

Clasificación Jerárquica de Huellas Dactilares con Selección de Características

Daniel Peralta¹, Isaac Triguero²³, Yvan Saeys²³, Salvador García¹⁴, Jose M. Benitez¹, and Francisco Herrera¹

¹ Departamento de Ciencias de la Computación e Inteligencia Artificial, CITIC-UGR, Universidad de Granada, 18071 Granada, España
{dperalta, salvag1, J.M.Benitez, herrera}@decsai.ugr.es

² VIB Inflammation Research Center, 9052 Zwijnaarde, Bélgica
isaac.triguero@irc.vib-ugent.be, yvan.saeys@ugent.be

³ Department of Respiratory Medicine, Ghent University, 9000 Gante, Bélgica

⁴ Department of Information Systems, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Arabia Saudí

Resumen La identificación mediante huellas dactilares es un problema de gran actualidad debido al número y tamaño creciente de las bases de datos de huellas. Por tanto, se hace necesario el diseño de sistemas que permitan identificar huellas en grandes bases de datos en un tiempo razonable. La clasificación, en combinación con métodos de extracción de características a partir de las imágenes, suele utilizarse para reducir el espacio de búsqueda en la base de datos. Sin embargo, los extractores más precisos suelen rechazar algunas de las huellas. En este trabajo se propone un sistema jerárquico completo para clasificar huellas, donde cada nivel aplica una selección de características sobre la fusión de varios algoritmos de extracción de características. De este modo se elimina el rechazo y se aumenta la tasa de clasificación respecto a los métodos aislados. Los resultados obtenidos sobre una base de datos de huellas de baja calidad confirman el buen rendimiento de la propuesta.

Palabras clave: Clasificación · Huellas dactilares · Biometría · Selección de Características

Agradecimientos Este trabajo ha sido financiado por los proyectos de investigación TIN2011-28488, TIC-6858 y P11-TIC-7765. D. Peralta tiene una beca FPU del Ministerio de Educación y Ciencia (FPU12/04902). I. Triguero tiene una beca postdoctoral BOF en la Universidad de Gante.

1. Introducción

La identificación de personas es uno de los grandes retos en las últimas décadas. De entre todas las técnicas posibles, el uso de huellas dactilares es una de las más extendidas [17]. Una huella dactilar es un patrón de crestas en la yema de un dedo. A partir de la imagen de una huella se pueden aplicar distintos métodos para extraer características, que generalmente se categorizan en tres niveles [17]:

- **Nivel 1 (global):** flujo de crestas (orientaciones y puntos singulares).
- **Nivel 2 (local):** minucias (bifurcaciones y finales de cresta).
- **Nivel 3 (detalle):** poros, anchura de crestas, etc.

Las características de nivel 2 son las más utilizadas para el reconocimiento de huellas [19], tanto en su vertiente de verificación (comparación de dos huellas) como en la de identificación (búsqueda de una huella en una base de datos). El problema de identificación es por definición el más complejo, especialmente cuando la base de datos es grande, dado que el tiempo de búsqueda aumenta de forma lineal respecto al tamaño de la base de datos.

Una solución ampliamente utilizada para paliar este problema es la clasificación [4,5], que suele utilizar características de nivel 1. Consiste en agrupar las huellas en clases, de forma que una huella de entrada solamente se compara con las huellas de la misma clase, reduciendo así el número de comparaciones en la búsqueda (tasa de penetración en la base de datos). La mayor parte de las propuestas actuales se basan en la propuesta de Henry [9], que divide las huellas en cinco tipos de acuerdo con el patrón de las líneas: *Arch*, *Left Loop*, *Right Loop*, *Tented Arch* y *Whorl*. De forma natural las huellas se distribuyen entre estas clases con los porcentajes 3,7%, 33,8%, 31,7%, 2,9% y 27,9% respectivamente.

