

Identificación de Cámara Fuente a partir del Patrón de Ruido del Sensor extraído mediante Transformada Wavelet no Diezmada

Source Camera Identification from Sensor Noise Pattern obtained with the Undecimated Wavelet Transform

C. Muñoz-Bermúdez  ; J. A. Cortes-Osorio 

DOI: <https://doi.org/10.22517/23447214.24829>

Artículo de investigación científica y tecnológica

Resumen— La identificación de la cámara de fuente (SCI) es una herramienta útil para respaldar las investigaciones forenses cuando los expertos necesitan pruebas de cámaras digitales. Una de las formas de realizarlo es analizando el patrón de ruido del sensor (SPN). Este documento presenta la identificación del fabricante y la individualización de dispositivos móviles similares que utilizan el SPN para obtener características de imágenes con marcos claros y oscuros. Posteriormente, se realiza la extracción de características en el espacio de frecuencia, utilizando la Transformada Wavelet No Desintegrada (UWT) y el filtro Wiener. Se llevó a cabo una clasificación multiclase utilizando el árbol de decisiones, bayesiano, vecino más cercano y máquinas de soporte de vectores. La evaluación utilizó 500 imágenes de la base de datos de Dresden y 3200 fotografías de marcos claros y oscuros en JPEG de una base de datos propia. Para la evaluación, el método de validación cruzada se estableció en K (pliegues) igual a 5. Además, se calcularon la matriz de confusión, la exactitud y la precisión. Por otro lado, también se consideró el tiempo de entrenamiento. El clasificador Máquinas de Soporte Vectorial (SVM) obtuvo los mejores resultados, con una precisión promedio del 97% y un tiempo de entrenamiento promedio de 1681.5 segundos empleando el SPN extraído de imágenes de fondo oscuro. Este estudio mostró la usabilidad del SPN extraído del campo oscuro para resolver el problema de identificación de la cámara de origen.

Palabras claves— Identificación de Cámara Fuente, Máquinas de Soporte Vectorial, marco claro, marco oscuro, Patrón de ruido del sensor, Transformada Wavelet no Diezmada.

Abstract— The source camera identification (SCI) is a helpful tool in supporting forensic investigations when experts need evidence from digital cameras. One of the ways to carry it out is by analyzing the sensor noise pattern (SPN). This paper introduces the manufacturer identification and the individualization of similar mobile devices using the SPN to obtain features from dark and light frame pictures. Subsequently, the feature extraction was achieved in frequency space, using the Undecimated Wavelet Transform (UWT) and

the Wiener filter. A multiclass classification was carried out using Decision Tree, Bayesian, Nearest Neighbor, and Vector Support Machines. The evaluation used 500 images from the Dresden database and 3,200 light and dark frame photographs in JPEG from our database. For the evaluation, the Cross-Validation method was set at K (folds) equal to 5. Additionally, the confusion matrix, the accuracy, and the precision were calculated. On the other hand, the training time was also considered. Support Vector Machines classifier obtained the best results, with an average accuracy of 97% and an average training time of 1681.5 seconds employing the SPN extracted from clear light-frame images. This study showed the usability of the SPN extracted from the dark field to solve the source camera identification issue.

Index Terms— Clear frame, dark frame Source Camera Identification, Support Vector Machines, Sensor Pattern Noise, Undecimated Wavelet Transform.

I. INTRODUCCIÓN

EL uso y compra de dispositivos digitales, como cámaras digitales, videocámaras, y scanner, se ha ido facilitando gracias al agitado desarrollo tecnológico y bajo costo de su producción [1]. En consecuencia, los delitos que involucran imágenes o fotografías digitales han incrementado, sea porque las imágenes hacen parte del crimen, como el intercambio de fotografías de pornografía infantil, o porque sus análisis podrían ayudar a revelar pistas importantes sobre el autor de algún crimen [2].

Por lo mencionado anteriormente, en muchos casos de delitos, las imágenes y películas son usados en los juicios como evidencia para resolver un caso [3]. Es por ello que los encargados de realizar los análisis forenses en imágenes digitales, han buscado estudios y estrategias que ayuden en las investigaciones, cuando se cuenta con evidencia proveniente de dispositivos digitales. Unas de estas propuestas, es el estudio de la identificación de cámara fuente, que consiste en determinar, a partir de una imagen dada, el dispositivo usado en el momento de la toma de la imagen [2].

Este manuscrito fue enviado el 21 de enero de 2021 y aceptado el 23 de noviembre de 2021.

C. Muñoz-Bermúdez, estudiante de maestría de la universidad Tecnológica de Pereira (camilobermudez@utp.edu.co)

J. A. Cortes Osorio, docente de plata del Departamento de Física de Universidad Tecnológica de Pereira. (jacoper@utp.edu.co)



Existen varias propuestas para realizar la identificación de cámara fuente (SCI); entre ellas se encuentran: la identificación de cámara mediante las características estadísticas de la imagen [4] - [7]; otra alternativa es a partir del estudio de la matriz de filtro de color (CFA) [8]–[10]; y por último, se encuentra la propuesta de identificación de cámara de disparo a partir del patrón de ruido del sensor (SPN) [11]–[13], en el cual se enfocó esta investigación.

La SCI a partir del patrón de ruido del sensor (SPN), tiene en cuenta que cuando el obturador de una cámara fotográfica se abre, el sensor empieza a recolectar los fotones incidentes con el objetivo de medir su intensidad, la señal eléctrica producto de la conversión de la luz, trae consigo fluctuaciones que en metrología se le denomina ruido [11]. Estas perturbaciones pueden ser asignadas como una propiedad de cada sensor, permitiendo así la identificación de las cámaras a partir del estudio del SPN [12].

Entre los trabajos donde se utiliza el método SPN como solución al problema de la identificación de cámara fuente, se encuentra el estudio de Ahmad Ryad Soobhany et al. [14], en el cual, realizaron la extracción del SPN por medio de campo claro mediante la técnica basada en el filtrado en el dominio de la frecuencia. Ofreciendo así, una ventaja en la identificación del SPN, ya que lograron extraer las sub-bandas después de un nivel de descomposición, lo cual era una ventaja sobre los otros métodos de extracción del SPN.

También se encuentra la investigación de David Freire-Obregón et al. [15], en la que afrontaron el problema de la identificación de cámara fuente, desde el espacio de convolución. Para esto diseñaron una Red Neuronal Convolucional, usando la función de activación Unidad Lineal Rectificada mejorada (Leaky ReLU). Finalmente, los autores obtuvieron una exactitud en la identificación de dispositivo fuente superior al 98%, y una exactitud en la identificación de cámara fuente superior al 91%.

Por otro lado, Farah Ahmed et al. [16] realizaron una comparativa entre las dos convenciones para la extracción del patrón de ruido del sensor, filtrado en frecuencia y redes neuronales convolucionales. Los autores concluyeron que el método de filtrado en frecuencia logra un mayor número de verdaderos positivos con un menor tiempo de entrenamiento respecto a las redes neuronales.

Otro trabajo fue el de Lukas et al. [17], donde se propuso un método para la extracción de la huella del sensor de una fotografía, y mencionaron que este contiene en gran medida, el tipo de ruido utilizado mayoritariamente como huella del sensor, llamado ruido de foto respuesta no uniforme (PRNU); por esto aportó características deterministas en la identificación de cámara. Para la extracción del SPN, tomaron una imagen de campo plano, y a esta le aplicó un filtro f eliminador de artefactos en el dominio de la frecuencia. Finalmente separaron el SPN, de la imagen original,

restándole la imagen filtrada a la original.

