



Universidad Autónoma de Ciudad Juárez

**Instituto de Ingeniería y Tecnología
Departamento de Eléctrica y Computación
Maestría en Cómputo Aplicado**

Redes generativas adversarias (GAN) para la creación de imágenes con transformaciones en atributos faciales basadas en la edad de mujeres mexicanas desaparecidas

Tesis para obtener el grado de
Maestra en Cómputo Aplicado

Blanca Elena Cazares Salais

“Becada por el Consejo Nacional de Ciencia y Tecnología”

**Bajo la dirección de
Dr. Rogelio Florencia Juárez
Y la codirección de
Dr. Vicente García Jiménez**

Ciudad Juárez, Chihuahua 24 de mayo de 2024

Índice de contenido

Introducción	6
Capítulo I	8
1.1. Planteamiento del problema.....	9
1.2. Antecedentes	11
1.3. Objetivos	16
1.3.1. Objetivo general.....	16
1.3.2. Objetivos específicos	16
1.4. Justificación	17
1.5. Metodología	18
1.6. Impacto del proyecto.....	19
1.8. Alcances y limitaciones.....	19
Capítulo II. Marco teórico.....	21
2.1. Desaparición de mujeres en México	22
2.2. Situación de desaparición de personas en México	23
2.3. Redes Generativas Adversarias (GAN).....	24
2.3.1. Preparación de los datos de entrada.....	25
2.3.2. Modelado	27
2.3.3. Evaluación.....	27
2.4. Transformaciones en imágenes	29
2.4.1. Transformaciones basadas en manipulaciones básicas de la imagen	29
Capítulo III. Generación de rostros sintéticos.....	32
3.1. Conjunto de imágenes para generación de rostros	34
3.2. Arquitectura GAN para generación de rostros	35
Capítulo IV. Propuesta de solución	38
4.1. Comprensión del negocio y los datos.....	39
4.2. Preparación de los datos.....	41
4.3. Modelado.....	44
4.3.1. Envejecimiento de rostros con Red Generativa Adversaria Condicional (<i>conditional Generative Adversarial Network, cGAN</i>)	44
4.3.2. Envejecimiento en rostros con modelo StyleGAN-based Age Manipulation (SAM). 46	

4.4. Evaluación.....	50
4.5. Despliegue.....	53
Capítulo V. Resultados y discusiones.....	56
5.1. Modelo cGAN.....	57
5.2. Modelo pre entrenado basado en <i>StyleGAN</i>	59
Capítulo VI. Conclusiones y trabajo futuro.....	73
6.1. Conclusiones.....	74
6.2. Trabajos a futuro.....	75
Referencias.....	77
Anexos.....	80

Índice de figuras

Figura 1. Distribución porcentual de niñas, niños y adolescentes reportados como extraviados, desaparecidos o no localizados, según sexo y grupo de edad (2010-2017)	17
Figura 2. Edad de hombres y mujeres desaparecidos hasta abril de 2018 [23]	22
Figura 3. Personas desaparecidas y no localizadas en México [5]	23
Figura 4. Personas desaparecidas y no localizadas por entidad federativa [5]	23
Figura 5. Ejemplos de imágenes generadas con GAN [25]	24
Figura 6. Transformaciones geométricas [4]	30
Figura 7. Transformaciones fotométricas [4].....	31
Figura 8. Arquitectura GAN para generación de rostros	33
Figura 9. Conjunto de imágenes CelebA [27]	35
Figura 10. Arquitectura GAN	37
Figura 11. Imágenes resultantes de la GAN para generar rostros artificiales	37
Figura 12. Conjunto de imágenes IMDB-Wiki.....	40
Figura 13. Muestras de los metadatos contenidos en el conjunto UTKFaces.....	42
Figura 14. Gráfico de barras de distribución por género de UTKFaces	42
Figura 15. Distribución de edad en UTKFaces.....	43
Figura 16. Alineación y recorte	44
Figura 17. cGAN para envejecimiento en rostros.....	46
Figura 18. Arquitectura de modelo SAM	47
Figura 19. Bucket Amazon S3	52
Figura 20. Entrenamiento cGAN	58
Figura 21. Índice de similitud, media y mediana por grupo etario	70
Figura 22. Índice de similitud, media, mediana y desviación estándar por grupo etario	72

Índice de tablas

Tabla 1. Análisis comparativo sobre literatura DA.....	15
Tabla 2. Etiquetas de clase.....	45
Tabla 3. Google colab, T4 GPU, RAM acelerado	57
Tabla 4. Google colab pro, V100 GPU, RAM alta capacidad.	58
Tabla 5. Resultados mujer 43 años	60
Tabla 6. Resultados mujer 47 años	63
Tabla 7. Resultados adolescente 14 años	66
Tabla 8. Estadística descriptiva de similitud por grupo etario	69

Introducción

La desaparición de personas es una problemática crítica en muchas regiones del mundo, incluido México. Según cifras estadísticas publicadas por organizaciones gubernamentales [1], [2], la mayoría de los casos pertenecientes a mujeres, son relacionados a crímenes de perspectiva de género. En dichos casos, sus destinos van directamente al crimen organizado, donde son mantenidas con vida, privadas de su libertad y forzadas a trabajos de índole sexual. Pensando en esta situación, se da búsqueda por parte de autoridades por meses, incluso años. En dicho tiempo, muy seguramente la apariencia de la mujer se ha visto modificada por la edad. Enfrentarse a este problema es un desafío para las autoridades y familias afectadas, que requieren herramientas eficaces para la generación de nuevos retratos con las características actualizadas, con la finalidad de seguir dando difusión al caso, con la esperanza de encontrar a la persona. Una de las estrategias más prometedoras en el contexto generación de imágenes, es el uso de herramientas de Inteligencia Artificial (*Artificial Intelligence*, AI).

Las Redes Generativas Adversarias (*Generative Adversarial Network*, GAN) [3] han emergido como una de las técnicas más poderosas en el campo de la generación de imágenes [4]. Estas son capaces de aprender y replicar distribuciones complejas en datos, permitiendo la generación de imágenes con modificaciones de la edad realistas, sin requerir conjuntos de imágenes que reflejen distintas etapas etarias de la misma persona.

El presente trabajo de investigación está enfocado en la implementación de modelos GAN que generan rostros y cambios en la edad. La motivación viene dada por el área de oportunidad, donde se puede aplicar IA para expandir el conocimiento científico y tecnológico del país.

El objetivo principal de este estudio es la implementación de un modelo GAN que sea capaz de generar imágenes realizando transformaciones faciales relacionadas con rasgos de la edad a partir de fotografías de mujeres mexicanas para simular cómo podría lucir su rostro en años posteriores.

Se implementa como primer acercamiento un modelo GAN para generación de rostros artificiales. Se continúa con una cGAN que realiza envejecimiento en rostros, pero se obtienen los resultados finales con el modelo SAM (*Style-based Age Manipulation*), con el que se resaltan tres grupos principales con la evaluación realizada. Esta incluye la utilización de la herramienta en la nube

Amazon Rekognition y su función '*compare_faces*' para medir la similitud entre las imágenes generadas y la original de un rostro. Esta evaluación cuantitativa permite determinar si las transformaciones de envejecimiento preservan la identidad de la persona.

En el capítulo I se establece el planteamiento general que comprende de antecedentes, objetivos, justificación, metodología, impacto, así como sus alcances y limitaciones. Posteriormente en el capítulo II se desarrollan los temas que pertenecientes a la problemática y su solución propuesta. Iniciando con la teoría sobre la desaparición de mujeres en México, después un contexto general sobre la situación de desaparición de personas. Se da un análisis del funcionamiento de las GAN y se finaliza con las transformaciones en imágenes que se han realizado a través de los años mediante diversas técnicas.

En el capítulo III se despliega la implementación de un modelo de generación de rostros sintéticos, explicando su arquitectura y su funcionamiento de manera general. Se procede con el capítulo IV, donde se presentan formalmente los dos modelos aplicados para el envejecimiento en rostros. Más adelante, en el capítulo V se dan resultados y discusiones del modelo que generó los envejecimientos faciales. Se finaliza con el capítulo VI donde se presentan las conclusiones y trabajo futuro.

Capítulo I

En este capítulo se introduce la problemática que guía la investigación, antecedentes que presentan algunos trabajos relacionados a la temática propuesta, objetivo general y específicos a cumplir al culminar el proyecto y la justificación que motivó el desarrollo del tema, además se presenta su impacto.

1.1. Planteamiento del problema

En México, desde el año 1964 hasta la fecha, se han reportado aproximadamente 250,000 personas desaparecidas. Hasta 2022, más de 87,000 de estos casos permanecen sin resolverse, lo que representa cerca del 35% del total. De este porcentaje, se estima que alrededor de 21,000 personas corresponden al sexo femenino, es decir casi un 25% [5].

La Comisión Nacional de Búsqueda (CNB) en México, ha destacado que la desaparición forzada y la cometida por particulares impactan de manera diferenciada a mujeres y niñas. En respuesta a esta situación, se han incorporado procesos específicos que adoptan una perspectiva de género al abordar estos casos [6].

Desde 1962 hasta el 15 de junio de 2023, se han registrado 26,764 casos de niñas, adolescentes y mujeres desaparecidas, de los cuales 22,162 corresponden al período desde 2007 hasta la actualidad. Es relevante señalar que la mayoría de las mujeres desaparecidas se encuentran en el rango de edad de 15 a 19 años, rango que también se observa en las mujeres que han sido localizadas [7].

La cifra de personas desaparecidas en México aumenta día tras día. Es habitual observar un incremento en la tasa de desapariciones, dado que se registran aproximadamente 100 nuevos reportes diariamente ante las autoridades [5]. Ante esta creciente problemática, se considera necesario adoptar medidas y utilizar herramientas que ayuden a enfrentar este tipo de delitos que, al estar catalogados bajo una perspectiva de género, comúnmente están asociados con fenómenos como el reclutamiento forzado por parte de la delincuencia organizada, diversas formas de violencia de género (incluyendo violencia sexual, delitos sexuales, feminicidios o transfeminicidios), así como tráfico humano y trata de personas con fines de explotación, secuestro, desaparición forzada o por particulares [6].

Ante el caso de desaparición de una persona, la Secretaría de Gobernación (SEGOB) en conjunto con la CNB ha implementado un método tecnológico para facilitar la presentación de denuncias. Este método consiste en el uso de un formulario electrónico, disponible en el sitio web mencionado en [8] bajo la sección “Inicio de Reporte de Persona Desaparecida o No Localizada”. En este se solicitan los datos generales, tales como nombre, apellidos, nacionalidad, estado civil, escolaridad,

entre otros. Además de lo anterior, se solicita una descripción breve pero detallada de los hechos, incluyendo fecha, hora y circunstancias de la desaparición.

De igual manera, es posible presentar una denuncia en cualquier fiscalía. En dicho caso, la persona que la realiza debe proporcionar la misma información requerida en el formulario electrónico. Además, es altamente recomendable llevar una fotografía reciente y de buena calidad de aquella persona reportada como desaparecida. Este material visual tiene como objetivo facilitar la difusión nacional de la información a través de diversos medios de comunicación. Así, en el evento en que la persona desaparecida sea vista, se pueda notificar a las autoridades pertinentes para proceder con su localización.

A pesar de la inclusión de fotografías en los reportes de personas desaparecidas, la apariencia del individuo se ve alterada con el paso del tiempo si no es localizado. Para abordar estos casos, se utilizan retratos hablados. Estos son bosquejos que proyectan cómo podría lucir el rostro después de cierto tiempo, elaborados a partir de las fotografías recientes. Para su creación se requiere la intervención de especialistas conocidos como morfólogos, ya que su ocupación gira en torno a la morfología facial. Para ello, estos deben completar cursos avanzados de antropología física para la reconstrucción gráfica y plástica, además de estudios en anatomía, osteología y dibujo, orientados hacia la disciplina forense. Sin embargo, a pesar de su formación, la generación de este tipo de retratos implica una complejidad considerable, así como una inversión significativa de tiempo y recursos. Para optimizar esto, se explora la utilización de un modelo de aprendizaje profundo (*Deep Learning*, DL) para automatizar y acelerar el proceso.

Bajo el contexto anterior, se propone la implementación de un modelo GAN para transformar fotografías faciales y proyectar su apariencia a diferentes edades. Este cuenta con la capacidad de tomar como entrada una fotografía facial para realizar las transformaciones pertinentes en diversos grupos etarios. Dando como salida conjuntos de imágenes envejecidas correspondientes a la imagen de entrada, las cuales deben ser sometidas a un proceso de evaluación donde se implementará una herramienta de análisis de imágenes basado en la nube que utiliza tecnologías de DL y visión por computadora [9].

La AI es un campo de estudio que tiene por objetivo comprender y simular la capacidad de aprender que tenemos los seres humanos. Esta ha comenzado a proveer importantes aportaciones en años recientes en el área de la seguridad pública, ya que posee herramientas muy poderosas

para reconocimiento de patrones y extracción de características a partir de entradas que a simple vista parecieran contener información irrelevante ya que son imperceptibles por los humanos. Este tipo de estudios no han sido enfocados únicamente al análisis de datos, como estadísticos y predicciones en comercios, sino que también se hace uso de estos para reconocimiento de voz o huellas dactilares, análisis de texto y por supuesto, reconocimiento facial [10]. En el área de la seguridad pública tienen gran impacto, por ejemplo, al buscar identificar a delincuentes con fotografías faciales captadas por cámaras de seguridad.

Hasta el momento no se tiene conocimiento de algún proyecto en México que plantee similar temática a la propuesta. A continuación, se presentan los antecedentes sobre las técnicas ya estudiadas en la generación de imágenes artificiales, un proceso también conocido como DA.

1.2. Antecedentes

A continuación, se presentan y se describen brevemente algunos trabajos previos relacionados con el presente proyecto que han sido identificados en la literatura.

En [11] se analiza cómo el avance del reconocimiento facial ha sido impulsado significativamente por la expansión de los conjuntos de datos de entrenamiento, que ahora incluyen millones de imágenes faciales descargadas y etiquetadas para la identificación. Este estudio cuestiona la necesidad de acumular un volumen tan extenso de imágenes y propone alternativas más accesibles y económicas para ampliar los conjuntos de entrenamiento, como la generación de datos artificiales. Además, se describen métodos para enriquecer las colecciones originales mediante variaciones significativas en la apariencia facial, manipulando las características de los rostros en las imágenes. Los resultados del estudio indican que la incorporación de imágenes sintéticas generadas durante los experimentos, utilizando herramientas de modelado facial 3D para modificar la pose, forma y expresiones, mejora la precisión de los modelos de reconocimiento facial comparativamente con los estándares del estado del arte. No obstante, el trabajo también señala limitaciones de este tipo de técnicas, como la incapacidad para controlar variaciones en la oclusión y calidad de la imagen, entre otros factores que pueden afectar el desempeño del modelo.

En [12] se expone que las Redes Neuronales Convolucionales Profundas (*Deep Convolutional Neural Network*, DCNN) han demostrado un alto desempeño en el reconocimiento facial. El proceso de entrenamiento estas redes requieren una gran cantidad de datos etiquetados, cuya

obtención es costosa y consume demasiado tiempo. En respuesta a esto, el artículo propone cinco métodos para la generación de datos artificiales dedicados específicamente a imágenes faciales, ajustando elementos como la posición, el estilo de cabello, el uso de lentes, las poses y la iluminación. Esta necesidad surge debido a que las transformaciones tradicionales de DA, tales como el recorte, cambio de iluminación y efecto de difuminación, entre otros, no son suficientes para enriquecer las imágenes faciales. Los métodos propuestos, enfocados en la apariencia física, han incrementado eficazmente el conjunto de entrenamiento, como lo demuestran los resultados experimentales. Evaluaron la precisión del modelo en la tarea de reconocimiento, con y sin el uso del método de DA propuesto, confirmando que el aumento de datos reduce la problemática de identificación de rostros ante cambios de apariencia. Finalmente, se realizó una comparación con otros métodos de generación de datos artificiales encontrados en la literatura, aunque no descritos detalladamente en el documento, enfocándose en atributos de imágenes faciales. Los resultados evidenciaron una mayor precisión en las predicciones bajo diversas condiciones, destacando un alto nivel de confiabilidad.

En [4] se afirma que la calidad y el tamaño de un conjunto de datos de entrenamiento influyen significativamente en los resultados de tareas relacionadas al reconocimiento facial que emplean métodos de DL. El documento también aborda la complejidad y el alto costo asociado con la recolección y etiquetado de muestras de alta calidad, y bien balanceados, destacando el uso de técnicas de DA como una solución para enriquecer los conjuntos de datos. Adicionalmente, el artículo revisa trabajos existentes sobre la generación de datos artificiales orientada a rostros, examinando los tipos de transformaciones y métodos utilizados. Se presentan los principios subyacentes a estas investigaciones y se realiza una comparación entre ellas. Se concluye planteando los desafíos y oportunidades asociadas con estos métodos de DA.