Algunos métodos de extracción de características rechazan algunas huellas y no extraen las características necesarias para la clasificación, de modo que no se puede asignar ninguna clase a tales huellas [10]. Por tanto, aunque la tasa de error en clasificación suele ser muy buena para estos métodos, la búsqueda no se restringe para esas huellas y la tasa de penetración promedio aumenta. En aplicaciones de identificación es fundamental reducir la tasa de rechazo para así minimizar la penetración de la búsqueda.

Una de las vías para mejorar los resultados de la clasificación es aunar distintas características de las huellas [15,10], dado que el uso de múltiples fuentes de información poco correladas puede resultar en modelos más robustos y precisos. En este contexto los métodos de preprocesamiento [6], y más concretamente de Selección de Características (SC) [7,20], pretenden simplificar los datos de entrada para eliminar información ruidosa o redundante. Técnicas de este tipo, así como otras de reducción de dimensionalidad, ya se han aplicado con éxito sobre algunos aspectos del problema de clasificación de huellas [3].

En este trabajo se propone un clasificador jerárquico de huellas que busca aumentar la precisión y eliminar el rechazo, para así minimizar la tasa de penetración de la búsqueda. El proceso jerárquico utiliza distintas características y clasificadores en función de qué métodos rechazan una huella. Cada clasificador se basa en la fusión y selección de características de distinto tipo, todo ello para aumentar la precisión y limitar o reducir el coste computacional en el consiguiente proceso de identificación.

Esta contribución se estructura como sigue. La Sección 2 describe el problema tratado y las técnicas actuales. La Sección 3 presenta la propuesta de clasificación jerárquica, que se evalúa empíricamente en la Sección 4. Finalmente, la Sección 5 plantea las conclusiones del estudio y las líneas de trabajo futuro.

2. Preliminares

Esta sección presenta el trabajo previo en la literatura científica relacionada con la propuesta. Primero, la Sección 2.1 explica la clasificación de huellas dactilares. A continuación, la Sección 2.2 introduce la selección de características.

2.1. Clasificación de Huellas Dactilares

Un problema de clasificación parte de un conjunto de instancias x_1, \dots, x_n , cada una de las cuales está definida por un conjunto de atributos a_1, \dots, a_D y por una clase $y_i \in \mathbb{C} = \{c_1, \dots, c_p\}$. El objetivo de la clasificación es definir un modelo capaz de determinar la clase a la que pertenece una instancia nueva [4].

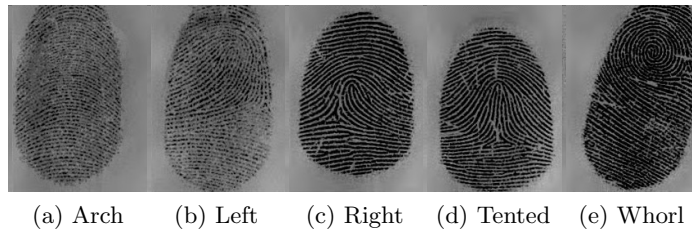


Figura 1. Cinco clases de huellas definidas por Henry [9]

En el ámbito de las huellas dactilares, generalmente se utilizan las cinco clases definidas por Henry [9], basadas en el patrón visual de las crestas como muestra la Figura 1. Existen dos grandes alternativas para la clasificación de huellas dactilares [5]:

- **Clasificación fija:** la clase se determina directamente desde la imagen [12,21].
- **Clasificación tras extracción de características:** la clasificación se realiza en dos pasos. Primero, se extrae una serie de características de la imagen, que se codifican en un vector. Después se utiliza el vector de características para clasificar la imagen. Aunque para este paso algunos autores proponen su propio clasificador, se pueden utilizar clasificadores de propósito general. Algunos métodos de extracción de características se definen en [3,10,16].

Algunas de las características más utilizadas para la clasificación se derivan de las siguientes [17]:

- **Mapa de orientaciones:** matriz donde cada posición contiene la dirección promedio de un bloque de píxeles de la imagen.
- **Puntos singulares:** puntos en torno a los cuales se organizan las crestas. Se dividen en *loops* y *deltas*.
- **Mapa de frecuencias:** matriz donde cada posición contiene el número de crestas que pasan por un bloque de píxeles de la imagen.

