Neurocomputing 188 (2016) 167–177

Contents lists available at ScienceDirect

Neurocomputing

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

RABOC: An approach to handle class imbalance in multimodal biometric authentication ☆

Quang Duc Tran *, Panos Liatsis

Department of Electrical and Electronic Engineering, City University London, Northampton Square, London EC1V 0HB, UK

ARTICLE INFO

Article history:
Received 30 September 2014
Received in revised form
11 November 2014
Accepted 1 December 2014
Available online 4 December 2015

Keywords:
Biometric authentication
Class imbalance problem
One-class classification
Classifier ensemble
Real AdaBoost

ABSTRACT

Class imbalance poses serious difficulties to most standard two-class classifiers, when applied in performing classification in the context of multimodal biometric authentication. Most conventional classifiers assume equally balanced classes. They do not work well when impostor samples vastly outnumber the samples of the genuine user class. In this paper, we propose an algorithm, called RABOC, which inherits the natural capabilities of one-class classification and Real AdaBoost algorithm to handle the class imbalance problem in biometric systems. Particularly, we develop a weak classifier, which consists of one-class classifiers and is trained using data from both classes. We then exploit Real AdaBoost to combine the multiple weak classifiers in order to improve their performance without causing overfitting. Unlike the conventional Real AdaBoost, the weak classifiers in the proposed schema are learned on the same data set, but with different parameter choices. This not only generates the diversity necessary to make RABOC work, but also reduces the number of user-specified parameters. Extensive experiments were carried out on the BioSecure DS2 and XM2VTS benchmark databases, which involve data with extremely imbalanced class distribution. They demonstrate that the proposed RABOC algorithm can achieve a relative performance improvement of 28%, 24%, and 22% as compared to other state-of-the-art techniques, specifically the sum of scores, likelihood ratio based score fusion, and Support Vector Machines.

1. Introduction

Biometric verification is the process of authenticating a human identity using his/her behavioral and physiological characteristics. It is well-known that multimodal biometric systems can further improve the verification accuracy by combining information from multiple biometric traits at various levels, namely, i.e., sensor, feature, matching score, and decision levels [1]. Fusion at matching score level is generally preferred due to the trade-off between information availability and fusion complexity [1,2].

A common practice in many reported works on multimodal biometrics is to view fusion at matching score level as a two-class classification problem, where the vector of matching scores is treated as a feature vector, and can be classified into one of two-classes: genuine user/impostor [1]. Many classifiers, including HyperBF [3], k-Nearest Neighbours using vector quantization [4], C4.5 decision tree, Fisher linear discriminant, Bayesian classifier, Multilayer Perceptron, and Support Vector Machines (SVM) [5] have been applied to render to biometric decision in an identity verification system. However, recent studies have indicated that most conventional two-class classifiers are inadequate, when applied to problem, characterized by class imbalance [6–9]. Class imbalance is a common problem in multimodal biometric authentication, where the samples of the impostor class greatly outnumber the samples of the genuine user class; it is not unusual for class imbalance to be in the order of 500:1. Conventional two-class classifiers assume or expect balanced class distributions, and generally create suboptimal classification models, when presented with complex imbalanced data sets [7].

Class imbalance has received limited attention in the biometric literature. Under-sampling has been widely applied to handle the problem [10,11]. An obvious shortcoming of under-sampling is that it may cause the classifiers to miss important aspects in the data, pertaining to the impostor class, since the optimal class distribution is usually unknown [9]. In [12], a small percentage of Gaussian noise with respect to the largest magnitude of the match score was included in the genuine user class to increase the

☆Preliminary version of this manuscript has been selected in International Conference on Intelligent Computing (ICIC), 2014 (Paper ID: 591, Title: Improving Fusion with One-class Classification and Boosting in Multimodal Biometric Authentication) and its extended version is sub-selected for the Neurocomputing journal.

* Corresponding author. Tel.: +44 7424185310.
E-mail addresses: quang.tran.1@city.ac.uk (Q.D. Tran), p.liatsis@city.ac.uk (P. Liatsis).

http://dx.doi.org/10.1016/j.neucom.2014.12.126

training sample size and reduce the proportion of imbalanced data for different classes. However, using a single Gaussian component can potentially reduce the applicability of this approach for combining non-Gaussian match scores. These factors indicate the need for additional research efforts to advance the classification performance of imbalanced biometric data sets.

