

Método de Identificación de la Fuente de Captura de una Imagen Unitono

Marco Antonio Acevedo Mosqueda¹, Ulises Noé Villegas Millán¹, Sandra Luz Gomez Coronel², María Elena Acevedo Mosqueda¹.

¹ Instituto Politécnico Nacional, SEPI ESIME Zacatenco Sección Telecomunicaciones, CDMX México, macevedo@ipn.mx, uvillegasm1400@alumno.ipn.mx, eacevedo@ipn.mx

² Instituto Politécnico Nacional, UPIITA. CDMX México, sgomez@ipn.mx

Resumen— La captura de imágenes digitales actualmente es muy sencillo, debido a la facilidad con la que se puede tener acceso a un dispositivo que cuente con una cámara fotográfica, por ejemplo, con una cámara digital, tabletas, computadoras o laptops, cámaras Web, dispositivos inteligentes, etc. De igual forma, la distribución de las imágenes capturadas, con o sin autorización del propietario, es sencilla. En el caso de la distribución de imágenes sin autorización del propietario, o en su defecto, toma de fotografías de individuos que no lo autorizan, cobra relevancia poder identificar el dispositivo con el cual se han capturado las imágenes, ya que al hacer esta identificación se puede asociar al propietario de dicho dispositivo. En este trabajo se propone un método para identificar la fuente de captura de imágenes unitono (imágenes con colores y texturas planas). El proceso consiste en tener una imagen unitono “imagen en disputa” de la cual se desea determinar su fuente de captura, un conjunto de imágenes promediadas (huella), y un conjunto de datos con las “huellas” de los posibles dispositivos que pudieran haber capturado la imagen en disputa. Cada huella está compuesta por un conjunto de imágenes capturadas por cada dispositivo, a partir de las cuales se obtiene un promedio de la distribución estadística del ruido Photo Response Non-uniformity (PRNU) de dichas imágenes. Se realiza la comparación entre las posibles huellas involucradas y la imagen en disputa calculando la similitud que existe entre ambas variables por medio de dos métodos: la distancia Mahalanobis y la correlación. Los resultados obtenidos muestran una efectividad positiva en la identificación del dispositivo que capturó la imagen en disputa.

Palabras clave— identificación, fuente de captura, PRNU, imagen unitono, distancia Mahalanobis.

I. INTRODUCCIÓN

Hoy en día existen diversos dispositivos que permiten capturar imágenes, lo que hace que la toma de una fotografía sea una tarea sencilla. Dentro de los diferentes dispositivos que se pueden usar para la captura de imágenes se tienen los dispositivos inteligentes, cámaras digitales, cámaras web, computadoras, etc., siendo fácil para cualquier persona poder capturar fotografías de cualquier objeto, paisaje, persona (con o sin autorización) y poder almacenarlas y compartirlas a través de cualquier medio digital.

Una cámara digital agrega automáticamente a las imágenes datos para su identificación, como la fecha, hora, ubicación, fuente de captura, etc. Sin embargo, estos datos pueden modificarse por cualquier persona, de forma intencional o no intencional. Por ejemplo, cuando una imagen capturada con una

cámara fotográfica se copia a una computadora, o cuando las imágenes se comparten o envían a otros dispositivos por diferentes métodos, o a través de las redes sociales.

La captura de una fotografía y su distribución, no necesariamente siempre se hace con el consentimiento del propietario. Es decir, la toma de una fotografía va acompañada del consentimiento de la persona a quien se le está tomando, y por ende el permiso de poder distribuirla por cualquier medio digital. Realizar un análisis en imágenes digitales permite determinar su origen y autenticidad, además de identificar la fuente de captura, y relacionar la fuente con su dueño. Esto es, determinar de forma específica que dispositivo se usó para capturar una imagen o imágenes determinadas, e identificar al dueño del dispositivo.

En este trabajo se presenta un algoritmo para identificar la fuente de captura de una imagen unitono, es decir, una imagen de una sola tonalidad (con colores y texturas planas). Para ello, se realizó un conjunto de datos usando cámaras de dispositivos inteligentes (celulares) y cámaras web. El proceso consistió en comparar la huella digital de una cámara digital contra el ruido de una imagen digital unitono en disputa (de la que se desea conocer su origen).

