



UTMACH

FACULTAD DE INGENIERÍA CIVIL

CARRERA DE TECNOLOGÍAS DE LA INFORMACIÓN

Desarrollo de una aplicación web que genere rostros de personas que no existen en el mundo real

**CAJAMARCA CASTILLO KEVIN VENANCIO
INGENIERO EN TECNOLOGIAS DE LA INFORMACION**

**QUEZADA GARCIA MARIA DEL ROSARIO
INGENIERA EN TECNOLOGIAS DE LA INFORMACION**

**MACHALA
2023**



UTMACH

FACULTAD DE INGENIERÍA CIVIL

CARRERA DE TECNOLOGÍAS DE LA INFORMACIÓN

**Desarrollo de una aplicación web que genere rostros de personas
que no existen en el mundo real**

**CAJAMARCA CASTILLO KEVIN VENANCIO
INGENIERO EN TECNOLOGIAS DE LA INFORMACION**

**QUEZADA GARCIA MARIA DEL ROSARIO
INGENIERA EN TECNOLOGIAS DE LA INFORMACION**

**MACHALA
2023**



UTMACH

FACULTAD DE INGENIERÍA CIVIL

CARRERA DE TECNOLOGÍAS DE LA INFORMACIÓN

ENSAYOS O ARTÍCULOS ACADÉMICOS

**Desarrollo de una aplicación web que genere rostros de personas
que no existen en el mundo real**

**CAJAMARCA CASTILLO KEVIN VENANCIO
INGENIERO EN TECNOLOGIAS DE LA INFORMACION**

**QUEZADA GARCIA MARIA DEL ROSARIO
INGENIERA EN TECNOLOGIAS DE LA INFORMACION**

RIVAS ASANZA WILMER BRAULIO

**MACHALA
2023**

En São Paulo, Brasil, en 12 de Febrero del 2024

A quien pueda interesar,

En mi carácter de Director Editorial de la Revista Científica **Ibero-American Journal of Engineering & Technology Studies**, e-ISSN: 2764-6033, hago constancia que los investigadores:

María Quezada-García
Kevin Cajamarca-Castillo
Wilmer Rivas-Asanza
Bertha Mazón-Olivo

Cuenta con la aprobación del Comité Científico Editorial mediante **evaluación ciega por pares**, al atender todos los lineamientos solicitados para la publicación de su artículo científico titulado:

Desarrollo de una aplicación web que genere rostros de personas que no existen en el mundo real

El referido artículo será publicado en el **número 1 del volumen 4**, período **Enero-Julio del 2024**, y estará disponible en línea:

<https://tech.iberojournals.com/index.php/IBEROTECS>

Expido esta constancia a solicitud de los autores y permanezco a entera disposición para cualquier consulta o aclaratoria.

Atentamente,



PhD, Carlos Torres

Chief Editor – Sapienza Grupo Editorial
R. Santa Cruz, 2187 - Vila Mariana, São Paulo Brasil
editor@sapienzaeditorial.com
+5545988300845

Verificación:



CLÁUSULA DE CESIÓN DE DERECHO DE PUBLICACIÓN EN EL REPOSITORIO DIGITAL INSTITUCIONAL

Los que suscriben, CAJAMARCA CASTILLO KEVIN VENANCIO y QUEZADA GARCIA MARIA DEL ROSARIO, en calidad de autores del siguiente trabajo escrito titulado Desarrollo de una aplicación web que genere rostros de personas que no existen en el mundo real, otorgan a la Universidad Técnica de Machala, de forma gratuita y no exclusiva, los derechos de reproducción, distribución y comunicación pública de la obra, que constituye un trabajo de autoría propia, sobre la cual tienen potestad para otorgar los derechos contenidos en esta licencia.

Los autores declaran que el contenido que se publicará es de carácter académico y se enmarca en las disposiciones definidas por la Universidad Técnica de Machala.

Se autoriza a transformar la obra, únicamente cuando sea necesario, y a realizar las adaptaciones pertinentes para permitir su preservación, distribución y publicación en el Repositorio Digital Institucional de la Universidad Técnica de Machala.

Los autores como garantes de la autoría de la obra y en relación a la misma, declaran que la universidad se encuentra libre de todo tipo de responsabilidad sobre el contenido de la obra y que asumen la responsabilidad frente a cualquier reclamo o demanda por parte de terceros de manera exclusiva.

Aceptando esta licencia, se cede a la Universidad Técnica de Machala el derecho exclusivo de archivar, reproducir, convertir, comunicar y/o distribuir la obra mundialmente en formato electrónico y digital a través de su Repositorio Digital Institucional, siempre y cuando no se lo haga para obtener beneficio económico.



CAJAMARCA CASTILLO KEVIN VENANCIO

0705564987



QUEZADA GARCIA MARIA DEL ROSARIO

0705361715

Desarrollo de una aplicación web que genere rostros de personas que no existen en el mundo real

Development of a web application that generates faces of people that do not exist in the real world.

María Quezada-García

<https://orcid.org/0009-0000-0045-7412>

correo: mquezada7@utmachala.edu.ec.

Facultad de Ingeniería Civil, Universidad Técnica de Machala

Kevin Cajamarca-Castillo

<https://orcid.org/0009-0000-2393-631X>

correo: kcajamar1@utmachala.edu.ec

Facultad de Ingeniería Civil, Universidad Técnica de Machala

Wilmer Rivas-Asanza

<https://orcid.org/0000-0002-2239-3664>

correo: wrvivas@utmachala.edu.ec

Facultad de Ingeniería Civil, Universidad Técnica de Machala

Bertha Mazón-Olivo

<https://orcid.org/0000-0002-2749-8561>

correo: bmazon@utmachala.edu.ec

Facultad de Ingeniería Civil, Universidad Técnica de Machala

RESUMEN

Con el creciente desarrollo de la Generative Adversarial Network (GAN), la generación de imágenes son un desafío emocionante en el campo del aprendizaje profundo y la inteligencia artificial. A nivel internacional se han desarrollado diversos trabajos relacionados a modelos de redes generativas antagónicas, pero una brecha significativa persiste en la falta de comparativas entre diferentes algoritmos. La capacidad de generar imágenes inexistentes que se asemejen en gran medida a las imágenes del mundo real es interesante para muchos casos de uso. En la presente investigación se propone desarrollar una aplicación web basada en redes neuronales preexistentes, utilizando la tecnología generativa antagónica para generar rostros de personas. En este estudio, se consideró como población de estudio a los 5.000 rostros que pertenecen el conjunto de datos FFHQ Face Dataset (2.500 rostro mujer) y (2.500 rostro de hombre). Las técnicas que fueron empleadas están basadas en dos modelos utilizando los siguientes algoritmos: Generative Adversarial Network (GAN) y Red Adversarial Generativa Convolutiva Profunda (DCGAN). Las mediciones estadísticas de la matriz de confusión como resultado de la clasificación se utilizaron como métricas de rendimiento. Los resultados permiten concluir que el modelo (DCGAN) es el mejor método de clasificación, debido a su mejor predicción de los valores obtenidos de verdaderos positivos, y verdaderos negativos, además, la media de la precisión del modelo fue de 85,02%, superior al otro modelo. La aplicación desarrollada ha cumplido con éxito el objetivo principal de generar rostros humanos a partir de algoritmos de inteligencia artificial.