Over the years, the machine learning community has addressed the issue of class imbalance in many different ways. Among others, the two most perspective approaches are one-class classification and boosting. One-class classification is naturally quite robust to the class imbalance problem by defining the decision boundary using single class samples (target samples) rather than distinguishing between samples of the two classes [6,13]. In [14–16], the authors suggested that one-class classifiers are particularly useful in handling extremely imbalanced data sets in feature spaces of high dimensionality, while two-class classifiers are more suitable for data sets with relatively moderate degrees of imbalance. On the other hand, boosting is a technique, which can be used to further improve the classification performance of any classifier regardless of whether the data is imbalanced [6,7]. The aim of boosting is to combine multiple (*weak*) classifiers in order to develop a highly accurate (*strong*) classifier system. It is known to reduce bias and variance errors as it focuses on the samples, which are harder to classify [8]. Particularly, boosting weighs each sample to reflect its importance, and places the most weights on those samples, which are most often misclassified by the preceding classifiers [8,9]. Boosting is effective at handling the class imbalance problem because the small class samples are most likely to be misclassified. Real AdaBoost [17] is considered as the most representative boosting algorithm. It is also resistant to classification noise, which appears very naturally in biometric applications [18].

The basic motivation of this paper is to address the problem of highly imbalanced class distributions in biometric data sets. Towards this end, we propose a novel hybrid boosting algorithm, called RABOC, which is capable of exploiting the natural capabilities of both Real AdaBoost (RAB) and one-class classification (OC). RABOC works by first developing a weak classifier, based on Bayes Decision Theory, for the fusion of one-class classifiers to effectively and efficiently use the training data from both the genuine user and impostor classes. The paradigm of Real AdaBoost is then applied to improve so far the performance of this classifier without causing overfitting. It has been recognized that diversity is a key requirement for the success of Real AdaBoost. Conventional Real AdaBoost generates diversity by training weak classifiers on different data subsets, constructed from the original training data. In the proposed paradigm, a new training procedure is introduced to train these classifiers on the same data set, but with different parameter choices. The target is to reduce the number of user-specified parameters, while still generating the diversity necessary to enable the classifier ensemble to perform well. Extensive experiments are carried out on the BioSecure DS2 [19] and XM2VTS [20] benchmark databases. They demonstrate that the proposed RABOC algorithm achieves significantly better results in terms of Half Total Error Rate (HTER) as compared to state-of-the-art solutions, namely the sum of scores [21], likelihood ratio based score fusion [22] and Support Vector Machines [23].

The remainder of the paper is organized as follows. Following the introduction, Section 2 presents an overview of reported imbalanced learning solutions. Section 3 provides a thorough discussion on the weak classifier algorithm and the mechanism of the proposed RABOC. Section 4 reports our experimental results. Section 5 is dedicated to conclusions and future work.

2. Imbalanced learning solutions: an overview

The significant difficulty of the class imbalance problem has been attracting an increased research interest. A large number of solutions have been developed to address the problem. There is however a lack of systematic study on the applicability of these solutions to advance the classification performance of highly imbalanced biometric data sets.

2.1. Data level approaches

Solutions at data level involve many different forms of sampling in order to rebalance the class distributions in the training set, and thus, decreasing the possible effects, caused by imbalance in the learning process. Oversampling and undersampling change the training set by replicating samples in the small class and removing samples from the prevalent class, respectively. Oversampling is time-consuming. It also leads to overfitting since it appends replicated data to the original training data [6]. Undersampling requires shorter training time. It, however, may risk losing important information for the prevalent class [9,24].

Synthetic Minority Oversampling Technique (SMOTE) [25] is a powerful method, which has shown a great success in various applications. SMOTE aims to find the k nearest neighbors of a small class sample and generate artificial samples in the direction of some or all the nearest neighbors, depending on the amount of oversampling required. Although SMOTE has many promising benefits, its drawbacks include over generation and variance, which potentially increases the occurrence of overlapping between classes [6].

2.2. Algorithm level approaches

Algorithm level solutions try to adapt existing classifier learning algorithms to strengthen the learning towards the small class. This is achieved by either adjusting the class boundary based on kernel modification methods or assigning different cost matrices, which corresponds to the costs for misclassifying any particular data sample [6]. Algorithm level solutions are both application-dependent and classifier-dependent. They may fail to produce a satisfactory performance on data with uneven class distributions [9].

One-class classification has attracted much attention in the community [6]. In [14,16], one-class classification was observed to be particularly useful in addressing extremely imbalanced data sets with high feature space dimensionality. Particularly, it offered a viable solution when the small class constitutes approximately 4% or less of the training data. In [15], it was reported that the one-class approach may be superior to the two-class approach under certain conditions, such as multimodal domains. In [26], the authors demonstrated that the standard two-class SVM is inferior to the one-class counterpart, when applied in performing classification in multimodal biometric verification systems.

2.3. Classifier ensemble learning

Classifier ensemble learning is the method of constructing multiple classifiers and then aggregating their predictions when classifying unknown samples. The basic motivation of combining classifiers is to improve their general capabilities: each classifier is known to make errors; however, the patterns that are misclassified by the different classifiers are not necessarily the same [9]. In the literature, the effect of combining classifiers is studied in terms of statistical concept of bias–variance decomposition [27]. In general, bias is associated with underfitting, while variance is associated with overfitting. The improved performance of a