Se han desarrollado diversos trabajos con el objetivo de identificar la fuente de captura de una imagen digital. En [1] se describe un método para identificar la cámara digital que capturó una imagen digital, con base en el patrón de ruido del sensor que tienen las cámaras digitales. Se propone generar un patrón de ruido de referencia para cada cámara digital, que se toma como huella digital de identificación única. Dicha huella se obtiene promediando el ruido obtenido de múltiples imágenes mediante un filtro de eliminación de ruido. Este trabajo se basa principalmente en el ruido no uniforme de pixel (PNU, Pixel Nonuniformity), ya que es una componente del patrón de ruido que se mantiene mejor después del procesamiento de la imagen a diferencia del patrón de ruido fijo (FPN, Fixed Pattern Noise). De acuerdo con las pruebas realizadas, se sugiere obtener al menos 50 ruidos residuales. Dentro de las ventajas que se destacan de este método para obtener el patrón de referencia se tienen, 1) que no es necesario el acceso a la cámara digital, y 2) se puede aplicar a todas las

cámaras digitales, independientemente de que la cámara permita o no el acceso a la salida del sensor.

El trabajo descrito en [2] propone un algoritmo que mejora la identificación de la fuente de una imagen proveniente de dispositivos móviles. La propuesta del algoritmo de identificación de la marca y modelo de la cámara digital se basa en el método de Goljan [1], así como el uso de la transformada Wavelet para la eliminación de ruido de las imágenes. De acuerdo con los resultados reportados, se tiene un 93,65 % de identificación entre la marca y el modelo de la cámara digital.

Otro método, propuesto con el objetivo de identificar dispositivos a partir de las texturas y canales de color de las imágenes, se describe en [3]. Este método utiliza Local Binary Patterns (LBP) y Local Phase Quantization (LPQ) extraídos de las imágenes, imágenes de ruido residual y la transformada Countourlet. Toma como referencia el supuesto de que todas las imágenes capturadas con el mismo dispositivo presentan las mismas características internas, como la distorsión de la lente y el modo Color Filter Array (CFA). Las pruebas realizadas se hacen con 14 modelos de cámaras aleatorias consideradas de un conjunto de datos público. Las imágenes empleadas tienen formato “jpeg” y, de acuerdo con los resultados publicados, presentan una efectividad del 98% al 99% en promedio, considerando que todos los dispositivos son diferentes. Cuando se tienen dispositivos del mismo modelo, la efectividad baja aproximadamente a 72%.

Otro método desarrollado es el que se describe en [4], en este trabajo se plantea el uso de los filtros de extracción PNRU en los componentes de baja frecuencia y alta frecuencia de la imagen por separado, utilizando un filtro Mihcak. Posteriormente, se calcula la correlación entre el residuo de ruido y el PNRU de la cámara considerada sospechosa y se determina si la imagen proviene de dicho dispositivo. Para las pruebas, se emplearon 14594 imágenes de 57 cámaras digitales, estos datos se obtuvieron de una base de datos pública. Las efectividades obtenidas fueron del 54% y del 70%.

Ambuj Mehrish [5] propuso un algoritmo para estimar de forma robusta el ruido Photo Response Non-uniformity (PNRU) a partir de datos brutos obtenidos de manera probabilística. Debido a que no todas las cámaras proporcionan datos sin procesar como salida, se determinan los datos sin procesar a partir de la salida JPEG a través de un método de desrenderización de color probabilístico. Los datos estimados sin procesar se modelan como un proceso de Poisson y se utiliza la estimación de máxima verosimilitud (MLE) para estimar el PNRU. Para la evaluación de las pruebas, se utilizan cámaras digitales de diferentes modelos y marcas, usando diferentes tamaños de sensor, resoluciones y formatos de imagen. De cada dispositivo se usan 200 imágenes: 50 para la estimación del PNRU y 150 para la identificación de la cámara digital. Los resultados obtenidos muestran una efectividad superior al 90%

con imágenes de 512×512 , mientras que con imágenes de 128×128 la efectividad se reduce a aproximadamente el 70%.

En este trabajo se presenta un algoritmo para identificar la fuente de captura de imágenes unitono. Se desarrolló un conjunto de datos propio utilizando 17 dispositivos de captura digital diferentes, tanto dispositivos inteligentes (celulares) como cámaras web. Se diseñó un algoritmo de extracción de ruido de imágenes JPEG unitono para conformar huellas digitales de diferentes cámaras, cámaras de teléfonos celulares y de cámaras web, donde se considera un criterio de comparación entre la huella digital de una cámara y el ruido intrínseco de una imagen digital en disputa.

