Identity in the Blood Relation: Unraveling the Complexity of Morph Detection in Kinship Biometrics

Shalvika Srivastav[1]
shalvika21@iiserb.ac.in
Poulomi Bhattacharya[1]
poulomi24@iiserb.ac.in
Akshay Agarwal[1]
akagarwal@iiserb.ac.in
Nalini Ratha[2]
nratha@buffalo.edu

[1] Trustworthy BiometraVision Lab, IISER Bhopal, India
[2] Department of Computer Science and Engineering, University at Buffalo, USA

Abstract

Face morphing attacks present a significant challenge to biometric authentication systems by allowing multiple individuals to share a single identity. While much existing research has focused on detecting morphs between unrelated individuals, a critical gap remains in understanding morphs generated from genetically related subjects, such as parents and children. These kinship-based morphs closely mimic natural familial resemblance, making them difficult to detect using conventional approaches. This work explores the problem of kinship-based morphing by introducing a large-scale synthetic dataset generated using both Open-CV and latent autoencoder-based techniques across various blending ratios. A hybrid detection framework is proposed that leverages identity features extracted from both FaceNet and ResNet50, aiming to capture nuanced facial inconsistencies that may arise in morphs with substantial familial overlap. The study also considers performance across multiple morphing ratios and investigates generalization under unseen synthetic conditions. Additionally, a face recognition experiment is performed across different morph ratios to reflect how the recognition score varies as the morph ratio varies. This research opens new directions for improving the robustness of morph detection systems in the context of realistic kinship-based attacks.

1 Introduction

Face morphing attacks present a significant security risk to biometric authentication systems by enabling multiple individuals to impersonate a single identity. Although existing research primarily addresses morphs generated from unrelated individuals, the detection of kinship-based morphs, particularly parent-child morphs, remains largely underexplored. Due to inherent facial similarities, these morphs closely resemble genuine familial images, effectively concealing morphing artifacts and increasing the likelihood of false acceptances. This subtle challenge poses a critical threat to the reliability of biometric systems in practical deployments. To address this gap, a large-scale synthetic kinship-based morph dataset is constructed using OpenCV-based and autoencoder-based techniques across multiple blending ratios to simulate a range of attack strengths and realism levels. A proposed hybrid detection framework combines embeddings from FaceNet [19], and ResNet50 [7], leveraging their complementary strengths to capture low-level facial texture anomalies and high-level identity inconsistencies. The framework undergoes evaluation across various morphing ratios and is further tested for its cross-ratio generalization capability to assess robustness in unseen blending conditions. This work provides new insights into the limitations of traditional morph detection approaches in familial resemblance. It highlights the importance of incorporating synthetic kin-based data into detection pipelines to improve the reliability and generalizability of biometric verification systems. Major contributions of this work are as follows: ✪ A large-scale synthetic kinship-based morphed dataset is constructed using OpenCV and autoencoder techniques with varying blending ratios to simulate realistic parent-child morphing attacks, ✭ A hybrid morph detection framework is proposed by combining FaceNet and ResNet50 embeddings to capture complementary facial identity cues and morphing artifacts, ✳ Comprehensive intra-ratio and cross-ratio evaluation is performed to analyze performance across morphing strengths and assess generalization to unseen ratios, ✴ Face recognition uses varying morph ratio images to reflect the morphing needed to fool the state-of-the-art face recognition networks, and ✷ The study provides empirical evidence that kin-based morphs are significantly more complex to detect than general morphs due to natural familial resemblance.

2 Related Work

Face morphing attacks enable multiple individuals to be authenticated under a single identity, presenting a significant vulnerability to biometric authentication systems. These attacks rely on synthetically blending the facial features of two subjects, creating hybrid identities that remain visually plausible and often evade detection by face recognition models. While substantial progress is made in general-purpose morph detection [16, 21], kinship-based morphs remain largely overlooked. Kinship-based morphing introduces a unique challenge, as parent-child facial similarities naturally exist due to shared genetic traits, which reduce distinguishable morph artifacts and hinder the efficacy of traditional detection methods.

