Pattern Recognition 56 (2016) 63–76

Contents lists available at ScienceDirect

Pattern Recognition

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

Incremental granular relevance vector machine: A case study in multimodal biometrics

Hunny Mehrotra [a], Richa Singh [b], Mayank Vatsa [b,*], Banshidhar Majhi [a]

[a] National Institute of Technology Rourkela, India
[b] Indraprastha Institute of Information Technology (IIIT), Delhi, India

ARTICLE INFO

Article history:
Received 24 September 2013
Received in revised form
3 September 2015
Accepted 15 November 2015
Available online 15 December 2015

Keywords:
Incremental learning
Granular computing
Relevance vector machine
Biometrics

ABSTRACT

This paper focuses on extending the capabilities of relevance vector machine which is a probabilistic, sparse, and linearly parameterized classifier. It has been shown that both relevance vector machine and support vector machine have similar generalization performance but RVM requires significantly fewer relevance vectors. However, RVM has certain limitations which limits its applications in several pattern recognition problems including biometrics such as (1) slow training process, (2) difficult to train with large training samples, and (3) may not be suitable to handle large class imbalance. To address these limitations, we propose *i*GRVM which incorporates incremental and granular learning in RVM. The proposed classifier is evaluated in context to multimodal biometrics score classification using the NIST BSSR1, CASIA-Iris-Distance V4, and Biosecure DS2 databases. The experimental analysis illustrates that the proposed classifier can be a good alternative for biometric score classification with faster testing time.

1. Introduction

Classifiers are an integral component of a pattern classification system. In order to determine the class of any query, the data is processed, a representation is computed, and the classifier classifies it into one of the classes. Before testing, the classifier learns a model using the given training data. For instance, in a biometric verification problem, there are two classes, *genuine* and *imposter*. The task is to match the probe image with the corresponding gallery image and determine whether the probe is a genuine match or imposter. Existing biometric recognition algorithms have used different classifiers such as linear threshold, Bayesian classification, and Support Vector Machine (SVM) [1].

For training an accurate classification model, it is generally assumed that sufficient and representative training data is available during the training stage. However, in real world applications, there are several challenges in ensuring the availability of good quality training data:

- There exists the possibility that the entire training data is not available simultaneously. For example, in the case of India's Aadhaar project [2] or US-VISIT program [3], users are enrolled on a continuous basis. In such a scenario, training data is available only in an incremental manner. Training the classifiers in batch mode with every incremental update can be computationally expensive.
- Training databases can be highly unbalanced where data from one class is over populated compared to other class(es). In a biometric system that has n users in the database each having m samples ($n \gg m$), the number of genuine scores available for training is $nm(m-1)/2$ in comparison to $n(n-1)m^2/2$ impostor scores.
- Some classifiers are inherently computationally expensive, they perform well if the training size is small but on large training data they may require significant computational time or become intractable.

To address some of these challenges, researchers have proposed multiple solutions. The availability of sequential training data is addressed by incremental learning and online learning algorithms [4]. In incremental learning, classifiers are trained with new batches of data, as they arrive, while preserving the knowledge of previous learning. Some incremental learning approaches are incremental Principal Component Analysis (IPCA) [5], incremental learning of Bidirectional Principal Component Analysis [6], incremental Linear Discriminant Analysis (ILDA) [7], incremental Subclass Discriminant Analysis (ISDA) [8], and incremental and decremental SVM [9,10].

* Corresponding author. Tel.: +91 9654653404; fax: +91 11 26907405.
E-mail addresses: hunny@nitrkl.ac.in (H. Mehrotra), rsingh@iiitd.ac.in (R. Singh), mayank@iiitd.ac.in (M. Vatsa), bmajhi@nitrkl.ac.in (B. Majhi).

http://dx.doi.org/10.1016/j.patcog.2015.11.013
0031-3203/

In the literature, several researchers have also explored the challenge of class imbalance [11,12]. Chawla et al. [13] have stated that class imbalance problem is handled either by assigning distinct cost to training data [14–16] or by resampling the entire database [17]. The resampling approaches work by either oversampling the minority class and under-sampling the majority class, or by combining the under-sampling and oversampling approaches [18,19]. To balance class distributions, random under-sampling may lead to information loss whereas random oversampling can increase the chances of overfitting. Tang et al. [20] have proposed an under-sampling approach using granular learning. Granular learning divides the data into granules represented as either classes, clusters or subsets and solves the problem in each information granule locally [21]. The challenge of large training database for learning computationally expensive classifiers has also been addressed by granular computing approaches [22].