El resto del artículo se divide de la siguiente manera. En la sección II se presentan todos los elementos necesarios para el desarrollo del método, desde la generación del conjunto de datos hasta la descripción de la distancia de Mahalanobis y del ruido de respuesta no uniforme (PRNU), elementos fundamentales para la identificación de la fuente de captura. La descripción del método se presenta en la sección III, con énfasis en el cálculo del PRNU. Las pruebas y resultados que se llevaron a cabo para determinar la efectividad del método se incluyen en el apartado IV, y las conclusiones de dichos resultados se presentan en la sección V.

II. METODOLOGÍA

A continuación, se describen los elementos considerados para desarrollar el algoritmo propuesto.

A. Conjunto de Datos

Se generó un conjunto de datos con 17 dispositivos, de los cuales 10 fueron teléfonos celulares y 7, cámaras web. En la Tabla 1 se muestran los 17 dispositivos.

De la Tabla 1, se puede observar que hay dispositivos de marca y modelo repetidos, tanto en los celulares como en las cámaras web, con la intención de determinar si es posible distinguir con el algoritmo el dispositivo que capture una imagen, a pesar de tener otro dispositivo de la misma marca y modelo y que pudiera confundir al algoritmo. Con cada uno de los 17 dispositivos se capturaron 100 imágenes, obteniendo un total de 1700 imágenes, tomadas bajo las siguientes condiciones.

- Estudio semiprofesional de fotografía, con dimensiones de 50×50 cm, con iluminación led interna que ilumina uniformemente las paredes del estudio. Asimismo, estas paredes internas presentan una superficie metálica, lo que permite contar con un total de 6 paredes que evitan la interferencia de sombras o de luminosidad externa.
- Orientación: horizontal. Para la captura de imágenes se utilizó un trípode en el que se fijaron los dispositivos para

mayor estabilidad. Sin embargo, al colocar los celulares de manera vertical, estos capturaban parte de las paredes del estudio fotográfico debido a que la superficie gris que se utilizó era más pequeña que la pared del estudio; por lo que, se decidió que la orientación fuera horizontal.

- Distancia del dispositivo a la superficie: 12 cm - 15 cm. Ya que no todas las cámaras cuentan con el mismo angular, se determinó una distancia promedio para que, sin importar el tipo de dispositivo a usar, la imagen capturada no incluyera parte de las paredes.
- Superficie: lámina de fotografía en color gris. Se consideró este tipo de superficie unitono, ya que el PRNU se obtiene mejor en imágenes con una sola tonalidad; es decir, que no existan sombras o relieves, ni algún rayo de luz ni manchas.
- Modo de captura: ráfaga y manual. El modo de captura de ráfaga solo aplica a los dispositivos celulares ya que las cámaras web no cuentan con esta opción. Dicho modo consiste en mantener presionado el botón de captura para capturar imágenes de forma continua, no una por una. Dependiendo el tipo de celular, la ráfaga es de distintas cantidades de imágenes. El modo manual se utilizó en las cámaras web, por lo que se tomó imagen por imagen. Ambos métodos de captura se realizaron de forma alámbrica para no interferir con el movimiento y causar perturbaciones en la imagen. Los dispositivos nunca se tocaron en el momento de la captura, en el caso de los celulares se realizó con la ayuda de unos audífonos y en el de las cámaras conectándolas a una computadora.
- Iluminación: focos de fotografía. Ésta se controló con las luces integradas del estudio fotográfico, ya que iluminan internamente de manera uniforme todas las paredes del estudio fotográfico, lo cual ayuda a que no exista ninguna sombra que pueda interferir en la captura de la imagen. Además, se utilizaron unas lámparas de fotografía que iluminaban por el exterior al estudio y de esta manera evitar en su totalidad algún tipo de sombra.

B. Imágenes Unitono

Una imagen unitono se caracteriza por su orientación a un concepto minimalista, buscando un estilo de simplicidad, con colores y texturas planas, de manera que se perciban visualmente limpias. Estas imágenes no deben incluir sombras ni relieves y deben ser capturadas en un ambiente controlado. En la Fig. 1 se muestra un ejemplo de imagen unitono.