Kinship verification explores biological relationships such as parent-child and sibling connections, with practical applications in immigration screening, forensic analysis, and familial photo organization [4, 23]. Conventional kinship verification utilizes either handcrafted features or deep learning-based representations. However, kinship-based morphs introduce synthetically blended faces that closely resemble both contributors, making it difficult to differentiate legitimate biological resemblance from manipulated imagery. In these scenarios, biometric verification systems become more vulnerable to high false acceptance rates, particularly in security-critical environments such as border control and law enforcement [21, 24]. Early approaches to morph detection rely on handcrafted features, including Local Binary Patterns (LBP) [12], Binarized Statistical Image Features (BSIF) [8], and frequency-domain analysis through Fourier transform [17, 21]. While effective for detecting basic morphs, these methods struggle against synthetically generated morphs with fewer artifacts and better-blending fidelity. Deep learning approaches using CNNs such as ResNet50 [7] capture spatial texture inconsistencies, and Vision Transformers (ViTs) [5] employ self-

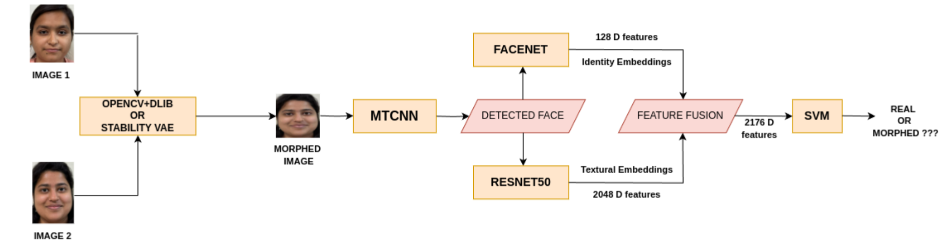

Figure 1: Proposed kinship morph generation and detection architecture.

attention to detect subtle distortions [1]. Despite their effectiveness, ViTs often require large-scale annotated datasets and high computational resources. Hybrid techniques emerge, fusing handcrafted and learned features to balance interpretability and performance. Multi-level and multi-modal fusion strategies combine local texture descriptors like Histogram of Oriented Gradients (HoG) with deep identity embeddings, improving morph detection accuracy across varying datasets [16, 18, 22].

MagNet [1] introduces the Weighted Local Magnitude Pattern (WLMP), which captures imperceptible local distortions in synthetic images more effectively than traditional descriptors. Other works integrate embeddings from networks like AlexNet [10] and ResNet50 [7], leveraging low-level and high-level representations for robust classification. MADation [3] presents the first adaptation of foundation models to morphing attack detection by fine-tuning CLIP with LoRA and a classification head, achieving state-of-the-art performance across diverse evaluation scenarios while maintaining generalization capabilities. Despite these advancements, morph detection between genetically related individuals remains an open problem, as natural similarity reduces the contrast between real and synthetic samples. To mitigate these limitations, the proposed work focuses on kinship-aware synthetic morph detection using a hybrid approach that combines FaceNet for identity-aware embeddings and ResNet50 for texture-aware representations. These complementary features are fused and classified using an SVM to evaluate robustness across morphing ratios and kinship-specific attack scenarios.

