Detecting Morphing Attack for Secure Digital Transactions

Yi Ci Kek, Wai Lam Hoo
Faculty of Computer Science and Information Technology
Universiti Malaya
Kuala Lumpur, Malaysia
22102539@siswa.um.edu.my, wlhoo@um.edu.my

***Abstract*—The rapid adoption of biometric authentication in online banking and eKYC systems has increased exposure to face morphing attacks, where multiple faces are blended to create a single image that can trick both humans and facial recognition algorithms. These attacks exploit weaknesses in current systems that focus only on identity matching without verifying image authenticity. To address these challenges, this research proposes a hybrid deep learning framework that combines S-MAD and D-MAD techniques. The S-MAD module, built on the EfficientNet-B3 architecture and trained on the SMDD dataset, identifies pixel- and texture-level morphing artifacts in single images. The D-MAD module utilizes a Siamese network with ArcFace embeddings and cosine similarity to compare a reference ID with a selfie image, which is trained on the FEI dataset and evaluated on FRLL for cross-dataset performance. The Decision Fusion Layer integrates the outputs of both models to enhance detection accuracy and reliability. The model is deployed as a FastAPI–Streamlit web application that enables users to upload and verify images in real-time, displaying confidence scores and visual explanations of detected morphing regions. Evaluation using accuracy, precision, recall, F1-score, and AUC metrics shows strong generalization and consistent performance. Overall, this research strengthens biometric authentication by providing a scalable, interpretable, and secure solution to prevent morph-based identity fraud in digital financial systems.**

***Index Terms*—Face Morphing Attack Detection, Biometric Authentication, Deep Learning, Siamese Network, eKYC Security**

I. Introduction

Fake images used for identity fraud are not a new phenomenon. Since the rise of digital authentication systems, cybercriminals have exploited image manipulation to bypass biometric verification workflows. In the past, creating forged identification photos required advanced editing tools and technical expertise, making such attacks relatively rare. However, the rapid advancement of artificial intelligence and open-source image synthesis tools has made it significantly easier to fabricate realistic biometric forgeries (Damer et al., 2022).

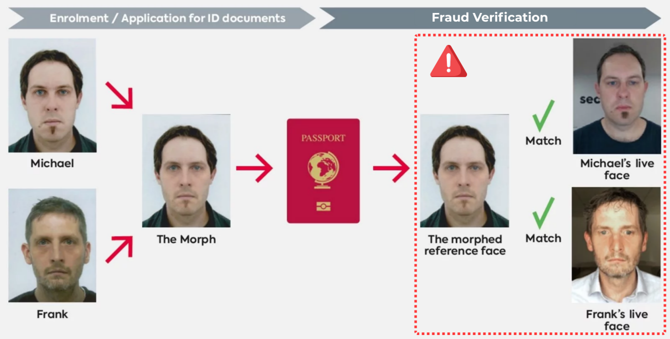

Fig. 1. Example of a face morphing attack where a morphed image matches both individuals during automated verification and fails to detect subtle feature inconsistencies (Secunet, 2021).

One of the most critical threats in this domain is the face morphing attack. In this technique, facial images are algorithmically blended to create a single composite that shares biometric similarities with multiple individuals. These morphed images can deceive both human observers and automated facial recognition systems, allowing attackers to fraudulently pass identity checks or obtain digital credentials without physical verification. As illustrated in Figure 1, a morphed image can simultaneously match two different individuals during automated verification, making it difficult to detect subtle feature inconsistencies that indicate tampering.

In the context of digital finance and online identity verification, such as eKYC (electronic Know Your Customer) and secure digital transactions, the risk is even greater. Current facial recognition systems are designed primarily to confirm identity matches but lack mechanisms to detect synthetic morphing patterns, making them vulnerable to impersonation-based fraud (Venkatesh et al., 2020). While conventional recognition methods are fast and widely used, they overlook subtle texture or structural inconsistencies. In contrast, existing morphing detection models can recognize such anomalies but often require large, labeled datasets and significant computational resources.

To address this gap, this research proposes a hybrid deep learning–based face morphing attack detection model designed to enhance the security of digital transactions. The framework integrates Single-image Morphing Attack Detection (S-MAD) and Differential Morphing Attack Detection (D-MAD) to ensure comprehensive protection. By leveraging convolutional neural networks (CNNs) and transfer learning, the S-MAD module detects morphing artifacts from a single facial image. In contrast, the D-MAD module compares a reference ID image with a selfie to identify inconsistencies between facial representations. Through this dual-stage approach, the model aims to accurately differentiate between bona fide and morphed facial inputs, providing a scalable and reliable safeguard against identity fraud and enhancing trust in modern biometric authentication systems.

