Detection of Facial Spoofing Attacks in Uncontrolled Environments using ELBP and Color Models

W. Valderrama, A. Magadan, O. Vergara, *Senior member, IEEE*, J. Ruiz, R. Pinto and G. Reyes

***Abstract*— Distance education has become an alternative in teaching derived from the Covid-19 pandemic. However, distance education has led to bad practices for some students. For example, it was detected that some students spoofed the teacher in a class or exam. Therefore, facial biometrics can be used to solve, in real-time, the spoofing problem. However, the solution is not exempt from presentation attacks that undermine the reliability of the systems. Other challenges that must be considered are lighting, resolution, and variable size of the faces, among others. In this paper, we present a methodology to address the problem of facial spoofing attacks. We combine the Extended Local Binary Patterns (ELBP) descriptor and YCbCr, HSV color models to highlight the saturation and illumination of an image. For the experiments, we present a comparison of our proposal against other state-of-the-art methods. We obtain an error of 2.45% with the Half Total Error Rate (HTER) metric in the MSU image bank. The results revealed that for environments where the camera resolutions are not controlled, our proposal provides a feasible solution reducing the costs of acquiring specific hardware.**

***Index Terms*— Biometrics, Facial Spoofing, Distance Education, ELBP, Colors model.**

I. Introducción

La educación a distancia se ha convertido en una alternativa en el ámbito de la enseñanza-aprendizaje derivado de la pandemia por covid-19. Sin embargo, la seguridad en el proceso de autenticación de los sistemas utilizados para la educación a distancia es todavía un reto para las instituciones educativas. El rostro es uno de los rasgos biométricos más utilizados en los sistemas de autenticación [1]. A diferencia de otros rasgos biométricos como la huella dactilar y el iris que necesitan de hardware específico para su funcionamiento, los algoritmos para detectar rostros pueden ser implementados y ejecutados en tiempo real casi en cualquier procesador [2].

Dado que el objetivo de los sistemas de reconocimiento facial es verificar la identidad de los usuarios, pueden ser vulnerables a los ataques de suplantación de identidad. Es decir, no pueden determinar si la muestra biométrica presentada al sensor es real o falsa.

W. Valderrama, A. Magadan, J. Ruiz, R. Pinto, and G. Reyes, Centro Nacional de Investigación y Desarrollo Tecnológico (TecNM/CENIDET), Cuernavaca, Morelos, México, (e-mail: wendy.valderrama17ca@cenidet.edu.mx, {andrea.ms, jose.ra, raul.pe, gerardo.rs}@cenidet.tecnm.mx).

O. Vergara, Universidad Autónoma de Ciudad Juárez, Ciudad Juárez, Chihuahua, México, (e-mail: overgara@uacj.mx).

La suplantación de identidad se puede realizar mediante la superposición, frente a la cámara de la computadora de una fotografía impresa o la reproducción de video usando una pantalla electrónica (a través de una Tablet o teléfono celular). Además, se debe considerar que la fotografía se puede obtener por redes sociales u otros medios de difusión.

Los ataques de suplantación de identidad son también conocidos como ataques de presentación (*Presentation Attacks* (PA)) [3]. Un método de detección de PA puede distinguir automáticamente entre los rasgos biométricos reales presentados al sensor y los artefactos producidos sintéticamente. Por lo tanto, sería recomendable contar con algoritmos que identifiquen si la persona detrás de la cámara es una persona real o ha sido suplantada.

En la última década (2011-2021), en la literatura se han publicado investigaciones que proponen técnicas de detección de ataques de presentación (*Presentation Attack Detection* (PAD)) o, también conocidas en inglés como *anti-spoofing* [4]–[11]. Sin embargo, a pesar de los avances, todavía sigue siendo un área de investigación abierta debido a los desafíos por resolver al implementar los sistemas en entornos reales, como los ambientes en las clases en línea. Los retos incluyen: mala iluminación en los rostros, fondos complejos, variabilidad en la resolución de los dispositivos, distintos anchos de banda de la red, imágenes con menor calidad, entre otros.

De acuerdo con el estándar ISO/IEC 30107-1, la característica biométrica utilizada en un PA se denomina Instrumento de Ataque de Presentación (*Presentation Attack Instrument* (PAI)) y puede clasificarse en dos tipos [8]:

1) Artificial: Se refiere a un medio para generar el PAI. Por ejemplo, un video de una cara, una máscara facial 3D, una impresión facial 2D, etc.
2) Características humanas: implica el uso de humanos como un PAI y puede ser: una parte de la cara de un cadáver, cara de un humano inconsciente, entre otras.

De los dos diferentes tipos de PAI, el artificial es el más utilizado en la literatura para estudiar las vulnerabilidades de los sistemas de reconocimiento facial. Detectar dichos ataques, a veces es una tarea desafiante, incluso para los humanos.

El objetivo del artículo es proponer una solución al problema de la suplantación de identidad por medio del rostro sin utilizar hardware específico. Las principales aportaciones son:

1. Se presenta un sistema para la detección de suplantación facial en entornos no controlados, considerando variaciones en la iluminación, en la resolución y tamaño de la imagen.
2. El entrenamiento y las pruebas del sistema se realizan

con diferentes bancos de imágenes.

3. La metodología de solución se basa en el tratamiento digital de la imagen por medio de los canales de color YCbCr (Y es el componente de luminancia, CB y CR son los componentes de crominancia de diferencia azul y diferencia roja) y HSV (Hue, Saturation, Value – Matiz, Saturación, Valor) en combinación con el descriptor de textura LBP extendido.
4. Se revisan las variaciones que ofrecen los bancos de imágenes públicos para estudios de suplantación facial en entornos no controlados y los resultados obtenidos con algoritmos de visión por computadora.

El trabajo está organizado de la siguiente manera. En la sección II, se muestra un análisis de los trabajos relacionados; en la sección III, se presenta la descripción de la metodología propuesta. En la sección IV, se describe la experimentación y resultados. Finalmente, en la sección V se muestran las conclusiones y trabajos futuros.

II. Trabajos relacionados

En la literatura sobre suplantación de identidad, son escasos los estudios que se enfocan en abordar la problemática en entornos no controlados. Además, la mayoría de los estudios utilizan hardware especializado, que suele ser caro. Mohammadi et al. [12] realizaron un estudio de los métodos con mejores resultados en la detección de posibles ataques. Presentaron las características de los bancos de imágenes más utilizados, las métricas y un caso de estudio basado en el análisis de color y textura. También, presentaron una comparativa de su propuesta con los resultados reportados en el estado del arte. En los resultados explicaron que los principales factores que afectan la efectividad son la normalización de la imagen, la insuficiente iluminación, las sombras y los reflejos que pueden aparecer en los lentes. Por otro lado, Li et al. [3] realizaron un análisis de los desafíos que presenta la suplantación facial. Revisaron las técnicas que se han propuesto en diversos estudios para dar solución al problema que incluyeron, el análisis de movimiento, de textura, calidad de la imagen y los basados en hardware. Con los bancos de imágenes Replay-Attack obtuvieron un error de hasta 18% y con CASIA-FASD un error entre el 37% y 39% con la métrica HTER, con lo que se muestra que la complejidad del banco de imágenes CASIA-FASD es mayor a la de Replay-Attack.