3 Proposed Kinship Face Morph Dataset

The morphed dataset used in this work is derived from the TSKinFace dataset [13], a widely adopted benchmark for kinship verification tasks. TSKinFace includes images of biologically related individuals grouped into father (F), mother (M), son (S), and daughter (D) categories, captured in controlled conditions to maintain consistency in pose, illumination, and background. Using classical morphing techniques, kinship-based morphed samples are generated by synthetically blending facial images of parent-child pairs from the same family. The combinations used include father-daughter, mother-daughter, father-son, and mother-son pairs. The morphing process primarily utilizes OpenCV and Dlib. It begins with facial landmark detection using Dlib's 68-point landmark detector, which is used for facial alignment. This is followed by geometric warping and feature blending through OpenCV's morphing tools, ensuring structural consistency in the resulting morphed images. Morphs are generated at five different blending ratios 90-10, 80-20, 70-30, 60-40, and 50-50, where, for example, a 90-10 blend contains 90% of the child and 10% of the parent's features. Lower-ratio morphs tend to preserve the child's appearance, making detection more difficult, whereas higher-ratio blends introduce more visible artifacts, making detection comparatively easier. The dataset comprises 5,176 images, with 2,588 authentic (real) and 2,588 morphed samples across each ratio. In addition to traditional OpenCV-based morphs, a new autoencoder-based kinship morphing dataset is generated using stable variational-autoencoder (Stability-VAE) [15]. This VAE encodes high-resolution images into a compressed latent representation. Morphs are generated by linear interpolating latent vectors with a weighted average of 0.9, followed by decoding into image space.

The autoencoder-based dataset contains 2,030 images for each ratio and mirrors the same morphing ratio splits (90-10 through 50-50) to enable consistent and fair comparison with classical morphs. To understand the increased difficulty of detecting kinship-based morphs, we compare our TSKinFace-based morph detection results against general-purpose morph detection performance on the FRLL-Morphs dataset [18]. FRLL-Morphs include high-quality frontal morphs generated using multiple algorithms such as OpenCV warping, FaceMorpher, WebMorpher, and StyleGAN, but without kinship relations between source subjects. Our evaluation reveals that kinship-based morph detection is substantially more rigid due to the inherent facial similarity between related individuals, which reduces the presence of detectable artifacts. Overall, the proposed dataset offers a rich and diverse benchmark for kinship-aware morph detection, featuring: ❶ A novel kinship-based morphed dataset derived from TSKinFace includes 2588 real, 2,588 OpenCV-based, and 2,030 autoencoder-based morph images for each of five morphing ratios (90-10 to 50-50), ❷ Introduction of autoencoder-based morphs using Stability-VAE, providing highly realistic kinship morphs with minimal artifacts, ❸ Direct comparison with the FRLL-Morphs dataset to demonstrate the increased challenge of kinship morph detection over general morph detection, and ❹ Comprehensive morph generation pipeline using traditional warping and latent-space interpolation for broader attack diversity.

4 Proposed Face Morph Detection Algorithm

The proposed morph detection algorithm adopts a hybrid feature extraction strategy by integrating identity-aware and texture-focused representations, as shown in Figure 1. FaceNet, specifically the InceptionResnetV1 architecture pre-trained on the VGGFace2 dataset, extracts 128-dimensional embeddings optimized for identity verification. These embeddings capture critical facial features typically shared in morphed images, often blending identity traits between two individuals. Simultaneously, ResNet50, pre-trained on the ImageNet dataset, extracts 2048-dimensional deep features from the final convolutional layer, effectively capturing local texture artifacts and blending inconsistencies introduced during the morphing process. We combine the identity-focused embeddings from FaceNet with the texture-sensitive features extracted by ResNet50 into a single 2,176-dimensional vector. This fusion of FaceNet's identity-discriminative representation and ResNet50's artifact-sensitive descriptors yields complementary information that enhances our ability to detect subtle morphing artifacts. This comprehensive feature vector is fed into a Support Vector Machine (SVM) classifier with a radial basis function (RBF) kernel, which is well-suited for high-dimensional, nonlinear classification tasks. The SVM model is trained on a balanced set of real and morphed samples using SMOTE for oversampling and undersampling for test balancing, with stratified 3-fold cross-validation applied to ensure performance stability and generalizability. This hybrid model effectively captures global semantic and localized visual cues essential for accurate morph detection, forming the core of our proposed approach.

Table 1: Morphing detection performance across different models on 90-10 ratio.

