TED UNIVERSITY

Faculty of Engineering

Department of Computer Engineering

CMPE 491 – High Level Design Report

SIMA

27.12.2024

İbrahim Ataberk Kabasakal – 17032700808

Furkan Tosun – 64312119164

Ahmet Tunç – 30274207140

Nusret Mert Yaşar - 52786056504

Contents

1. Introduction

1.1 Purpose of the System

The **Secure Image Masking Algorithm (SIMA)** is a comprehensive software solution designed to ensure privacy in images and videos by anonymizing facial features. Leveraging advanced artificial intelligence techniques, SIMA modifies key facial landmarks—such as eyes, nose, and mouth—to make individuals unrecognizable while retaining the contextual integrity of the image. This solution is particularly suited for applications in social media, surveillance, news media, and other sectors requiring compliance with privacy regulations like GDPR.

Beyond its immediate functionality of anonymizing faces in images and videos, SIMA aims to become a versatile tool that supports privacy across various domains. In the long term, the system could integrate seamlessly with social media platforms to automatically anonymize user-uploaded content, protecting individuals' privacy without manual intervention. Additionally, SIMA has the potential to be adopted by governmental and non-governmental organizations for protecting the identities of individuals in sensitive scenarios, such as public protests, humanitarian crises, or whistleblower situations. This extended vision underlines SIMA's role as a pioneer in the intersection of AI and privacy ethics, ensuring compliance with evolving global privacy standards while advancing societal trust in anonymization technologies.

1.2 Design Goals

1. **Accuracy:** Achieving precise detection and anonymization of facial features is a cornerstone of the SIMA system. The system aims to ensure that all facial features, such as eyes, nose, and mouth, are accurately identified and anonymized without introducing unnatural artifacts that might disrupt the image's visual integrity. This is particularly important in scenarios where maintaining the overall aesthetics and context of the image or video is essential, such as news broadcasts or public events. By leveraging advanced machine learning models and iterative training processes, the system seeks to minimize errors, even in challenging conditions like low-light environments or extreme facial poses.

2. **Real-Time Performance:** Real-time performance is critical for supporting high-throughput environments, such as live video streams or bulk image processing on social media platforms. The SIMA system is designed to process data at high speeds without compromising on the quality of anonymization. For instance, it targets the ability to anonymize up to 30 frames per second for live video feeds. Optimization techniques, such as model pruning and hardware acceleration with GPUs, ensure that latency is minimized. This allows for seamless integration with applications that require immediate anonymization, such as surveillance systems or live event streaming.

3. **Compliance:** The SIMA system is developed with strict adherence to global privacy standards, including the General Data Protection Regulation (GDPR) in Europe and the California Consumer Privacy Act (CCPA) in the United States. These regulations require robust safeguards for personal data, ensuring that individuals' identities are protected during and after processing. The system integrates compliance mechanisms such as audit trails, secure data handling, and consent management features to align with these legal frameworks. This ensures that organizations using SIMA can maintain full regulatory compliance while protecting user privacy.

4. **Scalability:** Scalability is a key focus of the SIMA system, ensuring that it can handle the growing demands of large datasets and high-resolution media. The architecture is designed to accommodate horizontal and vertical scaling, enabling it to process thousands of images or videos simultaneously without significant degradation in performance. For example, cloud-based deployment options and distributed processing frameworks ensure that the system remains responsive even under heavy workloads. Scalability also includes support for future expansion, allowing additional features or processing capabilities to be integrated without requiring major system overhauls.

5. **Security:** Security is paramount in protecting sensitive information handled by the SIMA system.To this end, the system employs end-to-end encryption standards, such as AES-256, to secure data during storage and transmission. Role-based access control (RBAC) and multi-factor authentication (MFA) are implemented to prevent unauthorized access. Additionally, comprehensive logging and monitoring mechanisms are in place to detect and respond to potential security breaches. The system's security features are designed to instill confidence in users and organizations that their data is being processed and stored securely.

6. **Sustainability:** Sustainability is a core design goal of the SIMA system, reflecting its commitment to minimizing environmental impact. By optimizing computational resources, the system reduces energy consumption during both training and inference processes. Techniques such as model compression, use of energy-efficient hardware, and batch processing ensure that resource usage is kept to a minimum. This approach not only lowers operational costs but also contributes to the global effort to reduce the carbon footprint of AI technologies. Sustainability ensures that the SIMA system remains viable for long-term use in a variety of applications.

7. **Adaptability**: Ensure that the system can be seamlessly integrated across diverse platforms and devices, ranging from cloud servers to local edge devices, to accommodate various user needs and operational environments. This adaptability will enhance SIMA's versatility and broaden its application domains.
8. **User-Centric Design**: Focus on delivering an intuitive interface that caters to users with varying levels of technical expertise. The design will prioritize ease of use, with features such as drag-and-drop uploads, real-time feedback on processing status, and interactive comparisons between original and anonymized outputs.

9. **Error Management**: Implement a robust error-handling mechanism that logs system failures and provides clear feedback to users. For example, in scenarios where face detection fails due to low-quality images, the system will notify users and offer suggestions for resolution. This ensures a smoother user experience while aiding in continuous system improvement.

1.3 Definitions, Acronyms, and Abbreviations

GANs (Generative Adversarial Networks)

Generative Adversarial Networks are a class of machine learning frameworks that consist of two neural networks: a generator and a discriminator. The generator creates synthetic data that closely resembles the training data, while the discriminator evaluates whether the data is real or generated. This adversarial process helps produce highly realistic outputs, such as anonymized faces, by training the generator to improve continuously. GANs are particularly useful in applications where realism and contextual accuracy are critical, such as in the SIMA system.

API (Application Programming Interface)

An API is a set of protocols, tools, and definitions that allow different software components to communicate with each other. In the context of SIMA, the API facilitates the interaction between the web interface, backend services, and AI models. It enables seamless operations such as image uploads, anonymization requests, and result retrieval, ensuring a smooth and efficient user experience.

GDPR (General Data Protection Regulation)

The General Data Protection Regulation is a legal framework enacted by the European Union to safeguard personal data and privacy. It mandates strict compliance measures, including data encryption, secure storage, and the right to be forgotten. The SIMA system adheres to GDPR guidelines by anonymizing facial data, ensuring that personally identifiable information (PII) is protected throughout the processing pipeline.

SIMA (Secure Image Masking Algorithm)

SIMA is the core system being developed to anonymize faces in images and videos using advanced AI techniques. The algorithm focuses on modifying key facial landmarks to render individuals unrecognizable while maintaining the overall integrity of the media. Its primary applications include social media, surveillance, news reporting, and compliance with privacy regulations such as GDPR.

FDF (Flickr Diverse Faces Dataset)

The Flickr Diverse Faces Dataset is a publicly available dataset used for training and evaluating AI models in tasks such as facial detection, recognition, and anonymization. It contains a wide variety of facial images with variations in pose, lighting, and background, making it an ideal resource for improving the robustness and accuracy of the SIMA system.

