

INNOVATE Perú
Programa Nacional de Innovación para la Competitividad y
Productividad

Proyecto N° 176-PNICP-PIAP-2015

Reporte Técnico R01-CONCYTEC/176-PNICP-PIAP-2015
Análisis comparativo de eficiencia de verificación automática de firmas
off-line en grande escala

Luis A. Rivera, David Mauricio, Gleen Rodrigues, Luis Vilchez, Luis
A. Estrada

Mayo 2016

Análisis comparativo de eficiencia de verificación automática de firmas off-line en grande escala

Luis A. Rivera, David Mauricio, Gleen Rodrigues, Luis Vilchez, Luis A. Estrada

Resumen

La demanda de la humanidad cada vez más exigente, a raíz de la evolución de la tecnología de computación, hace que científicos de computación dediquen esfuerzos a la verificación automática de firmas en documentos. Firmas que son digitalizadas, purificadas, en algunos casos, mejoradas a través de operaciones morfológicas, para luego ser descritas como elementos vectoriales numéricas fáciles de operar para su clasificación como firmas válidas o inválidas. Con ese objetivo, muchos métodos, en cada caso, fueron utilizados y validados respecto a la precisión de no confundir las firmas válidas como inválidas e inválidas como válidas que aún persisten. En este trabajo son analizados los diferentes métodos de verificación automática de firmas offline en los tres aspectos: caracterización, clasificación, validación con los databases usados y las métricas usadas para la precisión.

1. Introducción

La firma es un atributo biométrico de característica comportamental de cada persona. Un sistema biométrico reconoce a la persona en base de su carácter fisiológico y comportamental (Rivard et al, 2013). Esos atributos son inherentes a las personas, como tal no puede ser robado, perdido, ni olvidado como ocurren con códigos u objetos (Daramola y Ibiyemi, 2010). El proceso de reconocimiento es usado en la identificación, consulta y verificación de personas por sus firmas. La “identificación” establece la identidad de la persona por sus características que anteriormente fueron asociadas en un banco de datos, función que son bien aprovechadas por entidades para identificar a sus funcionarios y clientes. “Consulta”, que si el dato biométrico se aproxima a personas de una lista de identidades. La “verificación” constata la autenticidad del dato biométrico basado en la comparación de informaciones del banco de datos (Rivard et al, 2013) (Pushpalatha et al, 2014). En ese contexto, las firmas generalmente son consideradas por muchas instituciones jurídicas como medio legal de identificación de personas. Por tanto, existen personas que falsifican firmas para conseguir beneficios propios, grupales o institucionales.

Deore y Handore (2015) consideran ocho tipos de firmas falsas: aleatorio, casual, forzada, simulada, trazada, corte-colado, electrónico y libre. Pero firmas en papel con bolígrafo, generalmente son consideradas tres tipos de firmas falsas: aleatoria, casual y forzada. La aleatoria es realizada por el farsante sin la preocupación de cómo es la firma que se desea sustituir. La casual es cuando el farsante generalmente escribe el nombre de la persona, y la forzada es cuando el farsante imita la firma genuina. La firma forzada es más difícil de detectar porque la imitación puede ser tan real a simple vista. Sin embargo, las firmas genuinas también son confundidas como siendo falsas, porque no existen firmas idénticas realizadas por una persona.

Las investigaciones en reconocimientos de firmas parten del hecho de que firmas de diferentes autores tienen algunas características que las diferencian una de las otras, mientras firmas de la misma persona tienen características similares pero nunca son iguales. El problema es que no existe una regla general que las personas adopten para realizar sus firmas. Por otro lado, los estilos de firmas que dependen de idiomas e culturas adaptadas a sus formas de escribir – como los estilos de firmas de hindús, diferentes a los estilos de firmas latinos (Pal et al, 2015) –, hacen difícil encontrar un método para extraer las características representativas (Neamah et al, 2014).

Uno de los procesos bastante conocidos en verificación de firmas es la depuración de planillones de firmas colectadas en las calles de ciudadanos activos con deberes cívicos. Ese

proceso es común para petitorios de inscripciones de partidos políticos, pedido de consultas populares, pedidos de alteración de leyes por parte de la comunidad, entre otros. Las firmas colectadas en las calles deben ser verificadas para validar el pedido. Ese proceso es practicado por muchos países en sistemas de gobierno. En esos casos, la verificación es comparar cada firma recogida con las firmas genuinas de la persona que están en el banco de datos de la institución de identidades. Generalmente, esos bancos de datos almacenan varias versiones de firmas genuinas de cada persona.

En la verificación manual, los peritos expertos en grafotecnia analizan los elementos estructurales de las firmas que son: dimensión, dirección, inclinación, presión, velocidad, formas de enlaces, ornamentación, habilidad, y proporcionalidad o tamaño de las grafías. Consideran también, la formación de las firmas, clasificación de los movimientos sucesivos, y elementos especiales de las firmas como caracteres intrínsecos y extrínsecos. Ellos consideran que una firma falsa, en particular la imitada, tiene trazo lento, levantamientos del bolígrafo, retomadas, repasos, paradas, trazos temblorosos, mala interpretación del desenvolvimiento de los trazos, retoques, presión y velocidad variada. Pero, ni siempre un perito está disponible en todas las actividades de verificación de firmas, ni están en todos los casos que una institución requiera, mucho menos en la verificación de volúmenes de firmas como el caso de pedido de consultas populares, el proceso sería demorado y con un número grande de peritos. En este caso, es requerido un sistema de verificación automática de firmas

1.1. Sistema de verificación automática de firmas

Un sistema de verificación automática de firmas (ASVS) es un proceso realizado por un computador siguiendo las tres etapas típicas abordadas en la literatura: captura de la firma digital, pre-procesamiento, extracción de atributos y clasificación. La captura de firma digital, depende del tipo de ASVS, en general es la imagen digitalizada de la firma a partir de firmas en papel y bolígrafo. En el pre-procesamiento, la imagen digital debe ser preparada para una buena extracción de sus características. La firma en la imagen ofrece muchas informaciones para atributos descriptores numéricos que son elementos fundamentales para la operación de clasificación. En práctica, es imposible representar un espacio de atributos de todas las firmas para conseguir un modelo de un clasificador universal (Eskander et al, 2013).

El proceso de clasificación es la verificación, mediante operaciones matemáticas de los atributos, de una firma en relación a otras de referencia para decidir si la firma es genuina o falsa. Las operaciones matemáticas constituyen los clasificadores. Un clasificador puede ser una operación de comparación por similitud, por ejemplo, basada en una distancia Euclidiana como templates (Ghosh et al, 2014), o por cosenos de vectores (Abdelrahman et al, 2013), o por los más elaborados basados en modelos estocásticos como HMM (Hidden Markov Model) (Daramola y Ibiyemi, 2010), NN (Neural Net), SVM (Support Vector Machine), por k-NNs (k-Nearest Neighbor) (Abdelrahman et al, 2013), entre otros. Generalmente, el proceso de verificación debe ser basado en el enfoque de los verificadores humanos de firmas (forensic documents examinations) que consiste en enfocar en atributos locales y establecer criterios secuenciales para decidir.

Existen enfoques variados en selección de métodos de extracción de atributos y formas de clasificación, en particular en relación de complejidad computacional y el grado de precisión de los resultados. Los métodos buenos usan información local o combinados con global, y los métodos puramente global son lentos (Arathi et al., 2014). La comparación de templates por similitud de vectores característicos es más rápida porque son basadas en atributos locales, pero no necesariamente son bueno.

En la literatura, es común dividir los sistemas de verificación de firmas en dos tipos: online e offline. El diseño del sistema de verificación de firmas offline es más dificultoso que de la firma online, debido a la ausencia de las características dinámicas en las firmas. La verificación, en este caso, solo depende únicamente de la imagen, el objetivo es identificar la validez de una firma en formato de imagen de la firma en relación a las otras firmas; o sea, en sistemas estáticos solo se tiene la firma final, de documentos escaneados, por tanto se tiene menos

información directa del firmante. En sistemas online, las firmas son capturadas por dispositivos electrónicos, como lápiz tablets, digitalizadores, pantallas sensibles en PDAs, celulares, generados por bolígrafos electrónicos (Pirlo e Impedovo, 2012). Eso permite adquisición de informaciones dinámicas como presión, dirección orden, tiempo de duración de la firma, estilos de movimiento, velocidad, atributos de aceleración, levante del bolígrafo y otros. También existen otras categorías de informaciones dinámicas, como propuesta por Alsulaiman et al (2013), como la biometría basada en haptica para verificación de firmas on-line de usuarios. La información haptica, toque manual, incluye trayectorias de deslizamiento, información cutánea y kinésica que generalmente consiste de posición, velocidad, torque, orientación y fuerza provenientes de interfaces hapticas. Sistemas de verificación on-line pueden ser implementados con herramientas usadas en firmas off-line, como Alsulaiman et al (2013) que utilizan la programación genética.

La firmas offline aún pueden clasificarse, como indicado por Ferrer et al (2012), en dos tipos: la estática y pseudo dinámica. La estática demanda atributos geométricos de la firma, mientras que pseudo dinámico trata de estimar la información dinámica de la imagen estática por tres enfoques: a) métodos matemáticos, que estima el orden temporal de la producción de la marca; b) inspirado por teoría de control de motor, que recupera atributos temporales en base de trazos geométricos tal como curvaturas; c) método de análisis de firmeza de trazos y variaciones de intensidad de trazos. Ese método es generalmente usado para clasificación de firmas de fondo complejo en tonalidad gris, en firmas en documentos, con texturas y textos de fondo.

Los verificadores, desde el enfoque de usuarios, también se subdividen en writer-depend (WD) y writer-independent (WI). En la primera, se crea un modelo clasificador por cada persona. El principal problema del enfoque WD es que se necesita crear modelos nuevos por cada nueva persona, y se necesitan muchas muestras de firmas verdaderas por persona para crear un buen clasificador. Según Berolini et al (2010), en la vida real no hay muchas firmas disponibles y el modelo tiene grandes errores. Algunos han creado firmas deformadas (de las verdaderas) para aumentar las muestras. Mientras WI se desarrolla un sistema con un conjunto de firmantes (Guerbai et al, 2015)(Zois et al, 2015). Por eso, la WI emplea transformación dicótoma para suavizar las dificultades del conjunto de firmas que representa un número grande de usuarios.

Sea cual tipo fuera el verificador, existen varios enfoque de encarar el complejo problema de verificación. Velez et al (2009) enfatizan que, la variabilidad de forma y tamaño de las firmas del mismo firmante demanda una consideración de tratamiento tolerante de precisión y uso de registro de su configuración inicial de forma. Lo cierto es que las firmas genuinas de la misma persona siempre tienen variaciones entre sí, conocida como variabilidad intra-clase (Swanepoel et al, 2013)(Pirlo et al, 2013), que es deseada. También contiene variabilidad inter-clase (Kumar y Puhan, 2014), que es no deseada. Los problemas inherentes a las variaciones intra-clase son debido a factores como el momento, edad, enfermedad, ánimos, orientación del espacio a firmar, calidad del bolígrafo, ilegibilidad (Shekar et al, 2015), también debido a la posición del firmante. También la variabilidad intra-clase puede tender a variabilidad extra-clase, variación a lo largo de tiempo, debido estado emocional y estado de salud permanentes, y edad (Kekre et al, 2010). En virtud de eso, el verificador debe tener en cuenta el alto grado de variación relacionado a las características morfológicas del patrón de las firmas.

1.2. Problemas de verificación de firmas

Debido a la complejidad de los estilos de firmas, motivo de la dificultad de representación de atributos descriptores, las operaciones de selección ni siempre van a dar resultados deseados. En la literatura de ASVS son considerados cuatro tipos de errores consideradas, que son FRR (false rejection rate), FAR (false acceptance rate), AER (average error rate) y EER (equal error rate) cuando ocurre $FAR = FRR$. El FRR es la tasa de firmas genuinas que fueron consideradas como falsas, FAR es la tasa de que las firmas falsas pasaron a ser consideradas como genuinas. En realidad, esos dos parámetros deben ser minimizados para decir que el ASVS es eficiente.

El problema que se enfrenta en estos trabajos, según Powik et al (2015), es la variedad de uso de coeficientes de performance (FAR, FRR, EER, AER) con diversos focos en factores (accuracy,

sensitivity, specificity) y diferentes métricas (Kudlacik et al, 2012). Tampoco se explicitan del cómo se obtuvieron las firmas falsas random, simple y skilled, y sus tratos diferenciados, tan solo hacen referencia a las firmas falsas como algo genérico. Bajo ese criterio, dos firmas genuinas, de diferentes autores son genuinas, pero una firma en relación a otra sería falsa aleatoria.

