Knowledge-Based Systems 81 (2015) 98–116

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

A survey of fingerprint classification Part II: Experimental analysis and ensemble proposal

Mikel Galar [a,*], Joaquín Derrac [g], Daniel Peralta [b], Isaac Triguero [e,f,b], Daniel Paternain [a], Carlos Lopez-Molina [a], Salvador García [b,c,d], José M. Benítez [b], Miguel Pagola [a,h], Edurne Barrenechea [a,h], Humberto Bustince [a,h], Francisco Herrera [b]

[a] Departamento de Automática y Computación, Universidad Pública de Navarra, 31006 Pamplona, Spain
[b] Department of Computer Science and Artificial Intelligence, University of Granada, 18071 Granada, Spain
[c] Faculty of Computing and Information Technology – North Jeddah, King Abdulaziz University, 21589 Jeddah, Saudi Arabia
[d] Department of Computer Science, University of Jaén, 23071 Jaén, Spain
[e] Department of Respiratory Medicine, Ghent University, 9000 Gent, Belgium
[f] VIB Inflammation Research Center, 9052 Zwijnaarde, Belgium
[g] Affectv, Exmouth House, 1 Pine Street, London EC1R 0JH, United Kingdom
[h] Institute of Smart Cities (ISC), Universidad Pública de Navarra, 31006 Pamplona, Spain

ARTICLE INFO

Article history:
Received 7 October 2014
Received in revised form 10 January 2015
Accepted 8 February 2015
Available online 23 February 2015

Keywords:
Fingerprint classification
Feature extraction
Classification
Fingerprint recognition
SVM
Neural networks
Ensembles
Orientation map
Singular points
Experimental evaluation

ABSTRACT

In the first part of this paper we reviewed the fingerprint classification literature from two different perspectives: the feature extraction and the classifier learning. Aiming at answering the question of which among the reviewed methods would perform better in a real implementation we ended up in a discussion which showed the difficulty in answering this question. No previous comparison exists in the literature and comparisons among papers are done with different experimental frameworks. Moreover, the difficulty in implementing published methods was stated due to the lack of details in their description, parameters and the fact that no source code is shared. For this reason, in this paper we will go through a deep experimental study following the proposed double perspective. In order to do so, we have carefully implemented some of the most relevant feature extraction methods according to the explanations found in the corresponding papers and we have tested their performance with different classifiers, including those specific proposals made by the authors. Our aim is to develop an objective experimental study in a common framework, which has not been done before and which can serve as a baseline for future works on the topic. This way, we will not only test their quality, but their reusability by other researchers and will be able to indicate which proposals could be considered for future developments. Furthermore, we will show that combining different feature extraction models in an ensemble can lead to a superior performance, significantly increasing the results obtained by individual models.

1. Introduction

Fingerprint classification has been a hot topic since 1975. In the first part of this series of two papers [27], we reviewed the different approaches that have been presented in the specialized literature in order to address this problem from a double perspective. First, we considered the problem of feature extraction (FE) [57] that aimed at obtaining a suitable representation of the fingerprint for its posterior processing. Second, we dealt with the classification problem from the machine learning point of view [19], where a classifier capable of classifying new fingerprints represented by their extracted features should be learned from a set of previously labeled fingerprints (represented by the same features).

In this context, we presented a taxonomy of FE methods and additionally other two of Singular Point (SP) and Orientation Map (OM) extraction methods. Similarly, we grouped the learning models into different categories. The revision of those works led us to try to investigate which one would perform better in a real implementation, i.e., would be more accurate in its predictions. However, currently, it is extremely difficult to answer this

* Corresponding author. Tel.: +34 948 166048; fax: +34 948 168924.
E-mail addresses: mikel.galar@unavarra.es (M. Galar), jderrac@decsai.ugr.es (J. Derrac), dperalta@decsai.ugr.es (D. Peralta), Isaac.Triguero@irc.vib-UGent.be (I. Triguero), daniel.paternain@unavarra.es (D. Paternain), carlos.lopez@unavarra.es (C. Lopez-Molina), sglopez@ujaen.es (S. García), J.M.Benitez@decsai.ugr.es (J.M. Benítez), miguel.pagola@unavarra.es (M. Pagola), edurne.barrenechea@unavarra.es (E. Barrenechea), bustince@unavarra.es (H. Bustince), herrera@decsai.ugr.es (F. Herrera).

http://dx.doi.org/10.1016/j.knosys.2015.02.015
0950-7051/

question by simply reviewing the existing literature. This fact is due to the different experimental set-ups that have been used in fingerprint classification. As we discussed in the first part of the paper, there are even papers using the same database (for example NIST-4 [77]) for testing their model but in different ways [35,32,10,82,72,59,29,42], which make them not comparable, even though some authors continue comparing algorithms among papers by simply taking the results from them [35,10,67,59,29,42,43] (despite the differences in the evaluation procedures).