Fig. 1 Ejemplo de imagen unitono.

TABLA I
DISPOSITIVOS EMPLEADOS PARA GENERAR EL CONJUNTO DE DATOS.

Marca	Modelo	Tipo
1. Samsung	Galaxy A13	Celular
2. Samsung	Galaxy A11	Celular
3. Samsung	Galaxy A01	Celular
4. Samsung	Galaxy Note 9 1	Celular
5. Samsung	Galaxy Note 9 2	Celular
6. Samsung	Galaxy S22	Celular
7. Samsung	GalaxyS21Fe	Celular
8. Apple	Iphone 13 Pro	Celular
9. Huawei	Y9	Celular
10. Motorola	G8 Plus	Celular
11. Asus	C3 1	Cámara Web
12. Asus	C3 2	Cámara Web
13. Logitech	C920 Pro 1	Cámara Web
14. Logitech	C920 Pro 2	Cámara Web
15. Genérico	Full HD Web cam	Cámara Web
16. Logitech	StreamCam Plus 1	Cámara Web
17. Logitech	StreamCam Plus 2	Cámara Web

C. Ruido de Respuesta No Uniforme (PRNU; Photo Response Non Uniformity).

Las imágenes digitales son propensas a distintos tipos de ruido. Dicho ruido puede deberse a la transmisión electrónica de datos de la imagen, al sensor de la cámara o al lente de esta. Hay diferentes fuentes de imperfecciones y ruido introducidas en las etapas del proceso de generación de la imagen en la cámara. Esto se puede deber al ruido de disparo que es aleatorio y, generalmente, al patrón de ruido que es determinista y se mantiene aproximadamente igual si se toman varias fotografías de la misma escena. Dicho patrón de ruido en una imagen se refiere a cualquier patrón espacial que no cambia de una imagen a otra y está compuesto por el ruido espacial debido a la diferencia de respuesta de cada pixel a la señal incidente o ruido de respuesta no uniforme (PRNU). Se genera de la diferencia de sensibilidad a la intensidad de la luz de los foto-sensores, debido al proceso de fabricación, generando un patrón de ruido para cada sensor que imprime en cada fotografía digital que se genera [6]. El ruido PRNU es la componente dominante del patrón de ruido de las imágenes y es un ruido multiplicativo dependiente. Está formado principalmente por la uniformidad de pixel, Pixel Non-Uniformity (PNU), y los defectos de baja

frecuencia como la configuración del zoom y la refracción de la luz en las partículas de polvo y lentes. Es importante mencionar que la uniformidad de sensibilidad es diferente entre los canales o capas de la imagen (Rojo, Verde y Azul, RGB por sus nombres en inglés, respetivamente), en la región de respuesta lineal. Su singularidad asegura que los ruidos de los sensores extraídos de diferentes cámaras no estén correlacionados, incluso cuando provienen del mismo modelo de cámara. Por ello, se considera al PRNU como la forma más eficaz de asignar una cámara a una imagen.

D. Distancia de Mahalanobis

Es una medida de distancia alternativa a la distancia euclídea. La distancia euclídea entre dos puntos $A(x_1, y_1)$ y $B(x_2, y_2)$ a la longitud del segmento de la recta que tiene por extremos A y B , y está definida por la expresión (1):

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (1)$$

La utilidad de la distancia de Mahalanobis radica en que es una medida de similitud entre dos variables aleatorias multidimensionales, que considera la correlación entre ellas [7]. La distancia Mahalanobis entre dos variables aleatorias con la misma distribución se define de acuerdo con (2):

$$d_m = (\vec{x}; \vec{y}) = \sqrt{(\vec{x} - \vec{y})^T \Sigma^{-1} (\vec{x} - \vec{y})} \quad (2)$$

donde:

x es el vector de datos.

y es el vector de medias de la distribución.

Σ^{-1} es la matriz de covarianza inversa de la distribución.