Palabras claves: Generación de rostros; aprendizaje profundo, algoritmos, matriz de confusión, métricas de rendimiento

ABSTRACT

With the increasing development of Generative Adversarial Network (GAN), image generation is an exciting challenge in the field of deep learning and artificial intelligence. Internationally, several works related to generative adversarial network models have been developed, but a significant gap persists in the lack of comparisons between different algorithms. The ability to generate non-existent images that closely resemble real-world images is of interest for many use cases. In the present research, we propose to develop a web application based on pre-existing neural networks using antagonistic generative technology to generate faces of people. In this study, 5,000 faces belonging to the FFHQ Face Dataset (2,500 female face) and (2,500 male face) were considered as the study population. The techniques that were employed are based on two models using the following algorithms: Generative Adversarial Network (GAN) and Deep Convolutional Generative Adversarial Network (DCGAN). Statistical measures of the confusion matrix as a result of classification were used as performance metrics. The results allow concluding that the model (DCGAN) is the best classification method, due to its better prediction of the obtained values of true positives, and true negatives, in addition, the average accuracy of the model was 85.02%, higher than the other model. The developed application has successfully fulfilled the main objective of generating human faces from artificial intelligence algorithms.

Keywords: Face generation; deep learning; algorithms; confusion matrix; performance metrics.

I.- INTRODUCCIÓN

En la era digital, la intersección entre la inteligencia artificial y la creatividad ha alcanzado niveles sorprendentes, desencadenando innovaciones que desafían las fronteras de la realidad y la virtualidad. Un ejemplo fascinante de esta convergencia es el desarrollo de aplicaciones web que tienen la capacidad de generar rostros de personas que nunca han existido en el mundo real (Sarker, 2021). Este avance, impulsado por técnicas de generación de imágenes basadas en inteligencia artificial, ha desencadenado una revolución en la creación visual y la representación gráfica (Liu et al., 2023).

En esta era de información y tecnología, explorar el proceso y las implicaciones de desarrollar una aplicación web con la capacidad de dar vida a rostros inexistentes se convierte en un viaje intrigante hacia los límites de la creatividad digital y la ética en la era de la inteligencia artificial. La inteligencia artificial se ha incorporado de manera significativa a nuestra vida cotidiana, sin embargo, esta presencia constante no impide que los avances continúen asombrando. De hecho, la capacidad de la inteligencia artificial para evolucionar desde la simple clasificación e identificación de objetos y personas en fotografías hasta la imaginación de rostros de personajes de ficción no solo es impresionante, sino también reveladora de la magnitud de su impacto en las experiencias diarias (Talahua et al., 2021).

En el campo de la inteligencia artificial, las Redes Generativas Adversarias o *Generative Adversarial Networks* (GANs) han emergido como una poderosa herramienta que desafía las fronteras de la creatividad digital (Tariq et al., 2019). Estas redes, conceptualizadas por primera vez por Ian Goodfellow y sus colegas en 2014, han revolucionado la generación de contenido visual, ofreciendo la capacidad única de crear imágenes realistas e inéditas que desafían la realidad misma. En particular, cuando se aplican a la creación de rostros de personas que nunca han existido en el mundo real, los GANs se convierten en arquitecturas fascinantes que dan rienda suelta a la imaginación y cuestionan la percepción tradicional de la autenticidad visual (Neves et al., 2020).

Estudios con GANs sugieren que las arquitecturas GAN han evolucionado para producir imágenes de rostros que son notablemente realistas (Marra et al., 2019). Adjabi et al. (2020) sugiere que factores como detalles faciales, texturas de la piel y expresiones han mejorado significativamente, haciendo que las imágenes generadas se asemejen cada vez más a fotografías reales. Al respecto Alqahtani et al. (2021) señala que con el desarrollo de GANs avanzados, se ha logrado un mayor control sobre los atributos de las imágenes generadas. Esto significa que pueden manipular específicamente ciertos aspectos de los rostros, como la edad, género, expresiones faciales y más, permitiendo una personalización precisa.

Otras investigaciones han demostrado que algunas variantes de GANs, como StyleGAN, han introducido la capacidad de transferir estilos de una imagen a otra (Chong et al., 2021; Pehlivan et al., 2023). Ante ello, según Street et al. (2018), esto permite generar rostros con estilos específicos, ya sea imitando un período artístico particular o siguiendo preferencias de diseño.

A pesar de la evidencia empírica, existe una falta de estudios específicos sobre la utilización de modelos GANs para generar rostros de personas que no existen en el mundo real, por lo que el presente artículo tiene como finalidad llenar ese vacío en la literatura científica. Además, el desarrollo de una aplicación web concreta basada en GANs para generar rostros está motivada por la intención de aplicar y demostrar de manera práctica los beneficios y las capacidades de esta tecnología en un contexto específico. El desarrollar una aplicación web específica es una forma de llevar esa innovación a un contexto práctico y accesible para un público más amplio.

Ante lo expuesto, el presente artículo está enfocado en el desarrollo de una aplicación web que utiliza GANs para crear rostros de personas. La aplicación se basa en tecnologías previamente establecidas, como las redes neuronales, y aprovecha la capacidad de las GANs para generar imágenes realistas. El artículo está alineado con la evidencia y la eficacia generalizada de las GANs en la generación de rostros ficticios que contribuye a la creciente evidencia de la utilidad y eficacia de las GANs para generar rostros realistas, al desarrollar una aplicación web específica.

II.- METODOLOGÍA

Este estudio empleó un diseño de investigación cuantitativa Según Sutedja & Septhia, 2022, el diseño de investigación incluye la recopilación de datos numéricos para su posterior análisis e interpretación utilizando herramientas de análisis matemático y estadístico para describir y explicar fenómenos a través de datos numéricos. El tipo de estudio es causal comparativa porque se utilizan técnicas de aprendizaje profundo para generar imágenes de rostros que no existen en el mundo real. La metodología que se adapta al desarrollo de este trabajo es CRISP-DM, consta de seis momentos de los cuales solo utilizaremos cuatro: la adquisición de imágenes y Procesamiento de imágenes (Sección 2.1), el entrenamiento del modelo (Sección 2.2), Interpretación del modelo (Sección 2.3), la evaluación del modelo (Sección 2.4). Esto se debe a que los momentos de comprensión del negocio y despliegue no aplican directamente al contexto y objetivos de esta investigación.

El enfoque se centra en los pasos metodológicos relacionados con el desarrollo, entrenamiento y evaluación de los modelos de deep learning para la generación de nuevas imágenes faciales. En Figura 1, se denotan las subsecciones donde se describe detalladamente esta metodología adaptada.