Por otro lado, la verificación de firmas, tanto offline como online, enfrenta la diversidad de estándares de comparaciones de resultados, conjuntos de datos usados y la selección de números de firmas genuinas y falsas para los testes. Muchos trabajos usan banco de datos de firmas privadas sin un criterio de estándar de generación y selección de las firmas a usarse en la validación del modelo propuesto. Eso hace que métodos de extracción de atributos y verificación sean evaluados engañosamente, con dificultades de comparaciones de resultados (Kudlacik et al, 2012). Algunos papers reportan resultados mejores, pero no son confiables por usar dataset muy pequeños, o training hecho tanto con firmas verdaderas como falsas (Fabregas et al, 2009).

Las operaciones de pre-procesamiento no apropiadas, también, pueden afectar la eficiencia de métodos verificadores, como señalado por Wen et al (2009), tal caso de la imagen de la firma al ser normalizada, como la operación de giro para conseguir una alineación horizontal, altera algunas características naturales de la firma. Proponen, por eso, el método RPF (Ring-Peripheral Features), u otro método independiente de rotación e independiente de la periodicidad.

Un enfoque natural de abordar un problema es la suposición basada en observaciones. Por ejemplo, Kumar et al (2010) abordan la verificación de firmas bajo supuesto que una persona realiza su firma genuina como un trazo casi continuado y rápido, mientras que el falsificador lo realiza más lento y de trazos no tan continuados. Kovari et al (2013) y Fabregas et al (2009) trabajan bajo la premisa que cualquier característica fisiológica, incluyendo los movimientos de escritura y firmas, de la persona tiene distribución normal Gaussiana. Otra suposición, dada por Pirlo et al (2013) y Kekre et al (2010), es que las firmas genuinas de un mismo firmante puede ser considerada como versiones destorcidas del modelo ideal único. Las suposiciones, o premisas, son buenas justificativas para el uso de determinados métodos de extracción de características. Por ejemplo, Pirlo et al (2013) justifican el uso del flujo óptico para análisis de estabilidad local en regiones en la firma, por considerar que las firmas de un mismo firmante son versiones destorcidas de un ideal. El enfoque viene de la consideración que algunas variantes son típicas del firmante mientras otras variaciones no lo son. Con esos criterios definen la variabilidad intra-clase y inter-clase, respectivamente.

También no sería correcto que un ASVS implementado con técnicas de reconocimientos de firmas occidentales, también pueda funcionar con las firmas de otros continentes e culturas donde imperan alfabetos y grafías de diferentes estilos. En esos casos, las suposiciones son fundamentales, tal como el trabajo abordado por Pal et al (2013) -firmas basados en alfabetos bengales con métodos tratados por Pal et al (2012) para firmas hindú-. La firma hindú es diferente en forma y estructura en relación a los de los occidentales. Las palabras de caracteres que se tocan y con base horizontal superior, por tanto consideran variaciones en relación a líneas horizontales.

Por fin, en esta presentación de ASVS, se observa que las falsificaciones forzadas son los que generalmente pasan como genuinas, por tanto necesita mejoras en técnicas de verificación, y aún el campo está abierto para las investigaciones. Kovari et al (2013), basado en trabajos anteriores, afirman que mejor EER es alrededor de 9% en verificación offline con firmas falsas forzadas, mientras que en online se acerca al 2%.

1.3. Objetivos y organización

En este trabajo hacemos una revisión de la literatura en verificación automática de firmas, siguiendo las recomendaciones de Kitchenham (2009), para colocar en evidencia los métodos enfocados para una eficiente verificación automática de firmas en grande volumen en lo que

respecta a la representación de características, clasificación, métricas usadas para eficiencia y banco de datos usados.

Para alcanzar ese objetivo, la secuencia seguida en resto del trabajo es dada de la siguiente forma: En la Sección 2 se formula el método de la revisión, en la Sección 3 se presenta el sistema de verificación automática de firmas off-line, en la Sección 4 se hace el análisis de elementos usados para una verificación eficiente de firmas, y por fin, en la Sección 5 se concluye.

2. Método de revisión

Una buena parte del éxito en la verificación de firmas es la representación de los atributos descriptores. El descriptor debe mostrar, en forma cuantitativa, el detalle que esconde una firma de una persona para ser diferente de la firma realizada por otra persona, a pesar que las dos firmas se parezcan en forma. Al mismo tiempo, el descriptor debe mostrar la similaridad entre las firmas genuinas de una misma persona. Por otro lado están los clasificadores, que manipulan los descriptores para mostrar que una dada firma es genuina o falsa. Dada que las firmas esconden las características de genuinidad o falsedad de las firma, los clasificadores deben considerar la esencia de como las firmas son generadas, y no sea una mera operación de clasificación. Debido a que el problema de reconocimiento de firmas es un problema complejo, los resultados esperados son dados en términos de indicadores de precisión. Se debe de respetar parámetros de medir la precisión definidos internacionalmente. Finalmente, los datos utilizados son fundamentales para la credibilidad de los métodos de verificación de firmas formulados. Universalidad de los datos evitan la parcialización en la discusión de los resultados.

Frente estas motivaciones, se establecen las siguientes cuestiones que deben ser analizadas y respondidas en este trabajo.

- Q1: ¿En la formulación de los métodos de representación son consideradas la esencia de las firmas genuinas e falsas?
- Q2: ¿En que se basan los diferentes métodos de extracción de atributos de representación?
- Q3: ¿Los métodos de clasificación utilizados en la verificación son considerados la esencia de generación de firmas?
- Q4: ¿Los diferentes conjuntos de datos son los suficientemente universales, de forma que los métodos formulados ofrezcan resultados confiables?
- Q5: ¿La precisión del reconocimiento de firmas utilizadas son universalmente aceptables?

Esas cuestiones serán respondidas en el desarrollo de esta trabajo siguiendo, en parte, el método de Kitchenham (2009), considerando las fases de *Planning the review*, *conducting the review* and *Reporting the review*. Las cinco cuestiones, Q1, Q2, Q3, Q4 y Q5, corresponden a *Planning the review*. La Sección 3 corresponde a la *conducting the review*, y en la Sección 4 analizamos los casos abordados para *Reporting the review*.

Las fuentes de información para responder las preguntas planteadas se realizaron en los bancos: IEEE Xplore, ACM Digital Library, Springer, Science Direct, DOAJ, Taylor and Francis, Emerald y otros. La investigación cubre el período comprendido desde el año 2009 a 2016, considerando que el trabajo del estado arte de Impedovo et al. (2008) es uno de los trabajos más exhaustivo sobre verificación automática de firmas. Las claves usadas para la búsqueda de fuentes en esos bancos fueron con la combinación de las palabras “signature verification”, “signature recognition”, “offline signature”, “handwriting recognition”.

Los criterios usados para la selección y exclusión de fuentes, incluidos journals con factor de impacto SJR, son indicados en la Table 1.

Table 1: criterios considerados para inclusión y exclusión de fuentes de referencia.

Selección	Exclusión
-----------	-----------

<ul style="list-style-type: none"> • Presentan modelos, algoritmos, herramientas para la verificación de firmas off-line • Proponen caracterización y/o formas de obtenerlos • Responden directamente a las preguntas de investigación • Proponen métricas para la medición de la precisión de verificación de firmas 	<ul style="list-style-type: none"> • Trabajos verificación de firmas que no aportan a su automatización o la extracción de atributos • Trabajos dedicados solo a reconocimiento de firmas on-line • Trabajos de revisión de literaturas genéricas • Posters, editoriales, libros y tesis
---	--

En *Conducting the review*, las informaciones potenciales identificadas en el proceso de búsqueda, con un resultado de 11577 fuentes, fueron sometidos a un proceso de selección de acuerdo a los criterios de inclusión y exclusión establecidos en la Table 1, previa una revisión de contenidos afín de determinar su relevancia para el presente estudio, seleccionándose 56 trabajos de investigaciones relacionados con la verificación automática de firmas (ver Table 2).

Table 2. Potentially eligible studies y selected studies.

Source	Potentially eligible staudies	Selected studies
ACM Digital Library	280	1
IEEE Xplore	2551	18
Science Direcct	7160	16
Springer	162	9
DOAJ	333	8
Taylor and Francis	1012	3
Esmerald	79	1
Total	11577	56

Los estudios y análisis de las fuentes seleccionadas, a fin de responder las preguntas de investigación en la Sección 3 y sus respectivas discusiones y *Reporting the review* son abordadas en la sección 4.

3. Sistema de Verificación de firmas off-line

Los ASVSs offline tienen como entrada imágenes digitalizadas conteniendo la firma de la persona y como salida una información binaria de que la firma es falsa o genuina. Para ese propósito, las imágenes de entrada son manipuladas con operaciones apropiadas para colocar en evidencia la región de la firma (ROI: region of interest). En esa situación son aplicadas la extracción de sus características representativas, que es fundamental para una buena clasificación de la información afín de verificar la autenticidad de la firma. Todas esas operaciones son realizadas tradicionalmente en las fases captura, pre-procesamiento, caracterización y clasificación, como ilustra la Figura 1. Algunos, como Pirlo e Impedovo (2012), consideran solo tres fases sin incluir la captura, a pesar que la calidad de digitalización de las firmas es fundamental para el suceso de verificación de las firmas. Abroug et al (2014) presentan un estudio claro del estado de arte ASVSs offline, considerando las cuatro fases. En la caracterización enfocan criterios predominantes como representación geométrica y transformaciones matemáticas, geométricas, estadísticas y grafométricas. La clasificación basados en NN, HMM, SDC (simple-distance classifiers), SVM, template matching, DTW (dynamic time warping), modelos Fuzzy, RBF-SVM (support vector machine with RBF kernek) y MLP (multilayer perceptron). En los experimentos son considerados algunos de los errores (FAR, FRR, EER, AER) con diferentes datasets como GPDS-39, GPDS-100, MCYT-75, 4NSigComp2010, SIGMA y GPDS960. Otros autores, como Bertolini et al (2010), defienden la idea de usar FAR o FRR no es una buena métrica para el problema porque la cantidad de firmas falsas disponibles es mucho mayor que la de firmas verdaderas. También existe los data set privados, generados por los mismos autores, como el de Zhu et al (2009) con una data set definida por 40 personas cada uno con 20 firmas. Verificación con datasets de referencia dará mismo resultados?.

Buena parte de los trabajos en verificación de firmas offline consideran que la detección de firma falsa aleatoria es menos dificultoso que la forzada. Los métodos son usualmente orientados a detección de aleatorios y simples, mientras que las forzadas es más dificultoso a ser detectado (Pirlo e Impedovo, 2012). Existen enfoques, también, la detección de la combinación

de los tres tipos de firmas al mismo tiempo (Jaiwal et al, 2014). En todo eso, la esencia debe ser la variabilidad de las firmas del mismo firmante (variabilidad intra-clase) y variabilidad entre firmas de diferentes personas (variabilidad inter-clase) tratados con métodos estadísticos (Aranakatha et al (2015) (Velez et al, 2009).

En esencia, los métodos aplicados son enfocados en las fases de extracción de atributos y las clasificaciones, con algo de variantes en detalles en el pre-procesamiento; tópicos que merecen ser analizados en detalle en las siguientes sub-secciones.

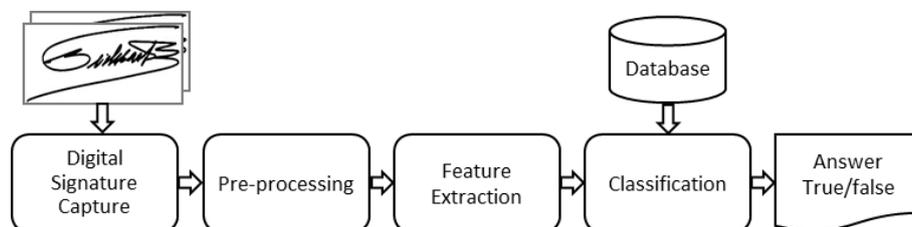


Figure 1: Automatic signature verification process.

3.1. Pre-procesamiento

En pre-procesamiento la imagen es preparada para una eficiente caracterización. Generalmente la imagen es binarizada por Tsu, pero algunos enfoques operan en tonalidad de gris como Arathi et al (2014). Son eliminados los ruidos usando operaciones de filtro morfológica, es capturado ROI, normaliza por una serie de dilataciones y ajustes de la imagen (Shekar et al, 2015).

La tonalidad de tinta usada en las firmas es siempre variada, que al digitalizar se registra en formato RGB. Son convertidas a la escala gris para luego pasar para binario (Radhika et al., 2011). Ferrer et al (2012) remueven el fondo de gris distorsionado y binarizan por posterización. Estandarizan el espesor del trazo de la firma con operaciones topológicas, remoción de puntos aislados y de fondo. Wen et al (2009) y Telios et al (2012), aparte de esos pasos, también adelgazan el trazo de la firma para un pixel (esqueletización). Algunos enfoques demandan muestreo (Pham et al, 2015; Neamah et al, 2014). Abdoli et al (2014) convierten la tonalidad del trazo a escala de 4 grises y fondo blanco. Tselios et al (2012) cortan la imagen a un 80% centrado en el centro de gravedad, para eliminar trazos externos y obtener un ROI.