II. Problem Statement

Recent advances in artificial intelligence and open-source image processing tools have made face morphing attacks increasingly realistic and challenging to detect. These attacks create composite facial images that resemble multiple individuals, enabling identity fraud in digital authentication systems. Traditional facial recognition models rely on surface-level similarity and often overlook subtle texture morphing artifacts (Hamza et al., 2022).

Existing S-MAD methods effectively identify morphing traces within a single image; however, they are limited when both a reference ID and a live selfie are available, as in eKYC and digital banking scenarios. D-MAD approaches address this by comparing two images, but often lack integration with single-image verification, reducing their general applicability (Tapia & Busch, 2023). To overcome these limitations, this research proposes a hybrid morphing attack detection framework that combines S-MAD and D-MAD. The S-MAD module, based on EfficientNet-B3 and trained on the SMDD dataset, detects morph artifacts at pixel and texture levels. In contrast, the D-MAD module, implemented as a Siamese network trained on FEI and tested on FRLL, utilizes cosine similarity to compare embeddings between reference and selfie images. This hybrid approach enhances detection accuracy across both single-image and paired-image verification, improving the reliability of eKYC and secure transaction systems against morphing-based identity fraud.

III. Objective

To address the problems stated above, this project aims to achieve the following objectives:

1) To develop a deep learning–based face morphing attack detection model capable of distinguishing between bona fide and morphed facial images.
2) To evaluate the performance of the proposed morphing detection model
3) To develop and present a web-based application that integrates the trained model into a simulated digital transaction environment.

IV. Data Science Methodology

This project adopts the CRISP-DM methodology to guide the data science lifecycle. Introduced in 1999, CRISP-DM provides a structured and widely adopted framework for analytics projects, ensuring consistency and reproducibility across industries (Saltz, 2021). It consists of six iterative phases: Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation, and Deployment, that collectively represent the complete process for developing and implementing data-driven solutions.

A. *Business Understanding*

The primary goal of this project is to enhance biometric security in digital transactions by developing a hybrid morphing attack detection framework that integrates Single-image Morphing Attack Detection (S-MAD) and Differential Morphing Attack Detection (D-MAD). The S-MAD module detects morphing artifacts in a single facial image, while the D-MAD module compares reference and selfie pairs to identify inconsistencies through feature similarity. This hybrid design ensures accurate detection in both single-image and paired verification scenarios, strengthening fraud prevention in eKYC and online banking systems.

B. *Data Understanding*

This phase focuses on exploring and analyzing the datasets used to train and evaluate the hybrid S-MAD and D-MAD frameworks. Five facial datasets are employed in this study: the Synthetic Morphing Attack Detection Development (SMDD) dataset from the Machine Intelligence Lab (Damer et al., 2022), the FEI Face Database (Thomaz & Giraldi, 2010) and its derived FEI Morph Dataset (Batskos et al., 2023; Di Domenico et al., 2023), and the AMSL Face Morph Image Dataset (Neubert et al., 2018) derived from the FRLL dataset (DeBruine & Jones, 2017) hosted on Figshare.

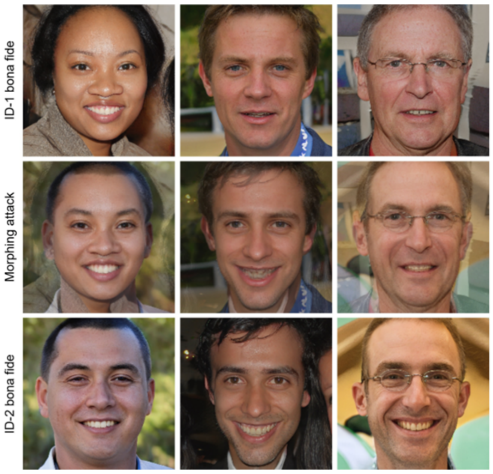

Fig. 2. Sample of the synthetic SMDD dataset (Damer et al., 2022). The top and bottom rows are bona fide samples. The middle row is samples of the morph created from the corresponding images above and below.

In Figure 2, the SMDD dataset serves as the foundation for S-MAD training, containing a large set of bona fide and morphed facial images generated using multiple morphing techniques to support single-image morphing detection. For D-MAD training, the FEI Face Database provides bona fide facial images, while the FEI Morph Dataset contributes corresponding morphed pairs created from the same source images. The AMSL Face Morph Image Dataset is generated using FRLL. It is employed for testing and cross-dataset evaluation, providing realistic, high-resolution morph and bona fide pairs that assess generalization performance.