Pose Estimation

Pose estimation is the process of predicting the spatial configuration of key facial landmarks, such as eyes, nose, and mouth, to determine the orientation of a face. This is a crucial step in the SIMA system, as accurate pose estimation enables precise anonymization even in challenging conditions, such as extreme angles or partial occlusions.

Progressive GAN Training

A training method for GANs that starts with low-resolution images and progressively increases the resolution as training progresses. This approach enhances the stability and quality of the generated outputs, making it particularly useful for tasks that require high-resolution media processing, such as facial anonymization in the SIMA system.

Temporary Storage

A secure and time-limited storage mechanism for uploaded media files during the processing phase. In the SIMA system, temporary storage ensures that user data is automatically deleted after anonymization, reducing the risk of unauthorized access and lowering storage costs.

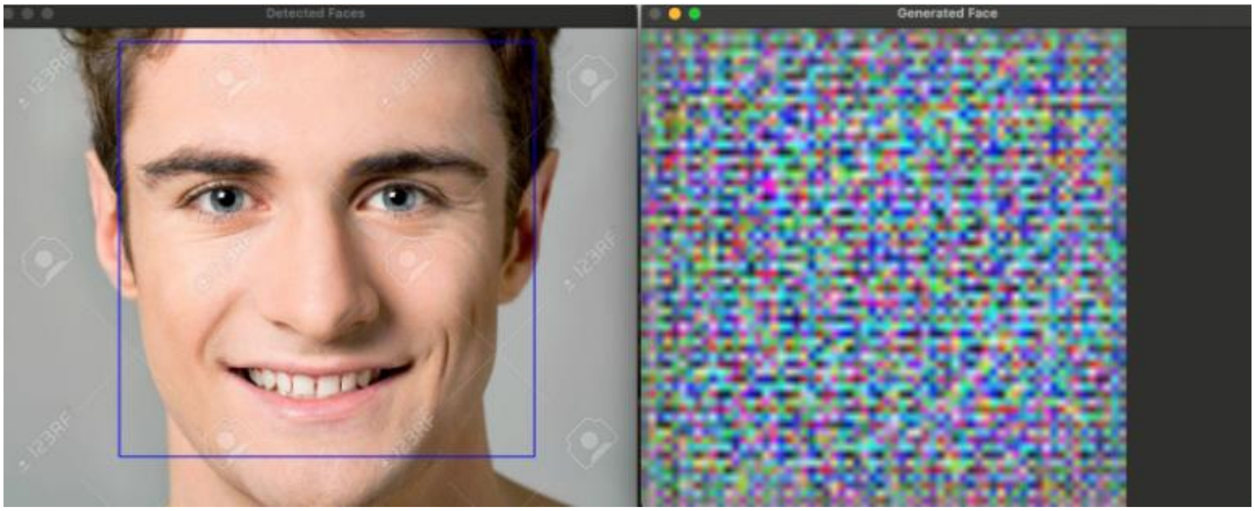
1.4 Overview

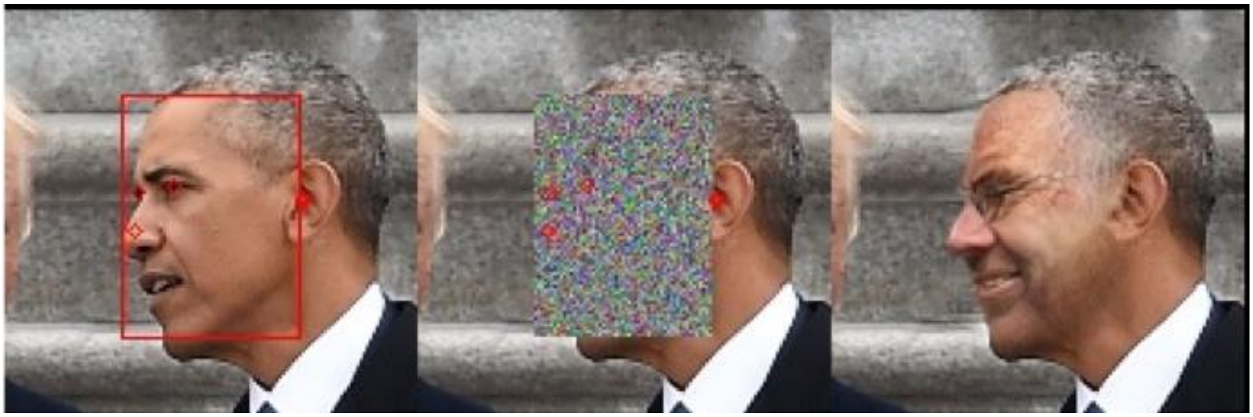
The Secure Image Masking Algorithm (SIMA) is designed to address the growing need for privacy protection in digital media by providing an efficient and user-friendly solution for facial anonymization. The system leverages state-of-the-art artificial intelligence techniques, such as Generative Adversarial Networks (GANs) and convolutional neural networks (CNNs), to detect and modify facial features in images and videos. By ensuring compliance with global privacy standards like GDPR and CCPA, SIMA serves as a reliable tool for various applications, including social media platforms, surveillance systems, and news media.

The project is structured into several phases to ensure systematic development and delivery. The initial phase focuses on research and requirement analysis, which includes exploring existing anonymization techniques and collecting diverse datasets like CelebA and FaceForensics for model training. The second phase involves designing and training AI models, followed by extensive testing to optimize the system for real-world scenarios. The final phase integrates the solution into a web-based platform, offering users an intuitive interface to upload their media files, process them, and download the anonymized results.

SIMA is built on a modular architecture that separates the AI processing, backend, and web interface subsystems, ensuring scalability and ease of maintenance. The system aims to provide real-time performance for high-throughput environments while minimizing computational costs and energy consumption. With features such as real-time feedback, secure file handling, and adaptive anonymization, SIMA is a step forward in protecting individual privacy in an increasingly connected digital world.

2. Current Software Architecture





Top Image (Current System):
The current system focuses on face detection and facial landmark identification. Using OpenCV and MediaPipe, the key facial features are detected, and basic anonymization is applied. Random offsets are added to the landmark points to distort recognizable features of the face. This demonstrates the system's capability to perform basic anonymization tasks. However, the current solution does not yet provide advanced anonymization.

Bottom Image (Target System):
The target system aims to produce fully anonymized and context-appropriate faces. With models like DeepPrivacy GANs, faces will be realistically yet unrecognizably modified while preserving the overall structure of the image. Additionally, the system is expected to handle complex scenarios such as masked faces, varying angles, and low-light conditions successfully.

Conclusion:
The current system delivers basic anonymization, but achieving the desired level requires implementing GAN-based anonymization, edge case handling, and system optimizations.

