COMPUTACIÓN E INFORMÁTICA
Recibido 16 Sep. 2025 *ReCIBE, Año 14 No. 3, Dic. 2025*
Aceptado 03 Nov. 2025

Protección de Datos Biométricos en Vídeos Disponibles en Datasets Mediante Face Swapping

Biometric Data Protection in Videos Available in Datasets through Face Swapping

Héctor Caballero Hernández[1]
hcaballeroh240@profesor.uaemex.mx

Vianney Muñoz Jiménez[1]
vmunozj@uaemex.mx

Marco Antonio Ramos Corchado[1]
maramosc@uaemex.mx

[1] Universidad Autónoma del Estado de México

Resumen

Los datos biométricos como el rostro, la voz o las huellas dactilares son vulnerables a ataques con herramientas de inteligencia artificial (IA), ya que contienen características irremplazables en un individuo. Este trabajo presenta la aplicación de una técnica de anonimización facial basada en *face swapping* y la eliminación del fondo de las escenas en vídeos para proteger la privacidad de los participantes que aparecen en el dataset experimental LSM-VMX, el cual consta de 180 señas de la Lengua Mexicana de Señas (LSM). El proceso de anonimización ha sido desarrollado en Python para generar un cambio de rostro de los participantes empleando el modelo *inswapper_128*, mientras que la eliminación del fondo de las escenas se ha realizado mediante la librería de *rembg* de U2Net. Para probar la funcionalidad de los vídeos modificados, se entrenaron modelos basados en MediaPipe y una máquina de soporte vectorial (SVM) empleando los vídeos del dataset A (original) y del dataset B (modificado) para generar un modelo de reconocimiento de señas LSM, los resultados mostraron que la exactitud (*accuracy)* promedio para los datasets A y B fue de 0.975 y 0.983, respectivamente, lo cual demuestra que los cambios no repercutieron en el desempeño del modelo. Por otra parte, se ejecutaron pruebas de simetría facial para revisar que el proceso de anonimización fuera exitoso, empleando el modelo de VGG-face y la métrica SSIM, los resultados arrojaron que los rostros mostrados en el dataset B eran distintos a los del dataset A.

Palabras Clave: Datos biométricos, Lengua de Señas Mexicana, privacidad, datasets, anonimización, ética de la inteligencia artificial.

Abstract

Biometric data such as face, voice, and fingerprints are vulnerable to attacks using artificial intelligence (AI) tools, as they contain irreplaceable characteristics of an individual. This work presents the application of a facial anonymization technique based on face swapping and background removal from video scenes to protect the privacy of participants appearing in the LSM-VMX experimental dataset, which consists of 180 signs of the Mexican Sign Language (LSM). The anonymization process was developed in Python to generate a face swap of the participants using the inswapper_128 model, while the background removal of the scenes was performed using the U2Net rembg library. To test the functionality of the modified videos, models based on MediaPipe, and a support vector machine (SVM) were trained using the videos from dataset A (original) and dataset B (modified) to generate an LSM sign recognition model, the results showed that the average accuracy for datasets A and B was 0.975 and 0.983, respectively, which demonstrates that the changes did not impact the performance of the model. Furthermore, facial symmetry tests were run to check that the anonymization process was successful, using the VGG-face model and the SSIM metric, the results showed that the faces shown in dataset B were different from those in dataset A.

Keywords: Biometric data, Mexican Sign Language, privacy, datasets, anonymization, artificial intelligence ethics.

1. Introducción

En la actualidad, la masificación de dispositivos electrónicos con la capacidad de adquirir imágenes y vídeo ha permitido que grandes volúmenes de este tipo de datos estén disponibles en diversas plataformas en internet (redes sociales) (Nesterova, 2020). Debido a la falta de leyes que garanticen la privacidad de estos datos (Sánchez, 2020), todo el contenido relativo a exposición de datos biométricos en redes públicas es un blanco potencial para ser explotado por entidades maliciosas, cuyo objetivo principal es hacer mal uso de las características biométricas disponibles (Kumar et al., 2024), empleando algoritmos avanzados de IA es posible extraer los rasgos biométricos únicos o comportamientos conductuales para ejecutar acciones como suplantación de identidad (Lee et al., 2023).

Debido al desarrollo intensivo de algoritmos en el área de la visión por computadora y la IA para realizar tareas como clasificación y reconocimiento de objetos, para su uso en diversas áreas de la ciencia, industria y actividades cotidianas, han permitido extraer distintas características y parámetros sobre los objetos de estudio (Voulodimos et al., 2018). Estos algoritmos de IA generalmente emplean grandes volúmenes de datos ordenados, a los cuales se les conoce como datasets. Los datasets contienen objetos digitales agrupados en distintas clases semánticas, que pueden ir desde archivos de audio hasta imágenes y vídeo (Goodfellow et al., 2016). Con la aparición de conceptos como el modelo del metaverso y el desarrollo de modelos de IA generativa se ha incrementado el riesgo de un uso inadecuado de datos biométricos, los cuales no pueden ser modificados fiscalmente en las personas (Far & Rad, 2022).

Diversas aplicaciones de seguridad se han centrado en emplear distintos tipos de datos biométricos, (Chowdhury & Imtiaz, 2022), pero su recopilación conlleva riesgos significativos de privacidad que pueden derivar en robo de identidad y otras formas de daño si no se gestionan adecuadamente (Gichoya et al., 2023; Ciocca et al., 2023). Para proteger la información biométrica relacionada con rasgos faciales existen distintos métodos, entre los que se encuentran los de ofuscación directa, los cuales incluyen la pixelación, el desenfoque y el enmascaramiento (Puussaar et al., 2017), aunque estas modificaciones provocan de forma irreversible la perdida de información (Hukkelås et al., 2019; Blanton & Murphy, 2024, 2020; Hanisch et al., 2025).

Generalmente para la evaluación de los procedimientos de protección de datos biométricos se emplean las siguientes métricas.