Los ruidos causados por brillos y sombras en el escaneado deben ser reducidos con aplicación de técnicas como filtro mediana (Pirlo et al, 2013) (Radhika et al, 2011). Kumar et al (2012) aplican en la imagen binaria el análisis de componentes conectados, con un limiar apropiado, así las partículas aisladas de tamaño menor que el limiar son removidos. En algunos casos, para mantener buena conectividad de los trazos de la firma, se espesan los trazos, los niveles de intensidades son intercambiados con la operación de XOR y después se adelgazan e invierte la operación para una firma de líneas suavemente conectadas (Radhika et al, 2011).

Algunos autores, como Jena et al (2010), Kumar et al (2010), Rico et al (2012) y Arathi et al (2014), ajustan la imagen de ROI al menor tamaño con el método bounding box mínimo, en particular orientado en función de ejes principales que permite la corrección de inclinación. Otros autores, como Wen et al (2009), consideran que corrección de inclinación altera algunos detalles naturales de las firmas y prefieren usar con orientación original.

3.2. Extracción de atributos

Los atributos oculto en los trazos de las firmas deben ser extraídos eficazmente, e forma las firmas genuinas de una misma persona sean similares y firmas de personas diferentes sean no similares. El análisis de la distribución de los pixeles de las imágenes permite capturar atributos basados en parámetros geométricos, como área, dimensiones, relaciones entre dimensiones, presencia de puntos, como vértices y terminaciones, cruzamientos, cúspides, lazos, huecos, porciones aislados, y otros. También ofrece otras informaciones, como densidad de distribución, ángulos de vectores, texturas, matrices de forma, distribución de tonalidades gray, histogramas, variaciones topológicas, energías, transformaciones, entre otros. En general, las características

son extraídas en función de informaciones local, global e híbridas (Figura 2), considerando el aspecto estructural, geométrica, y estadística (Jena et al, 2010).

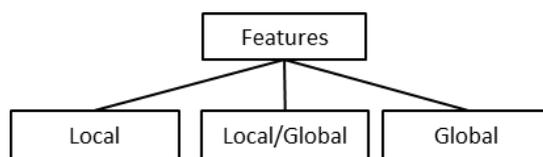


Figure 2: Atributos en categorías de informaciones.

La construcción de características para un ASVS eficiente, a pesar de que el estado de arte de los clasificadores ha ganado la atención, aún es un mar amplio que se debe explorar más (Zois et al, 2015). Muchos investigadores proponen diferentes enfoques de como extraer los elementos de características a partir de una imagen de firma. Las características son una representación de una entidad numérica, generalmente un vector de elementos numéricos que representa la esencia de la imagen de firma. Por eso, el objetivo es extraer con eficacia las características más representativas de la imagen. Para eso debe considerarse las características de generación de las firmas conceptuados desde varios ángulos. Por ejemplo, Batista et al (2015) apela a la esencia de la firma escrita de izquierda para la derecha, mostrando los detalles de variaciones en forma horizontal de las diferentes firmas de un único firmante.

3.2.1. Informaciones locales

Las informaciones locales tienden ofrecer elementos más detallados de la firma (Kumar y Puhan, 2014) (Tselios et al, 2012). Ellas usan variaciones posicionales, gradientes, informaciones periféricas, descripciones de cambios internos y externos de las firmas, atributos geométricos basados en envolventes y similitud de cosenos, líneas, arcos y ángulos. Son técnicas que utilizan operaciones de entidades numéricas en base de una vecindad de pixels de la imagen.

Para forensic document examiners (FDEs), las informaciones locales fornecen atributos observables por un verificador humano, tales como las propiedades de tipos de líneas y lazos. En líneas, son considerados puntos extremos, longitud e inclinación; en los lazos, son considerados perímetro, área, factor de forma, diámetro máximo, ángulo del diámetro máximo, redondez, diámetro del círculo inscrito, extensión, densidad, círculo bounding, ángulo del eje de momento, convexidad, solidez y ratio de aspecto. Los trabajos de caracterización basados propiamente en FDEs son pocos. Kovari et al (2013) inspiran su trabajo parcialmente en los métodos de FDEs combinando con conceptos de la dinámica.

Las divisiones de imagen permiten simplificar el análisis local de informaciones en sectores menores. La justificativa es que las firmas concentran algunos detalles en sectores menores. Los sectores pueden ser rectangulares, angulares, circulares o combinación de esas categorías. Por ejemplo, Neamah et al (2014) dividen la imagen en regiones angulares en relación del centro de gravedad e círculos concéntricos. Jena et al (2010) usan centros geométricos de imagen binaria para dividir iterativamente en sectores verticales y horizontales, de izquierda para derecha y de abajo para arriba. Pushpalatha et al. (2014) generan vector de atributos sub-dividiendo la imagen en rectángulos concéntricos, abordados previamente por Randhawa et al. (2013). Das et al (2009) usan features equal size distribution de la imagen de la firma divide en cajas verticales u horizontales de igual número de pixels de trazo de la firma. Rivard et al (2013) usan extracción de múltiples atributos en diferentes escalas de división de la imagen de la firma, y en cada celda analiza las medidas locales y selecciona las más apropiadas.

En general, trabajos de extracción local en la literatura, un tanto utilizando parte de FDEs, son basados en análisis de la imagen, y en la mayoría de los casos en sectores divididos de la imagen. Los elementos de análisis son basados en intersecciones de líneas con trazos de firmas, determinación de ángulos, análisis de intra-clases, densidad, métodos estadísticos, textura, morfología, energía, y variaciones tonales de gris.

Ángulos e intersecciones

En el modelo de Pirlo et al (2013a) la imagen es particionada en regiones siguiendo el enfoque de equimass (la división de histograma de masa horizontal y vertical de la imagen de la firma definen las líneas divisorias). Números de pixels negros del trazo de la firma interceptado por rectas paralelas trazadas, igualmente espaciadas, en direcciones de 0, 90, 45 y -45 grados, componen los atributos. Kudlacik et al (2012) definen features con distancias normalizadas de rayos disparados a cada 30 grados, del centro de gravedad de la imagen al punto de intersección con el trazo de la firma. Elrajubi et al (2015) dividen la imagen en sectores iguales, y en cada sector un atributo es la media de ángulos, establecidos en el vértice inferior izquierdo del sector, entre el eje horizontal y el segmento de recta que pasa por cada pixel del sector.

Otro enfoque, propuesta por Kumar y Puhan (2014), son las cuerdas limitantes superior e inferior de la firma contenida en una caja orientada. Las cuerdas son líneas que unen un punto-extremo del envolvente con el punto-extremo vecino inmediato. Un punto-extremo es el punto máximo o mínimo de las intersecciones de cada línea vertical imaginaria con el trazo de la firma. Un atributo es el ángulo de la cuerda con eje horizontal de la caja envolvente. Para Ghosh et al (2014), los atributos son ángulos de aristas en matriz de adyacencia de puntos de centro de gravedad generados recursivamente hasta un número significativo de veces, de sub-regiones divididas por líneas imaginarias horizontal y vertical, alternadamente, pasando por un punto de una sub-región.

Variabilidad y estabilidad

Wen et al (2009) extraen Ring-Internal Feature (RIF) y Ring-External Feature (REF) de N sectores circulares trazados de centro de gravedad de la imagen. Un círculo mínimo en el centro de gravedad contiene la firma. El RIF de un sector es el número de pixeles blancos del sub-área limitada por trazo de la firma y el origen, mientras que REF es el número de pixeles del sub-área limitada por trazo de la firma más externa y borde externo limitado por radio mayor. Eso genera 2N atributos. Para N grande, los atributos son casi similares, variando solo las posiciones, que en práctica es la variación intra-clase. Para las comparaciones deben ser alineadas eliminando desplazamientos de fases utilizando FFT del vector de elementos consecutivos en pares RIF y REF, para usar el vector solo con los N primeros, debido a los espectros simétricos en relación al punto central. La estabilidad local de regiones de trazos de firma es otro enfoque seguido por Pirlo et al (2013b) vía vectores de flujo óptico. El enfoque viene de la consideración que existen algunas variantes típicas del firmante y otras no. Con eso, una firma es verificada en relación a la compatibilidad de las variaciones observables de la firma genuina. Para eso, consideran un conjunto de firmas genuinas, que combinadas por cada par, definen un vector de variabilidad intra-clase, y otro conjunto de firmas falsas que combinadas en pares con las genuinas definen vectores de variabilidad inter-classes.

Densidad

Número de pixel del trazo de la firma en relación al número de pixels de la región ofrece información local de compactitud y concentración de trazo de la firma. Batista et al (2015) genera el vector característico utilizando la técnica de densidad de celdas definidos por 40x16 pixels, siendo la imagen de la firma definida por 10 x 62 celdas. Arunakatha et al (2015) dividen ROI en 800 celdas de 5x5. Por cada celda, computan la densidad de pixeles y la distancia del centro geométrico al vértice inferior izquierdo. Con eso obtienen 1600 atributos. Batista et al (2010) dividen la imagen binaria en 60 celdas horizontales, y los atributos son la densidad de puntos del trazo en cada celda, centro de gravedad, inclinación axial y curvatura del trazo. Bertolini et al (2010) dividen la imagen en celdas rectangulares, para cada celda son calculados atributos del trazo que son cuatro atributos grafométricos: densidad, inclinación, distribución de pixeles y curvatura. La distribución es dada por la altura y ancho de la proyección del contorno de trazo en relación a cada lado de la celda. A curvatura del trazo asociado a una curva cúbica Bezier básico. Ellos generan varias situaciones variando dimensión de celdas.

Estadísticos y entropía

Uso explícito de operaciones estadísticas en función del comportamiento de pixels de la imagen es una herramienta bastante usada en la literatura. Para Zois et al (2015) los atributos son

definidos por posicionamiento y operaciones combinadas de máscaras de forma a atender las variaciones intra-clases de las firmas. 32 símbolos binarios son delineados con una máscara 5x5, considerado como alfabético de una fuente probabilística. Kumar et al (2012) extraen la información de naturaleza curvilínea y densidad local de los pixels del trazos de la firma. Con esos criterios, componen un conjunto de atributos basado en surroundedness de los pixels vecinos. Para cada pixel negro del trazo es medido surroundedness a diferentes distancias. Es calculada la distancia Chebyshev en círculo de radio en pixels con centro en el pixel candidato. Por cada distancia son obtenidos entropía, momentos de orden 1, 2 y 3.

Arathi et al (2014) dividen la imagen en sectores de columnas, de izquierda para derecha. Generan la secuencia de números de pixeles negros de cada sector, considerada como serie de tiempo. Analizan los posibles shapelets (funciones de corta duración) de tamaño 3 hasta tamaño de longitud de la serie de tiempo. Selecciona las mejores shapelets basadas en contenido de información usando entropía. Pal et al (2015) extraen características de imagen en cuatro grupos: connected components (separación en palabras y la distancia entre ellas); enclosed regions (atributos geométricos de lazos); basic features (orientación y dimensión de la firma); curvelet feature (componentes de transformada de curvelet). Con esos valores definen modelo simbólico de intervalo para cada persona, parametrizados por indicadores estadísticos, como la media y desviación estándar. Para Angadi et al (2014) las características son formadas por las dimensiones del área útil de la firma binaria, centroide, media y desvío estándar del vector procedente de la transformada de Gabor tomada cada treinta grados de la horizontal.

Textura

Atributos de textura y geométricos son juntados por Abdoli et al (2014). La imagen es particionada en 4x3 bloques con 60% de overlap. Para cada bloque son calculados los atributos Local Derivative Pattern (LDerivP) y la distancia geodésica (distancia mínima entre dos puntos) del trazo de la firma. El LDerivP ofrece informaciones de atributos pseudo-dinámicas, como variaciones de direcciones de derivadas de primer orden calculadas respecto a un pixel en relación a sus vecinos en direcciones 0, 24, 90 y 130 grados. Mientras que la distancia geodésica ofrece información estática. Son obtenidos el grado de similaridad de cada caso, y después son usados para calcular la similaridad de ambos. Esteban et al (2012) aplican la extracción de trazos en una lista de contornos, posteriormente separan en segmentos de curvas parametrizables de tamaño limitado. Se calculan el ángulo de la tangente y curvatura en punto medio de cada segmento. La acumulación de evidencias busca segmentos de firma limpia que coincidan con segmentos de textura.