A. Libraries and Technologies Used

A.1 OpenCV

OpenCV is extensively used in the current SIMA system for real-time image processing and geometric transformations. It serves as the backbone for detecting and localizing faces within images and videos. OpenCV's CascadeClassifier is utilized to perform face detection, providing bounding boxes around detected facial regions. Additionally, OpenCV is used to preprocess images by resizing, cropping, or converting them to grayscale, which is essential for improving the efficiency and accuracy of subsequent facial landmark detection and anonymization processes. The library also facilitates geometric transformations such as affine warping, which is critical for modifying and repositioning facial features during the anonymization process.

A.2 MediaPipe

MediaPipe plays a crucial role in the SIMA system by enabling highly accurate facial landmark detection. Using its pre-trained models, MediaPipe identifies critical facial points, such as the eyes, nose, mouth, and jawline, which are vital for the anonymization process. The landmarks detected by MediaPipe are used as input for further modifications, ensuring that anonymization is precise and contextually appropriate. For example, MediaPipe's ability to handle challenging conditions, such as tilted faces or partial occlusions, significantly enhances the system's robustness. This integration allows the system to operate effectively across diverse scenarios, from low-light environments to high-resolution images.

A.3 NumPy

NumPy is employed in the SIMA system to perform mathematical operations on the detected facial landmarks. Once MediaPipe identifies the facial landmarks, NumPy is used to calculate distances, angles, and offsets required for modifying these landmarks. For instance, random offsets are added to the coordinates of specific landmarks to create realistic distortions while ensuring that the modified points remain within the image boundaries. NumPy also supports efficient handling of large datasets during batch processing, enabling the system to scale for high-throughput use cases such as anonymizing multiple images simultaneously.

A.4 DeepPrivacy GANs

DeepPrivacy GANs are integrated into the current SIMA system to provide advanced anonymization capabilities. This generative model replaces facial regions with synthetic yet realistic outputs, ensuring that the original identity is unrecognizable. GANs are particularly effective in scenarios where traditional methods like blurring or pixelation fail to preserve contextual accuracy. By training on diverse datasets, DeepPrivacy GANs generate anonymized facial features that seamlessly blend into the surrounding image. In the current system, this model operates as an optional layer, allowing users to choose between simple geometric transformations and generative anonymization for enhanced privacy.

A.5 Flask/Django

Flask and Django provide a flexible and robust backend API architecture for the SIMA system. These web frameworks manage the communication between the user interface and the AI model through RESTful APIs, enabling efficient data transfer and integration. Flask offers lightweight and modular functionality, making it ideal for rapid prototyping and specific microservices. On the other hand, Django is employed for more comprehensive features such as user authentication, database management, and deployment. Together, these frameworks support file uploads, real-time processing requests, and secure handling of user data, ensuring the backend operates seamlessly and efficiently.

B. Limitations of the Current System

- **Accuracy:** Errors occasionally occur in detecting facial features under complex poses or partial occlusions. This is especially problematic in images with faces taken from various angles or partially covered.
- **Latency:** Processing delays are experienced in scenarios involving video streams or batch uploads. This negatively impacts user experience, particularly in real-time video processing applications.
- **Adaptability:** The system lacks sufficient support for extreme poses or challenging conditions like low light. This hinders consistent performance across diverse environments.
- **Energy Efficiency:** The current system consumes high energy due to computationally intensive operations. This poses sustainability challenges, especially in large-scale applications.
- **Integration Challenges:** The current system faces difficulties integrating with various platforms and applications. This limits the widespread adoption of SIMA.

3. Proposed Software Architecture

3.1 Overview

The updated architecture focuses on modularity, efficiency, and scalability, dividing the system into four main subsystems. By leveraging advancements in GAN models and cloud infrastructure, the system aims to provide acceptable performance across diverse use cases. However, it's crucial to understand that actual performance may vary significantly depending on factors such as image complexity, dataset size, hardware limitations, and server load, and may fall short of initially defined targets.

3.2 Subsystem Decomposition

AI Processing Subsystem

- **Components:** DeepPrivacy GANs (or similar) for anonymization and custom-trained neural networks for landmark detection.
- **Input/Output:** Processes images/videos to generate anonymized versions while attempting to preserve contextual accuracy. The degree of accuracy will depend on model training, input data, and the limitations of the models used. Initially, errors and artifacts may be observed, especially in complex scenes and with diverse facial expressions.
- **Enhancements:** Use of progressive GAN training with the goal of achieving high-resolution outputs and incorporating conditional information like background and pose. These are challenging areas and may require iterative development and significant research. Achieving these goals may take time and cannot be guaranteed.
- **Face Detection**: Using OpenCV's CascadeClassifier to detect faces in input media.
- **Landmark Processing**: Leveraging MediaPipe to extract and modify facial features for precise anonymization.
- **Anonymization**: Employing DeepPrivacy GANs to generate synthetic, anonymized faces when required.
- **Scalability**: Supports batch processing and high-throughput environments, allowing seamless integration with larger systems like cloud-based platforms or local edge devices.

Web Interface Subsystem

- **Components:** React.js for frontend development.
- **Functionality:** Attempts to provide users with a user-friendly interface for uploading images, viewing results, and accessing logs. User experience will be a key focus during development. However, it may be challenging to provide a perfect user experience across all different user profiles and usage scenarios.
- **Enhancements:** Real-time progress updates and responsive design are planned. The performance of these features may vary depending on network conditions, user hardware, and browser compatibility issues.
- **File Upload and Management**: Allows users to upload images and videos through a web-based interface.
- **Real-Time Status Updates**: Displays the progress of processing tasks and provides immediate feedback.
- **User Authentication**: Implements secure login and registration features to ensure that only authorized users can access the system.

- **Result Delivery**: Facilitates easy download of anonymized results once processing is complete.

Backend Subsystem

- **Components:** Flask/Django backend with PostgreSQL database and AWS S3 storage.
- **Functionality:** Manages file uploads, API requests, and data storage with a focus on security. However, absolute security cannot be guaranteed, and there is always a risk of security breaches.
- **Enhancements:** Includes RESTful APIs for potential external integrations. The stability and performance of these APIs will be improved over time through testing and optimization.

Administrator Subsystem

- **Components:** Administrative dashboards for monitoring performance and auditing tasks.
- **Functionality:** Aims to enable compliance verification, system diagnostics, and feedback tracking. Specific tools and metrics will be refined during development. Initially, the scope and functionality of these tools may be limited.

3.3 Hardware/Software Mapping

- **Frontend:** Runs on browsers, developed with React.js.
- **Backend:** Hosted on cloud servers, employing Flask/Django (or similar).
- **GPU Infrastructure:** NVIDIA GPUs (or similar) will be used for AI model execution. While real-time processing performance is a goal, model complexity, image size, and instantaneous server load can significantly impact performance. Hardware failures can also negatively affect performance.
- **Database:** PostgreSQL for data management.