En cuanto a esta investigación, se plantea la identificación de marca de dispositivos móviles e individualización de dispositivos tecnológicamente iguales, a partir del análisis del ruido, extraído desde fotografías de marco oscuro y fotografías de marco claro. Para ello, se propone la extracción de SPN basado en la técnica implementada por los autores en las investigaciones [14], [17] con algunas modificaciones que resulta en un aporte significativo al estado del arte.

Teniendo en cuenta que el sensor mide la intensidad de los fotones incidentes [11], la novedad presentada por esta investigación, es el uso de fotografías de campo oscuro para la extracción del ruido, de forma que si el ruido se compone de corrientes, provenientes de actividades de los fotodiodos por la temperatura en escenas con poca luz [11]; es considerado hipotéticamente que la imagen no tiene una gran saturación, por lo tanto permite extraer características con mayor componente de ruido.

Para la identificación de los dispositivos a partir de la extracción de SPN, se realizó la clasificación de dichas características mediante un análisis multiclase, propuesta que no fue tenida en cuenta en [14], utilizando clasificadores de tipo bayesiano, arboles de decisión, vecino más cercano, máquinas de soporte vectorial, y de conjuntos. Al usar esta variedad de clasificadores, se logró contrastar los tiempos de entrenamiento y exactitud de los resultados de la clasificación multiclase.

Conviene señalar que este artículo es derivado del trabajo de grado [18], en el cual se estudió a profundidad la temática de la identificación de cámara fuente, implementando técnicas basadas en la extracción del ruido en el dominio frecuencial, como es el uso del filtro invariante, y en el dominio espacial, con la utilización de redes neuronales convolucionales.

El artículo se encuentra organizado así: en la sección II se expone el desarrollo metodológico de las técnicas clásicas usadas en la identificación de cámara fuente; después, en la sección III se presentan los resultados obtenidos; posteriormente, en la sección IV se realiza un análisis de los resultados más relevantes, y se compara los resultados obtenidos usando la extracción del SPN desde imágenes de marco oscuro y marco claro; finalmente en la sección V se exponen usabilidad del marco oscuro en la extracción de la huella del sensor como solución al problema de la identificación de cámara fuente.

A. Componentes principales de SPN

El ruido en el sensor se produce por imperfecciones a la hora de la fabricación del transductor [13]. Este ruido puede ser una señal descompuesta en dos componentes: el primero, es el ruido de patrón fijo (FPN), el cual es producido por corrientes oscuras, que son las diferencias de voltajes entre

píxeles en escenas con poca iluminación [17]. El segundo es una señal que representa la excitación de los fotodiodos por efecto de la temperatura, creando así, un patrón de actividad eléctrica que también se almacena como información binaria periódica, conocida como Patrón de Ruido de Foto-respuesta no Uniforme (PRNU) [17].

B. Métodos de extracción de SPN

Para la extracción del SPN se implementaron dos técnicas: la primera con base en la extracción del ruido en el dominio de la frecuencia, con el uso de un filtro invariante; y la segunda es en el dominio espacial, con la utilización de redes neuronales convolucionales.

En el espacio de frecuencias se toma una fotografía de entrada, se le aplica la Transformada Wavelet no Diezmada, con el objetivo de realizar una descomposición de la señal, y así obtener los coeficientes de aproximación y de detalle de la imagen. El SPN es extraído mediante la implementación de un filtro de paso bajo en forma de Wiener en las tres sub-bandas que contienen los coeficientes de detalle. Después de esto, el ruido es obtenido restando las sub-bandas filtradas a las originales [17].

Por otro lado, para la extracción del SPN en el dominio del espacio, las redes neuronales convolucionales (CNN), reciben como entrada una imagen y mediante nodos interconectados entre sí, llamados capas, extraen las características de interés y logran realizar una clasificación de dichas características. El diseño de una CNN para la extracción del SPN, se implementa como solución al problema del aumento en los falsos positivos de los sistemas basados en visión por computadora [15].

C. Métodos de clasificación

El clasificador de tipo bayesiano, considera que un objeto a clasificar se compone de un conjunto de características y que a su vez, cada una de ellas contribuye de manera independiente a la asignación de una variable en particular, de modo que la suma de todas las probabilidades individuales dan como resultado una asignación de clase exitosa [19].

La clasificación a través de árboles de decisión, se realiza mediante una evaluación estadística a cada característica que se tenga en el conjunto a clasificar. Este conjunto de atributos es dividido y a cada uno de ellos se le cuantifica la entropía, denominada ganancia de información, y mediante ella se estima el atributo con mayor ganancia para establecerlo como nodo principal del árbol de decisión [20]. A partir de la elección del atributo con mayor ganancia, se desciende hacia los nodos y a las ramas, las cuales serán las posibles respuestas en la clasificación [20].

En cuanto al clasificador del vecino más cercano o KNN por sus siglas en inglés (k-Nearest Neighbor), se ubica el elemento a clasificar en un plano y mide la distancia hasta cada uno de los elementos conocidos que hay a su alrededor,

eligiendo así el vecino que se encuentre a una distancia más corta. Las métricas comúnmente utilizadas en este tipo de clasificador son, la distancia Euclidiana y la distancia Chebychev [20].

Por otro lado, las Máquinas de Soporte Vectorial (SVM) es un método multidimensional, que encuentra la dimensión más alta existente entre los atributos de entrada, y respecto a ellos, aumenta en uno la dimensionalidad del plano que contiene los datos. Con esto busca identificar un hiperplano, que permita mapear las características y discriminarlas. Si los atributos de entrada se encuentran en R^n , las máquinas de soporte vectorial los discriminarán en R^{n+1} [20].

En cambio, en la clasificación por medio de conjuntos, se realiza una reducción de un problema multiclase a una clasificación binaria [21]. La clase que se arroja como resultado de esta clasificación, es elegida bajo una ponderación de cada entrenamiento individual que realizaron los clasificadores para la misma tarea [21].

II. MÉTODO

El método definido es conducente a realizar una comparación entre las clasificaciones de SPN extraídos desde fotografías de marco claro con fotografías de marco oscuro. Tal como, resaltar con base en los resultados la usabilidad del marco oscuro para la disminución del tiempo que se emplea en la identificación de cámara fuente.

Las técnicas seleccionadas para desarrollar esta investigación, sobre la identificación de cámara fuente, se obtuvieron con trabajos publicados en revistas científicas, que manifestaron una exactitud mayor al 75%. La construcción y prueba de los algoritmos se hizo con MATLAB R2017b, en un equipo de 64bits, 8GB de RAM con Windows 10, un procesador Intel® Core™ i7-45100 y una tarjeta gráfica NVIDIA GEFORCE 2048 MB.

El algoritmo implementado para la extracción de SPN, se basó en los propuestos en [14], [17] donde el SPN, se obtiene a partir de marcos claros, sin embargo, en este estudio se usó también para la extracción del ruido desde fotografías de marco oscuro, además al algoritmo se le añadió un recorte de las fotografías a 180×180 píxeles desde el centro, con el objetivo de reducir el tamaño de la matriz de características, y así adaptar el algoritmo a las especificaciones del equipo usado en las pruebas. De este modo, en la extracción del SPN de las fotografías, presentes en las bases de datos usadas, se siguió el ALGORITMO I.

