

# ECOC 多类分类研究综述

雷 蕾<sup>1</sup>, 王晓丹<sup>1</sup>, 罗 玺<sup>2</sup>, 周进登<sup>3</sup>, 陈 琴<sup>4</sup>

(1. 空军工程大学防空反导学院, 陕西西安 710051; 2. 空军工程大学信息与导航学院, 陕西西安 710077;  
3. 空军装备软件测评中心, 北京 100076; 4. 93704 部队, 北京 101149)

**摘 要:** 纠错输出编码能有效地将多类问题转化为二类问题进行求解, 已受到国内外从事机器学习的研究者们的重视, 并使其成为多类分类领域的研究热点. 本文首先分析了 ECOC 多类分类的原理和框架, 指出解决 ECOC 多类分类问题的关键在于解码策略和编码策略的确定; 然后从这两个关键点出发综述了 ECOC 多类分类的最新进展和应用领域; 最后指出了目前存在的问题以及下一步研究方向. 论文研究成果将为基于 ECOC 多类分类方法在实际应用过程中起借鉴和参考作用.

**关键词:** 多类分类, 纠错输出编码, 机器学习

**中图分类号:** TP391

**文献标识码:** A

**文章编号:** 0372-2112 (2014)09-1794-07

**电子学报 URL:** <http://www.ejournal.org.cn>

**DOI:** 10.3969/j.issn.0372-2112.2014.09.020

## An Overview of Multi-Classification Based on Error-Correcting Output Codes

LEI Lei<sup>1</sup>, WANG Xiao-dan<sup>1</sup>, LUO Xi<sup>2</sup>, ZHOU Jin-deng<sup>3</sup>, CHEN Qin<sup>4</sup>

(1. *The Air and Missile Defense Institute, Air Force Engineering University, Xi'an, Shaanxi 710051, China*; 2. *The Information and Navigation Institute, Air Force Engineering University, Xi'an, Shaanxi 710077, China*; 3. *Air Force Equipment Software Testing Center, Beijing 10076, China*; 4. *Air Force 93704, Beijing 101149, China*)

**Abstract:** Multi-classification has been one of the research hotspot in pattern recognition, and there are many solutions to it. As a common way to model multi-classification to design a set of binary classifiers and fuse them, Error-correcting output codes (ECOC) represents a successful framework to deal with this type of problems and is attracting more and more attention of researchers. In this paper, the framework of ECOC is concluded at first. Then the two keys of multi-classification based on ECOC, i.e., the coding strategies and decoding strategies are proposed. The main part focuses on the research of the two keys and the application of ECOC. Finally, the still existing problems of ECOC are pointed out and the promising research fields are given. The analysis of the paper will provide reference and advice in the practical application of multi-classification based on ECOC.

**Key words:** multi-classification; error-correcting output codes; machine learning

## 1 引言

多类分类是模式识别的重要研究领域. 目前常用的解决方法有两种: (1) 直接构造多类分类器. 此方法看似最为直接, 然而事实证明由于现实环境中模式种类繁多、结构复杂、区别难度大, 很难直接构造出满足要求的多类分类器, 即使存在也通常由于其泛化能力弱而缺乏实用性. (2) 采用分而治之的办法. 纠错输出编码 (Error-Correcting Output Codes, ECOC) 作为一种多类分解框架, 从一提出便受到了众多学者的关注<sup>[1]</sup>. 其来源于通信领域中的信号传输问题, 即把多类分类问题建模为信号在

信道传输的问题, 多类分解为二类的过程可以看成是信源段对待传输信号的编码过程, 多类的分类过程为信号在信道的传输过程, 多类分类的决策融合过程则可以看成是信号在信宿段的解码过程, 由于信号在传输过程中会受噪声的影响而产生误码, 为了使传输的信号具有抗噪能力, 在信号的编码过程中往往会某些易受噪声影响的码位设计冗余编码, 这样在解码过程中当产生误码时冗余码可提供纠错能力. 而在多类分解框架中为了克服单个二分器所带来的误差也往往设计多个二分类器, 并在决策过程中通过某种融合规则达到容错的能力.

因此, 在研究基于 ECOC 进行多类分类过程时, 一

般把该多类分类问题分解为编码、训练、解码三个阶段进行研究,这也是文献[1]首次提出该理论时所使用的专业术语,可以看出这和通信学领域中所说的信号的编、解码还是有区别的,确切地说在研究多类的分解框架时,ECOC 方法只是借用通信领域中的相关概念,与其本身的含义有本质的区别.事实上,随着该理论的逐步发展,基于 ECOC 的多类分类研究已经从最开始的注重分解框架纠错能力提高方面逐步转移到如何构造符合问题域的输出编码矩阵.