III. MÉTODO DE IDENTIFICACIÓN DE FUENTE DE CAPTURA DE UNA IMAGEN DIGITAL

A. Extracción de Ruido de una Imagen Digital

Para identificar la fuente de captura de una imagen son necesarias dos variables, la imagen unitono y la huella compuesta de un conjunto de imágenes, donde a cada una de las imágenes del conjunto de datos, se le extrae el ruido para realizar dicha comparación. El proceso de extracción de ruido comienza con la selección de una imagen, almacenando los datos de dicha imagen en una matriz del tamaño de la resolución por las tres capas de la imagen. Posteriormente, se selecciona la capa verde, ya que aporta más información según los filtros de color. Después, se obtiene el ruido de la imagen, tomando como parámetro los valores de la capa verde. Finalmente, se obtiene la imagen sin ruido, esto es, a los valores de la capa verde se le resta el ruido, obteniendo las intensidades de los píxeles sin ruido; a esto se le denomina $I0$. Por otro lado, para obtener el

ruido PRNU, se aplica un filtro adaptativo, lo que permite preservar los detalles importantes de la imagen.

B. Huellas Digitales de las Cámaras Empleadas

Para la generación de la huella digital de cada uno de los 17 dispositivos utilizados para generar el conjunto de datos (Tabla 1), se toman como referencia 15 imágenes unitono capturadas con cada uno de ellos. Se calculan los valores $I0$ y PRNU de cada imagen unitono. Se agrupan en matrices y se unen en un vector para tener el Patrón de Huella. Este proceso se repite para cada dispositivo (cámaras web y teléfonos celulares).

Para asociar una imagen digital con la cámara con la que fue tomada, se compara la huella digital de las diferentes cámaras por las que pudo haber sido capturada con el ruido de la imagen. Esta comparación se realiza mediante la distancia de Mahalanobis, ya que el resultado indicará la similitud entre la huella digital de una cámara en particular y el ruido de una imagen en disputa ($I0$ y PRNU). De los valores obtenidos en cada comparación, el más pequeño indicará el dispositivo con mayor probabilidad de corresponder a la fuente de captura de la imagen en disputa analizada.

IV. PRUEBAS Y RESULTADOS

Se realizaron pruebas con el conjunto de datos creado para evaluar la efectividad de identificar una fuente de captura mediante la distancia de Mahalanobis. El proceso se realizó en cada uno de los 17 dispositivos. Esto es, se eligió una imagen unitono del conjunto de datos y se comparó con las huellas digitales (Patrón) de cada fuente de captura. Este procedimiento se realizó en varias ocasiones para determinar la efectividad del método y realizar pruebas con cada dispositivo. En la Fig. 2 se muestra un ejemplo de los valores obtenidos al realizar una prueba seleccionando imágenes unitono en disputa del dispositivo *Samsung Galaxy S22*. Cada uno de los valores obtenidos de la distancia Mahalanobis se agrupó en celdas para identificar los valores más pequeños y obtener un promedio de eficiencia en cada prueba. Las 10 filas representan la cantidad de dispositivos sospechosos posibles de haber capturado la imagen, y las 80 columnas son las diferentes imágenes en disputa (en la figura solo se representan de la imagen unitono 21 a la imagen unitono 32).

De la Fig. 2, se resaltan en verde los valores más pequeños, lo que indica que las imágenes unitono en disputa pertenecen a la fuente de captura *Samsung Galaxy S22*. Asimismo, se resaltan en amarillo los valores correspondientes a otro dispositivo Samsung, con la intención de mostrar que, a pesar de ser del mismo modelo, con el método sí se distingue la fuente de captura que generó la imagen en disputa.

	21	22	23	24	25	26	27	28	29	30	31	32
1. Samsung Galaxy A13	14.993	15.018	14.966	14.993	15.056	15.009	15.028	15.127	15.070	15.021	15.084	15.037
2. Samsung Galaxy A11	23.607	23.670	23.538	23.606	23.766	23.647	23.695	23.543	23.800	23.676	23.834	23.715
3. Samsung Galaxy A01	49.321	49.377	49.256	49.320	49.467	49.357	49.400	49.629	49.496	49.383	49.528	49.421
4. Samsung Galaxy Note 9_1	5.634	5.608	5.660	5.635	5.560	5.617	5.596	5.479	5.547	5.606	5.532	5.583
5. Samsung Galaxy Note 9_2	6.732	6.694	6.783	6.734	6.622	6.707	6.675	6.504	6.634	6.690	6.582	6.657
6. Samsung Galaxy S22	0.028	0.027	0.026	0.023	0.005	0.030	0.022	0.007	0.015	0.022	0.013	0.028
7. Samsung Galaxy S21Fe	1.149	1.160	1.138	1.149	1.175	1.156	1.164	1.205	1.181	1.161	1.187	1.167
8. iPhone 13 Pro	10.687	10.706	10.672	10.687	10.726	10.699	10.710	10.773	10.738	10.708	10.748	10.714
9. Huawei Y9 2019	11.288	11.314	11.270	11.289	11.338	11.304	11.320	11.401	11.356	11.318	11.369	11.323
10. Motorola G8 Plus	14.035	14.053	14.008	14.035	14.096	14.051	14.085	14.165	14.109	14.052	14.123	14.077

Fig. 2 Resultados del método de identificación de fuente de captura de una imagen unitono usando la distancia Mahalanobis.

