

Evaluación de rostros faciales para la expedición de documentos de identificación.

Fernando Broncano Morgado
Universidad de Extremadura
Escuela Politécnica de Cáceres
fbroncan@alumnos.unex.es

Victoria Amores Chaparro
Universidad de Extremadura
Escuela Politécnica de Cáceres
kamoresc@alumnos.unex.es

Óscar Mogollón Gutiérrez
Universidad de Extremadura
Escuela Politécnica de Cáceres
oscarimg@unex.es

Alberto López Trigo
Universidad de Extremadura
Escuela Politécnica de Cáceres
albertolt@unex.es

José Carlos Sancho Núñez
Universidad de Extremadura
Centro Universitario de Mérida
jcsanchon@unex.es

Resumen—La evaluación de rostros humanos es una cuestión que ha estado presente desde la introducción de la identificación personal. Cualquier órgano acreditador de identidad exige una imagen facial para facilitar el reconocimiento de las personas y la validación documental. La democratización del uso de las nuevas tecnologías ha digitalizado los procesos de identificación, pues se han propuesto diferentes modelos de reconocimiento, clasificación y validación, ya que acreditar la identidad mediante elementos fotográficos vía internet es una acción cada vez más común. Con el fin de garantizar la calidad de los rostros usados en la identificación, se propone un proceso de evaluación basado en tres etapas, seis dimensiones de clasificación, y acorde a diferentes características basadas en los estándares internacionales y normativas nacionales. El proceso de evaluación consiste en la extracción de métricas de diferentes puntos faciales, su posterior análisis, y en función de su cumplimiento, proporcionar una puntuación 0 – 100.

Index Terms—Evaluación, Rostros Faciales, Documentos de Identidad, Calidad de Caras, Reconocimiento Facial.

I. INTRODUCCIÓN

Uno de los sistemas que tradicionalmente se ha empleado para validar y reconocer la identidad de las personas son las fotos de las caras de las personas. La mayoría de documentos de identificación personales contienen una imagen facial del usuario propietario del documento, pudiendo validar por analogía la persona portadora del documento con los datos del sujeto recogido en este. Actualmente, muchos dispositivos electrónicos emplean el reconocimiento facial como principal elemento biométrico de seguridad.

Los esfuerzos de la Unión Europea van en la dirección de conseguir una Identidad Digital Única [1], pues esto implica que toda persona que pueda optar a un documento nacional de identidad, obtenga una identidad digital reconocida en cualquier lugar de la Unión Europea. Esta identidad única europea conseguiría una mejora en los servicios sanitarios y de la administración recibidos por los ciudadanos. Por ello, es fundamental que los trabajos vayan dirigidos en conseguir una calidad óptima en los rostros que identifiquen a los sujetos.

Por otra parte, los avances en materia de inteligencia artificial han conseguido que se puedan construir diferentes fotografías y vídeos en base a una única foto personal. Por ello, cada vez está cobrando más relevancia el análisis de

la calidad de los rostros faciales, recogidos en varios estándares internacionales propuestos por el *National Institute of Standards and Technology* [2], o el *International Organization of Standardization*. Asimismo, los organismos expedidores de documentos de identidad, siempre han solicitado unos requisitos mínimos de calidad en las fotografías de rostros empleadas para la expedición de estos documentos, como por ejemplo el Reino de España y el Real Decreto 1553/2005 [3].

De misma forma, los avances en las redes generativas [4] permiten a los usuarios generar rostros sintéticos, pero estos no suelen cumplir determinados estándares marcados para fotografías que identifiquen a los individuos, pues en ocasiones llevan gafas de sol, tienen determinadas zonas de la cabeza cortada, o no tienen el rostro completamente frontal a la toma de la captura. En este sentido, para usar estas imágenes en documentos de identidad, han de ser evaluadas para comprobar su grado de cumplimiento.

La principal contribución de este trabajo se centra en generar un proceso de evaluación basado en métricas faciales para obtener el grado de cumplimiento de la captura para la generación de un documento nacional de identidad. El proceso consta de tres fases: extracción de métricas faciales, análisis de métricas y evaluación final de puntuaciones.

Esta contribución permite añadir este proceso a uno superior que emplee el uso de imágenes de personas en la generación sintética de otro tipo de documentos, como Documentos Nacionales de Identidad –DNI–, o carnés de conducir.