- **Filtros:** muchos métodos aplican filtros sobre la imagen de la huella. Uno de los filtros más utilizados es el de Gabor, que realza las crestas en una determinada orientación. La imagen se transforma en un vector extrayendo las desviaciones típicas de los niveles de gris en la imagen filtrada [10].

Algunos métodos de extracción de características combinan varias de las anteriores para elaborar vectores con más información [4]. La Tabla 1 presenta algunos de los métodos más relevantes de la literatura.

Tabla 1. Métodos de extracción de características

Nombre	Características	Tamaño del vector	Rechazo	Ref.
Cappelli	Orientaciones	357	No	[3]
Hong	Puntos singulares, Gábor y Ridge tracing	202	Sí	[10]
Leung	Gábor	816	No	[14]
Li	Puntos singulares y Orientaciones	58	No	[15]
Liu	Puntos singulares	64	No	[16]
Nyongesa	Puntos singulares y Orientaciones	7	No	[18]

2.2. Selección de Características

Los métodos de selección de características (SC) buscan la reducción de un conjunto de datos eliminando atributos irrelevantes o redundantes. El objetivo es obtener un conjunto mínimo de características, de modo que la aplicación de métodos de minería de datos sobre el conjunto reducido produzca resultados muy próximos (o incluso mejores) a los obtenidos con el conjunto de datos completo [7]. Se reduce así el tiempo de procesamiento de los datos y facilitando su interpretación, así como la de los modelos construidos a partir de ellos.

Los métodos de SC se clasifican en tres categorías. Los métodos *wrapper* son métodos de aprendizaje automático cuya función de evaluación incluye el criterio de selección de los atributos [13]. Los métodos de filtrado utilizan métricas relacionadas con los datos, tales como la separabilidad o el *crowding* [7]. Finalmente, en los métodos empotrados se utiliza un algoritmo de minería de datos que incluya algún criterio de ponderación de las características, para extraer del modelo la información necesaria para la selección [20].

La SC, así como otros métodos de reducción de dimensionalidad, se suele utilizar en los métodos basados en orientaciones para reducir el vector de características extraído del mapa [3].

3. Clasificación Jerárquica de Huellas Dactilares con Selección de Características

En este trabajo, se propone un sistema de clasificación de huellas que no rechaza ninguna huella, al tiempo que mantiene una alta tasa de clasificación,

con vistas a reducir la tasa de penetración de la búsqueda en un sistema de identificación. Los diversos problemas que surgen en la clasificación se resuelven mediante distintos componentes:

- **Tasa de clasificación:** el sistema fusiona las características de varios métodos de extracción para mejorar el acierto.
- **Tasa de rechazo:** si alguno de estos métodos de extracción rechaza la huella, el sistema utiliza una combinación de métodos diferente que no tenga rechazo.
- **Crecimiento del vector de características y redundancias:** son consecuencia del proceso de fusión y conllevan un aumento del tiempo de procesamiento, al tiempo que dificultan el entrenamiento del clasificador. Para paliar estos problemas se aplica una SC.

Supóngase que se fusionan m métodos diferentes, cada uno de los cuales produce un vector de características $\mathbf{x}_i = \{x_1^{(i)}, \dots, x_{D_i}^{(i)}\}$ (con $i = 1, \dots, m$). El proceso de fusión es el siguiente:

- Si ningún método rechaza la huella, se concatenan todos los vectores de características $\mathbf{x}^{(i)}$ como se muestra en la Ecuación 1, obteniendo un vector \mathbf{x} de tamaño $D = \sum_{i=1}^m D_i$.

$$\mathbf{x} = \{x_1^{(1)}, \dots, x_{D_1}^{(1)}, \dots, x_1^{(i)}, \dots, x_{D_i}^{(i)}, \dots, x_1^{(m)}, \dots, x_{D_m}^{(m)}\} \quad (1)$$

- Si alguno de los métodos rechaza la huella, no se genera vector fusionado y la huella se rechaza.