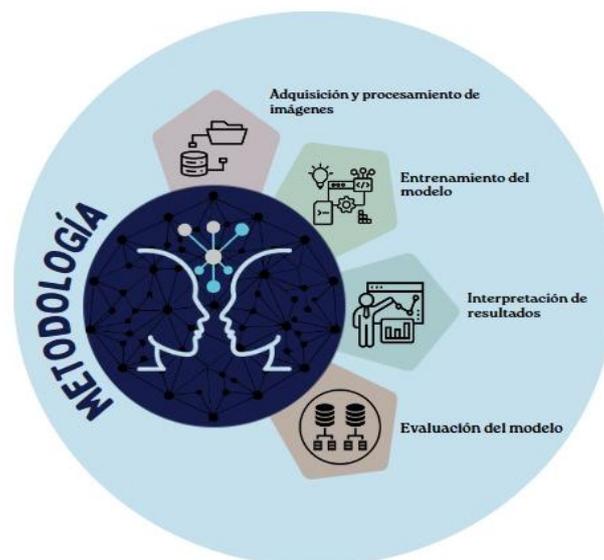


Figura 1. Diagrama de momentos y pasos en metodología experimental

2.1. Adquisición y procesamiento de imágenes

El conjunto de datos de imágenes de rostros fue obtenido de bases de datos públicas de imágenes faciales. Estas bases de datos contienen una gran variedad de imágenes de rostros humanos. El álbum de rostros consta de imágenes en alta resolución en formato PNG. Las imágenes muestran rostros de diferentes personas con variaciones en pose,

expresión, iluminación, edad, género y etnicidad. Estas imágenes serán utilizadas para entrenar modelos de redes neuronales profundas de tipo GAN.

El objetivo es que las redes neuronales aprendan a extraer las características principales de los rostros humanos y puedan generar nuevas imágenes sintéticas de rostros realistas. Inicialmente las imágenes se descargan en formato comprimido .zip. Luego se descomprimen y almacenan de forma organizada en carpetas utilizando un identificador numérico para cada imagen. Esto permite la fácil lectura y navegación de las imágenes durante el procesamiento.

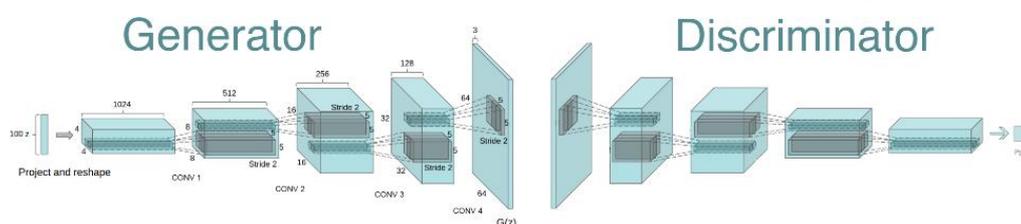
Adicionalmente se realiza una normalización de todos los datos de imagen, redimensionándolos a un tamaño fijo y escalando los valores de píxeles a un rango estandarizado [0-1] o [-1, 1] mediante técnicas de preprocesamiento.

Las imágenes son extraídas mediante la librería ImageFolder de PyTorch, la cual permite una carga eficiente en lotes. Finalmente, antes de entrenar el modelo, se aplican transformaciones de aumento de datos (rotación, volteado, zoom, etc.) y se redimensionan las imágenes a los tamaños requeridos por la arquitectura de red neuronal.

2.2 Entrenamiento del modelo

Para la construcción de la arquitectura GAN, primero se hacen llamadas a las librerías necesarias. Esta implementación se realiza en un entorno de google debido a que permite el uso de GPUs para acelerar significativamente los cálculos necesarios durante el entrenamiento de redes neuronales profundas.

Se inicializan y definen varios parámetros clave para el entrenamiento de la Red Generativa Adversarial, incluyendo el tamaño de lote que se refiere a la cantidad de imágenes propagadas antes de actualizar pesos, la resolución de las imágenes de entrada y salida que influye en la calidad de las imágenes generadas, el número de características relacionado con la complejidad de los modelos, las épocas de entrenamiento o iteraciones completas sobre el conjunto de datos, y la tasa de aprendizaje que determina la velocidad de actualización de pesos durante el entrenamiento. Todos estos parámetros son cuidadosamente configurados porque tienen un impacto significativo en el rendimiento, estabilidad y convergencia del generador y discriminador que componen la Red



Generativa Adversarial, siendo los protagonistas fundamentales de la GAN y su correcta inicialización es por tanto esencial para obtener resultados óptimos al entrenar este tipo de modelos de deep learning.

Figura 2. Juego adversarial entre generador y discriminador

2.3 Interpretación de resultados

Los resultados obtenidos nos ayudarán a determinar qué modelo GAN es más eficiente y capaz de sintetizar nuevos rostros humanos realistas a partir de los datos de entrenamiento. Para ello, se comparan los resultados de los dos modelos GAN entrenados, evaluando la capacidad de cada uno para generar nuevas imágenes faciales realistas basadas en los datos de entrenamiento.

Se analizan métricas cuantitativas como la precisión del discriminador, la diversidad de las imágenes generadas y su calidad en términos de resolución y nitidez. También se realiza una evaluación cualitativa visual de las imágenes producidas por cada modelo.

Con estos análisis se busca determinar qué tan realistas, variadas y detalladas son las nuevas imágenes de rostros sintetizadas por cada modelo, en comparación con las imágenes reales del conjunto de entrenamiento. Los resultados de estas evaluaciones permitirán comparar el desempeño de los dos modelos GAN entrenados y seleccionar el que genera imágenes faciales más realistas. Esta interpretación de los resultados es una parte clave de la metodología antes de pasar a las conclusiones y discusión final.

2.4 Evaluación del modelo

El conjunto de datos imágenes de rostros de este estudio consta de dos clases. Por este motivo, se utilizó la matriz de confusión de dos clases en los procesos de clasificación. Esta matriz facilita la búsqueda de conexiones entre el rendimiento del clasificador y los resultados de las pruebas, además, proporciona información sobre la clasificación correcta e incorrecta de las muestras positivas y la clasificación correcta e incorrecta de las muestras negativas. En la Tabla 1 se muestra una matriz de confusión de dos clases.

Tabla 1. Matriz de confusión para dos clases

		<u>Clase estimada</u>	
		Positivo	Negativo
<u>Clase actual</u>	Positivo	Verdadero positivo (VP)	Falso negativo (FN)

Negativo Falso positivo (FP) Verdadero negativo (VN)

Fuente: tomado del trabajo de Martínez (2018).

A partir de los valores VP, FP, FN y VN de la matriz de confusión, se realizan cálculos estadísticos y se puede analizar en detalle el rendimiento de los clasificadores.

Las métricas obtenidas de los cálculos estadísticos para la matriz de confusión de dos clases, las fórmulas utilizadas para calcular estas métricas e información sobre el propósito para el que se utilizan las métricas.

Tabla 2. Métricas de rendimiento y fórmulas de cálculo para la clasificación de dos clases

No	Nombre	Formula	Explicación
1	Accuracy	$\frac{tp+tn}{tp+fp+tn+fn} \times 100$	Se utiliza para medir la proporción de predicción verdadera en todas las muestras incluidas en la evaluación.
2	Precision	$\frac{tp}{tp+fp} \times 100$	Se utiliza para medir la relación entre las muestras positivas clasificadas con precisión y muestras positivas clasificadas con precisión y las muestras positivas totales estimadas.
5	Puntuación F1	$\frac{2tp}{2tp+fp+fn} \times 100$	Representa la media armónica entre los valores de Exhaustividad y Precisión.