Models	Accuracy (↑)	MACER (↓)	BSCER (↓)	D-EER (↓)
MagNet [1]	57.51	43.20	36.54	39.87
FaceNet + MLP	71.48	39.00	18.08	28.54
AlexNet + ResNet + SRKDA [22]	74.13	**13.16**	43.93	28.55
FaceNet + ECA + SVM	71.00	28.41	28.31	28.36
ResNet + SVM	73.64	27.90	27.00	27.50
ResNet (fine-tuned) on unseen data	68.64	34.42	27.64	31.03
FaceNet + MLP (Optuna Tuned)	74.18	35.91	15.77	25.84
FaceNet + SVM [6]	75.43	36.49	12.69	24.59
FaceNet + HoG + Fourier Transform + SVM	76.40	34.36	12.88	23.62
CLIP (ViT) + SVM	76.89	33.42	14.21	23.12
CNN + ECA [11]	77.41	32.18	13.46	22.59
FaceNet + ViT-b/16 + SVM	**77.78**	32.26	**12.17**	22.22
AdaFace + SVM	64.65	35.96	35.78	35.87
AdaFace + ResNet + SVM	76.81	21.51	21.49	21.50
AdaFace + ViT-b/16 + SVM	76.41	**21.46**	21.38	**21.42**
FaceNet + ResNet + SVM (Proposed)	**78.83**	30.33	**12.02**	**21.17**

5 Experimental Results and Analysis

To rigorously assess the effectiveness of the proposed hybrid morph detection model, we evaluate its performance using a range of experiments under diverse conditions. The evaluation is conducted using four widely accepted metrics: Accuracy, which measures the proportion of correctly classified samples; Morphing attack classification error rate (MACER), indicating the proportion of morphed samples incorrectly classified as bona fide samples in a specific scenario; Bona Fide Sample Classification Error Rate (BSCER), representing proportion of bona fide samples incorrectly classified as morphed samples in a particular scenario; and Detection Equal Error Rate (D-EER), which captures the trade-off point where MACER equals BSCER, serving as a balanced error indicator for biometric systems.

5.1 Comparison with Existing Models

Table 1 presents a comprehensive comparative analysis conducted to evaluate the effectiveness of the proposed hybrid morph detection model against several existing and baseline approaches. The first baseline is the FaceNet + SVM model [6], which leverages 128-dimensional identity embeddings from FaceNet and uses an SVM for classification. This setup is extended in multiple directions. A Multi-Layer Perceptron (MLP), both in its standard form and optimized using Optuna [2], is evaluated to model complex non-linear relationships in the feature space. A hybrid FaceNet + ViT-CLIP + SVM model is explored, where the ViT-based CLIP encoder [14] provides high-level semantic embeddings fused with FaceNet representations and classified using an SVM. Additionally, a standalone CLIP (ViT-B/16) + SVM configuration is evaluated, utilizing CLIP's vision encoder to extract global

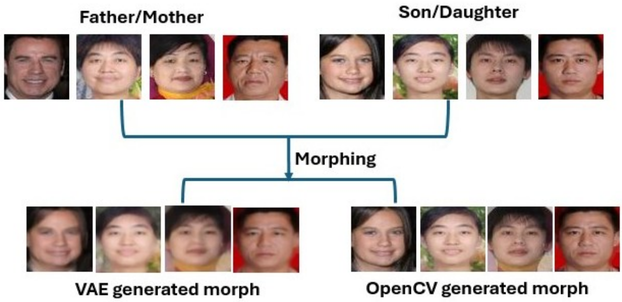

Figure 2: The morphed faces are created using an autoencoder and OpenCV. In each case, a parent is blended with their child to produce realistic, child-like morphed images.

visual features, followed by SVM classification. While CLIP is not explicitly designed for morph detection, its strong generalization helps distinguish between authentic (real) and morphed images. Another strong baseline is the AlexNet + ResNet50 + SRKDA model [22], which performs multilevel deep feature fusion and applies Spectral Regression Kernel Discriminant Analysis (SRKDA) for improved class separability. We also include attention-driven models such as Patch CNN + Efficient Channel Attention (ECA) [11], where local facial patches are processed to capture fine-grained morphing artifacts, and ECA modules enhance facial regions through channel-wise attention. The integration of FaceNet embeddings with ECA-based attention (FaceNet + ECA + SVM) is also examined to assess the fusion of local attention and identity-aware features.