Por tanto la fusión rechaza las huellas que son rechazadas por cualquiera de los métodos que se fusionan, lo que puede conllevar un aumento de la tasa de rechazo. La reducción de esta tasa se lleva a cabo mediante un clasificador jerárquico en el que cada nivel utiliza una combinación de extractores de características diferente. Dada una huella a clasificar se intentan extraer las características para el primer nivel. Si la huella se rechaza, se intenta con el segundo nivel, y así sucesivamente hasta alcanzar el primer nivel en el que la huella no se rechace y por tanto se obtenga un vector de características fusionado con el que clasificar la huella. Si la última de las combinaciones incluye solamente métodos sin rechazo, se asegura que el clasificador jerárquico no rechaza ninguna huella. En general, es deseable que los primeros niveles tengan mayor rechazo y mayor acierto para mantener la máxima precisión.

Para completar el sistema, se aplica una selección de características sobre los vectores fusionados. De entre los diversos algoritmos de SC propuestos en la literatura, se ha escogido un algoritmo empotrado [11]. Este método entrena un clasificador Random Forest [1] sobre los datos para sumar la ganancia de información en cada una de las bifurcaciones de los árboles. De este modo, se obtiene un vector de pesos que indica la importancia global de cada característica. El conjunto final de características seleccionadas se forma escogiendo las D' características con mayor peso, lo que permite una aplicación muy flexible de la selección de características estableciendo manualmente el valor de D' .

El entrenamiento del clasificador jerárquico completo se esquematiza en la Figura 2, incluyendo los pasos de extracción de características, fusión de los métodos, selección de características y entrenamiento del clasificador de cada nivel de la jerarquía. En este ejemplo se fusionan cinco métodos de extracción, tres de ellos con rechazo (A, B y C) y dos sin rechazo (D y E), y se hace una clasificación en una jerarquía de dos niveles. El primer nivel fusiona cinco tipos de características; el segundo nivel fusiona dos. La clasificación de una huella nueva se presenta en la Figura 3.

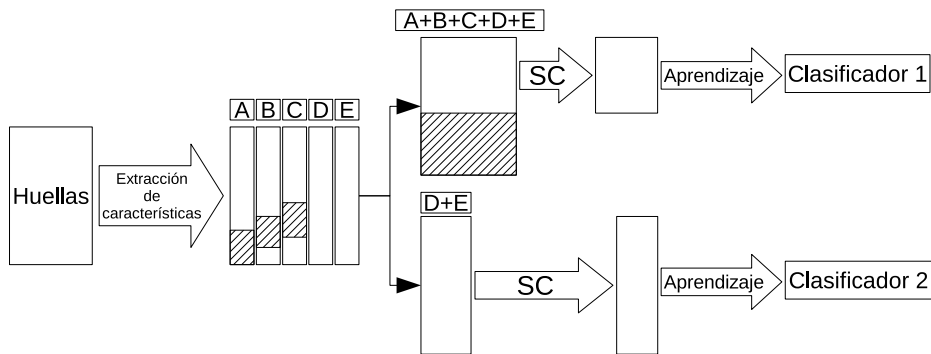


Figura 2. Entrenamiento del método jerárquico de clasificación con SC.

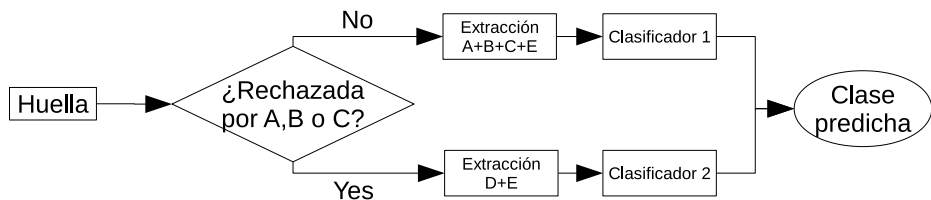


Figura 3. Clasificación de una huella.