II. TRABAJOS RELACIONADOS

Desde la democratización y aumento del uso de las inteligencias artificiales generativas, la creación sintética de rostros humanos también ha ido en aumento. Inicialmente, los modelos solo eran capaces de generar imágenes estáticas de rostros humanos. Pero con el incremento en la capacidad de cómputo, la obtención de modelos más eficientes y el avance tecnológico, se ha conseguido que estos modelos sean capaces de generar vídeos de rostros sintéticos que contengan guiños, cabeceos y movimiento facial acompañando a la voz, partiendo tan solo de la imagen de un rostro.

En 2014, Taigman et al. [5], desde la corporación Meta, anteriormente Facebook, publicaron un proceso que permitía

el reconocimiento facial aplicado en cuatro etapas: la detección, la alineación, la representación y la clasificación. Este modelo, basándose en diferentes imágenes de un rostro, es capaz de determinar si una nueva imagen corresponde a la persona ya clasificada o a una nueva con una precisión del 97,35 % sobre el conjunto de datos *Labeled Faces in the Wild* [6], lo que se aproxima al rendimiento humano.

Un año después, en 2015, Schroff et al. [7] publican el modelo *FaceNet*. Este modelo permite la verificación y reconocimiento facial. Para ello, aprende directamente un mapeo de imágenes faciales a un espacio compacto de Euclides, donde las distancias reflejan la similitud entre rostros. Este enfoque permite la implementación de la verificación, el reconocimiento y la agrupación de rostros. Este modelo emplea una red convolucional profunda entrenada. Logra un rendimiento de 99,63 % sobre el conjunto de datos *Labeled Faces in Wild* [6].

La técnica más comunmente empleada en la actualidad en el reconocimiento de rostros faciales humanos es *ArcFace*. Deng et al. [8] presentan un método novedoso para el reconocimiento facial. Los resultados experimentales destacan la robustez y precisión de esta técnica, especialmente en condiciones de datos ruidosos, y su capacidad para mantener un rendimiento competitivo frente a métodos recientes de reconocimiento facial.

En base a los modelos que permiten el reconocimiento, la verificación y la clasificación, se ha ido avanzando en la creación de nuevos conjuntos de datos que permiten detectar si se ha generado un vídeo del rostro de una persona a partir de una única captura de un rostro. Estas actuaciones corresponden con las manipulaciones faciales en vídeos, dentro de las usurpaciones de identidad. Agarwal et al. [9] realizan un modelo para detectar suplantaciones de identidad de figuras públicas analizando patrones de expresión característicos, como los movimientos de cabeza o expresiones faciales. Para ello, aplican técnicas de intercambio de rostros, sincronización de labios y manipulación de gestos. Muestra un resultado más fuerte en comparación con redes neuronales convolucionales.

Rössler et al. [10] presentan un conjunto de datos novedoso para la detección de manipulaciones faciales en vídeos. Este conjunto de datos incluye más de 500,000 imágenes editadas provenientes de 1004 vídeos. Las manipulaciones se generaron utilizando la técnica avanzada *Face2Face* [11], permitiendo transferir expresiones faciales de un vídeo fuente a uno objetivo, recreando expresiones faciales. También, se aportan resultados para tareas forenses clásicas como la clasificación y segmentación de imágenes manipuladas, evaluando su rendimiento en vídeos comprimidos a diferentes niveles de calidad. Además, el conjunto de datos permite evaluar métodos generativos refinados en un entorno supervisado, proporcionando una base sólida para el desarrollo y evaluación de algoritmos de detección de manipulaciones faciales en un contexto realista.

La generación sintética de rostros humanos es una cuestión vinculada con el reconocimiento, la verificación y la clasificación de estos. Uno de los nuevos paradigmas a clasificar es la identificación de si un rostro ha sido generado con inteligencia artificial o corresponde con una toma real. Karras et al. [4] mejoran significativamente el modelo generativo de imágenes

StyleGAN solucionando problemas relacionados con la calidad de las imágenes y sus limitaciones.

Con ello, varios autores han ido describiendo el problema que supone las suplantaciones de identidad. Estas suplantaciones de identidad van desde la transferencia e intercambio de rostros a la generación de pruebas de vida en base a una fotografía dada. Es vital que el reconocimiento de rostros faciales permita diferenciar entre pruebas de vida reales y generadas por computador.

Yuezun et al. [12] proponen un método para la detección de suplantaciones de identidad basada en la ausencia de parpadeos en los vídeos generados por modelos de inteligencia artificial, pues estos modelos son entrenados con ausencia de imágenes de ojos cerrados. Por otra parte, Yang et al. [13] proponen una alternativa basada en la posición de la cabeza para comprobar si las imágenes son generadas por un modelo generativo. Nanda et al. [14] estudian la necesidad de implementar sistemas de prueba de identidad con altos niveles de seguridad frente a ataques de presentación, como la manipulación y suplantación.