La importancia de la información de textura a veces se descuida o incluso se minimiza en modelos de visión. En biometría facial, existen diversos modelos de PAD basados en textura [13]–[18] y la información que aportan se considera una característica crucial contra los ataques de suplantación. Una de las técnicas tradicionales de textura es el patrón binario local (*Local Binary Pattern* LBP). Chan et al. [19] realizan la detección de suplantación facial mediante una combinación de LBP con imágenes iluminadas mediante el flash de una cámara. La luz resalta la textura y ayuda a tener una mejor descripción con LBP. Los resultados reportados alcanzan un 2.73% de error con la métrica HTER en un banco de imágenes creado por los autores del artículo. Existen trabajos que utilizan variaciones de LBP como lo muestran X. shu, et al. [18] que proponen la técnica llamada patrón binario local de diferencia de equilibrio (*Equilibrium difference local binary pattern ED-LBP*) aplicado a tres diferentes modelos de color, con el fin de obtener las texturas de cada uno de ellos y concatenarlas para conformar el vector de características. su propuesta logra obtener un 2.59 % con la métrica HTER aplicado al banco de imágenes de CASIA-FASD.

Se pueden encontrar diversos trabajos en la literatura que utilizan el banco de imágenes NUAA [20], debido a que cuenta con imágenes impresas en diferentes condiciones de iluminación lo que lo hace útil para pruebas con sistemas que se enfocan en textura, como lo es el caso de T. Das, et al. [21] y A. Gunay, et al. [22], los cuales reportan resultados de 9.5% y 12.18% respectivamente, con la métrica HTER.

Angadi et al. [4], usaron tanto la información estructural como la información de magnitud de la vecindad 3 × 3 de cada píxel, con un valor de nivel de gris central para lograr ser discriminatorio obteniendo porcentajes entre 98.85% y 98.97% de exactitud con el banco de imágenes NUAA (*Nanjing University of Aeronautics and Astronautics*). De la misma manera Kartika et al. [23] calcularon el histograma de LBP convencional y combinaron la descripción con la variante LBP. La unión del histograma convencional y el histograma de la variante de LBP son la entrada al clasificador del vecino más cercano (k-Nearest Neighbor (KNN)). El clasificador KNN no crea un modelo de clasificación por lo tanto el vector de características debe ser lo suficientemente discriminante y robusto para obtener buenos resultados. En el trabajo se reporta un 87.22% de exactitud con NUAA. Song et al. [24] combinan LBP con el color proporcionado por imágenes RGB (Red, Green, Blue) y fortalecen la descripción de la textura con filtros Gabor obteniendo buenos resultados en dos bancos de imágenes con características diferentes. Con NUAA alcanzan un 0.02% de error mientras que con CASIA-FASD alcanzan un 0.07% de error, ambos resultados con la métrica HTER. Una propuesta diferente es la de Tsitiridis et al. [25] donde siguen un método inspirado en el funcionamiento del ojo humano al percibir los colores y toman las ventajas que ofrece cada tono (del modelo RGB) por separado. Posteriormente, extraen la textura en cada tono, e integran el vector de características con la información de los tres canales. Después, identifican la presencia o no de una posible suplantación de identidad. El banco de imágenes que utilizaron fue de su creación con una cámara infrarroja y alcanzaron un porcentaje de 13.83% con la métrica FRR (*False Rejection Rate*).

Para abordar la suplantación facial, en la literatura, se encuentran otros enfoques que consideran distintos aspectos que complementan a los descriptores de color y textura, como es la calidad de imagen. El enfoque se basa en la premisa de que la captura de la imagen de una fotografía no tiene la misma calidad que la imagen de una persona real. Simanjuntak et al. [26] emplea el enfoque de la calidad de imagen en combinación con el espacio de color HSV. En la imagen resultante se calculan las funciones estadísticas como la media, la desviación estándar, la asimetría y los porcentajes de histograma mínimo y máximo para cada canal de HSV. A las características extraídas les aplican una transformación mediante el Análisis de Componentes Principales (ACP) y se quedan con los mejores componentes principales. Los resultados que obtuvieron en el entrenamiento y evaluación con los bancos de imágenes son de 19.8% de error con CASIA-FASD y 21.6% con MSU-MFSD con la métrica HTER. La propuesta es interesante; sin embargo,

tiene una falla ya que las imágenes con baja resolución tienen pérdida de información y la detección de vida se puede confundir con un ataque de suplantación.

Por su parte, Ali et al. [11] consideraron describir el rostro de manera local al analizar sólo las zonas más discriminantes o que proporcionan más información para ser descritas mediante Fourier. Obtuvieron resultados de 0% de error en el banco de imágenes Replay-attack con la métrica HTER, mientras que con el banco de imágenes UVAD (*Unicamp Video-Attack Database*) obtienen un 26% con la métrica EER. Si se considera la imagen completa surgen varias interrogantes: ¿Ayuda el detectar los límites de la fotografía o dispositivo utilizado en el reconocimiento del ataque de la suplantación? ¿La información del fondo aporta datos de textura o produce ruido y con ello una mayor complejidad en la detección de suplantación facial? ¿El cabello aporta información? Arini, et al. [27] aportan información sobre dichas interrogantes ya que mostraron que la detección de figuras geométricas en el fondo de la imagen es una alternativa interesante para delimitar la zona que se desea analizar y que contribuye en la detección de ataques. Reportan resultados de 13.0% en el banco de imágenes MSU y 23.8% en el banco de imágenes NUAA con la métrica HTER. Edmunds, et al. [28] optaron por el análisis de movimiento, el cual resulta muy útil para identificar ataques estáticos como el realizado con imágenes impresas. Reportan resultados de 5.7% en el banco de imágenes Replay-attack, 19% con el banco CASIA y 17% con el banco MSU, los tres con la métrica EER. sin embargo, el sistema presenta dificultades cuando se utilizan ataques mediante reproducciones de video o máscaras 3D.

La Tabla 1 muestra un concentrado de los bancos de imágenes utilizados en las investigaciones y el rendimiento reportado.