3.4 Persistent Data Management

- **Temporary Media Storage:** Securely holds uploaded files for a limited period with automatic deletion policies. Although data loss risk is minimized during this period, unforeseen system errors or hardware failures could lead to data loss.
- **Feedback Data:** Stores anonymized usage logs for system improvement. The usefulness of this data will depend on its quality, volume, and whether it's properly anonymized.
- **User Data:** Retains minimal information under strict encryption standards, with specific retention policies to be determined based on legal and operational requirements. Absolute data security cannot be guaranteed, and there is always a risk of unauthorized access.

3.5 Access Control and Security

- **Authentication:** Multi-factor authentication and role-based access control are planned. However, the effectiveness of these methods depends on users adopting strong passwords and safeguarding their credentials.

- **Encryption:** End-to-end AES-256 encryption (or a similarly robust standard) will be used for file transfers and storage. Nevertheless, if encryption keys are compromised or weaknesses are discovered in the encryption algorithms, data may be at risk.
- **Audit Trails:** Comprehensive logging is intended for monitoring activities and detecting anomalies. The effectiveness of anomaly detection will depend on the sophistication of the logging and analysis methods, and on accurately identifying suspicious behavior patterns in the system.
- **Authentication and Authorization**: Utilizes multi-factor authentication (MFA) for user login and API key-based access for secure interactions between subsystems.
- **Data Retention Policies**: Implements strict rules for temporary storage, ensuring that uploaded media files are automatically deleted after processing is complete.

3.6 Global Software Control

The SIMA system's global software control strategy aims for flexibility and scalability.

- **Microservices Architecture:** Each subsystem (AI Processing, Web Interface, Backend, and Admin) is designed as an independent microservice. This structure aims to allow each subsystem to scale independently based on its specific requirements. However, ensuring data consistency and synchronization between microservices can be complex and may require additional development effort.
- **Auto Scaling:** Automatic adjustment of resources based on system load is planned. While system performance and availability during peak usage will be attempted to be preserved, sudden and unpredictable spikes in load may result in performance degradation.
- **Containerization:** Docker containers will be used to create a consistent development and deployment environment. This aims to reduce issues like "it works in development but not in production." However, container orchestration and management can introduce added complexity.
- **Continuous Integration/Continuous Deployment (CI/CD):** A CI/CD pipeline is planned to automate the testing and deployment of code changes. This aims to facilitate early error detection and provide rapid feedback loops. However, the CI/CD pipeline itself can be susceptible to errors and outages.

3.7 Boundary Conditions

The boundary conditions for the SIMA system are defined as cautious expectations established at the beginning of the project and may be updated throughout the development process. These conditions provide a framework for evaluating the system's performance and capabilities:

- **Processing Capacity:**

  - Target: To be able to process an average of 20-30 images per second.
  - Reality: The system will be designed to handle 20-30 images per second. However, the actual processing speed may vary significantly depending on hardware specifications, the complexity of the processed images (resolution, level of detail, etc.), model optimization, and the instantaneous system load. In

cases of heavy usage and complex images, performance may drop below this target, even below 10 images per second. While optimization efforts and performance testing will aim to exceed this value, it cannot be guaranteed.

- **Maximum Image Size:**

    - Target: To be able to process images up to HD resolution (1280x720 pixels) primarily.
    - Reality: The core functionality of the system will focus on images up to HD resolution. Support for higher resolutions (e.g., Full HD or 4K) may be added gradually in later development stages and as the system is optimized. It should be considered that processing time will be much longer at higher resolutions, and the system may become unstable.

- **System Availability:**

    - Target: 75.0% uptime.
    - Reality: Necessary infrastructure and backup systems will be put in place for high availability. However, hardware failures, software bugs, planned maintenance, or unforeseen external factors may cause system downtime. Response plans and procedures for such situations will be developed. The 90.0% value is set as an achievable target, but it should not be disregarded that downtime may exceed this rate.

- **Data Protection Compliance:**

    - Target: To comply with applicable data protection regulations (GDPR, etc.).
    - Reality: Necessary technical and administrative measures will be taken to meet the requirements under relevant legal regulations. Compliance will be ensured and maintained through legal counsel and regular audits. However, the complexity of legal regulations and evolving interpretations may pose compliance risks.

4. Subsystem Services

A. AI Processing Subsystem

GAN-Based Anonymization

- Utilizes DeepPrivacy GANs to produce realistic and context-appropriate anonymized outputs.
- Modifies facial features (e.g., eyes, nose, mouth) to ensure individuals are unrecognizable.
- Generates high-resolution outputs by considering conditional information like background and pose.
- Employs a U-net architecture to preserve background integrity and enhance image quality.

CNN-Based Feature Extraction

- Conducts detailed facial feature analysis using Convolutional Neural Networks (CNNs).
- Achieves high accuracy in face recognition and classification tasks.
- Ensures robust feature extraction under varying pose and lighting conditions.

Continuous Model Improvement

- Continuously evaluates and monitors model accuracy with benchmark datasets, such as the Flickr Diverse Faces Dataset (FDF).
- Updates models to improve performance in challenging scenarios, including complex poses and low-light environments.

Face and Landmark Detection

- Uses advanced neural networks for precise face and facial landmark detection.
- Integrates MediaPipe's high-accuracy pre-trained models.
- Identifies seven key facial points, including eyes, ears, shoulders, and nose.

Edge Case Handling

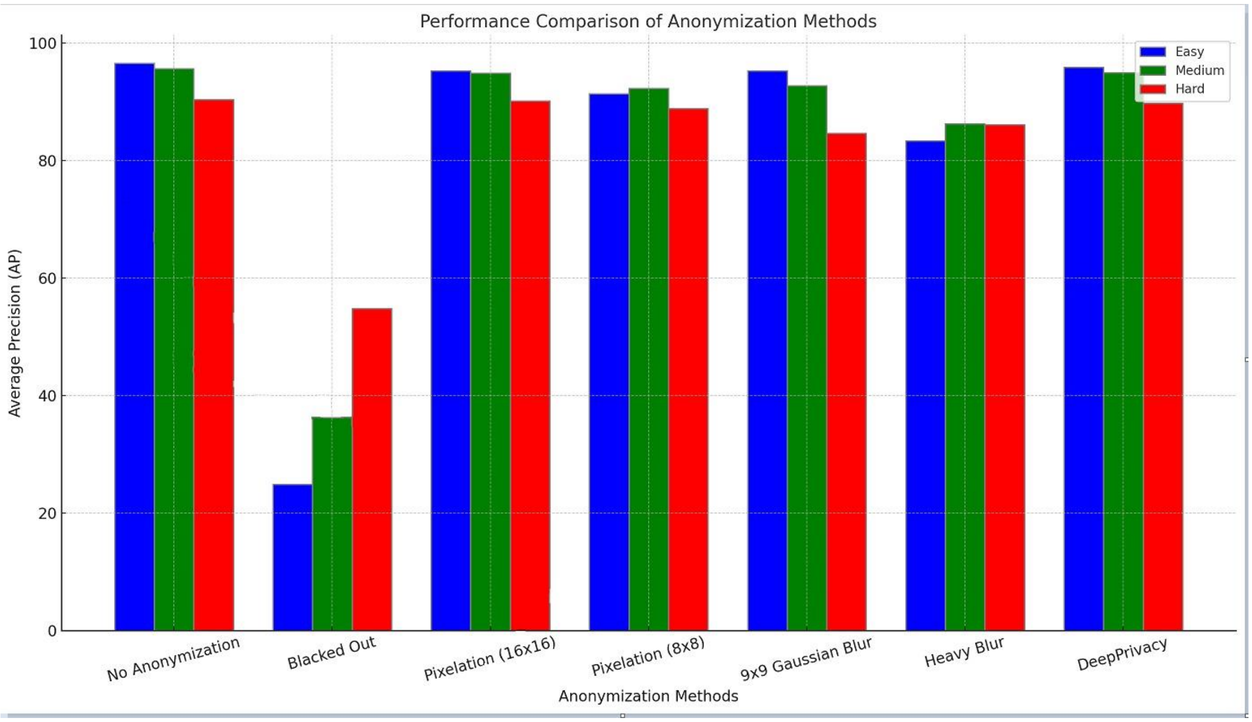
- Implements specialized models and algorithms to handle challenging scenarios such as occluded facial features (e.g., masks or sunglasses).
- Enhances accuracy for extreme head poses, ensuring consistent anonymization in non-ideal conditions.
- Improves system reliability across diverse environments by adapting to edge cases effectively.