La literatura ha sugerido diferentes alternativas a la clasificación, reconocimiento y validación de rostros, y cada vez resulta más relevante las cuestiones acerca de la consideración de si los rostros faciales cumplen determinadas condiciones como su veracidad humana.

III. DEFINICIÓN DE LA EVALUACIÓN

La evaluación de rostros humanos es una dimensión poco explorada en la literatura, pues los esfuerzos se concentran en el reconocimiento, validación y clasificación de rostros. A este respecto, el proceso de evaluación propuesto realiza una estimación de diferentes características faciales para estudiar el cumplimiento de diferentes estándares de calidad y legales exigidos para las fotografías requeridas en la expedición de documentos de identidad.

III-A. Evaluación de la calidad del rostro

La primera dimensión que debe explorar el algoritmo de evaluación de imágenes faciales es la calidad de la fotografía y del rostro.

El Instituto Nacional de Estándares y Tecnología de Estados Unidos de América –NIST– define el estándar de evaluación de la tecnología de análisis facial –FATE– [15] que contiene una serie de documentos relacionados con la tecnología de análisis facial. El NIST define un documento conceptual con las características básicas para la detección de defectos en imágenes [2], así como, una evaluación de diferentes algoritmos en función del cumplimiento de estas características [16]. La norma ISO/IEC 29794-5: *Information technology, Biometric sample quality, Part 5: Face image data* [17] define una serie de características relacionadas con la toma de la imagen y del sujeto para verificar la calidad de una imagen facial. El conjunto de características a emplear en el algoritmo propuesto viene nutrido de ambos estándares.

Las características a evaluar pueden dividirse en dos grupos: características acerca del sujeto, y características acerca de la captura. El primer grupo viene definido por ocho características extraídas de las normas de estandarización.

- Número de caras. En la imagen facial solo ha de contar con una rostro reconocible en la captura.

- Pose de la cabeza. La imagen del rostro facial debe ser de frente y centrada en la captura, siendo los ángulos de rotación los mínimos posibles.
- Visibilidad de los ojos. El rostro no puede portar ningún tipo de elemento que comprometa el reconocimiento de los ojos, como gafas de sol.
- Apertura de la boca. La cara contenida en la captura debe minimizar la apertura de la boca, priorizando aquellas imágenes que tengan bocas cerradas sobre aquellas que tienen bocas abiertas.
- Oclusión facial. Todas las zonas del rostro deben estar al descubierto, sin que cualquier tipo de prenda o corte de pelo pueda disminuir la visibilidad completa de la cara.
- Apertura de los ojos. El rostro debe contar con los ojos abiertos, que puedan ser bien identificados en la fotografía.
- Distancia entre ojos. Deberá evaluarse la distancia entre ojos para estimar la resolución de la toma y el tamaño del rostro dentro de la imagen.
- Distancia de los ojos a los bordes. Las medidas entre los ojos y los bordes de la imagen permiten comprobar la equidistancia del rostro a los bordes de la captura. De la misma forma, deben usarse dichas medidas para evaluar si la captura corta determinadas zonas del rostro.

El segundo grupo de características a evaluar serán aquellas que guardan relación con la captura del rostro facial.

- Resolución y nitidez. Las dimensiones de la captura deben ser suficientes como para que las diferentes zonas del rostro sean nítidas y contar con el detalle suficiente.
- Fondo. La evaluación del fondo de la captura debe conseguir facilitar la distinción entre el rostro humano del contexto de la imagen.
- Exposición. La cantidad de luz contenida en la captura debe ser suficiente como para permitir la distinción de las zonas faciales, pudiendo apreciar correctamente los detalles.
- Ruido. Han de evaluarse los posibles efectos de aplicación de difuminados, patrones de *moiré* o artefactos de compresión de la captura.

III-B. Adecuación a los requisitos legales

Para la expedición de diferentes documentos de identificación, los estados expedidores regulan unos requisitos que deben cumplir las fotografías de los rostros faciales para la identificación personal en estos documentos. En España, el RD1553/2005 [3] regula los requisitos necesarios para la expedición del documento nacional de identidad. Concretamente, el artículo 5.1. de dicho Real Decreto recoge los requisitos relativos a la documentación a aportar.

La normativa dicta el fondo que ha de tener la captura, *fondo uniforme blanco y liso*; la pose de la cabeza, *tomada de frente*; la oclusión facial, *con la cabeza totalmente descubierta*; la visibilidad de los ojos, *sin gafas de cristales oscuros o cualquier otra prenda que impida la identificación*; así como, la resolución, nitidez y exposición, *fotografía reciente en color del rostro del solicitante*. Todas las cuestiones indicadas por el Real Decreto son estudiadas en base a las definiciones de los estándares para la calidad de imágenes faciales.