TABLA I
RESUMEN DE BANCOS DE IMÁGENES UTILIZADOS EN LAS ETAPAS DE ENTRENAMIENTO, VALIDACIÓN Y EL RENDIMIENTO LOGRADO

Artículo	Banco de imágenes de entrenamiento	Banco de imágenes de pruebas	Resultados
Angadi et al. [4]	NUAA	NUAA	Exactitud = 98.97 %
Kartika et al. [23]	NUAA	NUAA	Exactitud = 87.22%
Song et al. [24]	NUAA	NUAA	HTER= 0.0216%
	Casia- FASD	Casia- FASD	HTER= 0.0722%
Tsitiridis et al. [25]	BIOPAD	CASIA	ACER=92.75%
Simanjuntak et al. [26]	CASIA	MSU	TPR=68.3%
	MSU	CASIA	TPR=90.8%
Ali et al. [11]	Replay-Attack	Replay-Attack	HTER=0.00%
	CASIA	CASIA	EER=0.00%
	UVAD	UVAD	EER=26%
Arini, et al. [27]	MSU	CASIA	HTER=38.30%
	NUAA	MSU	HTER=37.90%
	CASIA	NUAA	HTER=54.40%
	Replay-Attack	MSU	EER=40.8%
Edmunds, et al. [28]	CASIA	Replay-Attack	EER=33.7%
	MSU	Replay-Attack	EER=30.6%
	3DMAD	CASIA	EER=51.5%

Como se puede observar, mientras el banco de imágenes empleado para entrenamiento y pruebas sea el mismo los porcentajes son buenos. Sin embargo, en los artículos en donde se entrena y evalúa con diferentes bancos de imágenes los resultados no son buenos, lo cual muestra la dificultad de generalización ante las variaciones en los bancos de imágenes. Los resultados mostrados en la tabla 1, se dan en términos de las métricas utilizadas para evaluar el desempeño de los sistemas en la detección de suplantación y que se explican en la sección III.C del presente artículo.

Existen trabajos que dan solución al problema de suplantación de identidad con Deep Learning como: Arora et al. [29], Kumar et al. [30], Nagpal et al. [31], Yu et al. [32], Fatemifar et al. [33], Kavitha et al. [34], Hashemifard et al. [35], Chen et al. [36]. Los resultados de estos trabajos no se consideraron en las tablas comparativas ya que el enfoque propuesto es basado en visión artificial.

III. METODOLOGÍA PROPUESTA

Se propone una metodología para realizar la detección de ataques en entornos no controlados, con imágenes adquiridas con cámaras web tradicionales (hardware no especializado). La metodología se basa en el análisis del color y la textura, mediante el uso de los espacios de color YCbCr [37] y HSV [38], la descripción se realiza con la técnica ELBP [39]. En la Fig. 1 se muestra un diagrama de las etapas que componen la metodología, las cuales se describen en las siguientes subsecciones siguientes.

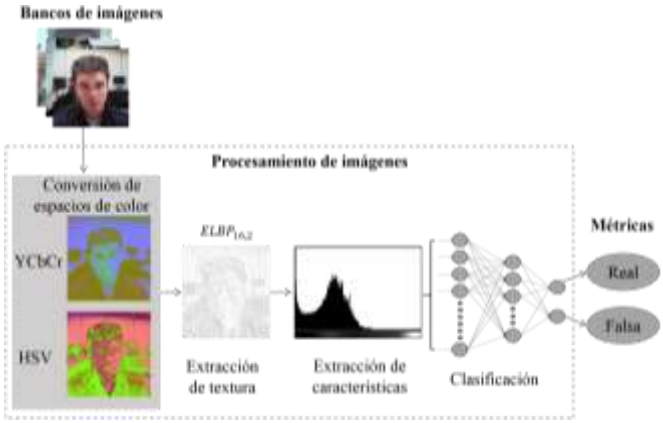

Fig. 1. Metodología propuesta.

A. Bancos de Imágenes para la Detección de Ataques de Suplantación

La disponibilidad de los bancos de imágenes públicos desempeña un papel importante en el desarrollo de nuevos esquemas de PAD frontal y en la reproducción de los resultados informados. Existen bancos de imágenes utilizados en la literatura para abordar la suplantación facial. En la Tabla 2 se muestra una lista de los bancos más utilizados, en su mayoría se enfocan a imágenes impresas planas en un ambiente controlado. Sin embargo, derivado del rápido avance de las tecnologías los sistemas actuales deben ser robustos a diferentes tipos de ataques como, la reproducción de video por medio de celulares, pantallas, etc.

TABLA III
BANCOS DE IMÁGENES CON AMBIENTE CONTROLADO REPORTADOS EN LA LITERATURA

Nombre	Tipo de ataque	Sujetos	No. Imágenes Reales/Falsas
MOBIO [40]	Video y audio	152	3990/147630
Yale-Recaptured [41]	Foto impresa plana	28	640/1920
Print-Attack [42]	Foto impresa plana	50	200/200
BERC Webcam [43]	Foto impresa plana	25	1408/7461
UVAD [44]	Video	404	344/3528
Replay-Attack [45]	Fotografía impresa y reproducción de video	40	390/640

Los bancos de imágenes que se utilizan para la experimentación de la presente investigación abordan el problema de la suplantación de identidad en entornos no controlados y a continuación, se describen brevemente.

1) MSU-MFSD [46]
Consta de 440 videoclips de intentos de ataque con fotografías y video de 55 sujetos como se muestra en la Fig. 2. Se utilizaron dos tipos de cámaras: i) cámara incorporada en MacBook Air 13, denominada cámara de portátil; ii) cámara frontal en el teléfono Android Google Nexus 5, denominada cámara Android. Para la cámara de la computadora portátil, los videos se capturan usando QuickTime en la plataforma Mac OS X Mavericks, con una resolución de 640 × 480. Para la cámara de Android, los videos se capturan usando el software integrado de Google en Android 4.4.2, con una resolución de 720 × 480. El banco de imágenes permite evaluar los algoritmos de detección de falsificaciones faciales en diferentes cámaras y condiciones de iluminación con dispositivos móviles.

2) CASIA-FASD [47]
Cuenta con tres diferentes calidades: baja calidad que es capturada por una cámara USB con dimensiones de 640 x 480, calidad normal la cual es obtenida de otra cámara USB con dimensiones de 480 x 640 y alta calidad la cual se obtienen de una cámara Sony NEX-5 con resolución de 1920 x 1080. El banco de imágenes consta de tres tipos de ataques: fotografía impresa deformada, las fotografías se obtienen con una cámara de calidad alta y se imprimen en papel de cobre, otro tipo de ataque son las fotografías impresas con ojos cortados, para el cual las fotografías impresas se les recortan los ojos y por último el ataque de video para el cual los videos de alta resolución se muestran usando un iPad. Las características del banco de imágenes permiten evaluar la efectividad del sistema con imágenes de baja resolución. En la Fig. 3 se puede observar un ejemplo de los diferentes ataques.

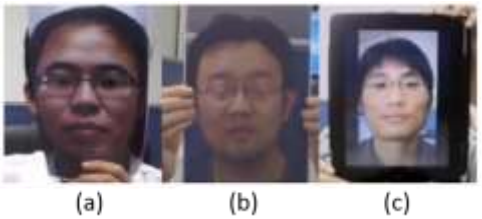

Fig. 3. Ejemplo de caras falsas. (a) fotografía impresa con ojos recortados; (b) fotografía impresa deformada y (c) fotografía por medio de iPad [47].