A continuación, se presentan los resultados de diferentes pruebas realizadas.

Se determinó realizar una prueba con el conjunto de datos de los 10 teléfonos móviles. El número de imágenes unitono por dispositivo de captura fue de 100, de las cuales 15 se usaron como referencia y 80 en disputa. Dentro de estos dispositivos se encuentran un par de gemelos, es decir, que ambos dispositivos son de la misma marca y modelo pero siendo 2 dispositivos físicos diferentes.

Se realizó la extracción de la intensidad de los píxeles (IO) y PRNU de todas las imágenes, y se obtuvieron un total de 1000 ruidos. Como se explicó en el apartado III, con la intensidad de los píxeles y el ruido PRNU de las 15 imágenes de referencia se conformó la huella digital de cada dispositivo, teniendo un total de 10 huellas. Posteriormente, se calculó la distancia de Mahalanobis de todas las imágenes en disputa respecto de cada huella digital. En la Tabla II se muestran las efectividades promediadas para las imágenes en disputa de cada dispositivo, donde las filas representan el conjunto de imágenes en disputa y las columnas los dispositivos de captura (la numeración de las columnas corresponde a los primeros 10 dispositivos presentados en la Tabla I), con una efectividad general del **94.75%**.

TABLA II
PORCENTAJE DE ACIERTOS DE IMÁGENES UNITONO DEL CONJUNTO DE DATOS DE LOS TELÉFONOS CELULARES.

	1	2	3	4	5	6	7	8	9	10
1	76.25	0	0	0	0	0	0	0	0	23.75
2	0	100	0	0	0	0	0	0	0	0
3	0	0	100	0	0	0	0	0	0	0
4	0	0	0	100	0	0	0	0	0	0
5	0	0	0	0	100	0	0	0	0	0
6	0	0	0	0	0	100	0	0	0	0
7	0	0	0	0	0	0	100	0	0	0
8	0	0	0	0	0	0	0	100	0	0
9	2.50	0	0	0	0	0	16.25	26.25	82.75	15
10	11.25	0	0	0	0	0	0	0	0	88.75

De la Tabla II, se puede inferir que la alta efectividad, es por el uso de imágenes unitono, ya que están capturadas en ambiente controlado, sin la interferencia de rayos de luz o sombras.

Este mismo procedimiento se repitió utilizando las cámaras web indicadas en la Tabla I. Para este caso se consideran 7 fuentes de captura, por cada dispositivo se consideraron 100 imágenes en disputa, de las cuales 15 son de referencia y 80 en disputa. Dentro de estos dispositivos se encuentran 3 pares de

gemelos; esto quiere decir que son de la misma marca y modelo, pero son dispositivos físicos diferentes.

Se realizó la extracción de la intensidad de los píxeles (IO) y PRNU de todas las imágenes; en este caso, se obtuvo un total de 700 ruidos.

En la Tabla III se muestran las efectividades promediadas de las imágenes en disputa para cada dispositivo, donde las filas representan el conjunto de imágenes en disputa y las columnas, los dispositivos de captura (la numeración corresponde a los dispositivos presentados en la Tabla I). Se logró una efectividad general del **96,78 %**.

Como en el caso de los teléfonos celulares, la efectividad es alta. Sin embargo, en este caso también se puede inferir que la efectividad es mayor porque la cantidad de dispositivos es menor.

TABLA III
PORCENTAJE DE ACIERTOS DE IMÁGENES UNITONO DEL CONJUNTO DE DATOS DE LAS CÁMARAS WEB.

