

Procedimiento para la Verificación Automática de Firmas Manuscritas Off Line

David Santos Mauricio Sánchez, Arthur Mauricio Delgadillo, Luis Alberto Estrada Huamán

Sector de la técnica

La presente invención sobre la verificación automática de firmas manuscritas *off-line*, es de amplio ámbito de aplicación, éste va desde la verificación (cotejo) en cheques, documentos bancarios, notariales, contratos en general, a la masiva verificación en el sector electoral. En todas estas actividades se necesita validar que la firma manuscritas registradas en un soporte físico (hoja de papel) o electrónico sea genuina, es decir, firmada por el titular (firmante); mediante la cual se le identifica nominativamente, expresa su consentimiento o no, aun cuando se encuentre distante o ausente.

Estado de la técnica

La firma ológrafa o simplemente firma es definida legalmente como el nombre y/o un gráfico manuscrito de un individuo, y considerada como uno de los atributos personalísimos del individuo. Desde hace siglos las sociedades utilizan la firma como medio para identificar al firmante y/o expresar su consentimiento a través del cual se validan distintos actos, soportados en documentos, en muchos de los cuales su registro es obligatorio (cheques bancarios, contratos, tarjetas de crédito, procesos electorales, notariales, entre otros). Así mismo los documentos validados por la firma, se pueden conservar, tratar y transmitir, considerándose su contenido válido en tiempo, lugar y modo (Soporte físico o digital, incluyendo micro-formas legisladas por el Decreto Legislativo N° 681). Por todo ello la firma ha sido y sigue siendo un instrumento atractivo para la usurpación de identidad o fraude en los contenidos desde tiempos inmemoriales.

Para controlar o disminuir los riesgos, la autoridad como los comerciantes han apelado a múltiples disciplinas científicas. La grafotecnia es una de ellas. Ella tiene por finalidad el estudio de documentos físicos, a fin de determinar su autenticidad o falsedad, incluyendo el análisis de la escritura, la tinta, el papel, las impresiones, los sellos y las firmas. En particular la grafotecnia nos permite examinar y cotejar la autenticidad o autoría de los grafismos personales, como son la firma y la escritura. La verificación de la firma manuscrita es una modalidad biométrica conductual abordada tradicionalmente en el área forense y realizada por lo general por peritos expertos en grafotecnia. Pero por diversos factores, este puede incluir fallas, y ciertamente éstas se incrementan en ausencia de expertos. Además, un perito no siempre se encuentra disponible en las instituciones que lo requieren, aún más tratándose de verificación de volúmenes

importantes de contenidos o firmas, por ejemplo en consultas populares que requieren de este tipo de verificación. De otro lado, asistimos a un incremento exponencial de la verificación de firmas requeridas por la sociedad debido a la expansión y desarrollo de aquello que ha devenido el Modelo de desarrollo dominante: la Sociedad de la Información y del Conocimiento, que hacen que la demanda de estos procesos sea cada vez mayor a la oferta brindada por los Peritos. Todo ello propuestas de nuevas formas o procedimientos para la verificación automática de firmas *off line* y de aplicativos que lo implementen.

El proceso de verificación de firmas manuscritas se divide en verificación estática (*off-line*) y verificación dinámica (*on-line*). En la verificación estática solo se considera las características de la firma ya realizada, en tanto, la verificación dinámica incluye, además, otras características que se pueden obtener durante el proceso de realización de la firma. Actualmente el proceso de verificación en el Perú es solo estática. La presente invención se refiere a la verificación estática u *off line*, que en adelante llamaremos **verificación de firmas**.

El proceso de verificación de una firma actualmente se realiza de forma manual, en ella, una persona física, con conocimientos en grafotecnia determina su validez, siguiendo un procedimiento probado por la experiencia. Primero, revisa varias firmas genuinas del firmante, también denominado titular de la firma. Segundo, extrae características de dichas firmas (V.g.: fineza de las curvas, presión, espaciamiento, inclinación, entre otras). Tercero, revisa la firma a verificar y extrae sus características. Cuarto, compara las características de las firmas genuinas con las características de la firma a verificar. Quinto, concluye: sí todas las características corresponden entonces la firma es genuina (es decir la firma corresponde al firmante); en caso contrario, es falsa.

Sin embargo, desde hace décadas, existen tecnologías disponibles para la verificación en forma automática a través del uso de programas informáticos, capaces de extraer características grafotécnicas de las firmas del firmante, y a través de algoritmos, autenticar su validez. No obstante, estas tecnologías no son usadas en el Perú, y en general su precisión, no es confiable.

Descripción detallada de la invención

En la descripción de la invención usaremos las notaciones siguientes:

P: identifica el procedimiento,

D: identifica entradas o salida de los procedimientos, pueden ser imágenes (impresas o digitales) o información digital o impresa.

P01: Procedimiento general de verificación automática de firmas

Este procedimiento tiene por finalidad realizar la verificación automática de la firma, y consta de 2 procedimientos secuenciales (Fig. 1): “Construcción del modelo de verificación de un firmante” (P02), y “Verificación Automática” (P03).

P01 propone los siguientes pasos: Primero, se reciben las “Firmas genuinas de un firmante” (D01) que se encuentra en forma impresa o digitalizada. Segundo, se aplica P02 sobre D01 y se obtiene un “Modelo de verificación” (D02). Tercero, se aplica P03 usando D02 sobre la “Firma a verificar” (D03), obteniéndose un reporte con alguna de dos respuestas: “Genuina o Falsa” (D04).

P02: Construcción del Modelo de Verificación de un Firmante

Este procedimiento tiene por finalidad construir D02, y consta de 6 procedimientos (Fig. 2): “Digitalización” (P04), “Pre-procesamiento” (P05), “Extracción de características” (P06), “Existen firmas” (P07), “Registra características” (P08) y “Genera modelo de Verificación” (P09). P04 no se aplica cuando las firmas de D02 se encuentran digitalizadas.

P02 propone los siguientes pasos: Primero, aplica P04 sobre D01 y obtiene la “Firma digitalizada” (D05). Segundo, aplica P05 sobre D05 y obtiene la “Firma pre-procesada” (D06). Tercero, aplica P06 sobre D06 y obtiene las “Características de la firma” (D07). Cuarto, aplica P07 que genera una de dos respuestas “Si” o “No”. Quinto, si la respuesta del paso anterior es “Si” entonces aplicar P08 y repetir todos los pasos anteriores; de lo contrario, saltar al siguiente paso. Sexto, aplicar P09 usando D09 y obtener D02 (Modelo de verificación para el firmante).

P03: Verificación automática.

Este procedimiento tiene por finalidad verificar si una firma dada es “Genuina o Falsa. Consta de 4 procedimientos (Fig. 3): P04, P05, P06, y “Verificador” (P10).

P03 propone los siguientes pasos: Primero, se aplica P04 sobre D03 y se obtiene la “Firma digitalizada V” (D09). Segundo, se aplica P05 sobre D09 y se obtiene la “Firma pre-procesada V” (D10). Tercero, se aplica P06 sobre D10 y se obtiene “Características V” (D11). Cuarto, se aplica P10 sobre D11 considerando D02, obteniendo el reporte de “Genuina o Falsa” (D04). Si hay más de una firma a verificar P08 deberá extraer el modelo del firmante correspondiente a través de un identificador del firmante, información que deberá ser introducida en D03.

P04: Digitalización

Este procedimiento tiene por finalidad obtener la firma digitalizada a partir de una firma impresa, y sólo se aplica si ésta no está digitalizada. Por ejemplo, P04 recibe D01 y genera D05.

P04 consiste en escanear una firma impresa a través de un escáner. Para este fin, la imagen debe ser colocada en la misma dirección de sus ejes, y debe usarse en lo posible el mismo tipo de escáner para todo el procedimiento. Es deseable que todas las imágenes digitalizadas tengan la misma resolución y cuanto mayor sea ésta mejor será la imagen a procesar, al menos 300 ppp (puntos por pulgadas, en ingles dpi – dots per inch).

P05: Pre-procesamiento

Este procedimiento tiene por finalidad obtener la firma pre-procesada. Una firma pre-procesada es un conjunto de imágenes de firmas del cual pueden extraerse características representativas. Por ejemplo, P05 recibe D09 y genera D10.

P05 consta de 6 procedimientos: “Eliminar ruido fondo” (P11), “Rectángulo mínimo” (P12), “Eliminar ruido marco” (P13), “Eliminar ruido isla” (P14), “Afinar” (P15), y “Normalizar” (P16).

P05 sigue las siguientes acciones para cada firma (D12) (Fig. 4): Primero, aplica P11 sobre D12 y obtiene “Firma sin ruido de fondo” (D13). Segundo, aplica P12 sobre D13 y obtiene la “Firma en rectángulo mínimo” (D14). Tercero, aplica P13 sobre D14 y obtiene “Firma sin ruido marco” (D15). Cuarto, aplica P14 sobre D15 y obtiene “Firma sin ruido isla” (D16). Quinto, aplica P12 sobre D16 y obtiene “Firma sin ruido y en rectángulo mínimo” (D17). Sexto, aplica P15 sobre D17 y obtiene “Firma afinada en rectángulo mínimo” (D18). Séptimo, aplica P16 sobre D17 y obtiene “Firma sin ruido y normalizada” (D19). Octava, aplica P15 sobre D19 y obtiene “Firma afinada y normalizada” (D20).

La Fig. 15 ilustra el resultado de aplicar este procedimiento.

P06: Extracción de Características

Este procedimiento tiene por finalidad obtener los valores de las características de una firma pre-procesada. Por ejemplo, P06 recibe D06 y obtiene D07.