3) NUAA Photograph Imposter Database [20]
El banco de imágenes se construyó utilizando una cámara web genérica. La ubicación y las condiciones de iluminación de cada sesión son variadas. El conjunto de imágenes consta de 15 sujetos en cada sesión, las fotografías de los sujetos (rostro real) y sus fotografías de ataque se capturan con una velocidad de 20 fps. Las fotografías se toman de dos formas: La primera forma es imprimirlos en un papel fotográfico con el tamaño general de 6.8 cm × 10.2 cm (trivial) y con un tamaño de 8.9 cm × 12.7 cm (más prominente), respectivamente y la segunda forma es imprimirlos en un papel de tamaño A4 de 70 g con una impresora HP de color estándar. En la Fig. 4 se puede observar un ejemplo de los diferentes ataques.

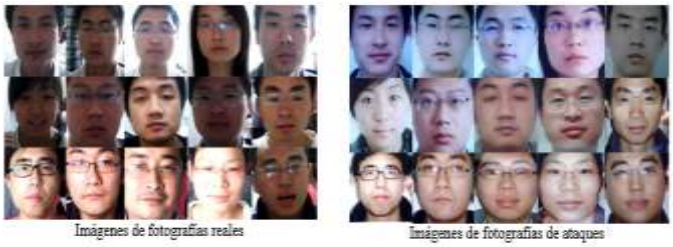

Fig. 4. Imágenes de ejemplo contenidas en el banco NUAA [20].

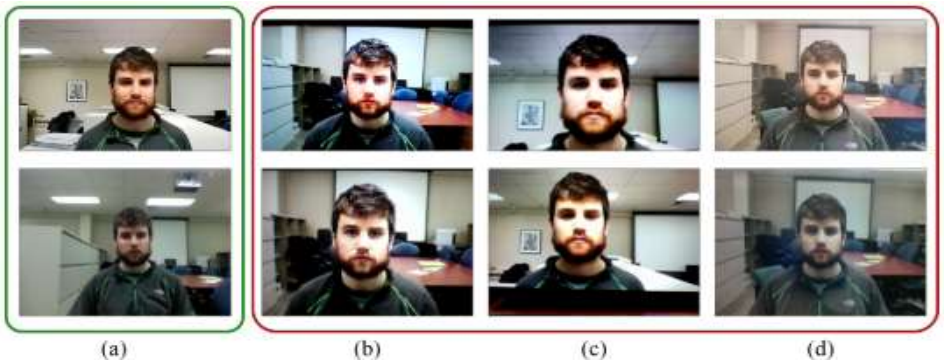

Fig. 2. Imágenes de ejemplo de caras reales y falsas de uno de los sujetos del banco de imágenes capturadas con la cámara del teléfono inteligente Google Nexus 5 (fila superior) y MacBook Air 13'' cámara portátil (fila inferior). (a) Caras reales; (b) Caras falsas generadas por iPad para ataques de reproducción de video; (c) caras falsas generadas por el iPhone para el ataque de reproducción de video; (d) Caras falsas generadas para el ataque fotográfico impreso [46].

B. Procesamiento de Imágenes

1) Conversión de espacios de color

La metodología propuesta es denominada YHE por la combinación de YCbCr + HSV + ELBP. En la propuesta se considera que la imagen completa (fondo y rostro) proporciona más información respecto a las variaciones de textura. Por lo tanto, la imagen original en RGB se cambia al espacio de color YCbCr [37] para resaltar los componentes de brillo que existan en la imagen utilizando las ecuaciones siguientes:

$$Y = \left(\frac{1}{256}\right) * [(16 * 256 + 129 * G) + (66 * R + 25 * B)] \quad (1)$$

$$Cb = \left(\frac{1}{256}\right) * [(128 * 256 + 112 * B) - (38 * R + 74 * G)] \quad (2)$$

$$Cr = \left(\frac{1}{256}\right) * [(128 * 256 + 112 * R) - (94 * G + 18 * B)] \quad (3)$$

donde R, G y B son valores RGB con corrección de gamma y las señales de entrada y salida son de valores de 8 bits. Posteriormente, la imagen resultante se transforma al espacio de color HSV cónico [38] para resaltar la saturación, por medio de las siguientes ecuaciones:

$$R' = \frac{R}{255}; G' = \frac{G}{255}; B' = \frac{B}{255} \quad (4)$$

$$C_{max} = MAX(R', G', B') \quad (5)$$

$$C_{min} = MIN(R', G', B') \quad (6)$$

$$\Delta = C_{max} - C_{min} \quad (7)$$

$$H = \begin{cases} 60° * \left(\frac{G' - B'}{\Delta} mod 6\right), C_{max} = R' \\ 60° * \left(\frac{B' - R'}{\Delta} + 2\right), C_{max} = G' \\ 60° * \left(\frac{R' - G'}{\Delta} + 4\right), C_{max} = B' \end{cases} \quad (8)$$

$$S = \begin{cases} 0, \Delta = 0 \\ \frac{\Delta}{C_{max}}, \Delta <> 0 \end{cases} \quad (9)$$

$$V = C_{max} \quad (10)$$

donde *H* es el tono y su rango varia de 0 a 360 grados, *S* es la intensidad del color en el rango de 0-100% y *V* es el brillo del color y su rango se encuentra de 0-100%.

2) Extracción de textura

A la imagen resultante de las trasformaciones de los espacios de color, se le aplica el descriptor de textura ELBP [39] cuyo funcionamiento es similar al LBP original propuesto por Ojala et al. [48]. Mientras que LBP codifica sólo la relación entre un punto central y sus vecinos, ELBP está diseñado para codificar relaciones espaciales distintivas en una región local y, por lo tanto, contiene más información espacial. ELBP consta de tres descriptores similares a LBP ELBP_CI, ELBP_NI y ELBP_RD que exploran información de la intensidad del píxel central de sus píxeles vecinos, diferencias radiales y diferencias angulares, respectivamente. La intensidad del píxel central se establece como umbral frente a β que es la media de toda la imagen.

$$ELBP\_CI(x_c) = s(x_c - \beta) \quad (11)$$

donde x_c son las coordenadas del pixel central, de tal caso que si las coordenadas de x_c son (0,0), entonces las coordenadas de $x_{r,p,,n}$ en la ecuación 12, vienen dadas por $(-r \sin(2\pi n/p), r \cos(2\pi\, n/p))$. Mientras que los valores de gris $x_{r,p,n}$ de los vecinos que no caen exactamente en el centro de los píxeles se estiman por interpolación.