本文首先分析了 ECOC 多类分类的原理和框架,指出解决 ECOC 多类分类问题的关键在于解码策略和编码策略的确定,然后结合现有文献针对 ECOC 多类分类研究进行了详细的分析;同时讨论了 ECOC 多类分类的典型应用;最后指出了现有方法存在的问题和需要进一步解决的研究方向.

## 2 ECOC 多类分类框架

ECOC 多类分类框架用一种二元或三元的编码矩阵实现多类类别分解和基分类器集成.在其编码矩阵中,二元码用  $\{-1, +1\}$  表示,三元码用  $\{-1, 0, +1\}$  表示,“-1”代表一类,“+1”代表另一类,“0”表示该码字位所对应的类在其列所形成的二类划分中被忽略(即不参与由该列所产生的基分类器的训练).图 1 给出了四种常见的 ECOC 分类系统示意图,以编码矩阵来区分,分别是:图 1(a)“一对多”编码阵(one-versus-all)、图 1(b)“一对一”编码阵(one-versus-one)、图 1(c)密集随机阵(Dense Random)、图 1(d)稀疏随机阵(Sparse Random).

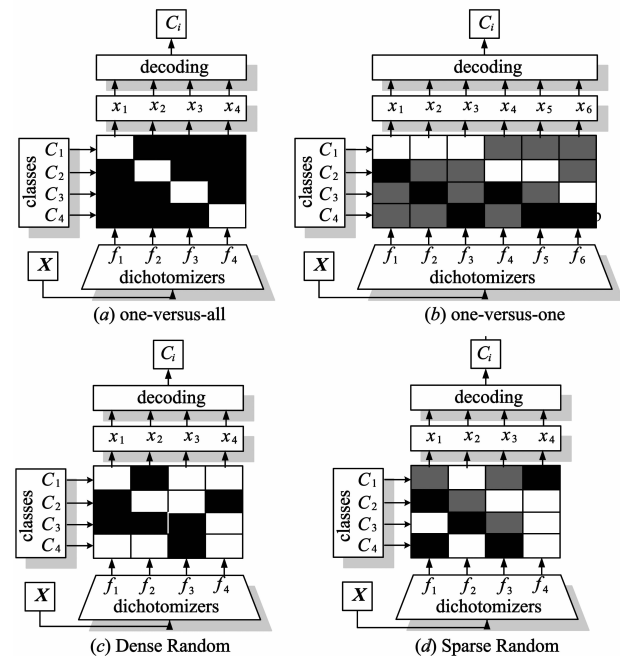


图 1 四种常见的 ECOC

图 1 中所有编码阵的每一行代表某一类  $C_i$  ( $i = 1, 2, 3, 4$ ) 的码字,每一列代表样本的一种二类划分,码元“1”、“-1”和“0”分别用白色、黑色和灰色表示.在训练阶段,每一个基分类器  $f_i$  ( $i = 1, 2, \dots, 6$ ) 的训练样本先根据其所在编码阵对应的列重新划分把多类转化成二类,然后分别训练得到与该列对应的二类分类器.例如,在图 1(d)中对基分类器  $f_3$  进行训练时,白色对应的  $C_2$  为一类,黑色对应的  $C_4$  为另一类,而灰色对应的  $C_1$  和  $C_3$  不参与该基分类器的训练.由此依次训练得到四个二类分类器  $\{f_1, f_2, f_3, f_4\}$ .在测试阶段,给定一个测试样本  $X$ ,同时利用这四个二类分类器对其进行分类,结果为一码字向量  $(x_1, x_2, x_3, x_4)$  (其中  $x_i \in \{-1, +1\}$ ),最后根据某种解码规则(即融合策略)对其进行解码即可得最终分类结果.

基于 ECOC 的多类分类问题简化了多模式分类问题的复杂性,使多类分类问题能利用较为成熟的二类模式分类方法加以解决;同时继承纠错码特有的纠错能力,使得利用一定的解码规则能对由二类分类器产生的错误具有一定的纠错能力.

利用 ECOC 进行多类分类的过程通常是先确定编码矩阵和解码策略,进而挑选最有利于问题解决的二类分类器,使得能最大可能的挑选出有利于问题解决的最优方案.其关键在于如何确定有效的分解框架(即编码矩阵的构造)和融合方法(即解码策略的确定).