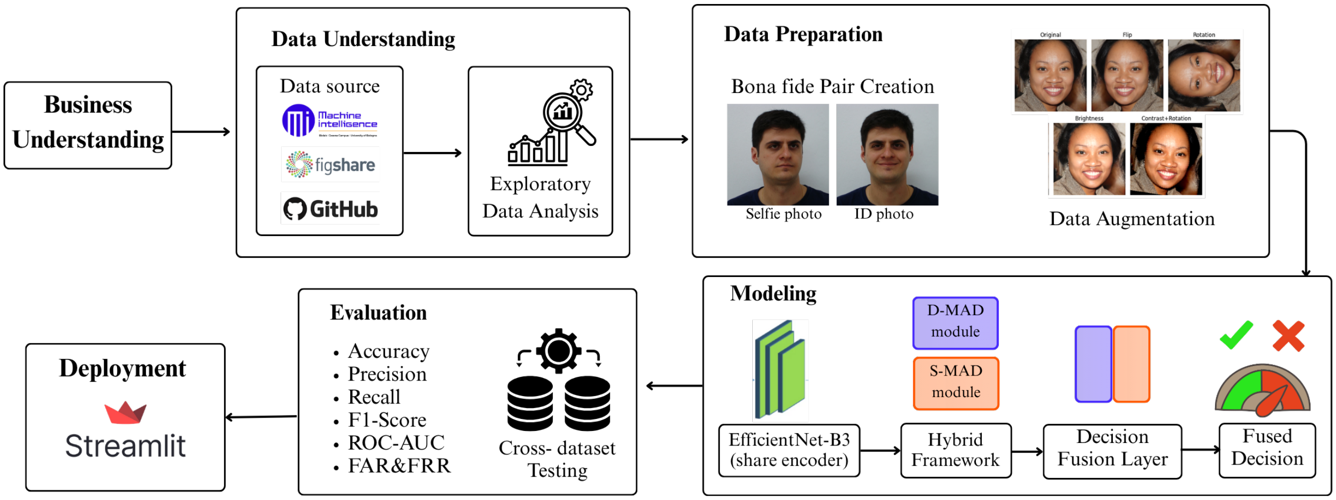

Fig. 3. Flowchart of Data Science Methodology

Fig. 4. Examples of bona fide image pairs from the FEI dataset captured at slightly different angles (Thomaz & Giraldi, 2010).

Exploratory Data Analysis (EDA) was conducted to assess image balance, facial alignment, and illumination consistency across all datasets. This combination of synthetic (SMDD), controlled (FEI), and realistic (FRLL) sources enables a comprehensive model learning and reliable performance evaluation for both S-MAD and D-MAD components.

C. Data Preparation

The Data Preparation phase ensures that all datasets are standardized and suitable for training and evaluating hybrid models. Each image first undergoes face detection and alignment to crop the facial region based on key landmarks such as the eyes and nose, ensuring consistent positioning across samples. The images are then resized and normalized to maintain uniform dimensions and pixel intensity distributions, providing a stable input for the neural network. In Figure 3, the flowchart of data preparation shows that data augmentation techniques, such as horizontal flipping and slight rotation, are applied to improve generalization and minimize overfitting (Shorten & Khoshgoftaar, 2019).

For the D-MAD module, additional preprocessing involves constructing pairs using the FEI Face Database and the FRLL dataset to generate bona fide pairs, where two authentic images of the same individual, captured at slightly different angles, are paired together. Figure 4 shows that the FEI Face Database with frontal and three-quarter poses is used. Morph pairs are directly utilized from the FEI Morph Dataset and the AMSL Face Morph Image Dataset, which contain pre-generated blended facial images. This structured preprocessing pipeline ensures that both the S-MAD and D-MAD modules receive balanced, high-quality inputs, enabling robust and generalizable morph detection performance across diverse datasets.

D. Modeling

In this phase, two deep learning models are developed to form the hybrid morphing attack detection framework: one for S-MAD and another for D-MAD.

The S-MAD module utilizes an EfficientNet-B3 backbone, pretrained on ImageNet and fine-tuned using the SMDD dataset, to classify bona fide and morphed images. It focuses on pixel- and texture-level features to identify subtle morphing artifacts that are not visible to the human eye (Jia et al., 2023; Li et al., 2019).