En lugar de utilizar el valor de gris del píxel central como el valor de umbral, como se usa en LBP, ELBP_NI utiliza el promedio de las intensidades de los píxeles vecinos para generar el patrón binario. ELBP_NI se define como el umbral frente a la media local $\beta_{r,p} = \frac{1}{p}\sum_{n=0}^{p-1} x_{r,p,n}$.

$$ELBP\_NI_{r,p}(x_c) = \sum_{n=0}^{p-1} s(x_{r,p,n} - \beta_{r,p}) 2^n \quad (12)$$

En paralelo a los descriptores basados en intensidad ELBP_NI y ELBP_CI, ELBP_RD se deriva de las diferencias de píxeles en direcciones radiales:

$$ELBP\_RD_{r,r-1,p}(x_c) = \sum_{n=0}^{p-1} s(x_{r,p,n} - x_{r-1.,p.n}) 2^n \quad (13)$$

El operador ELBP permite un análisis multiescala mediante la variación de los parámetros (r, p); es decir, cualquier radio y número de píxeles en la vecindad. Para el presente trabajo se modificaron los valores por default de ELBP y se optó por utilizar 16 vecinos con un radio de tamaño 2. El tener un mayor número de vecinos y de radio provoca que la muestra se degrade significativamente. Por lo tanto, se busca que, con este degradado de la imagen, se haga notoria la diferencia entre una imagen real de una falsa. La Fig. 5 muestra un ejemplo del resultado de aplicar a una imagen un $ELBP_{8,1}$ y $ELBP_{16,2}$.

Fig. 5. Ejemplo de los resultados con $\mathbf{ELBP_{8,1}}$ y $\mathbf{ELBP_{16,2}}$.

3) Extracción de características y clasificación

Finalmente, se crea un histograma para recopilar las ocurrencias de diferentes patrones binarios, dicho histograma se ocupa como vector de características con un tamaño de 256, el cual es la entrada al clasificador, en este caso se utiliza un perceptrón multi-capa (*Multi-Layer Perceptron* MLP) con C capas (C-2 capas ocultas) y n_c neuronas en la capa *c,* para *c=1,2, …, C.* la arquitectura que se ilustra en la Fig. 6.

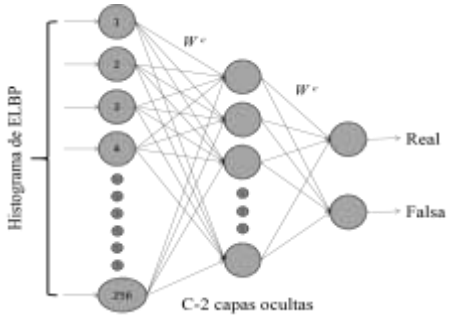

Fig. 6. Arquitectura de MLP para la detección de ataques de suplantación.

$W^c = (w_{ij}^c)$ es la matriz de pesos, donde w_{ij}^c representa el peso de la conexión de la neurona *i* de la capa *c* para *c* = 2, …, C. Las activaciones de las neuronas se calculan de la siguiente manera:

$$a_i^1 = x_i \, para \; i = 1,2, \dots, n_1 \tag{14}$$

donde a_i^1 es la activación de las neuronas de la capa de entrada y $X = (x_1, x_2, \dots, x_{n1})$. Otro aspecto importante es la función de activación, para lo cual se utiliza la función *sigmoid* (ver ecuación 15) que arroja como salida un máximo de 1 y un mínimo de 0, valores que se tomaran de la siguiente manera, 1 para rostros reales y 0 para rostros falsos.

$$f(x) = \beta * (1 - e^{-\alpha x})/(1 + e^{-\alpha x}) \tag{15}$$

donde β=1 y $\alpha = 1$.

C. Métricas Empleadas para la Metodología

Un sistema de detección de suplantación de identidad tiene que lidiar con dos tipos de eventos:

1. La persona que se encuentra frente a la computadora es quien debe ser (en cuyo caso, la imagen corresponde a una persona real).
2. No lo es (en cuyo caso, es un impostor).

Por lo tanto, una parte importante dentro de los sistemas PAD es verificar que tan eficientes son los algoritmos anti-spoofing, para ello son empleadas diferentes métricas. A continuación, se presentan las métricas de evaluación de PAD utilizadas en la literatura.

1) APCER [49]

La tasa de error de clasificación de presentación de ataques (*Attack Presentation Classification Error Rate* APCER) se define como la probabilidad esperada de un ataque exitoso y se define de la siguiente manera [49]:

$$APCER = \frac{\#\ de\ ataques\ detectados}{\#\ de\ ataques} \tag{16}$$

2) BPCER [49]

La tasa de error de clasificación de presentación de buena fe (*Bona fide Presentation Classification Error Rate* BPCER) se define como la probabilidad esperada de que un acceso de buena fe se declare falsamente como un ataque de presentación. El BPCER se calcula como [49].

$$BPCER = \frac{\#\ de\ accesos\ reales\ rechazados}{\#\ de\ accesos\ reales} \tag{17}$$

3) HTER [50]

Para proporcionar un número único para el rendimiento, los resultados se presentan típicamente utilizando la tasa de error total medio (*Half Total Error Rate* HTER), que es básicamente el promedio entre APCER y BPCER:

$$HTER = \frac{APCER + BPCER}{2} \tag{18}$$

4) Exactitud [51]

Porcentaje general de los datos clasificados correctamente. Se refiere a la dispersión del conjunto de valores, cuanto menor es la dispersión mayor es la exactitud. Se representa por la proporción entre el número de predicciones correctas (tanto positivas VP, como negativas VN) y el total de predicciones (VP, VN, Falsos Negativos (FN) y Falsos Positivos (FP)).

$$Exactitud = \frac{VP + VN}{VP + VN + FN + FP} \tag{19}$$

IV. Experimentación y Resultados

Se realizaron dos casos de prueba con el objetivo de comparar la propuesta planteada en la investigación con respecto a los resultados reportados en la literatura. Los casos de experimentación se dividen en:

a) Entrenamiento y evaluación con el mismo banco de imágenes.
b) Entrenamiento y evaluación con diferentes bancos de imágenes.

En la Tabla 3 se observan los resultados, respecto a la prueba "a", del algoritmo YHE con los bancos de imágenes CASIA y MSU, en función de las métricas APCER, BPCER, HTER y exactitud, mientras que en la Tabla 4 se observa una comparación de los resultados de la literatura en función de la métrica HTER.

Tabla IV
Resultados obtenidos con YHE con el mismo banco de imágenes

Métrica	NUAA	CASIA	MSU
APCER	4.23	6.85	0.82
BPCER	4.79	50.59	4.07
HTER	4.51	28.72	2.45
Exactitud	95.57	82.43	98.29

Como se observa en la Tabla 3 la diferencia entre los dos bancos de imágenes es significativa, esto se debe a que en las imágenes del MSU no es perceptible, a simple vista, que se trate de un ataque; es decir no se ven rastros de los bordes de las hojas

de papel, de *tablets* o celulares. Los ataques de este tipo son los más comunes en el sector de la educación ya que los estudiantes buscan que no sea perceptible el engaño. En este caso, el sistema obtiene 98.29% de exactitud. Por el contrario, en las imágenes que integran el conjunto de CASIA son notorios los bordes del dispositivo del ataque, lo que agrega textura al fondo y perjudica el análisis cuando se toma la imagen completa y no sólo la región del rostro, como en nuestra propuesta. Como resultado se tiene un alto porcentaje de error al detectar rostros genuinos como se muestra en la métrica BPCER, obteniendo un 50.59% de error.