- Métricas de anonimato. Verifican que el anonimato es completo, entre las que se encuentran los k-anonimato a otros cuasi-identificadores (ropa, fondo, etc.) (Sweeney, 2002).
- Riesgo de reidentificación. Busca verificar la identificación facial mediante ataques de reidentificación basados en otras señales contextuales (Tanuwidjaja et al., 2020).
- Utilidad de los datos. Se basa en evaluar de forma preliminar entrenando un modelo base de reconocimiento de señas sobre el dataset anonimizado.
- Consistencia y fidelidad visual. Evalúa cualitativamente la fluidez de las transiciones y el realismo de los rostros anonimizados, aplicando métricas cuantitativas como SSIM (Structural Similarity Index) (Wang et al., 2004) o FVD (Ciftci et al., 2023).

En este contexto, la importancia de los datasets permite generar diversas aplicaciones, entre las que encuentra la detección de la lengua de señas, debido a que en el mundo existen más de 475 millones de personas con deficiencias asociadas a la audición y al habla (FMS) (OMS, 2021; FMS, 2025).

En México hay más de 2.3 millones de personas con problemas de la audición y del habla, de este grupo de personas miles de ellas emplean la Lengua Mexicana de Señas para comunicarse, (INEGI, 2021), aunque existe una barrera digital significativa debido a la escasez de datasets de vídeo de alta calidad y a gran escala, lo cual dificulta el desarrollo de sistemas de interpretación automática (Rastgoo et al., 2021), a diferencia de otras lenguas de señas como la Americana (ASL) o la Británica (BSL).

Aunque en la actualidad se han desarrollado diversos esfuerzos para crear datasets para la LSM (Rodriguez et al., 2025; Martínez-Sanchez et al., 2023; Lara-Ortiz et al., 2025; Espejel et al. 2024: Trujillo-Romero et al., 2023), aún sigue siendo una tarea pendiente, el desarrollo de un dataset solido que permita crear modelos altamente eficientes en tareas de comunicación bidireccional, entre signantes y parlantes, así como la integración de una serie de procedimientos que garantice la protección de datos biométricos contenidos en él. Por tal motivo, esta investigación presenta una metodología enfocada a determinar la viabilidad del proceso de anonimización de los participantes que se encuentran en un dataset experimental denominado como LSM-VMX, basándose en la modificación de rostros para la protección de datos biométricos, así como la eliminación del fondo de las escenas para descontextualizar el escenario en donde se ejecutó la grabación, disminuyendo la probabilidad de una reidentificación, así como eliminar ruido para el algoritmo de detección de señas.

En la siguiente sección se presentan investigaciones dedicadas a la protección de datos biométricos empleando diversas técnicas de anonimización.

2. Estado del Arte

El proceso de anonimización es un campo de estudio ampliamente estudiado, en diversas investigaciones se puede observar el uso de técnicas orientadas al reemplazo de rostros, como en Zhu et al. (2021) propusieron MegaFS, el cual es un método para el intercambio de rostros en alta resolución a nivel de megapíxeles utilizando una representación jerárquica de rasgos faciales. Por su parte, Liu et al. (2023) presentaron DeepFaceLab, un marco de trabajo flexible e integrado que facilita a los usuarios la creación de deepfakes de alta calidad, buscando mejorar la fidelidad. En Groshev et al. (2022) desarrollaron GHOST, un enfoque que perfecciona arquitecturas existentes mediante una función de pérdida basada en los ojos y estabilización de video, mientras que Huang et al. (2024) se centraron en la preservación de la identidad del rostro fuente a través de modelos generativos de sustitutos duales. La investigación de Xu et al. (2022) presenta el diseño de un marco único tanto para la recreación facial como para el intercambio, basado en la separación de identidad y atributos. En Xu et al. (2022) se introdujo MobileFaceSwap, el cual es un sistema ligero para el intercambio de rostros en video en tiempo real en dispositivos móviles.

En Huang et al. (2023) abordaron la detección mediante un método que identifica inconsistencias entre la identidad explícita e implícita de un rostro, mientras que Ding et al. (2021) exploraron la anti-forensia, creando un modelo para generar videos que engañen a los detectores. En Perea- Trigo et al. (2025), utilizaron el reemplazo de rostros como técnica para el aumento de datos para mejorar significativamente la precisión de los modelos de reconocimiento de lengua de signos.

En las técnicas relativas a la privacidad diferencial se tienen distintos trabajos como en Dong, Roth, y Su (2022) propusieron la privacidad diferencial gaussiana (GDP por sus siglas en inglés), la cual facilita el análisis de la composición de algoritmos privados y la interpretación de sus garantías. En Yao et al. (2024) implementaron un marco con privacidad diferencial para proteger la autenticación de identidad mediante reconocimiento facial en sistemas de entrega con drones. Por otra parte, Bozkir et al. (2021) abordaron las vulnerabilidades en los datos de seguimiento ocular, proponiendo un mecanismo de privacidad diferencial que maneja las correlaciones temporales para proteger la información biométrica en dispositivos de realidad virtual (RV) y realidad aumentada (RA). En Rot et al. (2023) desarrollaron el sistema PrivacyProber, capaz de evaluar y revertir técnicas de mejora de la privacidad en datos biométricos faciales. En Hassanpour et al. (2022) desarrollaron una metodología para aplicar la privacidad diferencial al aprendizaje continuo, logrando un equilibrio entre la privacidad y la utilidad del modelo.

En Sharma, Das, y Chaudhury (2025) desarrollaron un sistema descentralizado basado en realidad extendida (XR, por sus siglas en inglés) y un criptosistema biométrico híbrido para la visualización segura y colaborativa de datos médicos en 3D. En Jeremiah et al. (2024) propusieron PrivacyGuard, la cual es una arquitectura colaborativa para la anonimización facial en grabaciones de CCTV. Por su parte, Jamil y Jamil (2024) presentaron un enfoque dual denominado IncepX- Ensemble, se encarga de realizar el reconocimiento étnico y la anonimización facial. El trabajo de Mishra, Pagare, y Sharma (2025), desarrollaron un método híbrido que combina PNL (procesamiento del lenguaje natural) basado en reglas y aprendizaje automático para detectar y anonimizar con precisión información personal identificable en documentos financieros.
Finalmente, en Xu et al. (2024) diseñaron un esquema de anonimización reversible para imágenes faciales utilizando aprendizaje cíclico, lo que permite desidentificar los datos para proteger la privacidad y reidentificarlos de forma segura cuando sea necesario.