En [13], se analiza un estudio reciente sobre GAN y sus efectivas aplicaciones, centrándose en una extensión de esta tecnología conocida como Redes Generativas Antagónicas Condicionales (*Conditional Generative Adversarial Network*, cGAN). El propósito es enriquecer el modelo cGAN mediante la integración de un codificador que optimiza el tratamiento de imágenes reales. Para este desarrollo se combinó una cGAN con un codificador, facilitando la transformación de imágenes a un espacio caracterizado por múltiples atributos. Utilizando el modelo *Invertible conditional Generative Adversarial Network* (IcGAN), se experimentó con los conjuntos de datos

públicos MNIST, que comprende cerca de 70 mil imágenes de dígitos manuscritos para entrenamiento y 10 mil destinadas para pruebas, y CelebA, que incluye más de 200 mil imágenes faciales de celebridades con sus respectivas etiquetas. Esto permitió la ejecución de ediciones realistas modificando los datos de entrenamiento. Además, se añadió un codificador a la estructura de la GAN para comprimir una imagen real en un espacio latente y un vector condicional, mejorando significativamente la precisión de las visualizaciones generadas hasta alcanzar un 86%. La integración del vector y en la primera capa convolucional del discriminador y en la entrada del generador maximizó la precisión. Este estudio priorizó la calidad visual de las muestras reconstruidas como un criterio de evaluación y realizó comparaciones cuantitativas de diferentes configuraciones usando la mínima función de pérdida, revelando la eficacia de la red en base a los valores promedio obtenidos del conjunto de entrenamiento. El experimento concluyó con la implementación del modelo IcGAN, que superó las limitaciones de las GAN permitiendo un control detallado en la manipulación de atributos complejos mediante información condicional. Los resultados con CelebA fueron evaluados como satisfactorios y prometedores.

En [14] se examina la problemática asociada a los sistemas de reconocimiento facial al tratar con imágenes de rostros que utilizan productos cosméticos que ocultan características distintivas como cicatrices o tatuajes. Para abordar estas variaciones, se han investigado diversos métodos incluyendo la eliminación de distorsiones transferencia de aprendizaje y DA. Este estudio se enfoca en la creación de estilos de maquillaje de celebridades, con el objetivo de entrenar a las DCNN para reconocer características robustas que puedan diferenciar las variaciones cosméticas, y así poder desempeñar sus tareas específicas, incluso cuando el conjunto de datos disponible no sea extenso. Se explora la aplicación de seis estilos distintos de maquillaje utilizando taaz, una herramienta pública de síntesis de maquillaje. Además, se incrementa el volumen de datos mediante técnicas como destexturizar, decolorar y mapeo de bordes, generando así nueve variantes de cada imagen original, lo que resulta en un conjunto de datos considerablemente aumentado. Se evaluó el rendimiento del enfoque propuesto mediante cinco experimentos aplicados en dos conjuntos de datos estándar: *Youtube Makeup Database (YMU)* y *Virtual Makeup Database (VMU)*. Estos incluyeron el entrenamiento de una *Convolutional Neural Network (CNN)* tanto con la técnica propuesta como sin ella, evaluando en ambos casos su nivel de precisión. El propósito de estas pruebas fue comparar los resultados obtenidos con el enfoque propuesto frente a otros métodos eficaces en DA, como las GAN. Los resultados revelaron un aumento

significativo de la precisión, alcanzando un 90.04% en YMU y un 92.99% en VMU con la estrategia de aprendizaje propuesta, demostrando que esta técnica puede competir efectivamente con las demás metodologías evaluadas.

En [15] se analizan los desafíos enfrentados en la clasificación facial cuando se utilizan modelos como *Support Vector Machines* (SVM) y *Softmax*, especialmente en escenarios donde se dispone de pocas imágenes de un mismo sujeto, resultando en un rendimiento subóptimo. Para abordar esta limitación, se propone un método de generación de muestras virtuales, basado en análisis facial con conjuntos Bayesianos, junto con un algoritmo diseñado para reducir la duración del proceso. Se llevaron a cabo experimentos utilizando características de *Local Binary Pattern* (LBP) de alta dimensión y características extraídas de una CNN desarrollada por los autores. Los experimentos iniciales, bajo el protocolo sin restricciones del conjunto de datos *Labeled Faces in the Wild* (LFW), no superaron en eficacia a los algoritmos de alta precisión, alcanzando una precisión de 94.21% en LFW. No obstante, en términos de velocidad, el método propuesto fue más eficiente, procesando imágenes a una velocidad promedio de 100 milisegundos (ms) por imagen en una *Central Processing Unit* (CPU), en comparación con 240 ms por imagen del algoritmo *High dimensional LBP features* (HighDimLBP). Sin embargo, con conjuntos de datos más simples y utilizando una sola muestra de cada clase para el entrenamiento, el método alcanzó solo un 60.20% de precisión, indicando un rendimiento limitado en escenarios de datos restringidos.

La presente revisión de la literatura revela que las GAN han demostrado recientemente su eficacia en la generación de imágenes artificiales de alta calidad. Dada esta capacidad, se considera que, con un alto espectro de algoritmos, ofrecerán resultados prometedores en el marco del proyecto descrito en este documento.

En **Tabla I** se presenta un resumen de los trabajos relacionados en el contexto del presente tema de tesis, generación de imágenes artificiales. Se presentan investigaciones previas sobre DA y técnicas de procesamiento de imágenes, detallando los objetivos, transformaciones aplicadas y métodos utilizados. Cada entrada destaca los enfoques que contribuyen al campo, desde cambios de pose y expresión facial hasta la síntesis de imágenes con diferentes efectos cosméticos y estilos. La presente investigación se distingue por su enfoque hacia la generación de imágenes con transformaciones en la edad mediante un modelo GAN que genera envejecimiento a partir de una fotografía facial.

Tabla 1. Análisis comparativo sobre literatura DA

Autores	Año	Objetivo	Transformaciones	Método
I. Masi, A. T. Trán, T. Hassner, J. T. Leksut y G. Medioni	2016	Analizar DA dentro del dominio correspondiente para aumentar el conjunto de imágenes.	<ul style="list-style-type: none"> • Cambio de pose • Cambio de forma del rostro • Cambio de expresión 	Modelo 3D
G. Perarnau, J. van de Weijer, B. Raducanu, y J. M. Álvarez	2016	Implementar un modelo llamado IcGAN, conformado por una GAN y un codificador.	<ul style="list-style-type: none"> • Mejora de calidad de la imagen 	IcGAN
B. Leng, K. Yu y J. QIN	2017	Aumentar conjuntos de imágenes mediante muestras virtuales y reducir el tiempo de ejecución del algoritmo	Imágenes virtuales a partir de características del espacio	CNN
J. J. Lv, X. H. Shao, J. S. Huang, X. D. Zhou y X. Zhou	2017	Aplicar 5 técnicas de DA dedicadas a imágenes para intentar enriquecer y mejorar la calidad de un conjunto de entrenamiento.	<ul style="list-style-type: none"> • Perturbación de puntos en el rostro • Síntesis de estilo de cabello • Síntesis de anteojos • Síntesis de pose • Síntesis de iluminación 	Modelo 3D con procesamiento de imagen
M. Sajid	2018	Evaluar la capacidad de discriminación de las DCNN cuando las imágenes faciales tienen	<ul style="list-style-type: none"> • Síntesis de maquillaje 	Visión por computadora

		distintos efectos cosméticos		
X. Wang, K. Wang, y S. Lian	2020	Hacer revisión sobre los trabajos existentes de aumento de conjuntos de imágenes.	• -----	-----
<i>Este trabajo</i>	2024	Aplicar un modelo GAN a fotografías faciales de mujeres mexicanas para generar nuevas imágenes conforme a transformaciones en el rostro producidas por la edad.	• Aumento de edad	GAN

1.3. Objetivos

En la presente sección se describen tanto el objetivo general como los específicos.

1.3.1. Objetivo general

Implementar un modelo GAN que sea capaz de generar imágenes realizando transformaciones faciales relacionadas con rasgos de la edad a partir de fotografías de mujeres mexicanas para simular cómo podría lucir su rostro en años posteriores.

1.3.2. Objetivos específicos

- Analizar los tipos de modificaciones en las que se enfocan distintas variantes de GAN, encontradas en la literatura, que puedan dar mayor aportación en los resultados del proyecto.
- Realizar pruebas de programación de GAN para comprensión de los atributos y estructura de este tipo de redes.

- Entrenar las redes neuronales (*Neural Network*, NN) del generador y discriminador, ajustando sus parámetros de manera que produzca imágenes de mayor calidad al aplicar las transformaciones.
- Determinar la forma de evaluar si el rostro de las imágenes generadas por la GAN luce similar a aquellas reales de la persona en la edad simulada.

1.4. Justificación

La problemática de personas desaparecidas en México es una crisis humanitaria de gran magnitud, en esta investigación se estará abordando principalmente del género femenino. Esto debido a que dicho grupo presenta un impacto diferenciado en mujeres y niñas, presentando que 1 de cada 4 personas desaparecidas pertenece al género femenino [16]. En el caso de las personas desaparecidas menores de 18 años, este porcentaje aumenta a más de la mitad, alcanzando su punto más alto en el grupo de 12 a 18 años (Figura 1).

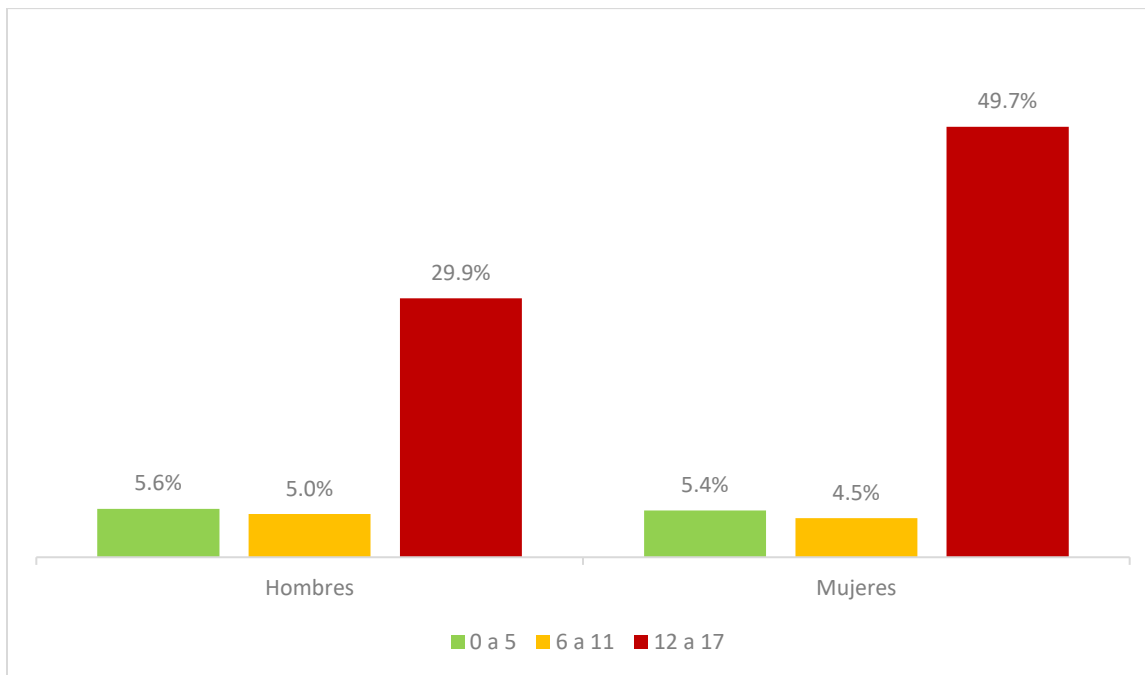


Figura 1. Distribución porcentual de niñas, niños y adolescentes reportados como extraviados, desaparecidos o no localizados, según sexo y grupo de edad (2010-2017)

A menudo, la búsqueda se ve obstaculizada por falta de imágenes actualizadas que reflejen la apariencia actual de la persona, considerando la progresión del tiempo desde su desaparición. Un ejemplo de esto se ve reflejado en las cifras presentadas en [1], donde se menciona que a finales del 2017 había casos que llevaban casi 8 años sin resolverse. Cabe mencionar que las personas, con el paso del tiempo, cambian su aspecto físico por naturaleza, por lo que, si una persona no es localizada de manera rápida, las probabilidades de ser encontrada se ven reducidas. Aquí es donde se plantea la posibilidad de creación de imágenes artificiales, basadas en la imagen original, de tal manera que disminuya la necesidad de bosquejos. En este contexto, el presente proyecto se justifica por la necesidad de incorporar herramientas tecnológicas de DL que den actualización y mejora de calidad a las imágenes disponibles para el apoyo en la localización de mujeres desaparecidas.

La implementación de GAN para el aumento de edad en rostros, enfocada específicamente en mujeres mexicanas desaparecidas, ofrece una alternativa de apoyo a este problema, ya que han demostrado ser altamente efectivas en la generación de imágenes realistas y en la modificación de atributos faciales en imágenes existentes. Al aplicar esta tecnología para simular envejecimiento se pretende proporcionar imágenes que representen de manera eficaz cómo podrían lucir estas personas en la actualidad.

Esta aproximación podría mejorar las tasas de éxito en la identificación y localización de mujeres desaparecidas, al igual que contribuir al campo académico y práctico del área de reconocimiento facial y de AI en el contexto mexicano. Además, el proyecto aborda una brecha en la aplicación de tecnologías de reconocimiento facial adaptadas a la diversidad demográfica de América Latina.

En conclusión, este proyecto tiene una aplicación directa en un problema social del día a día y hace uso de tecnologías avanzadas para la justicia y la equidad social, justificándose desde una perspectiva tecnológica y humanitaria.

1.5. Metodología

Para estructurar la investigación a lo largo del proyecto se siguieron diversas etapas como: comprensión del negocio y datos, preparación de los datos, modelado y evaluación. Estas incluyeron la definición de objetivos, recolección y verificación de los datos, así como su

selección y limpieza, generación de características, estandarización, desarrollo de modelos y evaluación.

1.6. Impacto del proyecto

Según los datos disponibles en el sitio oficial de la Secretaría de Gobierno [5], actualmente se reportan más de 90 mil personas desaparecidas en México, que aún no han sido localizadas. En el año 2020 solamente, se registraron 19,932 nuevos reportes de desaparición. Con el tiempo, las características físicas de estas personas desaparecidas habrán cambiado inevitablemente, complicando aún más los esfuerzos de localización, ya que las imágenes disponibles para su búsqueda pueden no reflejar su aspecto actual.

El delito de privación de la libertad, incluso sin violencia directa, presenta un área significativa de oportunidad para desarrollo de tecnologías que apoyen a diferentes instituciones dedicadas a este tipo de crímenes.

Actualmente, se tiene acceso a varios enfoques de AI que pueden utilizarse para facilitar diversos procesos sociales y culturales. El presente proyecto busca contribuir al Programa Nacional Estratégico (PRONACE) seguridad humana, relacionado con las violencias estructurales, mediante la investigación y aplicación de un modelo de DL, combinado con diversas herramientas tecnológicas, para apoyar con soluciones automáticas la situación de inseguridad en el país.

1.8. Alcances y limitaciones

El presente proyecto tiene delimitados los logros que se pretenden alcanzar mediante la implementación del modelo GAN.

- Implementación de un modelo GAN para transformar fotografías faciales y simular el envejecimiento en diversos grupos etarios.
- Uso de métrica cuantitativa para evaluar la similitud entre la imagen original y las que fueron generadas por el modelo.
- Asegurar la preservación de la identidad de la mujer en las fotografías.
- Documentación de fases del proyecto para plantear futuras investigaciones y desarrollos en el área.

A continuación, se detallan las restricciones y desafíos que afectan el desarrollo y resultados del proyecto.

- Disponibilidad y calidad de los datos.
- El modelo pre entrenado puede estar limitado a las características faciales específicas de su conjunto original de entrenamiento.
- Precisión del envejecimiento simulado puede variar dependiendo en factores individuales.
- Entrenar y generar imágenes con GAN es computacionalmente costoso, requiere hardware sofisticado y tiempos de procesamiento significativos.

Capítulo II. Marco teórico

En el presente capítulo se desarrolla la teoría sobre la problemática, soluciones basadas en tecnología y la explicación general del modelo seleccionado para trabajar a lo largo del proyecto.

2.1. Desaparición de mujeres en México

A partir de 2007, la tasa de desapariciones en México ha mostrado un notable incremento, según los datos del Registro Nacional de Datos de Personas Extraviadas o Desaparecidas (RNPED) de 2019. Este aumento ha sido más pronunciado entre los hombres en comparación con las mujeres. Sin embargo, en el caso de los menores de edad, se observa una tendencia contraria, y con un número significativamente mayor de desapariciones del género femenino en comparativa con el género masculino [17]. Esto puede visualizarse en la Figura 2. Edad de hombres y mujeres desaparecidos hasta abril de 2018 [23].

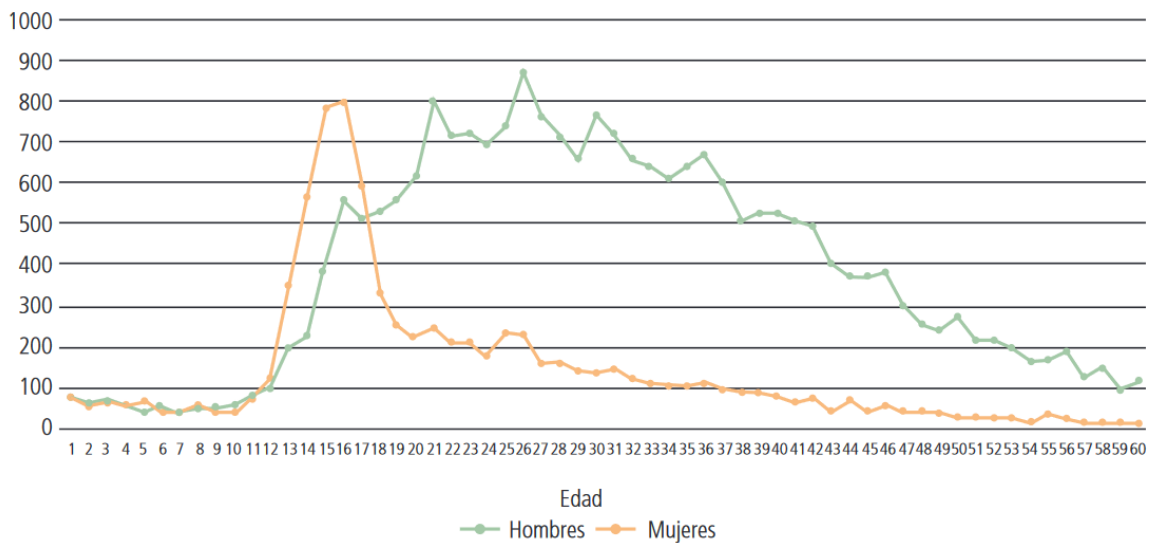


Figura 2. Edad de hombres y mujeres desaparecidos hasta abril de 2018 [23]

En los últimos años, diversas investigaciones y proyectos feministas han centrado sus esfuerzos en explorar las causas y los procedimientos para la localización de mujeres desaparecidas. A pesar de estos esfuerzos, persisten numerosos casos de mujeres jóvenes que aún no han sido encontradas, como se destaca en los datos proporcionados por las participantes en [18]. Este estudio se enfocó en analizar las situaciones de vulnerabilidad específicas de género que contribuyen a estas desapariciones.