La presentación sobre el documento nacional de identidad se realiza con el rostro de la persona en escala de grises y con

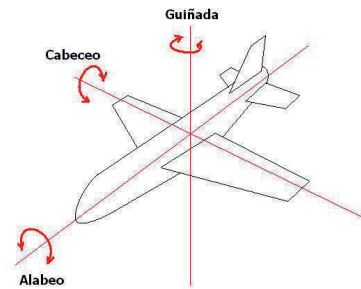


Figura 1. Ángulos de giro

un pequeño posprocesado. Esta dimensión consiste en aplicar detección del rostro sobre la captura, eliminar cualquier tipo de fondo presente fuera del rostro, y aplicar un filtro en escala de grises.

IV. IMPLEMENTACIÓN DEL ALGORITMO DE EVALUACIÓN

El algoritmo de evaluación propuesto en este trabajo permite analizar los rostros reconocidos en fotografías y estimar su calidad en función de las características establecidas en la Sección III. Este devuelve la puntuación asociada al rostro encontrado en la captura. La puntuación se encuentra entre el rango 0–100 puntos, siendo la puntuación 100 la imagen más fiable y de mayor calidad para ser empleada en documentos de identidad.

El proceso algorítmico se constituye en tres fases: extracción de métricas, análisis de las métricas y evaluación conforme al análisis. La extracción de métricas se realiza mediante una serie de técnicas que permiten medir diferentes características faciales acorde a las características presentadas en la sección III. Se toman seis dimensiones para representar las doce características consideradas en la evaluación de calidad. Para cada una de ellas realizará una serie de medidas sobre el rostro, estima los resultados obtenidos con una métrica, y en base a ese dato se obtiene una puntuación.

IV-A. Extracción de métricas

En primer lugar, el algoritmo emplea el uso de la librería *TensorFlow* [18] y el modelo *MTCNN* [19] [20] para detectar los rostros que aparecen en la imagen. Se le indica al modelo que los rostros deben tener un tamaño mínimo, con el fin de filtrar aquellos que aparezcan en el fondo de la imagen. Con la ejecución de este modelo, se devuelve el número de caras, la distancia y el radio de la cara normalizados con respecto a la boca y los ojos, la determinación de si la cara está frontal, la imagen alineada con respecto a los puntos que contienen exclusivamente la cara, y varios puntos claves de la cara, entre ellos: los puntos izquierdo y derecho de la boca, los centros de los ojos derecho e izquierdo y el centro de la nariz.

Una vez obtenidos los diferentes puntos de referencia faciales, se estiman los ángulos de la cara, siendo estos el balanceo –rotación sobre eje anterior-posterior de la cabeza–, el cabeceo –rotación sobre eje izquierdo-derecho de la cabeza– y la guiñada –rotación sobre eje superior-inferior de la cabeza–. En la Figura 1 se puede observar los diferentes ángulos de posiciones.

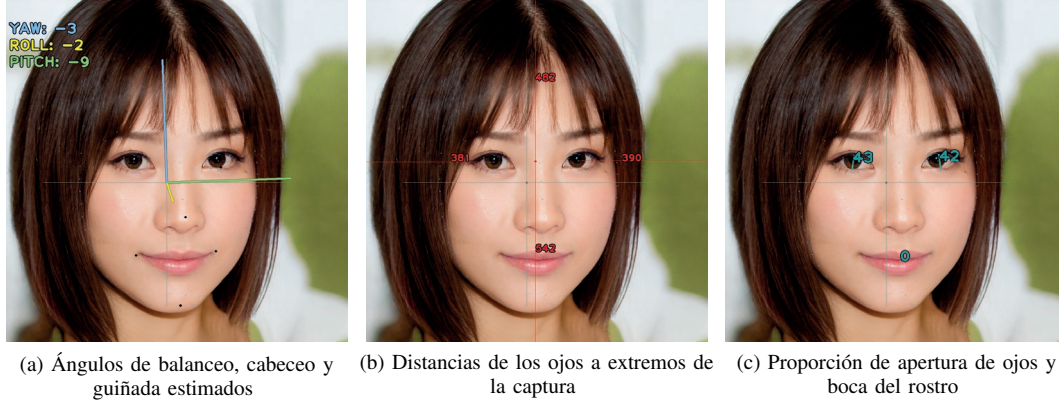


Figura 2. Métricas marcadas sobre el rostro.