La firma posee información representativa sobre ella, denominada características, que permiten diferenciar la pertenencia sobre la firma, es decir, diferenciar entre la información representativa de un firmante a otro. En la literatura especializada existen más de cuarenta y cuatro (44) características de las firma¹, y sus definiciones precisan el proceso de obtenerlo.

P06 es dado por el proceso implícito o explícito expresado en las definiciones de las características sobre una firma pre-procesada y obtiene los valores de sus características. No es recomendable usar todas las características dadas en la literatura, se sugiere considerar características que sean complementarias o independientes entre sí, o usar un método para la selección de características, por ejemplo, el método de análisis de componentes principales.

¹Rivera L., Mauricio D. and Rodriguez G., “Analysis of Efficiency of Automatic Verification in Offline Signatures”, Reporte Técnico 01/2016/CONCYTEC-176-PNICP-PIAP-2015. 2016.

P07: Existen firmas

Es un condicionante, que genera una de dos posibles respuestas “Si” o “No”.

P07 verifica si hay firmas en D01 no procesadas aún, si hubiera, la respuesta es “Si” de lo contrario es “No”.

P08: Registra características

Este procedimiento consiste en registrar D07 en un repositorio denominado “Características” (D08).

P09: Genera modelo de Verificación

Este procedimiento tiene por finalidad construir D02 (modelo de verificación) a partir de D08. Un Modelo de verificación es un conjunto de datos y parámetros de estructuras de machine learning y/o algoritmo de machine learning asociado a las firmas de un firmante. Los modelos de verificación deben ser tal que permitan realizar la verificación con altos niveles de precisión.

P09 consta de 5 procedimientos: “Aprendizaje algoritmo de machine learning” (P17), “Condicionante” (P18), “Ajustar parámetros” (P19), “Aprendizaje algoritmo de machine learning híbrido” (P20), y “Registra modelos” (P21).

P09 consta de los siguientes pasos (Fig. 5): Primero, aplica P17 sobre D08 usando los parámetros $(\lambda_1, \lambda_2, \dots)$ dados por P19 y determina “M” y “p”, y registra estos y los parámetros en D22. Segundo, aplica P18 y obtiene una respuesta “si” o “no”. Tercero, si el resultado del paso anterior es “No” se aplica P19, y se retorna al primer paso, de lo contrario continuamos al paso siguiente. Cuarto, repetir los pasos anteriores cuantos algoritmos de machine learning (AMLx) considere; se sugiere al menos un número impar de algoritmos, deseable 5 o 7. Quinto, aplicar P20 sobre los datos conformados por D08 y los D22 generados por los P17, con los parámetros $(\lambda_1, \lambda_2, \dots)$ dados por P19, y determine M, p, y registre estos y los parámetros en D23. Sexto, aplicar P18 y obtenga una respuesta “si” o “no”; Séptimo, si el resultado del paso anterior es “no” se aplica P19, y se retorna al quinto paso, de lo contrario continuar con el siguiente paso. Octavo, aplicar P21 considerando todos los D22 y el D23, y obtener D02.

P10: Verificador

Este procedimiento tiene por finalidad determinar si las características de una firma a verificar (D11) corresponde a un supuesto firmante. Es decir, determina si la firma a verificar es “Genuina o Falsa” (D04) a partir de un Modelo de Verificación (D02).

P08 está conformado por dos procedimientos: “Aplica Algoritmo de Machine Learning” (P22) y “Aplica Algoritmo de Machine Learning Híbrido” (P23).

P08 consta de los siguientes pasos: Primero, considerando D02, aplicar P22 sobre D11, y obtener D24. Segundo, repita el primer paso con todos los AMLx considerados en P09, y obtenga los correspondiente D24. Tercero, considerando D02 aplicar P23 sobre los datos conformados por D11 y los D24 obtenidos en el paso anterior, y obtener el resultado “Genuina o Falsa” (D04).

P11: Eliminar ruido de fondo.

Este procedimiento tiene por finalidad obtener la imagen “Firma sin ruido de fondo” (D13) a partir de una firma digitalizada (D12). La Fig. 7 ilustra el resultado de este procedimiento.

P11 consta de dos pasos: Primero, colocar la imagen en escala de grises (gs), mediante una combinación lineal de los colores RGB (red-green-blue) asociado a cada pixel; se sugiere usar la combinación $gs = 0.2126 * red + 0.7152 * green + 0.0722 * blue$. Segundo, colocar cada pixel con gs inferior a un umbral dado un valor 0 (anula el fondo), y 1 en caso contrario, la imagen resultante es la imagen de la firma en blanco y negro, es decir D13. Se sugiere usar el umbral de 210.

P12: Rectángulo-Mínimo.

Este procedimiento tiene por finalidad delimitar una firma sin ruido de fondo al menor rectángulo que la contiene y que tiene ejes paralelos a ésta. Por ejemplo, P12 recibe D13 y genera D14.

El rectángulo mínimo es el menor rectángulo de lados paralelos a la firma que contiene a la firma.

P12 consta de dos pasos: Primero se determina las filas (superior e inferior) y columnas extremas (izquierda y derecha) del rectángulo mínimo. La fila superior (inferior) extrema es la primera fila de arriba hacia abajo (abajo hacia arriba) que contiene un pixel con valor 1 (pixel negro), y la columna izquierda (derecha) extrema es la primera columna de izquierda a derecha (derecha a izquierda) que contiene un pixel de valor 1. Segundo, se recorta la imagen de la firma sin fondo con las filas y columnas extremas dado en el paso anterior, generando el rectángulo mínimo.

La Fig. 8 ilustra el resultado de este procedimiento.

P13: Eliminar-Ruido-Marco.

Este procedimiento tiene por finalidad eliminar el ruido marco de una firma en blanco y negro, en rectángulo mínimo y sin ruido (de) fondo.

El ruido marco es dado por los pixeles con valor 1 que se encuentran al borde del rectángulo mínimo pero que no son parte de la firma. En general las líneas de la firma tienen un grosor máximo de pixeles, por lo que toda línea paralela a los lados de D14 con grosor mayor a

este máximo se puede considerar un ruido marco. En general las líneas de una firma tienen grosor menor a 10 píxeles, por lo que se sugiere usar como máximo el valor de 10.

P13 consta de tres pasos: Primero, se identifica si existe un ruido marco en cada línea (fila o columna) de D14 que se encuentra a distancia de 0 píxeles de su borde. Segundo, si existe dicho ruido marco, este se elimina; es decir, sus píxeles son fijados con valor 0. Tercero, se repiten los dos pasos anteriores incrementando la distancia en 1 píxel, hasta que no haya ruido marco. La imagen resultante es “Firma sin ruido marco” (D15).

La Fig. 9 ilustra el resultado de aplicar este procedimiento.

P14: Eliminar-Ruido-Isla.

Este procedimiento tiene por finalidad eliminar el ruido isla de D15, de forma a generar una imagen de la firma sin ruido isla, y que denominamos “Firma sin ruido isla” (D16).

El ruido isla es dado por una aglomeración de píxeles con valor 1 y que se encuentran aislados; es decir, estas tienen borde píxeles con valor 0 y/o el borde de D15.

P14 consta de cinco pasos: Primero, coloca como punto de partida las coordenadas del píxel superior e izquierdo de D15. Segundo, identifica un ruido isla (si existe) analizando cada píxel de D15 considerando el punto de partida del paso anterior; es decir, se analizan los píxeles que se encuentran en la fila del punto de partida hacia la parte inferior de D15, y en la dirección de izquierda a derecha. Tercero, si existe un ruido isla, todos sus píxeles son fijados con valor 0. Cuarto, se fija como punto partida el punto dado en el paso ?? (si hubiera); Quinto, se repiten los pasos dos, tres y cuatro hasta recorrer todos los píxeles de D15, la imagen resultante es D16.

La Fig. 10 ilustra el resultado de aplicar este procedimiento.

P15: Afinar.

Este procedimiento también denominado adelgazamiento o esqueletización, tiene por finalidad afinar una firma sin ruido (de fondo, marco y de isla). Decimos que una firma está afinada cuando el ancho de sus líneas es de 1 píxel.

P15 es un procedimiento bastante conocido en la literatura y es dado por los algoritmos de afinamiento. Sugerimos usar el algoritmo de Zhang and Fu².

Las Fig. 12 y 14 ilustran el resultado de la aplicación de este procedimiento sobre una firma en blanco y negro, de una firma normalizada.

²Zhang S. and Fu K.S, “A Thinning Algorithm for discrete Binary Images”, in “Algorithms for Image Processing and Computer Vision”, JR Parker- John Wiley & Sons, Inc. 1997.

P16: Normalizar.

Este procedimiento también es conocido como escalonamiento, consiste en ampliar o reducir (transformar) una imagen de forma que sus dimensiones de alto y ancho sean normalizadas (igual para todas), por ejemplo, transforma D17 a D19. Este procedimiento es muy importante porque permite reducir las distorsiones de algunas características de la firma ocasionada por la variación de los tamaños de la firma de un mismo firmante.

Las dimensiones normalizadas pueden ser establecidas antemano, por ejemplo, siguiendo las proporciones de la regla dorada. Sugerimos usar el promedio de las dimensiones dadas de las imágenes de las firmas de un firmante.

P16 es un procedimiento bastante conocido en la literatura y está dado por métodos bien conocidos, sugerimos usar la interpolación bilineal y que está disponible en java³.

La Fig. 13 ilustra el resultado de aplicar este procedimiento.

P17: Aprendizaje Algoritmo de Machine Learning.

Es un procedimiento que corresponde a un proceso de aprendizaje de machine learning. P17 consta de dos pasos: Primero, aplicar un algoritmo de machine learning (AMLx) sobre D08 usando los parámetros $(\lambda_1, \lambda_2, \dots)$ proporcionado por P19 y obtenga "M" un modelo de machine learning, "p" una precisión de clasificación. Segundo, registre "M", "p", y $\lambda_1, \lambda_2, \dots$ en D22.