Morfología

La forma de los trazos consideradas como morfología de contorno es usada por Kekre et al (2010), como consecuencia de dilataciones y mapas de sobre-posición. La dilatación es realizada con 4 círculos de diferentes radios y colores, generando siluetas de 4 colores que se superponen, definiendo campos de atributos basados en pixels. Kumart et al (2010) buscan las morfologías que puedan entrar en una ventana de $n \times n$ pixeles, y calcular la frecuencia de cada una de las morfologías como segmentos de rectas o curvas de 1 pixel de ancho con n de largo. Con $n = 7$ generan 2.720 morfologías posibles, siendo ese el tamaño del vector de atributo que debe ser reducido para 44 usando MLP para atributos más importantes. Las características morfológicas del patrón de la firma, representadas en términos de histogramas, son aprovechadas para la extracción de características por Shekar et al (2015). Ellos dividen la imagen de la firma en ocho sectores verticales, y en cada sector es calculado un disco máximo inscrito para cada pixel de la firma, obtenido un espectro patrón local que es convertido en un histograma de 10 columnas. Con eso se obtiene un vector de características de 80 elementos.

Variaciones de Gris

La firma digitalizada en tonalidad de gris es importante para un análisis efectivo de informaciones para las características, para algunos investigadores. Para Ohshima et al (2015), las intensidades de presión ejercida son determinadas por variaciones de escala de gris de cada pixel, y los trazos son dilatados por una operación morfológica. La imagen es particionada por

una línea vertical pasando por centro de gravedad de la región. El centro de gravedad en relación al eje horizontal es calculado en función de la posición y la intensidad del trazo. En cada región es calculada la gradiente. Los atributos son definidos por histograma direccional de la gradiente y distancia de Mahalanobis. Xu et al (2014) usan el Local Binary Pattern (LBP) en escala gris de la firma, que es robusto contra variación por iluminación pero no respecto al ruido, compara con Local Direction Pattern (LDP). Sedouk et al (2015) extraen atributos con GLBP (gradient local binary pattern) de la línea de procesamiento de imágenes. Calcula LBP, luego busca la distancia 1 de unos para cada pixel central para determinar el ángulo de Freeman de esa región de unos y la gradiente. Así obtienen atributos longitud, ángulo y gradiente.

Topologías, histogramas y energías

Uso de topologías, histogramas, geometría y energías son también usadas en la determinación de características local. Serdouk et al (2014) usan características topológicas ortogonal combination local binary pattern (OC-LBP) y el Long-Run Feature (LRF). Con OC-LBP mejoran a información de color en regiones de imagen limitando el tamaño de histograma del LBP, para obtener histogramas sobre diferentes vecinos ortogonales. LRF cuenta píxeles de trazo en dirección horizontal, vertical y diagonales. La suma total de los valores en cada dirección define un LRF, por tanto se obtienen cuatro características.

Contornos deformables, basado en variaciones de energías internas y externas, envueltiendo la firma con Snakes son utilizados por Velez et al (2009). Las energías internas ofrecen las restricciones de elasticidad y rigidez del snake, y la externa ofrece algunos atributos como bordes, esquinas, intensidades de pixels y otros. Esas energías son minimizadas usando programación dinámica. La energía de snake, con centros de gravedad coincidentes, se ajusta iterativamente a la forma de la firma por una curva paramétrica, de forma que el snake no sea demasiado deformado. Los parámetros del snake final son los atributos deseados.

3.2.2. Informaciones globales

Las informaciones globales son abstraídas como una síntesis de la imagen en forma general, sea en espacio de frecuencia o en espacio de imagen.

El DCT (Discrete Cosine Transform) convierte la imagen para el dominio de frecuencias, usado por Daramola e Ibiyemi (2010), Bharathi y Shekar (2014). Extraen 100 características representativas, de una pequeña porción de la información concentrada en la parte superior izquierdo de una región, recorriendo los pixels de horizontal-diagonal-vertical de superior izquierdo para lado inferior izquierdo. Bharathi et al (2014), al igual que Bharathi e Shekar (2014), obtienen los 10x10 componentes de bajas frecuencias, reducen dimensión menor con 2D2FLD y escogen un grupo de auto-vectores de mayores auto-valores.

Shirdhonkar e Kokare (2012) usan el duo DT-DWT (Dual Tree Discrete Wavelet Transformation) y DT-RCWF (Dual Tree Rotated Complex Wavelet Filters) para extraer detalles en 12 direcciones diferentes, inspirados en recuperación de texturas (Kokare et al, 2005). Por las desventajas de DWT - Haar, como poca capacidad de direccionamiento y sensibilidad variante, es complementada con CWF (Complex wavelet Filter), en descomposición hasta nivel 6. En cada nivel son calculados dos valores del vector característico: energía (e) y desviación estándar (s). Obteniéndose los 12 valores que define el vector de atributos (s1, s2, ..., s6, e1, e2, ..., e6). Ribeiro et al (2014) estudian e implementan varios métodos de caracterización de firmas, entre ellos Best Fit, DCT, Geometric parameters, histogram parameters, modified direction feature (MDF), six-fold-surface, wavelet transformation, e algunas combinaciones. Curvelet, que é una extensión del concepto de Wavelet, es usado por Guerbai et al (2014) para construir los atributos a partir de los coeficientes de la transformada. Con los coeficientes, que son una representación de aristas y otras singularidades a lo largo de la curva, calcularon las energías y desviación estándar.

Modelos usando matriz de Discrete Radon Transform (DRT), que consiste en la proyección de la imagen en relación un lado definido en un ángulo, obteniéndose histogramas en columnas

(Abdelrahman et al, 2013). Tambien, Swanepoel et al (2013) transforman las firmas en conjuntos de atributos por DRT.

3.2.3. Informaciones local-global

En (Pham et al, 2015) consideran dos tipos de atributos: Apariencia global de la firma por adopción de forma descriptor de contexto, y características más finas, o locales, son obtenidas con la adaptación del descriptor gradiente de posición e histogramas de orientación. Atributos finos son comparados por distancia Euclidiana, mientras que los atributos globales requieren de solución de problemas de correspondencia entre puntos amostrados de las dos firmas, resuelto por la aplicación de “problema de asignación lineal” interactuando con técnica iterativo de “consenso muestra aleatorio”. Bharathi y Shekar (2014) enfatizan las características globales, estadísticas y geométricas, donde el global describe la imagen como un todo, la estadística es extraída de la distribución de los pixels de la imagen, y la geométrica es la descripción topológica de la firma. Una combinación global y local, en función de entropía y geometría, es abordada por Jaiwal et al (2014). Ellos sacan 13 atributos de esa mezcla, compuesto por número de pixels negros, ratio ancho/alto, máximo entre vertical y horizontal de la imagen, área de la firma, entropía de la distribución de pixeles en gris, energía horizontal y vertical, diagonal derecha e izquierda, longitud de la trayectoria, número de puntos de intersección y de puntos de borde.

Histograma es una forma de sintetizar los componentes de los elementos de la imagen de forma global. Zhang (2010) caracteriza la firma por Pyramid Histogram of Oriented Gradient (PHOG), que representa forma por histogramas de orientación de aristas computadas por cada subregión de la imagen. Cada barra de histograma representa el número de aristas que tiene orientaciones dentro de un rango angular de 0 a 180 grados. La forma piramidal es construida por la división de cada región de la imagen en subregiones menores, hasta 3 niveles para evitar la saturación. El vector característico es la concatenación de los vectores de cada nivel.

Los momentos de invariantes de Hu de imágenes binarias en estructura quadtree son aplicados por Radhika et al (2010, 2011). Para cada quadtree components (QC) es calculado los momentos de orden menor 1^o-2^o y 7^o de Hu que derivan las características de forma. Calculan las varianzas de cada tipo de momentos y seleccionan una cuarta parte del número de los QCs por el criterio de minimum variance quadtree component (MVQC). Los autores afirman descartar las partes de las firmas con variaciones mayores que ocultan los aspectos inter-clases. Rico et al (2012) dividen ROI de la imagen binaria en subregiones menores para definir 11 vectores característicos. El primer vector contiene masa de la firma de cada subregión; los siguientes 4 vectores son definidos por los acumulados de impactos de líneas disparadas para cada lado por los pixels de fondo de la imagen con los pixeles de la firma. Cada dirección de disparo genera un vector. El sexto vector es definido por número de pixels de fondo aislados por trazo de la firma. Los 4 restantes so obtenidos por la forma de vecindad del contorno de la firma, utilizando las 4 direcciones de un pixel del borde en relación de los 8-vecinos. El último vector contiene el número del resto del trazo de la firma con erosión y número de trazos de la imagen del nivel anterior a la última erosión.

3.3. Clasificación

La clasificación es dividida, según Hu et al (2014), en tres grupos: template-matching, estadísticos y estructurales. Las combinaciones de métodos de esos grupos, los híbridos, también dan buenos resultados. Por tanto, los modelos clasificadores de reconocimiento de firmas pueden ser agrupados en cuatro grupos (Figura 2). En template-matching una firma se compara en relación a una genuina, un caso común es dynamic time warping (DTW). Los estadísticos usan modelos estadísticos son los que usan métodos estadísticos en sus operaciones, como redes neurales y modelos. Las estructurales separan en clases, como el SVM.

Los modelos HMM, según Doere y Handore (2015) y Batista et al (2015), son adecuados para firmas que se propagan de izquierda para la derecha. Las redes neuronales clasifican la firma como siendo genuina e falsa. Los modelos basados en semejanza de templates, consideradas

como semejanzas óptimas y elásticas, determinan la falsedad de las firmas cuando comparadas con una genuina son distantes. Las semejanzas elásticas usan atributos bidimensionales, mientras las semejanzas óptimas usan atributos unidimensionales. Los métodos estadísticos analizan la correlación de los coeficientes para encontrar alguna relación entre los elementos del conjunto de firmas. El SVM usa atributos de alta dimensión, y trata de encontrar las diferencias entre clases de datos.

Algunos enfoques de clasificación son inspirados en los procedimientos de los expertos en caligrafía forense, que comparan dos firmas en dos niveles: apariencia global de semejanzas y análisis en detalles. Caso no haya semejanza confiable, entonces las dos firmas son de diferentes personas. Solo en caso de semejanzas evidentes es realizado un análisis detallado para coleccionar más evidencias de prueba. Es una preposición de modelos de basados en informaciones geométricas globales y detalles locales de las firmas.

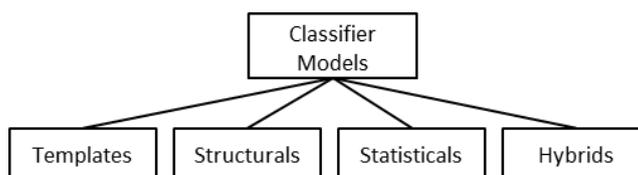


Figure 2: Modelos de clasificadores utilizados en verificación de firmas.

Templates o directos

Basado en el grado de similitud entre dos firmas por operaciones de sus respectivos vectores de atributos. Por ejemplo, Shirdhonkar e Kokare (2012) usan la distancia Canberra como métrica de comparación entre dos vectores. Pirlo et al (2013) enfocan la clasificación desde la óptica de estabilidad, considerando que firmas genuinas tienen estabilidad menor. Para eso, los cosenos de los vectores de las respectivas regiones de dos imágenes, divididas en regiones en función de equimass de firmas, definen un conjunto de valores reales que determina la variabilidad entre esas imágenes. Modelos basados en k-NNs son regularmente utilizados en combinación de distancias de similitud (Abdelrahman et al, 2013). La idea es que las firmas se agrupan en clases de firmas próximas en características, así el análisis de 7 vecinos bien próximos minimiza el tiempo de comparaciones. Neamah et al (2014) utilizan comparación por similitud de grafos no orientados, de elementos característicos, determinado por algoritmo dinámico.

El método seguido por Jena et al (2010) usa la distancia Euclidiana entre cada par de vectores en relación a una tolerancia obtenida en función de la media de los datos de prueba y sus respectivas varianzas. Rico et al (2012) validan las firmas, en una primera instancia, usando distancia Euclidiana. Después, para refinar las comparaciones, definen 11 clasificadores individuales basados en vecinos más próximos. Kekre et al (2010) comparan los atributos no comunes en las siluetas del trazo de cada par de firmas dilatadas en relación a un umbral asociado a la firma de cada persona. El umbral es registrado junto con firma genuina. Pal et al (2015) establecen comparación por similaridad entre firma a verificarse y la data simbólica del firmante, en relación al intervalo que puede ser dentro, en el borde o fuera.

El refuerzo en esta categoría, es el enfoque asumido por Kovari et al (2013) que features de cada persona tienen una distribución normal, consideran que los atributos de las categorías de firmas deben caer dentro del intervalo definido por la media y desviación estándar determinado por atributos seleccionados de un conjunto de firmas de entrenamiento. Por tanto, la distancia de similitud estaría en función de la media y desviación estándar de firmas genuinas.