Los trabajos presentados en el área de la anonimización se han desarrollado diversos métodos que van desde el reemplazo de rostros, pasando por la privacidad diferencial, hasta la combinación de plataformas hibridas basadas en la nube y métodos que garantizan la reversibilidad de este proceso de forma segura. En la siguiente sección se presenta una metodología para la anonimización de un dataset experimental de la LSM, con la finalidad de analizar la viabilidad de la protección de datos biométricos sin perder las características esenciales de los movimientos que permita generar modelos de IA para el reconocimiento del lenguaje de señas LSM.

3. Metodología

En el presente capítulo se detalla el proceso metodológico de esta investigación, el cual se centra en la construcción de un dataset experimental de la LSM denominado como LSM-VMX. Este dataset servirá para corroborar los efectos de la implementación de *face swapping*, y comparar el rendimiento en el proceso de evaluación de cada modelo generado (videos originales versus modificados), empleando las métricas de desempeño de exactitud, puntuación F1, entre otras. En la Figura 1 se muestran las etapas de la metodología empelada para llevar a cabo la modificación de rostros en datasets.

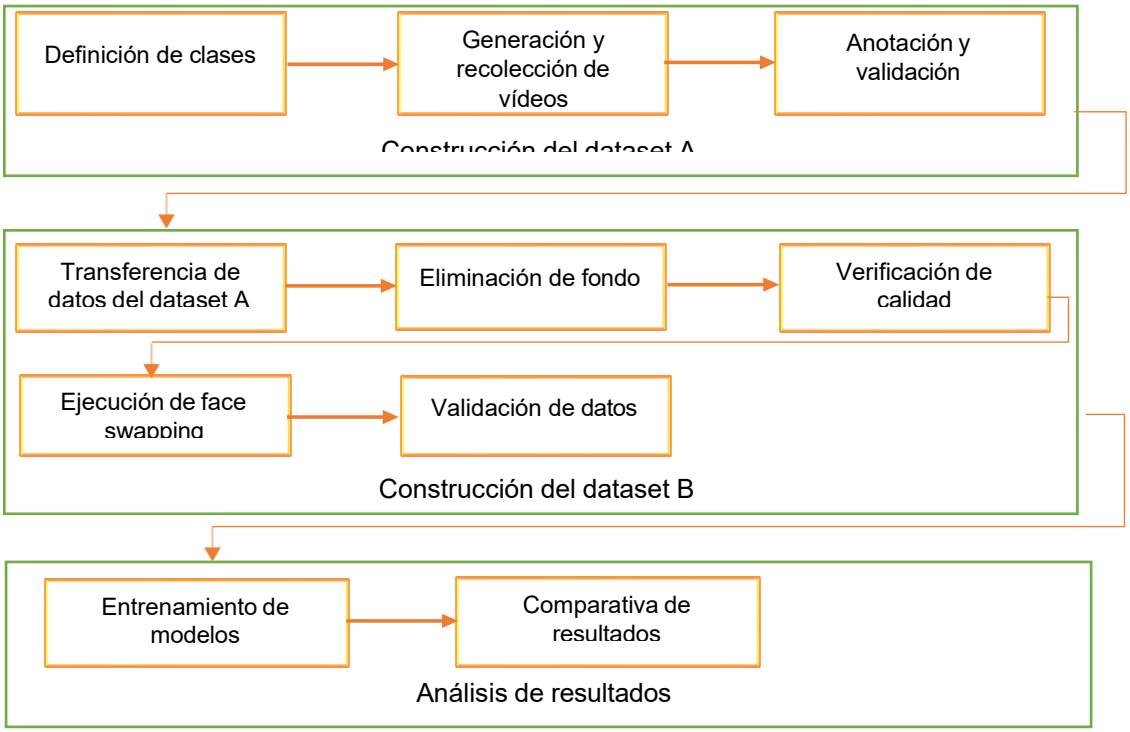


Figura 1. Flujo metodológico para el proceso de anonimización de vídeos

El flujo metodológico expresado en la Figura 1 cuenta con 3 etapas importantes, las cuales son: la construcción de dataset A y B, y el análisis comparativo.

- Definición del dataset A:
    1. Definición de clases. Se seleccionaron 180 señas de uso frecuente en la LSM, agrupadas en 9 categorías semánticas: animales, colores, preguntas comunes, saludos y despedidas, prendas de vestir, objetos de casa y oficina, dígitos y letras.
    2. Generación y recolección de vídeos. Se reclutaron 10 personas para realizar las señas. Además de las grabaciones controladas, se incluyeron vídeos extraídos de internet con licencia para reproducción para aumentar la diversidad de signantes y entornos.
    3. Anotación y validación. Cada vídeo fue etiquetado con la glosa en español, la categoría semántica y si la persona que grabó el vídeo pertenece la sexo masculino o femenino. El dataset final contiene 3600 vídeos.

- Construcción del dataset B:

    4. Transferencia de datos del dataset A. Terminando el proceso de construcción de dataset A se procede a transferir una copia de este, sin modificar los nombres de los vídeos contenidos en las clases.

    5. Eliminación de fondo. Los vídeos MP4 del dataset B son modificados para eliminar el fondo, limitando la posibilidad del reconocimiento del lugar en donde se realizó la escena, así como eliminar ruido que afecte el algoritmo de reconocimiento de señas.

    6. Verificación de calidad. De manera manual y visual se verifica que los vídeos a los que se les ha eliminado el fondo no presenten aberraciones visuales.

    7. Ejecución de *face swapping*. Una red neuronal convolucional (CNN) se encarga de detectar de la segmentación del rostro, para posteriormente emplear una imagen de reemplazo (dependiendo del sexo del participante), en este caso, el rostro de reemplazo debe de ser una figura pública o generada por IA.

    8. Validación de datos. Los vídeos son revisados manualmente para comprobar que no presente anomalías visuales que impidan su reproducción o aparición de ruido por la eliminación del fondo.

- Comparativa de resultados:

    9. Entrenamiento. Los modelos son entrenados por el algoritmo MediaPipe (Lugaresi et al., 2019) y SVM (support vector machine). Al terminar el proceso se obtienen los resultados de las métricas de exactitud, precisión y recuperación.

    10. Análisis de resultados. Se verifica el rendimiento de los modelos para determinar si los cambios de eliminación de fondo y cambio de rostro no han provocado modificaciones sensibles en la interpretación de las señas, así como el cálculo de la posibilidad de la identificación de las personas que han pasado por un proceso de anonimización digital.