Se sugiere usar para AMLx: Árbol de Decisión, Random Forest, Support Vector Machine, Naives Bayes, K vecinos más próximos, Perceptron Multicapa.

P18: Condicionante

Es un condicionante que genera una de dos posibles respuestas "Si" o "No". P18 recibe de P17 "p" y genera una respuesta "Si" si se verifica $p > u$, caso contrario la respuesta es "No". Se sugiere usar $u=0.7$ (umbral), un valor alcanzable por los algoritmos de machine learning para la verificación de la firma. El valor de "p" depende de la cantidad de firmas para un firmante y de la representatividad de las características, cuantas más firmas existan y mayor sea la representatividad de las características, mejor será el valor de "p".

P19: Ajustar parámetros.

Este procedimiento tiene por finalidad hacer los ajustes de los parámetros (si hubiera) de un algoritmo de machine learning de forma a mejorar la precisión de sus resultados. El ajuste o calibración de parámetros es usual en el área de machine learning. Se sugiere usar el criterio greedy para la calibración. Cuando P19 se aplica por primera vez, se sugiere usar los valores

³<http://tech-algorithm.com/articles/bilinear-image-scaling/>

de los parámetros dados por defecto en las librerías de machine learning; en otra situación se usará otro criterio para calibrar los parámetros.

P20: Aprendizaje Algoritmo de Machine Learning Híbrido.

Es un procedimiento que corresponde a un proceso de aprendizaje de machine learning. P20 consta de dos pasos: Primero, aplicar un algoritmo de machine learning (AML) sobre los datos conformados por los D22 y D08, usando los parámetros ($\lambda_1, \lambda_2, \dots$) proporcionado por P19 y obtenga "M" un modelo de machine learning, "p" una precisión de clasificación, y "r" el resultado de clasificación. Segundo, registre "M", "p", y $\lambda_1, \lambda_2, \dots$ en D23.

Se sugiere usar para AML: Vector Machine, Naives Bayes, K vecinos más próximos, o Perceptron Multicapa que son los que proporcionan mejores resultados conforme a las pruebas, o un criterio de votación ponderada.

AML es denominado híbrido porque aprende de los resultados generados por algoritmos de machine learning (AMLx)

P21: Registra modelos

Este procedimiento consiste en registrar los D22 y D23 en un repositorio denominado "Modelo de verificación" (D02).

P22: Aplica Algoritmo de Machine Learning.

Es un procedimiento que corresponde a la fase de aplicación (también denominado uso o producción) de machine learning.

P22 es dado por aplicar un AMLx considerado en P17 con su correspondiente modelo de machine learning y parámetros dado en D02 sobre D11, y obtener el resultado "r" (D24).

P17 y P22, corresponden al mismo algoritmo de machine learning (AMLx) pero en fases diferentes, la primera corresponde a la fase de aprendizaje (construcción del modelo de machine learning) y la segunda a la fase de aplicación (producción).

P23: Aplica Algoritmo de Machine Learning Híbrido.

Es un procedimiento que corresponde a la fase de aplicación (también denominado uso o producción) de machine learning.

P23 es dado por aplicar un AML considerado en P20 con su correspondiente modelo de machine learning y parámetros dado en D02 sobre los datos conformados por D11 y los D24, y obtenga el resultado que único que puede ser Genuino o Falso (D04).

P20 y P23, corresponden al mismo algoritmo de machine learning híbrido (AML) pero en fases diferentes: la primera, corresponde a la fase de aprendizaje (construcción del modelo de machine learning) y la segunda, a la fase de aplicación (producción).

D01: Firmas de un firmante.

Es un conjunto de al menos tres (3) firmas genuinas y recientes de un firmante. La precisión de los métodos de verificación depende de las firmas genuinas, así cuanto más firmas genuinas y recientes se tengan mejor será la precisión de los métodos de verificación. Las firmas genuinas pueden estar en soportadas en forma impresa o digitalizada.

D02: Modelo de Verificación.

Es un repositorio que almacena el modelo de verificación de un firmante, y consta de: modelos de machine learning "M", precisión de clasificación "p", y parámetros $\lambda_1, \lambda_2, \dots$ para cada algoritmo de machine learning (AMLx y AML) considerado en P09.

D03: Firma a Verificar.

Esta dada por la firma a verificar, sea ésta en un documento impreso o en forma digital. Para más de una firma a verificar, se asocia a cada una de ellas un identificador del supuesto firmante, que será usado para extraer el modelo de verificación del firmante en el proceso de verificación.

D04: Genuina o Falsa.

Es el reporte en forma digital o impresa, que sólo adopta una de las alternativas: "Genuina" o "Falsa"; es decir, que la firma a verificar corresponde o no al firmante en cuestión.

D05: Firma digitalizada.

Es un repositorio en donde se almacena una firma digitalizada.

D06: Firma pre-procesada.

Es un repositorio en donde se almacena una firma pre-procesada, es decir cuatro (4) imágenes: "Firma sin ruido y en rectángulo mínimo" (D17), "Firma afinada y en rectángulo mínimo" (D18), "Firma sin ruido y normalizada" (D19), y "Firma afinada y normalizada" (D20).

D07: Características de firma.

Es un repositorio en donde se almacena los valores de las características de una firma.

D08: Características.

Es un repositorio en donde se almacena los valores de las características de las firmas de un firmante. Es decir, D08 almacena todos los D07 asociado a las firmas genuinas de un firmante.

D09: Firma digitalizada V.

Es un repositorio que almacena en forma digital la firma a verificar. Equivale a D03 pero en forma digital.

D10: Firmas pre-procesada V.

Es un repositorio en donde se almacena la firma pre-procesada a verificar, éste está compuesto de cuatro (4) imágenes: D17, D18, D19, y D20.

D11: Características V.

Es un repositorio en donde se almacena los valores de las características de la firma a verificar. Las características consideradas deben ser las mismas dadas en D07.

D12: Firma digitalizada.

Es un repositorio en donde se almacena una firma en forma digital.

D13: Firma sin ruido de fondo.

Es un repositorio que almacena en forma digital la imagen de la firma sin ruido de fondo; es decir, en blanco y negro (fondo blanco). El ruido de fondo es todo pixel que no es parte de la firma por que presenta una escala de gris diferente a ella.

D14: Firma en rectángulo mínimo.

Es un repositorio que almacena en forma digital la imagen de la firma delimitada al rectángulo mínimo. El rectángulo mínimo es el menor rectángulo de lados paralelos al eje de la firma y que la contiene.

D15: Firma sin ruido marco.

Es un repositorio que almacena en forma digital la imagen de la firma sin ruido marco. El ruido marco es el ruido de forma lineal que aparece en los lados del rectángulo mínimo, y que en general es ocasionado por el uso de recuadros para firmar. Su eliminación es importante pues de lo contrario, puede confundirse como si fuera parte de la firma.

D16: Firma sin ruido isla.

Es un repositorio que almacena en forma digital la firma sin ruido de isla. El ruido de isla, son aquellas imágenes pequeñas y aisladas (limitadas por píxeles en blanco), que se encuentran en la imagen de la firma pero que no son parte de ella; en general se forman por el polvo que hay en el papel a escanear o manchas en el panel del escáner.

D17: Firma sin ruido y en rectángulo mínimo.

Es un repositorio que almacena en forma digital y en blanco y negro (blanco para el fondo y negro para la firma) la imagen de la firma, sin ruido (ruido isla, ruido marco) y delimitado al rectángulo mínimo que contiene a la firma.

D18: Firma afinada en rectángulo mínimo.

Es un repositorio que almacena en forma digital, en blanco y negro (blanco para el fondo y negro para la firma) la imagen de la firma, sin ruido (ruido isla, ruido marco), en rectángulo mínimo y además afinada. Una firma está afinada (también denominada adelgazada o esqueletizada), cuando el ancho de sus líneas es de 1 pixel, y es importante porque permite determinar algunas características representativas como la masa.

D19: Firma sin ruido y normalizada.

Es un repositorio que almacena en forma digital, en blanco y negro (blanco para el fondo y negro para la firma) la imagen de la firma, sin ruido (ruido isla, ruido marco), en rectángulo mínimo, afinada y en un tamaño normalizado. El tamaño normalizado se refiere a las dimensiones (alto y ancho) estándar que deben tener todas las imágenes de las firmas para un firmante: dichas dimensiones pueden ser establecidas antemano, por ejemplo siguiendo las proporciones de la regla dorada. Sugerimos usar el promedio de las dimensiones de las imágenes de las firmas de un firmante.

D20: Firma afinada y normalizada.

Es un repositorio que almacena en forma digital, en blanco y negro (blanco para el fondo y negro para la firma) la imagen de la firma, sin ruido (ruido isla, ruido marco), normalizada y afinada.

D21: Firmas pre-procesada.

Es un repositorio en donde se almacena una firma pre-procesada, es decir cuatro (4) imágenes asociadas a una firma digitalizada (D12): "Firma sin ruido y en rectángulo mínimo"

(D17), “Firma afinada y en rectángulo mínimo” (D18), “Firma sin ruido y normalizada” (D19), y “Firma afinada y normalizada” (D20).

D22: Modelo AMLx

Es un repositorio que almacena un modelo de aprendizaje de machine learning “M”, una precisión “p” y parámetros $\lambda_1, \lambda_2, \dots$ (si hubiera).

D23: Modelo AML

Es un repositorio que almacena un modelo de aprendizaje de machine learning híbrido “M”, una precisión “p” y parámetros $\lambda_1, \lambda_2, \dots$ (si hubiera).

