



Deepfake-Proof eKYC System for Secure Identity Verification

As deepfake technology advances at an unprecedented pace, digital identity verification faces critical vulnerabilities. Fraudsters increasingly exploit AI-generated synthetic media to bypass traditional security measures, threatening financial systems and personal data protection.

Our solution: an intelligent AI pipeline combining deepfake detection, facial identity verification, and liveness validation—leveraging the comprehensive Sentinel_FaceV1 dataset to build a robust defense against sophisticated identity fraud.

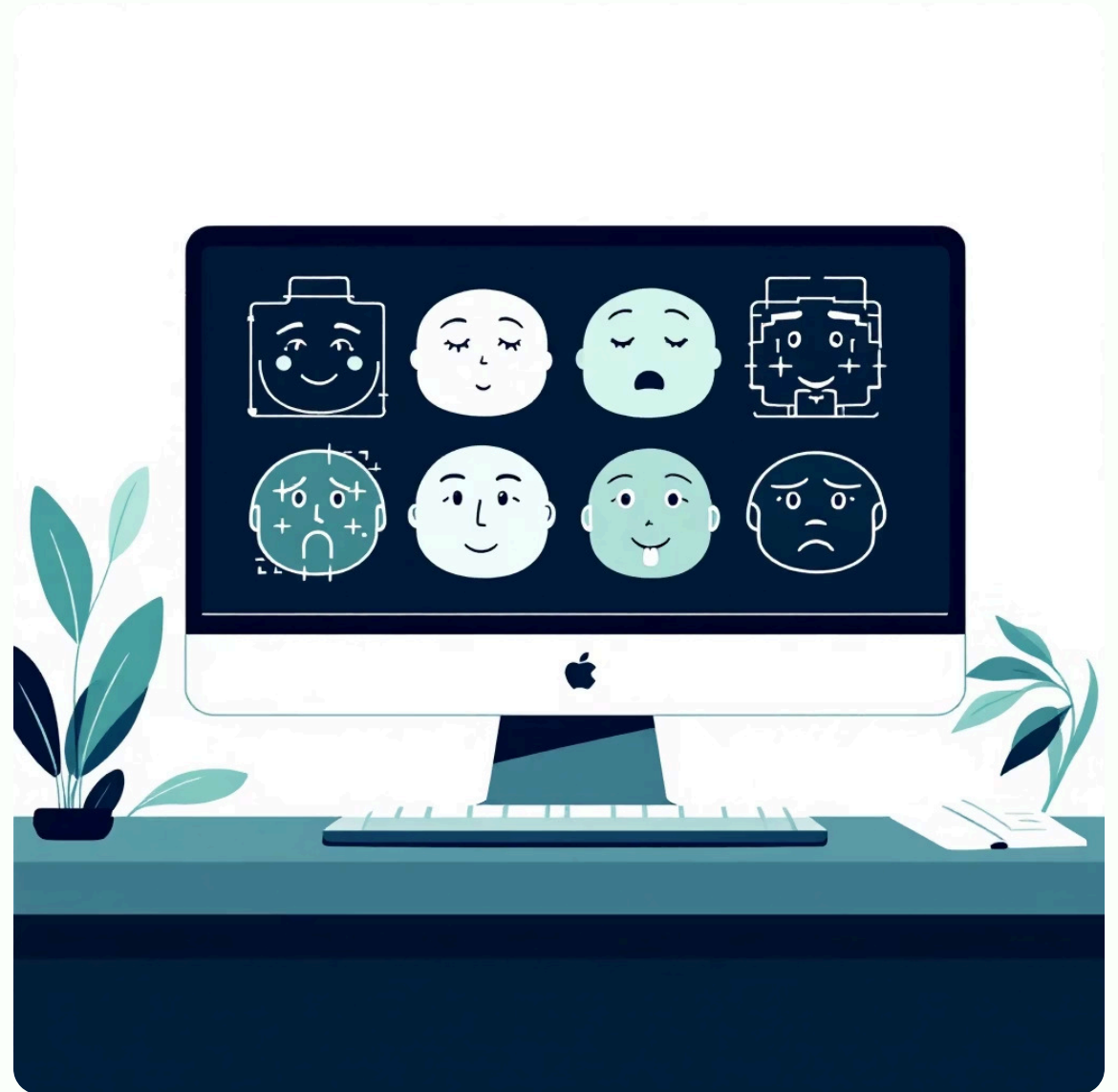
Dataset Engineering and Advanced Preprocessing

Sentinel_FaceV1 Dataset

A diverse collection balancing real and synthetic faces across multiple demographics, lighting conditions, and deepfake generation techniques—providing the foundation for robust model training.

Data Augmentation Strategy

- Dynamic lighting adjustments
- Color space variations
- Geometric transformations
- Noise injection patterns



01

Face Detection

MTCNN identifies facial landmarks with precision, ensuring accurate region extraction even in challenging conditions.