Since the formulation of every classifier is different, the extension of an existing classifier that operates in batch mode to the corresponding incremental version is also different. In designing the incremental or granular variant of an existing classifier, it is important to ensure that the updated variants do not reduce the accuracy while reducing the training time or computational complexity. Therefore, researchers have proposed specific formulations for individual classifiers, such as SVM.

SVM has been shown to yield good results in several pattern classification problems including biometrics. It avoids overfitting and leads to good generalization by finding the separating hyperplane that maximizes the margin width. The subset of training data points used to represent the hyperplane are denoted as support vectors. Several formulations have been proposed for online training of SVM and addressing the class imbalance problem [10,20,22,23]. However, SVM suffers from the following limitations [24]:

1. The number of support vectors required for classification is relatively large,
2. In classical SVM, there is a need to fine tune the regularization parameter (C) during the training phase, and
3. The kernel function must satisfy the Mercer conditions [25].

Relevance vector machine (RVM) [24], on the other hand, is a probabilistic classifier which introduces a prior over each weight governed by the set of hyper-parameters. RVM is a sparse linearly parameterized model like SVM and it has been shown that the generalization performance of RVM is comparable to that of SVM with significantly fewer relevance vectors [24]. Another advantage of RVM is that it has very few parameters to be optimized while training. Along with these advantages, RVM has the following challenges owing to which it has not been well explored particularly in biometrics.

1. The native formulation of RVM requires expensive matrix inversion which makes it difficult to learn conventional RVM with very large training databases. Further, the amount of memory required to store the product of basis functions also limits its utilization for considerably large training databases.
2. RVM is trained in batch mode and if new batch of data arrives, the classifier has to be re-trained with new as well as old data. This is not feasible for many real-time applications such as biometrics where it may be required to continuously update the classifier to adjust the changes (in data and template) that happen over time.
3. RVM may not be suitable to handle large class imbalance in the training data and may get biased towards the class with more number of training samples.

To address these challenges, in this paper, we propose an incremental granular RVM that can be trained with large unbalanced training data to perform efficient classification. As shown in Fig. 1, the learning process starts by considering batches of training data which are divided into granules. An RVM is trained on each granule independently and the results are amalgamated to obtain a robust boundary for classification. For online learning, the knowledge from the previous training is carried forward to learn the next batch of training database. The major contributions of this research are:

1. *Incremental RVM* (*i*RVM) is proposed which is scalable with new enrollments and also reduces the training time.
2. *Granular RVM* (GRVM) handles the class imbalance problem by training the classifier locally for each granule.
3. *Incremental Granular RVM* (*i*GRVM) combines the advantages of both incremental and granular learning into RVM.

The proposed variant provides a good alternative to existing classifiers and overcomes the limitations of native RVM classifier. The performance of incremental granular RVM is evaluated using a case study in multimodal biometrics with two classes (genuine and imposter). The match scores obtained from different modalities, units and algorithms are normalized followed by incremental granular RVM classification. Experiments performed on three match score databases show that the proposed classifier is comparable to existing approaches in terms of classification performance and provides significant reduction in computational time.

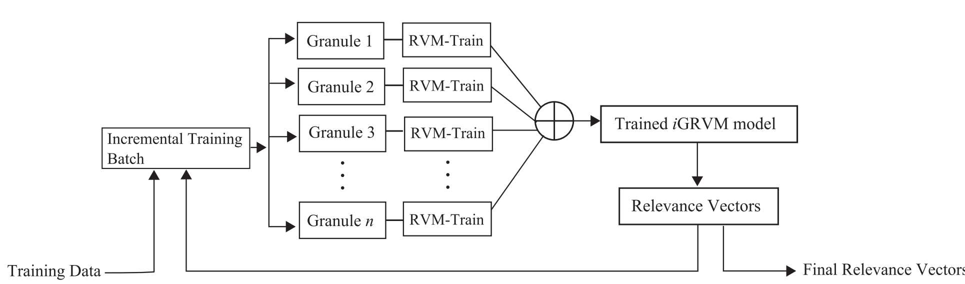

Fig. 1. Block diagram of incremental granular relevance vector machine (*i*GRVM).