ALGORITMO I ALGORITMO PARA LA EXTRACCIÓN DEL SPN. MODIFICADO CON BASE EN [14].

1. Toma una imagen de entrada.
2. Recorta la imagen a 180×180 desde el centro.

3. Aplica una descomposición UWT de primer nivel.
4. Aplica un filtrado Wiener 2D a las sub bandas que contienen información de alta frecuencia (LH, HL, HH).
5. El SPN es obtenido tomando las sub bandas extraídas y restando las filtradas.
6. Realiza una reconstrucción de la imagen que corresponde puramente al SPN.
7. Aplica un filtrado Wiener 2D al SPN para atenuar cualquier artefacto que pueda ser insertado por la función eliminadora de ruido.
8. La imagen del SPN es concatenada para formar un vector que representará la huella del sensor de la cámara fuente.

Teniendo las firmas de los sensores de cada dispositivo existente en la base de datos *Dresden* [22] y la base de datos propia del grupo de investigación, se procede a ejecutar la metodología para llevar a cabo la clasificación de los ruidos. En esta se incluyó el desarrollo e implementación de pruebas para la clasificación no consideradas en [14]. En la TABLA I se presentan todos los tipos de clasificadores utilizados en los ensayos.

Técnica de aprendizaje de máquina	Tipo de clasificador
Decision Trees	Fine
	Medium
	Coarse
NAIVE BAYES	Gaussian
	Kernel
	Linear
	Quadratic
SVM	Cubic
	Fine Gaussian
	Medium Gaussian
	Coarse Gaussian
	Medium
KNN	Coarse
	Cosine
	Cubic
	Weighted
ENSEMBLE	Boosted Trees
	Bagged Trees
	Subspace KNN
	RUSBoosted Trees

A. Base de datos Dresden

Esta base de datos, se compone de gran variedad de fotografías tomadas con cámaras digitales compactas y Réflex. Al momento de realizar las pruebas, la base de datos contenía más de 14000 imágenes, provenientes de alrededor de 73 cámaras y 25 modelos [22].

Para este estudio, se usaron fotografías en formato JPEG provenientes de 10 modelos diferentes de cámaras, para la cual, por cada modelo, se eligieron 50 fotografías llegando así a un total de 500 imágenes usadas. La TABLA II muestra los dispositivos elegidos y las resoluciones de las fotografías tomadas de la base de datos Dresden.

TABLA II
DISPOSITIVOS DE LA BASE DE DATOS *DRESDEN* [24] ELEGIDOS PARA LAS PRUEBAS

Nombre de dispositivo	Tipo de cámara	Resolución de las fotografías (px)	Identificador
Agfa DC-733s_0	Cámara digital	3072x2304	Agf_DC_1
Canon Ixus 70_0	Cámara digital	3072x2304	Can_Ixus_1
Canon Ixus 70_1	Cámara digital	3072x2304	Can_Ixus_2
Canon Ixus 70_2	Cámara digital	3072x2304	Can_Ixus_3
Rollei RCP 7325XS_0	Cámara digital	3072x2304	Rol_RC_1
Samsung L74wide_0	Cámara digital	3072x2304	Sam_L74_1
Samsung Galaxy S2_A	Cámara de telefono	3262x2448	Sam_S2_1
Samsung Galaxy S2_B	Cámara de telefono	3262x2448	Sam_S2_2
Zte Orange Sanfrancisco_A	Cámara de telefono	1536x2048	Zte_Or_1
Zte Orange Sanfrancisco_B	Cámara de telefono	1536x2048	Zte_Or_2

B. Base de datos propia del grupo de investigación

En la realización de la galería propia, se recolectó información sobre los dispositivos móviles más utilizados en Colombia. Se tuvieron en cuenta informes del Ministerio de Tecnologías de la Información y las Comunicaciones (TIC) [23] y organizaciones que exponen datos con las marcas que tienen mayor impacto en el mercado colombiano [24] [25].

Posteriormente, se consultaron las bases de datos de las empresas que prestan servicios de telecomunicaciones en el país, con el fin de identificar cuáles eran los dispositivos más vendidos en el año 2019. Después, se buscaron estas marcas entre los estudiantes de la Universidad Tecnológica de Pereira, y se inició la creación de la galería de fotografías propia del grupo de investigación de Robótica Aplicada, en la línea Visión por Computador y Aprendizaje de Máquina.

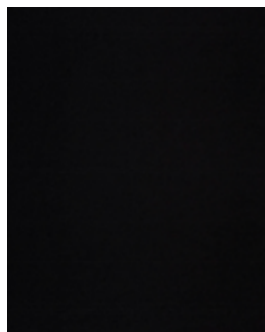
La creación de esta galería, fue realizada dentro del laboratorio del Grupo de Investigación Robótica Aplicada. Para ello se usaron 32 cámaras de dispositivos móviles, de 22 modelos distintos y 6 marcas diferentes, con las cuales se tomaron 3200 fotografías en formato JPEG. A continuación, en la TABLA III, se presentan las características de los dispositivos usados y sus respectivas resoluciones de la base de datos propia.

TABLA III
MARCAS Y MODELOS DE CELULARES USADOS EN LA
CAPTURA DE LA IMÁGENES DE LA GALERÍA PROPIA

Marca	Modelo	Identificador	Resolución
Asus	Zenofone Max	Asu_ZeMax	2448x3264
	Zenofone 4	Asu_Ze4	1088x1920
	Zenofone 4+5	Asu_Ze4+	3120x4160
iPhone	6	iPhone 6	1668x2968
	6 plus	iPhone 6plus	2448x3264
	7	iPhone 7	3024x4032
Motorola	Moto G5s plus	Mot_G5s	3120x4160
	Moto G6 plus	Mot_G6ps	3024x4032
	Moto G2	Mot_G2	1836x3264
	Moto G6	Mot_G6	2304x4096
	Moto G6 play	Mot_G6py	3120x4160
Samsung	J7 prime	Sam_J7p	3096x4128
	Edge S6	Sam_S6	2988x5312
	S5	Sam_S5	2592x4608
	Galaxy J4	Sam_J4	3096x4128
	J7 neo	Sam_J7n	3096x4128
	J2 prime	Sam_J2	2448x3264
	A30	Sam_A30	3456x4608
	J6	Sam_J6	3096x4128
A6	Sam_A6	4608x3456	
Xiaomy	Note 5	Xia_N5	3000x4000
Huawei	Y9	Hua_Y9	3120x4160

Con cada dispositivo se tomaron 50 fotografías en ráfaga de un marco oscuro, las cuales corresponden a imágenes capturadas en una caja cerrada y con su interior completamente pintado de negro, en el cual se introdujo el dispositivo, véase Fig. 1a. En esta etapa, la temperatura que se tuvo en cuenta para analizar el ruido de los sensores asociado a la variación de dicha medida, fue la promedio de la Universidad Tecnológica de Pereira [26].