AI Optimization for Low Resources

- Incorporates optimized AI models for deployment on low-resource environments such as mobile devices and edge computing systems.
- Ensures efficient anonymization tasks without requiring extensive computational resources.
- Extends usability to platforms with limited hardware capabilities, broadening the system’s applicability.

Dataset Usage

- Utilizes the Flickr Diverse Faces Dataset (FDF) as the primary resource for training AI models.
- Ensures diversity in facial features, lighting conditions, and poses, enhancing model generalization.
- Improves performance in real-world scenarios by training on a variety of challenging conditions.



B. Web Interface Subsystem

User Interaction

- Provides a responsive and accessible design built with React.js.
- Offers an intuitive interface for uploading images/videos, viewing results, and accessing logs.

Real-Time Updates

- Displays processing progress in real-time for user convenience.
- Presents anonymization results with side-by-side comparisons of original and processed media.

Responsive Design

- Ensures a flexible user interface adaptable to various devices and screen sizes.

Accessibility Features

- Ensures compliance with accessibility standards, such as the Web Content Accessibility Guidelines (WCAG).
- Implements screen reader support for visually impaired users.
- Provides high contrast modes to improve readability for users with low vision.
- Enables keyboard navigation for seamless interaction without a mouse.
- Focuses on creating an inclusive experience for users with disabilities.

Enhanced User Management

- Allows users to customize their accounts by saving preferences for frequently used settings.
- Enables users to review previous anonymization results for better task management.
- Streamlines workflows for frequent users by retaining personalized configurations.

C. Backend Subsystem

Secure File Management

- Manages file uploads and API requests securely using a Java Spring backend.
- Provides temporary media storage integrated with secure storage systems.

Scalability Optimization

- Handles concurrent tasks efficiently with an optimized backend structure.
- Ensures effective data management using a PostgreSQL database.

API Integration

- Delivers RESTful APIs for seamless integration with external systems.
- Ensures secure and fast data exchange through standardized protocols.

D. Administrator Tools

System Performance Monitoring

- Tracks system performance metrics in real-time, including processing times, resource utilization, and error rates.

Compliance Checks

- Ensures adherence to global privacy regulations like GDPR and CCPA.
- Conducts regular compliance audits and generates comprehensive reports.

User Feedback Integration

- Collects and analyzes user feedback to guide iterative system improvements.

Security and Auditing

- Implements multi-factor authentication and role-based access control for enhanced security.
- Provides comprehensive logging to monitor activities and detect anomalies effectively.

Customizable Metrics Dashboards

- Provides interactive dashboards for administrators to monitor key performance metrics in real time.
- Includes graphical representations of system performance for better visualization.
- Tracks user activity to identify usage patterns and optimize the system.
- Displays task completion rates and error trends for continuous improvement.
- Offers customizable views to focus on specific metrics as needed by administrators.

Proactive Security Measures

- Implements real-time anomaly detection to flag unusual activities.
- Generates automated alerts to promptly respond to potential threats.
- Enhances system security through proactive monitoring and threat mitigation.
- Helps maintain system integrity by addressing vulnerabilities as they occur.

Detailed Audit Trail

- Maintains a comprehensive audit trail of user actions, API calls, and system events.
- Supports regulatory compliance by logging all relevant activities.
- Facilitates debugging by providing detailed logs of system operations.
- Aids in forensic analysis by capturing a detailed history of system interactions.
- Ensures transparency and accountability through thorough activity tracking.

E. Reporting and Analytics Subsystem

Anonymization Reports

- Generates detailed reports that include success rates, error analysis, and processing times for anonymization tasks.
- Helps administrators evaluate the effectiveness of the system in various conditions.
- Offers exportable reports in common formats (e.g., PDF, Excel) for further analysis.

Insights into Trends and User Behavior

- Provides insights into anonymization trends, including frequently processed media types and user activity patterns.
- Analyzes user behavior to guide system optimizations and improve user experience.

- Identifies common issues or bottlenecks, allowing for targeted improvements in future updates.

Performance Tracking

- Tracks system-wide performance metrics, including resource utilization and task throughput.
- Highlights areas for improvement to maintain high operational standards.
- Supports long-term planning by providing historical data for performance comparisons.

5. Visual Demonstration of System Components

5.1 Facial Landmark Detection and Anonymization

The following images illustrate the functionality of the SIMA system in detecting and anonymizing facial features using advanced AI algorithms:

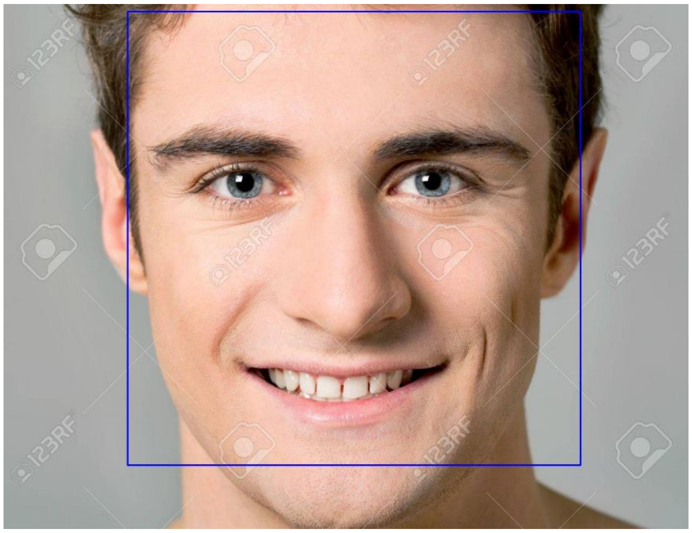
Original Image with Facial Landmarks Detected

The system begins by processing the input image using OpenCV for initial face detection. This is followed by MediaPipe's high-accuracy pre-trained models to identify precise facial landmarks. These landmarks correspond to critical facial points, including the eyes, nose, mouth, and jawline.