D24: Resultado Parcial

Es un repositorio que almacena el resultado “r” (solo puede ser genuino o falso) generado por el procedimiento P22.

D24: Firma con ruido de fondo

D25: Firma

D27: Ruido de Fondo

D28: Firma sin ruido de fondo

D29: Firma en blanco y negro

D30: Firma en blanco y negro y en rectángulo mínimo

D31: Firma en blanco y negro y con ruido de marco

D32: Ruido de marco

D33: Firma en blanco y negro y sin ruido de marco

D34: Firma “AR” en píxeles “01” con ruido isla

D35: Ruido isla

D36: Firma “AR” en píxeles “01” sin ruido isla

D37: Firma sin pre-procesar

D38: Firma en blanco y negro, sin ruido, y en rectángulo mínimo.

D39: Firma en blanco y negro, sin ruido, en rectángulo mínimo y afinada.

D40: Firma en blanco y negro, sin ruido, en rectángulo mínimo y normalizado.

D41: Firma en blanco y negro, sin ruido, en rectángulo mínimo, normalizado y afinado.

D42: Firma pre-procesada (cuatro firmas: D38, D39, D40, D41).

Descripción de las figuras:

Fig. 1 Es un diagrama de bloques que muestra el “Procedimiento general de la verificación automática de la firma” (P01), que está conformado por dos grandes procedimientos: la “Construcción del modelo de verificación de un firmante” (P02), y la “Verificación automática” (P03). P01, recibe firmas de un firmante (D01), y la firma a verificar (D03), y genera un resultado “Genuino o falso” (D04).

Fig. 2 Es un diagrama de bloques que muestra el “Procedimiento de construcción del modelo de verificación” (P02), que consta de seis (6) procedimientos (P04, P05, P06, P07, P08, P09), cuya aplicación genera la Firma digitalizada (D05), las Firmas pre-procesadas (D06), las Características de firma (D07), el repositorio de Características (D08) y finalmente el Modelo de verificación (D02).

Fig. 3 Es un diagrama de bloques que muestra el “Procedimiento de verificación automática de firmas” (P03), que está conformado por cuatro (4) procedimientos (P04, P05, P06, P10), tres repositorios (D09, D10, D11). P03 recibe la firma a verificar (D03), y el Modelo de verificación (D02) obtenido por P02, y genera el reporte Genuino o Falso (D04).

Fig. 4 Muestra en diagrama de bloques el “Procedimiento de pre-procesamiento” (P05). P05 consta de seis (6) procedimientos (P11, P12, P13, P14, P15, P16,); recibe una firma digitalizada (D12) y genera la firma pre-procesada (D21) dada por cuatro imágenes (D17, D18, D19, D20).

Fig. 5 Muestra en diagrama de bloques el procedimiento “Genera modelo de verificación” (P09). P09 consta de cinco (5) procedimientos (P17, P18, P19, P20, P21); recibe D08, conjunto de características de las firmas de un firmante y genera el modelo de verificación (D02).

Fig. 6 Es un diagrama de bloques que muestra el procedimiento “Verificador” (P10), que consta de dos (2) procedimientos (P22, P23), uno de ellos (P22) se replica cuantas veces se consideren algoritmos de machine learning. P10 recibe D02 (Modelo de verificación) y D11 (Características V) y genera el reporte D04 (Genuino o Falso).

Fig. 7 Muestra una imagen de una firma con ruido de fondo (A), y la misma imagen sin ruido fondo (B) obtenido por la aplicación del procedimiento P11.

Fig. 8 Muestra la imagen de una firma en blanco y negro en un rectángulo (A), y la misma imagen en un rectángulo mínimo (B), obtenido por la aplicación del procedimiento P12.

Fig. 9 Muestra una imagen de una firma en blanco y negro en un rectángulo y con ruido marco (A), y la misma imagen sin ruido marco (B), obtenido por la aplicación del procedimiento P13

Fig. 10 Muestra la imagen de una firma “AR” con pixeles “0” (blanco) y “1” (negro) con ruido de isla (A), y la misma imagen sin ruido de isla (B), obtenido por la aplicación del procedimiento P14.

Fig. 11 Muestra la imagen de una firma con ruido (de fondo, de marco, isla) y en rectángulo mínimo (A), y la misma imagen sin ruido (B).

Fig. 12 Muestra la imagen de una firma en blanco y negro, sin ruido, en rectángulo mínimo, y sin afinar (A) y la misma imagen con la firma afinada (B), obtenida al aplicar P15 sobre A.

Fig. 13 Muestra la imagen de una firma en blanco y negro, sin ruido, en rectángulo mínimo, y sin normalizar (A) y la misma imagen normalizada (B), obtenida al aplicar P16 sobre A.

Fig. 14 Muestra la imagen de una firma en blanco y negro, sin ruido, en rectángulo mínimo, y normalizada y sin afinar (A) y la misma imagen afinada (B), obtenida al aplicar P15 sobre A.

Fig. 15 Muestra la imagen de una firma sin pre-procesar (A), y la misma imagen pre-procesada (B, C, D, E), esto se obtiene al aplicar P05 sobre A.

Modo de realización de la invención.

La presente invención presenta varios modos de realización, algunas de ellas son:

- ***Verificación automática de firmas de adherentes para la inscripción de organizaciones políticas.***

Actualmente la verificación de firmas de adherentes para la inscripción de partidos políticos es realizada por el Registro Nacional de Identificación y Estado Civiles (RENIEC), de forma manual por peritos grafólogos. Es un proceso que requiere de grandes recursos humanos logísticos y tecnológicos dado los grandes volúmenes de firmas a procesar las mismas que pueden superar más de un millón de firmas por cada evento. Además, es un proceso de alto costo, que requiere tiempo y presenta un alto porcentaje de error humano debido a la exigencia de cumplimientos en tiempos, carga natural de trabajo, concentración y la naturaleza repetitiva y compleja de la verificación.

La implementación de la presente invención consiste en la realización de un sistema de verificación automática de la firma a partir de los planillones escaneados, presentados por los partidos y/o agrupaciones políticas, la información contenida en el CD adjunto a los planillones y el banco de firmas genuinas registradas en la RENIEC. El sistema realizará la verificación automática de las firmas de los planillones, reduciendo significativamente el empleo de recursos humanos, logísticos y costos. Además, la invención permitirá reducir las tasas de error en la verificación manual (errores por rechazo de firmas genuinas o aceptación de firmas falsas).

De manera similar la presente invención es aplicable a la verificación de firmas de adherentes en las *consultas populares para revocatorias del mandato de autoridades*, y también en la *demarcación territorial*.

La implementación de la presente invención permitirá que el proceso de verificación sea realizado de forma automática, normalizada, rápida y con altos índices de confiabilidad, reduciendo las amenazas y actos de usurpación de identidad, fraude como permitirá la disuasión, investigación y sanciones de las autoridades respectivas.

Otras realizaciones de la presente invención son:

- ***Verificación automática de firmas en cheques bancarios***

Actualmente la verificación de firma en cheques bancarios se realiza de forma manual. El operador recibe el cheque y mediante un identificador del emisor del cheque extrae de una base de datos la imagen de la firma registrada (considerada genuina) del emisor, y realiza el cotejo y verificación en forma visual. En general el operador no es un perito en grafotecnia, por lo que puede incurrirse en altas tasas de error.

- ***Verificación automática de firmas manuscritas en registros públicos.***

Actualmente la verificación de la firma manuscrita en determinadas operaciones comerciales y procedimientos administrativos como de compra-venta de inmuebles, constitución de empresas, entre otros, realizados en la Superintendencia Nacional de los Registros Públicos (SUNARP) se practica de forma manual por parte de operadores o peritos grafólogos. Estos procesos demandan ingentes recursos humanos, logísticos, tecnológicos, económicos, dada la cantidad de actos y operaciones a practicarse en la SUNARP, que requieren de tiempo y concentración, pudiéndose incurrir en errores u omisiones, debido a la naturaleza compleja de la verificación de las firmas.

- ***Verificación automática de firmas manuscritas en Notarías.***

Actualmente la verificación de firma manuscritas para determinadas operaciones comerciales, procedimientos administrativos y procesos judiciales, entre otros, realizados en la Notarías se realiza en forma manual por parte de un operador, un perito en grafotecnia o el propio Notario; proceso que requiere recursos humanos, logísticos, económicos y tecnológicos dada la cantidad de firmas que se verifican a nivel nacional por las Notarías. Al igual que en las realizaciones anteriormente descritas, son procesos

que insumen tiempo y pueden ser generadores de errores u omisiones, debido a la naturaleza compleja de la verificación de las firmas.

- **Verificación automática de firmas en contratos laborales y servicios de salud**

Actualmente la verificación de firma manuscritas en contrato laborales se realiza en forma manual, por un operador, en muchos casos no se realiza el cotejo ni la verificación de las firmas de los contratos, que junto con la usurpación de identidad y el fraude permiten la contratación de trabajadores fantasmas o la prestación de servicios de salud a personas que usurpan una identidad.

En general la presente invención se puede aplicar a todo proceso de verificación de firmas manuscritas, en donde la firma a verificar se encuentre en forma impresa o digital, y que existan firmas genuinas referentes del firmante en cuestión.

Reivindicaciones

1. **El procedimiento de verificación automática** de firmas manuscritas en documentos en soporte papel o en micro formas propuesto, se organiza en veintitrés (23) procedimientos que detallan como realizar la verificación automática de firmas manuscritas y establece los lineamientos lógicos para implementar un sistema computacional con este fin. Existen en la literatura especializada propuestas sobre la automatización de la verificación de la firma manuscrita, sin embargo, pocos autores describen el sistema computacional propuesto, y escasamente se detallan los procedimientos para la verificación automática.