El balanceo es calculado como la media del ángulo entre la línea que une ambos ojos y el *eje Y* de la imagen, junto con el ángulo entre la línea vertical de la nariz y el *eje X* de la captura. El cabeceo es calculado mediante la proyección de los puntos de la cara, si estuviera completamente equidistante de los márgenes izquierdo y derecho, sobre la imagen dada como entrada, correspondiendo el ángulo con el ángulo entre los dos planos, el proyectado y el original, empleando una librería para la estimación de la pose [21]. El ángulo de guiñada es calculado entre la perpendicular al rostro y una paralela al *eje Y* de la imagen por el centro de la nariz. En la Figura 2a se pueden observar los tres ángulos medidos para un rostro.

Tras obtener los ángulos de la cara, se analiza si el rostro es portador de gafas de ver, de sol o no lleva gafas. Este algoritmo se apoya la librería *Keras* [22] empleando el modelo entrenado para determinar dicha característica. Este modelo se ha generado a partir de *MobileNetV2* [23], empleando los pesos de *ImageNet* [24] y entrenado con el conjunto de rostros faciales *Glasses or Not Glasses* [25]. Devuelve la probabilidad de que el rostro esté en la clase de rostros sin gafas, rostros con gafas de ver, y rostros con gafas de sol.

La distancia entre los bordes de la captura y los ojos, así como, la distancia entre ojos también son varias de las métricas a extraer. Para la obtención de las primeras, se obtienen los bordes de la imagen y se resta las posiciones de los ojos, mientras que para la segunda se debe calcular la distancia entre los centros de ambos ojos y aplicar el ángulo de guiñada. En la Figura 2b se puede observar la distancia de los ojos a los bordes de la captura.

Por otro lado, se debe comprobar que los ojos y boca estén abiertos. Estas comprobaciones se realizan obteniendo los puntos superiores e inferiores de dichas partes y estimando cual es la proporción que corresponde a esta apertura sobre la cara. En la Figura 2c se puede observar el porcentaje de apertura de ojos y boca.

Para estudiar la exposición, iluminación, contraste y desenfoco se realizan cuatro operaciones sobre la imagen en escala de grises. Para la exposición se obtiene el mayor valor de la escala de grises que tiene presencia en la captura; para la iluminación se realiza la media de todos los valores en escala de grises de la captura; para el contraste se calcula la desviación típica de los valores de los píxeles de la imagen; y, para el desenfoco, se aplica el efecto *Canny* para obtener

TABLA I
RESUMEN DE MÉTRICAS EXTRAÍDAS PARA ANALIZAR.

Métrica	Valor
Ángulos de balanceo, cabeceo y guiñada	Entre -90 y 90
Lleva gafas de sol, de ver o no lleva	Si/No
Distancia a bordes	Valor numérico
Apertura de ojos y boca	Entre 0 y 100
Iluminación media de la imagen	Entre 0 y 255

los bordes y se calcula la media de dichos valores. La Tabla I recoge un resumen de las métricas analizadas.

IV-B. Análisis de métricas

Una vez obtenidas las métricas faciales, se realiza una evaluación de consecución de dichas medidas. Estas métricas son analizadas en seis dimensiones, que posteriormente serán ponderadas en una evaluación final.

IV-B1. Ángulos de balanceo, cabeceo y guiñada: Con los tres ángulos de la pose facial, se estima la puntuación en el rango 0 – 100 conforme a la Ecuación 1.

$$D_1 = \max \left[\left(45 - \frac{|\bar{a}| + \max |a|}{2} \right) \cdot \frac{100}{45}, 0 \right] \quad (1)$$

En la Ecuación 1, $|\bar{a}|$ corresponde con la media de los valores absolutos de los tres ángulos, mientras que $\max(|a|)$ corresponde con el máximo de los valores absolutos de los ángulos. Cuando los tres ángulos de rotación son cero, la puntuación obtenida es 100, correspondiente a que el rostro esté centrado y de frente al objetivo, lo evaluado en el algoritmo. Si el cálculo realizado con los ángulos fuese mayor a 45°, la puntuación es negativa, entendida como 0.

IV-B2. Apertura de ojos: A partir de la proporción de apertura que tiene cada ojo, se realiza un cálculo para otorgar la máxima puntuación a cada ojo cuando estos estén abiertos más de un 20 %. Para cada ojo se efectúa el cálculo de la Ecuación 2.

$$D_2 = \min(a \cdot 500, 100) \quad (2)$$

En la Ecuación 2, el valor a corresponde con la proporción de apertura del ojo. Si el valor a asciende a 0,2, el 20 % de la apertura máxima, la puntuación será máxima. De esta forma se da una mayor evaluación a los ojos abiertos, con lo que

se cumplen los requisitos solicitados. La puntuación final de esta dimensión será la media aritmética de la puntuación para el ojo izquierdo y el ojo derecho.