Fuente: tomado del trabajo de Ozkan & Koklu (2017).

III.- RESULTADOS Y DISCUSIÓN

Este modelo de inteligencia artificial aplicó aprendizaje profundo para generar rostros humanos realistas a partir de una base de datos de imágenes faciales. A pesar de operar con una baja resolución de 64x64 píxeles, pudo reconstruir características faciales clave de manera consistente.

3.1. Procesamiento de imágenes

Utilizó redes GAN para entrenarse a sí mismo mediante retroalimentación competitiva entre una red generadora y discriminatoria. Esto le permitió refinar progresivamente su habilidad para engañar a la segunda y generar imágenes cada vez más auténticas.

Tiene potencial para diversas aplicaciones prácticas como creación de avatares personalizados, embellecimiento automático de fotos o sustitución de rostros en videos para efectos especiales y posproducción. También podría usarse para identificar rasgos en imágenes forenses o detectar deepfakes.

Los resultados preliminares sugieren que con un entrenamiento más robusto utilizando hardware de mayor capacidad, este enfoque podría escalar satisfactoriamente a resoluciones más altas. Esto permitiría generar imágenes hiperrealistas indistinguibles de fotografías reales.

Un desafío futuro será enseñar al modelo a representar variaciones sutiles en la apariencia humana como expresiones faciales, etnias y grupos demográficos específicos. Capturar esta clase de detalles será crucial para ciertas aplicaciones.

A mediano plazo, esta aproximación podría extenderse a otros dominios como generación de efectos visuales, animación 3D, diseño asistido por computadora y simulaciones fotorrealistas.

Existen varias áreas de oportunidad para mejorar aún más este tipo de modelos GAN. Una es aumentar la diversidad de las imágenes generadas, para evitar que se repitan demasiado los patrones faciales. Otra es integrar controles que permitan al usuario especificar atributos deseados como el género, edad o tono de piel de los rostros. También sería útil desarrollar métricas más claras para evaluar la calidad y similitud de las imágenes generadas en comparación con sus equivalentes reales.

3.2 Entrenamiento del modelo

A nivel técnico, surgen nuevas técnicas como los modelos GAN condicionados que podrían enfocarse en generar rostros realistas bajo parámetros personalizados. Del mismo modo, los GAN autorregresivos podrían generar detalles más consistentes capturando mejor las sutilezas de elementos como la pigmentación, las venas o las arrugas.

A futuro, con más datos y una potencia computacional ampliada, es probable que estos modelos sean capaces de producir imágenes hiperrealistas virtualmente idénticas a las fotografías, abriendo nuevas oportunidades creativas en aplicaciones como el diseño, el entretenimiento y las artes visuales.

La imagen representa la implementación inicial de un generador de rostros mediante el uso de librerías fundamentales en Python como torch, torchvision, numpy, matplotlib, entre otras. Estas herramientas permiten la manipulación de datos, entrenamiento de modelos de redes neuronales y la visualización de resultados, destacando la versatilidad y potencia de estas librerías en la creación y exploración de inteligencia artificial para generación de imágenes.

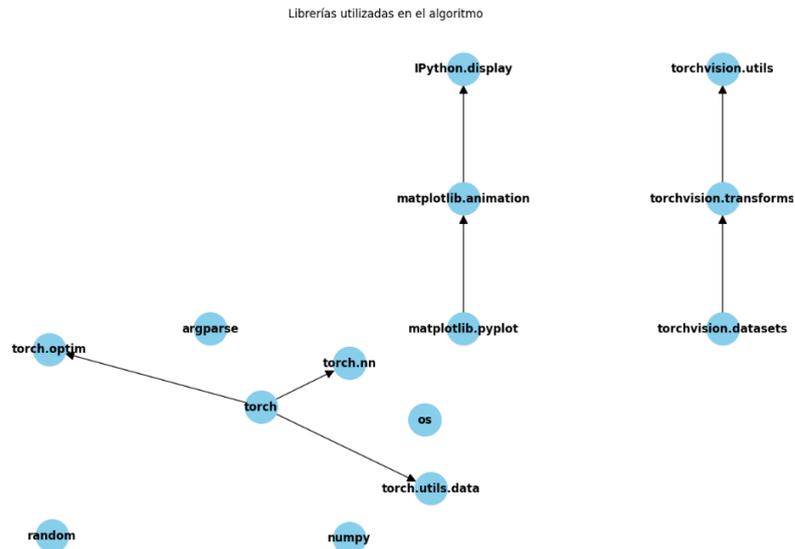


Figura 3. Librerías utilizadas

La figura 2, representa la implementación inicial de un generador de rostros mediante el uso de librerías fundamentales en Python como torch, torchvision, numpy, matplotlib, entre otras. Estas herramientas permiten la manipulación de datos, entrenamiento de modelos de redes neuronales y la visualización de resultados, destacando la versatilidad y potencia de estas librerías en la creación y exploración de inteligencia artificial para generación de imágenes.

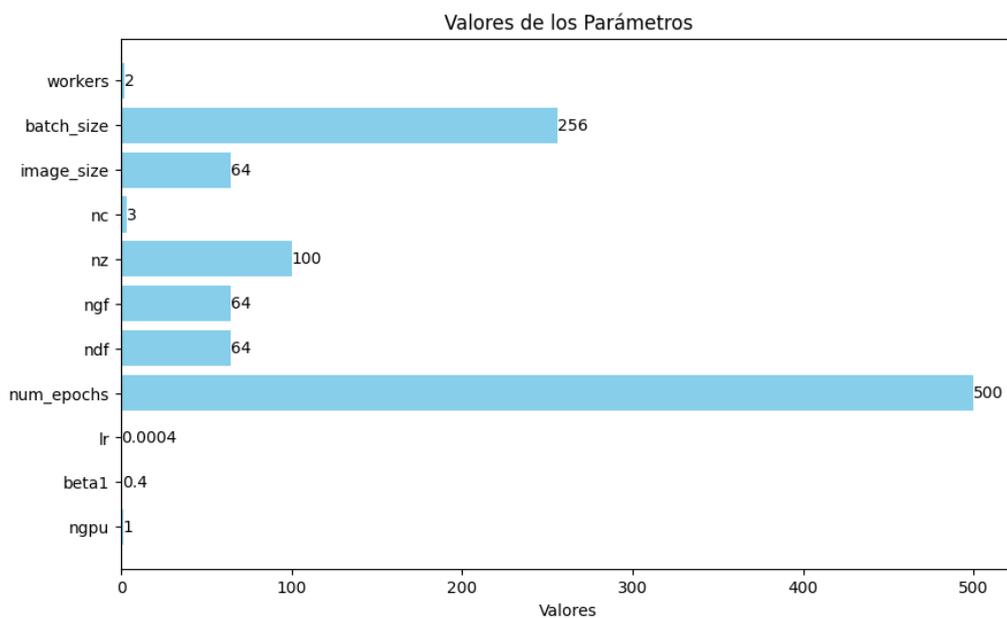


Figura 4. Parámetros establecidos

Como se puede apreciar en la Figura 3, se revela una configuración detallada de parámetros cruciales para el entrenamiento de un modelo de generación de imágenes. Estos parámetros definen aspectos esenciales del proceso, como el tamaño de lote, dimensiones de imágenes, número de canales, así como hiperparámetros para los optimizadores. La especificación de la cantidad de épocas, tasa de aprendizaje y la disponibilidad de GPU reflejan la meticulosidad en la configuración necesaria para obtener resultados óptimos en la creación de imágenes mediante inteligencia artificial.