For this reason, our aim in this second part is to experimentally study the performance of different FE methods and learning models used for fingerprint classification. We want to investigate which method performs better in a common experimental framework, which can thereafter be used as a baseline for comparing new algorithms presented by other researchers. We are not doubting on the results presented in the corresponding papers, but since they are in many cases not comparable, we will carefully implement and test them in a common experimental framework with different databases. Moreover, carrying out these implementations would show which methods are not only better but easier to be reproduced by other researchers, which many times is overlooked despite its importance. Therefore, this study will allow us to extract meaningful conclusions about the fingerprint classification literature that will also show which methods can be recommended to practitioners when facing the problem of developing a fingerprint classification system.

In order to perform such a deep experimental study, we have implemented some of the most relevant FE methods covering all the categories in our taxonomy for FE (orientations, SPs, ridge structure and filter responses), including works from 1995 and taking into account their relevance in terms of citations received, but always with a previous analysis of the possibility of their implementation (see Discussion section in [27]). Finally, fourteen FE methods have been considered, which we show in Table 1. The works of Tan [72] and Park [59] were also implemented but due to different reasons they were finally leave out of this comparison (poor results, generate too big data-sets and the implementation may not follow the real one proposed by the authors due to a lack of details even though in our initial analysis they seemed to be reproducible). [1]

Since most of the FE methods are proposed in conjunction with a learning model by the authors, we have also implemented these models and included in the comparison in addition to the three classical classifiers that we have selected for carrying out the comparison: Support Vector Machines (SVMs) [73], C4.5 decision tree [62] and *k*-Nearest Neighbor (*k*NN) algorithm [1].

All these methods are tested in two different types of fingerprint databases: NIST-4 database [77], where most of the methods were originally tested; SFinGe (Synthetic Fingerprint Generator) tool[2] [13,49] based fingerprint databases, which will allow us to simulate different real-world scenarios with varying fingerprint qualities.

In addition to this study, we propose the usage of an ensemble in order to improve the performance of the individual models. To

Table 1
Fingerprint classification works considered in the experimental study.

Name	Year	Features	Orientations	Singular points	Reference point	Classification technique	Test databases	Reference
Candela	1995	Orientations (Registered) Ridge structure (Ridge-tracing (Generic))	Slits sum (16 × 16)		Rule-based (R92)	Neural Networks	NIST-14	[8]
Karu	1996	Singular points (Number/positions)	Slits sum (8 × 8)	Poincarè		Fixed Classification	NIST-4/9	[35]
Cappelli99a	1999	Orientations (Segmentation)	Slits sum (16 × 16)			Graph Matching	NIST-4/14	[10]
Jain	1999	Filter responses (Gabor)	Gradient (16 × 16)	Poincarè	Covariances	Multiple Techniques	NIST-4	[32]
Cappelli02	2002	Orientations (Registered) Orientations (Segmentation)	Slits sum (16 × 16)	Poincarè	Rule-based (R92)	Nearest Neighbor	NIST-14	[12]
Zhang04	2004	Singular points (Number/positions) Ridge structure (Ridge-tracing (from SPs))	Gradient (16 × 16)	Poincarè		Fixed Classification	NIST-4	[82]
Nyongesa	2004	Singular points (Number/positions) Singular points (Relative measures)	Gray-level Consistency/variance (5 × 5)	Poincarè		Neural Networks	NIST-4	[58]
Shah	2004	Orientations (Direct/unaltered) Orientations (Registered)	Line detection (16 × 16)		Entropy	Nearest Neighbor	Other	[67]
Wang07	2007	Singular points (Number/positions) Ridge structure (Ridge-tracing (from SPs))	Gradient (16 × 16)	Poincarè		Fixed Classification	NIST-4	[74]
Hong08	2008	Singular points (Number/positions) Ridge structure (Ridge-tracing (from SPs)) Filter responses (Gabor)	Gradient (16 × 16)	Poincarè	Covariances	Multiple Techniques	NIST-4	[29]
Li	2008	Orientations (Registered) Singular points (Number/positions) Singular points (Relative measures)	Gradient (8 × 8)	Complex filters		Support Vector Machines	NIST-4	[42]
Liu10	2010	Singular points (Relative measures)	Gradient (10 × 10)	Complex filters		Structural Models	NIST-4	[43]
Leung	2011	Filter responses (Gabor)	Gradient (16 × 16)		Rule-based (R92)	Other	FCV,NIST-4	[40]
Le	2012	Filter responses (Gabor)	Gradient (6 × 6, 12 × 12, 18 × 18)	Complex filters		Nearest Neighbor	FCV2004	[39]

[1] In [72], 2496 original features were generated that should be reduced by a Genetic Programming model. However, following the implementation given by the authors the dimensionality reduction has been intractable (which may be due to the lack of some key details of this algorithm such as initialization of the chromosomes, number of generations, etc.). In the case of [59], even though it only represents OMs of 21×21 blocks, they were codified with gray scale orientations in an image, leading to 11,025 features, which would lead to a high-dimensional problem. Moreover, we were not able to reduce such large dimensionality following the explanations given in the source paper.

[2] Synthetic Fingerprint Generator: http://biolab.csr.unibo.it/research.asp?organize=Activities&select=&selObj=12&pathSubj=111——12&.