## 3 ECOC 多类分类研究

自 1995 年 Dietterich 和 Bakiri 首次提出应用 ECOC 解决多类分类问题以来<sup>[1,2]</sup>,众多学者对其在多类分类问题的表现产生了浓厚兴趣,也因此产生了一系列与此相关的各种研究和结论.Kong 在 1995 年首次通过偏差方差分析(bias and variance analysis)对该方法能有效解决多类分类问题的原因进行了分析<sup>[3]</sup>.F. Masulli 和 G. Valentini 分析了 ECOC 的结构后认为:编码矩阵的结构和基于该编码所产生的基分类器之间的相关性是影响此类方法性能的主要因素<sup>[4]</sup>.李建武从最小封闭球几何描述的角度出发,对 binary ECOC 多类分类器进行了几何解释,并在实数编码空间中进行了推广,进一步探讨了给出后验概率估计的 ECOC 多类分类器实现策略<sup>[5]</sup>.到目前为止已有大量的相关文献针对 ECOC 进行了深入研究和扩展,一下将分为编码策略和解码策略分别进行阐述.

### 3.1 编码策略

编码矩阵的构造是完成 ECOC 多类分类的第一步,目前的工作主要是对现有编码方法的改进和重新提出新的编码方法.事前编码(predefined coding)、基于样本数据编码(data depended coding)(也称基于问题域编码

(problem-depended encoding)) 和基于基分类器编码 (based dichotomizes coding) 是目前主要的编码方法<sup>[6]</sup>.

事前编码是指编码不依赖样本的编码方法. 一个分类问题给出前编码矩阵就已确定, 此类编码即我们通常所熟悉的 one-versus-all、one-versus-one、随机编码等. Terry Windeatt 和 Reza Ghaderi 在对基于 ECOC 解决多类分类问题的原理进行深入分析后, 研究了一种 circular ECOC 编码方法<sup>[7]</sup>. Elizabeth 等在文献<sup>[8]</sup>中提出了一种回归纠错输出编码 (recursive ECOC). Simeone 和 Tax 针对现有经典的 ECOC 提出了一种能提高编码矩阵分类效果的快速提升法<sup>[9]</sup>, 该方法具有通用性可对任意事前给定的 ECOC 进行分类性能的提升. Schapire 提出将集成学习中的 Boosting 方法与 ECOC 方法结合构成 Adaboost.ECC 方法, 该方法综合 Boosting 和 ECOC 两者的优点, 提高了 ECOC 方法对基分类器选择的灵活性, 使得基于 Boosting 方法训练得到的二分器具有良好的互补性, 从而提高了 ECOC 的整体分类性能<sup>[10]</sup>. 此外, 文献<sup>[11]</sup>还研究了当基分类器为决策树时如何设计剪枝方法使 ECOC 的分类性能得到有效提高. 文献<sup>[12]</sup>讨论了一种新的编码设计准则—最大化三符号编码距离准则 (ternary distance), 通过该准则作者提出了一种新的稀疏编码矩阵, 实验结果表明基于该准则获得的 ECOC 分类效果有显著提高, 且与解码方法的选取相关性较小. 文献<sup>[13]</sup>基于最小化代价敏感函数 (cost-sensitive function) 来设计一种敏感纠错输出编码 (sensitive error correcting output codes). 文献<sup>[14]</sup>针对类别数较少时利用 ECOC 方法进行分类可能产生由于基分类器的决策边界相同而导致过拟合的问题, 讨论了利用多层感知器 (multi-layer perceptrons) 作为二分器来获得差异性大的基分类器, 从而能有效的减少过拟合带来的误差. 尹安容等利用 Hadamard 纠错码来简化 ECOC 的构造, 并得出相比其它编码矩阵的构造方法, 该方法的实现简单快捷, 且更容易构造出性能优越的纠错码本<sup>[15]</sup>.