- **Detection Process**:

OpenCV's CascadeClassifier identifies the bounding box around the face.

MediaPipe extracts up to 256 detailed facial landmarks for precise mapping.



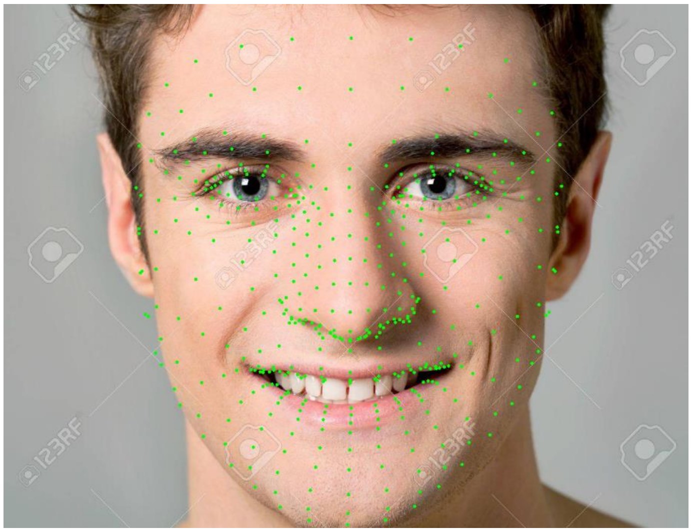
Modified Landmarks for Anonymization

The detected facial landmarks are then modified using algorithms implemented in the SIMA system. These modifications involve adding random offsets to landmark coordinates to ensure realistic yet anonymized distortions.

- **Modification Details**:

The modify_landmarks function uses NumPy to calculate and apply random offsets to the landmark points while keeping them within the image boundaries.

Key facial features such as eyes, nose, and mouth are adjusted to render individuals unrecognizable, while maintaining overall facial symmetry.



Final Anonymized Result

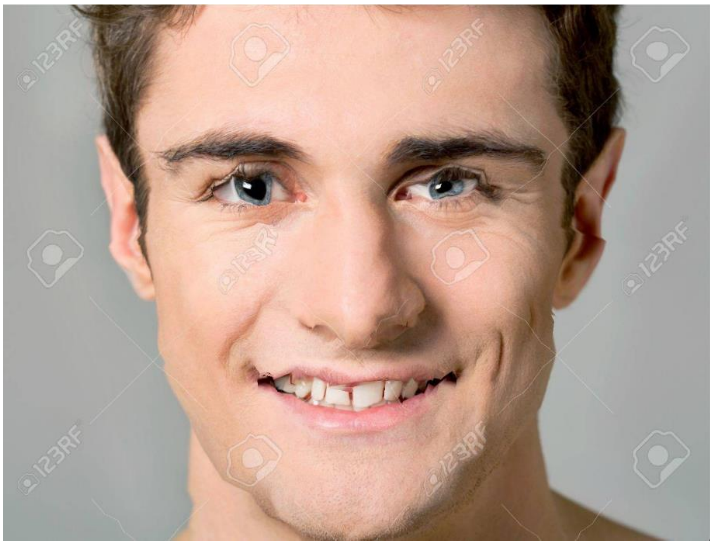
Using the modified landmarks, the system applies geometric transformations and pixel adjustments to the original image, creating the final anonymized output. The approach ensures that the anonymized face blends seamlessly into the image, preserving its background integrity.

- **Anonymization Process**:

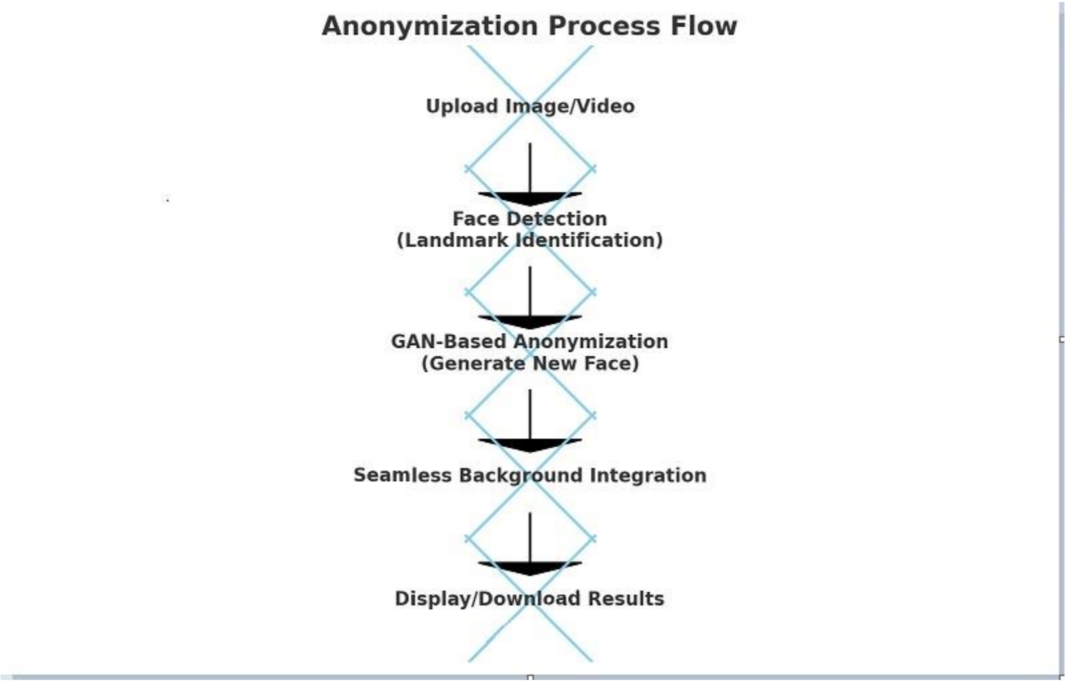
The warp_triangle function is employed to map triangular sections of the face to their new positions based on the modified landmarks.

Optional integration with DeepPrivacy GANs further enhances realism by generating high-quality synthetic face features if required.

The anonymized result is optimized to ensure no artifacts are introduced during the process.