Así mismo existen patentes para la verificación automática de firmas manuscritas, pero ellas se refieren a dispositivos y muy escasamente se orientan a procedimientos. Además, los procedimientos existentes cubren algunas actividades de la verificación automáticas, y siguen combinaciones de actividades diferentes a la propuesta, por lo que la invención propuesta es única, innovativa.

2. El procedimiento general de verificación de la firma (P01), que consta de los procedimientos de construcción de un modelo de verificación de un firmante (P02) y de verificación automática (P03), incluyendo entre ambos a otros veinte (20) procedimientos, es único, por lo expuesto en el párrafo anterior (1), **reivindicándose P01**.

3. P02 consta de los procedimientos: Digitalización (P04), Pre-procesamiento (P05), Extracción de características (P06), Existen firmas (P07), Registra Características (P08) y Genera modelo de verificación (P09).

Varios de los procedimientos que consta P02 son conocidos en la literatura, pero ninguno de ellos integra los integra ni combina de la forma propuesta, por lo que **se reivindica P02**. P04 es usual y conocido mundialmente y P06 es muy conocido en la literatura relacionada sobre la verificación automática de firmas manuscritas por lo que ambos no

son reivindicables. P05 se revisa en los párrafos siguientes. P07 y P08 se refieren a actividades elementales que se realizan en diversos métodos y procedimientos en varios problemas, por lo que tampoco no son reivindicables.

4. P03 consta de los procedimientos: Digitalización de la imagen de la firma (P04), Pre-procesamiento (P05), Extracción de características (P06) y Verificador (P10). Varios de los procedimientos de los cuales consta P03 son conocidos en la literatura pero ninguno los integra ni los combina de la forma propuesta, por lo que **se reivindica P03**.
5. P05 consta de los procedimientos: Eliminar-ruido-fondo (P11), Rectángulo-mínimo (P12), Eliminar ruido marco (P13), Existen-ruido-isla (P14), Afina (P15) y Normalizar (P16). Varios de los procedimientos que consta P05 son conocidos en la literatura pero ninguno los integra ni los combina de la forma propuesta, por lo que **se reivindica P05**. Los procedimientos P11, P12, P15 y P16 son conocidos en la literatura y usados en los sistemas de verificación por lo que no son reivindicables. Sin embargo, los procedimientos **P13 y P14** son nuevos procedimientos **por lo que se reivindican ambos**.
6. P09 consta de los procedimientos: Aprendizaje algoritmos de machine learning (P17), Condicionante (P18), Ajustar parámetros (P19), Aprendizaje algoritmo de machine learning híbrido (P20) y Registra modelos (P21). Los procedimientos para generar modelos de verificación son presentados en forma implícita en la literatura especializada, sin embargo, ninguna integra los procedimientos que tiene P09 y no presentan la secuencia lógica propuesta, por lo que **se reivindica P09**. P17 y P19 se refieren a procedimientos conocidos en la literatura por lo que no se reivindican. P18 y P21 se refieren a actividades de condicionante y registro respectivamente, que son elementales en diversos métodos y procedimientos por lo que no se reivindican. **P20 describe una invención para la verificación automática por lo que se reivindica**.
7. P10 consta de los procedimientos: Aplica Algoritmo de Machine Learning (P22) y Aplica Algoritmo de Machine Learning Híbrido (P23). Procedimientos para verificar en forma automática si una firma es genuina o falsa a partir de un modelo verificador no son presentado en la literatura especializada, por lo que **se reivindica P10**. P22 es un procedimiento que corresponde a la aplicación de un algoritmo de machine learning en su fase de uso, lo que es conocido en la literatura por lo que no se reivindica. Sin embargo, **P23** es un procedimiento usado en otros tipos de problemas, pero es único para la verificación automática de la firma, **por lo que se reivindica**.
8. En síntesis, se reivindican los procedimientos: **P01, P02, P03, P05, P09, P10, P13, P14, P20, y P23** por ser únicos e innovativos. No se reivindican los procedimientos restantes.

Resumen

La presente invención establece un procedimiento para la verificación automática de firmas manuscritas off-line en soporte impreso o digital a través de técnicas de machine learning. Es decir, identifica si una firma es genuina o falsa. El procedimiento propuesto a su vez consta de veintitrés (23) procedimientos que establecen los lineamientos lógicos para el desarrollo de un *software* inteligente para la verificación automática, y que están agrupados en procedimientos para la construcción de un modelo de verificación personalizado para cada firmante, y procedimientos para la verificación automática.

La propuesta de invención presenta diversas realizaciones dado que la firma es considerada como un atributo personalísimo del individuo y desde hace siglos las sociedades la usan como un medio para identificar al firmante, expresar consentimiento a través del cual se validan actos soportados en documento, siendo obligatoria en distintos documentos (cheques bancarios, contratos, tarjetas de crédito, procesos electorales, notariales, entre otros). Además, los documentos validados por la firma se pueden conservar, tratar y transmitir, y su contenido se considera válido en el tiempo y lugar, sea en soporte físico o en digital, por ello la firma es un instrumento muy atractivo para la práctica de usurpación o fraude desde hace mucho tiempo, de ahí también su importancia. La implementación del procedimiento propuesto permitirá reducir recursos humanos y materiales, costos, tiempos, así como disuadir y reducir la usurpación y el fraude en todos los procesos que involucran la verificación de firmas manuscritas.

Del total de los veintitrés (23) procedimientos descritos en el presente documento, se reivindican ocho (8) procedimientos por ser únicos e innovativos.