Estructurales

Modelos basados en árboles de decisión y SVM constituyen el tipo de clasificador estructural. Pirlo et al (2013) utilizan una estructura de árbol de decisión alterna (ADT) generadas en base de vectores de atributos de variabilidad de flujo óptico. El ADT contiene nodos de decisión que especifican predicados de condición, y nodos de predicción que contienen números. En la clasificación, cada vector es comparado con los nodos de ADT, y si resulta verdadera entonces se incrementa el voto de ese vector. Al final, el vector que obtuvo mayor voto se define como

variabilidad compatible, validando la firma en análisis. Arathi et al (2014) usan el árbol de decisión para la clasificación, donde los nodos no terminales son los shapelets y la distancia de Mahalanobis, y las terminales la clasificación.

Los modelos SVM y sus variantes son bastante usados en verificación de firmas, evidenciados por Bharathi y Shekar (2014), Kumar y Puhan (2014), Ohyama et al (2015), Swanepoel et al (2013), Sedouk et al (2015) y Radhika et al (2010). Una variante de SVM, presentada por Guerbai et al (2014), es el modelo OC-SVM (one-class support vector machine) basado en una hiper-esfera en que al menos un grupo de elementos de entrenamiento está dentro. Ohyama et al (2015) proponen un sistema basado en SVM combinando resultados de verificaciones obtenidas con técnicas on-line y offlines de firmas japonesas. En pre-procesamiento las firmas son preparadas para verificación online u offline. Usan la programación dinámica para relacionar los datos de series de tiempo. Al final, un modelo SVM decide si la firma genuina o falsa. Guerbai et al (2015) construyen un modelo one-class SVM (OC-SVM) solo con firmas genuinas, justificando que las instituciones que identifican firmas, como los bancos, solo registran firmas genuinas. Asumen que cada usuario tiene su propio modelo, pero todos los usuarios de ASVS tienen los mismos parámetros. Cuando un nuevo usuario es presentado al sistema, el modelo es creado sin calcular sus parámetros, así el sistema es considerado como WI.

Zhang (2010) utilizan un SVM multi-clase basada en la relación binaria “un-todos”, en que es necesario un SVM para cada clase de los k usados. Consideran que el enfoque “un-un” es más intensa computacionalmente, ya que los resultados de cada par de SVM deben ser computadas. Zhang compara ese método en relación a otros métodos de clasificación como kNN, regresión logística, Fisher Discriminating Linear (FLD) y MLP. Kumar et al (2010) entrenan el verificador SVM con kernel RBF. Xu et al (2014) usan, con atributos en LBP y LDP, clasificador SVM con kernel lineal. Bertolini et al (2010) definen un conjunto de clasificadores de decisión parcial basado en SVM, uno para cada clase de firma genuina de un conjunto de personas, alimentados por el grado de disimilaridad definido por la norma de la diferencia de par de vectores de atributos, una genuina e la otra cualquiera (puede ser genuina de la misma persona o de otra considerada como falsa). La decisión final ofrece la fusión de las decisiones parciales por mayoría de votos. En el entrenamiento consideran casos positivos (combinación de firmas genuinas por persona) y negativos (combinación de firmas de pares de personas diferentes).

Tselios et al (2012) crean un clasificados basados en SVM por cada firmante con un umbral común para todos, validado por varios datasets. Randhawa et al (2012) modelan un clasificador basado en SVM con entrada generadas por invariante de Hu de la firma particionada en tres regiones. Para los entrenamientos, testeos y validaciones es generado un banco de firmas de 300 firmas genuinas y 200 falsas. Ribeiro et al (2014) compara a eficacia de atributos de un grupo de firmas obtenidos por varios métodos con clasificador SVM en GPU que incluye librería GPUMLib y usando por CUDA de NVIDIA.

Estadísticos

Técnicas de Artificial Immune Systems (AIS), inspirado en biología y computación natural, son usadas por Serdouk et al (2014) para construir el sistema de reconocimiento artificial inmune (AIRS), que es un algoritmo supervisionado de recursos limitados. El núcleo de clasificación es inspirado en K-NN. Los autores afirman que el sistema se comporta mejor que SVM. Posteriormente, ellos mismos, Serdouk et al (2015), formulan una versión mejorada de AIRS sustituyendo K-NN por un SVM. Las características seguidas para entrenamiento y teste de modelo son llamados antígenas, mientras que las unidades del sistema son llamados células B que compiten unos con otros por un número fijo de recursos.

El Artificial Neural Network (ANN) es una alternativa eficiente en reconocimiento de firmas. Kumar et al (2012) examina la eficiencia de los atributos de las firmas usando dos clasificadores populares: multilayer perceptron y SVM. Radhika et al (2011) utiliza radial basis function (RBF) alimentados por suma de momentos de MVQC, cuyos resultados alimenta, en segunda fase, el clasificador forward neural network (FNN). Jaiwal et al (2014) emplean una ANN feed

forward back propagation y entrenada con Levenberg-marquardt. Para validar el modelo usaron dataset privado generado por ellos. Bharathi et al (2014) clasifican las firmas con red neural MLP de 2 capas ocultas usando data set CEDAR y GPDS-160 y Angadi et al (2014), con red neural back propagation de 2 capas ocultas, con datasets propio.

En (Batista et al, 2010) los vectores de atributos son clusterizados para generar el codebook como alfabeto de símbolos para entrenar varios HMMs left to right, con diferentes estados. Daramola e Ibiyemi (2010) modelan un HMM con cuatro estados para cada persona. Kudlacik et al (2012), debido a la dimensión variante de vectores de atributos basados en distancias de intersección, proponen un modelo difuso para cada persona como un conjunto fuzzy de datos. Por cada firma nueva a verificar, es realizada la operación de membresía fuzzy de cada elemento y de todo el conjunto. Aplican un umbral por persona para aceptar o rechazar la firma, ajustando el umbral al peor caso de membresía de la firma. Arunalatha et al (2015) calculan varianza de componentes principales de los vectores característicos, colocados en una matriz, de las firmas de entrenamiento. En la comparación, los nuevos vectores característicos se adicionan en la última columna de la matriz de entrenamiento. Si la varianza es menor que el ideal de entrenamiento, entonces la firma en cuestión es clasificada como falsa.

Híbridos

Batista et al (2015) proponen un modelo híbrido de HMM (como clasificador generativo) y selección dinámica como clasificador discriminativo, un modelo compuesto por clasificadores en dos etapas. Modelos híbridos HMM y SVM son usados por Pushpalatha et al (2014) que clasifica en primera etapa con HMM y complementa con SVM en segunda etapa. Batista et al (2015) proponen un modelo híbrido de HMM (como clasificador generativo) y selección dinámica como clasificador discriminativo. Velez et al (2009) usan métricas de coincidencia y distancia, y energía (snake). La primera usa mapa potencial de la imagen, que es 1 si el snake pasa exactamente por medio de los trazos más gruesos de la firma, y baja cuando se desvía de ese trazo. La distancia es la suma de los pixel en función del ancho del trazo. También usa el enfoque difuso de primer orden de Takagi-Sugeno.

4. Análisis de los elementos usados

Un método eficiente de reconocimiento de firmas offline envuelve, en primera instancia, una eficacia en análisis para describir las características de las firmas. Eso significa que los elementos usados en la definición de los descriptores deben ser debidamente justificados, si posibles con las debidas demostraciones. En la revisión de la literatura, Sección 3.2, fue raro observar alguna justificación del por qué se adopta un método de caracterización de la firma. Los métodos usados aparecen como siendo hecho al azar, solo vistos desde los objetivos de obtener atributos numéricos y algo diferente de los otros métodos ya estudiados. Esta deducción responde a la cuestión Q1. En la Table 3 es mostrado los descriptores de atributos usados que responde a la cuestión Q2, basado en el análisis de los métodos de extracción de atributos en la sección anterior. Caso similar ocurre con los clasificadores, la esencia de análisis de firmas por su originalidad es confiada a la comparación mecánica basada generalmente en los métodos de machine learning, para responder la cuestión Q3. Otro elemento fundamental, que da un valor de eficiencia de los métodos usados, son los datos para las validaciones. Algunos datos parcializados, posiblemente, favorecería a ciertos métodos y desfavorecería a otros, no sería posible establecer el principio de inducción para generalizar la validez del método en un universo mayor. Ese análisis permite responder la cuestión Q4. Un método de inducción también requiere de variables formales debidamente establecidas. En este caso son las métricas establecidas para medir en forma universal la eficiencia del método, para responder la cuestión Q5. Como descrito a seguir, e las siguientes subsecciones, existen cantidad de trabajos, pero no hay eficiencia universal en verificación de firmas offline se refiere.

4.1. Descriptores de atributos usados

Los diferentes descriptores de atributos que son posibles de cuantificar, usados en los tres tipos de atributos –sea para generar atributos locales, globales e híbridos– en verificación de firmas off-lines, tal como analizados en las subsecciones anteriores son sintetizados en la Table 3. Los descriptores envuelven cálculos y operaciones de ángulos y sus proyecciones concebidos entre segmentos de rectas y gradientes de variaciones de píxeles, arcos, curvaturas y convexidades de trazos de segmentos de firmas. Las operaciones morfológicas usadas en análisis de conectividad, tipos de distancias, la esqueletización, dilataciones, determinación de lazos, variaciones de tonalidades en caso de entropías y texturas, intersecciones o impactos de rayos con píxeles de firmas. Las operaciones geométricas con entidades de retas, arcos e vértices identificados permitiendo cálculos de número de Euler, áreas, masa, centros de gravedad, intersecciones, máximos horizontal y vertical, perímetros, relación altura y ancho e orientaciones. También se observan uso operadores de transformación como DCT, DRT, DWT, FFT, Curvelet y Gabor cuyos coeficientes y parámetros generan elementos de caracterización. Son usados, también, parámetros estadísticos, probabilísticos, desvío estándar, varianza, distribución de pixels, histogramas, flujo óptico, momento de Hu, entre otros. Esa variedad de descriptores detectados en la literatura detalla la respuesta de la cuestión Q2.

A Figure 3 ilustra la frecuencia de uso de los descriptores que las fuentes analizadas en la definición de los atributos. Se observa que los descriptores basados en ángulos son más usados, seguidos por densidad y gradientes. Preferencias por histogramas, intersecciones, masa y distribución de píxeles están en tercer lugar. En cuarto lugar son usados atributos basados en distancias, DCT y morfología, seguidos por descriptores basados en energía, entropía, centros de gravedad y LBP. En los últimos lugares de preferencias, pocos usados, están los descriptores basados en círculos, arcos, conectividades, curvatura, curvelet, DRT, número de Euler, centro geométrico, variación de gray, Hu momento, LDP, loop, orientación, perimeter, ratio (high/wide), standard deviation y stroke curvature; seguidos por menos preferidos, angle projection, área, convexity, DWT, FFT, geodetic distance, Gabor transformation, LDerivP, LRF, maximal horizontal, optical flow, perimeter ring, rope momento y texture.

Table 3: Feature descriptors used in off-line signature verification.

ID	Feature descriptors	References	Num
D01	Angle	Das et al (2009), Pirlo et al (2013a), Elrajubi et al (2015), Ghost et al (2014), Kovari et al (2013), Kumar and Puhan (2014), Esteban et al (2012), Serdouk et a (2015)	8
D02	Angular projection	Rico et al (2012)	1
D03	Area	Kovari et al (2013)	1
D04	Arcs	Tselios et al (2012), Kumar et al (2012),	2
D05	Circle	Kovari et al (2013), Neamah et al (2014)	2
D06	Connectivity	Pal et al (2015), Kovari et al (2013)	2
D07	Convexity	Kovari et al (2013)	1
D08	Curvature	Bertolini et al (2010), Esteban et al (2012)	2
D09	Curvelet	Guerbai et al (2014), Pal et al (2015)	2
D10	Density	Batista et al (2015), Arunakatha et al (2015), Batista et al (2010), Bertolini et al (2010), Pushpalatha et al (2014), Kumar et al (2012)	6
D11	Distance	Kudlacik et al (2012), Arunakatha et al (2015), Ohyama et al (2015), Serdouk et a (2015)	4
D12	Cosine Transform (DCT)	Daramola e Ibiyemi (2010), Bharathi y Shekar (2014), Bharathi et al (2014), Ribeiro et al (2014)	4
D13	Radon Transform (DRT)	Abdelrahman et al (2013), Swanepoel et al (2013)	2
D14	Wavelet Transformation (DWT)	Shirdhonkar e Kokare (2012)	1
D15	Energy	Velez et al (2009), Guerbai et al (2014), Jaiwal et al (2014)	3
D16	Entropy	Jaiwal et al (2014), Kumar et al (2012), Arathi et al (2014)	3
D17	Euler number	Pushpalatha et al (2013, 2014)	2
D18	Fast Fourier Transformation (FFT)	Wen et al (2009),	1
D19	Geodetic distance	Abdoli et al (2014)	1
D20	Geometric center	Jane et al (2010), Angadi et al (2014)	2
D21	Gravity center	Neamath et (2014), Das et al (2009), Batista et al (2010)	3
D22	Gray variation	Ohyama et al (2015), Xu et al (2014),	2