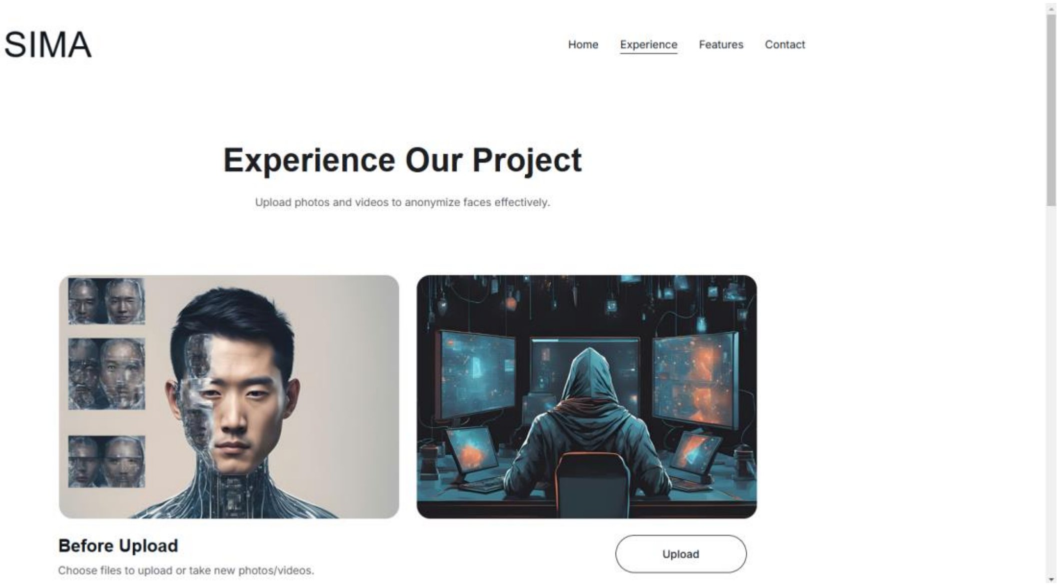


5.2 Workflow Diagram



A detailed workflow of the SIMA system:

1. **Upload Image/Video:** Users initiate processing by uploading files through the web interface.

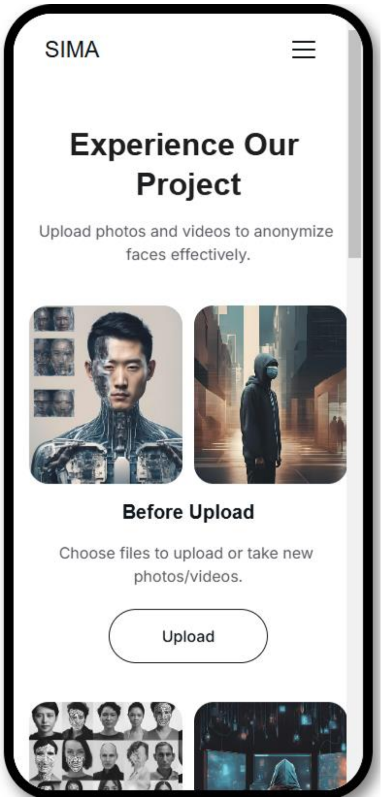


Experience Our Project

Upload photos and videos to anonymize faces effectively.

Before Upload

Choose files to upload or take new photos/videos.



Experience Our Project

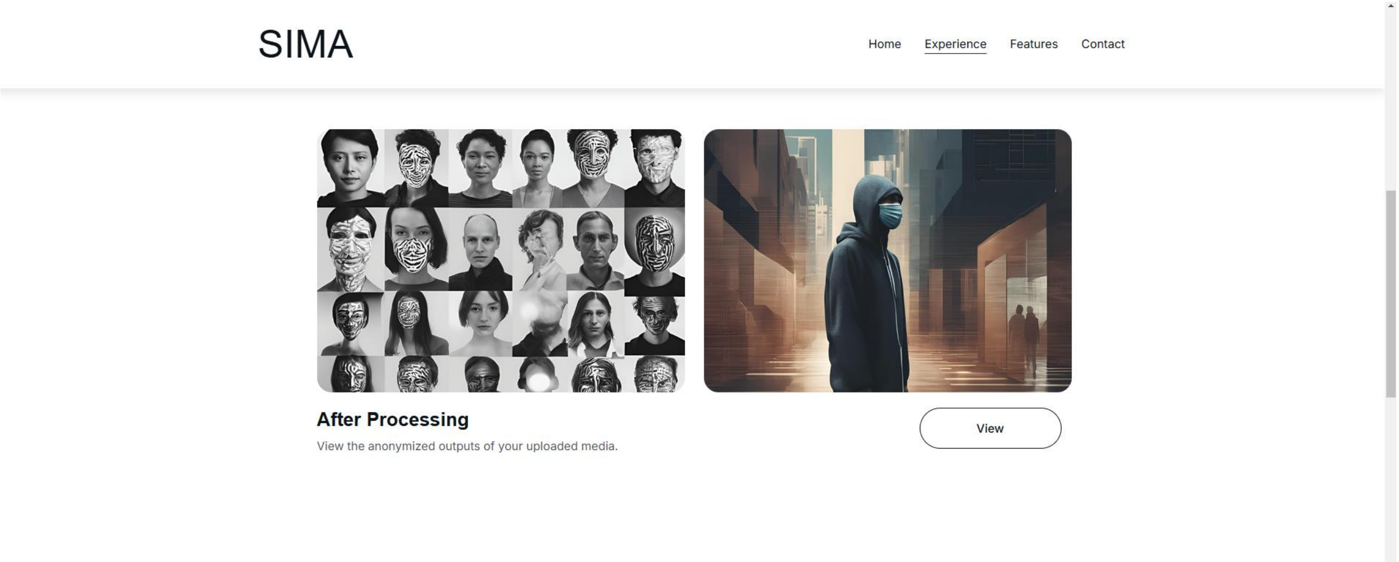
Upload photos and videos to anonymize faces effectively.

Before Upload

Choose files to upload or take new photos/videos.

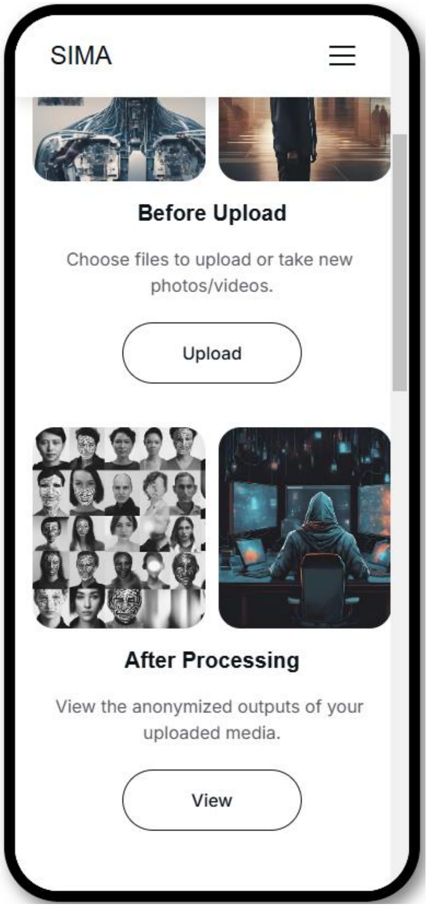
2. **Face Detection:** The AI subsystem identifies key facial regions for processing.
3. **Landmark Modification:** Detected landmarks are shifted to anonymize facial features effectively.