De manera análoga, se capturaron 50 fotografías de marco claro. Para lograrlas se exponía una cartulina completamente blanca, se tomaron las imágenes con condiciones controladas de fuentes de luz provenientes del exterior y se tomaron como referencia los datos meteorológicos de [26]. La intensidad de luz en el ambiente fue medida con el luxómetro digital LX1010B, teniendo una intensidad promedio de 315 ± 15.8 Lux, véase Fig. 1b.



(a) Fotografía de marco oscuro



(b) Fotografía de marco claro

Fig. 1 Fotografías contenidas en la galería propia del grupo.

C. Extracción del SPN

Para la extracción del ruido en el espectro de frecuencias; a las fotografías se le aplicó un filtro Wiener que sirvió como eliminador de artefactos interpuestos por los fabricantes. Este paso se realizó con el fin de disminuir los falsos positivos. Después se efectuó una descomposición no diezmada de Wavelet a la imagen para obtener las sub bandas de frecuencia de la imagen.

Luego, con base en [14], se identificaron las sub bandas Bajas-Altas, Altas-Bajas y Altas-Altas (LH, HL, HH). A estas se les aplicó un filtro de paso bajo en forma de Wiener en dos dimensiones. En seguida, el SPN se obtuvo restando las sub bandas filtradas a las originales [17]. Finalmente, se realizó una reconstrucción de la imagen utilizando las sub bandas filtradas y la sub banda Baja-Baja (LL), con el objeto de representar el ruido como una matriz.

D. Clasificación del SPN para la identificación de Marcas de dispositivos

La clasificación del SPN para identificación de marca, contó con la elección de dos dispositivos por cada fabricante, con el fin de utilizar sus ruidos como SPN general de la marca, eligiendo así los dispositivos presentados en la TABLA IV.

TABLA IV
MARCAS Y MODELOS DE CELULARES ELEGIDOS PARA LA PRUEBA 1.

Marca	Modelo	Identificador
Asus	Zenofone Max	Asus
	Zenofone 4+5	
Huawei	Y9	Huawei
	Y9	
iPhone	6	iPhone
	6	
Motorola	Moto G6 plus	Motorola
	Moto G2	
Samsung	Edge S6	Samsung
	J6	

E. Clasificación del SPN para la identificación entre dispositivos de la misma marca

La clasificación de los SPN entre dispositivos de la misma marca, se realizó en tres etapas donde se validaron las marcas iPhone, Motorola y Samsung; eligiendo así los dispositivos presentados en las TABLA V, TABLA VI y TABLA VII.

TABLA V
MODELOS DE CELULARES IPHONE ELEGIDOS PARA LA PRUEBA 2.1.

Marca	Modelo	Identificador
iPhone	6	iPhone 6_1
	6 plus	iPhone 6plus
	7	iPhone 7_1
	6	iPhone 6_2

TABLA VI
MODELOS DE CELULARES MOTOROLA ELEGIDOS PARA LA PRUEBA 2.2.

Marca	Modelo	Identificador
Motorola	Moto G5s plus	Mot_G5s
	Moto G6 plus	Mot_G6ps
	Moto G2	Mot_G2
	Moto G6	Mot_G6
	Moto G6 play	Mot_G6py

TABLA VII
MODELOS DE CELULARES MOTOROLA ELEGIDOS PARA LA PRUEBA 2.3.

Marca	Modelo	Identificador
Samsung	J7 prime	Sam_J7p
	Edge S6	Sam_S6
	S5	Sam_S5
	Galaxy J4	Sam_J4
	J7 neo	Sam_J7n
	J2 prime	Sam_J2
	A30	Sam_A30
	J6	Sam_J6
	A6	Sam_A6

F. Evaluación de los clasificadores

En la realización de las pruebas de validez; el Toolbox Classification Learner (CL) de Matlab permitió usar el método llamado validación cruzada [20], el cual consiste en partir aleatoriamente el paquete principal de características en k secciones iguales, para usar cada una de estas, como entrenamiento y la otra en la evaluación del desempeño de los clasificadores [20]. El resultado de la validación general, será entonces el promedio de los resultados de cada paquete k . De modo que, para el proceso de validación de los clasificadores usados en este estudio, el paquete principal de características fue dividido en 5 partes iguales.

Culminado el entrenamiento y validación cruzada de cada clasificador, se realizó la matriz de confusión, y a partir de ella, con (1) se obtuvo la exactitud, que evalúa la cantidad de verdaderos positivos y verdaderos negativos, obtenidos en la clasificación. También se calculó la precisión con base en [27] donde, utilizando (2), se obtiene el porcentaje de verdaderos positivos, conseguidos dentro de los datos identificados como positivos. Además, también se consideró el tiempo de entrenamiento.

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

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

Donde:

TP: Son los verdaderos positivos arrojados en la clasificación.
TN: Son los verdaderos negativos arrojados en la clasificación.

FP: Son los Falsos positivos arrojados en la clasificación.
FN: Son los Falsos negativos arrojados en la clasificación.

III. RESULTADOS

A. Validación de la implementación UWT con base de datos Dresden

En la TABLA VIII se presentan los resultados arrojados al evaluar la técnica UWT con las familias de cada clasificador. Se puede notar que con una exactitud de 99,2%, una precisión de 99,0% y un tiempo de entrenamiento de 7134,9 segundos; el clasificador con mejor desempeño fue el SVM con núcleo cuadrático. En la Fig. 2 se muestra su matriz de confusión. Allí mismo se evidencia que el clasificador con menor desempeño fue el de tipo bayesiano con núcleo gaussiano teniendo una exactitud del 59,6%, una precisión 49,5% y un tiempo de entrenamiento de 4603,8 segundos.

TABLA VIII
RESULTADOS DE LA EVALUACIÓN CON BASE DE DATOS DRESDEN.

Clasificador		Resultados		
Familia	Kernel	Exactitud (%)	Precisión (%)	Tiempo de Entrenamiento (s)
Decision Trees	Fine	83,6	79,5	2176,7
	Medium	83,6	79,5	2708,5
	Coarse	71,4	64,3	3864,1
NAIVE BAYES	Gaussian	59,6	49,5	4603,8
	Kernel	73,0	66,3	5447,5
SVM	Linear	98,8	98,5	6291,2
	Quadratic	99,2	99,0	7134,9
KNN	Medium Gaussian	96,8	96,0	7978,6
	Fine	79,6	74,5	8822,3
	Cubic	75,4	69,3	9666,0
Emsemble	Weighted	78,8	64,7	10509,7
	Boosted Trees	91,8	89,8	11353,4
	Bagged Trees	87,4	84,25	12197,1
	Subspace Discriminant	71,2	64,0	13040,8

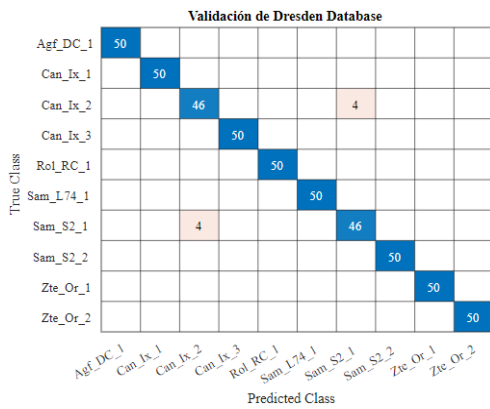


Fig. 2 Matriz de confusión del clasificador SVM Quadratic, validación con base de datos Dresden.