D23	Gradient	Xu et al (2014), Ohyama et al (2015), Serdouk et a (2015), Pham et al (2015), Zhang (2010)	6
D24	Gabor transformation	Angadi et al (2014)	1
D25	Histogram	Zhang (2010), Pirlo et al (2013a), Shekar et al (2015), Serdouk et al (2014), Pham et al (2015)	5
D26	Hu moment	Radhika et al (2010, 2011)	2
D27	Intersections	Pushpalatha et al (2014), Randhawa et al (2013), Pirlo et al (2013a), Kudlacik et al (2012), Rico et al (2012)	5
D28	Local binary pattern (LBP)	Xu et al (2014), Serdouk et a (2015), Serdouk et al (2014)	3
D29	Local direction pattern (LDP)	Xu et al (2014), Serdouk et a (2015)	2
D30	Local derivative pattern (LDerivP)	Abdoli et al (014)	1
D31	Long-Run Feature (LRF)	Serdouk et al (2014)	1
D32	Loop	Kovari et al (2013), Pal et al (2015)	2
D33	Mass	Jena et al (2010), Jaiwal et al (2014), Pirlo et al (2013a), Pushpalatha et al (2014), Rico et al (2012)	5
D34	Maximal (vertical and horizontal)	Jaiwal et al (2014)	1
D35	Morphology	Kekre et al (2010), Kumar et al (2010), Shekar et al (2015), Rico et al (2012)	4
D36	Optical flow	Pirlo et al (2013b)	1
D37	Orientation	Batista et al (2010), Bertolini et al (2010)	2
D38	Perimeter	Kovari et al (2013), Velez et al (2009)	2
D39	Perimeter ring	Wen et al (2009)	1
D40	Pixel distribution	Pushpalatha et al (2014), Bertolini et al (2010), Arathi et al (2014), Bharathi y Shear (2014), Rico et al (2012)	5
D41	Probability	Zois et al (2015)	1
D42	Ratio (high/wide)	Jaiwal et al (2014), Das et al (2009)	2
D43	Rope moment	Kumar and Puhan (2014)	1
D44	Standard deviation	Angadi et al (2014), Guerbai et al (2014)	2
D45	Stroke curvature	Batista et al (2010), Bertolini et al (2010)	2
D46	Texture	Abdoli et al (2014)	1

Significa que los descriptores menos usados aún no están siendo bien explorados o dejaron de ser usados o no dan buenos resultados. Pero según los trabajos analizados, una buena combinación de esos descriptores, considerando la formación de las firmas y la experiencia de los FDEs, pueden dar buenos resultados en verificación de firmas off-line.

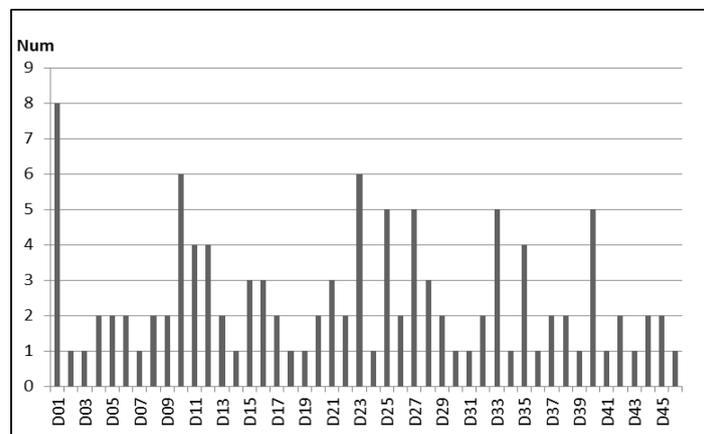


Figure 3: Descriptors used in features.

4.2. Clasificadores

Los métodos de clasificación mayormente usados en los trabajos revisados son basados en SVM, superando largamente a los otros métodos, tal como ilustra Table 4. La mayoría de esos trabajos usan SVM y variantes, como OC-SVM por Guerbai et al (2014), SVM multi-clase (Zhang et al 2010). También SVM combinando en secuencia con otros métodos, como el trabajo de Pushpalatha et al (2014) que SVM complementa al HMM, Serdouk et al (2015) que formulan AIRS con kernel SVM sustituyendo K-NN. El método de comparación por similitud, o DTW, es el segundo método más utilizado después del SVM. Similitud usa métrica de distancia de Camberra (Shirdhonkar e Kokare, 2012), distancia Euclidean (Jena et al, 2010), cosenos (Pirlo et al, 2013), entre otros de templates. HMM es el tercero de los más

usados en los trabajos de clasificación, como Batista et al (2010) y Daramola & Ibiyemi (2010) que usan varios HMMs en el sistema de clasificación, como también HMMs combinados con otros métodos (Batista et al, 2015). Se observa que los clasificadores basados en AIS, ANN, DT, NN, Fuzzy y statistical también son preferidos menos que HMM pero más que los métodos basados en grafos, correlation, minimal patch, PCA, stability y variance.

El método global DCT aplicado en sub-bloques por Daramola e Ibiyemi (2010) es interesante, porque lo global aplica en regiones menores. El enfoque de espacio y tiempo de la firma, bajo supuesto que se realiza de izquierda a derecha, para uso de HMM como clasificador no sería universal, porque no toda firma termina en el lado derecho, pues existen firmas con trazos terminando en el extremo izquierdo. La extracción de atributos basado en el par de energía y desviación estándar de los coeficientes wavelets en multiresolución, propuesta de Shirdhonkar e Kokare (2012), es uno de los nuevos enfoques globales en esta materia. Pero el proceso de recuperación secuencial en clasificación lo hace pesado, siendo cuadrático para grandes cantidades de firmas. El método propuesto de Pirlo e Impedovo (2013) teóricamente puede funcionar eficientemente, en el caso que las variaciones de las firmas del mismo firmante no sean morfológicas, por ejemplo cuando el número de las oscilaciones periódicas no varían. Pero el sistema puede funcionar solo para un número limitado de firmas utilizadas para entrenamiento de ADT, mientras que las firmas no incluidas en el entrenamiento tendrían menos chance de ser considerada genuina. El entrenamiento de un clasificador es más eficiente cuando se usan varias firmas de la misma persona (Wen et al, 2009). Caso de pocas firmas se puede adaptar por estimación, por ejemplo en el caso específico de Fourier, por desviación.

El principio de que características de los atributos sea una distribución Gaussiana, dado por Kovari et al (2013), no siempre es correcto porque no hay garantía de encontrar tantos atributos sin violar el principio de que sean casi independientes entre sí. La propuesta de Pal et al (2015) de establecer un umbral por cada persona en lugar de uno único, parece interesante. El problema es que no llega a probar el concepto en su totalidad, y no hay evidencia de como sería para grande masa de personas. La idea de Berolini et al (2010) de usar un solo clasificador entrenado con muy pocas muestras de firmas verdaderas por persona, por ejemplo 4 firmas que ellos lo usan, que si bien no es “mucho”, tampoco es “poco”. La muestra usada es un agregador simple por fusión de reglas (mayoría de votos?). Es posible que se mejore el error si se dan pesos variables a cada clasificador de la muestra.

Table 4: Classifier methods.

Cod	Classifier method	References	Num
C01	Artificial Inmune System (AIS)	Serdouk et al (2014), Serdouk et al (2015)	2
C02	Artificial Neural Network (ANN)	Kumar et al (2012), Randhika et al (2011)	2
C03	Decision trees (DT)	Arathi et al (2014), Pirlo et al (2013)	2
C04	Correlation	Bathatacharyya et al (2008)	1
C05	Graphs	Neamah et al (2014)	1
C06	Fuzzy Set	Kudlacik et al (012), Velez et al (2009)	2
C07	HMM	Batista et al (2010, 2015), Daramola & Ibiyemi (2010), Pushpalatha et al (2014),	4
C08	Nearest Neighbors (NN)	Rico et al (2012), Abdelrahman et el (2013)	2
C09	Principal Component Analysis (PCA)	Arunalatha et al (2015)	1
C10	Similarity Distance	Shirdhonker & Kokare (2012), Jena et al (2010), Rico et al (2012), Abdelrahman et el (2013), Kekre et al (2010), Pal et al (2015), Velez et al (2009), Pirlo et al (2013)	8
C11	Stability	Pirlo et al (2013)	1
C12	Statistical	Jena et al (2010), Kovari et al (2013)	2
C13	Suport Vector Machine (SVM)	Bharathi & Shekar (2014), Kumar & Puhan (2014), Ohyama et al (2015), Swanepoel et al (2013), Sedouk et al (2015), Radhika et al (2010), Zhang (2010), Kumar et a (2010), Bertolini et al (2010), Tselios et al (2012), Randhawa et al (2012), Ribeiro et al (2014), Damola & Ibiyemi (2010), Pushpalatha et al (2014), Guerbai et al (2015)	15

Se observa un defecto en esta parte, para complementar la respuesta de la cuestión Q3, que los autores escogen los métodos de clasificación solo pensando como operadores de entidades

numéricas. Son muy pocos, al decidir usar HMM, justifican que las firmas son realizadas de izquierda para derecha que el método atendería bien.

4.3. Off-line signature datasets

La diversidad de los bancos de datos de firmas usados para las validaciones de los métodos de verificación de firmas offline en la literatura, hacen difícil la comparación efectiva de los diferentes métodos propuestos (Tselios et al, 2012; Alsulaiman et al, 2013). Es lógico pensar que el conjunto de las firmas de un database no satisfaga las condiciones de otros databases. Esa variación de tipos de datos puede favorecer en la validación a ciertos métodos y a otros no. Otro elemento confuso para comparación es el uso de varias métricas, como EER y AER como media de FAR y FRR. En algunos casos la validación solo ofrece el porcentaje de precisión, como Abdelrahman & Abdallah (2013) que obtienen 80% precisión con su classificador k-NN y usando MCYT100 database. El uso de varios parámetros de medición de los métodos impide determinar cual método del universo todo esta siendo mas preciso; en todo caso, la precisión estará limitada a pocos trabajos de un universo mayor.

Los databases usados en los trabajos analizados son GPDS, referenciados a sus datasets GPDS300, GPDS160 y GPDS960GraySignatures. Otros databases regularmente utilizados en verificación de firmas son CEDAR, MUKOS, SVC2004, SigComp2011, MCYT100, Brazilian database; Hindi database; GAVAB, Greek database, Dolfing data set, Tobacco-800 y databases privados, tal como se muestra en la Tabla 5. En la misma tabla, se observa que mayor número de trabajos usan GPDS300, seguidos por trabajos usando databases privados. El CEDAR es tercer preferido por los trabajos. La preferencia por dataset privados posiblemente justifique que no todos los databases estan disponibles em forma gratuita o simplicidade de acceso.

Table 5: Signature data bases used.

Data bases	Users	Num.
GPDS300	Kovari & Charaf (2013); Kumar et al (2012); Serdouk et al (2015a, 2015b), Zois et al (2014), Tselios et al (2012), Serdouk et al (2015a), Ribeiro et al (2014), Pirlo et al (2013a), Pirlo et al (2013b), Zhang (2010); Neamah et al (2014); Guerbai et al (2014, 2015); Eskander et al (2013)	15
GPDS160	Batista et al (2015); Bharathi & Shekar (2014a, 2014b), Shekar et al (2015)	4
GPDS960	Pushpalatha y Gautan (2014); Ferrer et al (2012); Arunalatha et al (2015); Xu et al (2014); Abdoli & Hajati (2014), Xu et al (2014)	6
CEDAR	Shekar et al (2015); Serdouk et al (2014); Serdouk et al (2015a, 2015b); Guerbai et al (2014, 2015); Kumar & Puhan (2014); Kumar et al (2010); Kumar et al (2012); Bharathi & Shekar (2014a, 2014b)	11
MCYT100	Wen et al (2009), Rico e Iñesta (2012); Radhika et al (2010); Radhika et al (2011); Ferrer et al (2012); Abdelrahman & Abdallah (2013)	6
SVC2004	Porwik et al (2015); Kovari & Charaf (2013); Kudlacik et al (2012);	3
SigComp2011	Ohyama et al (2015); Pham et al (2015);	2
MUKOS	Shekar et al (2015); Bharathi & Shekar (2014a, 2014b))	3
Greek database	Zois et al (2014)	1
Tobacco-800	Shu et al (2009)	1
Hidu Sig. Database	Pal et al (2015); Pal et al (2013)	2
Brazilian Database	Batista et al (2010, 2015); Eskander et al (2013)	3
GAVAB	Esteban et al (2012);	1
Dolfing data set	Swanepoel & Coetzer (2013);	1
Private Database	Veles et al (2009), Tselios et al (2012), Rivard et al (2013), Randhawa et al (2012); Ghosh et al (2014); Jaiswal & Kasetwar (2014); Jena et al (2010); Kekre et al (2010); Bertolini et al (2010); Daramola y Ibiyemi (2010); Elrajubi & El-Feghi (2015); Arathi et al (2014); Kovari & Charaf (2013); Angadi et al (2014),	14

Son observados, en la generación de los diferentes database, el número de personas que dejan sus firmas (firmantes), número de firmas genuinas y falsas por cada persona. En ciertos casos, en particular en las privadas, son explicitados los periodos de recolección de las firmas. En la Tabla 6 se resume las características que constituyen los databases conocidas regulamente institucionalizados.