2.2. Situación de desaparición de personas en México

La desaparición de personas en México ha sido un problema grave y persistente. Actualmente, en el país se reportan como desaparecidas más de 90,000 personas que aún no han podido ser localizadas, como se muestra en la Figura 3. Personas desaparecidas y no localizadas en México [5]. Siendo los estados de Tamaulipas, México y Jalisco aquellos donde existe mayor concentración de este fenómeno, como se puede apreciar en la Figura 4. Personas desaparecidas y no localizadas por entidad federativa [5].



Figura 3. Personas desaparecidas y no localizadas en México [5]

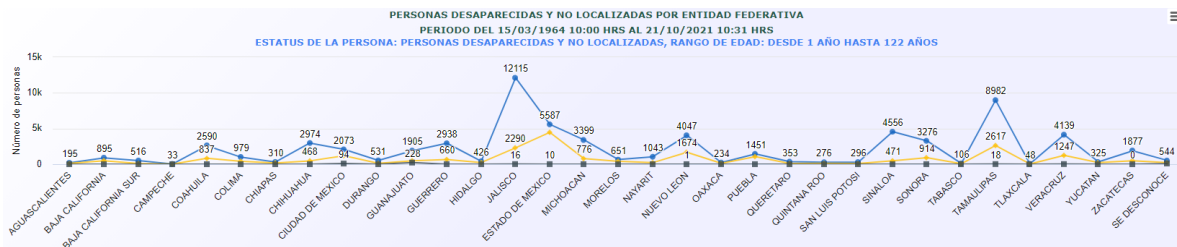


Figura 4. Personas desaparecidas y no localizadas por entidad federativa [5]

Se estima que, aproximadamente 200 reportes mensuales son pertenecientes al género femenino, de estos la mayoría se atribuyen a adolescentes del rango de edad 14-17 años [2].

A pesar de que el género masculino encabeza el estadístico de desapariciones, también constituyen la principal cantidad de personas localizadas. Mientras el femenino tiene una tasa muy baja según los datos sociodemográficos de este fenómeno.

La desaparición de mujeres en México es una problemática que ha aumentado en los últimos años. Hasta enero de 2023, un total de 22,571 mujeres fueron reportadas como desaparecidas. En ese mismo mes, desaparecieron 507 mujeres entre 14 y 60 años, lo que significa un promedio de, aproximadamente, 16 mujeres desaparecidas por día. Y, en el primer trimestre del mismo año se registró un aumento del 24% en desapariciones forzadas [2].

2.3. Redes Generativas Adversarias (GAN)

Las GAN [3] se basan en la teoría de juegos, en la que dos modelos de ML, generalmente compuestos por NN compiten entre sí. Uno de estos modelos, denominado Generador (G), aprende a crear muestras que imitan la distribución de un conjunto de datos proporcionado. Su principal objetivo es aprender la función $G(z)$ que transforma ruido z en datos realistas. El otro modelo, conocido como Discriminador (D), se encarga de examinar las muestras x para estimar si la entrada es un dato real o generado por G, actuando como un clasificador binario.

Las GAN se consideran un método para extraer información adicional de los datos. Aunque no son el único modelo generativo existente, han demostrado tener un alto nivel tanto en velocidad computacional como en la calidad de los resultados, como se muestra en la Figura 5. Gracias a su buen desempeño, son muy atractivas para tareas de DA, ya que pueden generar nuevos datos de entrenamiento que mejoran las respuestas de los modelos de clasificación.

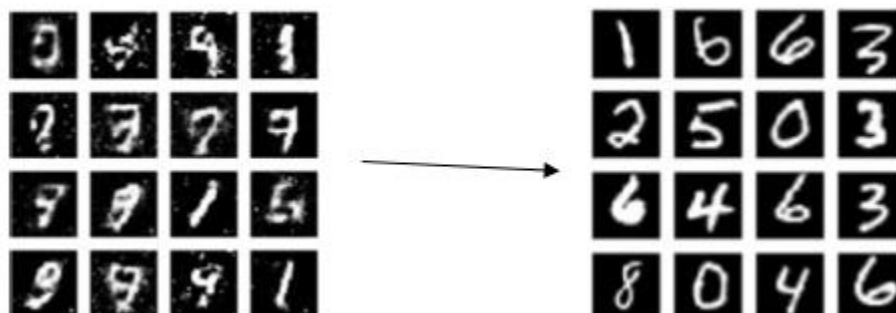


Figura 5. Ejemplos de imágenes generadas con GAN [25]

Los datos generados por la GAN dependen del conjunto de entrenamiento que se les proporcione, pero su funcionamiento es similar en todos los casos, con las redes intentando superarse continuamente. A medida que G mejora en la creación de datos convincentes, D debe esforzarse por distinguir entre datos falsos y reales.

Durante el entrenamiento de un clasificador, se mide su error con los conjuntos de entrenamiento y validación, y se detiene cuando su desempeño empeora para evitar un sobreajuste de los datos. En una GAN, los modelos tienen sus propios objetivos: mejorar mientras el otro empeora. Se dice que una GAN converge cuando alcanza el equilibrio *Nash*, un punto en el que ningún jugador puede mejorar su situación cambiando sus acciones. La GAN alcanza este equilibrio cuando:

- G produce datos falsos indistinguibles de los reales en el conjunto de entrenamiento.
- La mejor estrategia que sigue D para distinguir entre datos reales y falsos es adivinar al azar.

En la práctica, alcanzar este equilibrio es considerado imposible, y es uno de los principales desafíos en la investigación de GAN. A pesar de la ausencia de rigurosos procedimientos matemáticos que garanticen resultados óptimos, las GAN han demostrado un desempeño de muy buena calidad. Para su implementación fue necesario aplicar una metodología diseñada para proyectos de aprendizaje automático. Las tareas se detallan a continuación.

2.3.1. Preparación de los datos de entrada

En esta parte, se llevan a cabo diversas actividades esenciales para acondicionar los datos antes de su utilización en el entrenamiento del modelo. Esta etapa incluye la definición precisa de la dimensionalidad de los datos de entrada, así se asegura que cualquier nuevo conjunto de datos destinado a la predicción sea compatible y pueda integrarse con la estructura del modelo entrenado. La preparación de los datos es primordial para proceder a la etapa de modelado. Cabe aclarar que no es estática, se pueden repetir tareas anteriores tantas veces como sea necesario.

Selección: Al descartar o seleccionar muestras, es imperativo documentar y fundamentar cada decisión en criterios de calidad rigurosamente definidos para alcanzar los objetivos establecidos. Ante la presencia de clases desbalanceadas se podrán implementar diversas estrategias para optimizar la distribución, en caso de considerarse necesario. El conjunto de imágenes

seleccionado debe contener diversidad significativa, es decir, que se incluyan rostros de personas con una variedad de características faciales y poses.

Limpieza: En esta fase se concentra la remoción de cualquier ruido presente en el conjunto o bien, imputación de datos para tratar con valores especiales o campos nulos. Al trabajar con un conjunto de imágenes público, es fundamental considerar varios aspectos para asegurar que este sea adecuado para los objetivos del proyecto. Es esencial analizar meticulosamente los metadatos de las imágenes para realizar una selección prioritaria basada en criterios relevantes para el proyecto, tales como el género y la etnia, en este trabajo. Además, cada imagen debe ser inspeccionada para confirmar la presencia de un rostro en cada fotografía, lo que contribuye directamente al entrenamiento efectivo del modelo. Es necesario eliminar imágenes que puedan introducir ruido, como aquellas con múltiples rostros, para mantener la calidad y la consistencia del conjunto de datos.

Estandarización: La estandarización es un requisito esencial para muchas herramientas de ML, que a menudo requieren que los datos estén en un formato específico, como valores separados por comas (*comma separated values*, CSV). Además, la normalización de imágenes es un paso crítico en el preprocesamiento de datos. Los procedimientos comúnmente incluyen [19]:

- **Escalado de tamaño:** Las imágenes generalmente se redimensionan a tamaños estándar, tales como 64x64, 128x128, 256x256 píxeles, para facilitar el procesamiento.
- **Normalización:** Los valores de píxeles, que típicamente oscilan entre 0 y 255 en imágenes a color, se normalizan a una escala, comúnmente entre -1 y 1 o 0 y 1, lo cual facilita la convergencia del modelo durante el entrenamiento.
- **Corte o *padding*:** Para imágenes que no se ajustan al tamaño deseado después de redimensionar, se suele hacer cortes o agregar bordes para alcanzar el tamaño necesario sin distorsiona la relación de aspecto original.
- **Aumento de datos (DA):** Aunque no es estrictamente parte de la estandarización, DA puede mejorar la robustez del modelo al incrementar la diversidad del conjunto de datos.
- **Centrado:** Se asegura que el sujeto principal de la imagen esté centrado, eliminando información irrelevante alrededor, lo cual puede enfocar mejor al modelo en las características relevantes.

- **Normalización por lotes (*Batch normalization*):** Esta técnica puede aplicarse para estabilizar el aprendizaje y acelerar la convergencia.
- **Consistencia de canales:** Asegurar que los canales de color sean consistentes en términos de codificación y tratamiento a lo largo de todo el conjunto es vital para mantener la uniformidad en el aprendizaje.

2.3.2. Modelado

La selección de técnicas de modelado en ML se determinará directamente en función de los objetivos de negocio establecidos. El objetivo principal es desarrollar uno o varios modelos que cumplan con las restricciones y satisfagan los requerimientos del proyecto.

Una buena práctica es iniciar el proceso de modelado utilizando técnicas de baja complejidad. Este enfoque proporciona una base sobre la cual se puede incrementar gradualmente la capacidad del modelo, lo cual ayuda a minimizar el riesgo de desarrollar un modelo innecesariamente complejo.

Debido a la alta calidad de las imágenes generadas por las GAN, han sido seleccionadas como la técnica de modelado principal para este proyecto.

2.3.3. Evaluación

Los modelos GAN son particularmente desafiantes para evaluar debido a la naturaleza de cómo generan datos. Tradicionalmente, los métodos de evaluación para modelos de aprendizaje supervisado, como precisión, *recall* y *F1-score*, no se aplican directamente a las GAN porque no se trata de tareas de clasificación o regresión convencionales. Por lo tanto, se han desarrollado algunas métricas cuantitativas específicas para evaluar la calidad y la diversidad de las imágenes generadas por las GAN. Algunas de las más utilizadas se mencionan a continuación:

Métricas cuantitativas

Estas métricas proporcionan una forma de evaluar objetivamente los resultados de la GAN desde distintas perspectivas: calidad de imagen (FDI, KID), diversidad de las imágenes generadas (IS), comportamiento del modelo en el espacio latente (PPL), y balance entre precisión y *recall*. Seleccionar las métricas adecuadas depende de los objetivos específicos del proyecto y las características del conjunto de datos seleccionado.

- ***Inception Score (IS)***: Utiliza un modelo de red neuronal pre entrenado para evaluar las imágenes generadas. El puntaje se basa en dos criterios, el primero es la variedad y el segundo es la claridad, es decir, que cada imagen generada debe ser claramente identificable o tener alta confianza en las predicciones de clases [20].
- ***Distancia de Inicio de Fréchet (Fréchet Inception Distance, FDI)***: Se mide la similitud entre las distribuciones de las imágenes generadas y las reales, utilizando las características extraídas por una red *Inception* pre entrenada. Un valor bajo indicaría que las imágenes generadas son más similares a las reales, lo cual es indicativo de un mejor modelo. Esta técnica es actualmente una de las más populares y confiables para evaluar la calidad de las imágenes generadas por GAN [21].
- ***Perceptual Path Length (PPL)***: Utilizada especialmente en modelos como *Style-Based Generative Network* (StyleGAN), ya que mide la suavidad de la transición entre imágenes generadas mientras se mueve en el espacio latente del modelo. Una longitud de trayectoria perceptual menor sugiere que pequeños cambios en el espacio latente resultan en cambios graduales y predecibles en el espacio de la imagen, lo que es un indicador de un modelo generativo más estable [22].
- ***Precision and Recall***: En el contexto de las GAN, estas métricas han sido adaptadas para evaluar la precisión, es decir, qué porcentaje de las imágenes generadas son realistas. Y el *recall*, el cual indica qué tan bien las imágenes generadas cubren el espectro de las imágenes reales. De hecho, estas métricas pueden ayudar a evaluar si una GAN está sobreajustando o subajustando [22].
- ***Kernel Inception Distance (KID)***: Este es una mejora de FDI, midiendo la distancia entre las distribuciones de las imágenes reales y las generadas, utilizando un *kernel* en el espacio de características del modelo. A diferencia del FDI, KID tiene propiedades estadísticas mejoradas y computacionalmente es más eficiente [23].

Métricas cualitativas

Las evaluaciones cualitativas son esenciales para entender el desempeño de un modelo GAN, ya que aportan una dimensión de análisis que complementa a las métricas cuantitativas y permite captar la calidad perceptual de las imágenes generadas. Combinar evaluaciones cualitativas y cuantitativas proporciona una visión más completa del rendimiento del modelo.

- **Evaluación visual de expertos:** Esta puede considerarse la forma más directa de evaluación cualitativa mediante la inspección visual de las imágenes generadas. Los expertos del dominio o evaluadores entrenados revisan las imágenes generadas para determinar su realismo, calidad y la presencia de artefactos visuales. Esta evaluación puede ser estructurada en torno a criterios específicos como textura, color y relevancia del contexto.
- **Estudios de usuarios:** Estos se efectúan presentando las imágenes generadas a un grupo de personas, no necesariamente expertos, y recopilar sus opiniones.
- **Prueba tipo Turing:** En una prueba a ciegas, a los evaluadores se les muestra una serie de imágenes sin decirles cuáles son generadas por la máquina y cuáles son reales. Después, se les pide que identifiquen si la imagen es real o generada por una máquina. Un modelo GAN adecuadamente entrenado debería ser capaz de engañar a los evaluadores con frecuencia, indicando un alto nivel de realismo.

Las métricas descritas anteriormente no agotan todas las posibles evaluaciones que se pueden aplicar para verificar el desempeño de una GAN; sin embargo, representan las más comunes y son frecuentemente utilizadas en la práctica.

2.4. Transformaciones en imágenes

Las DCNN han mostrado un desempeño sobresaliente en diversas tareas de visión por computadora. No obstante, un desafío importante radica en la necesidad de grandes conjuntos de datos para su entrenamiento. Aquí es donde DA desempeña un papel crucial, proporcionando diversas técnicas para aumentar tanto el tamaño como la calidad de los conjuntos de datos, permitiendo la construcción de modelos de DL más precisos.

Además, estas técnicas tienen aplicaciones en la mejora de conjuntos de imágenes a través de diferentes enfoques de aumento de datos, descritos a continuación.

2.4.1. Transformaciones basadas en manipulaciones básicas de la imagen

Este tipo de transformaciones consiste en modificaciones simples como las siguientes:

2.4.1.1. Transformaciones geométricas

En esta sección se describen formas de aumentar los conjuntos de datos mediante técnicas geométricas, ejemplificadas en la Figura 6.



Figura 6. Transformaciones geométricas [4]

- **Voltear la imagen (flipping):** Esta técnica permite voltear la imagen original tanto en el eje X como el eje Y, siendo el primero el más comúnmente aplicado. Es una de las transformaciones más simples y eficientes, especialmente en conjuntos de imágenes como CIFAR-10 e *ImageNet*, que incluyen miles de fotografías de objetos y animales. Sin embargo, no es adecuada para tareas de reconocimiento de texto.
- **Color del espacio:** Modificar los canales de color de una imagen es una estrategia práctica a implementar. Las transformaciones de color muy simples intentan aislar un solo canal de color como R, G o B. Además, los valores RGB (*Red, Green, Blue*) pueden ser manipulados mediante operaciones con matrices para incrementar o disminuir el brillo.
- **Cortar:** Consiste en cortar partes de la imagen, ya sea del centro o al azar y de distintos tamaños. Hay que tomar en cuenta que con este tipo de modificación se reduce el tamaño de la entrada, por ejemplo (256, 256) a (224, 224).
- **Rotar:** El aumento mediante rotación se realiza hacia la izquierda o derecha en uno de los ejes, con una magnitud de 1° a 359° . Su eficiencia está fuertemente relacionada a su parámetro de grados al girar la imagen. Algunas rotaciones ligeras, es decir, entre 1° y 20° pueden ser de bastante utilidad en tareas de reconocimiento en conjuntos como MNIST. Sin embargo, si se van incrementando los grados es muy posible que las etiquetas de los datos no sean preservadas después de la transformación, por ejemplo, el número 6 y el número 9 podrían ser confundidos.

- **Trasladar:** Para conjuntos de imágenes centradas, como es el caso de imágenes faciales, se aplica traslación en una dirección. El espacio sobrante puede ser llenado con un valor constante entre 0-255 o bien, con ruido gaussiano.