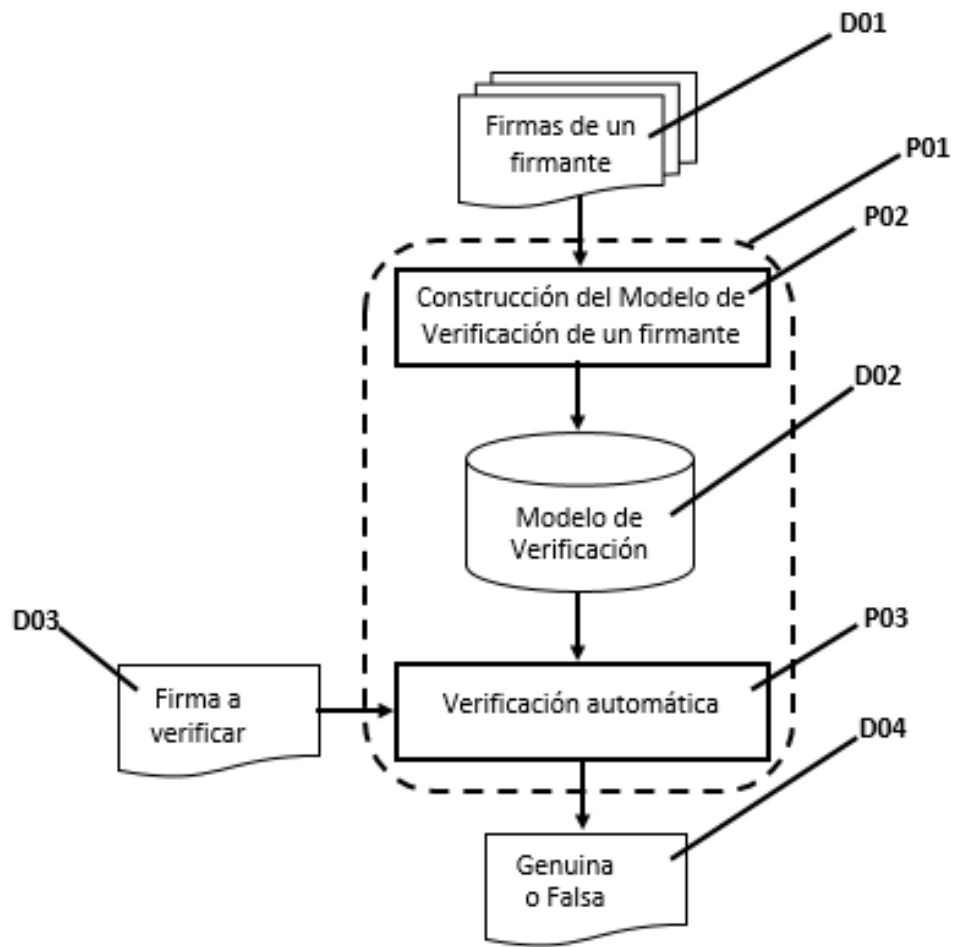


Fig. 1. P01: Procedimiento General de Verificación Automática de la Firma

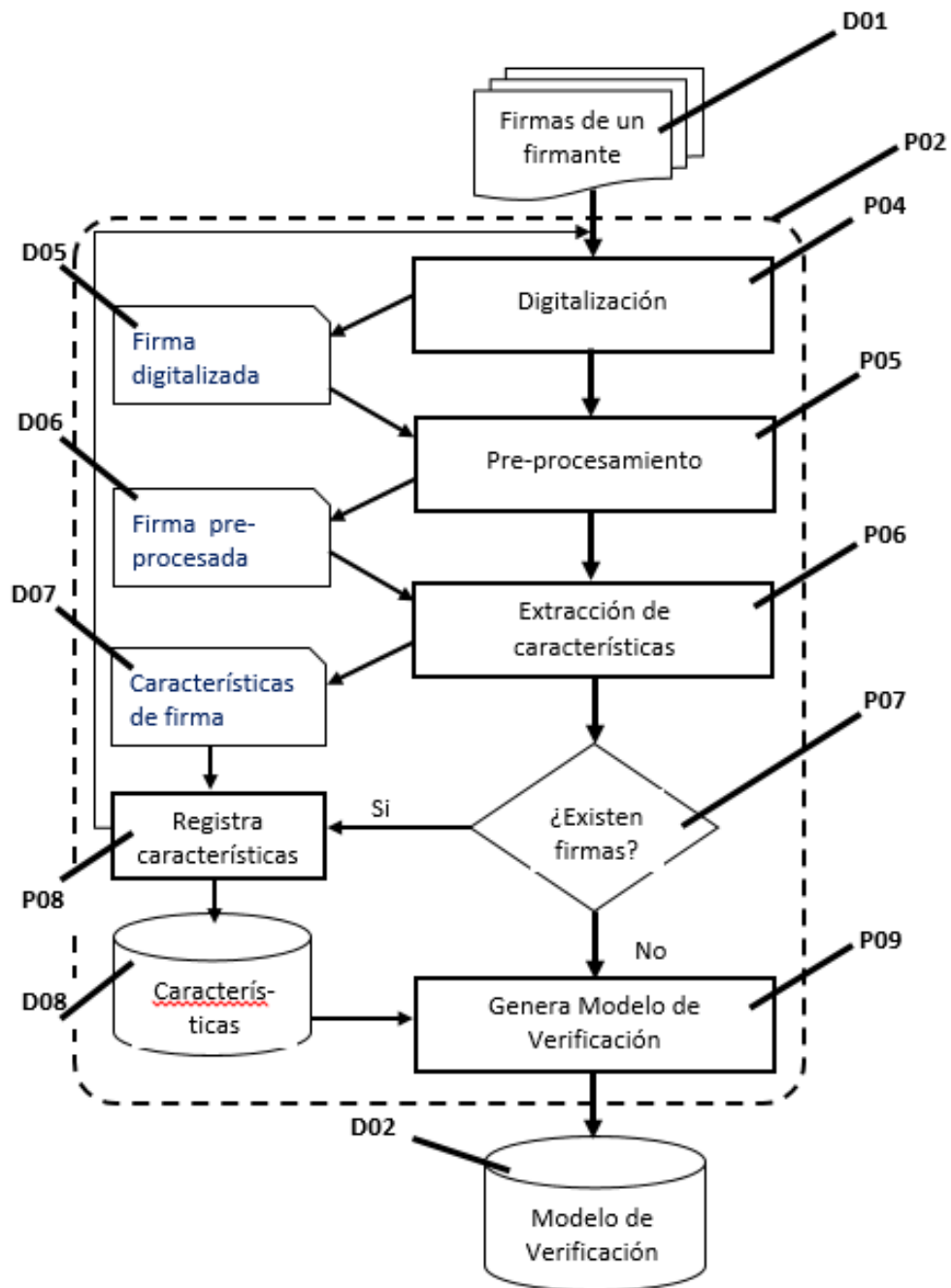


Fig. 2. P02: Procedimiento Construcción del Modelo de Verificación

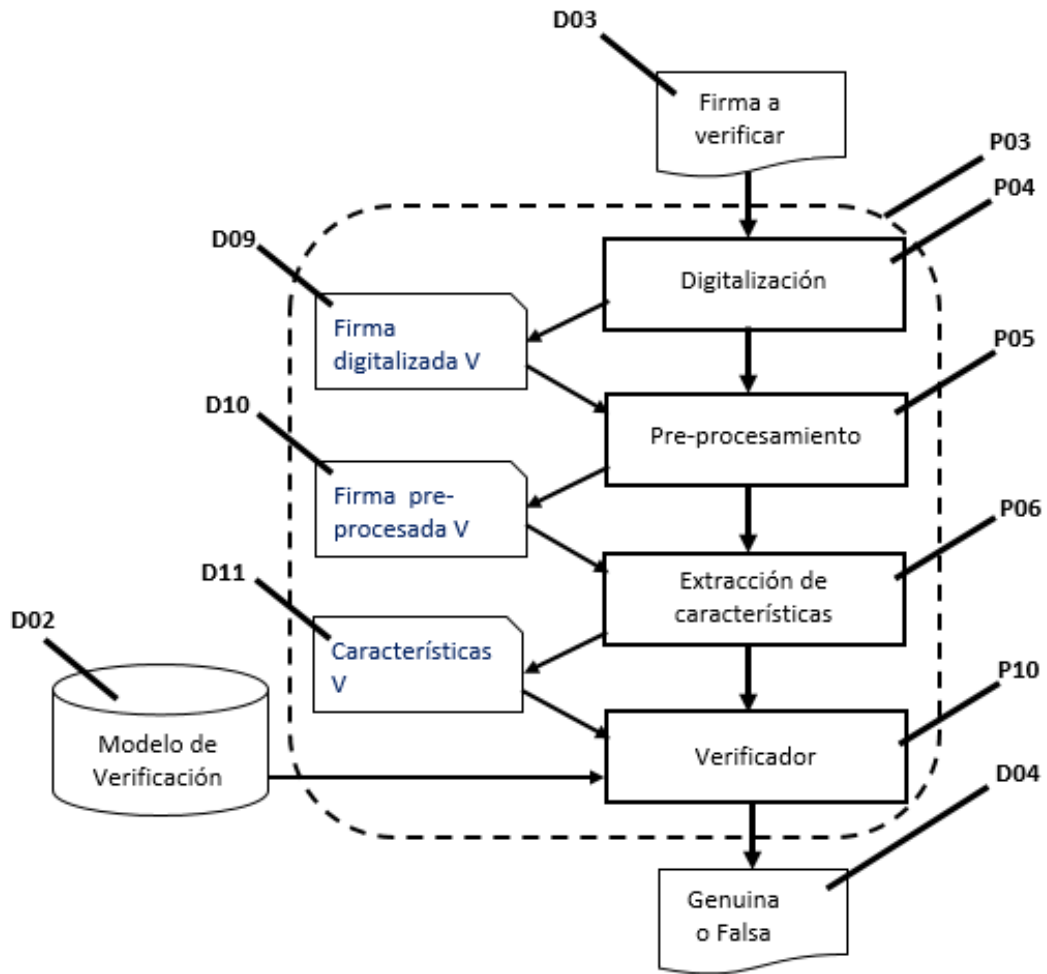


Fig. 3. P03: Procedimiento Verificación Automática de Firmas.

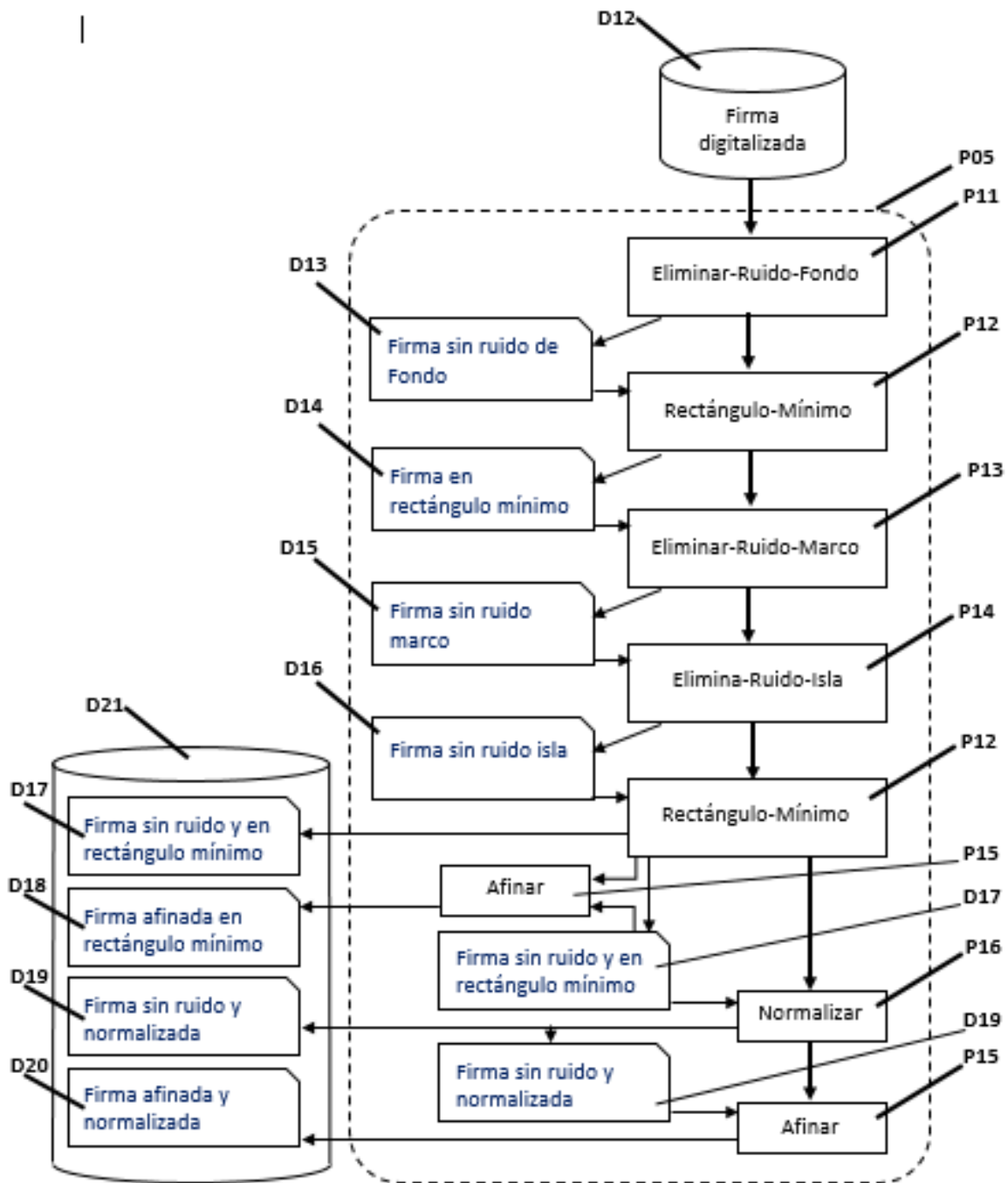


Fig. 4. P05: Procedimiento de Pre-procesamiento.