Furthermore, we assess handcrafted and anomaly-based methods. MagNet [1] applies a reconstruction-based anomaly detection strategy, flagging morphed samples as out-of-distribution inputs based on reconstruction loss. A handcrafted fusion method combining FaceNet, Histogram of Oriented Gradients (HoG), and Fourier transform with SVM classification is evaluated to exploit spatial and frequency domain characteristics. AdaFace [9], a face recognition model with adaptive margin learning, is used in multiple configurations, including AdaFace + SVM, AdaFace + ResNet + SVM, and AdaFace + ViT + SVM, to evaluate the adaptability of identity-aware embeddings to morph detection. We also investigate a fine-tuned ResNet model trained end-to-end on unseen kin-based data to benchmark the generalization capacity of single-model deep learning approaches.

Additionally, we study the robustness of our proposed hybrid FaceNet + ResNet + SVM model when augmented with Gaussian noise during testing to simulate real-world degradations. Under Gaussian noise (mean=0, standard deviation=10), the hybrid model achieves 77.68 % accuracy (slightly down from 78.83 %) and 21.95 % D-EER (up from 21.17 %), compared to Patch-CNN + ECA and FaceNet + ViT + SVM. These results demonstrate the hybrid model's superior robustness and make it a reliable choice for noisy, real-world biometric verification. Across all experiments, the proposed hybrid model consistently

Table 2: Performance of the proposed hybrid morph detection model across different morph ratios and cross-ratio configurations.

Detection across ratios					Cross ratio generalization (D-EER %)				
Morph Ratio	**Accuracy**	**MACER**	**BSCER**	**D-EER**	**90-10**	**80-20**	**70-30**	**60-40**	**50-50**
90-10	78.83	30.33	12.02	21.17	21.17	18.62	12.68	9.94	8.22
80-20	85.22	19.13	10.43	14.78	18.36	14.78	10.12	9.16	8.24
70-30	92.08	8.93	6.92	7.92	15.08	12.98	7.92	10.84	9.52
60-40	95.38	4.68	4.48	4.58	17.68	15.24	12.84	4.58	11.08
50-50	95.40	4.48	4.71	4.60	20.66	18.50	15.12	12.08	4.60

demonstrates superior performance in morph detection across various morphing techniques and datasets. Combining identity-sensitive and texture-focused features and classifying them with a kernel-based method, the hybrid model outperforms deep learning, handcrafted, transformer-based, attention-enhanced, and anomaly-based baselines, particularly in challenging cross-ratio generalization and kin-based scenarios. After the comparison, it is evident that the proposed hybrid model leads and is closely followed by attention-enhanced methods like Patch CNN + ECA. Transformer-based CLIP + ViT + SVM underperforms in this kinship context with a D-EER of 21.17 due to a limited focus on fine-grained identity cues. Anomaly detection-based methods like MagNet struggle to achieve high precision but a poor recall with a D-EER of 26.81. At the same time, AdaFace variants (e.g., AdaFace + ResNet + SVM) offer moderate results, with D-EERs around 21.50, reflecting their optimization for recognition rather than morph detection.

5.2 Age-Related Bias Examination

Figure 2 presents a series of morphed child faces generated using a variational autoencoder (VAE) in conjunction with OpenCV. Given that the chronological gap between parent and child typically ranges from 25 to 30 years, one might worry that age-related artifacts could simplify distinguishing genuine from morphed children. To address this, we subject both the original and morphed child images to an age-estimation analysis using the DeepFace Apparent Age Estimation model [20], the standard VGGFace (VGG-16) face recognition backbone with a custom convolutional (1x1) head. Across all samples, the mean estimated age difference between each morphed child and its genuine counterpart is only 2-3 years. This minimal discrepancy confirms that the morphing pipeline does not introduce significant age-related cues; also the improvements in detection accuracy cannot be attributed to spurious "aging" artifacts.