Sin embargo, en términos del error promedio (métrica HTER), como se observa en la Tabla 4, el sistema muestra ser competitivo en comparación a los resultados reportados en el estado del arte, con los mismos conjuntos de imágenes, utilizando técnicas de visión artificial en entornos no controlados. La propuesta YHE con el conjunto CASIA tiene un error mayor al alcanzado por Simanjuntak et al. [26] debido, al análisis holístico de la imagen de entrada y al ruido en la descripción que producen los contornos de las *tablets,* alcanzando un 19.8% de error con la métrica HTER. No obstante, con el conjunto de MSU logra un porcentaje del 2.45%, porcentaje menor a lo reportado por Simanjuntak et al. [26] y Arini, et al. [27].

TABLA V
COMPARACIÓN DE RESULTADOS CON LA LITERATURA EN TÉRMINOS DE LA MÉTRICA HTER

Artículo	NUAA	CASIA	MSU
T. R. Das, et al. [21]	9.5		
X.Shu, et al. [18]		2.59	
A. Gunay, et al. [22]	12.18		
Song et al. [24]	**0.0216**	**0.0722**	
Arini, et al. [27]		66.9	13
Simanjuntak et al. [26]		19.8	21.6
Propuesta YHE	4.51	28.72	**2.45**

Como se observa en la Tabla 5, para el inciso "b", la experimentación realizada se llevó a cabo considerando las combinaciones de los dos conjuntos, NUAA, CASIA y MSU. En la prueba, el rendimiento del sistema disminuyó, aumentando los porcentajes de error significativamente, obteniendo resultados inferiores al 50%; esto se debe principalmente a la diferencia de condiciones de los bancos de imágenes. En comparación con los resultados reportados por Song et al. [24], los cuales reportan para su experimentación que descartan imágenes con baja resolución. Sin embargo, esta evaluación presenta un ambiente más cercano a los desafíos que enfrentan los sistemas de detección de suplantación en la educación a distancia, ya que se está evaluando con imágenes en ambientes no controlados y con conjuntos distintos.

TABLA VI
RESULTADOS OBTENIDOS CON YHE CON ENTRENAMIENTO Y PRUEBAS CON DIFERENTES BANCOS DE IMÁGENES

Métrica	CASIA/MSU	MSU/CASIA
APCER	49.59	61.11
BPCER	51.27	42.04
HTER	50.43	51.57
Exactitud	49.95	43.56

En la Tabla 6 se muestra que los resultados obtenidos con la metodología YHE son bajos. Sin embargo, se encuentran cerca de los resultados reportados por otras investigaciones que realizaron una evaluación similar a la descrita.

A. Discusión de Resultados

En la literatura, la principal propuesta que se sigue en la detección de ataques de suplantación es analizar sólo la región perteneciente al rostro. Sin embargo, las imágenes completas aportan más información en términos de color y textura debido a que el fondo contiene información adicional para distinguir entre imágenes reales y falsas.

TABLA VII
COMPARACIÓN CON LOS RESULTADOS REPORTADOS EN LA LITERATURA CON ENTRENAMIENTO Y PRUEBAS CON DIFERENTES BANCOS DE IMÁGENES EN FUNCIÓN DE LA MÉTRICA HTER

Artículo	CASIA/MSU	MSU/CASIA
Edmunds, et al. [28]	**50**	47.7
Arini, et al. [27]	61.8	**44.9**
Propuesta YHE	50.43	51.57

La propuesta presentada analiza la imagen de manera holística y el algoritmo YHE obtiene mejores resultados en la prueba "a" con bancos de imágenes en los que no es perceptible el objeto de ataque, como es el caso del conjunto MSU. Con respecto a las imágenes que integran el conjunto CASIA, al contener los bordes del dispositivo del ataque, el descriptor LBP extrae características ruidosas que provocan un aumento en el error. Es importante mencionar que los mejores resultados se obtienen al detectar imágenes falsas; sin embargo, la detección de rostros genuinos sigue presentando desafíos con imágenes de baja resolución.

Los resultados con los bancos de imágenes utilizados en la experimentación reflejan la complejidad de los entornos no controlados y la necesidad de sistemas robustos ante entrenamientos y pruebas con bancos de imágenes con características diferente, lo que se ve reflejado en los resultados repostados en los trabajos relacionados con los bancos de imágenes de CASIA y MSU.

V. CONCLUSIÓN

En el artículo se presentó una solución diferente para la detección de suplantación facial. La propuesta YHE está basada en un enfoque holístico de la imagen en combinación de los modelos de color YCbCr, HSV [27] y una del descriptor de texura Patrón Binario Local Extendido (ELBP). El desempeño del método propuesto fue comparado con sistemas semejantes del estado del arte que aplican transformaciones de color y el descriptor clásico LBP, llevando a cabo dos evaluaciones. La experimentación realizada contempla los conjuntos CASIA y MSU que son conjuntos de imágenes en ambientes no controlados y las métricas utilizadas en la literatura para determinar la eficacia de los sistemas de detección de ataques de suplantación. Los resultados son mejores cuando la evaluación se realiza con el mismo conjunto de imágenes utilizado en el entrenamiento. Sin embargo, cuando los conjuntos de imágenes para el entrenamiento y evaluación son disjuntos, el rendimiento disminuye, logrando resultados semejantes a los reportados en la literatura.

Como trabajo futuro se propone complementar el sistema con los siguientes puntos: una arquitectura de aprendizaje profundo; la creación de un banco de imágenes con iluminación variable considerando que el objeto de ataque no sea visible frente a la cámara, debido a que cuando los alumnos están en una clase o se realiza un examen es difícil asegurar que las condiciones de iluminación sean adecuadas. Otra limitante de los bancos de imágenes actuales es la falta de variación en la resolución de las imágenes, en su mayoría la resolución es de una calidad aceptable; sin embargo, hay una ausencia de pruebas con imágenes de baja resolución.

RECONOCIMIENTO

Agradecemos al TecNM/CENIDET y al CONACYT por el apoyo económico brindado.

REFERENCIAS

[1] J. F. J. Galbally, S. Marcel, "Biometric antispoofing methods: A survey in face recognition," *IEEE Access,* vol. 2, pp. 1530–1552, 2014.

[2] M. Hassaballah and S. Aly, "Face recognition : challenges , achievements and future directions," *IET Computer Vision*, vol. 9(4), pp. 614–626, 2015.

[3] L. Li, P. L. Correia, and A. Hadid, "Face recognition under spoofing attacks: countermeasures and research directions," *IET Biometrics*, vol. 7, no. 1, pp. 3–14, 2017.