4. **Result Display:** Final outputs are displayed, with options for download and further analysis.



After Processing

View the anonymized outputs of your uploaded media.



Before Upload

Choose files to upload or take new photos/videos.

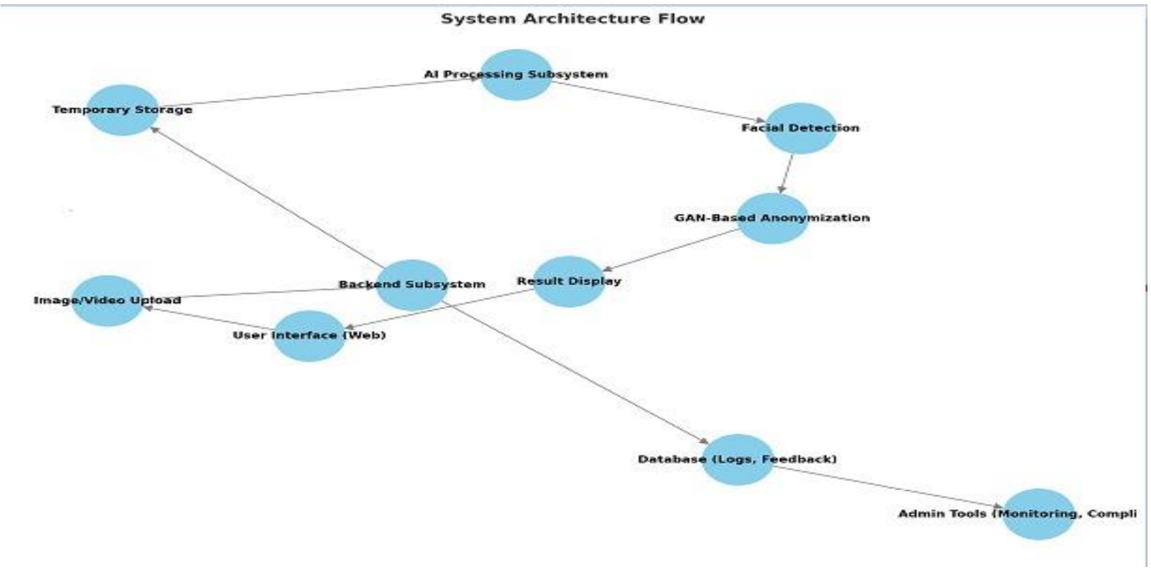
After Processing

View the anonymized outputs of your uploaded media.

5.3 Technical Comparison

- **DeepPrivacy GAN Performance:** Achieves 90.0% of original detection precision, outperforming traditional methods like pixelation or blurring.

- **Execution Efficiency:** Progressive GAN models reduce latency for high-resolution processing.
- **Realism:** Generates outputs with seamless integration into original backgrounds.



6. Glossary

1. GANs (Generative Adversarial Networks): Machine learning frameworks that consist of two neural networks—a generator and a discriminator—that compete in a zero-sum game to produce highly realistic data, such as images. GANs are widely used in image generation and modification tasks.

2. Scalability: The ability of a system to maintain performance or increase efficiency as workload or resource demand grows. Scalability is essential for large-scale applications to ensure consistent functionality under varying conditions.

3. Compliance: Conformance to legal and regulatory standards, such as the GDPR (General Data Protection Regulation) in Europe and CCPA (California Consumer Privacy Act) in the United States, which protect user data and privacy.

4. Pose Estimation: A process of predicting the spatial positions of key facial points (e.g., eyes, nose, and mouth) to determine the orientation of a face. Pose estimation enables systems to adapt anonymization techniques based on facial orientation.

5. Progressive GAN Training: A training method for GANs that gradually increases the resolution of generated images during the training process. This approach improves the stability and quality of generated outputs.

6. U-Net Architecture: A convolutional neural network architecture commonly used for tasks requiring detailed localization, such as image segmentation or background integration. It facilitates seamless integration of anonymized faces with their surrounding context.

7. GDPR (General Data Protection Regulation): A comprehensive legal framework established by the European Union to protect personal data and privacy for all individuals within the EU. It mandates strict guidelines for data handling and processing.

8. CCPA (California Consumer Privacy Act): A state-level data privacy law in the United States that gives California residents control over their personal information. It includes rights to know, delete, and opt-out of data sharing.

9. Role-Based Access Control (RBAC): A security mechanism that restricts system access based on the roles of individual users. This approach enforces the principle of least privilege and enhances system security.

10. Data Security: Practices and technologies designed to protect digital data from unauthorized access, corruption, or theft. Examples include encryption, multi-factor authentication, and secure storage solutions.

11. Conditional GANs: A variant of GANs where the generation process is conditioned on auxiliary information (e.g., pose or background). This enables targeted and context-aware image generation.

12. Anonymization: The process of removing or modifying personally identifiable information (PII) from data to protect individual privacy. In the context of images, this involves obscuring facial features while preserving overall context.

13. Temporary Storage: A short-term storage solution used to hold uploaded media files during processing. Ensures data is securely deleted after anonymization is complete, reducing storage costs and enhancing privacy.

14. FDF (Flickr Diverse Faces Dataset): A high-quality dataset used for training and evaluating face anonymization systems. It contains diverse facial images with variations in pose, occlusion, and background.

15. AES-256 Encryption: An advanced encryption standard that uses a 256-bit key length, ensuring a high level of data protection against unauthorized access.

7. References

1. Hukkelås, H., Mester, R., & Lindseth, F. (2019). DeepPrivacy: A Generative Adversarial Network for Face Anonymization. *arXiv preprint arXiv:1909.04538*. Retrieved from https://arxiv.org/abs/1909.04538
2. SIMA Specifications and Analysis Reports. (2024). *SIMA Development Team*. TED University, Faculty of Engineering.
3. CMPE 491 Syllabus. (2024). *High-Level Design Course*. TED University, Department of Computer Engineering.
4. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. *Advances in Neural Information Processing Systems*, 27, 2672–2680.
5. European Union. (2018). General Data Protection Regulation (GDPR). Retrieved from https://gdpr-info.eu
6. California Consumer Privacy Act (CCPA). (2018). State of California. Retrieved from https://oag.ca.gov/privacy/ccpa
7. Ren, Z., Lee, Y. J., & Ryoo, M. S. (2018). Learning to anonymize faces for privacy preserving action detection. *European Conference on Computer Vision (ECCV)*, 639-655. Springer International Publishing.
8. Karras, T., Aila, T., Laine, S., & Lehtinen, J. (2018). Progressive growing of GANs for improved quality, stability, and variation. *International Conference on Learning Representations (ICLR)*. Retrieved from https://openreview.net/forum?id=Hk99zCeAb