4. Estudio Experimental

Para evaluar el comportamiento de las tasas de clasificación y rechazo al utilizar el sistema propuesto, éste se ha evaluado utilizando una base de datos de 10.000 huellas y seis extractores de características diferentes. La Sección 4.1 describe la base de datos utilizada y los algoritmos de clasificación y extracción. La Sección 4.2 detalla y analiza los resultados obtenidos.

4.1. Base de Datos y Algoritmos

Se ha utilizado una base de datos de 10.000 huellas, denominada VQAndPert, generadas con el *software* SFinGe [2,17] con los parámetros mostrados en [5]. Se puede ver que las huellas generadas incluyen ruido, perturbaciones, cicatrices y otros factores que dificultan la clasificación.

Una única impresión por cada dedo se seleccionó para la base de datos de 10.000 imágenes sobre la que se ha realizado el estudio experimental. Esta base de datos se dividió en 5 subconjuntos de 2.000 huellas naturalmente distribuidos para aplicar una validación cruzada de cinco iteraciones, que se utiliza en todos los resultados presentados. Los conjuntos de datos resultantes de la aplicación de los métodos de extracción de características utilizados en este estudio están disponibles públicamente en <http://sci2s.ugr.es/fingerprintClassification>.

El sistema propuesto se basa en tres tipos de algoritmos: un conjunto de extractores de características, un clasificador y un selector de características. Se han utilizado los métodos de extracción de la Tabla 1. Cada método se denomina mediante el apellido del primer autor del trabajo en el que se propone. Los clasificadores utilizados para la experimentación son Random Forest [1] y SVM [8] (con *kernel* RBF). Random Forest también se ha utilizado para la SC. Los parámetros de estos algoritmos se muestran en la Tabla 2, donde D' es el número de características seleccionadas.

Tabla 2. Parámetros de los clasificadores

Algoritmo	Parámetros
Random Forest	Árboles: 1.000; Variables por split: $\sqrt{D'}$
SVM	Kernel RBF; $C = 1$; Tolerancia = 0,001; Epsilon = 0,1; $\gamma = 1/D'$

4.2. Resultados

Como punto de partida para este estudio, la Tabla 3 muestra el porcentaje de clasificación y de rechazo de cada método de extracción aislado y cada clasificador, como promedio en una validación cruzada de 5 iteraciones. El resultado más preciso con y sin rechazo se resalta en negrita y subrayado, respectivamente. Como puede observarse, el método de Hong es el más preciso, aunque se produce un importante rechazo. El más preciso de los métodos sin rechazo es el de Liu.

Los mejores aciertos obtenidos tras fusionar los distintos métodos de extracción con y sin rechazo se muestran en la Tabla 4. El mejor valor para cada combinación de extractores aparece en negrita. La diferencia de precisión entre métodos con y sin rechazo es solamente de un 1%. Además, la fusión mejora el acierto respecto al mostrado en la Tabla 3.

La Figura 4 describe la evolución de la tasa de clasificación en función del número de características seleccionadas tras aplicar el algoritmo empotrado de SC basado en Random Forest sobre los conjuntos de datos fusionados. Por simplicidad sólo se incluyen los resultados con los dos mejores clasificadores: Random

Tabla 3. Tasas de clasificación y rechazo para cada extractor y clasificador

Extractor	Random Forest	SVM-RBF	Rechazo
Cappelli	0,8908	0,8494	0,00 %
HongFe	0,9359	0,9490	15,90 %
Leung	0,9027	0,9064	0,00 %
Li	0,7981	0,6763	0,00 %
Liu	0,9259	0,9130	0,00 %
Nyongesa	0,7033	0,7353	0,00 %