2.4.1.2. Transformaciones fotométricas

Estas suelen modificar los canales de color de las imágenes alterando sus pixeles a distintos valores. Las principales son perturbación de color o iluminación, escalado en grises y filtros. Algunos ejemplos pueden ser visualizados en la Figura 7.



Figura 7. Transformaciones fotométricas [4]

2.4.1.3. Transformaciones basadas en modelos

En este enfoque, se ajusta un modelo a la imagen de entrada para sintetizar imágenes con apariencia distinta al realizar variaciones en los parámetros del modelo. Estos son llamados generativos, y pueden ser clasificados tanto en perspectiva 2D como 3D.

2.4.1.4. Transformaciones basadas en modelos generativos

Este tipo de modelos proveen una poderosa herramienta para crear datos nuevos conservando la distribución de los originales al aprender del conjunto de entrenamiento respectivo. Generalmente se alcanza esto mediante el uso de NN y modificando sus parámetros.

Capítulo III. Generación de rostros sintéticos

En este capítulo se describe un modelo dedicado a la generación de rostros sintéticos, aún sin cambios en las características etarias. Esto con la finalidad de comprender la estructura de un modelo dedicado a construcción de imágenes faciales, sus componentes, funciones y limitaciones.

Modelo *Face Aging GAN*

Como primer acercamiento a modelos GAN, dada su complejidad, se optó por analizar y dar seguimiento a su estructura para mayor comprensión del funcionamiento de estos al generar imágenes faciales, aún sin cambios de edad o de otro tipo. La arquitectura propuesta para generar rostros sintéticos se presenta en la Figura 8.

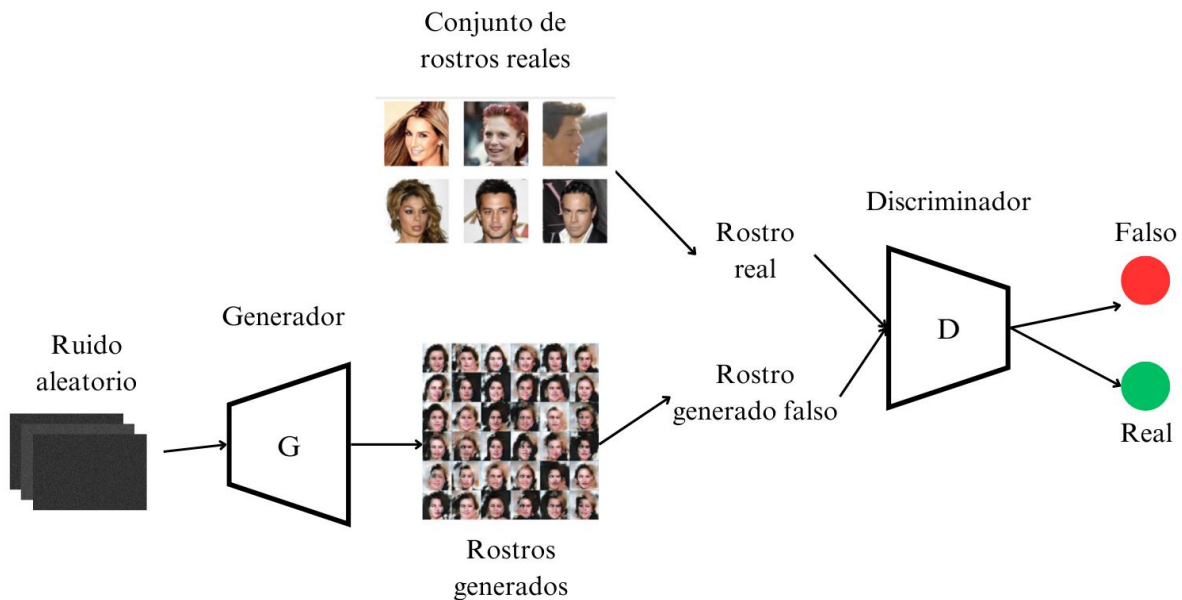


Figura 8. Arquitectura GAN para generación de rostros

En el diagrama se aprecian los componentes de una GAN integrada por dos redes, G y D, así como el flujo que sigue durante su entrenamiento. La entrada de G es un vector de ruido aleatorio, el cual está compuesto por un conjunto de números aleatorios. Se puede describir como una representación que transformará a una imagen de un rostro. Esta última se pasará a D junto con una muestra del conjunto real para que pueda analizar ambas y tratar de determinar cuáles son reales y aquellas que no lo son. Posterior a la clasificación, G recibirá retroalimentación basada en el rendimiento de D. Si D detecta que una imagen es falsa, G ajusta sus pesos para tratar de mejorar la calidad de sus imágenes en la siguiente iteración. Por otra parte, D se entrena para mejorar su precisión en la clasificación de imágenes reales y falsas. Recibe retroalimentación sobre su desempeño y ajusta sus parámetros en consecuencia. A través de múltiples iteraciones de entrenamiento, G mejorará la calidad de sus imágenes hasta que D no pueda distinguir de manera

efectiva entre las imágenes reales y las generadas. En este punto se puede determinar que la GAN ha alcanzado un equilibrio.

A continuación, la sección 3.1 describe el conjunto de datos que pasó por preprocesamiento. La sección 3.2 presenta la implementación de la arquitectura GAN.

3.1. Conjunto de imágenes para generación de rostros

El conjunto de imágenes público CelebA fue utilizado para el entrenamiento de la GAN dado que este es ampliamente aplicado en tareas de DL y visión, y una de las más desarrolladas es detección facial. Esta última con el propósito de detectar y reconocer patrones faciales o de apariencia. También tiene la ventaja de incluir variaciones en pose y diversidad entre individuos. Cuenta con un poco más de 200 mil imágenes faciales de celebridades, proporcionando una cantidad significativa de datos para entrenamiento. La gran cantidad de elementos permite que el modelo aprenda una amplia gama de características faciales, mejorando así su capacidad de generación realista. Otra ventaja es que incluye un archivo con las imágenes no procesadas y otro con las mismas ya recortadas y alineadas para su uso inmediato. Incluyen anotaciones de 40 atributos faciales, como la presencia de lentes, peinados y expresiones. Se observan algunas imágenes de ejemplo en su sitio, pero también se incluye para su visualización en la Figura 9. Obtenido del siguiente [enlace](#).



Figura 9. Conjunto de imágenes CelebA [27]

Se utilizaron 10,000 imágenes para el entrenamiento del modelo, 5000 épocas y un *batch* de tamaño 6. Primero se redimensionó cada elemento para estandarizar, quedando de un tamaño de 128 x 128 píxeles. Además, se trabajó con 3 canales de color, es decir RGB.

3.2. Arquitectura GAN para generación de rostros

Tras preparar el conjunto seleccionado para entrenamiento, se procedió a implementar una arquitectura GAN diseñada para la generación de imágenes faciales. Esta fue identificada después de examinar diversas estructuras capaces de ilustrar el funcionamiento de este tipo de redes en la creación de imágenes faciales.

Se optó por probar el código en un entorno de Anaconda, un software de distribución libre y abierta para poder aplicar lenguaje *Python*. Con fines prácticos y de estudio se utilizó un *Jupyter notebook*, para poder ir dando anotación a cada segmento de código.

Para la implementación de esta GAN se requirió una máquina virtual generada y configurada por el Laboratorio Nacional de Tecnologías de Información (LaNTI) en la Universidad Autónoma de Ciudad Juárez (UACJ). Se trabajó con el sistema operativo Ubuntu 20.04, doble núcleo, almacenamiento de 100 GB y RAM de 8 GB.

Debido a que se requiere una gran capacidad de cómputo para el entrenamiento de las GAN, se tuvo que realizar una serie de pruebas mediante la modificación de sus parámetros hasta que se lograra un entrenamiento que cumpliera con las expectativas.

G toma un vector de entrada de tamaño 32 y lo transforma en una imagen de tamaño 128x128 con tres canales de color (RGB) mediante capas convolucionales que juegan roles cruciales en la arquitectura ya que están diseñadas para transformar y refinar las características de las imágenes a lo largo del proceso generativo.

Posteriormente se implementa D. En este caso, dado que es una arquitectura para demostrar el funcionamiento de una GAN al generar imágenes artificiales de personas, D consiste en las mismas capas convolucionales que G. Toma una imagen de entrada de tamaños 128x128 con tres canales de color (RGB) y produce un único valor que indica si la imagen es real (1) o falsa (0). Al comportarse como un clasificador, utiliza *binary crossentropy* como función de pérdida, para medir la diferencia entre las probabilidades calculadas y las reales de las predicciones entre dos posibles clases. Durante su entrenamiento, recibe imágenes, tanto reales como generadas por G y aprende a distinguir entre ambas. Es importante resaltar que su objetivo es mejorar su precisión de clasificación, mientras que G intenta mejorar la calidad de las imágenes generadas para engañar a D. Este proceso adversarial y colaborativo entre G y D es lo que permite a las GAN generar datos sintéticos de alta calidad.

En la Figura 10 se presenta un diagrama que muestra la arquitectura del modelo GAN. Representa un modelo secuencial en el que se combina un generador y un discriminador. Comienza con una capa de entrada que toma un vector de tamaño 32, llamado ruido aleatorio. Este se pasa a G, una NN diseñada para transformar ruido a una imagen de tamaño 128x128 con tres canales de color. Esta última se pasa a D, diseñado con una NN dedicada a clasificar, y da como salida un único valor que representa la probabilidad de que la imagen sea real. Si esta es cercana a 1, cree que la imagen es real. Por el contrario, si es cercana a 0, cree que la imagen es falsa.

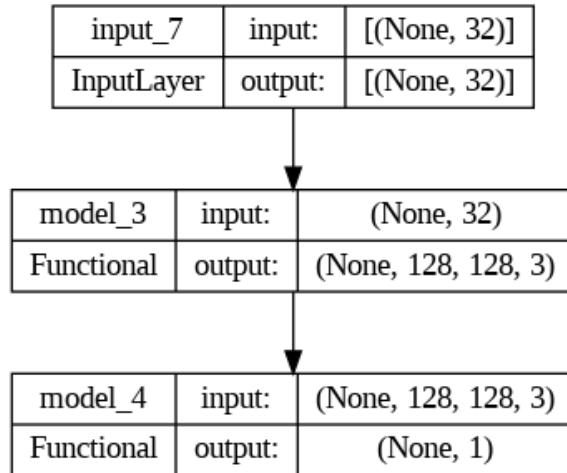


Figura 10. Arquitectura GAN

Con los parámetros asignados se obtuvieron las imágenes mostradas en la Figura 11.

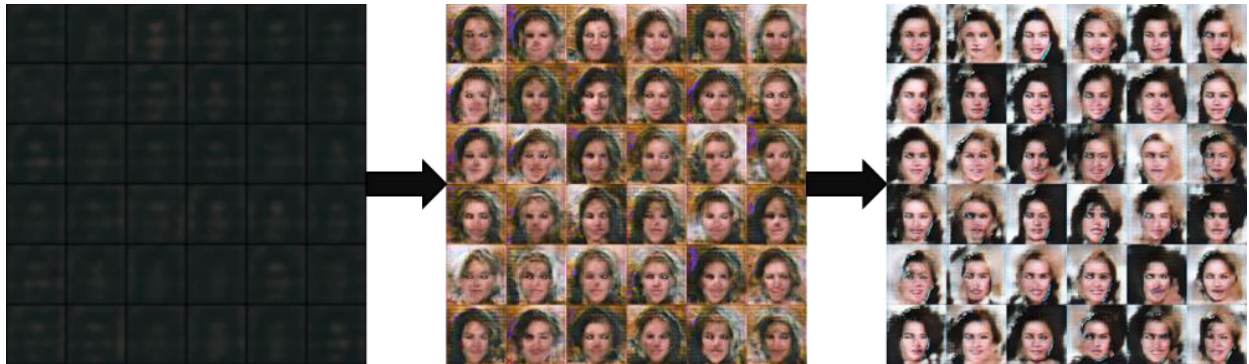


Figura 11. Imágenes resultantes de la GAN para generar rostros artificiales

Capítulo IV. Propuesta de solución

En este capítulo se describe la metodología implementada para estructurar el presente proyecto, así como las actividades realizadas en cada paso. Se describe la comprensión de los datos, la preparación de estos, el modelado que se siguió en cada arquitectura, así como el método de evaluación seleccionado para medir la similitud entre las imágenes generadas y la original de un rostro.

Metodología

Como se presentó en los trabajos relacionados, las GAN han demostrado un alto potencial en la generación de imágenes artificiales, mejorando la calidad de sus antecesores en la misma tarea, como se destaca en [26], aunque estos tienen el mérito de proporcionar las bases y avances en la generación de datos antes de la llegada de las GAN, las cuales han revolucionado el campo al permitir la creación de imágenes realistas mediante un novedoso enfoque adversarial [27].

4.1. Comprensión del negocio y los datos

El presente proyecto busca apoyar en la identificación y localización de mujeres mexicanas desaparecidas mediante el uso de tecnología avanzada de AI. El objetivo principal es desarrollar un modelo GAN que tenga la capacidad de generar imágenes de mujeres mexicanas desaparecidas con transformaciones en atributos faciales basados en la edad, con la finalidad de ayudar a autoridades y familias en la búsqueda y reconocimiento.

Los principales beneficiarios de los resultados obtenidos se concentran en las familias de mujeres desaparecidas, las autoridades y las organizaciones de búsqueda, ya que esta herramienta puede agilizar el proceso de generar envejecimiento en rostros de mujeres que no han sido localizadas en un tiempo prolongado. Para cumplir con los objetivos del proyecto, se identifican algunos requisitos esenciales, como la obtención de imágenes faciales de mujeres mexicanas desaparecidas, la generación de transformaciones de edad en dichas imágenes, así como la calidad y realismo de estas.

Existen también varios desafíos y consideraciones éticas que se enfrentan, dado el impacto social del proyecto. En el caso de las imágenes, se consideran las variaciones en calidad, resolución e iluminación, así como la dificultad en la recopilación y verificación de la información obtenida. El procedimiento para la obtención de imágenes de personas desaparecidas se concentró principalmente en organizaciones gubernamentales. En este sentido, se realizó contacto con la Fiscalía General del Estado de Chihuahua, específicamente con el Centro Estatal de Información, Análisis y Estadística Criminal en Ciudad Juárez, Chihuahua, sin embargo, la obtención de estas imágenes se dificultó debido a acuerdos de confidencialidad con las familias de las afectadas, teniendo comprometida la ética. Por lo que se comenzó una búsqueda de conjuntos de imágenes públicos para entrenamiento, pruebas y despliegue de resultados del modelo.

Uno de los conjuntos analizados para la implementación de la arquitectura fue “IMDB-Wiki” [26]. De carácter público, el cual contiene más de 500 mil imágenes etiquetadas con género y edad para entrenar modelos que puedan predecir ambos. Este es pensado debido a la alta demanda de imágenes que requieren algunos modelos de DL. Toma una lista de 100 mil actores en el sitio web IMDB y se utiliza su información pública, como género, edad y todas las imágenes posibles de la persona. Una estadística con su distribución es mostrada en el sitio, pero se puede visualizar en la Figura 12.

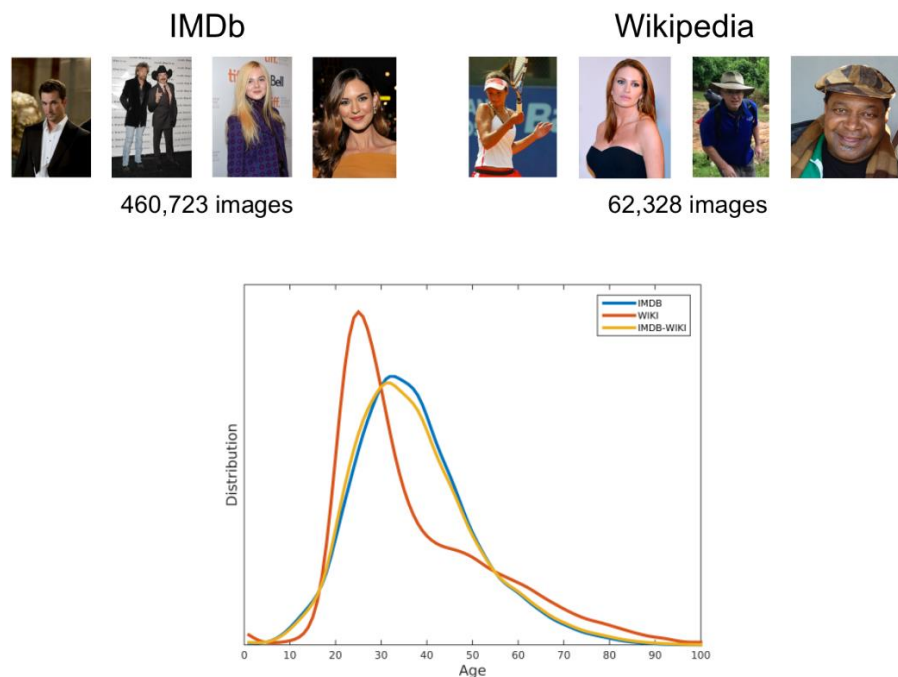


Figura 12. Conjunto de imágenes IMDB-Wiki

Este conjunto, al tener imágenes con múltiples rostros era complicado para modelar. Además, existían imágenes dentro del conjunto que contenían efectos, como modificación de tonos, blanco y negro, antiguos, entre otros. En resumen, su nivel de complejidad era muy alto para el preprocesamiento que requeriría en la generación de envejecimiento de rostros.