El Algoritmo 1 presenta el proceso de *face swapping* y eliminación de escenas de fondo. De forma general, se instalan las dependencias insightface, onnxruntime, *rembg* y opencv-python, posteriormente se emplea un modelo inswapper_128.onnx para el intercambio de rostros, se cargan las direcciones de los vídeos y los rostros modelos, los vídeos entran en un ciclo para su procesamiento y finalmente se crean nuevas carpetas para generar el dataset B.

Algoritmo 1. Modificación de vídeos para protección de datos
1. Inicio 2. Instalar las dependencias visión por computadora. 3. Descargar el modelo *face_swapper* de modificación de rostros de rostros. 4. Definir las rutas del *rostro_modelo* (la nueva cara, de dominio público y relacionada con el sexo del participante) y el vídeo de destino (del dataset LSM-VMX). 5. Cargar en memoria la librería rembg para la eliminación de fondo. 6. Ciclo para cada cuadro del vídeo: ▪ Identificación del objeto. ▪ Segmentación del fondo de la escena. ▪ Generación del canal alfa. ▪ Almacenamiento de *vídeo_intermedio*. 7. Cargar en memoria los modelos de IA para el análisis y el intercambio facial.
8. Procesar la imagen de origen para extraer el *rostro_modelo* que se utiliza para el reemplazo. 9. Leer el vídeo de destino cuadro por cuadro y configurar un archivo de vídeo para guardar el resultado. 10. Ciclo para cada cuadro de *vídeo_intermedio*: ▪ Se detectan todos los rostros en el cuadro actual. ▪ Si se detectan rostros, el modelo *face_swapper* reemplaza cada rostro detectado con el *rostro_modelo*. ▪ Si no se detectan rostros, el cuadro se mantiene sin modificar. ▪ Eliminar fondo de las escenas. ▪ El cuadro (original o modificado) se escribe en el archivo de vídeo de salida. 11. Procesados todos los cuadros, se libera la memoria y se guarda el vídeo final en dataset B. 12. Fin

El Algoritmo 2, presenta los pasos a desarrollar para la evaluación del modelo original y del modelo que recibió los vídeo modificados. Este algoritmo se divide en 3 etapas, en la etapa 1 para cada dataset (original y modificado) emplea MediaPipe para analizar cada fotograma de cada vídeo y detectar los 21 puntos clave de la mano. Posteriormente, se normalizan estos puntos tomando la muñeca como punto de referencia (el punto cero). Al final, cada vídeo se convierte en un vector que representa matemáticamente la seña.

Algoritmo 2. Entrenamiento de modelos de datasets
1. Inicio 2. Inicializar una lista vacía para los puntos clave del vídeo. 3. Para cada fotograma del vídeo (hasta un límite de 30): • Detectar los 21 puntos clave de la mano usando MediaPipe. • Si se detecta una mano, normalizar las coordenadas de todos los puntos restándoles las coordenadas de la muñeca. 4. Añadir los puntos normalizados a la lista del vídeo. 5. Aplanar todos los puntos clave del vídeo en un único vector numérico. 6. Almacenar todos los vectores numéricos (x) y sus etiquetas correspondientes (y). 7. Dividir el conjunto de datos (x, y) en un 80% para entrenamiento y un 20% para pruebas. 8. Definir una serie de hiperparámetros a probar para el SVM (valores de C y gamma). 9. Utilizar grid search con validación cruzada sobre los datos de entrenamiento para encontrar la combinación de hiperparámetros que genere el modelo más preciso. 10. Usar el modelo mejor evaluado para hacer predicciones sobre los datos de prueba. 11. Calcular y guardar las métricas de rendimiento: reporte de clasificación (precisión, recall, y puntuación F1). 12. Fin

Al finalizar el proceso de comparación entre ambos modelos se realiza un ciclo encargado de analizar el proceso de comparativo entre los vídeos del dataset A y el dataset B mediante el empleo de SSIM, debido a que es una métrica ampliamente eficaz para para determinar las diferencias entre imágenes.

En la siguiente sección se presentan los resultados al emplear la metodología mencionada en esta sección, con la finalidad de comprobar la vialidad de agregar protección a los datos biométricos (rostro) contenidos en vídeos para datasets públicos.

4. Resultados y Evaluación

En este apartado se describe el proceso empleado para el desarrollo de las pruebas a la metodología del proceso de anonimización de lo vídeos contenidos en el dataset LSM-VMX. En la Tabla 1 se presentan las condiciones experimentales para el desarrollo de las pruebas.

Tabla 1. Elementos considerados para el proceso de experimentación

Elementos considerados	Descripción
Software	Python 12 Google Colab Modelo pre-entrenado inswapper 128.onnx Librería *rembg*
Hardware	Computadora de escritorio iMac M4 con 16 GB de RAM y 256 GB de almacenamiento
Dataset	Conformado por 9 categorías semánticas: animales, colores, preguntas comunes, saludos y despedidas, prendas de vestir, objetos de casa y oficina, dígitos y letras
Métricas de desempeño	- Exactitud, precisión, puntuación F1 y mAP@0.5 para los modelos generados para reconocimiento de señas. - SSIM y VGG face para comprobación de diferencias estructurales con el cambio de rostros.

El proceso de anonimización se ha ejecutado en 4 etapas, la primera etapa ha consistido en la eliminación del fondo de los vídeos del d~~ataset A, esta etapa~~ se ha ejecutado en el computador iMac debido a su efectividad y velocidad para la eliminación de fondos, la segunda etapa se ejecutó en una máquina virtual de Google Colab para aprovechar la potencia de cálculo de la GPU disponible para el proceso de *face swapping* y solventar problemas de compatibilidad con algunas librería del proceso de *face swapping*. La tercera etapa ha consistido en realizar el entrenamiento de los modelos, empleando la librería de MediaPipe en combinación de una SVM, empleando los datos de los datasets A y B. Finalmente, la cuarta etapa consistió en la evaluación de las métricas de desempeño de los modelos, empleando exactitud, precisión, puntuación F1 (Jacob et al., 2021) y mAP@0.5 (Everingham et al., 2010) para que los resultados sean comparados entre ambos modelos, y verificar las diferencias entre ambos. Las ecuaciones para calcular la exactitud (1), precisión (2), puntuación F1 (3), SSIM (4), SSIM de fotograma medio (5) y SSIM facial medio (6) se presentan a continuación.