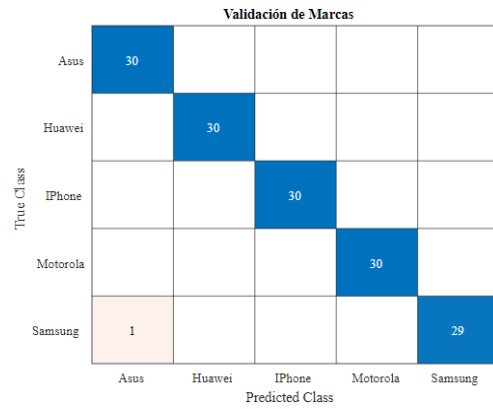


Fig. 3 Matriz de confusión del clasificador SVM Cubic para extracción de SPN desde el campo claro, en la identificación de marcas de dispositivos móviles.

B. Validación de la identificación de marcas de dispositivos

En la TABLA IX se exponen los resultados obtenidos para la identificación de marcas con SPN extraído desde fotografía de marco claro. Como se expone allí, el mejor resultado fue obtenido por el clasificador SVM con núcleo cúbico con exactitud de 99,2%, precisión de 99,0% y un tiempo de entrenamiento de 1913,3 segundos. En la Fig. 3 se muestra su matriz de confusión. Allí se muestra que el clasificador con menor desempeño fue de tipo bayesiano con núcleo gaussiano con exactitud de 59,6%, precisión de 49,5% y un tiempo de entrenamiento de 1762,6 segundos.

Mientras tanto, en la TABLA X se presentan los resultados de la clasificación para la identificación de marca con SPN extraído desde fotografías de marco oscuro. Muestra que el mejor clasificador fue el SVM con núcleo cuadrático con una exactitud de 92,8%, una precisión de 91,0 % y un tiempo de entrenamiento de 1929,2 segundos, mientras que el clasificador con menor desempeño fue el KNN con Fine, tuvo una exactitud de 49,8%, una precisión del 37,3% y un tiempo de entrenamiento de 1823,8 segundos.

TABLA IX
RESULTADOS DE LA EVALUACIÓN DE IDENTIFICACIÓN DE MARCA CON SPN EXTRAÍDO DESDE FOTOGRAFÍAS DE CAMPO CLARO.

Clasificador		Resultados		
Familia	Kernel	Exactitud (%)	Precisión (%)	Tiempo de Entrenamiento (s)
Decision Trees	Complex	83,6	79,5	1776,7
	Medium	83,6	79,5	1708,5
	Coarse	71,4	64,3	1864,1
	Gaussian	59,6	49,5	1762,60
SVM	Kernel	73,0	66,3	9504,50
	Quadratic	98,8	98,5	1963,4
	Cubic	99,2	99,0	1913,3
KNN	Fine Gaussian	96,8	96,0	2571,6
	Medium	79,6	74,5	1752,1
	Cubic	75,4	69,3	2359,9
Emsemble	Weighted	78,8	64,7	1280,9
	Bagged Trees	91,8	89,8	1813,2
	Subspace Discriminant	87,4	84,25	23794,0
	Subspace KNN	71,2	64,0	8022,1

TABLA X
RESULTADOS DE LA EVALUACIÓN DE IDENTIFICACIÓN DE MARCA CON SPN EXTRAÍDO DESDE FOTOGRAFÍAS DE CAMPO OSCURO.

Clasificador		Resultados		
Familia	Kernel	Exactitud (%)	Precisión (%)	Tiempo de Entrenamiento (s)
Decision Trees	Complex	60,6	50,8	1740,6
	Medium	61,6	52,0	1710,7
	Simple	55,4	44,3	1667,7
NAIVE BAYES	Gaussian	79,8	74,8	2533,70
	Kernel	79,8	74,8	11976,00
SVM	Quadratic	92,8	91,0	1929,2
	Cubic	91,2	89,0	1943,4
	Medium Gaussian	87,0	83,8	1919,7
KNN	Fine	49,8	37,3	1823,8
	Coarse	60,6	50,8	1676,5
	Cosine	50,0	37,5	1697,5
Emsemble	Boosted Trees	72,4	65,5	3262,8
	Bagged Trees	80,4	75,5	1319,9
	RUSBoosted Trees	63,4	54,3	3428,6

C. Validación entre dispositivos iPhone

En la TABLA XI se exponen los resultados obtenidos para la identificación de dispositivos iPhone con SPN extraído desde fotografía de marco claro. El clasificador que se comportó de mejor manera fue el de conjuntos con núcleo discriminantes en sub espacios, con una exactitud de 96,5%, una precisión de 95,3% y un tiempo de entrenamiento de 1953,3 segundos. Así mismo, el clasificador con menor desempeño fue el de conjuntos con núcleo vecino más cercano en sub espacios con exactitud de 86%, precisión de 81,3% y tiempo de entrenamiento de 1591,8 segundos.

Después de exponer los resultados de la evaluación con la base de datos de campo claro, se presentan los resultados de la evaluación con la base de datos de fotografías de campo oscuro en la TABLA XII. Donde se desempeñó como mejor clasificador el SVM con núcleo lineal con exactitud de 99%, precisión de 98% y un tiempo de entrenamiento de 605,1 segundos. En la Fig.4. Se muestra su matriz de confusión. En el mismo lugar se evidencia, el clasificador KNN ponderado con una exactitud de 50%, una precisión de 37,5% y un tiempo de entrenamiento de 1697,5 segundos.

TABLA XI

RESULTADOS DE LA EVALUACIÓN DE IDENTIFICACIÓN DE DISPOSITIVOS IPHONE CON SPN EXTRAÍDO DESDE FOTOGRAFÍAS DE CAMPO CLARO.

Clasificador		Resultados		
Familia	Kernel	Exactitud (%)	Precisión (%)	Tiempo de Entrenamiento (s)
Decision Trees	Complex	91,5	88,7	1200,2
	Medium	91,5	88,7	1313,8
	Coarse	87,0	82,7	1817,4
NAIVE BAYES	Gaussian	89,5	86,0	3238,6
	Kernel	91,5	88,7	1031,9
SVM	Linear	95,0	93,3	4233,8
	Cubic	94,5	92,3	5809,1
	Fine Gaussian	94,5	92,3	669,62
KNN	Medium	92,0	89,3	989,46
	Cubic	90,5	87,3	1169,3
	Weighted	94,5	92,7	1289,6
Emsemble	Bagged Trees	93,0	90,7	1359,2
	Subspace Discriminant	96,5	95,3	1953,3
	Subspace KNN	86,0	81,3	1591,8

TABLA XII

RESULTADOS DE LA EVALUACIÓN DE IDENTIFICACIÓN DE DISPOSITIVOS IPHONE CON SPN EXTRAÍDO DESDE FOTOGRAFÍAS DE CAMPO OSCURO.