上述编码方法都有一个共同的特点, 即编码矩阵的确定不依赖样本空间数据, 因此所得到的编码矩阵不能反映分类信息, 这在实际应用中将影响此类编码的应用效果, 为此 Alpaydin 和 Mayoraz 于 1999 年首次提出基于数据的反向传播编码确定方法 (back propagation algorithm)<sup>[16]</sup>. Utschick 和 Weichselberger 利用期望最大化法则 (expectation maximization algorithm) 通过对最大似然目标函数进行优化找出最适合样本空间的编码矩阵, 并得出结论: “一对多”策略仍是多数多类分类问题的最优选择<sup>[17]</sup>. 2006 年, O. Pujol 和 Radeva 等人提出一种判别式纠错输出编码方法—DECOC, 其针对样本集特征, 利用决策树结构试探性地逐步构造类间隔最大的子类, 从而确定编码矩阵<sup>[18]</sup>. 2008 年, Escalera 和 Tax 等

人提出针对样本集线性不可分问题, 提出对基类子集再分割的 Subclass ECOC 编码方法 (SECO)<sup>[19]</sup>. O. Pujol 等人提出一种基于数据的再学习编码矩阵, 在该方法中任意选定一个已知给定的编码矩阵作为初始化矩阵, 通过给定的数据集和学习方法对该编码矩阵进行学习, 使其在只增加少量基分类器个数的情况下成为一种新的基于数据的 ECOC 矩阵<sup>[20]</sup>. Bouzas 等人提出基于粒子群算法来寻找数据集最优子类群, 从而利用该子类群获得构造编码矩阵的二类划分并最终得到一种线性判别式纠错输出编码矩阵 (Linear Discriminant Error Correcting Output Codes, LDECOC)<sup>[21]</sup>. 钟国强提出了一种基于数据的编码矩阵和基分类器联合学习模型, 由问题域 (数据样本) 和基分类器分类性能共同确定编码矩阵, 基分类器在与编码矩阵的交互学习的同时得到训练<sup>[22]</sup>; 同时, 还提出将编码矩阵和基分类器的获取转化为一个以最小化数据集分类错误率为目标的最优化函数<sup>[23]</sup>. 2013 年, 为保证基分类器之间的独立性, Bagheri 提出子空间 ECOC 编码方法 (subspace ECOC), 引入特征空间形成三维的编码空间, 从而利用不同的特征子集训练基分类器<sup>[24]</sup>. S. Escalera 等针对经典的“一对一”三符号编码矩阵中符号“0”会引入分类偏差的问题, 利用训练样本分类结果对编码矩阵中的码字“0”进行再编码, 并将该分类结果作为权值融入到基于损失函数的解码过程中, 基于人脸数据的实验表明该方法能提高传统基于一对一和稀疏编码矩阵的 ECOC 分类性能<sup>[25]</sup>. Wang 等提出利用编码矩阵中二类划分的先验原始类结构信息可以提高 ECOC 分类性能, 并给出了在流形假设和聚类假设的情况下将先验结构信息融入基分类器决策函数的方法<sup>[26]</sup>. 周进登等提出以样本最小  $k$  近邻错分率作为评价准则, 把编码矩阵的构造问题转化为一个搜索问题并得到包含训练样本数据信息的编码方法<sup>[27]</sup>和基于混淆矩阵的自适应编码方法<sup>[28]</sup>.

基于基分类器编码, 即基分类器已选定, 如何找出与基分类器最优搭配的编码矩阵. Crammer 和 Singer 在 2002 年研究证明了在基分类器已选定的前提下, 计算出最优编码矩阵是 NP 难问题, 从理论上指出了基于分类器设计最优编码矩阵的困难<sup>[6]</sup>.

很多学者还研究了其他编码方法. S. Escalera 等提出了一种在线学习的 ECOC 多类分类方法<sup>[29]</sup>. M. A. Bautista 等针对码字长度最优问题, 提出了一种最小编码方法, 通过采用遗传算法提高编码和选择分类器参数的速度, 并证明对于  $N$  类分类问题, 只需要  $\lceil \log_2 N \rceil$  个基分类器 (即码字长度)<sup>[30]</sup>. 文献<sup>[31]</sup>提出了一种基于选择性集成的搜索编码方法, 通过删除初始编码矩阵中冗余和无关的列, 从而在保证 ECOC 分类精度的同

时提高基分类器的差异性.蒋艳凰等提出利用搜索编码方法,把编码矩阵的构造看成是在某一整数区间的搜索过程,通过对把编码阵每一行看成是某一整数的二进制码串,进而利用编码阵构造的约束条件搜索出满足要求的 ECOC<sup>[32]</sup>.