Donde:

$$\text{Exactitud} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}} \quad (1)$$

$$\text{Precisión} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (2)$$

$$\text{F1} = 2 \text{ x } \frac{\text{Precisión x Exhaustividad}}{\text{Precisión} + \text{Exhaustividad}} \quad (3)$$

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (4)$$

$$\text{SSIM de fotograma medio} = \frac{1}{V_{total}} \sum_{j=1}^{V_{total}} \left(\frac{1}{M_j} \sum_{i=1}^{M_j} S_{fotograma,ij} \right) \quad (5)$$

$$\text{SSIM facial medio} = \frac{1}{V_{total}} \sum_{j=1}^{V_{total}} \left(\frac{1}{O_j} \sum_{i=1}^{O_j} S_{rostro,ij} \right) \quad (6)$$

Verdadero positivo (TP). El modelo predijo correctamente la clase positiva.

Falso positivo (FP). El modelo predijo incorrectamente la clase positiva (error de tipo I).

Falso negativo (FN). El modelo predijo incorrectamente la clase negativa (error de tipo II).

Verdadero negativo (TN). El modelo predijo correctamente la clase negativa.

μ_x es la media de todos los valores de píxeles en la imagen x.

μ_y es el promedio (media) de todos los valores de píxeles en la imagen y.

σ_x^2 es la varianza de la imagen x.

σ_y^2 es la varianza de la imagen y.

σ_x y σ_y son las desviaciones estándar.

σ_{xy} es la covarianza entre las imágenes x y y.

C_1 y C_2 son constantes contantes para evitar la división entre 0.

M son pares de fotogramas

S es el valor de SSIM de los fotogramas pares

V_{total} total de vídeos.

M_j fotogramas analizados

O_j es el número de pares de rostros analizados exitosamente en el vídeo j

En la Figura 2 a) y Figura 2 b) se presentan los rostros empleados para ejecutar el *face swapping*, tomando en cuenta si eran hombre o mujeres los sujetos aparecen en los vídeos originales, debido a que los rasgos morfológicos entre los participantes de diferentes sexos cambian, en este caso se han elegido la figura pública como Salma Hayek y Fernando Colunga.

Rostro de reemplazo para mujeres

a) Rostro de Salma Hayek

Rostro de reemplazo para hombres

b) Rostro de Fernando Colunga

Figura 2. Rostros empleados para realizar *face swapping* en los vídeo disponibles en el dataset A.

La Tabla 2 muestra los resultados obtenidos al evaluar los modelos generados empleando los datasets A y B, donde es posible ver que las diferencias entre los resultados obtenidos no son significativas, debido a que los cambios realizados por la eliminación del fondo y del rostro no han afectado los movimientos que se observan en los vídeos, esto se puede observar en los puntajes obtenidos en las métricas de exactitud, puntuación F1 y precisión, donde las diferencias entre los modelos se encuentran en el orden de céntimas de unidad. A pesar de altos puntajes obtenidos en las pruebas, se puede inferir que existe un sobreentrenamiento en ambos modelos, este fenómeno será abordado en futuras iteraciones mediante el incremento de una mayor cantidad de vídeos para cada clase, ya que el objetivo principal de este estudio es evaluar las diferencias resultantes de la aplicación del proceso de anonimización mediante la modificación de rostros y la eliminación del fondo de las escenas.

Tabla 2. Comparación de resultados al evaluar los modelos obtenidos de los dataset A y B de detección de señas basándose en el dataset experimental LSM-VMX.

Tipo de modelo	Exactitud	Puntaje F1	Precisión
Basado en dataset original	0.975	0.971	0.976
Basado en dataset modificado por eliminación de fondo y *swapping*	0.973	0.978	0.983

En la Figura 3 a) se presenta el vídeo de la seña alacrán sin modificaciones, mientras que en la Figura 3 b) se muestra el vídeo con el proceso de anonimización con respecto al fondo y la modificación del rostro del participante, por otra parte, en las Figuras 3 c) y Figura 3 d) se presentan dos extractos de la seña amarillo, en estas es posible identificar que el proceso de anonimización ha modificado la calidad de vídeo, aunque es importante señalar que el reemplazo del rostro se realizó exitosamente, pero no se presentan algunos detalles visuales que contrastan con respecto a los de la Figura 3 b), donde el proceso de reemplazo y eliminación de fondo se han logrado sin dejar rastro de la eliminación del fondo principal.
Finalmente en las Figura 3 e) y Figura 3 f) y Figura 3 g) y Figura 3 h), el participante seleccionado es un hombre que realiza las señas avestruz y agua, respectivamente, donde es posible observar que en los vídeo de los incisos f y h de la Figura 3, el reemplazo de rostro se logró de manera exitosa, aunque en el inciso h, es posible observar que existe en el fondo un elemento que no se eliminó del todo, aunque el movimiento ejecutado respecto a la seña no presenta ningún cambio en relación al vídeo original, demostrando el éxito de la técnica de face swapping aplicada.

Video original

Video con *face swapping*

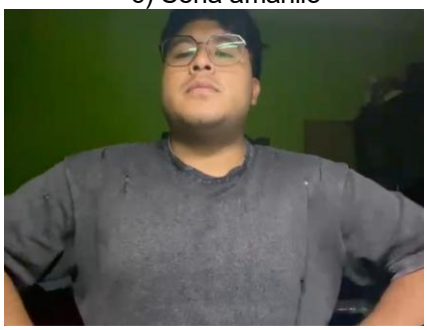
a) Seña alacrán

b) Seña alacrán, con rostro modificad

Video original

Video con *face swapping*

c) Seña amarillo



d) Seña amarillo

e) Avestruz

f) Seña avestruz

g) Seña agua

h) Seña agua

Figura 3. Resultados comparativos entre los resultados del dataset A versus dataset B

La Tabla 2 muestra los resultados respectivos al análisis efectuado al momento de calcular la distancia promedio facial mediante el modelo VGG-Face. El método propuesto para el reemplazo de rostro mediante *face swapping* ha cumplido con el objetivo de anonimización, debido a que el valor obtenido en la métrica de distancia facial ha permitido determinar que la morfología de rostro es completamente diferente, tomando en cuenta las pruebas ejecutadas por el modelo VGG-Face al considerar que la distancia entre rostros es mayor a 0.40. Al analizar el resultado de la evaluación de la métrica SSIM fácil medio es posible verificar que ha cambiado completamente sin dejar rastros estructurales con respecto a los originales con un puntaje menor a 0.290 puntos. El SSIM de fotograma medio demuestra que el proceso de eliminación de fondo y cambio de rostro ha cambiado completamente la escena original, teniendo en cuenta que el valor obtenido es inferior a 0.320 y considerando que lo valores aceptables del SSIM para indicar que una escena es similar a la otra varía entre 0.9 y 1.0, siendo 1.0 el puntaje que indica que los elementos comparados son idénticos. Por otra parte, es posible determinar que técnicamente la identificación del espacio en donde se grabó el vídeo no es identificable, dando como resultado un incremento en la privacidad de la escena.