02

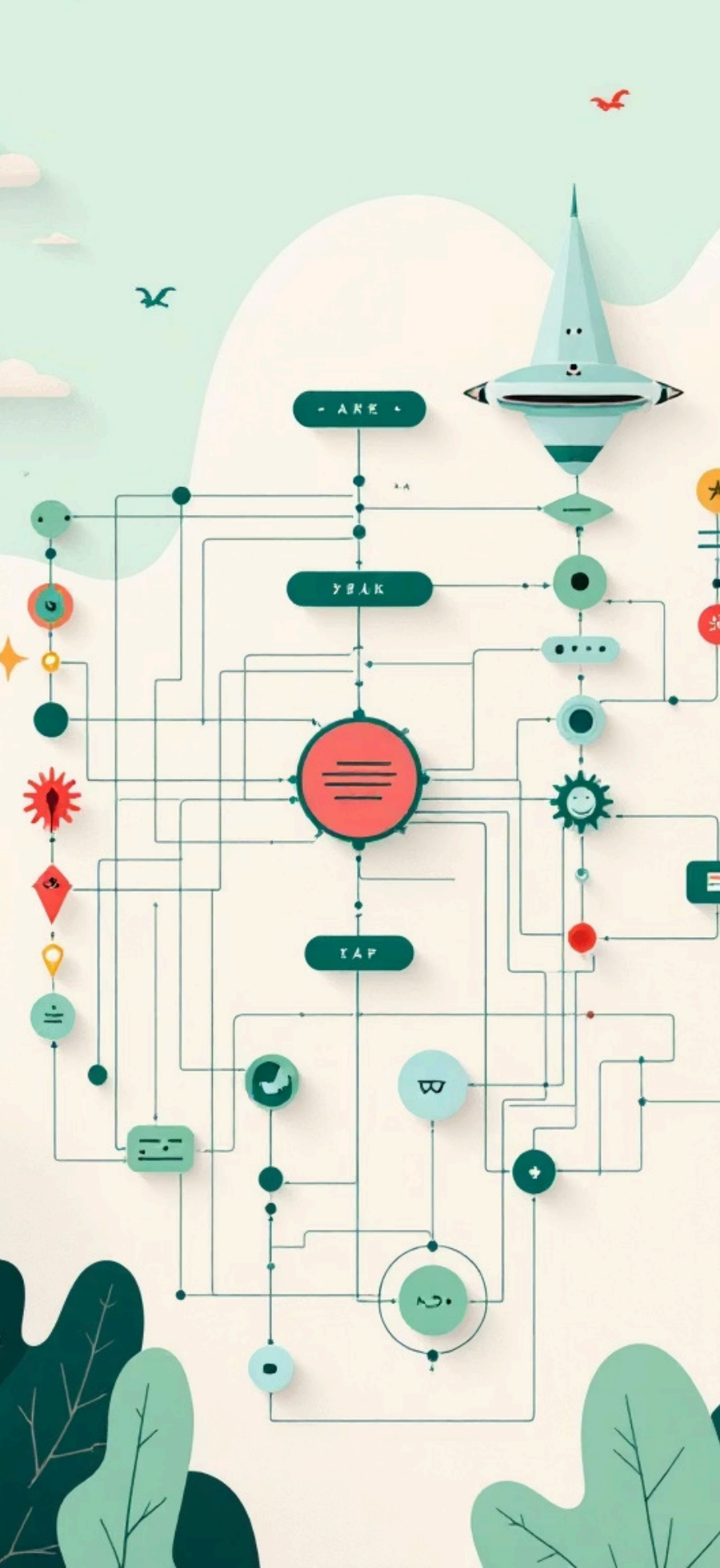
Alignment

Geometric normalization standardizes face orientation, scale, and position for consistent model input.

03

Frequency Analysis

DCT transforms reveal hidden artifacts in synthetic images that evade spatial domain detection—exposing deepfake signatures invisible to human eyes.



Intelligent Model Architecture and Training Pipeline

EfficientNet-B0 Backbone

Fine-tuned on DCT frequency features to detect subtle deepfake artifacts through efficient compound scaling and optimized parameter usage.

InceptionResnetV1

Generates discriminative face embeddings for identity verification, trained on millions of faces to capture unique biometric signatures.

Unified Pipeline

Seamlessly integrates detection and verification streams with shared preprocessing and coordinated inference logic.

Training Configuration

- **Loss Function:** Binary cross-entropy with focal loss for class imbalance
- **Optimizer:** AdamW with cosine annealing scheduler
- **Batch Size:** 32 images per iteration

Infrastructure

- **Hardware:** NVIDIA GPU acceleration
- **Training Duration:** 50 epochs with early stopping
- **Validation:** 80/20 train-test split with stratification

Inference Engine and Explainable AI



Input Processing

System accepts images and video streams, automatically extracting frames for comprehensive analysis.



Multi-Stage Analysis

Parallel deepfake detection and identity matching with confidence scoring for each prediction.



Verification Output

Binary authenticity classification with detailed match scores and liveness indicators.

Identity Match Scoring

Cosine similarity between face embeddings produces a confidence metric ranging from 0 to 100, with adaptive thresholds balancing security and user experience.

Liveness Detection

Frame-to-frame variance analysis identifies static or replayed content by measuring temporal inconsistencies across video sequences—distinguishing genuine human movement from artificial playback.



Grad-CAM Visualization

Gradient-weighted Class Activation Mapping highlights regions influencing model decisions—providing transparency and building trust through interpretable AI that shows *where* the model detects anomalies, not just *what* it predicts.

Results, Live Demonstration, and Future Directions

Could not deploy in time. Notebook for the training of model attached in zip file.

Roadmap for Enhancement

Advanced Liveness

Integrate 3D depth sensing and challenge-response mechanisms for enhanced spoofing resistance.

Cross-Dataset Validation

Test generalization on FaceForensics++, Celeb-DF, and DFDC datasets to ensure robustness.

Production Deployment

Containerize with Docker, implement API endpoints, and optimize for edge device inference.

Continuous Learning

Build feedback loops to retrain models on emerging deepfake techniques and adversarial examples.