Knowledge-Based Systems 81 (2015) 76–97

Contents lists available at ScienceDirect

Knowledge-Based Systems

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

A survey of fingerprint classification Part I: Taxonomies on feature extraction methods and learning models

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 9 January 2015
Accepted 8 February 2015
Available online 14 February 2015

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

ABSTRACT

This paper reviews the fingerprint classification literature looking at the problem from a double perspective. We first deal with feature extraction methods, including the different models considered for singular point detection and for orientation map extraction. Then, we focus on the different learning models considered to build the classifiers used to label new fingerprints. Taxonomies and classifications for the feature extraction, singular point detection, orientation extraction and learning methods are presented. A critical view of the existing literature have led us to present a discussion on the existing methods and their drawbacks such as difficulty in their reimplementation, lack of details or major differences in their evaluations procedures. On this account, an experimental analysis of the most relevant methods is carried out in the second part of this paper, and a new method based on their combination is presented.

1. Introduction

Classification in Machine Learning (ML) is the problem of extracting knowledge from a set of n input examples $x_1, \ldots, x_n$ characterized by i features $a_1, \ldots, a_i \in \mathbb{A}$, including numerical or nominal values, where each instance is labeled with a desired output class label $y_j \in \mathbb{C}$ (considering a m class problem $\mathbb{C} = \{c_1, \ldots, c_m\}$) and the aim is to learn a system capable of predicting this output for a new unseen example in a reasonable way (with good generalization ability) [26]. The system generated by the learning algorithm is a mapping function defined over the patterns $\mathbb{A}^i \rightarrow \mathbb{C}$ and it is called a *classifier*.

Therefore, fingerprint classification problem consists of learning a classifier from a set of labeled fingerprints, which should be able to classify new fingerprints in the corresponding class. The most commonly used fingerprint classification model was given by Henry [37]. Most of the classification approaches reviewed in this paper consider the five major classes shown in Fig. 1: *Arch, Tented Arch, Right Loop, Left Loop* and *Whorl*. These fingerprint classes are unevenly distributed in the population ($3.7\%, 2.9\%, 31.7\%, 33.8\%$ and 27.9%, respectively), which increases the difficulty of the classification problem from the ML point of view [33], but also makes the reduction of the search space class-dependent.

The fingerprint classification problem arises from the problem of fingerprint identification, which aims to claim the identity of a person by their fingerprint [77]. Unlike in the verification problem

* 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.2015.02.008
0950-7051/

(where the aim is to check whether two fingerprints are the same), the number of matchings that need to be carried out grow along with the number of individuals in the database. Hence, a reduction of the number of comparisons is required in order to maintain the response times as short as possible. This reduction is usually quantified with the ratio of penetration [99] in the database, which measures the percentage of the database that is searched before matching the fingerprint. Classification is the most extended method to reduce the ratio of penetration, which is also directly related with the classification accuracy (percentage of correctly classified examples) obtained by the classification methods. This paper focuses on this type of methods to reduce the search space even though other techniques have also been developed such as indexing [6,64,14], continuous classification [15,17,19] or clustering and classification [47,68,69].

Although classification in ML typically refers to learning a classifier from a set of examples characterized by several features, fingerprint classification usually stands for the problem as a whole, including feature extraction [89]. Feature extraction (FE) consists of obtaining a set of features that are able to properly characterize an object for its posterior processing. In fingerprint classification, FE aims to describe a fingerprint as accurately as possible in order to facilitate its classification among the predefined classes. FE is a key issue, since the classification problem directly depends on the quality of the features considered.

However, from our point of view, fingerprint classification can be divided into two well-differentiated steps, which is the viewpoint considered in this paper (although some proposals [46,104,38,61] consider strongly related models).

1. First, we deal with FE, that is, how to obtain a suitable representation of the fingerprint so as to achieve an accurate classification.
2. Second, we consider the classification problem as it is commonly done in ML, considering the construction of a classifier capable of classifying previously unknown fingerprints using the features extracted in FE phase.

Our aim is to review the different works proposed in the literature paying attention to both phases independently. As a result, we will put forward a taxonomy in which the different FE methods presented in the literature can be placed depending on the nature of the characteristics considered. Additionally, we will present two taxonomies in which the methods considered in fingerprint classification papers for the extraction of orientation maps and singular points can be classified (Sections 3 and 4). We should emphasize that these are the two most important features for fingerprint classification. Finally, we will make an overview of the techniques that have been used to address the classification problem; we consider different groups of algorithms and their evolution in the literature will be analyzed.

The study of all these methods has led us to a thorough discussion in Section 6, where a critical view of the reviewed works is presented regarding the lack of details in their descriptions, their reimplementability and the existing differences on the way they are evaluated, among others. On this account, we aim to experimentally show this problem in the second part of this paper [34], where several relevant methods have been implemented by the authors and an exhaustive experimental evaluation is carried out in a common experimental framework. This way, their results will be objectively analyzed and their validity for their usage by other researchers will be shown.

The rest of this paper is organized as follows. Section 2 introduces the fingerprint classification problem and recalls the most important concepts on this topic. Next, Section 3 deals with the two most important processes in fingerprint classification: the extraction of orientation maps and singular points, and presents their taxonomies. Then, Section 4 puts forward our taxonomy proposal for the classification of FE techniques and reviews the existing works in each one of the categories considered. Afterwards, Section 5 describes the different ML models that have been considered in the fingerprint classification literature. The discussion on the works reviewed along this paper is presented in Section 6, whereas Section 7 concludes this paper.

2. Background

Fingerprint features or characteristics are usually classified into three levels [77,29]:

- *Level 1 (Global)* – refers to the global ridge line flow (orientations) and the features derived from it (singular points).
- *Level 2 (Local)* – considers minutiae details extracted from the ridge skeleton.
- *Level 3 (Fine-detail)* – includes intra-ridge details such as width, shape, ridge contours, sweat pores, and creases.

Among these levels, only the first one is used for fingerprint classification (with few exceptions [95]), since fingerprint classes are intuitively defined from global characteristics. Otherwise, level 2 and 3 features are commonly considered for fingerprint matching [48,23,13,94] as they allow one to claim for the individuality of a fingerprint. Therefore, FE for classification is mainly carried out with level 1 features, that is, fingerprint features for classification are closely related to fingerprint orientations and Singular Points (SPs). Fingerprint orientations are represented in an Orientation Map (OM), which is the representation of the local ridge flow in the fingerprint. SPs are defined as the locations in the fingerprint with the greatest ridge orientation variance, i.e., where the ridges vary more abruptly. There are two types of SPs known as *cores* and *deltas*. Fig. 2 shows a fingerprint image (2a), its OM (2b) and

(a) Arch (b) Tented Arch (c) Right Loop (d) Left Loop (e) Whorl

Fig. 1. The five major classes defined by Henry [37] considered in the fingerprint classification problem.