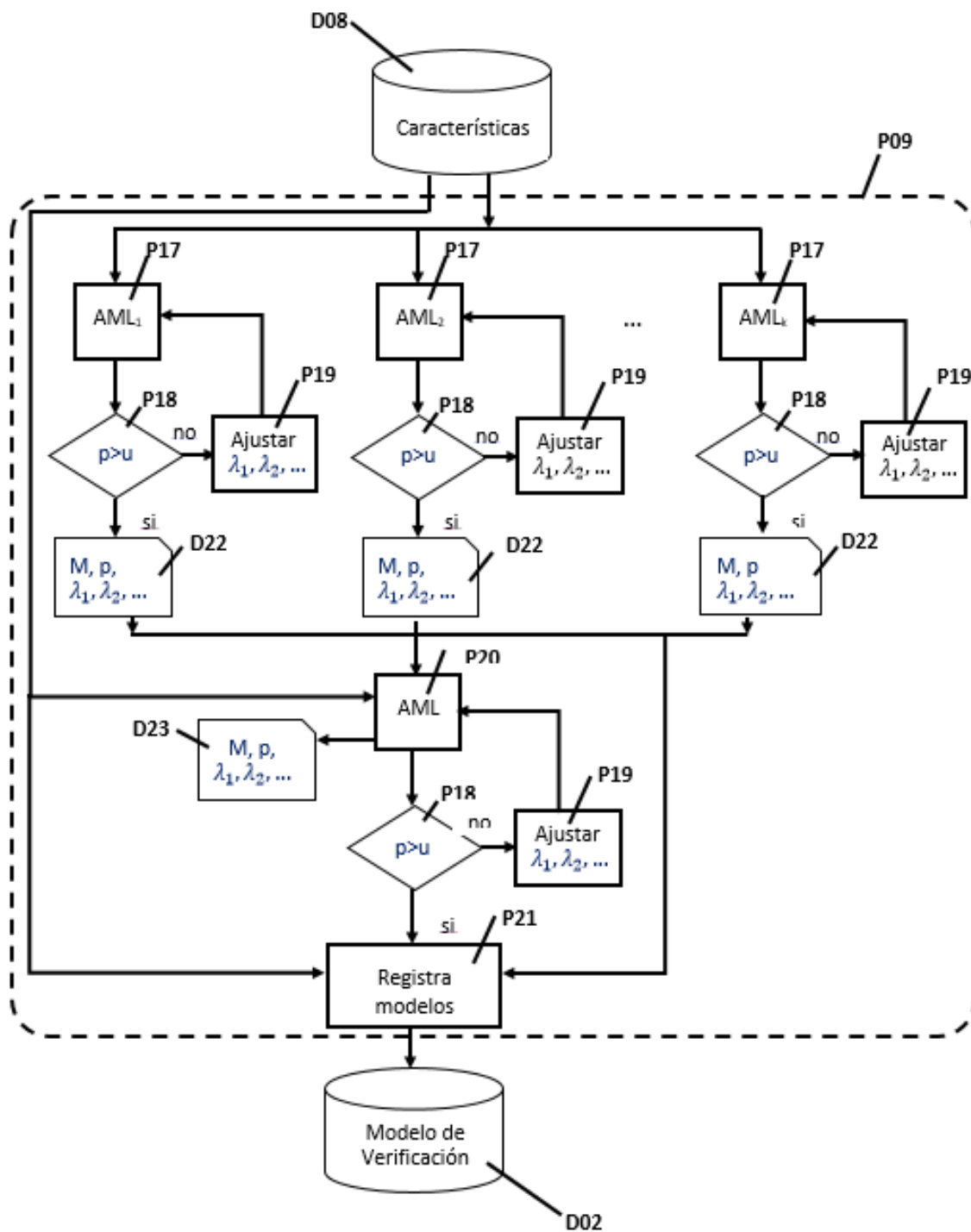


Fig. 5. P07: Procedimiento de Genera Modelo de Verificación.

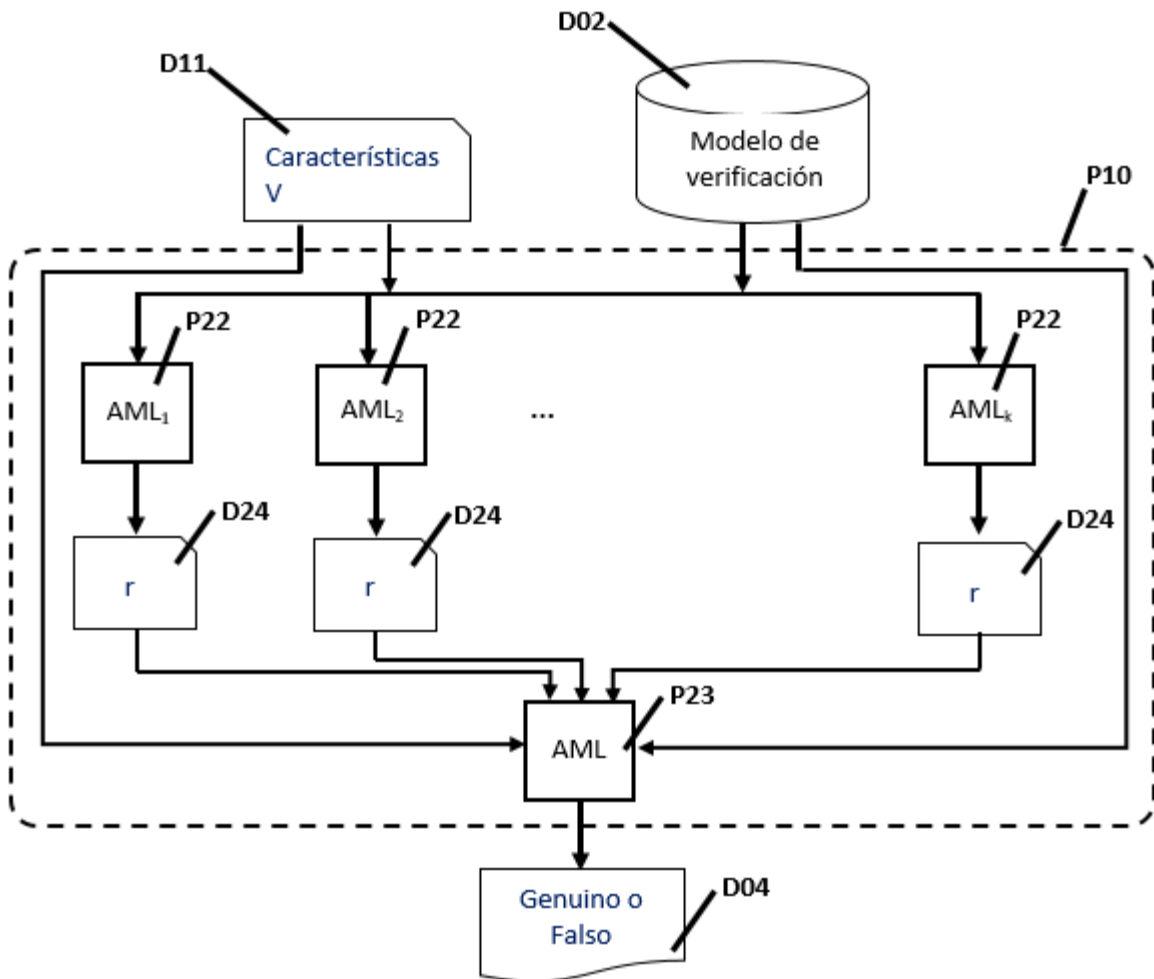


Fig. 6. P08: Procedimiento Verificador.

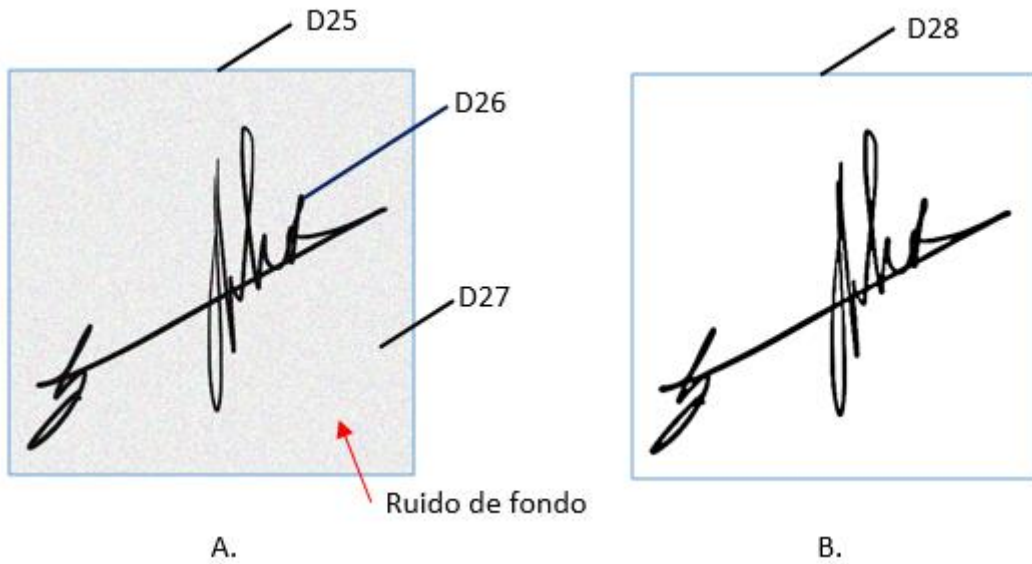


Fig. 7 Ruido Fondo: A) Firma con ruido fondo B) Firma sin ruido fondo.