Tomando en cuenta las métricas del rendimiento de modelo de reconocimiento de señas con las métricas de los modelos de IA es posible afirmar que no existen cambios significativos en el desempeñó entre el modelo generado empelando el dataset A o el modelo generado empleando el dataset B. Además, el proceso de eliminación de fondo ha creado diferencias visuales importantes entre los vídeos, además de contribuir ligeramente a obtener desempeño favorables en el puntaje F1 y en la precisión, sumado a la incorporación de la modificación del rostro.

Tabla 2. Comparación de valores de métricas de rendimiento en el proceso de anonimización

Distancia facial promedio calculada con VGG-Face	SSIM facial medio	SSIM de fotograma promedio
0.7312	0.2840	0.3149

Si bien, el proceso de anonimización ha sido exitoso, el principal desafío que se ha encontrado al analizar los resultados obtenidos es el equilibrio entre una privacidad robusta y la máxima utilidad de los datos, además de que en fases futuras será necesario implementar etapas de matizado y texturización sobre las zonas donde empieza el proceso de cambio de rostro, así como la evaluación adaptación de enfoques de fuentes de iluminación para crear una sensación de que el vídeo no presenta modificaciones, tomando en cuenta procesos como los seguidos en las investigaciones de la Zhu et al. (2021) y Liu et al. (2023). Por otra parte, el enfoque de la presente investigación es novedoso desde la perspectiva de resguardar la privacidad de los datos biométricos en vídeos referentes a lenguas de signos, en este caso para la LSM, aunque existen propuestas que han basado su metodología para el entrenamiento de modelos de lenguas de signos como se muestra en Perea-Trigo, et al. (2025) con el objetivo de incrementar los vídeos disponibles y su viabilidad en el proceso de entrenamiento. El proceso de *face swapping* demostró ser eficaz, pero se tiene que tomar en cuenta que la robustez de cualquier técnica de anonimización frente a ataques de reidentificación en constante evolución es una posibilidad que debe de ser tomada en cuenta (Wenger et al., 2022).

Como se ha observado en el análisis de los resultados numéricos y visuales del proceso de anonimización del dataset LSM-VMX, se ha cumplido con el objetivo de anonimizar a los participantes y lograr que los datos generados preserven sus características esenciales para lograr modelos de IA útiles en el reconocimiento de señas.

El proceso de creación y distribución de datasets con datos biométricos implícitos o explícitos plantea desafíos complejos desde la perspectiva ética y legal, por tal motivo este trabajo ha abordado el principio de la privacidad como parte fundamental del ciclo de vida del dataset (Cavoukian, 2011). Si bien es cierto que en México las leyes sobre protección de los datos biométricos no están ampliamente desarrolladas, como se estipula en regulaciones como el GDPR en Europa, la CCPA en California y la emergente Ley de IA de la UE (Voigt & von dem Bussche, 2017; European Parliament, 2024), este trabajo aboga por un manejo proactivo de la privacidad de los datos públicos que se encuentran disponibles en medios públicos, los cuales se encuentra disponibles para todo tipo de público, bajo este contexto, las personas que emplean estos datos no declaran la finalidad del uso de este tipo de información generando incertidumbre sobre el uso que se le dará, pero mediante el anonimato de los participantes en este tipo de datasets se reduce la posibilidad de un mal uso de este tipo de datos. Los resultados obtenidos en esta investigación representan un punto de partida para el desarrollo de nuevas propuestas para garantizar la integridad de los datos biométricos que se encuentran disponibles de forma pública en internet relativos a la LSM, así como dar pauta nuevos desarrollos que permitan generar nuevos datasets en donde la información biométrica no se vea comprometida.

6. Conclusiones

Este trabajo se ha centrado en demostrar la viabilidad de la implementación de un proceso de anonimización para el reemplazado de rostros mediante face swapping en vídeos provenientes de un dataset experimental de señas LSM. Los resultados han indicado que al momento de entrenar los modelos de reconocimiento de señas LSM demostraron que técnicamente no existe diferencia significativa entre los resultados obtenidos al evaluar los modelos con las métricas de exactitud, precisión y score F1, debido a que las variaciones están en orden de céntimas de unidad, así como valores de SSIM facial medio para SSIM por fotograma promedio son inferiores a 0.31 puntos, indicando diferencias drásticas entre los vídeos originales y los vídeos modificados en el proceso de anonimización.

De acuerdo con lo observado durante la ejecución de las pruebas, es posible verificar que los rasgos morfológicos de los rostros empleados se adecuan limpiamente a los rostros que no presentan alguna modificación, siendo factible el empleo de face swapping para proteger la identidad de las personas con respecto a la exposición de sus rasgos biométricos más sensibles a usurpación, tomando en cuenta que la eliminación del fondo de las escenas beneficio ligeramente los resultados de las métricas de puntaje F1 y precisión, además de que los datos obtenidos son 100% útiles para generar modelos de reconocimiento de señas LSM.

Trabajo Futuro

Como trabajo a futuro se proponen dos vertientes a ejecutar en paralelo, en primera instancia, se extenderá el proceso de anonimización para cambiar el color de la vestimenta, el fondo para crear otro tipo de escenario, así como la elección de un rostro de reemplazo que se ajuste al del participante en vídeo de manera automática de acuerdo a su morfología específica, en segunda instancia, se desarrollará un sistema con la capacidad de generar avatares que permitan replicar las señas LSM contemplando gestos faciales y manuales, con la finalidad conformar un dataset para el entrenamiento de modelos de IA con altos niveles de exactitud.