Table 6: Databases and theirs attribute numbers.

DATASET	N. Signers	N. Genuine signatures	N. skilled forgery	Total signatures
GPDS (Univ Las Palmas de Gran Canaria)	300	24	30	16200
GPDS-160	160	24	30	8640
GPDS960 GraySignature	75	24	30	4125
CEDAR (SUNY, Buffalo)	55	24	24	2640
Brazilian	168	80	20	7920
MUKOS (Mangalore University – Kannada)	30	30	15	1350
MCYT-100	75	15	15	2250
SigComp2011	64	773	1071	1844
Hindi Database	100	24	30	5400
SVC2004	40	20	20	1600
Dolfing data set	51	30	60	4530

Los diferentes databases utilizados por los diferentes trabajos para validar sus métodos propuestos son resumidos en la Tabla 7 en función de las métricas AER y EER. Trabajos que solo muestras FAR y FRR son calculados para AER con objetivo de uniformizar los valores de medición. En seguida son tratados los diferentes databases colocando en evidencia los valores de AER y EER mínimo y máximos de los respectivos trabajos.

Se observa que GPDS, de la Universidad las Palmas de Gran Canaria, posee 16200 firmas. Es comun ver en diferentes trabajos referenciando a grupos de este database, como el GPDS300, GPDS160. También GPDS960GraySignatures con 4125 firmas gray. Trabajos usando explícitamente GPDS300, como se observa en la Tabla 7, obtienen un AER mínimo de 2.17 por Ribeiro et al. (2014) y los trabajos que usan EER reportan un mínimo de 2.00 por Telios et al. (2012). Trabajos con GPDS160 registran AER un mínimo de 7.37 (Bharathi et al., 2014b) y un EER de 17.00 (Ferrer et al., 2012); mientras trabajos con GPDS960 registran un menor ERR de 11.00 (Xu et al., 2014) y un AER de 7.96 (Pushpalatha & Gautan, 2014). CEDAR database (SUNY, Buffalo), que posee 2649 firmas, es usado por los trabajos que reportan un AER menor 6.02 (Kumar & Puhan, 2014) y un EER de 3.56 (Serdouk et al., 2014). El MCYT100 signature CORPUS contains 2250 signatures, con AER de 7.65 (Bharathi & Shekar, 2014b) y un EER de 6.30 (Rico & Iñesta, 2012).

Los trabajos usando databases menos populares, como MUKOS (Mangalore University – Kannada) con 1350 firmas, reportan un AER menor 4.73 (Bharathi & Shekar, 2014a); con SVC2004, de 1600 firmas y creado para competencia de verificación de firmas 2004, se reportan AER mínimo de 1.81 (Porwik et al., 2015); con SigComp2011 (Signature Verification Competition for Online and Offline Skilled Forgeries 2011), que incluye 1844 firmas (entre chinos y holandeses), se tiene un AER de 12.20 (Pham et al., 2015) y EER de 4.29 (Ohama et al., 2015); con Brazilian database con 7920 firmas se obtienen un AER menor 7.34 (Eskander et al., 2013); Con Tobacco800 se obtiene 0.36 de AER (Shu et al., 2009); con Dolfing 3.78 de EER (Swanepoel & Coetzer, 2013); con Hindi obtiene un AER menor 8.15 (Pal et al., 2015); Finalmente, con Greek, Zois et al. (2014) obtienen un EER de 3.34.

Los databases privados, recolectados localmente en periodos y locales diferentes, también son regularmente usados por diferentes trabajos. El database usado por Daramola y Ibiyemi (2010) colectada de 250 students at Covenant University Ota, Nigeria, cada estudiante contribuye con 7 firmas genuinas; Elrajubi & El-Feghi (2015) generado por 40 personas diferentes con firmas 16 genuinas y 6 falsas de cada tipo cada uno, dando un total de 880 firmas; Arathi et al (2014) obtienen 1287 firmas combinadas entre genuinas y falsas; Angadi et al (2014) definen su database de 400 firmas registradas por 40 personas. Ghosh et al (2014) colectan 2 firmas de 100 diferentes personas; Jaiswal & Kasetwar (2014) usan database de 510 firmas, siendo 290 genuinas y 220 falsas, generados por 8 diferentes firmantes. Riverd et al (2013) generados por 168 firmantes con 40 firmas por cada uno y 10 firmas por vez. Telios et al (2012) generan Corpus-1 firmadas 105 genuinas y 21 falsas por cada uno de las 69 personas, pero también combinan con otros 2 bancos de datos como GPDS300 y otro Corpus-3 de origen privado usados por Huang & Yan (2002). Otros como Kekre et al (2010), Jena et al (2010), Bertolini et al (2010), Randhawa et al (2012) y Veles et al (2009) usan sus propios database sem especificar las características de total de las firmas explicitadas en las literaturas revisadas. En la Tabla 7,

son registrados solo algunos de ellos que informan métricas relacionados con AER y EER. En ese contexto, se observan trabajos con un AER mínimo de 3.00 (Randhawa et al., 2012), como también trabajos con EER mínimo de 2.00 (telios et al., 2012).

Table 7: Metrics used for several works and databases.

FAR	FRR	AER	EER	REFER	FAR	FRR	AER	EER	REFER
GPDS300					MCYT100				
3.87	3.98	3.93		Zhang et al. (2010)				17.00	Ferrer et al. (2012)
			15.95	Guerbai et al. (2014, 2015)				6.30	Rico & Iñesta (2012)
			12.50	Eskander et al. (2013)				15.30	Radhika et al. (2011)
			32.20	Kovari & Charaf (2013)	10.00	8.00	9.00		Radhika et al. (2010)
8.33	13.76	11.05		Kumar et al. (2012)	10.0	7.00	8.50		Radhika et al. (2011)
2.28	2.05	2.17		Ribeiro et al. (2012)	7.56	7.65	7.65		Bharathi & Shekar (2014b)
7.31	2.87	5.09		Zois et al. (2014)	MUKOS				
			10.00	Kovari & Charaf (2013)	7.08	6.58	6.83		Shekar et al. (2015)
			6.48	Sedouk et al. (2015b)	4.00	5.45	4.73		Bharathi & Shekar (2014a)
			13.19	Sedouk et al. (2015a)	4.00	5.50	4.75		Bharathi & Shekar (2014b)
			9.60	Pirlo et al. (2013a)	SVC2004				
			5.40	Pirlo et al. (2013b)				26.60	Kovari & Charaf (2013)
			2.00	Telios et al. (2012)	1.52	12.16	6.84		Kudlacik et al. (2012)
GPDS160					1.08	2.53	1.81		Porwik et al. (201)
			17.00	Ferrer et al. (2012)	SigComp2011				
8.94	9.24	9.09		Shekar et al. (2015)				4.29	Ohama et al. (2015)
7.93	8.04	7.99		Bharathi & Shekar (2014a)	12.05	12.35	12.20		Pham et al. (2015)
6.83	7.90	7.37		Bharathi & Shekar (2014b)	Brasilian				
GPDS960								7.34	Eskander et al. (2013)
			11.00	Xu et al. (2014)				7.79	Batista et al. (2010, 2015)
			13.90	Abdoli & Hajati (2014)	Tobacco-800				
			17.00	Arunalatha et al. (2015)	0.21	0.50	0.36		Shu et al. (2009)
11.02	4.90	7.96		Pushpalatha & Gautan (2014)	Dolfing				
CEDAR								3.78	Swanepoel&Coetzer (2013)
			7.83	Guerbai et al. (2014, 2015)	Hindi				
			3.56	Serdouk et al (2014)	7.12	11.50	9.31		Pal et al. (2013)
			8.46	Serdouk et al (2015a)	2.50	13.80	8.15		Pal et al. (2015)
			6.99	Serdouk et al (2015b)	Greek signers				
11.23	12.19	11.71		Kumar et al. (2010)	3.81	2.87	3.34		Zois et al. (2014)
8.33	12.76	10.55		Kumar et al. (2012)	Privates				
5.68	6.36	6.02		Kumar & Puhan (2014)				5.80	Arathi et al. (2014)
8.86	9.37	9.12		Shekar et al. (2015)				11.67	Elrajubi & El-Feghi (2015)
5.75	6.32	6.04		Bharati & Shekar (2014a)	5.80	5.00	5.40		Jaiswal & Kasetwar (2014)
					6.56	3.70	5.13		Kekre et al. (2010)
					2.00	4.00	3.00		Randhawa et al. (2012)
								12.50	Veles et al. (2009)
								2.00	Telios et al. (2012)

En conjunto, observando la Tabla 7, si todos los database fuesen uniformes, se tendrían trabajos con menores AER, usando database preferidos, 2.17 por Ribeiro et al. (2012) y definido por EER menor 2.00 por Telios et al. (2012); similarmente con los datasets privados se diría que menor AER 3.00 de Randhawa et al. (2012) y con menor EER 2.00 de Telios et al. (2012).

Para explicitar la respuesta a la cuestión Q4, la variedad de dataset, tal como mencionado en este seccion y ilustradas por Table 5 y Table 6, no necesariamete son universales y no garanten resultados confiables. Lo mismo se puede observar con cuestión de precisión (Q5), ilustrado por la Table 7, que métodos y datasets, con métricas variadas, es difícil convergir a un buen resultado.

5. Conclusiones

El análisis de los trabajos en verificación de las firmas off-line –realizado en las dimensiones de abstracción de las firmas en atributos de descriptores, la clasificación de tipo de firmas, confiabilidad de datasets de firmas usados y parámetros de precisión usados–, en base a las cinco cuestiones establecidas en la sección método- demandó un estudio crítico de los método

abordados en cada una de las fases de ASVS abordados por los diferentes autores en la literatura.

Fueron utilizados más de 46 tipos e descriptores, resumidos en la Table 3, para formular las características representadas en forma de vectores, grafos y árboles. Consideramos que la esencia de identificación de las firmas, como lo respaldan los peritos expertos en grafotecnia, está en forma como se identifican los detalle de las firmas, es decir, está en los forma de capturar los atributos de las firmas. Ese proceso debe ser realizado, sobre todo, previa justificación de lo que representa la generación de la firma. También, considerando que todas las firmas genuinas de una misma persona, generadas en diferentes instantes, son también diferentes, pero en esencia se trata del atributo biométrico de la misma persona. Encontramos que alto número de métodos desarrollados no consideran ese principio, hicieron solamente pensando en obtener elementos numéricos para definir los vectores de atributos. Sin embargo, son muy pocos trabajos que consideran las variaciones intra-clases asociadas a las variaciones de las firmas del mismo firmante, como también la tendencia de que esas firmas tengan una distribución normal, pero deben ser demostradas.

La clasificación, donde se identifica si la firma es genuina o falsificada, está implementada, generalmente, en la forma de operador numérico, basado en uno de los tantos métodos analizados (Table 4), para una respuesta adecuada. Consideramos que la elección de los métodos, o la propuesta de una arquitectura de clasificación basada en híbrido de métodos, sea adecuada a la esencia de generación de las firmas y los detalles inherentes a las firmas genuinas o falsas. También, en este aspecto, son pocos trabajos que consideran la esencia y detalles de las firmas, así como el enfoque seguido por peritos en grafotecnia.

También, en lo que respecta la confiabilidad y precisión del verificador eficiente de firmas, constatamos que existen una variedad de datasets usadas, unos generados por centros de pesquias y otros generados por los mismos autores de los trabajos (Table 5). Eso genera una desconfianza en la comunidad investigadores en la materia de reconocedores de firmas. La mayoría de los datasets de referencia, de centros de investigación, son indisponibles para externos y son pagos. Eso, posiblemente, obliga a los investigadores del mundo generar sus propios datasets de validación. Otro factor es la existencia de las métricas usadas para la validación de precisión, heterogeneidad por cierto (Table 7), que impide la confiabilidad en los supuestos mejores métodos de verificación de las firmas off-line.