Tabla 4. Combinaciones de extractores con mayor acierto, con y sin rechazo

Extractor	Random Forest	SVM-RBF	Rechazo
Hong-Liu-Nyongesa-Leung	0,9490	0,9568	15,90 %
Hong-Liu-Leung	0,9486	0,9564	15,90 %
Hong-Liu-Leung-Li	0,9493	0,9549	15,90 %
Hong-Liu-Nyongesa-Leung-Li	0,9505	0,9549	15,90 %
Hong-Liu-Nyongesa	0,9457	0,9540	15,90 %
Hong-Liu	0,9452	0,9539	15,90 %
Hong-Liu-Cappelli-Li	0,9534	0,9384	15,90 %
Hong-Liu-Cappelli-Nyongesa	0,9530	0,9416	15,90 %
Liu-Nyongesa-Leung	0,9424	0,9462	0,00 %
Liu-Cappelli-Leung-Li	0,9456	0,9394	0,00 %
Liu-Cappelli-Nyongesa-Leung-Li	0,9454	0,9400	0,00 %
Liu-Leung	0,9419	0,9454	0,00 %
Liu-Cappelli-Nyongesa-Li	0,9448	0,9193	0,00 %
Liu-Cappelli-Nyongesa-Leung	0,9444	0,9418	0,00 %
Liu-Cappelli-Leung	0,9441	0,9416	0,00 %
Liu-Leung-Li	0,9437	0,9425	0,00 %

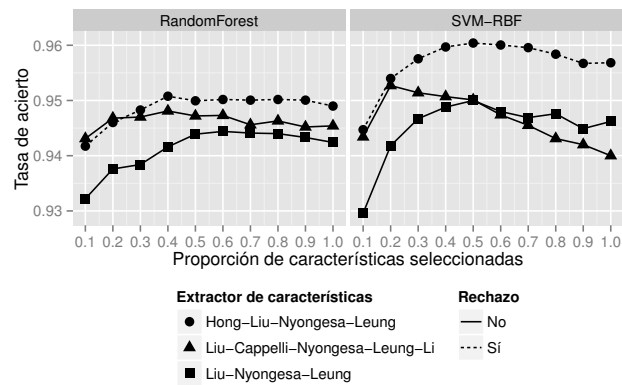


Figura 4. Tasa de clasificación al aplicar selección de características.

Forest y SVM-RBF. Se observa que utilizando Random Forest como clasificador, SC aporta una mejora muy pequeña del acierto debido al propio mecanismo de bifurcación de los árboles de Random Forest, que tiene en cuenta la importancia de cada atributo. De hecho, este mecanismo es el utilizado en el método de SC. En cambio, SVM-RBF sí ve su acierto mejorado de forma importante. Es especialmente destacable el resultado para la combinación de los cinco extractores sin rechazo, dado que la mejor precisión se obtiene utilizando un 20% de las características.

Finalmente, se han diseñado dos variantes del sistema jerárquico completo, cuyos métodos de extracción y resultados aparecen en la Tabla 5. Se observa que la SC aumenta el acierto en torno a un 0,4% al seleccionar un 50% y 60% de características en el primer y segundo nivel, respectivamente.

Tabla 5. Variantes para clasificador jerárquico (sin rechazo).

	Opción A	Opción B
Nivel 1	Hong-Liu-Nyongesa-Leung	Hong-Liu-Nyongesa-Leung
Nivel 2	Liu-Nyongesa-Leung	Liu-Cappelli-Nyongesa-Leung-Li
Clasificación (sin SC)	0,9497	0,9502
Clasificación (con SC)	0,9530	0,9540

5. Comentarios Finales

En este trabajo se ha propuesto un método de clasificación jerárquico para huellas dactilares, que combina una fusión de distintos métodos de extracción de características con un algoritmo de selección de características para eliminar el rechazo en la extracción y mejorar el acierto en clasificación. Los resultados experimentales muestran que la propuesta mejora la tasa de clasificación respecto al uso de extractores de forma aislada, incluso cuando éstos extractores rechazan huellas. Además, la selección de características elimina redundancias de forma que el acierto aumenta al tiempo que se reduce el tiempo de cómputo necesario para el entrenamiento y la aplicación de los clasificadores.

Como trabajo futuro se plantea integrar la clasificación dentro de un sistema completo de identificación de modo que ambos procesos sean escalables a bases de datos de tamaño arbitrario.