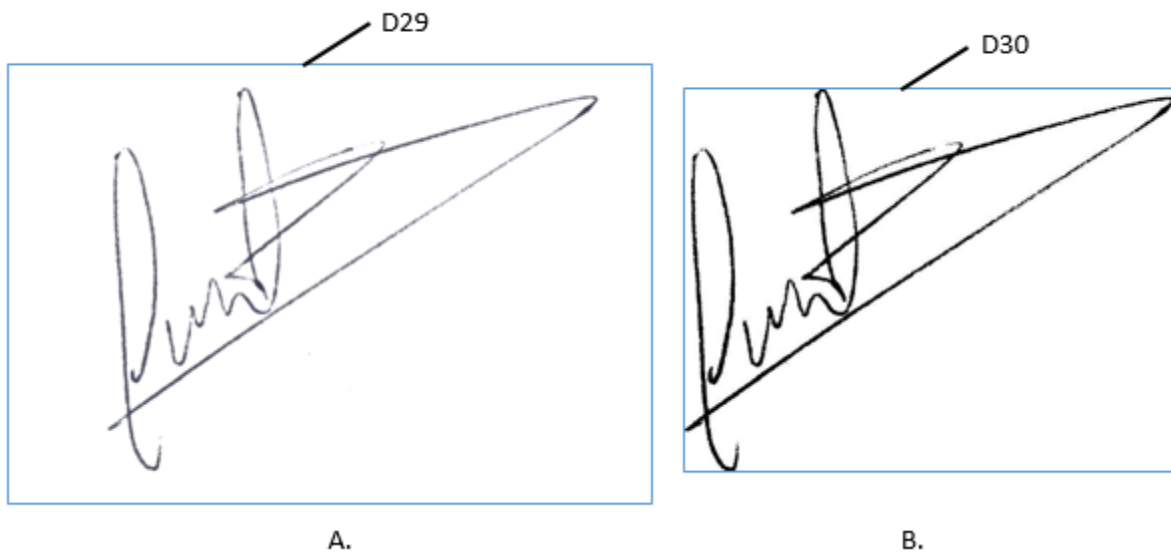


Fig. 8 Rectángulo mínimo: A) Firma en blanco y negro y un rectángulo, B) Firma en blanco y negro y en un rectángulo mínimo

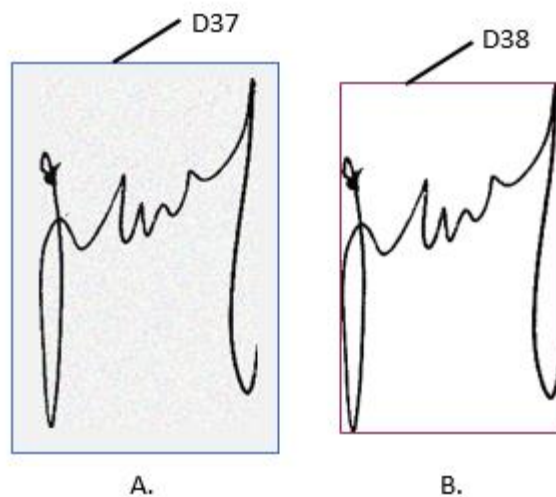


Fig. 11 Ruido y rectángulo mínimo: A) Firma sin pre-procesar B) Firma sin ruido, en blanco y negro y en rectángulo mínimo

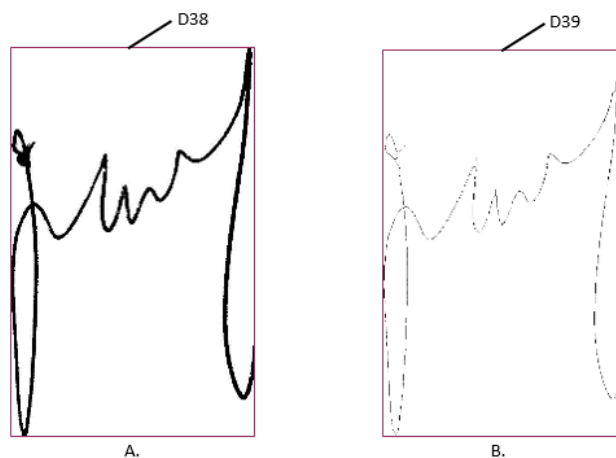


Fig. 12 Afinamiento: A) Firma en blanco y negro, sin ruido, en rectángulo mínimo, y sin afinar, B) Firma en blanco y negro, sin ruido, en rectángulo mínimo y afinado

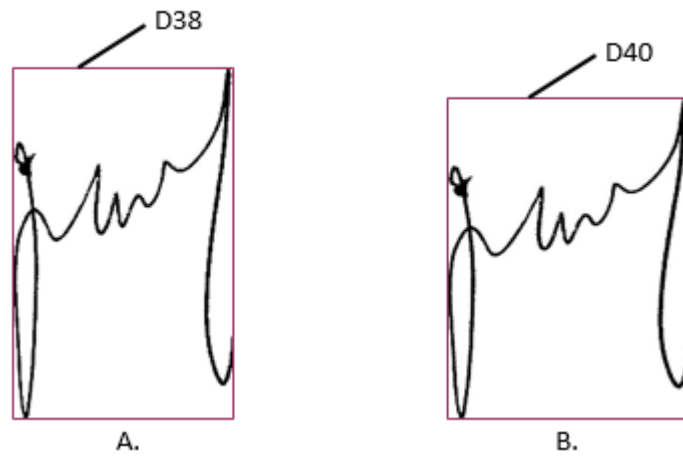


Fig. 13 Normalización: A) Firma en blanco y negro, sin ruido, en rectángulo mínimo y sin normalizar, B) Firma en blanco y negro, sin ruido, en rectángulo mínimo y normalizado.

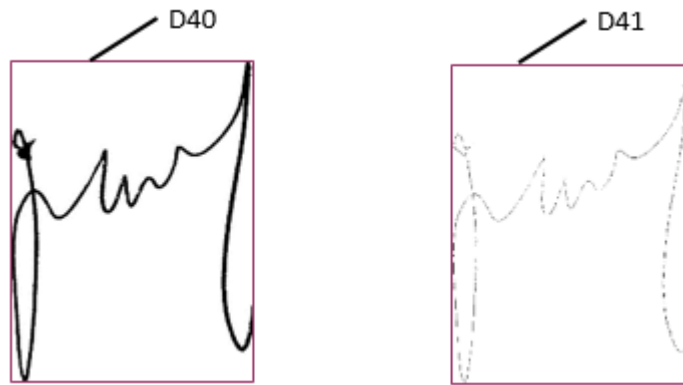


Fig. 14 Afinamiento de firma normalizada: A) Firma en blanco y negro, sin ruido, en rectángulo mínimo, normalizada, y sin afinar, B) Firma en blanco y negro, sin ruido, en rectángulo mínimo, normalizada y afinada.

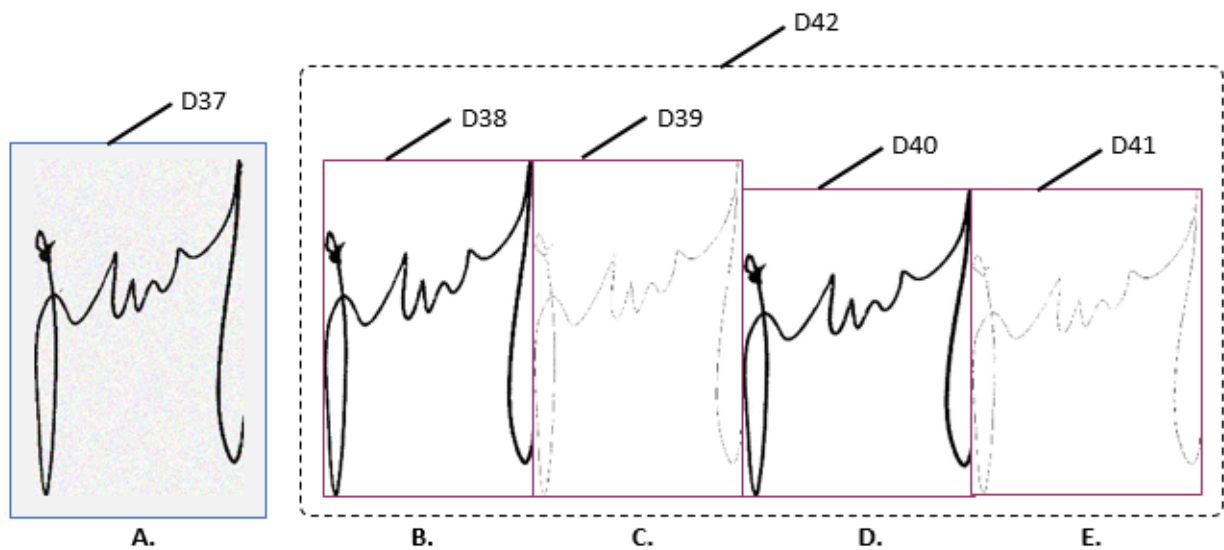


Fig. 15 Pre-procesamiento de una firma: A) Firma sin pre-procesar; B) Firma sin ruido, en blanco y negro, y en rectángulo mínimo; C) Firma sin ruido, en blanco y negro, en rectángulo mínimo y afinada; D) Firma en blanco y negro, en blanco y negro, sin ruido, en rectángulo mínimo y normalizada; E) Firma en blanco y negro, sin ruido, en rectángulo mínimo, normalizada y afinada.