Clasificador		Resultados		
Familia	Kernel	Exactitud (%)	Precisión (%)	Tiempo de Entrenamiento (s)
Decision Trees	Fine	52,0	36,0	1469,2
	Medium	52,0	36,0	1680,2
	Coarse	55,5	40,7	2069,0
NAIVE BAYES	Gaussian	98,5	98,0	3220,1
	Kernel	75,8	72,8	11778
	Linear	99,0	98,0	605,1
SVM	Cubic	45,6	32,0	807,74
	Fine Gaussian	83,6	79,5	1447,0
KNN	Medium	49,8	37,3	1823,8
	Cubic	60,6	50,8	1676,5
	Weighted	50,0	37,5	1697,5
Emsemble	Bagged Trees	91,8	89,8	1813,2
	Subspace Discriminant	87,4	84,25	23794
	Subspace KNN	71,2	64,0	8022,1

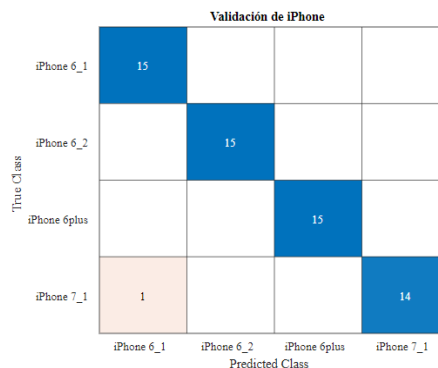


Fig.4 Matriz de confusión del clasificador SVM lineal para extracción de SPN desde el campo oscuro, en la identificación de dispositivos móviles iPhone.

D. Validación entre dispositivos Motorola

Los resultados de la clasificación de SPN extraído desde fotografías de marco claro, cuando estas son capturadas con dispositivos de marca Motorola, se muestran en la TABLA XIII. En cuanto al mejor clasificador, fue el SVM con núcleo cuadrático con una exactitud de 99,2%, una precisión de 99% y un tiempo de entrenamiento de 2331,6 segundos. En la Fig.5. Se muestra su matriz de confusión. También se muestra que el clasificador con menor desempeño fue el de conjuntos con vecinos más cercanos en un sub espacio, con una exactitud de 90,8%, una precisión de 88,5% y un tiempo de entrenamiento de 2527,7 segundos.

Validación de Motorola

True Class	Moto G2	15				
	Moto G5sps		15			
	Moto G6			15		
	Moto G6ps		1		14	
	Moto G6py				15	
		Predicted Class				
		Moto G2	Moto G5sps	Moto G6	Moto G6ps	Moto G6py

Fig.5 Matriz de confusión del clasificador SVM lineal para extracción de SPN desde el campo claro, en la identificación de dispositivos móviles Motorola.

TABLA XIII
RESULTADOS DE LA EVALUACIÓN DE IDENTIFICACIÓN DE DISPOSITIVOS MOTOROLA CON SPN EXTRAÍDO DESDE FOTOGRAFÍAS DE CAMPO CLARO.

Clasificador		Resultados		
Familia	Kernel	Exactitud (%)	Precisión (%)	Tiempo de Entrenamiento (s)
Decision Trees	Complex	94,8	93,5	2302,7
	Medium	94,8	93,5	2235,6
	Simple	94,8	93,5	2242,2
NAIVE BAYES	Gaussian	95,2	94,0	1695,6
	Kernel	95,6	94,5	6280,3
	Quadratic	99,2	99,0	2331,6
SVM	Cubic	98,8	98,5	2333,0
	Medium Gaussian	98,8	98,5	2332,9
	Medium	92,0	90,0	2272,9
KNN	Cubic	92,0	90,0	2901,5
	Weighted	91,2	89	2491,2
	Bagged Trees	99,2	99,0	1757,8
Emsemble	Subspace Discriminant	98,8	98,5	5401,3
	Subspace KNN	90,8	88,5	2527,7

Del mismo modo, en la TABLA XIV, se muestran los resultados obtenidos para la misma prueba pero realizada con el SPN extraído desde una fotografía de campo oscuro. Mostrando que el mejor clasificador es el SVM con núcleo lineal con una exactitud de 87,2%, una precisión de 84% y tiempo de entrenamiento de 635,52 segundos. Mientras que el clasificador de vecino más cercano fino tuvo una exactitud de 38,8%, precisión de 23,5% y tiempo de entrenamiento de 573,75 segundos.

TABLA XIV
RESULTADOS DE LA EVALUACIÓN DE IDENTIFICACIÓN DE DISPOSITIVOS MOTOROLA CON SPN EXTRAÍDO DESDE FOTOGRAFÍAS DE CAMPO OSCURO.

Clasificador		Resultados		
Familia	Kernel	Exactitud (%)	Precisión (%)	Tiempo de Entrenamiento (s)
Decision Trees	Fine	50,4	38,0	723,24
	Medium	50,4	38,0	828,37
	Coarse	45,6	32,0	807,74
NAIVE BAYES	Gaussian	83,6	79,5	1447,00
	Kernel	82,8	78,5	6091,00
	Linear	87,2	84,0	635,52
SVM	Coarse	86,0	82,5	608,24
	Quadratic	84,4	80,5	604,16
KNN	Cosine	52,0	40,0	568,54
	Coarse	47,2	34,0	563,34
	Fine	38,8	23,5	573,75
Emsemble	Boosted Trees	56,0	46,0	1216,20
	Bagged Trees	68,4	61,0	592,84
	Subspace Discriminant	78,0	72,5	1031,80

E. Validación entre dispositivos Samsung

Por último, en la TABLA XV se enseñan los resultados obtenidos para cada familia y cada clasificador implementado en la validación de SPN de dispositivos de la misma marca, Samsung, cuando son extraídos desde fotografía de campo claro. El clasificador con mayor desempeño fue el SVM con Kernel medio gaussiano con una exactitud de 88,7%, una precisión de 87,3% y un tiempo de entrenamiento de 637,95 segundos; mientras que el clasificador que obtuvo menor desempeño fue el árbol de decisión de tipo Coarse con una exactitud de 52,7%, una precisión de 46,8% y un tiempo de entrenamiento 576,01 segundos.

TABLA XV
RESULTADOS DE LA EVALUACIÓN DE IDENTIFICACIÓN DE DISPOSITIVOS SAMSUNG CON SPN EXTRAÍDO DESDE FOTOGRAFÍAS DE CAMPO CLARO.

Clasificador		Resultados		
Familia	Kernel	Exactitud (%)	Precisión (%)	Tiempo de Entrenamiento (s)
Decision Trees	Fine	72,0	68,5	595,28
	Medium	71,6	68,0	546,93
	Coarse	52,7	46,8	576,01
	Gaussian	68,7	64,8	2088,90
NAIVE BAYES	Kernel	75,8	72,8	11778,00
	Medium Gaussian	88,7	87,3	637,95
SVM	Linear	85,1	83,3	656,43
	Quadratic	83,1	81,0	681,64
	Fine	64,0	59,5	634,21
KNN	medium	65,6	61,3	567,98
	weighted	66	61,8	569,99
Emsemble	Boosted Trees	78,7	76,0	2582,9
	Bagged Trees	80,7	78,3	578,31
	Subspace Discriminant	84,9	83,0	7204,7

TABLA XVI
RESULTADOS DE LA EVALUACIÓN DE IDENTIFICACIÓN DE DISPOSITIVOS SAMSUNG CON SPN EXTRAÍDO DESDE FOTOGRAFÍAS DE CAMPO OSCURO.