### 3.2 解码策略

解码规则是 ECOC 的另一研究热点,可以分为三类.第一类是基于输出编码和目标编码距离的解码策略.汉明距离解码是最常见的此类解码方法之一,其解码规则简单,准确率较高,且具有纠错能力,Dietterich 在其首次提出 ECOC 方法时就利用了该解码方法.但随着分类问题复杂程度的增加,汉明距离解码有时得不到理想答案,出现解码正确率偏低甚至完全错误的情况,Ludmila 和 Kuncheva 在文献[33]中针对此问题进行了专门研究并指出了 Hamming 距离解码在面对复杂分类问题的弊病所在,同时提出引入差异性度量作为解码的新准则.文献[6]提出利用欧式距离代替汉明距离作为融合评价函数,提出了欧式距离解码(Euclidean decoding).文献[34]利用逆汉明距离作为各类别后验概率的近似值,并基于贝叶斯决策规则提出了逆汉明距离解码(inverse hamming decoding).文献[35]将样本输出空间映射为超立方体的各个顶点,提出一种以样本与各顶点的中心距为衡量规则的解码方法,简称为中心距解码(centroid decoding rule).文献[36]引入损失函数作为类别关键字之间的距离函数,基于此提出基于损失函数解码(loss-based decoding).Sergio Escalera 提出了一种加权损失函数解码(loss-weighted based decoding)<sup>[37]</sup>.

第二类是基于概率的解码策略.基于概率的解码策略是一种贝叶斯准则,通常能获得更好的分类效果.文献[38]提出了一种改进的估计条件概率的方法,它将条件概率的求解看作  $L$  个线性超定方程的求解.在此基础上,文献[39]提出了一种构造等距编码的修正框架,它能同时降低多类概率和二类概率的损失函数.Passerini 等提出了一种基于 ECOC 的后验概率建模方法,利用所得概率估计值得到最大似然概率解码(maximizing likelihood probability decoding),但经过分析发现 Passerini 方法会造成偏差的增加<sup>[40]</sup>.文献[41]提出了一种基于一阶明可夫斯基解码方法(first order minkowski decoding rule).J Zhou 提出一种基于证据理论的解码方法,将每个二分器的输出作为一条证据,将解码过程看成是证据的融合过程,通过证据的融合不断地缩小假设集的范围,从而使解码的正确率逐步提高;同时针对 Ternary ECOC 提出了两种基于后验概率的解码方法,并利用贝叶斯法则作为解码规则的最终决策依据<sup>[42]</sup>.文献[43]将特征提取概念融入到 ECOC 解码过程中,将二分器的概率输出作为特征对目标类进行再编码,采用

向量量化法进行融合决策.针对 ECOC 产生的基分类器不能实现对某些样本拒分导致错分代价较高的问题,P. Simeon 和 C. Marrocco 引入“拒绝域”的概念,提出了两种解码方法:一是针对包含拒绝域的基分类器软输出的解码方法(internal level);二是在解码的过程引入拒绝判定的方法(external level)<sup>[44]</sup>.第三类是基于模式空间的解码策略.文献[45]把基于 ECOC 的多类分类解码过程看成是一个矩阵博弈(matrix game)的一种特殊形式,并提出利用博弈论(Game Theory)作为解码规则的理论依据从而得到最终融合结果.

## 4 应用研究

应用 ECOC 解决实际分类问题也是 ECOC 多类分类研究的一个重要方面.到目前为止,ECOC 被广泛应用到生物学数据识别、疾病诊断、故障诊断、军事目标识别、机器视觉、智能交通等诸多领域,并取得很好的识别效果.

文献[46]将 ECOC 应用于自动语音识别,通过 ECOC 分解框架将原始语音特征空间映射到新的特征空间,在新的特征空间中每一维特征都能分辨不同的电话来源,利用线性 SVM 进行判别,实验结果表明该方法能降低传统光谱特征识别 10.5% 的误差率.文献[29]将在线 ECOC 学习用于视觉机器识别,即交通信号识别和人脸识别.文献[47]把 ECOC 方法应用到医学领域中的冠状血管超声波检测中,将复杂的血管样本数据(IVUS)通过子类划分得到不同的基分类器,再利用子类纠错输出编码对其进行判别,实验结果表明该方法能对 IVUS 数据进行有效的识别.文献<sup>[48]</sup>将 Boost ECOC 应用到人脸识别中,探讨了该类型 ECOC 在该领域的实际应用效果.文献<sup>[49]</sup>将 ECOC 用于多视角人脸检测,把人脸特征按视角不同划分为不同子类,利用 ECOC 在多类分类问题的优势进行求解并得到了良好的实验效果.文献<sup>[50]</sup>基于 ECOC 构造多类支持向量机(multiclass SVM)并应用于心电图,识别成功率达到 80%.陶晓燕、姬红兵等研究了基于 SOM 解码方法,并把它利用于人脸识别实验中取得很好的结果<sup>[51]</sup>.胡莹岑等将 ECOC 应用于模拟电路故障诊断方法中,实验表明诊断精度优于传统 SVM 和 BP 神经网络方法,并降低了故障诊断时间<sup>[52]</sup>.辛轶等提出了一种新颖的基于半监督技术的层析编码算法,通过同类簇编码与类层次编码,使编码矩阵更符合数据分布的特点,实验表明新算法有效地提高了分类精度<sup>[53]</sup>.郭巍等将 AdaBoost 与 ECOC 相结合,并用于合成孔径雷达目标识别,可以明显提高 SAR 识别系统的识别性能<sup>[54]</sup>.