The D-MAD module utilizes a Siamese network with shared EfficientNet-B3 encoders, which process the reference ID and selfie images in parallel. Their embeddings are refined using an ArcFace layer to increase angular margin separation and enhance discriminative ability. The cosine similarity is then applied to measure the distance between features, as larger distances indicate morphing inconsistencies (Tapia & Busch, 2023).

To improve efficiency and ensure consistent feature representation, the hybrid framework reuses a shared encoder across both modules. The ID image is processed once, and its embedding is cached for reuse when the D-MAD module compares it against the selfie input. This design minimizes redundant computation while maintaining robust feature alignment between S-MAD and D-MAD outputs.

Finally, a Decision Fusion Layer integrates the prediction probabilities from both the S-MAD and D-MAD modules to produce the final classification result. Figure 5 illustrates that the layer implements decision-level fusion by combining the confidence scores from the single-image and pairwise detec-

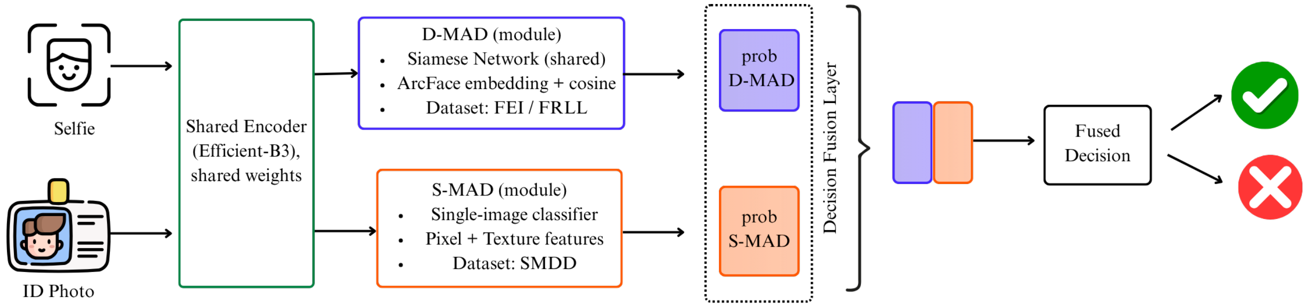

Fig. 5. Hybrid S-MAD and D-MAD architecture using a shared encoder for feature extraction and decision fusion.

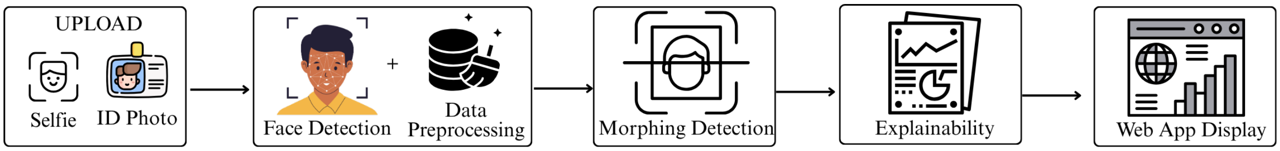

Fig. 6. Deployment workflow of the hybrid morphing attack detection system.

tion branches using a weighted or threshold-based scheme. The fused decision determines whether the verification instance is bona fide or morphed, ensuring higher reliability and robustness for secure digital authentication workflows

E. Evaluation

The evaluation phase measures the performance and reliability of the proposed hybrid framework using both standard and biometric metrics. Key metrics, including accuracy, precision, recall, F1-score, and ROC-AUC, are used to evaluate the classification performance of the models. Additionally, False Acceptance Rate (FAR) and False Rejection Rate (FRR) are applied to evaluate the model's ability to minimize false verification outcomes, which is crucial in biometric authentication.(Damer et al., 2022)

The S-MAD module is evaluated using the SMDD dataset. In contrast, the D-MAD module is validated by cross-dataset testing using FRLL to examine generalization between different facial sources (Tapia & Busch, 2023). Visualization tools such as confusion matrices and embedding distance plots are used to analyze prediction behavior and class separability. The results demonstrate that the hybrid S-MAD and D-MAD models maintain high accuracy and robustness under both controlled and real-world conditions, confirming their effectiveness for secure digital identity verification.

F. Deployment

Figure 6 illustrates the deployment of the workflow of the hybrid morphing attack detection system. The hybrid morphing attack detection framework is deployed as a web-based application integrating FastAPI for backend inference and Streamlit for user interaction. Users upload a selfie and ID photo, which undergoes face detection and preprocessing to align and normalize the inputs.

