhao, Y. (2024). Forgery-aware adaptive transformer for generalizable synthetic image detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 10770-10780).

[21] Flack, T. R., Ritchie, K. L., Cartledge, C., Fuller, E. A., & Kramer, R. S. (2023). Improving face morph detection with the pairs training effect. *Applied Cognitive Psychology*, *37*(6), 1158-1166.

[22] Afchar, D., Nozick, V., Yamagishi, J., & Echizen, I. (2018, December). Mesonet: a compact facial video forgery detection network. In *2018 IEEE international workshop on information forensics and security (WIFS)* (pp. 1-7). IEEE.

[23] Bondi, L., Lameri, S., Güera, D., Bestagini, P., Delp, E. J., & Tubaro, S. (2017, July). Tampering detection and localization through clustering of camera-based CNN features. In *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (pp. 1855-1864). IEEE.

[24] Mandisha, M. S., Hussien, M. A., Shalaby, A. K., & Fahmy, O. M. (2025). Wavelet Transform-based Methods for Forensic Analysis of Digital Images. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, *44*(1), 46-54.

[25] Chai, L., Bau, D., Lim, S. N., & Isola, P. (2020). What makes fake images detectable? understanding properties that generalize. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXVI 16* (pp. 103-120). Springer International Publishing.

[26] Atlam, E. S., Almaliki, M., Elmarhomy, G., Almars, A. M., Elsiddieg, A. M., & ElAgamy, R. (2025). SLM-DFS: A systematic literature map of deepfake spread on social media. *Alexandria Engineering Journal*, *111*, 446-455.

[27] Gorle, R., & Guttavelli, A. (2025). Enhanced Image Tampering Detection using Error Level Analysis and CNN. *Engineering, Technology & Applied Science Research*, *15*(1), 19683-19689.

[28] Chalini, G. R., & Kanimozhi, K. V. (2024, September). Evaluation Techniques to Detect Face Morphing Vulnerabilities for Differential Images. In *2024 5th International Conference on Smart Electronics and Communication (ICOSEC)* (pp. 1532-1537). IEEE.

[29] Seibold, C., Samek, W., Hilsmann, A., & Eisert, P. (2017). Detection of face morphing attacks by deep learning. In *Digital Forensics and Watermarking: 16th International Workshop, IWDW 2017, Magdeburg, Germany, August 23-25, 2017, Proceedings 16* (pp. 107-120). Springer International Publishing.

[30] Kadiri, P., Anusha, P., Prabhu, M., Asuncion, R., Pavan, V. S., & Suman, J. V. (2024, July). Morphed Picture Recognition using Machine Learning Algorithms. In *2024 Second International Conference on Advances in Information Technology (ICAIT)* (Vol. 1, pp. 1-6). IEEE.

[31] Sharma, P., Kumar, M. and Sharma, H.K., GAN-CNN Ensemble: A robust deepfake detection model of social media images using minimized catastrophic forgetting and generative replay technique, Procedia Computer Science, 2024, 235, 948–960.

[32] Agarwal, A. and Ratha, N., Face Morphing Detection in Social Media Content, 2024 IEEE International Conference on Image Processing (ICIP), 2024.

[33] Dua, F., Yu, M., Li, B., Chow, K.P., Jiang, J., Zhang, Y., Liang, Y., Li, M., and Huang, W., TAENet: Two-branch Autoencoder Network for Interpretable Deepfake Detection, Forensic Science International: Digital Investigation, 2024, 50, 101808.

[34] Dwivedi, R., Kothari, P., Chopra, D., Singh, M., and Kumar, R., An efficient ensemble explainable AI (XAI) approach for morphed face detection, Pattern Recognition Letters, 2024, 184, 197-204.

[35] Matakias, A. and Sotiropoulos, D.N., A Comparative Study of Face Morphing Algorithms, 2024 15th International Conference on Information, Intelligence, Systems & Applications (IISA), 2024.

[36] Tassone, F., Maiano, L., and Amerini, I., Continuous fake media detection: Adapting deepfake detectors to new generative techniques, Computer Vision and Image Understanding, 2024, 249, 104143.