5.3 Performance Across Morph Ratios

To evaluate the impact of morph complexity, we test the proposed hybrid model (FaceNet + ResNet50 + SVM) across five morphing ratios: 90-10, 80-20, 70-30, 60-40, and 50-50. In these ratios, the contribution of the child and parent to the morphed image varies, with 90-10 being the most visually similar to the child and 50-50 being the most evenly blended.

Table 3: Cross-dataset performance of the hybrid morph detection model trained on general (FRLL) and kinship (Proposed) datasets.

Test (↓)\Train (→)	**FRLL-Morphs**		**Proposed Dataset**	
	Accuracy	**D-EER**	**Accuracy**	**D-EER**
Proposed Dataset	59.89	38.96	78.83	21.17
FRLL-Morphs	82.46	22.54	68.98	24.32
UB KinFace	61.46	34.56	75.88	21.36

As presented in Table 2, the hybrid model's performance improves significantly as the morph becomes more balanced. Specifically, at the 90-10 ratio, where the child dominates the morph and facial artifacts are minimal, the model achieves an accuracy of 78.83% and a relatively high D-EER of 21.17%. This highlights the difficulty of detecting highly child-dominant kin morphs due to subtle visual cues. As the parent's contribution increases, morph artifacts become more detectable, and performance steadily improves. At the 70-30 ratio, the model surpasses 92% accuracy and reaches a D-EER of just 7.92%. The performance peaks at the 50-50 ratio with 95.40% accuracy and an exceptionally low D-EER of 4.60%, demonstrating the hybrid model's strength in identifying more evenly blended morphs.
Cross-ratio generalization is another critical real-world scenario where attackers may generate morphs at varying intensities to evade detection, particularly in low-ratio morphs that closely resemble genuine facial images. Our experiments demonstrated in Table 2 (last five columns) show that models trained on subtler morphs, i.e., trained on a 90-10 ratio, perform relatively better than those trained on ratios like 50-50. This performance disparity arises due to the visibility and consistency of morphing artifacts. In 50-50 morphs, facial features from both parent and child contribute equally, leading to prominent blending inconsistencies such as mismatched texture, symmetry distortion, and unnatural transitions around facial landmarks. These artifacts are easier for feature extractors and classifiers to learn and detect. In contrast, a 90-10 morph largely preserves the dominant identity (typically the child), resulting in images that appear nearly identical to genuine ones. The minimal contribution from the secondary identity introduces only subtle and sparse artifacts, which are harder to learn and distinguish. As a result, a model trained on 90-10 morphs often generalizes well when evaluated on more balanced or artifact-rich morphs. Overall, cross-ratio generalization serves as a comprehensive benchmark for evaluating the resilience of morph detection models. It ensures that models are not merely overfitting to specific morphing configurations but can effectively detect morphs created with different blending strengths, thereby enhancing their practicality for real-world biometric security applications.

5.4 Cross-Dataset Generalization

To assess the generalization ability of morph detection models across different datasets, we conducted cross-dataset evaluations using the proposed model in Table 3. The model was trained separately on a general morph dataset (FRLL) and the kinship-based morph dataset

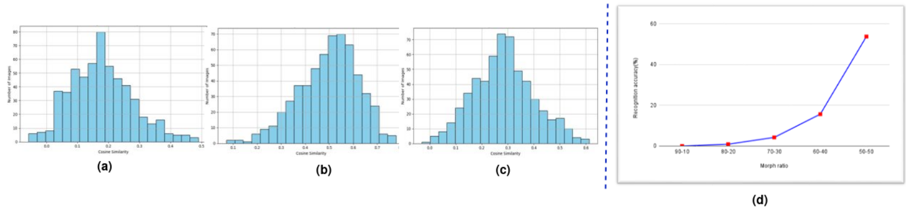

Figure 3: Three histogram panels in (a), (b), and (c) display the cosine-similarity distributions between each morphed image and its parent, while the line plot in (d) shows the recognition performance across varying morph ratios. Together, they reveal a clear upward trend in both cosine similarity and recognition accuracy as the proportion of parental features in the morphs increases.