## 5 存在的问题及研究方向

尽管近年来已有许多学者对 ECOC 相关问题进行

了研究,但仍存在以下几点不足:

(1)对基于数据编码矩阵的确定仍然缺乏有效的方法

与事前编码方法不同,基于数据编码方法在编码矩阵的确定时通常要考虑数据本身的属性即样本特征.一个好的分解框架必定是能反映样本数据本身分布特点的框架,因此构造与所要分类样本数据最匹配的编码矩阵往往是分类成功的关键所在.目前基于数据的编码方法成为主要的编码手段,但其仅能构造适合于基分类器数据划分的编码矩阵,如何研究出能有效提高分类准确率的基于数据编码矩阵仍是目前亟待解决的问题.

(2)解码策略过于简单导致融合阶段产生较大误差,同时缺乏使用解码策略的指导方法

现有的解码策略往往是建立的投票法及其改进版之上,虽然这些解码策略都具有纠错能力,但当其应用于较为复杂的多类分类任务中时,其带来的解码误差往往超过了由其纠错能力带来的分类性能的提高.目前,基于概率的解码策略成为众多研究者的首要选择.此外,在现有的研究 ECOC 多类分类解码策略的文献中,还未发现有对如何在不同分类问题中选择最为合适的解码策略进行讨论的研究.

(3)如何评价 ECOC 编码模型的优劣以及怎样确定最优的编码解码组合

基于概率的编码解码将会包含更大的信息量,目前已有许多基于此研究的模型,但对模型本身的优劣,如模式是否能获得多类后验概率的无偏估计没有统一的评价准则,同时如何在不同分类问题中构建编码和解码方法的最优组合也是亟待解决的问题.对于同一问题的不同建模得到的编码解码方法,如何评价其优劣性、适应性和鲁棒性也是一个难点. ECOC 多类分类在编码和解码过程中还涉及到基分类器的选择问题,而基分类器的学习也会影响利用 ECOC 解决多类分类问题的最终分类效果,如何将这三者的学习统一起来也是一个需要研究的问题

(4)ECOC 与集成学习的关系

ECOC 框架和集成学习有诸多相似之处,如都包含多个基分类器,且需要对每一个基分类器进行学习,在决策过程中同样都存在对各基分类器结果融合的问题.甚至可以将 ECOC 看成是一种特殊的集成学习框架,那么针对集成学习所获得的研究结论是否同样也适用于 ECOC 框架,以及如何扩展类似这样的结论将会是非常具有价值的研究方向.

## 参考文献

[1] T G Dietterich, G Bakiri. Solving Multi-class learning problems

via error-correcting output codes[J]. Journal of Artificial Intelligence Research, 1995, 34(2): 263 - 286.