IV-B3. Apertura de boca: Partiendo de la proporción de apertura de la boca, se realiza un cálculo discriminador, pues la boca debe permanecer cerrada para otorgar la máxima puntuación de la evaluación, manteniendo el óvalo de la cara definido. Para ello, la puntuación será 0 cuando la proporción de apertura de la boca sea 0,25 o superior. La puntuación para esta dimensión es calculada mediante la Ecuación 3.

$$D_3 = \max[(0,25 - a) \cdot 400, 0] \quad (3)$$

En la Ecuación 3, el valor a corresponde con la proporción de apertura de la boca. El resultado de discriminar este valor se multiplica por 400 para llegar a la puntuación 100 cuando a sea cero.

IV-B4. Clasificación en función de gafas: El rostro no debe portar gafas de sol, por ello, de haber sido etiquetado con mayor probabilidad como portador de gafas de sol penalizará el global de la puntuación obtenida por el rostro. Específicamente, la puntuación global se verá disminuida a la mitad en caso de tener esta casuística.

IV-B5. Iluminación de la imagen: La iluminación juega un papel fundamental en la calidad, pues si la luz que contiene es mínima o excesiva no se podrá identificar correctamente al rostro. Para ello, el valor medio de la iluminación de la imagen, calculado en la fase de extracción de métricas, debe estar en torno a 127, valor central de la escala de grises. Se aplica el cálculo de la Ecuación 4 para obtener la puntuación relativa a la iluminación.

$$D_4 = \min(\max(100 - |ilum - 127|, 0), 100) \quad (4)$$

La puntuación obtenida respecto a la iluminación será de 100 cuando los valores de iluminación estén en torno a los valores medios de la escala de grises, y cero cuando tienda la iluminación a los valores extremos.

IV-B6. Definición del rostro dentro de la captura: Con ayuda de una librería que permite el recorte del objeto principal de la captura [26], se obtiene el recorte del rostro del sujeto completo. Tras ello, se obtiene el canal alfa de la imagen, siendo este el grado de transparencia de cada punto. Con el canal alfa, se realiza la media de todos los puntos, p , siguiendo la Ecuación 5.

$$D_5 = \max \left[\left(\frac{127 - \bar{p}}{1,27} \right), 0 \right] \quad (5)$$

IV-C. Cálculo de la evaluación final

Tras evaluar cinco de las seis dimensiones puntuables en el rango 0-100 para las fotografías faciales acordes a la normativa, se calcula la puntuación final en base a la media ponderada de cada dimensión. La Ecuación 6 recoge mejor configuración de pesos para cada dimensión evaluada.

$$punt = 0,45 \cdot D_1 + 0,15 \cdot D_2 + 0,15 \cdot D_3 + 0,1 \cdot D_4 + 0,15 \cdot D_5 \quad (6)$$

En la Ecuación 6, D_1 corresponde con la puntuación de los ángulos de la cara, D_2 con la puntuación de la apertura de la boca, D_3 con la apertura de los ojos, D_4 con la exposición

TABLA II
EJEMPLOS DE EVALUACIONES DE ROSTROS

1 (92 puntos)	2 (88 puntos)	3 (15 puntos)	4 (65 puntos)
5 (74 puntos)	6 (60 puntos)	7 (82 puntos)	8 (56 puntos)
9 (52 puntos)	10 (68 puntos)	11 (33 puntos)	12 (75 puntos)
1. Rostro óptimo	2. Boca ligeramente abierta	7. Rostro óptimo.	8. Cabeza ladeada.
3. Cabeza muy ladeada; ojos cerrados; boca abierta.	4. Cortado superior; Boca abierta	9. Cabeza ladeada.	
5. Boca abierta.	6. Imagen oscura.	10. Ojos cerrados; boca abierta.	
		11. Gafas de sol; boca abierta.	
		12. Boca abierta.	

de la imagen, y D_5 con la alineación de la cara. Por último, debe aplicarse el factor de corrección si lleva gafas de sol de la Ecuación 7.

$$Punt\ final = \begin{cases} punt & \text{si no gafas de sol } (D_6) \\ 0,5 \cdot punt & \text{si gafas de sol } (D_6) \end{cases} \quad (7)$$

Con esta media ponderada, se obtiene la máxima puntuación $-100-$ para aquellos rostros que cumplen estrictamente con las condiciones, y la puntuación mínima $-0-$ cuando ninguna condición de las requeridas sea satisfecha.