[4] S. A. Angadi and V. C. Kagawade, "Detection of Face Spoofing using Multiple Texture Descriptors," in *International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS)*, 2019, pp. 151–156.

[5] M. Sajjad *et al.*, "CNN-based anti-spoofing two-tier multi-factor authentication system," *Pattern Recognit. Lett.*, vol. 0, pp. 1–9, 2018.

[6] X. Tu *et al.*, "Learning generalizable and identity-discriminative representations for face anti-spoofing," *ACM Trans. Intell. Syst. Technol.*, vol. 11, no. 5, 2019.

[7] O. Nikisins, A. George, and S. Marcel, "Domain adaptation in multi-channel autoencoder based features for robust face anti-spoofing," in *International Conference on Biometrics (ICB)*, 2019, pp. 1–8.

[8] I. J. S. Biometrics, "ISO/IEC 30107-1:2016. Information Technology Biometric Presentation Attack Detection," *Part 1 Fram. Int. Organ. Stand.*, 2016.

[9] J. V. C. I. R *et al.*, "Face liveness detection using convolutional-features fusion of real and deep network generated face images," *J. Vis. Commun. Image Represent.*, vol. 59, pp. 574–582, 2019.

[10] A. Parkin and G. Oleg, "Recognizing Multi-Modal Face Spoofing with Face Recognition Networks," in *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2019.

[11] Z. Ali and U. Park, "Face Spoofing Attack Detection Using Spatial Frequency and Gradient-Based Descriptor," *KSII Transactions on Internet and Information Systems*, vol. 13, no. 2, pp. 892–911, 2019.

[12] A. Mohammadi, S. Bhattacharjee, and S. Marcel, "Deeply vulnerable: a study of the robustness of face recognition to presentation attacks," *IET Biometrics*, vol. 7, no. 1, pp. 15–26, 2017.

[13] Z. Boulkenafet, J. Komulainen, and A. Hadid, "On the generalization of color texture-based face anti-spoofing," *Image Vis. Comput.*, vol. 77, pp. 1–9, 2018.

[14] J. Määttä, A. Hadid, and M. Pietikäinen, "Face spoofing detection from single images using texture and local shape analysis," *IET Biometrics*, vol. 1, no. 1, pp. 3–10, 2012.

[15] X. Zhao, Y. Lin, and J. Heikkila, "Dynamic Texture Recognition Using Volume Local Binary Count Patterns with an Application to 2D Face Spoofing Detection," *IEEE Trans. Multimed.*, vol. 20, no. 3, pp. 552–566, 2018.

[16] F. M. Chen, C. Wen, K. Xie, F. Q. Wen, G. Q. Sheng, and X. G. Tang, "Face liveness detection: Fusing colour texture feature and deep feature," *IET Biometrics*, vol. 8, no. 6, pp. 369–377, 2019.

[17] R. J. Raghavendra and R. Sanjeev Kunte, "A novel feature descriptor for face anti-spoofing using texture based method," *Cybernetics and Information Technologies*, vol. 20, no. 3, pp. 159–176, 2020.

[18] X. Shu, H. Tang, and S. Huang, "Face spoofing detection based on chromatic ED-LBP texture feature," *Multimed. Syst.*, no. 0123456789, pp. 1–16, 2020.

[19] P. P. K. Chan *et al.*, "Face Liveness Detection Using a Flash Against 2D Spoofing Attack," *IEEE Trans. Inf. Forensics Secur.*, vol. 13, no. 2, pp. 521–534, 2018.

[20] J. L. and L. J. X.Tan, Y.Li, "Face Liveness Detection from A Single Image with Sparse Low Rank Bilinear Discriminative Model," *Proc. 11th Eur. Conf. Comput. Vis.*, 2010.

[21] T. R. Das, S. Hasan, S. M. Sarwar, J. K. Das, and M. A. Rahman, "Facial spoof detection using support vector machine," *Adv. Intell. Syst. Comput.*, vol. 1309, pp. 615–625, 2021.

[22] A. Gunay, V. Nabiyev, and U. Turhal, "Effect of Feature Selection With Meta-Heuristic Optimization Effect of Feature Selection With Meta-Heuristic," *J Mod Technol Eng*, vol. 5, no. April, pp. 48–59, 2020.

[23] A. Kartika, I. B. Kusuma, T. Agung, B. Wirayuda, and K. Nur, "Image Spoofing Detection Using Local Binary Pattern and Local Binary Pattern Variance," *Int. J. Inf. Commun. Technol.*, vol. 4, no. December, pp. 11–18, 2019.

[24] L. Song and H. Ma, "Face Liveliness Detection Based on Texture and Color Features," *2019 IEEE 4th Int. Conf. Cloud Comput. Big Data Anal.*, pp. 418–422, 2019.

[25] A. Tsitiridis, C. Conde, B. G. Ayllon, and E. Cabello, "Bio-Inspired Presentation Attack Detection for Face Biometrics," *Front. Comput. Neurosci.*, vol. 13, no. May, pp. 1–17, 2019.

[26] G. D. Simanjuntak, K. N. Ramadhani, and A. Arifianto, "Face spoofing detection using color distortion features and principal component analysis," *2019 7th Int. Conf. Inf. Commun. Technol. ICoICT 2019*, 2019.

[27] D. D. Arini, K. N. Ramadhani, and F. Sthevanie, "Detection of face spoofing using low-level features and shape analysis," *J. Phys.*, vol. 1192, p. 012002, 2019.

[28] T. Edmunds and A. Caplier, "Motion-based countermeasure against photo and video spoofing attacks in face recognition," *J. Vis. Commun. Image Represent.*, vol. 50, pp. 314–332, 2018.

[29] S. Arora, M. P. S. Bhatia, and V. Mittal, "A robust framework for spoofing detection in faces using deep learning," *Vis. Comput.*, 2021.

[30] S. Kumar, S. Singh, and J. Kumar, "Face spoofing detection using improved SegNet architecture with a blur estimation technique," *Int. J. Biom.*, vol. 13, pp. 131–149, 2021.

[31] C. Nagpal and S. R. Dubey, "A Performance Evaluation of Convolutional Neural Networks for Face Anti Spoofing," in *Proceedings of the International Joint Conference on Neural Networks*, 2019, vol. 2019-July, no. July, pp. 1–8.

[32] Z. Yu, Y. Qin, X. Li, C. Zhao, Z. Lei, and G. Zhao, "Deep learning for face anti-spoofing: A survey," *Comput. Vis. Pattern Recognit.*, 2021.

[33] S. Fatemifar, S. R. Arashloo, M. Awais, and J. Kittler, "Client-specific anomaly detection for face presentation attack detection," *Pattern Recognit.*, vol. 112, p. 107696, 2021.

[34] K. R. Kavitha, S. Vijayalakshmi, A. Annakkili, T. Aravindhan, and K. Jayasurya, "Face Mask Detector Using Convolutional Neural Network," *Dr. Diss. Univ. Muhammadiyah Gresik*, vol. 25, no. 5, pp. 1979–1985, 2021.