	1	2	3	4	5	6	7
1	100	0	0	0	0	0	0
2	0	100	0	0	0	0	0
3	22.50	0	77.50	0	0	0	0
4	0	0	0	100	0	0	0
5	0	0	0	0	100	0	0
6	0	0	0	0	0	100	0
7	0	0	0	0	0	0	100

Finalmente, se realizó una prueba con todos los 17 dispositivos empleados en el conjunto de datos. De igual forma, el número de imágenes unitono por dispositivo de captura fue de 100, de las cuales 15 fueron de referencia y 80 en disputa. Se realizó la extracción de la intensidad de los píxeles (IO) y PRNU de todas las imágenes, es decir, que se obtuvieron un total de 1700 ruidos y 17 huellas.

En la Tabla IV se muestran las efectividades promediadas para las imágenes en disputa de cada dispositivo, donde las filas representan el conjunto de imágenes en disputa y las columnas, los dispositivos de captura, con una efectividad general del **88,16 %**.

De la tabla IV se puede observar que, a pesar de tener una mayor cantidad de fuentes de captura, la efectividad sigue siendo alta.

TABLA IV PORCENTAJE DE ACIERTOS DE IMÁGENES UNITONO DEL CONJUNTO DE DATOS DE TODOS LOS DISPOSITIVOS EMPLEADOS.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	72.25	0	0	0	0	0	0	0	0	23.75	0	0	0	0	0	0	0
2	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	77.50	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	65	8.75	0	0	0	0	0	12.50	13.75
10	11.25	0	0	0	0	0	0	0	0	88.75	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0	0	0
13	0	0	25	0	0	0	0	0	0	0	0	0	75	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	37.5	62.5	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100	0
17	0	0	0	0	0	0	0	0	0	68.75	0	0	0	0	0	0	31.25

A. Comparación con otros métodos similares.

No es posible hacer una comparación con los trabajos presentados en el estado del arte tal cual, dado que los conjuntos de datos empleados por otros métodos son diferentes. Sin embargo, a fin de mostrar la efectividad del método propuesto en este trabajo, se presentan las siguientes consideraciones al compararlo con otros trabajos.

El método presentado en [1] generó huellas digitales de los dispositivos empleados a partir de 50 imágenes de referencia. En este trabajo se disminuyó la cantidad de imágenes unitono a usar para obtener las huellas de cada fuente de captura, y aun así se obtuvieron efectividades altas.

El método descrito en [2] utilizó solo 8 dispositivos para el análisis de identificación de la fuente de captura y reportó una efectividad del 93,62%. En esta propuesta, la efectividad es del 94,75% con 10 dispositivos y del 96,78% con 7 dispositivos. Cuando se integraron los 17 dispositivos la efectividad disminuyó a 88.17%, que a pesar de ser menor a 90%, sigue siendo una efectividad alta y se obtiene ya con diferentes fuentes de captura, además de tener dispositivos gemelos. Considerando este último punto, se hace referencia al trabajo presentado en [3], ya que reporta una alta efectividad (98% - 99%) al usar diferentes fuentes de captura. Sin embargo, cuando usa dispositivos de la misma marca, la efectividad disminuye hasta el 72 %. En este caso, la propuesta descrita en este trabajo, a pesar de usar dispositivos de la misma marca y modelo, mantiene una alta efectividad (mayor al 90%)

En la Tabla V se presenta un resumen de los aspectos más sobresalientes de cada propuesta.

TABLA V
COMPARACIÓN CON OTROS TRABAJOS.

TRABAJO	IMÁGENES EMPLEADAS	DISPOSITIVOS EMPLEADOS	EFFECTIVIDAD (%)
[1]	50	-	-
[2]	-	8	93,62
[3]	-	14	99 - 98
[4]	14594	57	54 - 70
PROPUESTA	1700	17	88.17

Finalmente, se realizó una prueba utilizando el conjunto de datos VISION [8]. Este conjunto público contiene 34,427 imágenes de diferentes tipos y 1,914 videos provenientes de 35 teléfonos inteligentes de 11 marcas distintas. Se seleccionaron solo 10 dispositivos móviles, y de cada uno 50 imágenes, de las cuales 15 se tomaron como referencia y 30 en disputa. Se realizó el mismo proceso de identificación que se aplicó al conjunto de datos propio. El promedio general obtenido de identificación fue de 36,85%. Esto se atribuye a que las imágenes del conjunto de datos no son en su totalidad unitono, tal como se muestra en la Fig. 2 y Fig. 3. En la Fig. 2 se observa una superficie rugosa o con bordes, mientras que en la Fig. 3 es una escena completamente diferente, que muestra nubes y el cielo. Esto demuestra que las imágenes con alta dispersión, alto contraste y falta de uniformidad, no son útiles para la extracción de ruido.