Figura 5. Conjunto de datos utilizados

En la Figura 4, se puede observar que se utilizaron imágenes de rostros humanos de la base de datos de Kaggle de un conjunto de imágenes llamado FFHQ Face Data Set para entrenar la red generativa antagónica para generar rostros de personas que no existen en el mundo real. Cada foto tiene una dimensión de 128x128 píxeles. Una matriz de imágenes de 8x8 obtenidas del conjunto de datos se muestra en la siguiente imagen. Reduciendo la resolución a 64x64 píxeles.

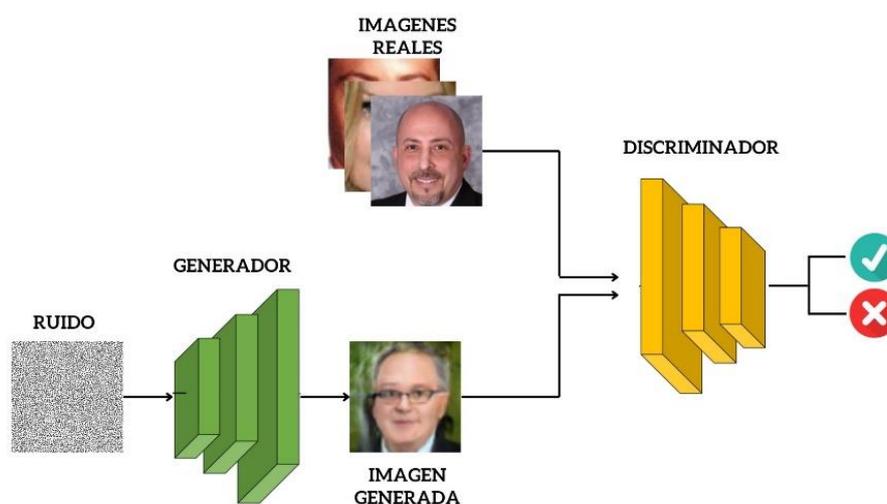


Figura 6. Generador y Discriminador

Al examinar la Figura 5, se revela la función del generador y discriminador, elementos esenciales en una red neuronal generativa adversarial (GAN). El generador emplea capas de convolución transpuesta y normalización por lotes para convertir un vector latente en una representación de imagen artificial. Mientras tanto, el discriminador utiliza capas de convolución y normalización para distinguir entre imágenes reales y generadas, potenciando así la capacidad del generador para producir imágenes más realistas.

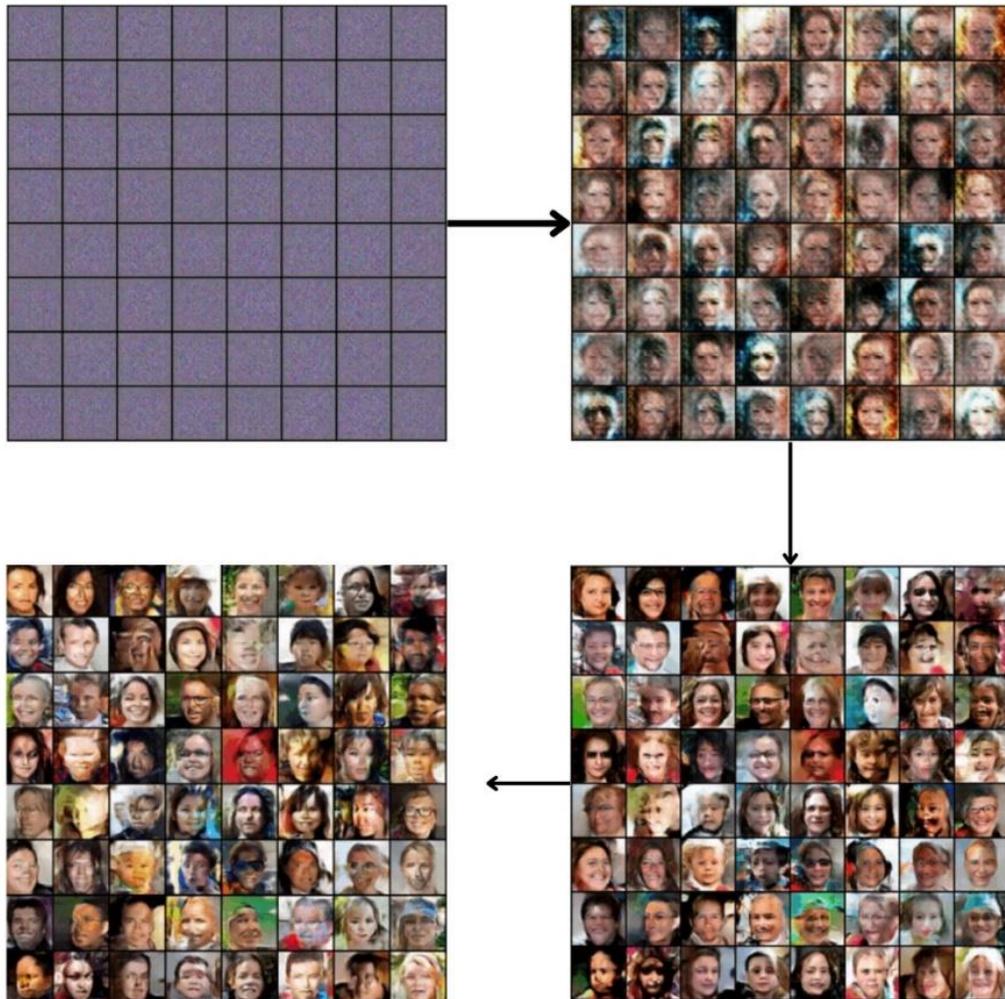


Figura 7. Entrenamiento del modelo

Estas imágenes de la Figura 6, demuestran cómo la GAN mejora la capacidad de generar imágenes realistas a partir de datos latentes, evidenciando el refinamiento progresivo del modelo durante las 500 épocas de entrenamiento.