V. RESULTADOS

El diseño de este algoritmo lleva asociado el lanzamiento de pruebas con rostros generados con redes generativas adversarias. Concretamente, se hace uso del servicio ThisPersonDoesNotExist [27] para la generación de rostros faciales sintéticos. La entrada del algoritmo va a ser un conjunto de imágenes que se pretenden filtrar con la evaluación del algoritmo. Como salida de la ejecución, se obtiene un conjunto de datos en el que se asocia la puntuación con cada captura. Adicionalmente, se limitan los resultados válidos al mejor percentil que se establezca.

Al proceso se le ha de indicar el número de imágenes que se quieren obtener como etapa final, o en su defecto, una puntuación de corte mínima. alguna de las puntuaciones obtenidas para diferentes rostros pueden verse en la Tabla II.

En la Figura 3 se recoge un histograma de las puntuaciones obtenidas de la ejecución del algoritmo sobre diez mil imágenes, así como, los cuartiles de puntuación en la Tabla III. Cuantas más imágenes de rostros faciales sean pasadas al algoritmo, mejores serán los resultados si, por ejemplo, se recoge el decil más alto. El histograma se sitúa bajo la curva de una campana de Gauss, aunque presentando una

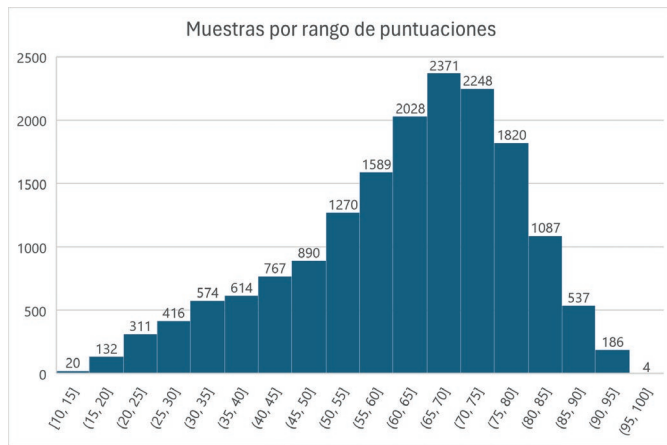


Figura 3. Distribución de las muestras por puntuación

cola a la izquierda mayor. Además, entre el rango (60 – 75] se concentran los dos tercios de las puntuaciones de las imágenes.

Una de las mejoras de resultados se daría descartando aquellas imágenes que incumplan las situaciones límites. Podrían desecharse todas los rostros que tuvieran gafas de sol, sin actuar de limitación, o eliminar aquellas imágenes que tengan aperturas de ojos inferiores al mínimo exigido, o apertura de boca mayor al máximo permitido.

TABLA III
CUARTILES OBTENIDOS EN LAS PRUEBAS LANZADAS

min. punt.	1 ^{er} cuartil	2 ^o cuartil	3 ^o cuartil	4 ^o cuartil / max. punt.
10	52	64	73	96

VI. CONCLUSIONES

Este trabajo presenta una metodología de evaluación de rostros humanos innovadora basada en un algoritmo en tres fases. El algoritmo evalúa las imágenes faciales en una puntuación numérica 0 – 100 conforme a seis dimensiones. Estas dimensiones son extraídas de las necesidades legales y de los diferentes estándares que permiten el aseguramiento de la calidad de los rostros faciales.

Con este proceso, se puede asegurar la obtención de conjuntos de imágenes faciales que tengan una calidad mínima asegurada para la expedición de diferentes documentos de identificación, permitiendo la generación de documentos de identificación sintéticos. De misma forma, este proceso puede ser empleado como filtro de seguridad para la verificación de la identidad en diferentes procesos.

AGRADECIMIENTOS

Destacar que esta iniciativa se realiza en el marco de los fondos del Plan de Recuperación, Transformación y Resiliencia, financiados por la Unión Europea (Next Generation) en el marco del proyecto con referencia C108/23 “Detección de Falsificación de Documentos de Identidad mediante Técnicas de Visión por Computador e Inteligencia Artificial”.

REFERENCIAS

[1] C. Europea. Identidad Digital Europea. [Online]. Available: https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/european-digital-identity_es

[2] N. Institute of Standards and Technology. Face analysis technology evaluation (fate) quality assessment - specific image defect detection - api and concept document. [Online]. Available: https://pages.nist.gov/frvt/api/FRVT_ongoing_quality_sidd_api.pdf

[3] M. del Interior. Gobierno de España, “Real decreto 1553/2005, de 23 de diciembre, por el que se regula la expedición del documento nacional de identidad y sus certificados de firma electrónica.” [Online]. Available: <https://www.boe.es/buscar/act.php?id=BOE-A-2005-21163&p=20150530&tn=1>

[4] T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, and T. Aila, “Analyzing and improving the image quality of stylegan,” 2020.