The model then performs morphing detection using hybrid S-MAD and D-MAD models. The S-MAD module verifies the authenticity of the ID image. In contrast, the D-MAD module compares the ID and selfie using a Siamese network with ArcFace and cosine similarity to detect morph inconsistencies. Their outputs are fused in the Decision Fusion Layer to produce a single final classification, bona fide or morphed.

An Explainability module presents confidence scores, similarity values, and heatmap visualizations highlighting detected morph regions. Finally, all results are displayed in the Streamlit web interface, providing clear, interpretable feedback for secure eKYC and digital verification applications.

V. Conclusion

This project develops a hybrid morphing attack detection model that combines S-MAD and D-MAD to improve the security of digital identity verification. The S-MAD model focuses on analyzing single images to detect fine morphing artifacts using an EfficientNet-B3 backbone trained on the SMDD dataset. The D-MAD model, on the other hand, compares the reference ID and selfie images through a Siamese network with ArcFace and cosine similarity to identify any inconsistencies. Both modules share a common encoder to maintain consistent feature extraction and reduce redundant computation. Their outputs are merged in a Decision Fusion Layer to produce the final classification result, determining whether the verification attempt is bona fide or morphed. The model is deployed through a FastAPI–Streamlit web interface that allows users to upload their images and view real-time detection results, including confidence scores and visual explanations. Overall, this project presents a practical, interpretable, and efficient solution to prevent morph-based identity fraud in eKYC and secure digital transactions.

References

Batskos, I., Spreeuwers, L., & Veldhuis, R. (2023). Visualizing landmark-based face morphing traces on digital images. *Frontiers in Computer Science*, *5*, 981933.

Damer, N., Lopez, C. A. F., Fang, M., Spiller, N., Pham, M. V., & Boutros, F. (2022). Privacy-friendly synthetic data for the development of face morphing attack detectors. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 1605–1616. https://doi.org/10.1109/cvprw56347.2022.00167

DeBruine, L., & Jones, B. (2017). Face Research Lab London Set. https://doi.org/10.6084/m9.figshare.5047666.v5

Di Domenico, N., Borghi, G., Franco, A., & Maltoni, D. (2023). Combining identity features and artifact analysis for differential morphing attack detection. *International Conference on Image Analysis and Processing*, 100–111.

Hamza, M., Tehsin, S., Humayun, M., Almufareh, M. F., & Alfayad, M. (2022). A comprehensive review of face morph generation and detection of fraudulent identities. *Applied Sciences*, *12*(24), 12545. https://doi.org/10.3390/app122412545

Jia, C.-K., Liu, Y.-C., & Chen, Y.-L. (2023). Face morphing attack detection based on high-frequency features and progressive enhancement learning. *Frontiers in Neurorobotics*, *17*. https://doi.org/10.3389/fnbot.2023.1182375

Li, Y., Yang, X., Sun, P., Qi, H., & Lyu, S. (2019). Celeb-DF: a large-scale challenging dataset for DeepFake forensics. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.1909.12962

Neubert, T., Makrushin, A., Hildebrandt, M., Kraetzer, C., & Dittmann, J. (2018). Extended stirtrace benchmarking of biometric and forensic qualities of morphed face images. *IET Biometrics*, *7*(4), 325–332. https://doi.org/10.1049/iet-bmt.2017.0147

Saltz, J. S. (2021). CRISP-DM for Data Science: strengths, weaknesses and potential next steps. *2021 IEEE International Conference on Big Data (Big Data)*, 2337–2344. https://doi.org/10.1109/bigdata52589.2021.9671634

Secunet. (2021, December). Man or morph? – protection against identity fraud at border control. https://secuview.secunet.com/en/blog-data/mensch-oder-morph

Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on Image Data Augmentation for Deep Learning. *Journal Of Big Data*, *6*(1). https://doi.org/10.1186/s40537-019-0197-0

Tapia, J., & Busch, C. (2023). Face feature visualisation of single morphing attack detection. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2304.13021

Thomaz, C. E., & Giraldi, G. A. (2010). A new ranking method for principal components analysis and its application to face image analysis. *Image and Vision Computing*, *28*(6), 902–913. https://doi.org/https://doi.org/10.1016/j.imavis.2009.11.005

Venkatesh, S., Ramachandra, R., Raja, K., & Busch, C. (2020). Face Morphing Attack Generation & amp; Detection: A Comprehensive Survey. *arXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2011.02045