Por lo anterior, se optó por seleccionar un conjunto de imágenes que contara con las clasificaciones necesarias, es decir, que incluyera fotografías de mujeres de etnia latina, lo más similar posible a

la temática de mujeres mexicanas que propone el presente proyecto. Para cumplir con estos requisitos se eligió el conjunto *UTKFaces*, obtenido del siguiente [enlace](#).

Su uso se ve justificado dado que contiene fotografías con un rostro identificable e incluyen información relevante como la edad y la etnicidad, indispensables para los objetivos de la investigación. El conjunto *UTKFaces* es una colección de imágenes que cubre un amplio espectro de edades, desde 0 hasta 116 años, etiquetadas con información como la edad, género y etnicidad. Consiste en más de 20,000 imágenes faciales que presentan una considerable variabilidad en términos de poses, expresiones, iluminación, oclusión y resolución. Ha sido utilizado en investigaciones relacionadas con la predicción de edad y reconocimiento facial, como se presenta en [28], [29] y [30].

4.2. Preparación de los datos

El preprocesamiento del conjunto *UTKFaces* se dirigió enteramente a los objetivos del proyecto, comenzando con un análisis estadístico del mismo para extracción de aquellas imágenes que cumplieran con los requerimientos. Se utilizaron los metadatos de cada fotografía para su clasificación y la preparación del conjunto. En la Figura 13 se presenta una muestra de la información disponible de cada imagen.

Se puede también dar visualización gráfica de la distribución según el género en la Figura 14 **Figura 14**. Dando un total de 12,583 para el género masculino y 11,529 del femenino. A su vez, se presenta un gráfico para mostrar la distribución de las edades en Figura 15.

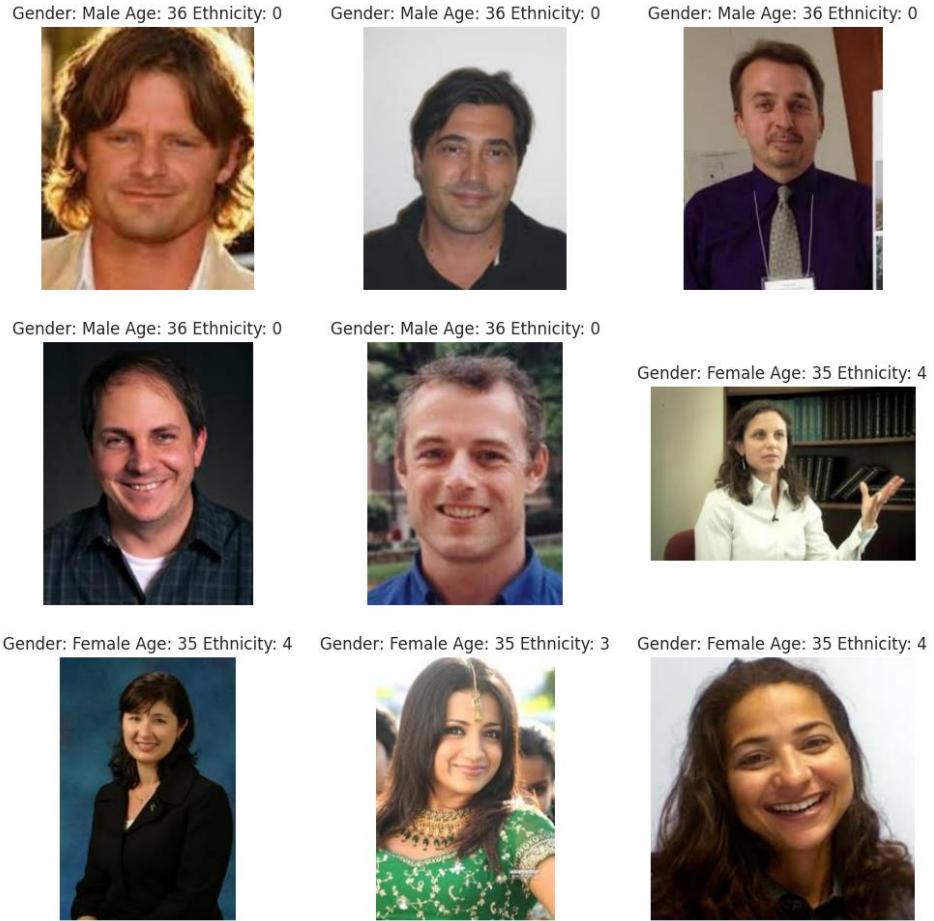


Figura 13. Muestras de los metadatos contenidos en el conjunto *UTKFaces*

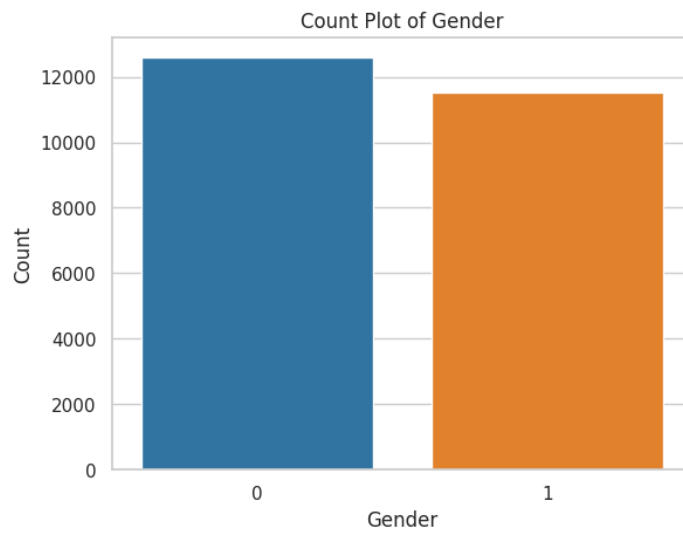


Figura 14. Gráfico de barras de distribución por género de *UTKFaces*

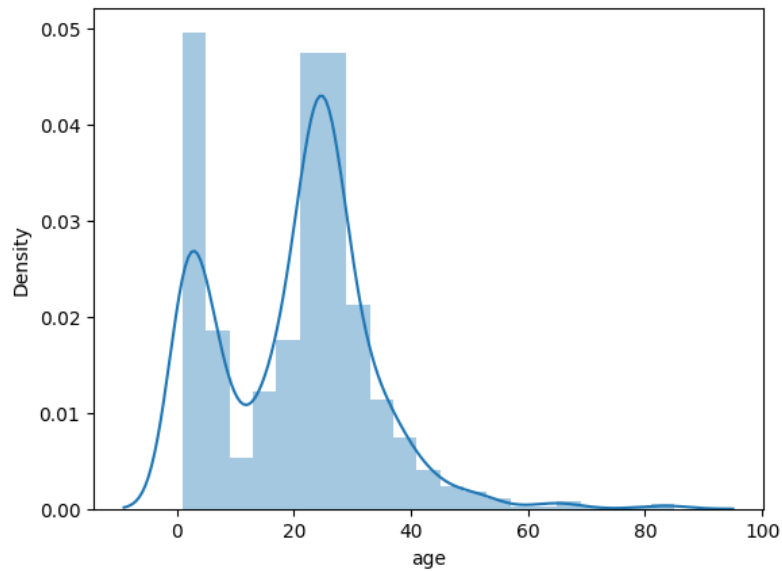


Figura 15. Distribución de edad en *UTKFaces*

La primera división del conjunto se basó en el género, almacenando únicamente las pertenecientes al género femenino. Posterior a ello se reservaron aquellas de etnia latina, debido a los intereses del proyecto y con el propósito de visualizar el comportamiento del modelo frente a los rasgos de dicho grupo.

Al tener el conjunto a utilizar, se continuó con su preparación. Ya que se utilizó un modelo pre entrenado de GAN [31], el conjunto debía tener un formato específico para poder ser agregado al entrenamiento. Para esto, se utilizó un modelo de detección de puntos faciales proveniente de la biblioteca `dlib` [32]. Este se encuentra público bajo el nombre `shape_predictor_68_face_landmarks.bz2`, utilizado ampliamente en aplicaciones de procesamiento de imágenes y visión por computadora para detectar características faciales como ojos, cejas, nariz, boca y la forma del contorno del rostro. Fue auxiliar para detectar los rostros y, en base a ello, se alineó la imagen y se recortó, quedando de un tamaño de 256x256. El procedimiento se presenta visualmente en Figura 16 **Figura 16.**

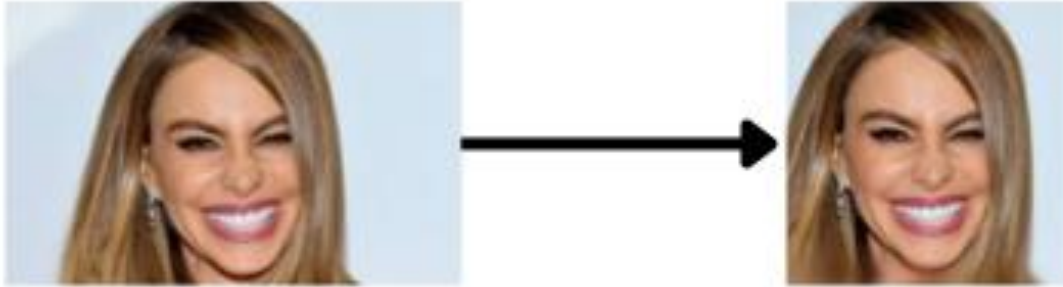


Figura 16. Alineación y recorte

También se utilizó el modelo de predicción de puntos faciales para encontrar y eliminar aquellas imágenes que no tuvieran un rostro, las cuales fueron un total de 12. Quedando un conjunto de 930 imágenes de mujeres de etnia latina.

Para el proceso de *fine-tuning* del modelo se hizo una división del conjunto de imágenes preprocesado, dedicando un 80% a *train* y el 20% restante a *test*.

4.3. Modelado

En los últimos años, las GAN se han convertido en una poderosa herramienta en el campo de DL, con su capacidad para generar datos sintéticos que son indistinguibles de los reales se ha destacado notablemente. Se han aplicado a una variedad de áreas, incluyendo la generación de imágenes, síntesis de voz, y la creación de datos sintéticos para entrenamiento de modelos. En el caso de envejecimiento en rostros es particularmente especial, dado que no requiere foto de antes y después de las personas en el conjunto de entrenamiento para aprender a generar sus características. El proceso de modelado de las GAN determina la eficiencia y calidad de los datos que genere.

La presente sección se enfoca en el modelado utilizando dos arquitecturas GAN distintas, proporcionando construcción y entrenamiento.

4.3.1. Envejecimiento de rostros con Red Generativa Adversaria Condicional (*conditional Generative Adversarial Network*, cGAN)

Para lograr los objetivos del proyecto se estudiaron algunas arquitecturas dedicadas al envejecimiento de rostros. Una de las primeras implementaciones fue utilizando una *conditional*

Generative Adversarial Network (cGAN), ya que modela en base a información adicional, la edad. En lugar de generar datos de forma completamente aleatoria, utiliza la información condicional para influir en el proceso de generación.

Se utilizó el conjunto de imágenes IMDB-WIKI, incluye una carpeta con los rostros previamente centrados y recortados. Además, cada imagen se encuentra etiquetada con la fecha de nacimiento y en la que fue capturada la fotografía, usando esto para calcular la edad de cada fotografía y extraer las características del respectivo grupo etario.

El modelo cGAN se compone de dos redes, G y D. G se construye con:

- Dos entradas concatenadas: vector de ruido aleatorio de tamaño 100 y la etiqueta de clase con tamaño 6, como se puede ver en
- . Estas se utilizan para condicionar el proceso de generación, siendo lo que permite controlar la clase de los datos generados.
- Se conforma de 22 capas. Cada una cumple un papel importante en la generación de imágenes. Las capas densas se utilizan para transformar el ruido y las etiquetas de clase en una representación, mientras que las convolucionales y los filtros ayudan a generar la estructura y los detalles de la imagen final.
- La salida tiene 3 filtros correspondientes a los canales de color RGB.

Rango de edad	Etiqueta
0-18	0
19-29	1
30-39	2
40-49	3
50-59	4
60+	5

En resumen, G toma un vector ruido y una etiqueta de clase, los combina y pasan a través de varias capas densas y convolucionales con funciones de activación y normalización. Tiene integradas capas, que van transformando el vector de entrada en una imagen 64x64x3 que representa un rostro sintético condicionado a la clase de edad proporcionada.

Rango de edad	Etiqueta
0-18	0
19-29	1
30-39	2

Tabla 2. Etiquetas de clase

40-49	3
50-59	4
60+	5

Por otra parte, D se constituye como el clasificador binario que se encarga de producir probabilidades:

- Recibe dos entradas, tanto la imagen como la etiqueta.
- Produce una salida que representa la probabilidad de que la imagen sea real o falsa.

La combinación de estas dos redes puede ser visualizada en la Figura 17.

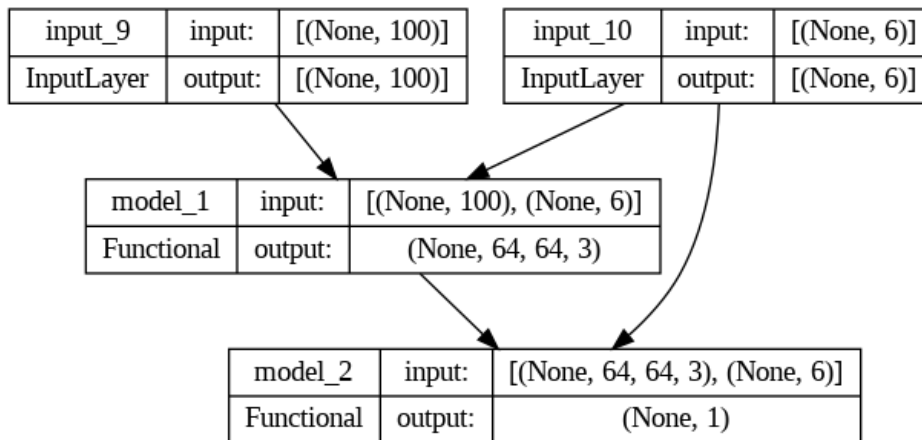


Figura 17. cGAN para envejecimiento en rostros

En esta GAN, G recibe como entrada el vector ruido y la etiqueta de clase para producir una imagen de 64x64x3 que se le pasará a D. Por su parte D recibirá la imagen y también la etiqueta de clase. Su objetivo es determinar si la imagen es real o falsa, para esto, evalúa la imagen recibida junto con la etiqueta y produce una salida binaria indicando su predicción.

4.3.2. Envejecimiento en rostros con modelo StyleGAN-based Age Manipulation (SAM)

En [31] se propuso un modelo que combina varios modelos y técnicas alcanzadas de DL. Entre ellos *StyleGAN2*, para generar imágenes realistas y manipulables, permitiendo modificar la apariencia de edad sin alterar otros aspectos de la identidad del rostro. Un modelo de regresión para aprender a transformar el estilo de la imagen de entrada a diferentes edades, utilizando un conjunto de imágenes faciales etiquetadas con la edad, permitiendo al modelo aprender características específicas asociadas con distintas etapas etarias. CNN para extracción de características faciales importantes de las imágenes para capturar detalles específicos de la apariencia que cambia con la edad, como arrugas, textura de la piel y cambios en la forma del rostro.

En la arquitectura presentada (Figura 18), no se implementa un D tradicional, pues se basa en la manipulación del espacio latente de un G pre entrenado (*StyleGAN2*) y el uso de varias funciones de pérdida para guiar el entrenamiento y garantizar la calidad de las imágenes generadas.

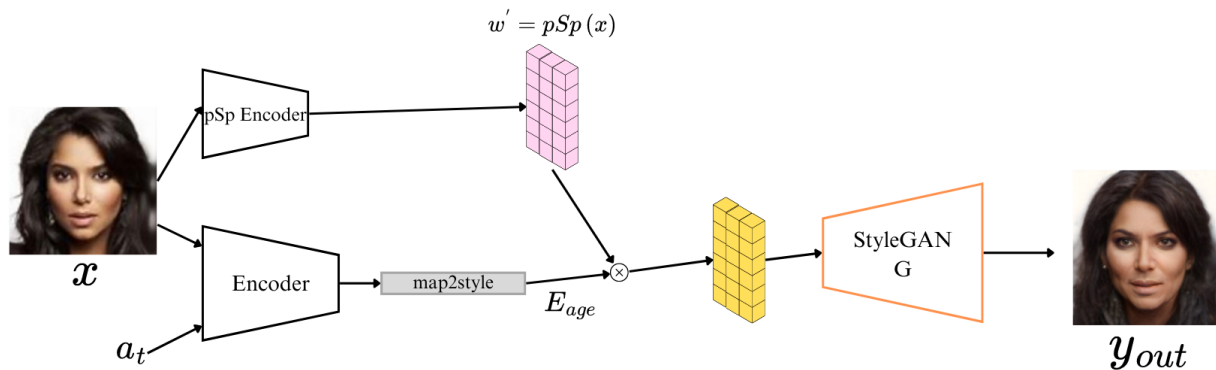


Figura 18. Arquitectura de modelo SAM

Para detallar sobre el modelado de la arquitectura es necesario explicar el funcionamiento de cada modelo empleado:

Pixel2style2pixel (pSp Encoder)

Pixel2style2pixel [33] es una red que toma una imagen de entrada y la convierte en un vector en el espacio latente del generador *StyleGAN2*. Captura las características esenciales de la imagen original, tales como rasgos faciales, expresiones, peinado, color de cabello, entre otros. Y, de esta manera que permite la manipulación de dichos atributos para generar nuevas versiones de la imagen utilizando a G.