to compare performance across domains. When trained on the FRLL dataset and tested on the kinship-based morphs, the model achieved a relatively low accuracy of 59.89% with a high D-EER of 38.96%. This performance degradation can be attributed to the model learning distinct artifacts from morphs generated using entirely unrelated individuals. As a result, it struggles to generalize to more challenging cases, such as morphs involving individuals with close familial resemblance, where the differences are more subtle and less distinguishable. In contrast, the model trained on FRLL struggled, achieving only 61.46% accuracy and a much higher D-EER of 34.56% when tested on UB KinFace. The CNN+ECA model trained on our proposed kinship-based dataset consistently outperforms the model trained on general FRLL morphs. This improvement stems from the model's ability to learn more complex and closely related features shared between parents and children. By training on these subtle familial resemblances, the model adopts fine-grained facial variations, which in turn enhances its generalization performance when applied to standard morph detection scenarios. When tested on the proposed dataset, it achieves a lower D-EER of 23.08% compared to 36.91% for the FRLL-trained model. On UB KinFace, the proposed-trained model achieves 22.93% vs. 36.73% on FRLL-Morphs; it records 26.14% compared to 24.18% for the FRLL-trained model. Results show that kin-trained models generalize to general morphs, but not vice versa, underscoring the distinct challenges posed by kinship-based morphs due to their natural facial similarity.

6 Impact of Morph Ratio in Face Matching

To evaluate how much morphed child images resemble their parents, we perform a face recognition task using morphed images (with blending ratios from 90-10 to 50-50) as probes and clean parent images as the gallery. We use the pre-trained AdaFace model [9] with a ResNet-100 backbone and cosine similarity (threshold = 0.5) for matching. As shown in Figure 3 (a), morphs with only 10% parental features (90-10) consistently score below the threshold, indicating no identity match and low recognition accuracy. As the parental contribution increases, similarity scores and recognition rates rise, with 50-50 morphs often exceeding the threshold [Figure 3 (b), (c), (d)]. These results suggest that low-contribution morphs are poor impersonators but harder to detect, while balanced morphs pose a greater security risk due to their potential for dual identity.

7 Conclusion

This work addresses a critical yet underexplored area in synthetic face analysis—kinship-based morphing attacks. Unlike conventional morphs or deepfakes generated between unrelated individuals, kinship morphs exploit natural familial resemblance, making them particularly difficult to detect. Higher morphing ratios amplify blending artifacts, improving detection and generalization across ratios. Our hybrid model, combining FaceNet and ResNet embeddings with a support vector machine classifier, outperforms traditional CNN-based methods and handcrafted features, establishing a robust baseline for kinship morph detection. External validation on unseen datasets further supports the generalizability of our approach. By revealing the vulnerabilities posed by kinship morphs in applications such as passport verification and identity screening, this study contributes meaningful insights to the synthetic data community. It highlights the need for specialized defenses in biometric systems exposed to familial manipulation threats. Future directions will aim to enhance detection performance against low-ratio morphs, which remain challenging due to the dominance of a single identity and subtle morphing artifacts.

Acknowledgement

A. Agarwal is partially supported through the PM-ECRG grant of ANRF, Gov. of India.

References

[1] Akshay Agarwal, Richa Singh, Mayank Vatsa, and Afzel Noore. Magnet: Detecting digital presentation attacks on face recognition. *Frontiers in Artificial Intelligence*, 4: 643424, 2021.

[2] Takuya Akiba, Shotaro Sano, Toshihiko Yanase, Takeru Ohta, and Masanori Koyama. Optuna: A next-generation hyperparameter optimization framework. In *25th ACM SIGKDD international conference on knowledge discovery & data mining*, pages 2623–2631, 2019.