- [2] T G Dietterich, G Bakiri. Error-correcting output codes: A general method for improving multiclass inductive learning programs[A]. Proceedings of the Ninth National Conference on Artificial Intelligence[C]. Menlo Park, San Francisco: AAAI, 1991. 572 - 577.
- [3] T G Dietterich, E Kong. Error correcting output codes corrects bias and variance[A]. Proceedings of the 21th International Conference on Machine Learning[C]. San Francisco: AAAI, 1995. 313 - 321.
- [4] F Masulli G Valentini. Effectiveness of error correcting output coding methods in ensemble and monolithic learning machines[J]. Pattern Anal Applic, 2003, 65(6): 285 - 300.
- [5] 李健武, 魏海周, 宋玉龙. ECOC 多分类器实现的最小封闭球模型[J]. 计算机研究与发展, 2011, 48(Suppl): 22 - 30.  
Li Jianwu, Wei Haizhou, Song Yulong. The minimum closed sphere model for multi-class classifiers using error-correcting output codes[J]. Journal of Computer Research and Development, 2011, 48(Suppl): 22 - 30. (in Chinese)
- [6] Crammer K, Singer Y. On the learnability and design of output codes for multiclass problems[A]. Proceedings of the Thirteenth Annual Conference on Computational Learning Theory[C]. Boston: Kluwer Academic Publishers, 2000. 896 - 909.
- [7] T Windeatt, R Ghaderi. Binary labelling and decision level fusion[J]. Information Fusion, 2001, 2(1): 103 - 112.
- [8] E Tapia, P Bulacio, L Angelone. Recursive ECOC classification[J]. Pattern Recognition letters, 2010, 31(3): 210 - 215.
- [9] Paolo Simeon, David M J Tax, Robert P W Duin, Francesco Tortorella. A fast approach to improve classification performance of ECOC classification systems[J]. Computer Science Trans Structural, Syntactic, and Statistical Pattern Recognition, 2008, 43(3): 459 - 468.
- [10] Robert E. Schapire. Using output codes to boost multi-class learning problems[A]. Machine Learning: Proceedings of the Fourteenth International Conference[C]. Nashville TN: IEEE, 1997. 313 - 321.
- [11] Terry Windeatt, G Ardeshtir. Tree pruning for output coded ensembles[A]. Proceedings 16th International Conference on Pattern Recognition[C]. Quebec City, Canada: IEEE Computer Soclos Alamitos, 2002. 92 - 96.
- [12] Sergio Escalera, Oriol Pujol, Petia Radeva. Separability of ternary error-correcting output codes[A]. Proceedings 19th International Conference on Pattern Recognition[C]. Tampa, FL, United States: IEEE, 2008. 1 - 4.
- [13] John Langford, Alina Beygelzimer. Sensitive error correcting output codes[J]. Lecture Notes in Computer Science, 2005, 2459(3): 21 - 49.