Referencias

Blanton, M., & Murphy, D. (2024, June). Privacy preserving biometric authentication for fingerprints and beyond. In *Proceedings of the Fourteenth ACM Conference on Data and Application Security and Privacy* (pp. 367-378). https://doi.org/10.1145/3626232.3653269.

Bozkir, E., Günlü, O., Fuhl, W., Schaefer, R. F., & Kasneci, E. (2021). Differential privacy for eye tracking with temporal correlations. *PLoS ONE, 16*(8), e0255979. https://doi.org/10.1371/journal.pone.0255979.

Cavoukian, A. (2011). Privacy by Design: The 7 Foundational Principles. Information and Privacy Commissioner of Ontario.

Ciftci, U. A., Yuksek, G., & Demir, I. (2023). My face my choice: Privacy enhancing deepfakes for social media anonymization. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision* (pp. 1369-1379).

Ciocca, G., Napoletano, P., & Schettini, R. (2023). A review on deep learning based biometric recognition. Pattern Recognition Letters, 173, 43-52. https://doi.org/10.1016/j.patrec.2023.08.012.

Chowdhury, A. M., & Imtiaz, M. H. (2022). Contactless fingerprint recognition using deep learning—a systematic review. *Journal of Cybersecurity and Privacy*, *2*(3), 714-730. https://doi.org/10.3390/jcp2030036.

Ding, F., Zhu, G., Li, Y., Zhang, X., Atrey, P. K., & Lyu, S. (2021). Anti-forensics for face swapping

videos via adversarial training. *IEEE Transactions on Multimedia, 24*, 3429-3441. 10.1109/TMM.2021.3098422.

Dong, J., Roth, A., & Su, W. J. (2022). Gaussian differential privacy. *Journal of the Royal Statistical Society Series B: Statistical Methodology, 84*(1), 3-37. https://doi.org/10.1111/rssb.12454.

Espejel, J., Jalili, L. D., Cervantes, J., & Canales, J. C. (2024). Sign language images dataset from Mexican sign language. *Data in Brief*, *55*, 110566. 10.1016/j.dib.2024.110566.

European Parliament. (2024, March 13). EU AI Act: First regulation on artificial intelligence. https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence

Everingham, M., Van Gool, L., Williams, C. K. I., Winn, J., & Zisserman, A. (2010). The PASCAL Visual Object Classes (VOC) Challenge. International Journal of Computer Vision, 88(2), 303–338. https://doi.org/10.1007/s11263-009-0275-4.

Far, S. B., & Rad, A. I. (2022). Applying digital twins in metaverse: User interface, security and privacy challenges. *Journal of Metaverse*, *2*(1), 8-15.

Federación Mundial de Sordos. (s.f.). Building a World Where Everywhere Deaf People Can Sign Anywhere! Recuperado el 20 de agosto de 2025, de https://wfdeaf.org.

Gichoya, J. W., Thomas, K., Celi, L. A., Safdar, N., Banerjee, I., Banja, J. D., ... & Purkayastha, S. (2023). AI pitfalls and what not to do: mitigating bias in AI. *The British Journal of Radiology*, *96*(1150), 20230023. 10.1259/bjr.20230023.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

Groshev, A., Maltseva, A., Chesakov, D., Kuznetsov, A., & Dimitrov, D. (2022). GHOST—a new face swap approach for image and video domains. *IEEE Access, 10*, 83452-83462. 10.1109/ACCESS.2022.3196668.

Hanisch, S., Arias-Cabarcos, P., Parra-Arnau, J., & Strufe, T. (2025). Anonymization techniques for behavioral biometric data: a survey. *ACM Computing Surveys*, *57*(11), 1-54. https://doi.org/10.1145/3729418.

Hassanpour, A., Moradikia, M., Yang, B., Abdelhadi, A., Busch, C., & Fierrez, J. (2022). Differential privacy preservation in robust continual learning. *IEEE Access, 10*, 24273-24287. 10.1109/ACCESS.2022.3154826.

Huang, B., Wang, Z., Yang, J., Ai, J., Zou, Q., Wang, Q., & Ye, D. (2023). Implicit identity driven deepfake face swapping detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 4490-4499).

Huang, Z., Tang, F., Zhang, Y., Cao, J., Li, C., Tang, S., ... & Lee, T. Y. (2024). Identity-preserving face swapping via dual surrogate generative models. *ACM Transactions on Graphics, 43*(5), 1-19. https://doi.org/10.1145/3676165.

Hukkelås, H., Mester, R., Lindseth, F. (2019). DeepPrivacy: A Generative Adversarial Network for Face Anonymization. In: Bebis, G., et al. Advances in Visual Computing. ISVC 2019. Lecture Notes in Computer Science, vol 11844. Springer, Cham. https://doi.org/10.1007/978-3-030-33720-9_44.

INEGI. (2021). Censo de Población y Vivienda 2020: Discapacidad. Instituto Nacional de Estadística y Geografía. https://www.inegi.org.mx/temas/discapacidad/.

Jamil, F., & Jamil, H. (2024, August). Toward Intelligent Ethnicity Recognition and Face Anonymization: An IncepX-Ensemble Model. In *International Conference on Computational Collective Intelligence* (pp. 243-255). Cham: Springer Nature Switzerland.

https://doi.org/10.1007/978-3-031-70819-0_19.

Jeremiah, S. R., Ha, J., Singh, S. K., & Park, J. H. (2024). Privacy guard: collaborative edge-cloud computing architecture for attribute-preserving face anonymization in CCTV networks. *Human-centric Computing and Information Sciences, 14*(43), 1e16. 10.22967/HCIS.2024.14.043.

Lee, P. Y. K., Ma, N. F., Kim, I. J., & Yoon, D. (2023). Speculating on risks of AI clones to selfhood and relationships: Doppelganger-phobia, identity fragmentation, and living memories. *Proceedings of the ACM on Human-computer Interaction*, 7(CSCW1), 1-28. https://doi.org/10.1145/3579524.

Kumar, T., Bhushan, S., Sharma, P., & Garg, V. (2024). Examining the vulnerabilities of biometric systems: Privacy and security perspectives. In *Leveraging Computer Vision to Biometric Applications* (pp. 34-67). Chapman and Hall/CRC.