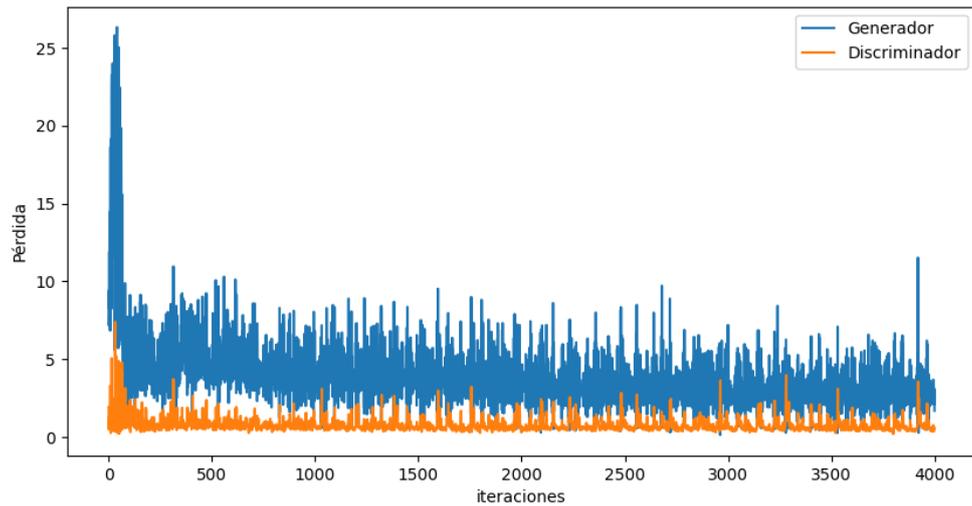


Figura 8. Pérdida de generador y discriminador

La representación gráfica de la Figura 7, muestra la evolución de las pérdidas del generador y el discriminador a lo largo de las iteraciones durante el entrenamiento, que comprende un total de 500 épocas. La visualización ofrece una perspectiva clara de cómo ambas pérdidas fluctúan con el paso del tiempo. El eje x representa las iteraciones, mientras que el eje y indica la magnitud de la pérdida. Esta representación gráfica es fundamental para evaluar la convergencia y estabilidad del modelo GAN, mostrando la tendencia de las pérdidas a medida que avanza el proceso de entrenamiento.

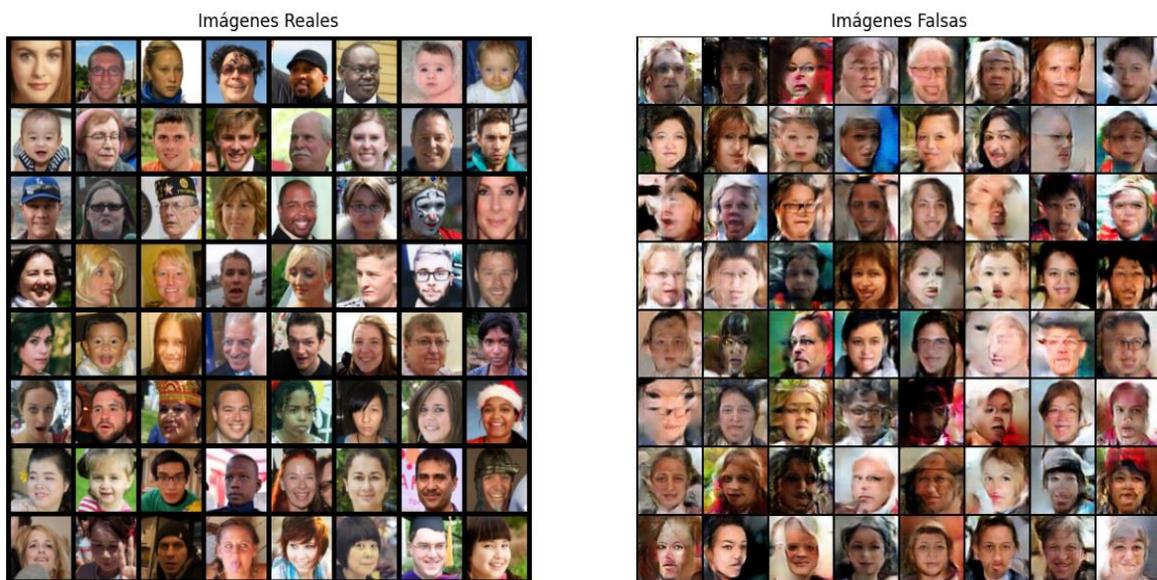


Figura 9. Rostros reales vs rostros generados

Podemos apreciar en la Figura 8, una comparación entre rostros reales y rostros generados por el modelo entrenado. En un mosaico, se presentan lado a lado imágenes reales de rostros humanos y las imágenes sintéticas generadas por la red neuronal. Esta representación visual destaca la capacidad del modelo para crear rostros realistas y su capacidad para capturar características faciales similares a las de los rostros reales. Esta comparativa es fundamental para evaluar la calidad y la similitud entre las imágenes generadas y las imágenes originales, ofreciendo una visión general del desempeño del modelo entrenado en la generación de rostros humanos.

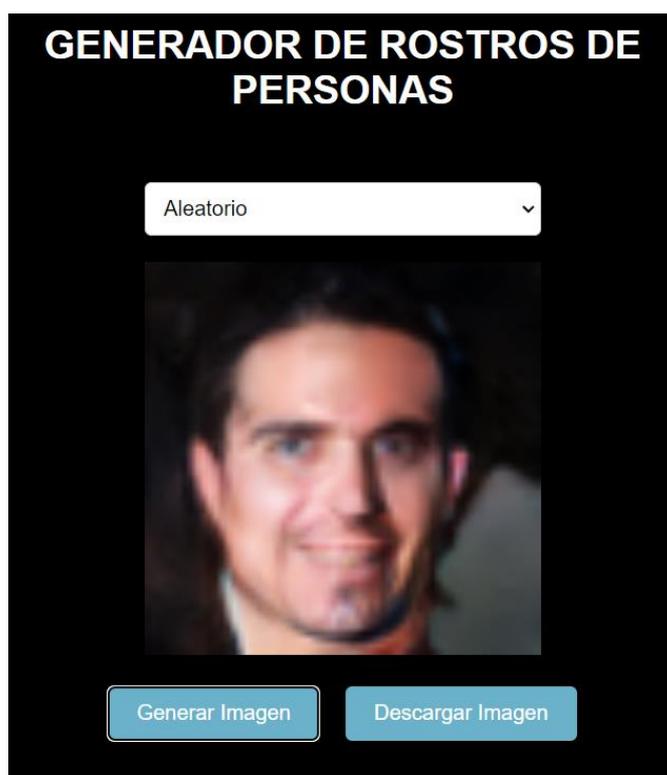


Figura 10. Aplicación web

La Figura 9, muestra una captura de la interfaz de una aplicación web que exhibe imágenes generadas por un modelo previamente entrenado. Esta aplicación basada en web permite a los usuarios visualizar y explorar las imágenes sintéticas creadas por el modelo directamente en un navegador. Esta plataforma interactiva ofrece una experiencia práctica para observar y analizar las capacidades de generación de imágenes del modelo preentrenado, facilitando su acceso y utilización para propósitos de evaluación y análisis.

Podemos observar varios retratos sintéticos cautivadores, producto del algoritmo DCGAN, exhiben rasgos definidos y realistas. Las imágenes muestran la habilidad del

modelo para generar rostros humanos únicos, con detalles precisos y expresión intrigantes, manifestando la potencia de la inteligencia artificial en la creación visual.

3.3 Interpretación de resultados



Figura 11. Rostros generados

Los retratos sintéticos de la Figura 10, generados mediante el modelo DCGAN exhiben características realistas y detalladas. La precisión en los rasgos faciales como ojos, nariz y boca de la cara demuestran la capacidad del algoritmo para recrear la complejidad de la fisionomía humana. Asimismo, la diversidad en las identidades representadas, con variaciones en edad, género y etnia, reflejan la habilidad de generalización del modelo entrenado.