Encoder

El *encoder* presentado en la arquitectura se encarga de transformar la imagen de entrada junto con la edad objetivo en una representación específica que captura la información necesaria para generar una imagen con la edad deseada. Aprende a modificar las características de la imagen para reflejar el cambio de edad mientras mantiene la identidad.

map2style

La salida del *encoder* es un vector latente que encapsula la información necesaria para transformar la imagen de acuerdo a la edad objetivo. Sin embargo, este necesita ser adaptado para ser utilizado por el G de *StyleGAN*, que opera en un espacio de estilo específico. Para esto, *map2style* toma el vector y lo transforma en una serie de vectores que G pueda usar.

Generador de *StyleGAN*

StyleGAN G es una red neuronal avanzada que recibe un vector latente que representa la imagen y los vectores de estilo generados por *map2style*, que modulan las características de cada capa del generador.

Ya explicada la funcionalidad de cada modelo se puede detallar el flujo de trabajo que sigue la arquitectura:

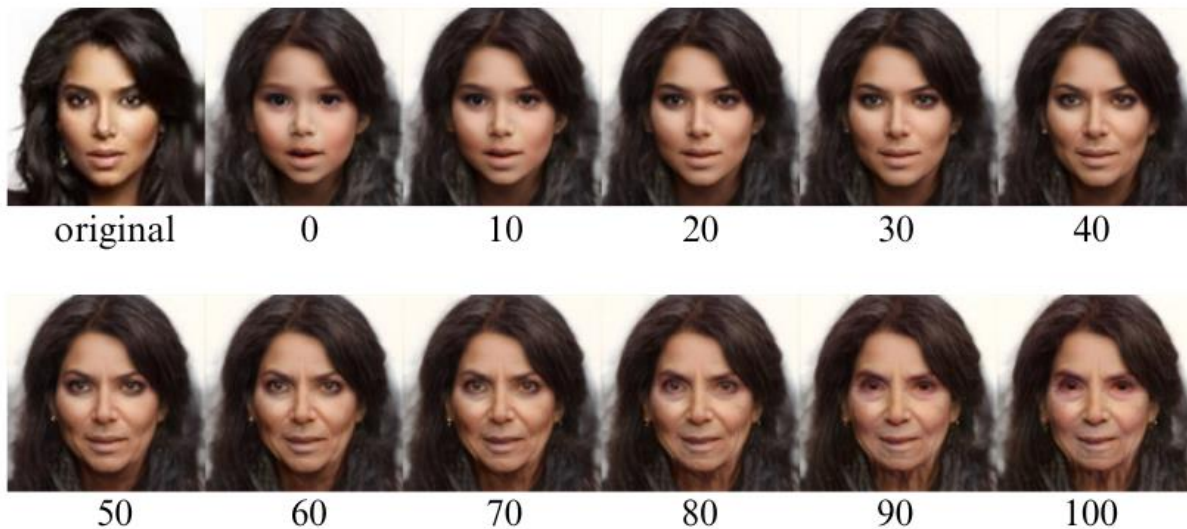
- La imagen de entrada x se mapea al espacio latente w' utilizando el encoder pSp. Esta representación captura las características esenciales de la imagen.
- A su vez, la imagen de entrada x y la edad objetivo a_t se pasan a través del encoder para generar una representación específica para la edad. Esta se mapea al espacio de estilo.
- La representación latente mapeada se pasa a *StyleGAN G*, que produce la imagen transformada y_{out} correspondiente a la edad objetivo.

En resumen, SAM es un enfoque basado en estilo para transformar la edad en imágenes faciales, aprovechando *encoders* y generadores pre entrenados. Dando como salida un rostro en diferentes etapas etarias (0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100).

Entrada:



Salida:



Las pruebas de esta arquitectura fueron desarrolladas en *Google Colab*: V100 GPU con RAM de 62.8 GB, almacenamiento en disco de 201.2 GB y velocidad de RAM aumentada.

El modelo basado en *StyleGAN2* es un sistema avanzado de manipulación de edad, utilizando componentes pre entrenados y técnicas de transformación de imágenes para generar resultados de alta calidad. Al no ser una GAN por sí mismo, los hiperparámetros con los que trabaja son distintos

a la tasa de aprendizaje o número de épocas. Este trabaja con las transformaciones aplicadas, configuraciones específicas del modelo y parámetros del experimento.

Parámetros del experimento en psp

Los parámetros se pasan a través de un objeto de configuración que contiene varios ajustes necesarios para el modelo.

1. Tamaño de salida (*output_size*): Define el tamaño de la imagen generada
2. Ruta del *checkpoint* (*checkpoint_path*): Ruta al archivo que contiene pesos pre entrenados del modelo psp.
3. Número de canales de entrada (*input_nc*).
4. Uso de encoder pre entrenado (*start_from_encoded_w_plus*): Indicador booleano que determina si se debe iniciar desde un encoder pre entrenado.
5. Promedio latente (*latent_avg*): Valor utilizado para mejorar la calidad de las imágenes generadas.
6. Tipo de encoder (*encoder_type*).

Parámetros de transformación de edad

Estos se utilizan para especificar las edades objetivo en la manipulación de la imagen.

1. Edades objetivo (*target_ages*): Lista de edades a las que se desea transformar la imagen.
2. *Age transformer*: Módulo que aplica las transformaciones de edad.

4.4. Evaluación

Como se menciona en la sección 2.3.3. Evaluación, existen numerosos métodos para evaluar la calidad, variación y similitud de imágenes generadas por GAN. En el presente proyecto se optó por emplear una herramienta en la nube, *AWS Rekognition*, que ofrece análisis de imágenes como detección de rostros y medición de similitudes. El uso de este tipo de plataformas es menos costoso en tiempo y recursos, ya que ofrecen modelos pre entrenados.

La elección de *AWS Rekognition* se fundamenta en sus capacidades avanzadas de análisis de imágenes y reconocimiento facial, cruciales para la tarea de evaluar similitud entre imágenes

generadas y originales. Además, permite identificar características faciales clave para evaluar cómo se preserva la identidad facial en las imágenes envejecidas generadas. También tiene la capacidad de comparar atributos faciales como la expresión, edad estimada, género y otros detalles que pueden ser de importancia para evaluar la consistencia de una imagen generada con la apariencia esperada según la edad objetivo. Utiliza un porcentaje de similitud para cuantificar qué tan similares son dos rostros, ayudando a determinar si la identidad se preservó adecuadamente. Otro punto importante por resaltar es que utiliza atributos faciales para asegurar que las características esenciales del rostro (forma, proporciones, rasgos) se mantuvieran consistentes, a pesar de los cambios debidos al envejecimiento.

A continuación, se presenta el procedimiento que se siguió para la evaluación de las imágenes.

- **Configuración del entorno**

- **Credenciales y configuración de AWS CLI**

- Para utilizar *Amazon Rekognition*, primero se configuraron las credenciales de AWS y CLI de AWS.

- 1. Crear una cuenta en AWS:

- Crear una cuenta en aws.amazon.com.

- 2. Instalar y configurar AWS CLI

- Se instaló AWS CLI siguiendo instrucciones oficiales.
 - Configuración de AWS CLI mediante el comando `'aws configure'`, donde se ingresan el `'AWS Access Key ID'`, `'AWS Secret Access Key'`, región y formato de salida por defecto.

- **Preparación de las imágenes**

- **Almacenamiento en Amazon S3**

- Las imágenes generadas y originales se almacenaron en un *bucket* de Amazon S3.

- 1. Crear un *bucket* en S3:

- Se creó un *bucket* en Amazon S3 llamado `'gan-bucket-aging'`.

- 2. Subir imágenes a S3:

- Utilizando un *script* de *Python* en *colab notebook*, se subieron las imágenes al *bucket* (Figura 19).

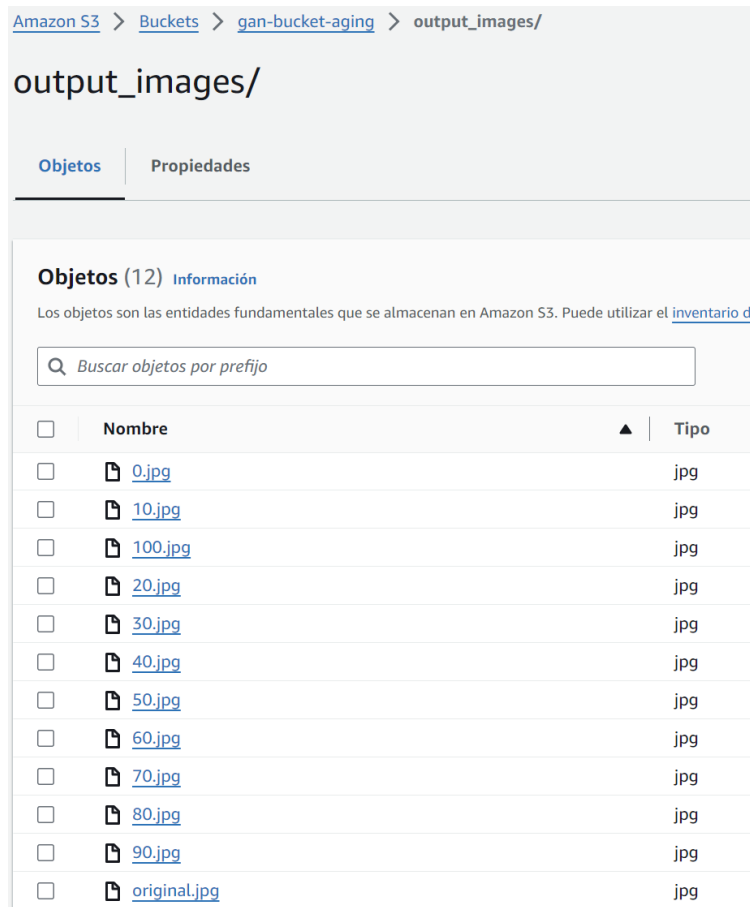


Figura 19. Bucket Amazon S3

- **Evaluación con Amazon Rekognition**

Se preparó el *script* en un *colab notebook*.

Configuración de Amazon Rekognition

Para evaluar la similitud de las imágenes, se configuró *Amazon Rekognition*.

1. Configurar el cliente de *Rekognition*:
 - Se creó un cliente de *Rekognition* usando las credenciales de AWS.

Comparación de imágenes

Se compararon las imágenes originales con las generadas usando la función ‘*compare_faces*’ de *Amazon Rekognition*.

1. Función para comparar rostros.
 - Se definió una función para comparar rostros y extraer la similitud.
2. Comparación de imágenes originales y envejecidas.
 - Se compararon todas las imágenes generadas con la imagen original.

Los resultados de la evaluación se explicarán a detalle en el Capítulo 5 del presente documento.

Procedimiento de AWS Rekognition para calcular similitud

En [9] se describen los cálculos de similitud que realiza *AWS Rekognition* para comparar dos imágenes faciales. Los pasos se resumen a continuación:

1. Detección de rostros: AWS detecta los rostros en ambas imágenes utilizando técnicas avanzadas de visión por computadora.
2. Extracción de características: una vez que se detectaron los rostros, el servicio extrae características faciales claves, como posición de los ojos, nariz, boca, forma del rostro.
3. Comparación de características: las características extraídas se comparan utilizando algoritmos de aprendizaje profundo, midiendo las diferencias entre los puntos faciales y las características extraídas.
4. Cálculo de similitud: la similitud se calcula en base a la diferencia entre las características faciales de dos imágenes, proporcionando un puntaje de similitud entre 0 y 100, donde un puntaje más alto indica mayor similitud.

4.5. Despliegue

Con la obtención de los resultados se definen posibles aplicaciones prácticas que puede tener una GAN dedicada a generar cambios en la edad de mujeres mexicanas desaparecidas.

- Reconstrucción de apariencia: Puede ayudar a reconstruir la apariencia de mujeres desaparecidas a medida que envejecen. Esto es de suma importancia para mantener actualizadas las imágenes en carteles de búsqueda y bases de datos.
- Comparación con imágenes recientes: Se puede realizar con mayor facilidad una comparativa de fotografías antes de la desaparición con imágenes recientes de mujeres que podrían ser vistas.
- Mejora de la identificación: Ayudar a las autoridades a identificar mujeres desaparecidas que han sido vistas después de un periodo prolongado.
- Búsqueda y seguimiento: Apoyo a las investigaciones proporcionando imágenes actualizadas que pueden ser distribuidas a través de redes policiales y comunitarias.
- Difusión efectiva: Imágenes actualizadas pueden ser utilizadas en campañas de sensibilización, redes sociales, y medios de comunicación, para captar la atención del público y obtener información de personas que puedan haber visto a las desaparecidas.

Las anteriores aplicaciones demuestran cómo las GAN pueden ser una herramienta poderosa en la búsqueda y recuperación de personas desaparecidas, así como en la mejora de estrategias de identificación y sensibilización pública.

Para difusión sobre el tema se redactó un capítulo publicado en el libro *Data Analytics & Computational Intelligence: Novel Models, Algorithms, and Applications* [34], titulado *Data augmentation techniques for facial image generation: a brief literatura review*. (véase Anexo VI). Además, se participó como ponente con el presente tema de investigación en diversos coloquios de posgrado organizados por la Universidad Autónoma de Ciudad Juárez. (véase Anexo I, II, III y IV). Y, hubo integración con el panel Mujeres In para destacar el impacto del proyecto en la investigación (véase Anexo V)

Durante el modelado, se debe destacar que una GAN, por más simple que se estructure, requiere una inversión significativa de tiempo y recursos. Al ser un modelo con tanto poder, también tiene un alto coste computacional debido a la complejidad y la cantidad de datos requeridos para obtener resultados de calidad. Puede variar también dependiendo de factores como tamaño del conjunto de imágenes, *hardware* utilizado, duración del entrenamiento, capas, funciones de pérdida, entre otros. Su entrenamiento puede requerir desde semanas hasta meses, implicando coste energético y de tiempo.

El estudio tiene una relevancia para la investigación en México por algunos puntos que se detallan a continuación.

- Mejora en la identificación y localización de desaparecidos: En México, el problema de personas desaparecidas es de gran escala. Según datos de la CNB, existen miles de personas reportadas como desaparecidas en el país [8]. Siendo especialmente útil la implementación del modelo GAN en aquellos casos donde la persona ha estado desaparecida por muchos años y su apariencia se ha visto modificada.
- El desarrollo y la implementación de tecnologías avanzadas como las GAN fortalecen las capacidades tecnológicas de las autoridades encargadas de la búsqueda y rescate.
- Contribuye al avance de la investigación científica y tecnológica en México. Fomenta la colaboración entre científicos de datos, ingenieros, y autoridades de seguridad, promoviendo el desarrollo de soluciones innovadoras para problemáticas complejas. Además, impulsa la formación de profesionales en el campo de la AI y ML, áreas de gran importancia para el desarrollo tecnológico del país.

En resumen, el estudio sobre el uso de las GAN para generar cambios en la edad de mujeres mexicanas desaparecidas es relevante no solo por su potencial para apoyar casos específicos de desaparición, sino también por su contribución al desarrollo tecnológico y fortalecimiento de las capacidades institucionales del país.

Capítulo V. Resultados y discusiones

En la presente sección se despliegan los resultados obtenidos del uso de un modelo GAN para generar imágenes que simulan el envejecimiento en rostros de mujeres mexicanas desaparecidas. Se describen las características del modelo implementado, los datos utilizados, el proceso de evaluación, y los resultados obtenidos en términos de similitud entre las imágenes originales y las generadas.

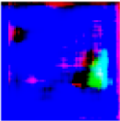
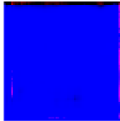
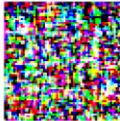
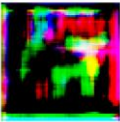
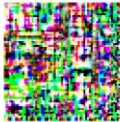

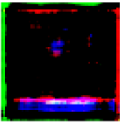
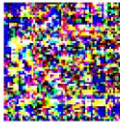
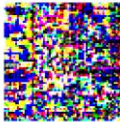
Resultados

En la sección 5.1. se presenta el primer modelo implementado cGAN, así como los resultados obtenidos durante las pruebas. La sección 5.2. se presenta el modelo SAM, que generó las imágenes con envejecimiento, con su evaluación y resultados.

5.1. Modelo cGAN

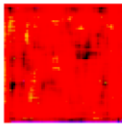
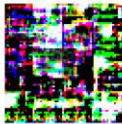

Las primeras pruebas que generaron imágenes se dieron en la arquitectura cGAN, como se describe en la sección 3.3.1. En esta arquitectura se aumentó significativamente el nivel de complejidad, por lo que se requería mayor capacidad de cómputo. Las primeras pruebas se hicieron en la misma máquina virtual que en la generación de rostros: sistema operativo Ubuntu 20.04, doble núcleo, almacenamiento de 100 GB y RAM de 8 GB. Hubo variedad de problemas de rendimiento, por esto no fue posible continuar en dicho entorno y se pasaron los experimentos a *Google Colab*: RAM de 12.7 GB, almacenamiento en disco de 166.8 GB. Aquí, con un acelerador T4 GPU y aumentando la velocidad de RAM se pudieron comenzar las pruebas obteniendo resultados únicamente con menos de 300 épocas, como se aprecia en Tabla 3, después de cierto tiempo de ejecución se desconectaba el entorno y se perdían las salidas obtenidas durante el proceso.

Tabla 3. Google colab, T4 GPU, RAM acelerado

GOOGLE COLAB, T4 GPU, RAM ACELERADO			
20 épocas			
100 épocas			
250 épocas			

Dada la arquitectura tan robusta, el tiempo que tomaba el entrenamiento continuaba siendo prolongado, por lo que se optó por adquirir *Google colab pro*: RAM de 51 GB, almacenamiento en disco de 166.8 GB. En esta versión se contaba con RAM de alta capacidad y V100 GPU. Uno de los beneficios en esta versión es que se extendió el tiempo que tardaba en desconectar el entorno sin actividad por parte del usuario. Esto permitió llegar a un máximo de 800 épocas, y algunas imágenes pueden visualizarse en Tabla 4.