[35] S. Hashemifard and M. Akbari, "A Compact Deep Learning Model for Face Spoofing Detection," 2021.

[36] B. Chen, W. Yang, and S. Wang, "Generalized Face Anti-spoofing by Learning to Fuse Features from High and Low Frequency Domains," *IEEE Multimed.*, no. c, pp. 1–1, 2021.

[37] B. Ahirwal, M. Khadtare, and R. Mehta, "FPGA based system for Color Space Transformation RGB to YIQ and YCbCr," *2007 Int. Conf. Intell. Adv. Syst. ICIAS 2007*, no. I, pp. 1345–1349, 2007.

[38] G. Saravanan, G. Yamuna, and S. Nandhini, "Real time implementation of RGB to HSV/HSI/HSL and its reverse color space models," *Int. Conf. Commun. Signal Process. ICCSP 2016*, pp. 462–466, 2016.

[39] L. Liu, S. Lao, P. W. Fieguth, Y. Guo, X. Wang, and M. Pietikäinen, "Median Robust Extended Local Binary Pattern for Texture Classification," *IEEE Trans. Image Process.*, vol. 25, no. 3, pp. 1368–1381, 2016.

[40] M. Chris, M. Pavel, M. Timo, and A. Jan, "Mobile biometry (mobio) face and speaker verification evaluation," *Idiap Res. Institute. Tech. Rep.*, no. May, p. 2652, 2010.

[41] B. Peixoto, C. Michelassi, and A. Rocha, "Face liveness detection

under bad illumination conditions," *18th IEEE Int. Conf. Image Process.*, pp. 3557–3560, 2011.

[42] A. Anjos and S. Marcel, "Counter-Measures to Photo Attacks in Face Recognition: a public database and a baseline," *Int. Jt. Conf. Biometrics*, pp. 1–7, 2011.

[43] G. Kim, S. Eum, J. K. Suhr, D. I. Kim, K. R. Park, and J. Kim, "Face Liveness Detection Based on Texture and Frequency Analyses," *2012 5th IAPR Int. Conf. biometrics*, pp. 1–6, 2012.

[44] A. Pinto, W. R. Schwartz, H. Pedrini, and A. Rocha, "Using Visual Rhythms for Detecting Video-based Facial Spoof Attacks," *IEEE Trans. Inf. Forensics Secur.*, vol. 10, no. 5, pp. 1025–1038, 2015.

[45] A. Costa-pazo, S. Bhattacharjee, E. Vazquez-fernandez, and S. Marcel, "The REPLAY-MOBILE Face Presentation-Attack Database," *2016 Int. Conf. Biometrics Spec. Interes. Gr.*, pp. 1–7, 2016.

[46] D. Wen, H. Han, and A. K. Jain, "Face spoof detection with image distortion analysis," *IEEE Trans. Inf. Forensics Secur.*, vol. 10, no. 4, pp. 746–761, 2015.

[47] Z. Zhang, J. Yan, S. Liu, Z. Lei, D. Yi, and S. Z. Li, "A Face Antispoofing Database with Diverse Attacks," *2012 5th IAPR Int. Conf. Biometrics*, pp. 26–31, 2012.

[48] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, 2002.

[49] A. A. A. Goshtasby, *Advances in Computer Vision and Pattern Recognition*. 2012.

[50] S. Bengio and J. Mari´ethoz, "A statistical significance test for person authentication," *Proc. Odyssey 2004 Speak. Lang. Recognit. Work.*, no. 2, 2004.

[51] D. Basso, "Propuesta de Métricas para Proyectos de Explotación de Información," *Rev. Latinoam. Ing. Softw.*, vol. 2, no. 4, p. 157, 2015.

Wendy Valderrama-Cardenas es Maestra en Ciencias Computacionales por el Centro Nacional de Investigación y Desarrollo Tecnológico (TecNM/CENIDET). Es Ingeniera en Informática por la Universidad Politécnica del Estado de Morelos. Actualmente es estudiante de Doctorado en Ciencias Computacionales en el (TecNM/CENIDET). Sus áreas de interés son la biometría, visión por computadora y aprendizaje profundo.

Andrea Magadán-Salazar es Doctora en Tecnologías de la Información y Sistemas Informáticos, por la Universidad Rey Juan Carlos, España. Maestra en Ciencias, en Ciencias de la Computación, por el TecNM/CENIDET, México. Actualmente labora como profesora-investigadora en el TecNM/CENIDET. Sus áreas de interés son en visión por computadora, aprendizaje de máquinas y aprendizaje profundo con aplicaciones en videovigilancia, biometría y agricultura de precisión.

Osslan Vergara (SM' 12) es ingeniero en sistemas computacionales por el Instituto Tecnológico de Zacatepec (2000); Maestro en ciencias en Ciencias Computacionales por el Centro Nacional de Investigación y Desarrollo Tecnológico (cenidet) (2003) y Doctor en Ciencias en Ciencias de la Computación también por cenidet (2006). Desde enero de 2007, es profesor de tiempo completo del Departamento de Ingeniería Industrial y Manufactura de la Universidad Autónoma de Ciudad Juárez (UACJ). Además, a partir de enero de 2014 es el director del laboratorio de visión por computadora y realidad aumentada de la UACJ. El Dr. Vergara es autor y coautor de más de 120 artículos en revistas, libros y congresos nacionales e internacionales. Es miembro del sistema nacional de investigadores (SNI) nivel I. En el año de 2012 recibió la distinción senior member por parte de la IEEE. Sus intereses en investigación incluyen: visión por computadora, procesamiento digital de imágenes, realidad aumentada y mecatrónica.

José Ruiz Ascencio es Físico egresado de la UNAM (1971). Es Maestro en Ciencias (1973) en Ingeniería Eléctrica de la Universidad de Stanford, E.U.A. en el área de sistemas digitales y Doctor en Ciencias (1989) por la Universidad de Sussex, Inglaterra, en control adaptativo. Es profesor investigador del Centro Nacional de Investigación y Desarrollo Tecnológico, Tecnológico Nacional de México desde 1995. Sus intereses actuales de investigación son la visión robótica y el control inteligente.

Raúl Pinto Elías es Doctor en Ciencias con especialización en Ingeniería Eléctrica por el Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, CINVESTAV, México. Actualmente labora como profesor-investigador en el TecNM/CENIDET. Sus áreas de interés son en visión por computadora, aprendizaje de máquinas tratamiento de lenguaje natural, aprendizaje automático cuántico.

Gerardo Reyes Salgado es Doctor en Ciencias Cognitivas por el Instituto Nacional Politécnico de Grenoble, Francia. Actualmente labora como profesor-investigador en el TecNM/CENIDET. Sus áreas de interés son en sistemas cognitivos, aprendizaje automático, aprendizaje profundo con aplicaciones en sistemas evolutivos bioinformática, optimización, tratamiento de lenguaje natural escrito.