Clasificador		Resultados		
Familia	Kernel	Exactitud (%)	Precisión (%)	Tiempo de Entrenamiento (s)
Decision Trees	Fine	58,2	53,0	611,16
	Medium	57,3	52,0	601,29
	Coarse	40,2	32,8	577,16
NAIVE BAYES	Gaussian	98,2	97,8	1906,40
	Kernel	96,9	96,5	11196,0
SVM	Linear	93,3	91,4	803,21
	Quadratic	87,8	86,3	803,60
	Medium Gaussian	96,2	95,8	860,80
KNN	Fine	50,0	43,8	641,81
	Cubic	42,4	35,3	1112,1
Emsemble	Weighted	46,9	40,3	625,44
	Boosted Trees	82,4	80,3	3106
	Bagged Trees	82,4	80,3	673,74
	Subspace Discriminant	83,3	81,3	2970,1

De manera similar, en la TABLA XVI se exponen los resultados de la evaluación de los clasificadores cuando son entrenados con el SPN proveniente desde fotografías de marco oscuro, donde se puede observar que el clasificador con mejor desempeño fue el de tipo bayesiano con Kernel gaussiano, el cual tuvo una exactitud de 98,2%, una precisión de 97,8% y un tiempo de entrenamiento en 637,95 segundos. En la Fig.6. Se muestra su matriz de confusión. A diferencia de este, los arboles de decisión de tipo Coarse tuvieron una exactitud de 40,2%, una precisión de 32,8% y un tiempo de entrenamiento de 576,1 segundos.

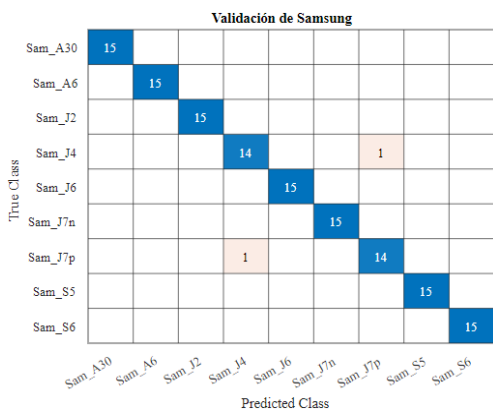


Fig.6 Matriz de confusión del clasificador Naive Bayes Gaussian para extracción de SPN desde el campo oscuro, en la identificación de dispositivos móviles Samsung.

IV. ANÁLISIS Y DISCUSIÓN

Los autores en [14] después de extraer el SPN de los dispositivos, realizan una correlación individual con la fotografía con la cual están contrastando el ruido. Para nuestro caso, la métrica utilizada se basó en la recopilación de la exactitud, precisión y tiempo de entrenamiento. En la TABLA XVII se contrastan los resultados propios con los de los autores, con el fin de presentar los valores promedios que demuestran que la evaluación con la clasificación multiclase, con la técnica de validación cruzada generan resultados superiores. Por consiguiente, el contenido obtenido en la extracción del ruido de las fotografías sí se puede tomar como una huella de cada sensor ya que existe un error considerable reducido en la clasificación.

TABLA XVII
RESULTADOS OBTENIDOS DE LA TÉCNICA UWT EN LA BASE DE DATOS PRESENTADA LA TABLA VIII.

Artículo	Métrica utilizada	Resultados		
A. R. Soobhany et al [14]	Coefficiente de correlación	Coefficiente de correlación 0,959		
Propios	Exactitud, precisión y tiempo de entrenamiento	Exactitud (%) 98,80	Precisión (%) 93,50	Tiempo de Entrenamiento (s) 7124,9

Por otro lado, como se evidencia en la TABLA XVIII, en la identificación de dispositivos, el marco claro presentó una exactitud promedio de 99,2%, mientras el marco oscuro obtuvo una exactitud promedio de 92,8%. Lo que demuestra que la exactitud de marco claro superó a la exactitud del marco oscuro por un 6,4%. De esta manera se presenta una superación de los resultados del trabajo [14] que indica de manera general un coeficiente de correlación de 0,959 sin efectuar pruebas exhaustivas como las realizadas en la presente investigación.

Sin embargo, en la TABLA XVIII, también se evidencia que, en la identificación de la misma marca, el marco oscuro presentó un menor tiempo de entrenamiento promedio, con un tiempo de 1049,0 segundos. Es decir que manifestó 400,7 segundos menos que la identificación con marco claro.

TABLA XVIII

COMPARACIÓN DE RESULTADOS PROMEDIO OBTENIDOS CON EXTRACCIÓN DE CARACTERÍSTICAS EN CAMPO CLARO Y CAMPO OSCURO.

Resultados propios	Campo de extracción del SPN	Exactitud promedio	Tiempo de Entrenamiento promedio (s)
Identificación de la marca de dispositivos	Claro	99,2	1913,3
	Oscuro	92,8	1929,2
Identificación de dispositivos de la misma marca	Claro	94,8	1449,7
	Oscuro	94,8	1049,0

V. CONCLUSIONES

A diferencia del trabajo [14], en este trabajo se propusieron dos cambios en la metodología. La extracción del SPN en fotografías de marco oscuro, además del uso de clasificación multiclase. Lo que permite tener un panorama más amplio de pruebas que se pueden realizar, como por ejemplo la discriminación entre marcas, dispositivos de la misma marca o individualización de dispositivos tecnológicamente idénticos.

La Utilización de la extracción del SPN en fotografías de marco oscuro, propuesto en esta investigación, genera resultados superiores tanto en exactitud, precisión como en tiempo de entrenamiento, comparables con la propuesta [14]; esto las hace viable para la solución del problema de identificación de cámara fuente.

Con los estudios realizados, se evidenció que las propiedades de la Transformada Wavelet No Diezmada permiten una extracción con mayor contenido del SPN, sin necesidad de realizar una descomposición multinivel de la imagen original.

También se pudo verificar que las fotografías de campo oscuro no tienen una saturación en la imagen. Por ende, tienen un contenido de ruido más puro y por lo tanto, con base en los tiempos promedio de entrenamiento, permiten una

clasificación con un buen tiempo de entrenamiento en la solución del problema de la identificación de cámara fuente.

Finalmente, los resultados obtenidos resaltan, que los clasificadores que tuvieron una mejor exactitud y precisión, no siempre fueron los que tenían cifras más bajas en el tiempo de entrenamiento o rapidez en la predicción.

AGRADECIMIENTOS

El siguiente artículo se desarrolló en el marco de la investigación "Propuesta metodológica para la identificación de imágenes digitales alteradas por copy-move, resampling y splicing" con código 3-18-8, financiada por la Universidad Tecnológica de Pereira. Por lo cual agradecemos a nuestra Universidad por fomentar la investigación desde los grupos adscritos a la vicerrectoría.