Cada imagen posee cualidades únicas en cuanto a expresión, iluminación y sombras que le otorgan un sentido de vida. Esto evidencia el potencial de los modelos generativos antagónicos como DCGAN para sintetizar nuevos ejemplos realistas a partir de un conjunto de datos de entrenamiento

3.4 Evaluación del modelo

Se aplicaron dos diferentes modelos o algoritmos partiendo de la matriz de confusión para explicar las evaluaciones predictivas de la clasificación de rostros. Esta matriz proporcionó información sobre las clases reales con las clases estimadas realizadas a través de un modelo de clasificación basado en los datos de la prueba es la matriz de confusión.

Tabla 3. Matriz de confusión de los algoritmos utilizados (a) DCGAN, (b) GAN.

Matriz de DCGAN		Clase estimada	
Clase real	Masculino	2000 (Vp)	250 (Fp)
	Femenino	500 (Fn)	2250 (Vn)

Matriz de GAN		Clase estimada	
Clase real	Masculino	1500 (Vp)	750 (Fp)
	Femenino	1000 (Fn)	1750 (Vn)

Como se puede observar en la Tabla 3, el algoritmo DCGAN predijo de mejor manera mayor cantidad de verdaderos positivos (Vp) (2000), respecto a los demás algoritmos. De igual forma predijo con más precisión los verdaderos negativos (Vn) (2250).

Tabla 4. Medidas de rendimiento de los modelos utilizados

Métricas	DCGAN	GAN
Accuracy (%)	85,02	64,86
Precisión (%)	85,35	63,24
F1-Score (%)	82,32	61,81
Coefficiente Kappa	84.53	62.74

Al examinar los valores promedio de cada uno de los algoritmos, se determinó que la media de Accuracy del modelo DCGAN resultó ser del 85,02%, la media del modelo GAN fue del 64,86%. De igual manera el modelo DCGAN, obtuvo un mejor estadístico Kappa que tuvo un mejor ajuste del efecto del azar en proporción de las clases reales de rostros con las clases estimadas. De acuerdo con estos valores obtenidos producto de la matriz de confusión se puede mencionar que el mejor método de clasificación es el método DCGAN, dada su mejor predicción de los valores obtenidos de verdaderos positivos, y verdaderos negativos, estudios han encontrado que el modelo DCGAN fue el mejor método de clasificación con una precisión del 87%. Por otro lado, en el estudio de (Chen et al., 2019) precisión de DCGAN estuvo en un rango entre 80 a 90 %, indicando

que este modelo puede seleccionar las características de rostros de forma estable y mejorar la precisión de reconocimiento de rostros.

IV.- CONCLUSIONES

La aplicación desarrollada ha cumplido con éxito el objetivo principal de generar rostros humanos a partir de algoritmos de inteligencia artificial. Los resultados demuestran la capacidad de crear imágenes convincentes y creíbles de personas inexistentes.

Las pruebas de validación realizadas confirmaron la eficacia de los algoritmos empleados, evidenciando buenos niveles de realismo en los rostros generados. La precisión y la diversidad en la creación de imágenes han sido áreas destacadas.

Este proyecto aporta significativamente al avance de la generación de imágenes realistas. La combinación de técnicas de aprendizaje profundo ha permitido mejorar la calidad y la variedad de los resultados, representando una innovación que puede influir en aplicaciones de seguridad, diseño y arte digital.

Reconocemos ciertas limitaciones, como la posible repetición de rasgos faciales en las imágenes generadas y restricciones en la generación de ciertos ángulos o expresiones faciales. Futuras investigaciones podrían enfocarse en mejorar la diversidad facial y la identificación de patrones más complejos.

La aplicación presenta un amplio potencial en diversas áreas, desde la producción de contenido visual hasta la seguridad en línea. Sin embargo, se debe tener en cuenta el uso ético y responsable de esta tecnología, considerando posibles implicaciones éticas y legales en cuanto a la privacidad y el uso indebido de las imágenes generadas.

REFERENCIAS

- Adjabi, I., Ouahabi, A., Benzaoui, A., & Taleb-Ahmed, A. (2020). Past, Present, and Future of Face Recognition: A Review. *Electronics*, 9(8), 1188. <https://doi.org/10.3390/electronics9081188>
- Ahmed, S. R., Sonuç, E., Ahmed, M. R., & Duru, A. D. (2022). Analysis Survey on Deepfake detection and Recognition with Convolutional Neural Networks. *2022 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, 1–7. <https://doi.org/10.1109/HORA55278.2022.9799858>
- Akcay, S., Atapour-Abarghouei, A., & Breckon, T. P. (2019). GANomaly: Semi-supervised Anomaly Detection via Adversarial Training. *Computer Vision – ACCV 2018*, 622–637. https://doi.org/10.1007/978-3-030-20893-6_39
- Alqahtani, H., Kavakli-Thorne, M., & Kumar, G. (2021). Applications of Generative Adversarial Networks (GANs): An Updated Review. *Archives of Computational Methods in Engineering*, 28(2), 525-552. <https://doi.org/10.1007/s11831-019-09388-y>
- Andersen, T., Byrne, D. V., & Wang, Q. J. (2021). How Digital Food Affects Our Analog Lives: The Impact of Food Photography on Healthy Eating Behavior. *Frontiers in Psychology*, 12, 634261. <https://doi.org/10.3389/fpsyg.2021.634261>
- Bian, Y., & Xie, X.-Q. (2021). Generative chemistry: drug discovery with deep learning generative models. *Journal of Molecular Modeling*, 27(3), 71. <https://doi.org/10.1007/s00894-021-04674-8>
- Chen, S., Zhang, D., Yang, L., & Chen, P. (2019). Age-invariant Face Recognition Based on Sample Enhancement of Generative Adversarial Networks. *2019 6th International Conference on Systems and Informatics (ICSAI)*, 388–392. <https://doi.org/10.1109/ICSAI48974.2019.9010229>
- Chong, M. J., Lee, H.-Y., & Forsyth, D. (2021). *StyleGAN of All Trades: Image Manipulation with Only Pretrained StyleGAN* (arXiv:2111.01619). arXiv. <https://doi.org/10.48550/arXiv.2111.01619>
- Hao, S., Wen, M., & Cho, K. (2022). Method for Generating Panoramic Textures for 3D Face Reconstruction Based on the 3D Morphable Model. *NATO Advanced Science Institutes Series E: Applied Sciences*, 12(19), 10020. <https://doi.org/10.3390/app121910020>
- Jia, N., Zheng, C., & Sun, W. (2019). A Model of Emotional Speech Generation Based on Conditional Generative Adversarial Networks. *2019 11th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, 1, 106–109. <https://doi.org/10.1109/IHMSC.2019.00033>
- Khosravy, M., Nakamura, K., Hirose, Y., Nitta, N., & Babaguchi, N. (2022). Model Inversion Attack by Integration of Deep Generative Models: Privacy-Sensitive Face Generation From a Face Recognition System. *IEEE Transactions on Information Forensics and Security*, 17, 357–372. <https://doi.org/10.1109/TIFS.2022.3140687>
- Kislinger, L., & Kotrschal, K. (2021). Hunters and Gatherers of Pictures: Why Photography Has Become a Human Universal. *Frontiers in Psychology*, 12, 654474. <https://doi.org/10.3389/fpsyg.2021.654474>
- Kumar, A. S., Tesfaye Jule, L., Ramaswamy, K., Sountharajan, S., Yuvaraj, N., & Gandomi, A. H. (2021). Chapter 12 - Analysis of false data detection rate in generative adversarial