Referencias

- Abdelrahman, A.; Abdallah, A. (2013), K-Nearest Neighbor Classifier for Signature Verification System, IEEE - International Conference on Computing, Electrical and Electronic Engineering (ICCEEE), pp 58-62.
- Abdoli, S.; Hajati, F. (2014), Offline Signature Verification using Geodesic Derivative Pattern. IEEE 22nd Iranian Conference on Electrical Engineering (ICEE), pp 1018-1023.
- Abroug, I. Amara, N. (2014), Off-line signature verification systems: recent advances. International Image Processing Applications and Systems Conference, pp. 1-6.
- Alsulaiman, F.; Sakr, N.; Valdes, J.; Saddik, A. (2013), Identity Verification Based on Handwritten Signatures with Haptic Information Using Genetic Programming. ACM Trans. Multimedia Computing, Communications and Applications, vol 9, No 2, pp. 11:1–11:21.
- Angadi, S.; Gour, S. Bajantri, G. (2014), Offline signature recognition system using radon transform. IEEE International Conference on Signal and Image Processing, pp 56-61.
- Arathi, M.; Govardhan, A. (2014), An Efficient Offline Signature Verification System, International Journal of Machine Learning and Computing, 4 (6), pp. 533-537.
- Arunalatha, J.; Prashanth, C.; Tejaswi, V.; Shaila, K.; Raja, K.; Anvekar, D.; Venugopal, K.; Iyengar, S.; Patnaik, L. (2015), PCVOS: principal component variances based off-line signature

verification. IEEE International Conference on Recent Trends in Information Systems (ReTIS), pp 19-199.

Batista, L.; Granger, E.; Abourin, R. (2010), Improving performance of HMM-based off-line signature verification systems through a multi-hypothesis approach. International Journal on Document Analysis and Recognition, 13, pp. 33-47.

Batista, L.; Granger, E.; Sabourin, R. Dynamic selection of generative–discriminative ensembles for off-line signature verification, Pattern Recognition, issue 45, 2015, pp 1326-1340.

Bertolini, D.; Oliveira, L.; Justino, E.; Sabourin, R. (2010), Reducing forgeries in writer-independent off-line signature verification through ensemble of classifiers. Pattern Recognition, 43, pp. 387-396.

Bharathi, B.K.; Shekar, B. (2014a), Discriminative DCT: An Efficient and Accurate Approach for Off-line Signature. IEEE - Fifth International Conference on Signal and Image Processing, pp. 179-184.

Bharathi, R.; Shekar, B. (2014b), Discriminative DCT-MLP based approach for off-line signature verification. IEEE International Conference on Advances in Computing, Communications and Informatics, pp 2309-2315.

Daramola, S.A.; Ibiyemi, T.S. (2010), Offline Signature Recognition Using Hidden Markov Model. International Journal of Computer Applications, V. 10, N 2, pp 17-22

Das, M.; Dulger, L. (2009), Signature verification (SV) toolbox: Application of PSO-NN. Engineering Applications of Artificial Intelligence, vol 22, pp. 688-694.

Deore, M.; Handore, S. (2015), A survey on offline signature recognition and verification schemes, IEEE: International Conference on Industrial Instrumentation and Control (ICIC), pp 165-169.

Elrajubi, O.M.; El-Feghi, I.S. (2015), Angle Features Extraction of Handwritten Signatures, IEEE International Conference on Computer Vision and Image Analysis Applications (ICCVIA 2015), pp-1-4.

Eskander, G.S.; Sabourin, R.; Granger, E. (2013), Hybrid writer-independent–writer-dependent offline signature verification system. IET Biometrics, vol 2, No 4, pp. 169-181.

Esteban, J.; Vélez, J.; Sanchez, A. (2012), Off-line handwritten signature detection by analysis of evidence accumulation. International Journal on Document Analysis and Recognition, (15), pp. 359-368.

Fàbregas, J.; Faundez-Zanuy, M. (2009), Biometric recognition performing in a bioinspired system. Cognitive Computation, vol 1, iss 3, pp. 257-267.

Ferrer, M.A.; Vargas, J.F.; Morales, A.; Ordoñez, A. (2012), Robustness of Offline Signature Verification Based on Gray Level Features. IEEE Transactions on Information Forensics and Security, v7, n3, pp. 966-977

Ghosh, B. R.; Banerjee, S.; Dey, S. Ganguli, S. Sarkar, S. (2014), Off-Line Signature Verification System Using Weighted Complete Bipartite Graph, IEEE 2nd International Conference on Business and Information Management (ICBIM), pp. 109-113.

Guerbai, Y.; Chibani, Y.; Hadjadji, B. (2015), The effective use of the one-class SVM classifier for handwritten signature verification based on writer-independent parameters, Pattern Recognition, 48, pp. 103-113.

Guerbai, Y.; Chibani, Y.; Hadjadji, B. (2014), Writer-Independent Handwritten Signature Verification based on One-Class SVM, IEEE: International Joint Conference on Neural Networks (IJCNN), pp 327-331.

Huang, K., Yan, H. (2002), Off-line signature verification using structural feature correspondence, Pattern Recognit., 35, pp. 2467–2477

- Jaiswal, S.G.; Kasetwar, A.R. (2014), Off-line signature verification using global and local features with neural networks. IEEE International Conference on Advanced Communication Control and Computing Technologies, pp. 1525-1531.
- Jena, D.; Majhi, B. Panigrahy, S.K.; Jena, S.K. (2010), Improved Offline Signature Verification Scheme Using Feature Point Extraction Method. IEEE International Conference on Cognitive Informatics (ICCI'10), pp 475-480.
- Kitchenham, B.; Brereton, O.p.; Budgen, A.B.; Turner, M.T.; Bailey, J.; Linkman, S. (2009), Systematic literature reviews in software engineering – A systematic literature review. Information and Software Technology, 51, pp 7-15.
- Kekre, H.B.; Bharadi, V.A.; Gupta, S.; Ambardekar, A.A.; Kullami, V.B. (2010), Off-line signature recognition using morphological pixel variance analysis. International Conference and Workshop on Emerging Trends in Technology (ICWET 2010), pp. 3-10.
- Kovari, B.; Charaf, H. (2013), A study on the consistency and significance of local features in off-line signature verification, Pattern Recognition Letters, 34, pp. 247-255.
- Kudlacik, P.; Porwik, P. (2012), A new approach to signature recognition using the fuzzy method. Pattern Analysis Applications, vol 17, pp. 451-463.
- Kumar, M.M.; Puhan, N.B. (2014), Off-line signature verification: upper and lower envelope shape analysis using chord moments. IET Biometrics, 3 (4), pp. 347-354.
- Kumar, R.; Kundu, L.; Shama, J.D.; Chanda, Bh. (2010), A writer-independent off-line signature verification system based on signature morphology. Proceedings of the International Conference on Intelligent Interactive Technologies and Multimedia, pp. 261-265.
- Kumar, R.; Shrama, J.D.; Chanda, B. (2012), Writer-independent off-line signature verification using surroundedness feature. Pattern Recognition Letters, 33, pp. 301-308
- Neamah, K.; Mohamad, D.; Saba, T.; Rehman, A. (2014), Discriminative Features Mining for Offline Handwritten Signature Verification, Springer: eD Research Center, 5:2, pp 1-6.
- Ohyama, W.; Ogi, Y. Wakabayashi, T. Kimura, F. (2015), Multilingual Signature – Verification by Generalized Combined Segmentation Verification, IEEE 13th International Conference on Document Analysis and Recognition (ICDAR2015), pp 811-815.
- Pal, S.; Alaei, A. Pal, U.; Blumenstein, M. (2015), Interval-value symbolic representation based method for off-line signature verification. IEEE International Joint Conference on Neural Networks, pp. 1-6.
- Pal, S.; Pal, U.; Blumenstein, M. (2013), Off-line verification technique for Hindi signature. IET Biometrics, vol 2, No 4, pp. 182-190.
- Pal, U.; Jayadevan, R.; Sharma, N. (2012), Handwriting recognition in Indian regional scripts: a survey of offline techniques. ACM Transactions on Asian Language Information Processing, Vol 11, No. 1, Art. 1, pp 1:1-1:35.
- Pham, T-A.; Le, H-H.; Do, N-T. (2015), Offline handwritten signature verification using local and global features, Springer: Ann Math Artif Intell, v 75, pp 231-247.
- Pirlo, G. Impedovo, D. (2013a), Cosine similarity for analysis and verification of static signatures, IET Biometrics, v.2, Iss.4, pp. 151-158.
- Pirlo, G. Impedovo, D. (2013b), Verifications on Static Signature by Optical Flow Analysis, IEEE Transactions on Human-Machine Systems, v. 43, Iss.5, pp. 499-504.
- Porwik, P.; Doroz, R.; Orczyk, T. (2015), The k-NN classifier and self-adaptive Hotelling data reduction technique in handwritten signatures recognition. Pattern Analysis Applications, vol 18, pp. 983-1001.

- Pushpalatha, K.N; Gautan, A.K. (2014), Offline Signature Verification Using Spatial Domain Feature Sets and Support Vector Machine, *International Journal of Emerging Technology and Advanced Engineering*, V. 4. Issue 7, pp 544-551.
- Radhika, K.R.; Venkatesha, M.K.; Sekhar, G.N. (2010), Off-line Signature Authentication based on Moment Invariants Using Support Vector Machine, *Journal of Computer Science*, 6(3), pp. 305-311
- Radhika, K.R.; Venkatesha, M.K.; Sekhar, G.N. (2011), Signature authentication base on subpattern analysis, *Applied Soft Computing*, 11, pp. 3218-3228
- Randhawa, M.K.; Sharma, A.K.; Sharma, R.K. (2012), Off-line signature verification based on hu's moment invariants and zone features using support vector machine. *International Journal of Latest Trends in Engineering and Technology (IJLTET)*, v. 1, Iss 3, pp. 16-23.
- Ribeiro, B.; Lopes, N.; Gonçalves, J. (2014), Signature identification via efficient feature selection and GPU-based SVM classifier. *IEEE International Joint Conference on Neural Networks (IJCNN)*, pp 1138-1145.
- Rico, J.R.; Iñesta, J.M. (2012), Confidence voting method ensemble applied to off-line signature verification. *Pattern Anal Application*, 35, pp. 113-120.
- Rivard, D.; Granger, E.; Sabourin, R. (2013), Multi-feature extraction and selection in writer-independent off-line signature verification, *Springer: IJDAR*, v16, pp. 83-103.
- Sedouk, Y.; Nemmour, H.; Chibani, Y. (2015a), New gradient features for off-line handwritten signature verification. *IEEE International Symposium on Inovations in Intelligent Systems and Applications (INISTA)*.
- Serdouk, Y.; Nemmour, H.; Chibani, Y. (2015b), An Improved Artificial Immune Recognition System for Off-line Handwritten Signature Verification, *IEEE 13th International Conference on Document Analysis and Recognition (ICDAR 2015)*, pp 196-200.
- Serdouk, Y.; Nemmour, H.; Chibani, Y. (2014), Topological and textural features for off-line signature verification based on Artificial Immune Algorithm. *IEEE International Conference of Sof Computing and Pattern Recognition*, pp. 118-122.
- Shekar, B.H.; Bharathi, R.K.; Kittler, J.; Vizilter, Y.V; Mestestkiv, L. (2015), Grid Structured Morphological Pattern Spectrum for Off-line Signature Verification, *IEEE: International Conference on Biometric (ICB2015)*, pp. 430-435.
- Shirdhonkar, M.S.; Kokare, M.B. (2012), Off-line Handwritten Signature Retrieval. *ACM: CUBE*, pp 290-293.
- Swanepoel, J.; Coetzer, J. (2013), A robest dissimilarity representation for writer-independent signature modelling, *IET Biometrics*, V.2, Iss. 4, pp. 159-168.
- Tselios, K.; Zois, E.; Siores, E.; Nassiopoulos, A.; Economou, G. (2012), Grid-based feature distributions for off-line signature verification. *IET Biometrics*, v 1(1), pp. 72-81.
- Velez, J.; Sanchez, A.; Moreno, B.; Esteban, J. (2009), Fuzzy shape-memory snakes for the automatic off-line signature verification problem. *Fuzzy Sets and Systems*, 160, pp 182-197.
- Wen, J.; Fang, B.; Tang, Y.; Zhang, T. (2009), Model-based signature verification with rotation invariant features, *Pattern Recognition*, 42, pp. 1458-1466.
- Xu, B.; Lin, D.; Wang, L.; Chao, H.; Li, W.; Liao, Q. (2014), Performance comparison of local directional pattern o local binary pattern in off-line signature verification system. *IEEE International Congress on Image and Signal Processing*, pp. 308-312.
- Zhang, B. (2010), Off-line signature verification and identification by pyramid histogram of oriented gradients. *International Journal of Intelligent Computing and Cybernetics*, vol 3, Iss 4, pp 611-630.

Zhu, G; Zheng, Y.; Doermann, D.; Jaege, S. (2009), Signature detection and matching for document image retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v 31, n 11, pp 2015-2031.

Zois, E.; Zervas, E.; Barkoula, K.; Economou, G. Fotopoulos, S. (2014), Poset Description of Grid Features and Application to Off-line Signature Verification. *IEEE 14th International Conference on Frontiers in Handwriting Recognition (ICFHR 2014)*, pp 55-60.