Lara-Ortiz, V., Fuentes-Aguilar, R. Q., & Chairez, I. (2025). Spanish to Mexican Sign Language glosses corpus for natural language processing tasks. Scientific Data, 12(1), 702. https://doi.org/10.1038/s41597-025-04871-7.

Liu, K., Perov, I., Gao, D., Chervoniy, N., Zhou, W., & Zhang, W. (2023). Deepfacelab: Integrated, flexible and extensible face-swapping framework. *Pattern Recognition, 141*, 109628. https://doi.org/10.1016/j.patcog.2023.109628.

Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., Zhang, F., Chang, C.-L., Yong, M. G., Lee, J., Chang, W.-T., Hua, W., Georg, M., & Grundmann, M. (2019). MediaPipe: A Framework for Building Perception Pipelines. https://doi.org/10.48550/arXiv.1906.08172.

Martínez-Sánchez, V., Villalón-Turrubiates, I., Cervantes-Álvarez, F., & Hernández-Mejía, C. (2023). Exploring a novel mexican sign language lexicon video dataset. Multimodal Technologies and Interaction, 7(8), 83. https://doi.org/10.3390/mti7080083.

Mishra, K., Pagare, H., & Sharma, K. (2025). A hybrid rule-based NLP and machine learning approach for PII detection and anonymization in financial documents. *Scientific Reports, 15*(1), 22729. 10.1038/s41598-025-04971-9.

Nesterova, I. (2020). Mass data gathering and surveillance: the fight against facial recognition technology in the globalized world. In *SHS web of conferences* (Vol. 74, p. 03006). EDP Sciences. https://doi.org/10.1051/shsconf/20207403006.

Organización Mundial de la Salud. (2021). Informe mundial sobre la audición. https://www.who.int/publications/i/item/9789240020481.

Perea-Trigo, M., López-Ortiz, E. J., Soria-Morillo, L. M., Álvarez-García, J. A., & Vegas-Olmos, J. J. (2025). Impact of face swapping and data augmentation on sign language recognition. *Universal Access in the Information Society, 24*(2), 1283-1294. https://doi.org/10.1007/s10209-024-01133-y.

Puussaar, A., Clear, A. K., & Wright, P. (2017, May). Enhancing personal informatics through social sensemaking. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 6936-6942). https://doi.org/10.1145/3025453.302580.

Rastgoo, R., Kiani, K., & Escalera, S. (2021). Sign language recognition: A deep survey. Expert Systems with Applications, 164, 113794. https://doi.org/10.1016/j.eswa.2020.113794.

Rodriguez, M., Oubram, O., Bassam, A., Lakouari, N., & Tariq, R. (2025). Mexican Sign Language Recognition: Dataset Creation and Performance Evaluation Using MediaPipe and Machine Learning Techniques. Electronics, 14(7), 1423. https://doi.org/10.3390/electronics14071423.

Rot, P., Grm, K., Peer, P., & Štruc, V. (2023). PrivacyProber: Assessment and detection of soft–biometric privacy–enhancing techniques. *IEEE Transactions on Dependable and Secure Computing, 21*(4), 2869-2887. https://doi.org/10.1109/TDSC.2023.3319500.

Sánchez, M. (2020). *Protección de datos personales biométricos*. Instituto Nacional de Transparencia, Acceso a la Información y Protección de Datos Personales (INAI).

Sharma, S., Das, D., & Chaudhury, S. (2025). A decentralized privacy-preserving XR system for 3D medical data visualization using hybrid biometric cryptosystem. *Scientific Reports, 15*(1), 28568. https://doi.org/10.1038/s41598-025-08784-8.

Sweeney, L. (2002). k-anonymity: A model for protecting privacy. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 10(05), 557-570. https://doi.org/10.1142/S0218488502001648.

Tanuwidjaja, H. C., Choi, R., Baek, S., & Kim, K. (2020). Privacy-preserving deep learning on machine learning as a service—a comprehensive survey. *IEEE Access*, *8*, 167425-167447. 10.1109/ACCESS.2020.3023084.

Trujillo-Romero, F., & García-Bautista, G. (2023). Mexican sign language corpus: Towards an automatic translator. *ACM Transactions on Asian and Low-Resource Language Information Processing*, *22*(8), 1-24. https://doi.org/10.1145/3591471.

Voigt, P., & von dem Bussche, A. (2017). The EU General Data Protection Regulation (GDPR). Springer. https://doi.org/10.1007/978-3-319-57959-7.

Voulodimos, A., Doulamis, N., Doulamis, A., & Protopapadakis, E. (2018). Deep learning for computer vision: A brief review. *Computational Intelligence and Neuroscience, 2018*, 7068349. https://doi.org/10.1155/2018/7068349.

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing, 13(4), 600–612. https://doi.org/10.1109/TIP.2003.819861.

Xu, C., Zhang, J., Han, Y., Tian, G., Zeng, X., Tai, Y., ... & Liu, Y. (2022, October). Designing one unified framework for high-fidelity face reenactment and swapping. In *European conference on computer vision* (pp. 54-71). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-19784-0_4.

Xu, S., Chang, C. C., Nguyen, H. H., & Echizen, I. (2024). Reversible anonymization for privacy of facial biometrics via cyclic learning. *EURASIP Journal on Information Security, 2024*(1), 24. https://doi.org/10.1186/s13635-024-00174-3.

Xu, Z., Hong, Z., Ding, C., Zhu, Z., Han, J., Liu, J., & Ding, E. (2022, June). Mobilefaceswap: A lightweight framework for video face swapping. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 36, No. 3, pp. 2973-2981). https://doi.org/10.1609/aaai.v36i3.20203.

Yao, A., Pal, S., Dong, C., Li, X., & Liu, X. (2024, March). A framework for user biometric privacy protection in UAV delivery systems with edge computing. In *2024 IEEE International Conference on Pervasive Computing and Communications Workshops and other Affiliated Events (PerCom Workshops)* (pp. 631-636). IEEE. 10.1109/PerComWorkshops59983.2024.10502849.

Zhu, Y., Li, Q., Wang, J., Xu, C. Z., & Sun, Z. (2021). One shot face swapping on megapixels. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 4834-4844).

Esta obra está bajo una licencia de Creative Commons Atribución-NoComercial-CompartirIgual 4.0 Internacional.