Tabla 4. *Google colab pro*, V100 GPU, RAM alta capacidad.

GOOGLE COLAB PRO, V100 GPU, RAM ALTA CAPACIDAD			
800 épocas			

Durante este último experimento las funciones de pérdida reflejaron el comportamiento mostrado en la Figura 20.

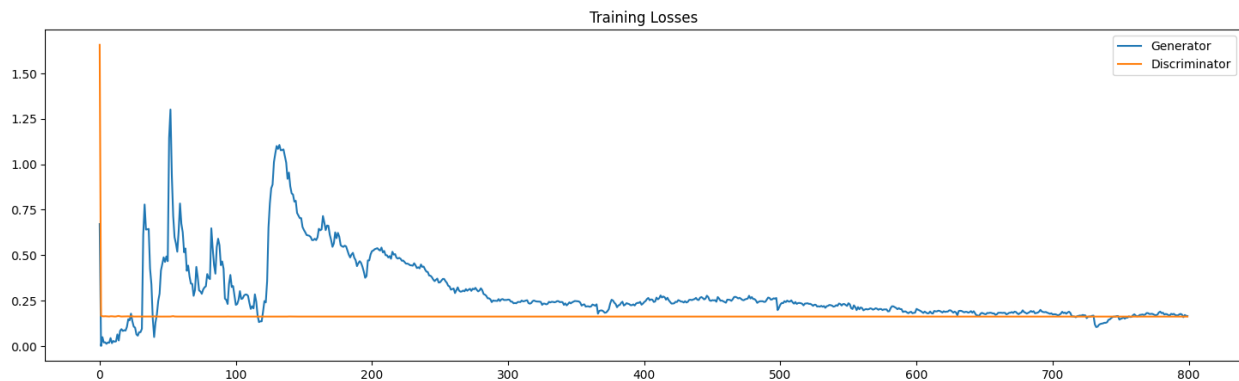


Figura 20. Entrenamiento cGAN

Al inicio del entrenamiento, la pérdida de D comienza muy alta y disminuye rápidamente, estabilizándose alrededor de un valor bajo después de 10-20 épocas. Lo cual sugiere que aprende veloz a distinguir entre imágenes reales y falsas. Por otra parte, G muestra alta variabilidad al

inicio, con picos pronunciados, lo cual es bastante común durante el entrenamiento de una GAN, ya que este intenta aprender a producir imágenes que puedan engañar a D.

Durante la estabilización inicial, la pérdida de D se mantiene constante en un bajo valor después de la caída inicial, lo que indica que aprendió fácilmente a diferenciar entre imágenes reales y generadas. La pérdida de G sigue siendo volátil, con una tendencia general a la baja, lo que indica que se encontró en un proceso de mejora gradual en la producción de imágenes que puedan engañar a D.

En la fase final del entrenamiento, alrededor de la época 300, la pérdida de G muestra una tendencia más estable y disminuye lentamente, aunque sigue presentando pequeñas fluctuaciones. La pérdida de D se mantiene estable en su valor bajo, indicando que sigue siendo efectivo en su tarea.






Hacia el final del entrenamiento, parece que el sistema alcanza el equilibrio buscado, donde ambas pérdidas son relativamente estables.





5.2. Modelo pre entrenado basado en *StyleGAN*




El envejecimiento en rostros fue dado gracias al modelo SAM [31] descrito en la sección 3.3.2. Se pretendía realizar *fine-tuning* con el conjunto UTKFaces, descrito en la sección 3.2. Sin embargo, esto no fue posible debido a que el modelo SAM, al utilizar algunos modelos pre entrenados, a su vez requería archivos de pesos pre entrenados que no fue posible obtener ni definir dada la estructura tan compleja. A pesar de esto, el modelo es entrenado originalmente con el conjunto de imágenes *FFHQ-Aging* [35]. El conjunto preprocesado se utilizó para generar los rostros envejecidos y evaluar los resultados.

Se presentan dos evaluaciones con la herramienta *AWS Rekognition*, descrita en la sección 4.4. La primera se realiza en una imagen facial de alta resolución, mujer de 43 años, cabello recogido, dejando su rostro completamente visible libre de cualquier posible oclusión, mirando directamente a la cámara con pose totalmente frontal. La pose más sencilla para las transformaciones se podría inferir (véase Tabla 5).

Tabla 5. Resultados mujer 43 años

Entrada	Edad objetivo	Salida	Similitud (%)
	0		54
	10		80
	20		97
	30		85



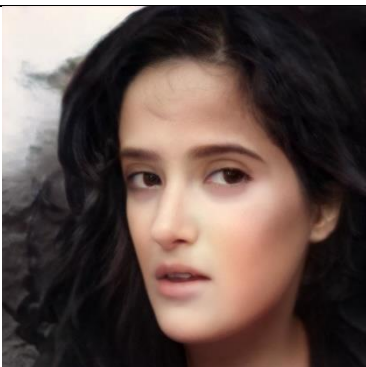


	40		75
	50		63
	60		78
	70		54





	80		32
	90		20
	100		0




Se puede observar que el modelo generó imágenes con mayor similitud en el rango etario 10-60 años. En el caso de edades más avanzadas, como 80, 90 y 100 años, las similitudes son mucho menores, 32%, 20% y 0% respectivamente, reflejando que es menos confiable al construir características de este rango.

Como segunda prueba se aplicó el proceso de envejecimiento a una fotografía facial con mayor reto para el modelo, aparentemente. La imagen muestra a una mujer, 47 años, con la cabeza ligeramente inclinada, cabello alrededor del rostro, presentando una ligera oclusión de este, parcialmente cubierto u obstruido. Véase Tabla 6.

Tabla 6. Resultados mujer 47 años

Entrada	Edad objetivo	Salida	Similitud (%)
	0		76
	10		93
	20		95
	30		95






	40		93
	50		89
	60		88
	70		85





	80		65
	90		25
	100		22


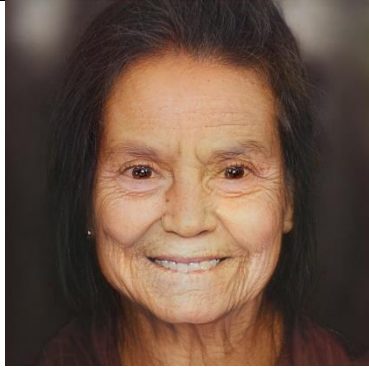
Muy parecido al experimento anterior, se observa una mejor tasa de similitud en el grupo de 10-70 años. Presentando una capacidad mucho menor para representar características de 90 y 100 años, con 25% y 22% respectivamente. Fuera de esto, el modelo parece haber plasmado las características en los rostros y conservando buena parte de su identidad en la mayoría de los casos.

Se presenta un experimento con una fotografía facial perteneciente a una adolescente de 14 años, mirando directamente a la cámara, cabeza y rostro completamente visibles, iluminación sin sombras significativas que cubran el rostro y cabeza sin inclinación notable (véase Tabla 7).

Tabla 7. Resultados adolescente 14 años

Entrada	Edad objetivo	Salida	Similitud (%)
	0		97
	10		99
	20		98
	30		80

	40		66
	50		56
	60		52
	70		44

	80		27
	90		0
	100		0

Se puede apreciar cómo los índices de similitud se observan muy altos en las edades 0, 10 y 20 años. A partir de los 30 años, la similitud comienza a disminuir significativamente. Esto sugiere que la imagen original tiene características faciales muy juveniles, donde parece que el modelo tiene mayor debilidad para predecir facciones de un adulto mayor en una adolescente o niña menor de edad.

Como los experimentos anteriormente presentados, se realizó el mismo procedimiento de evaluación con el conjunto preprocesado de *UTKFaces*. Se conformó por fotografías únicamente de mujeres de etnia latina, quedando un total de 942 imágenes. Todas fueron transformadas con el

modelo descrito en 4.3.2. Para cada grupo etario se calculó media, mediana y desviación estándar, presentados en Tabla 8. La media y la mediana pueden ser visualizados gráficamente en la Figura 21.

Tabla 8. Estadística descriptiva de similitud por grupo etario

Edad	Media (%)	Mediana (%)	Desviación estándar (%)
0	52.4	55.9	33.9
10	69.2	81.1	30.9
20	75.8	90.1	29.6
30	69.1	84.5	33.5
40	60.2	71.2	35.3
50	50.3	53.5	35.5
60	42.2	41.3	34.0
70	35.1	32.1	32.8
80	25.2	20.0	29.8
90	17.4	0.0	25.6
100	13.9	0.0	23.4

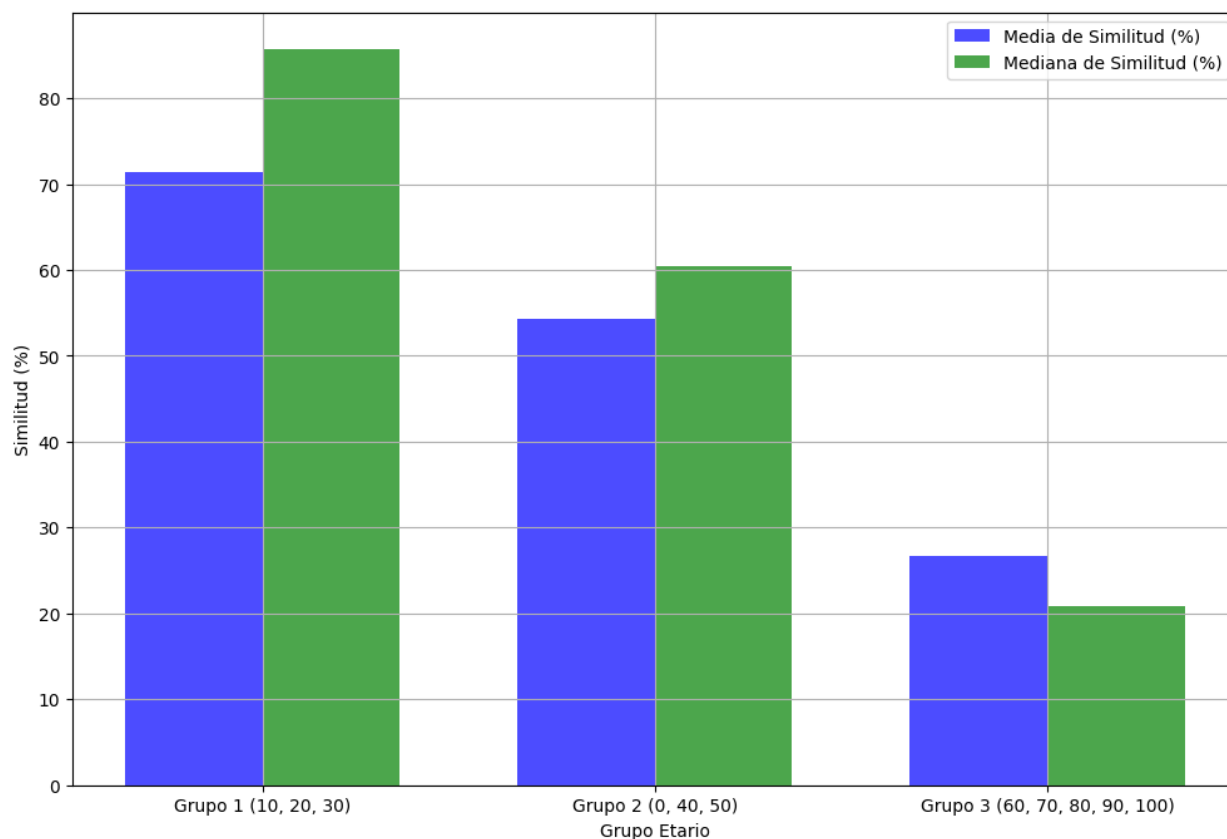


Figura 21. Índice de similitud, media y mediana por grupo etario

Las medias más altas son dadas en los grupos etarios 10, 20 y 30 años (69.2, 75.8 y 69.1, respectivamente). Esto sugiere que las imágenes generadas en estos son, en promedio, mayormente similares a las originales. Pudiendo indicar que el modelo de generación de imágenes funciona mejor preservando la identidad facial durante estos periodos de edad joven y media. Por otra parte, el valor de las medianas (81.1, 90.1 y 84.5) indican que la mayoría de las imágenes generadas en este grupo son altamente similares a las originales para este grupo etario. Y, por último, la desviación estándar (30.9, 29.6 y 33.5) indica una variabilidad razonable en los resultados, pero en general, las imágenes tienen a ser consistentemente similares a las originales.

Los grupos etarios siguientes con una media moderada son 0, 40 Y 50 años (52.4, 60.2 y 50.3, respectivamente) Estos valores sugieren que las imágenes generadas para estos grupos tienen similitud moderada con las originales. El modelo aún preserva ciertas características faciales, pero con menos precisión que aquellas obtenidas de los más jóvenes. También, la mediana moderada (55.9, 71.2 y 53.5) indica que las imágenes generadas son similares en al menos la mitad de los

casos, relativamente menor que en los grupos jóvenes. Por último, la desviación estándar (33.9, 35.3 y 35.5) muestra que existe una variabilidad más alta en los resultados, es decir, el modelo es menos consistente en estos grupos etarios.

Los grupos etarios con baja similitud son los restantes de los experimentos anteriores, 60, 70, 80, 90 y 100 años (42.2, 35.1, 25.2, 17.4 y 13.9, respectivamente). Los valores de media sugieren que las imágenes generadas para 60 años y más son, en promedio, menos similares a las imágenes originales. Pudiendo indicar que el modelo tiene dificultad en preservar la identidad a medida que las personas envejecen. Por otra parte, la mediana baja (41.3, 32.1, 20.0, 0.0 y 0.0) se traduce en que las imágenes generadas tienen muy poca similitud con las originales. La desviación estándar (34.0, 32.8, 29.8, 25.6 y 23.4) indica que hay consistencia en los índices de similitud obtenidos.

En general, se observó una disminución de la media conforme aumenta la edad del grupo etario objetivo. Esto puede reflejar la dificultad que presenta el modelo para capturar y preservar características faciales a medida que debe reflejar mayor edad. Y, la variabilidad en la desviación estándar relativamente alta en varios grupos indica que el modelo presenta algunas inconsistencias, pudiendo deberse a las diversas condiciones de cada fotografía (pose, iluminación, ángulo).

En la Figura 22 se muestra la evolución de la similitud en los principales tres grupos etarios, con los estadísticos de media, mediana y desviación estándar.

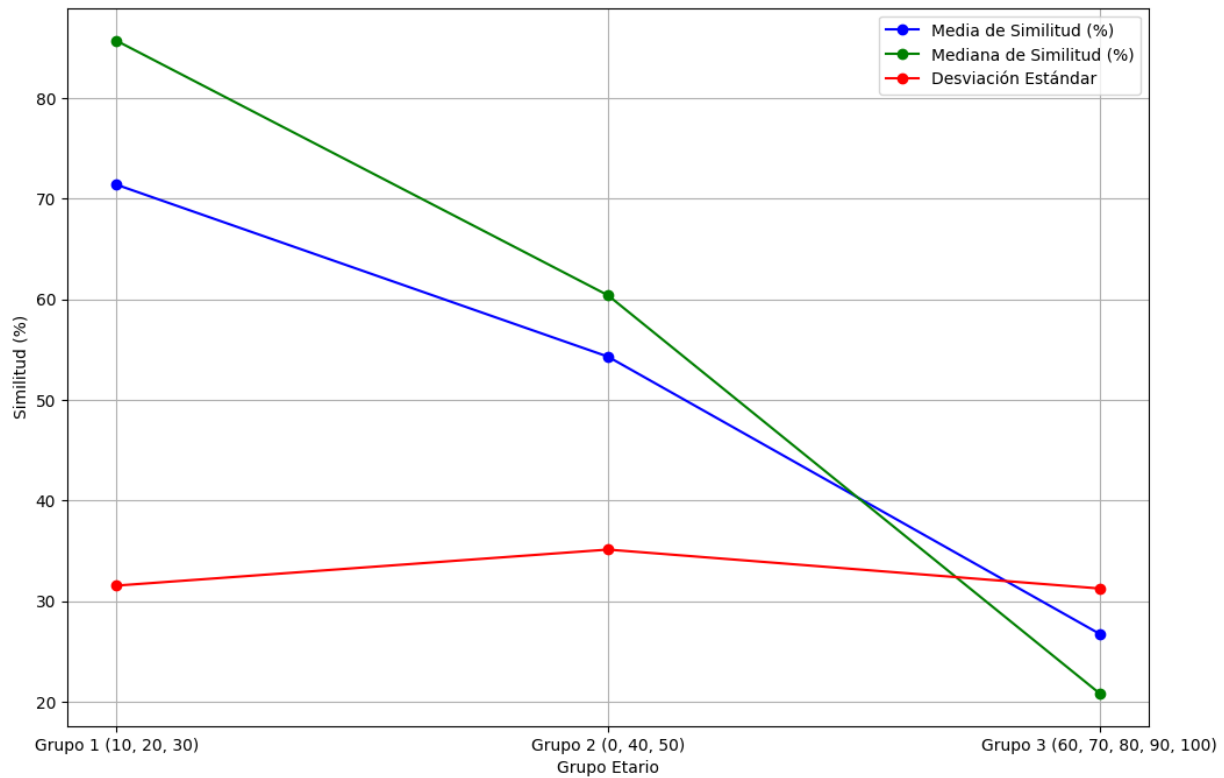


Figura 22. Índice de similitud, media, mediana y desviación estándar por grupo etario

El modelo de generación de imágenes parece funcionar mejor para edades jóvenes y medias, mientras su rendimiento disminuye conforme se acerca a edades más avanzadas. También se observa una considerable variabilidad en cómo se generan las imágenes, lo cual puede ser normal dada la naturaleza impredecible de las GAN. La similitud demuestra que, en las edades de 10, 20 y 30 años el modelo presenta mejor precisión, lo cual se traduce en una ventaja para la temática de mujeres desaparecidas que pueden llevar años sin ser encontradas. Apoyando las estadísticas de que la mayoría son jóvenes (14-17 años), como se vio en [2].