Referencias

1. Breiman, L.: Random forests. *Machine Learning* 45(1), 5–32 (2001)
2. Cappelli, R., Maio, D., Maltoni, D.: Synthetic fingerprint-database generation. In: *Proceedings of the 16th International Conference on Pattern Recognition*. vol. 3, pp. 744–747 (2002)

3. Cappelli, R., Maio, D., Maltoni, D.: A multi-classifier approach to fingerprint classification. *Pattern Analysis & Applications* 5(2), 136–144 (2002)
4. Galar, M., Derrac, J., Peralta, D., Triguero, I., Paternain, D., Lopez-Molina, C., García, S., Benítez, J., Pagola, M., Barrenechea, E., Bustince, H., Herrera, F.: A survey of fingerprint classification part I: Taxonomies on feature extraction methods and learning models. *Knowledge-Based Systems* 81, 76–97 (2015)
5. Galar, M., Derrac, J., Peralta, D., Triguero, I., Paternain, D., Lopez-Molina, C., García, S., Benítez, J., Pagola, M., Barrenechea, E., Bustince, H., Herrera, F.: A survey of fingerprint classification part II: Experimental analysis and ensemble proposal. *Knowledge-Based Systems* 81, 98–116 (2015)
6. García, S., Luengo, J., Herrera, F.: *Data Preprocessing in Data Mining*. Springer International Publishing, 1st edn. (2015)
7. Guyon, I., Elisseeff, A.: An introduction to variable and feature selection. *Journal of Machine Learning Research* 3, 1157–1182 (2003)
8. Hearst, M.A., Dumais, S.T., Osman, E., Platt, J., Scholkopf, B.: Support vector machines. *IEEE Intelligent Systems and their Applications* 13(4), 18–28 (1998)
9. Henry, E.: *Classification and Uses of Finger Prints*. London: Routledge (1900)
10. Hong, J.H., Min, J.K., Cho, U.K., Cho, S.B.: Fingerprint classification using one-vs-all support vector machines dynamically ordered with naive Bayes classifiers. *Pattern Recognition* 41(2), 662–671 (2008)
11. Huynh-Thu, V., Saeys, Y., Wehenkel, L., Geurts, P.: Statistical interpretation of machine learning-based feature importance scores for biomarker discovery. *Bioinformatics* 28(13), 1766–1774 (2012)
12. Karu, K., Jain, A.: Fingerprint classification. *Pattern Recognition* 29(3), 389–404 (1996)
13. Kohavi, R., John, G.: Wrappers for feature selection. *Artificial Intelligence* 97, 273–324 (1997)
14. Leung, K., Leung, C.: Improvement of fingerprint retrieval by a statistical classifier. *IEEE Transactions on Information Forensics and Security* 6(1), 59–69 (2011)
15. Li, J., Yau, W.Y., Wang, H.: Combining singular points and orientation image information for fingerprint classification. *Pattern Recognition* 41(1), 353–366 (2008)
16. Liu, M.: Fingerprint classification based on adaboost learning from singularity features. *Pattern Recognition* 43(3), 1062–1070 (2010)
17. Maltoni, D., Maio, D., Jain, A., Prabhakar, S.: *Handbook of fingerprint recognition*. Springer-Verlag New York Inc (2009)
18. Nyongesa, H.O., Al-Khayatt, S., Mohamed, S.M., Mahmoud, M.: Fast robust fingerprint feature extraction and classification. *Journal of Intelligent and Robotic Systems* 40(1), 103–112 (2004)
19. Peralta, D., Galar, M., Triguero, I., Paternain, D., García, S., Barrenechea, E., Benitez, J., Bustince, H., Herera, F.: A survey on fingerprint minutiae-based local matching for verification and identification: Taxonomy and experimental evaluation. *Information Sciences* 315, 67–87 (2015)
20. Saeys, Y., Inza, I., Larrañaga, P.: A review of feature selection techniques in bioinformatics. *Bioinformatics* 19, 2507–2517 (2007)
21. Wang, L., Dai, M.: Application of a new type of singular points in fingerprint classification. *Pattern Recognition Letters* 28(13), 1640–1650 (2007)