[3] Eduarda Caldeira, Guray Ozgur, Tahar Chettaoui, Marija Ivanovska, Peter Peer, Fadi Boutros, Vitomir Struc, and Naser Damer. Madation: Face morphing attack detection with foundation models. In *Winter Conference on Applications of Computer Vision*, pages 1650–1660, 2025.

[4] Hamdi Dibeklioglu, Albert Ali Salah, and Theo Gevers. Like father, like son: Facial expression dynamics for kinship verification. In *IEEE International Conference on Computer Vision*, pages 1497–1504, 2013.

[5] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020.

[6] Muhammad Hamza, Samabia Tehsin, Hanen Karamti, and Norah Saleh Alghamdi. Generation and detection of face morphing attacks. *IEEE Access*, 10:72557–72576, 2022.

[7] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 770–778, 2016.

[8] Juho Kannala and Esa Rahtu. BSIF: Binarized statistical image features. In *IEEE International Conference on Pattern Recognition*, pages 1363–1366, 2012.

[9] Minchul Kim, Anil K Jain, and Xiaoming Liu. Adaface: Quality adaptive margin for face recognition. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 18750–18759, 2022.

[10] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25, 2012.

[11] Min Long, Xuan Zhao, Le-Bing Zhang, and Fei Peng. Detection of face morphing attacks based on patch-level features and lightweight networks. *Security and Communication Networks*, 2022(1):7460330, 2022.

[12] Timo Ojala, Matti Pietikäinen, and Topi Mäenpää. Gray scale and rotation invariant texture classification with local binary patterns. In *European Conference on Computer Vision*, pages 404–420. Springer, 2000.

[13] Xiaoqian Qin, Xiaoyang Tan, and Songcan Chen. Tri-subject kinship verification: Understanding the core of a family. *IEEE Transactions on Multimedia*, 17(10):1855–1867, 2015.

[14] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, et al. Learning transferable visual models from natural language supervision. In *International conference on machine learning*, pages 8748–8763. PmLR, 2021.

[15] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10684–10695, 2022.

[16] Andreas Rossler, Davide Cozzolino, Luisa Verdoliva, Christian Riess, Justus Thies, and Matthias Nießner. Faceforensics++: Learning to detect manipulated facial images. In *IEEE/CVF International Conference on Computer Vision*, pages 1–11, 2019.

[17] Ulrich Scherhag, Luca Debiasi, Christian Rathgeb, Christoph Busch, and Andreas Uhl. Detection of face morphing attacks based on prnu analysis. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, 1(4):302–317, 2019.

[18] Ulrich Scherhag, Christian Rathgeb, Johannes Merkle, and Christoph Busch. Deep face representations for differential morphing attack detection. *IEEE transactions on information forensics and security*, 15:3625–3639, 2020.

[19] Florian Schroff, Dmitry Kalenichenko, and James Philbin. Facenet: A unified embedding for face recognition and clustering. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 815–823, 2015.

[20] Sefik Ilkin Serengil and Aykut Ozpınar. DeepFace: A python library for face recognition and facial attribute analysis. https://github.com/serengil/deepface, 2020.

[21] Sushma Venkatesh, Raghavendra Ramachandra, Kiran Raja, and Christoph Busch. Face morphing attack generation and detection: A comprehensive survey. *IEEE transactions on technology and society*, 2(3):128–145, 2021.

[22] Sushma Venktatesh. Multilevel fusion of deep features for face morphing attack detection. In *IEEE International Conference on Electrical, Computer, Communications and Mechatronics Engineering*, pages 1–7, 2022.

[23] Haibin Yan and Junlin Hu. Video-based kinship verification using distance metric learning. *Pattern Recognition*, 75:15–24, 2018.

[24] Haoyu Zhang, Sushma Venkatesh, Raghavendra Ramachandra, Kiran Raja, Naser Damer, and Christoph Busch. Mipgan—generating strong and high quality morphing attacks using identity prior driven gan. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, 3(3):365–383, 2021.