Fig. 2. Ejemplo de imagen del conjunto de datos VISION, tomada con un iPhone 4s.

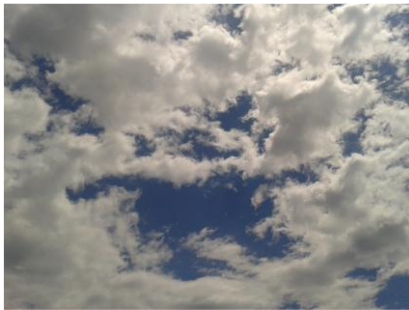


Fig. 3. Ejemplo de imagen del conjunto de datos VISION, tomada con un dispositivo Samsung S3 Mini 2.

IV. CONCLUSIONES

Una de las consideraciones principales de este trabajo es el tipo de imágenes empleadas, que son unitono (estilo simple, de una sola tonalidad y sin ningún tipo de marcas, relieves, puntos ni rayos de luz). Con el uso de estas imágenes se logró una alta efectividad en la identificación de 17 fuentes de captura diferentes, entre ellas dispositivos de la misma marca y modelo, y considerando que se usaron tanto teléfonos móviles como cámaras web. La cantidad de imágenes con las que se trabajó fue de 900, de las cuales se extrajeron 900 ruidos, con los que se formaron las huellas digitales de los dispositivos de captura.

El uso de la distancia Mahalanobis muestra resultados óptimos para comparar dos variables, es decir, determinar la similitud entre dos variables aleatorias multidimensionales: en este caso, el ruido PRNU de cada imagen y la huella digital de los posibles dispositivos que pudieron haber realizado la toma.

Como trabajo futuro, se debe estudiar más cómo hacer el tratamiento de imágenes que no son unitono y como extraer el ruido de ellas, para lograr una identificación exitosa del dispositivo que la haya capturado.

AGRADECIMIENTOS

Los autores agradecen al Instituto Politécnico Nacional todo el apoyo brindado para la realización de este trabajo. Asimismo, agradecen al Sistema de Becas por Exclusividad SIBE y a la Secretaría de Ciencia, Humanidades, Tecnología e Innovación a través del Sistema Nacional de Investigadoras e Investigadores (SNII).

REFERENCIAS

- [1] J. Lukas, J. Fridrich, and M. Goljan, "Digital camera identification from sensor pattern noise," *IEEE Transactions on Information Forensics and Security*, vol. 1, no. 2, pp. 205–214, 2006.
- [2] Rosales Corripio, J. Algoritmo de identificación de fuente en imágenes digitales de dispositivos móviles. (Universidad Complutense de Madrid, 2013).
- [3] J B. Xu, X. Wang, X. Zhou, J. Xi and S. Wang, "Source camera identification from image texture features," *Neurocomputing*, vol. 207, pp. 131–140, 2016.
- [4] GUPTA, Bhupendra; TIWARI, Mayank. "Improving source camera identification performance using DCT based image frequency components dependent sensor pattern noise extraction method," *Digital Investigation*, vol. 24, pp. 121–127, 2018.
- [5] MEHRISH, Ambuj; SUBRAMANYAM, A. Venkata; EMMANUEL, Sabu. "Robust PRNU estimation from probabilistic raw measurements" *Signal Processing: Image Communication*, vol. 66, pp. 30–41, 2018.
- [6] A. O. López-López, M. Nakano-Miyatake, and M. Cedillo-Hernández, "Identificación de dispositivo generador de fotografías digitales mediante la correlación del ruido de sensor," *CONAMTI*, pp. 40–45, 2018.
- [7] M. T. E. Portillo and J. A. S. Plata, "P. ch. mahalanobis y las aplicaciones de su distancia estadística," *CULCyT: Cultura Científica y Tecnológica*, vol. 5, no. 27, pp. 13–20, 2008.
- [8] D. Shullani, M. Fontani, M. Iuliani, O. Al Shaya, and A. Piva, "Vision: a video and image dataset for source identification," *EURASIP Journal on Information Security*, vol. 2017, no. 1, pp. 1–16, 2017.