REFERENCIAS

- [1] Ying-Chu Chen, Yongjian Hu, and Chang-Tsun Li, "Further studies on forensic features for source camera identification," pp. P22–P22, 2012, doi: 10.1049/ic.2011.0119.
- [2] A. Castiglione, G. Cattaneo, M. Cembalo, and U. Ferraro Petrillo, "Experimentations with source camera identification and Online Social Networks," *J. Ambient Intell. Humaniz. Comput.*, vol. 4, no. 2, pp. 265–274, 2013, doi: 10.1007/s12652-011-0070-2.
- [3] K. S. Choi, E. Y. Lam, and K. K. Y. Wong, "Source camera identification using footprints from lens aberration," *Digit. Photogr. II*, vol. 6069, no. 852, p. 60690J, 2006, doi: 10.1117/12.649775.
- [4] A. W. A. Wahab, A. T. S. Ho, and S. Li, "Inter-Camera Model Image Source Identification with Conditional Probability Features," *Proc. IEEEJ 3rd Image Electron. Vis. Comput. Work. (IEVC 2012)*, 2012.
- [5] M. Kharrazi, H. t. Sencar, and N. Memon, "Blind source camera identification," 2004, doi: 10.1109/icip.2004.1418853.
- [6] T. Ojala, M. Pietikainen, and T. Maenpaa, "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, Jul. 2002, doi: 10.1109/tpami.2002.1017623.
- [7] V. Ojansivu and J. Heikkila, "Blur Insensitive Texture Classification Using Local Phase Quantization," *Springer-Verlag Berlin Heidelberg*, 2008.
- [8] K. Wang, "Source camera identification using support vector machines," *IFIP Int. Conf. Digit. Forensics (pp. 107-118)*. Springer, Berlin, Heidelberg, 2009.
- [9] Y. Long and Y. Huang, "Image Based Source Camera Identification using Demosaicking," Oct. 2006, doi: 10.1109/mmisp.2006.285343.
- [10] J. S. Ho, O. C. Au, J. Zhou, and Y. Guo, "Inter-channel demosaicking traces for digital image forensics," Jul. 2010, doi: 10.1109/icme.2010.5582951.
- [11] Y. Sutcu, S. Bayram, H. T. Sencar, and N. Memon, "Improvements on Sensor Noise Based Source Camera Identification," Jul. 2007, doi: 10.1109/icme.2007.4284577.
- [12] Y. Tomioka and H. Kitazawa, "Digital camera identification based on the clustered pattern noise of image sensors," Jul. 2011, doi: 10.1109/icme.2011.6012060.
- [13] C.-T. Li, "Source camera linking using enhanced sensor pattern noise extracted from images," 2009, doi: 10.1049/ic.2009.0274.
- [14] A. R. Soobhany, A. Sheikh-Akbari, and Z. C. Schreuders, "Source camera identification using non-decimated wavelet transform" *Global Security, Safety and Sustainability - The Security Challenges of the Connected World*, Springer International Publishing, 2016, pp. 125–133.

- [15] D. Freire-Obregón, F. Narducci, S. Barra, and M. Castrillón-Santana, "Deep learning for source camera identification on mobile devices," *Pattern Recognit. Lett.*, vol. 126, pp. 86–91, Sep. 2019, doi: 10.1016/j.patrec.2018.01.005.
- [16] F. Ahmed, F. Khelifi, A. Lawgaly, and A. Bouridane, "Comparative Analysis of a Deep Convolutional Neural Network for Source Camera Identification," 2019, doi: 10.1109/icgs3.2019.8688260.
- [17] J. Luka, J. Fridrich, and M. Goljan, "Digital Camera Identification From Sensor Pattern Noise," *{IEEE} Trans. Inf. Forensics Secur.*, vol. 1, no. 2, pp. 205–214, Jun. 2006, doi: 10.1109/tifs.2006.873602.
- [18] C. M. Bermúdez, "Identificación de cámara fuente mediante patrón de ruido del sensor en fotografías digitales Identificación de cámara fuente mediante patrón de ruido del sensor en fotografías digitales," 2020.
- [19] L. E. Sucar, "Clasificadores Bayesianos: de Datos a Conceptos," *Eur. Conf. Mach. Learn. Princ. Pract. Knowl. Discov. Databases*, pp. 1–18, 2008.
- [20] I. J. Tobón-gonzalez and J. A. Cortés-osorio, "Artículo en prensa / Article in press Identificação de Instrumentos Musicais de Cordas," pp. 177–193, 2018.
- [21] M. Hosni, I. Abnane, A. Idri, J. M. Carrillo de Gea, and J. L. Fernández Alemán, "Reviewing ensemble classification methods in breast cancer," *Comput. Methods Programs Biomed.*, vol. 177, pp. 89–112, 2019, doi: 10.1016/j.cmpb.2019.05.019.
- [22] T. Gloe and R. Böhme, "The dresden image database for benchmarking digital image forensics," 2010, doi: 10.1145/1774088.1774427.
- [23] M. de tecnologia de la informacion y las comunicaciones, "Boletín trimestral de las tic," 2019. [Online]. Available: https://colombiatic.mintic.gov.co/679/articles-103108_archivo_pdf.pdf.
- [24] Deloitte., "Consumo movil en colombia," 2017. [Online]. Available: <https://www2.deloitte.com/content/dam/Deloitte/co/Documents/technology-media-telecommunications/Consumo>.
- [25] Tinalu, "Colombian smartphone market continues to grow in q1 2019," May 2019. [Online]. Available: <https://www.counterpointresearch.com/colombian-smartphone-market-continues-grow-q1-2019/>.
- [26] P. opad, "Boletín meteorológico," 2009. [Online]. Available: http://media.utp.edu.co/redhidroclimatologica/boletines/utp/mensual/2009/Informe_climatologico_de_Enero_09__UTP_.pdf.
- [27] W. Zhu, N. Zeng, and N. Wang, "Sensitivity , Specificity , Accuracy , Associated Confidence Interval and ROC Analysis with Practical SAS ® Implementations K & L consulting services , Inc , Fort Washington , PA Octagon Research Solutions , Wayne," *NESUG Heal. Care Life Sci.*, pp. 1–9, 2010.



Jimmy Alexander Cortes-Osorio nació en Pereira, Colombia, en 1968. Recibió el título de pregrado en ingeniería eléctrica y de maestría en Instrumentación Física en la Universidad Tecnológica de Pereira (UTP), Pereira, en 1999 y 2009, respectivamente. También es PhD. en Ingeniería Automática de la Universidad Nacional de Colombia (UNAL), Colombia en 2020. Desde 2003 a 2012, fue profesor en el Departamento de Física de la UTP, donde ha sido trabajado tiempo completo desde 2013. Sus intereses de investigación actuales incluyen procesamiento de imágenes, análisis de movimiento de imágenes, visión artificial, aprendizaje automático, aprendizaje profundo e instrumentación y medidas. El profesor Cortes-Osorio es autor de varios libros de texto y resultados de investigación y de decenas de artículos científicos. También ha participado en eventos científicos internacionales como ponente en varios países de Asia, Europa y América. Cortés-Osorio es a su vez Miembro Afiliado de la Sociedad de Matemáticas Industriales y Aplicadas (SIAM) y del Instituto de Ingenieros Eléctricos y Electrónicos (IEEE). Actualmente, es editor en jefe de la revista científica Scientia et Technica en Colombia, Sudamérica.

ORCID: <https://orcid.org/0000-0002-0413-807X>



Camilo Muñoz Bermúdez nació en Pereira, Colombia en 1997. Es Ingeniero Físico de la Universidad Tecnológica de Pereira (UTP), desde el 2020. Las áreas de interés son: ciencias básicas, robótica, visión por computadora, machine learning y Deep learning. Obtuvo una calificación sobresaliente en su trabajo de pregrado, el cual estuvo enfocado en informática forense donde aplicó técnicas de machine learning y Deep learning.

ORCID: <https://orcid.org/0000-0001-9710-4903>