Capítulo VI. Conclusiones y trabajo futuro

En el presente capítulo se describen las conclusiones a partir de los resultados del proyecto, así como trabajos futuros que surgieron durante el desarrollo de la implementación de modelos.

6.1. Conclusiones

En la sección 3.2. se presentó una arquitectura GAN para generación de rostros artificiales. Para su entrenamiento es utilizado el conjunto IMDB-Wiki, extraído de [24]. Se tomó como un primer acercamiento a este tipo de modelos con dichos objetivos y representó un nivel de complejidad alto, significando en un coste de tiempo y recursos muy alto. El siguiente modelo fue uno dedicado a la generación de rostros con envejecimiento implementando una cGAN [13]. Se hicieron variedad de experimentos para mejorar la calidad de los resultados, sin embargo, no se construyó ningún rostro identificable, como se puede apreciar en la sección 4.3.1. Al avanzar en investigación, se optó por utilizar el modelo SAM [31], que implementa algunos modelos pre entrenados, como *StyleGAN2*, y agrega un *encoder*. Es entrenado con el conjunto de imágenes *FFHQ-Aging*, extraído de [35]. Sus características y modelado se describen en la sección 4.3.2.

Para evaluar la precisión y similitud del modelo en la generación de rostros envejecidos, se seleccionó la herramienta *Amazon Rekognition*, específicamente la función '*compare_faces*', ya que esta proporciona una solución robusta y escalable para comparación de rostros. Lo cual es de alto valor en validar que las imágenes generadas mantienen la identidad de la mujer mientras se reflejan cambios etarios. Los resultados indican que el modelo es una estrategia eficaz para la generación de rostros con envejecimiento en los grupos de edad jóvenes y medios.

El proceso de investigación y codificación fue muy prolongado, pero conllevó un amplio aprendizaje para comprender desde la introducción de una GAN hasta la cantidad de parámetros, capas, incluso redes y modelos de los que pueden estar compuestas. Requieren una cantidad muy importante de tiempo de entrenamiento para poder generar imágenes de buena calidad. Sin mencionar el poder de cómputo que se necesita para ejecutar las pruebas y llegar a un modelo poderoso. El tiempo que tarda en entrenar una GAN varía considerablemente dependiendo de varios factores:

- Complejidad del modelo: Entre más parámetros o más profunda sea la arquitectura suele requerir más tiempo de entrenamiento.
- Tamaño del *dataset*: Entre más elementos contenga tomará más tiempo para pasar por cada época de entrenamiento.

- Calidad de los datos: Los datos que requieren menos preprocesamiento y están correctamente etiquetados pueden significar en un aceleramiento del proceso de entrenamiento.
- Capacidad computacional: El *hardware* que se utilice (CPU, GPU, TPU) afecta significativamente el tiempo de entrenamiento.
- Cantidad de épocas; El número de veces que el algoritmo trabajará a través del conjunto de datos completo. Más épocas se traducen en más tiempo, aunque después de cierto punto puede que comience un decrecimiento en términos de mejora del modelo.
- Optimización y técnicas de entrenamiento: El uso de técnicas de optimización y una buena selección de tasa de aprendizaje puede influir en la velocidad de convergencia del modelo.
- Tamaño del *batch*: Un tamaño de *batch* más grande puede acelerar el entrenamiento, pero también puede requerir más memoria y puede afectar la calidad del modelo entrenado.

6.2.Trabajos a futuro

Como trabajos a futuro se plantean las siguientes tareas:

- *Fine-tuning*: los resultados de la evaluación del modelo SAM indican una alta precisión en la generación de imágenes envejecidas para edades jóvenes y medias. Se propone la posibilidad de entrenarla con un conjunto de imágenes de mujeres mexicanas/latinas, para ver si logra capturar los cambios que, se asume, son diferentes por especificar la etnia.
- Agregar un discriminador: para evaluar los cambios generados y probar si pueden ser mejorados para cubrir todos los grupos etarios, quizás mejorar algunos.
- Integración con organizaciones de búsqueda: una dirección que podría tomar la investigación es la implementación de este modelo, desarrollo de interfaz de usuario y bases de datos, para las autoridades. Esto podría resultar en prueba de la eficacia de la herramienta en casos reales y su impacto en la localización de mujeres desaparecidas.
- Estudio de las implicaciones éticas y sociales: abordar la percepción pública del uso de tecnologías AI para la generación de envejecimiento facial, así como estudiar el impacto de esto en la privacidad y confianza pública.

- Evaluación de expertos: se considera la participación de un experto en el área de morfología forense, como evaluación cualitativa del modelo. Siendo importante la participación humana en una problemática del contexto social.

Referencias

- [1] M. G. Álvarez Gutiérrez and J. G. Castillo Koschnick, “Panorama estadístico de la violencia contra niñas, niños y adolescentes en México,” *Unicef*, p. 106, 2019.
- [2] Segob, “Estadística del RNPDO por filtros,” 2021.
<https://versionpublicarnpdno.segob.gob.mx/Dashboard/Sociodemografico>.
- [3] I. Goodfellow *et al.*, “Generative adversarial networks,” *Commun. ACM*, vol. 63, no. 11, pp. 139–144, 2014, doi: 10.1145/3422622.
- [4] X. Wang, K. Wang, and S. Lian, “A Survey on Face Data Augmentation,” no. 1, 2019, doi: 10.1007/s00521-020-04748-3.
- [5] S. de Gobernación, “Versión Pública RNPDO,” 2021.
<https://versionpublicarnpdno.segob.gob.mx/Dashboard/ContextoGeneral>.
- [6] C. N. de B. (CNB), “Búsqueda de personas desaparecidas con perspectiva de género.”
<https://comisionacionaldebusqueda.gob.mx/busqueda/>.
- [7] C. N. de B. (CNB), “Acciones de búsqueda.”
<https://comisionacionaldebusqueda.gob.mx/busqueda/>.
- [8] S. de G. (SEGOB) and C. N. de B. (CNB), “Inicio de Reporte de Persona Desaparecida o No Localizada,” 2021. <https://cnbreportadesaparecidos.segob.gob.mx/Pages/Procedimiento.aspx>.
- [9] AWS, “Amazon Rekognition.” <https://aws.amazon.com/es/rekognition/?nc=sn&loc=0>.
- [10] H. Galán and A. Martínez, “Inteligencia artificial . Redes neuronales y Aplicaciones,” *Univ. Carlos III Madrid, J.*, p. 8, 2015.
- [11] I. Masi, A. T. Trãn, T. Hassner, J. T. Leksut, and G. Medioni, “Do we really need to collect millions of faces for effective face recognition?,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9909 LNCS, pp. 579–596, 2016, doi: 10.1007/978-3-319-46454-1_35.
- [12] J. J. Lv, X. H. Shao, J. S. Huang, X. D. Zhou, and X. Zhou, “Data augmentation for face recognition,” *Neurocomputing*, vol. 230, pp. 184–196, 2017, doi: 10.1016/j.neucom.2016.12.025.
- [13] G. Perarnau, J. van de Weijer, B. Raducanu, and J. M. Álvarez, “Invertible Conditional GANs for image editing,” no. Figure 1, pp. 1–9, 2016.

- [14] M. Sajid *et al.*, “Data Augmentation-Assisted Makeup-Invariant Face Recognition,” *Math. Probl. Eng.*, vol. 2018, 2018, doi: 10.1155/2018/2850632.
- [15] B. Leng, K. Yu, and J. QIN, “Data augmentation for unbalanced face recognition training sets,” *Neurocomputing*, vol. 235, pp. 10–14, 2017, doi: 10.1016/j.neucom.2016.12.013.
- [16] O. M. México, “Búsqueda de Personas Desaparecidas con Perspectiva de Género,” 2021. <https://mexico.unwomen.org/es/digiteca/publicaciones/2021/05/busqueda-de-personas-desaparecidas-con-perspectiva-de-genero>.
- [17] M. de Lourdes Velasco-Domínguez and S. Castañeda-Xochitl, “Disappearance of women and girls in Mexico: Contributions of feminisms for understanding macrosocial processes,” *Iconos*, vol. XXIV, no. 67, pp. 95–117, 2020, doi: 10.17141/iconos.67.2020.4196.
- [18] K. Salazar Serna and R. Castro, “Nuestras hijas no volvieron a casa. Caminos de reconstrucción de mujeres que buscan mujeres desaparecidas,” *Rev. Via Iuris*, no. 30, pp. 1–35, 2021, doi: 10.37511/viaiuris.n30a7.
- [19] Tensorflow, “Red antagónica generativa convolucional profunda,” 2022. <https://www.tensorflow.org/tutorials/generative/dcgan?hl=es-419#:~:text=You will use the MNIST,1%2C 1>.
- [20] J. Zhang, K. Jia, J. Jia, and Y. Qian, “Improved Techniques for Training GANs,” *BMC Bioinformatics*, vol. 19, no. 1, pp. 1–9, 2018.
- [21] Y. Yu, W. Zhang, and Y. Deng, “Frechet inception distance (fid) for evaluating gans,” *Researchgate.Net*, no. September, pp. 1–7, 2021.
- [22] A. Borji, “Pros and cons of GAN evaluation measures: New developments,” *Comput. Vis. Image Underst.*, vol. 215, 2022, doi: 10.1016/j.cviu.2021.103329.
- [23] E. Betzalel, C. Penso, A. Navon, and E. Fetaya, “A Study on the Evaluation of Generative Models,” no. ii, 2022.
- [24] R. Rothe, R. Timofte, and L. Van Gool, “Deep expectation of real and apparent age from a single image without facial landmarks,” *Int. J. Comput. Vis.*, vol. 126, no. 2–4, pp. 144–157, 2018.
- [25] Kaggle, “Celeba dataset.” <https://www.kaggle.com/datasets/jessicali9530/celeba-dataset>.
- [26] E. Mohamed, O. Elharrouss, S. Al-Maadeed, and H. Tairi, “Image Generation: A Review,” *Neural Process. Lett.*, vol. 54, no. 5, pp. 4609–4646, 2022, doi: <https://doi.org/10.1007/s11063-022->

10777-x.

- [27] K. Wang, C. Gou, Y. Duan, Y. Lin, X. Zheng, and F. Y. Wang, “Generative adversarial networks: Introduction and outlook,” *IEEE/CAA J. Autom. Sin.*, vol. 4, no. 4, pp. 588–598, 2017, doi: 10.1109/JAS.2017.7510583.
- [28] N. Sharma, R. Sharma, and N. Jindal, “Face-Based Age and Gender Estimation Using Improved Convolutional Neural Network Approach,” *Wirel. Pers. Commun.*, vol. 124, no. 4, pp. 3035–3054, Jun. 2022, doi: 10.1007/s11277-022-09501-8.
- [29] V. Sheoran, S. Joshi, and T. R. Bhayani, “Age and Gender Prediction using Deep CNNs and Transfer Learning,” pp. 1–12.
- [30] M. A. H. Akhand, I. Sayim, S. Roy, and N. Siddique, “Human Age Prediction from Facial Image Using Transfer Learning in Deep Convolutional Neural Networks Human Age Prediction from Facial Image Using Transfer Learning in Deep,” no. May, 2020, doi: 10.1007/978-981-15-3607-6.
- [31] Y. Alaluf, O. Patashnik, and D. Cohen-Or, “Only a matter of style: age transformation using a style-based regression model,” *ACM Trans. Graph.*, vol. 40, no. 4, Jul. 2021, doi: 10.1145/3450626.3459805.
- [32] D. E. King, “Dlib-ml: A Machine Learning Toolkit,” *J. Mach. Learn. Res.*, vol. 10, pp. 1755–1758, Dec. 2009.
- [33] E. Richardson *et al.*, “Encoding in Style: a StyleGAN Encoder for Image-to-Image Translation.” 2021.
- [34] G. Zarate, L. Cruz-Reyes, B. Dorronsoro, and A. Rosete-Suárez, *Data Analytics and Computational Intelligence: Novel Models, Algorithms and Applications*. 2023.
- [35] R. Or-El, S. Sengupta, O. Fried, E. Shechtman, and I. Kemelmacher-Shlizerman, “Lifespan Age Transformation Synthesis.” 2020.

Anexos

Anexo I. Participación como ponente en 1er coloquio de posgrados IIT



1^{ER} COLOQUIO
DE POSGRADOS IIT

Coordinación de Apoyo al Desarrollo de la
Investigación
y el **Posgrado**
• IIT



El Instituto de Ingeniería y Tecnología
a través de la
Coordinación de Apoyo al Desarrollo de la Investigación y el Posgrado IIT
otorga la presente

CONSTANCIA

a:

BLANCA ELENA CAZARES SALAIS

por su participación como ponente del cartel

**“Aplicación de técnicas de aprendizaje profundo para generar imágenes
artificiales de mujeres desaparecidas”**

dentro del marco del Primer Coloquio de Posgrados IIT

“Por una vida científica, por una ciencia vital”

Ciudad Juárez, Chihuahua, 20 y 21 de mayo de 2021


Dra. Nelly Gordillo Castillo
Coordinadora CADIP IIT

Anexo II. Participación como ponente en 2do coloquio de posgrados IIT

UACJ

DO
2 COLOQUIO IIT
DE POSGRADOS
11 Y 12 DE NOVIEMBRE 2021
Investigación
y Posgrado
IIT

El Instituto de Ingeniería y Tecnología
a través de la
Coordinación de Apoyo al Desarrollo de la Investigación y el Posgrado
otorga el presente

RECONOCIMIENTO

a

Blanca Elena Cázares Salais

por su participación como **Ponente**
dentro del 2do Coloquio de Posgrados del IIT



Nelly Gordillo Castillo
Coordinadora de Investigación y Posgrado del IIT

Anexo III. Participación como ponente en 3er coloquio de posgrados IIT



El Instituto de Ingeniería y Tecnología
a través de la
Coordinación de Apoyo al Desarrollo de la Investigación y el Posgrado
otorga el presente

RECONOCIMIENTO

a

Blanca Elena Cazares Saláis

por su valiosa participación como **ponente** dentro del 3er Coloquio de Posgrados del IIT, llevado a cabo el 26 y 27 de mayo de 2022.

"POR UNA VIDA CIENTIFICA, POR UNA
CIENCIA VITAL"

Dra. Nelly Gordillo Castillo
Coordinadora de Apoyo al Desarrollo de la
Investigación y el Posgrado del IIT

Cd. Juárez, Chih., México, a 30 de mayo de 2022



Anexo IV. Participación como ponente en 4to coloquio de posgrados IIT



El Instituto de Ingeniería y Tecnología
a través de la
Coordinación de Apoyo al Desarrollo de la Investigación y el Posgrado IIT
otorga la presente

CONSTANCIA

a:

Blanca Elena Cázares Salais

por su participación como ponente en el

Cuarto Coloquio de Posgrados IIT

“Por una vida científica, por una ciencia vital”

Ciudad Juárez, Chihuahua, 22 al 25 de noviembre de 2022



Dra. Nelly Gordillo Castillo
Coordinadora CADIP IIT

Anexo V. Participación en panel Mujeres In



Anexo VI. Capítulo publicado en Data Analytics & Computational Intelligence: Novel Models, Algorithms and Applications

The Eureka Community

May 31, 2023

Who may concern:

By this letter, the editors certify that the acceptance of the following chapter was the result of a double-blind peer-review process:

Chapter title: Data augmentation techniques for facial image generation: a brief literature review
Authors: Blanca Elena Cazares Salais, Rogelio Florencia Juárez, Vicente García Jiménez and Julia Patricia Sánchez Solís
Book title: Data Analytics & Computational Intelligence: Novel Models, Algorithms and Applications
Editors: Laura Cruz-Reyes, Bernabé Dorronsoro, Gilberto Rivera, and Alejandro Rosete
Book series: Studies in Big Data (Springer)

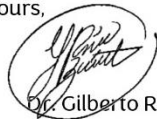
Furthermore, the editorial review process complied with the publishing agreement stated in the Springer Contract # 145403. *Studies in Big Data* is currently indexed in SCOPUS, SCImago, and EI Compindex.

This contributed book was an editorial initiative of the Eureka Community. Eureka is an international and multidisciplinary scientific research network that joins professionals in mathematics, computer sciences, engineering, administration, economics, and social sciences. It was founded in 2008 and is currently integrating more than 60 research groups in more than 20 countries, mainly in America and Europe. The submitted chapters were accepted only after a stringent review process by our collaborators worldwide, coordinated by the editors.

As evidence, the following documents are enclosed: (i) Initial version of the manuscript, (ii) review report, (iii) revision notes, (iv) revised manuscript, (v) decision letter, and (vi) preprint.

Please, do not hesitate to contact me with any doubts or questions regarding this letter.

Sincerely yours,



Dr. Gilberto Rivera

VICE PRESIDENT OF THE EUREKAS COMMUNITY
COORDINATOR OF PUBLICATION PROJECTS

with the editors' approval,



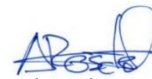
Dr. Laura Cruz-Reyes

TECNOLÓGICO NACIONAL DE
MÉXICO (MÉXICO)



Dr. Bernabé Dorronsoro

UNIVERSIDAD DE CÁDIZ
(SPAIN)



Dr. Alejandro Rosete

UNIVERSIDAD TECNOLÓGICA DE LA HABANA
"JOSÉ ANTONIO ECHEVERRÍA" (CUBA)