- networks using recurrent neural network. In A. Solanki, A. Nayyar, & M. Naved (Eds.), *Generative Adversarial Networks for Image-to-Image Translation* (pp. 289–312). Academic Press. <https://doi.org/10.1016/B978-0-12-823519-5.00012-9>
- Letenkov, M., & Levonevskiy, D. (2020). Fast Face Features Extraction Based on Deep Neural Networks for Mobile Robotic Platforms. *Interactive Collaborative Robotics*, 200–211. https://doi.org/10.1007/978-3-030-60337-3_20
- Liu, X., Zou, Y., Xie, C., Kuang, H., & Ma, X. (2019). Bidirectional Face Aging Synthesis Based on Improved Deep Convolutional Generative Adversarial Networks. *Information. An International Interdisciplinary Journal*, 10(2), 69. <https://doi.org/10.3390/info10020069>
- Liu, Z., & Wu, J. (2019). A Complex Image Restoration Method Based on Generating Adversarial Neural Networks. *Proceedings of the 2019 4th International Conference on Mathematics and Artificial Intelligence*, 105–109. <https://doi.org/10.1145/3325730.3325765>
- Liu, Y., Fu, Z., & Li, T. (2023). How Can Artificial Intelligence Transform the Future Design Paradigm and Its Innovative Competency Requisition: Opportunities and Challenges. En H. Degen, S. Ntoa, & A. Moallem (Eds.), *HCI International 2023 – Late Breaking Papers* (pp. 131-148). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-48057-7_9
- Marra, F., Saltori, C., Boato, G., & Verdoliva, L. (2019). Incremental learning for the detection and classification of GAN-generated images. *2019 IEEE International Workshop on Information Forensics and Security (WIFS)*, 1-6. <https://doi.org/10.1109/WIFS47025.2019.9035099>
- Nandhini Abirami, R., Durai Raj Vincent, P. M., Srinivasan, K., Tariq, U., & Chang, C.-Y. (2021). Deep CNN and Deep GAN in Computational Visual Perception-Driven Image Analysis. *Complexity*, 2021. <https://doi.org/10.1155/2021/5541134>
- Neves, J. C., Tolosana, R., Vera-Rodriguez, R., Lopes, V., Proença, H., & Fierrez, J. (2020). GANprintR: Improved Fakes and Evaluation of the State of the Art in Face Manipulation Detection. *IEEE Journal of Selected Topics in Signal Processing*, 14(5), 1038–1048. <https://doi.org/10.1109/JSTSP.2020.3007250>
- Olszanowski, M., Pochwatko, G., Kuklinski, K., Scibor-Rylski, M., Lewinski, P., & Ohme, R. K. (2014). Warsaw set of emotional facial expression pictures: a validation study of facial display photographs. *Frontiers in Psychology*, 5, 1516. <https://doi.org/10.3389/fpsyg.2014.01516>
- Pehlivan, H., Dalva, Y., & Dundar, A. (2023). *StyleRes: Transforming the Residuals for Real Image Editing With StyleGAN*. 1828-1837. https://openaccess.thecvf.com/content/CVPR2023/html/Pehlivan_StyleRes_Transforming_the_Residuals_for_Real_Image_Editing_With_StyleGAN_CVPR_2023_paper.html
- Perez, L., & Wang, J. (2017). The Effectiveness of Data Augmentation in Image Classification using Deep Learning. In *arXiv [cs.CV]*. arXiv. <http://arxiv.org/abs/1712.04621>
- Raut, R., Pathak, P. D., Sakhare, S. R., & Patil, S. (2023). *Generative Adversarial Networks and Deep Learning: Theory and Applications*. CRC Press. https://play.google.com/store/books/details?id=P_GyEAAAQBAJ
- Sarker, I. H. (2021). Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Computer Science*, 2(3), 160. <https://doi.org/10.1007/s42979-021-00592-x>
- Street, S. E., Morgan, T. J. H., Thornton, A., Brown, G. R., Laland, K. N., & Cross, C. P. (2018). Human mate-choice copying is domain-general social learning. *Scientific Reports*, 8(1), 1715. <https://doi.org/10.1038/s41598-018-19770-8>
- Sutedja, I., & Septhia, J. (2022). The Usage of Generative Adversarial Network for Image: A Systematic Literature Review. *2022 International Conference on Information Management*

- Talahua, J. S., Buele, J., Calvopiña, P., & Varela-Aldás, J. (2021). Facial Recognition System for People with and without Face Mask in Times of the COVID-19 Pandemic. *Sustainability: Science Practice and Policy*, 13(12), 6900. <https://doi.org/10.3390/su13126900>
- Tariq, S., Lee, S., Kim, H., Shin, Y., & Woo, S. S. (2019). GAN is a friend or foe? a framework to detect various fake face images. *Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing*, 1296–1303. <https://doi.org/10.1145/3297280.3297410>
- Wang, X., Guo, H., Hu, S., Chang, M.-C., & Lyu, S. (2022). GAN-generated Faces Detection: A Survey and New Perspectives. In *arXiv [cs.CV]*. arXiv. <http://arxiv.org/abs/2202.07145>
- Xiao, Y., Lu, M., & Fu, Z. (2020). Covered Face Recognition Based on Deep Convolution Generative Adversarial Networks. *Artificial Intelligence and Security*, 133–141. https://doi.org/10.1007/978-3-030-57884-8_12
- Xu, X., & Wang, X. (2022). Human face cartoon image generation based on CycleGAN. *5th International Conference on Computer Information Science and Application Technology (CISAT 2022)*, 12451, 516–520. <https://doi.org/10.1117/12.2656569>
- Yu, C., & Pei, H. (2021). Face recognition framework based on effective computing and adversarial neural network and its implementation in machine vision for social robots. *Computers & Electrical Engineering*, 92, 107128. <https://doi.org/10.1016/j.compeleceng.2021.107128>
- Zeng, J., He, X., Li, S., Wu, L., & Wang, J. (2022). Virtual Face Animation Generation Based on Conditional Generative Adversarial Networks. *2022 International Conference on Image Processing, Computer Vision and Machine Learning (ICICML)*, 580–583. <https://doi.org/10.1109/ICICML57342.2022.10009693>
- Zhang, P., Wang, Y., Kumar, N., Jiang, C., & Shi, G. (2022). A Security- and Privacy-Preserving Approach Based on Data Disturbance for Collaborative Edge Computing in Social IoT Systems. *IEEE Transactions on Computational Social Systems*, 9(1), 97–108. <https://doi.org/10.1109/TCSS.2021.3092746>