[5] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, “DeepFace: Closing the Gap to Human-Level Performance in Face Verification,” in *2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014. [Online]. Available: <https://research.facebook.com/publications/deepface-closing-the-gap-to-human-level-performance-in-face-verification/>

[6] G. B. Huang, M. Mattar, T. Berg, and E. Learned-Miller, “Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments,” in *Workshop on Faces in ‘Real-Life’ Images: Detection, Alignment, and Recognition*. Marseille, France: Erik Learned-Miller and Andras Ferencz and Frédéric Jurie, Oct. 2008. [Online]. Available: <https://inria.hal.science/inria-00321923>

[7] F. Schroff, D. Kalenichenko, and J. Philbin, “Facenet: A unified embedding for face recognition and clustering,” in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2015.

[8] J. Deng, J. Guo, N. Xue, and S. Zafeiriou, “Arcface: Additive angular margin loss for deep face recognition,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019.

[9] S. Agarwal, H. Farid, Y. Gu, M. He, K. Nagano, and H. Li, “Protecting world leaders against deep fakes,” in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2019.

[10] A. Rössler, D. Cozzolino, L. Verdoliva, C. Riess, J. Thies, and M. Nießner, “Faceforensics: A large-scale video dataset for forgery detection in human faces,” 2018.

[11] J. Thies, M. Zollhofer, M. Stamminger, C. Theobalt, and M. Niessner, “Face2face: Real-time face capture and reenactment of rgb videos,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[12] Y. Li, M.-C. Chang, and S. Lyu, “In ictu oculi: Exposing ai created fake videos by detecting eye blinking,” in *2018 IEEE International Workshop on Information Forensics and Security (WIFS)*, 2018, pp. 1–7.

[13] X. Yang, Y. Li, and S. Lyu, “Exposing deep fakes using inconsistent head poses,” 2018.

[14] A. Nanda, S. W. A. Shah, J. J. Jeong, R. Doss, and J. Webb, “Toward higher levels of assurance in remote identity proofing,” *IEEE Consumer Electronics Magazine*, vol. 13, no. 1, pp. 62–71, 2024.

[15] National Institute of Standards and Technology. Face Technology Evaluations - FRTE/FATE. [Online]. Available: <https://www.nist.gov/programs-projects/face-technology-evaluations-frtefate>

[16] J. Y. P. Grother, M. Ngan, K. Hanaoka, and A. Hom, “Face Analysis Technology Evaluation (FATE) Part 11: Face Image Quality Vector Assessment,” *National Institute of Standards and Technology, Gaithersburg, MD, NIST IR 8485*, 2023.

[17] I. S. Organization. Iso/iec 29794-5 information technology — biometric sample quality — part 5: Face image data. [Online]. Available: <https://www.iso.org/obp/ui/en/#iso:std:iso-iec:29794:-5:dis:ed-1:v1:en>

[18] M. Abadi *et al.*, “TensorFlow: Large-scale machine learning on heterogeneous systems,” 2015, software available from tensorflow.org. [Online]. Available: <https://www.tensorflow.org/>

[19] I. de Paz Centeno. MTCNN face detection implementation for TensorFlow. [Online]. Available: <https://github.com/ipazc/mtcnn>

[20] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, “Joint face detection and alignment using multitask cascaded convolutional networks,” *IEEE Signal Processing Letters*, vol. 23, no. 10, pp. 1499–1503, 2016.

[21] Jerry, “Jerryhouuu/Face-Yaw-Roll-Pitch-from-Pose-Estimation-using-OpenCV.” [Online]. Available: <https://github.com/jerryhouuu/Face-Yaw-Roll-Pitch-from-Pose-Estimation-using-OpenCV>

[22] F. Chollet *et al.*, “Keras,” <https://keras.io>, 2015.

[23] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, “Mobilenetv2: Inverted residuals and linear bottlenecks,” 2019.

[24] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, “Imagenet: A large-scale hierarchical image database,” in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248–255.

[25] Glasses or No Glasses. [Online]. Available: <https://www.kaggle.com/datasets/jeffheaton/glasses-or-no-glasses>

[26] D. Gatis, “Rembg.” [Online]. Available: <https://github.com/danielgatis/rembg>

[27] This person does not exist. [Online]. Available: <https://thispersondoesnotexist.com/>