- [14] Matthew Prior, Terry Windeatt. Over-fitting in ensembles of neural network classifiers within ECOC frameworks[J]. *Lecture Notes in Computer Science*, 2005, 3541(4): 286 – 295.
- [15] 尹安容, 谢湘, 匡镜明. Hadamard 纠错码结合支持向量机在多分类问题中的应用[J]. *电子学报*, 2008, 36(1): 122 – 126.  
Yi Anrong, Xie Xiang, Kuang Jingming. Application of Hadamard ECOC in multi-class problems based on SVM[J]. *Acta Electronica Sinica*, 2008, 36(1): 122 – 126. (in Chinese)
- [16] E Alpaydin, E Mayoraz. Learning error-correcting output codes from data[A]. *Proceedings of International Conference on Artificial Neural Networks (ICANN'99) (Volume: 2)* [C]. Edinbuergh: IEEE, 1999. 743 – 748.
- [17] W Utschick, W Weichselberger. Stochastic organization of output codes in multiclass learning problems [J]. *Neural Comput*, 2001, 13(5): 1065 – 1102.
- [18] O Pujol, P Radeva, J Vitria. Discriminate ECOC: A heuristic method for application dependent design of error correcting output codes[J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2006, 28(6): 1001 – 1007.
- [19] S Escalera, David M J Tax, O Pujol, P Radeva, Robert P W Duin. Subclass problem-dependent design for error-correcting output codes[J]. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 2008, 30(6): 1041 – 1054.
- [20] O Pujol, Escalera S., Radeva P. An incremental node embedding technique for error correcting output codes[J]. *Pattern Recognition*, 2008, 41(2): 713 – 725.
- [21] Dimitrios Bouzas, Nikolaos Arvanitopoulos, Anastasios Tefas. Optimizing linear discriminant error correcting output codes using particle swarm optimization[J]. *Lecture Notes in Computer Science*, 2011, 6792(4): 79 – 86.
- [22] Guoqiang Zhong, et al. Joint learning of error-correcting output codes and dichotomizers from data[J]. *Neural Comput & Applic*, 2012, 21(4): 715 – 724.
- [23] Guoqiang Zhong, Kaizhu Huang, Cheng-Lin Liu. Learning ECOC and dichotomizers Jointly from data[J]. *Lecture Notes in Computer Science*, 2010, 6443(3): 494 – 502.
- [24] M A Bagheri, GA Montazer. A subspace approach to error correcting output codes[J]. *Pattern Recognition Letters*, 2013, 34(1): 176 – 184.
- [25] Sergio Escalera, O Pujol. Re-coding ECOCs without re-training[J]. *Pattern Recognition Letters*, 2013, 31(5): 555 – 562.
- [26] Yunyun Wang, Songcan Chen, Hui Xue. Can under-exploited structure of original-classes help ECOC-based multi-class classification[J]. *Eurocomputing*, 2012, 89(15): 158 – 167.
- [27] 周进登, 王晓丹. 基于最小  $k$  近邻错分率编码确定方法及其在多类分类中的应用[J]. *控制与决策*, 2010, 26(9): 1296 – 1302.  
Zhou Jindeng, Wang Xiaodan. Designing of output codes based on minimal  $k$  nearest neighbor classifying error and its application in multi-class classification[J]. *Control and Decision*, 2010, 26(9): 1296 – 1302. (in Chinese)
- [28] 周进登, 王晓丹. 基于混淆矩阵的自适应纠错输出编码多类分类方法[J]. *系统工程与电子技术*, 2012, 34(7): 220 – 226.  
Zhou Jindeng, Wang Xiaodan. Multiclass classification of adaptive error-correcting output codes based on confusion matrix[J]. *Systems Engineering and Electronics*, 2012, 34(7): 220 – 226. (in Chinese)
- [29] S Escaler, David Masip, et al. Online error correcting output codes[J]. *Pattern Recognition Letters*, 2011, 32(3): 458 – 467.
- [30] Miguel Angel Bautista, Sergio Escalera, et al. Minimal design of error-correcting output codes[J]. *Pattern Recognition Letters*, 2012, 33(10): 693 – 702.
- [31] Nima Hatami. Thinned-ECOC ensemble based on sequential code shrinking[J]. *Expert Systems with Application*, 2012, 39(1): 936 – 947.
- [32] 蒋艳凰, 赵强利, 杨学军. 一种搜索编码法及其在监督分类中的应用[J]. *软件学报*, 2005, 45(6): 1081 – 1088.  
Jiang Yan-Huang, Zhao Qiang-Li, Yang Xue-Jun. A search coding method and its application in supervised classification [J]. *Journal of Software*, 2005, 45(6): 1081 – 1088. (in Chinese)
- [33] Ludmila I, Kuncheva. Using diversity measures for generating error-correcting output codes in classibler ensembles[J]. *Pattern Recognition Letters*, 2005, 26(5): 83 – 90.
- [34] T Windeatt, R Ghaderi. Coding and decoding strategies for multi-class learning problems[J]. *Information Fusion*, 2003, 43(4): 11 – 21.
- [35] James G. Majority Vote Classifiers: Theory and Applications [D]. CA: Stanford University, 1998.
- [36] Erin LAllwein, Robert E Schapire, Yoram Singer. Reducing multiclass to binary: A unifying approach for margin classifiers[A]. *Machine Learning: Proceedings of the Seventeenth International Conference* [C]. Palo Alto: Stanford University, 2000. 1545 – 1550.
- [37] S Escalera, O Pujol, P Radeva. On the decoding process in ternary error-correcting output codes[J]. *IEEE Trans Pattern Analysis and Machine Intelligence*, 2010, 32(1): 120 – 134.
- [38] E B Kong, T G Diettrich. Probability estimation via error correcting output coding[A]. *International Conference of Artificial Intelligence and Soft Computing* [C]. Banff, Canada: IEEE, 1997.
- [39] Reza Ghaderi, Terry Windeatt. Least squares and estimation measures via error correcting output code [A]. *2nd International Workshop on Multiple Classifier Systems* [C]. Cambridge, UK: Springer-Verlag, 2001. 148 – 157.

- [40] A Passerini, M Pontil, P Frasconi. New results on error correcting output codes of Kernel machines[J]. IEEE Trans Neural Networks, 2004, 15(1): 45 – 54.
- [41] Kittler J, Ghaderi R, Windeatt T, Matas J. Face identification and verification via ECOC[A]. Third Int Conf Audio and Video Based Biometric Person Authentication[C]. Halmstad, Sweden: IEEE, 2001. 1 – 13.
- [42] Zhou Jindeng, Wang Xiaodan, et al. Decoding design based on conditional probabilities in ternary error-correcting output codes[J]. Pattern Recognition, 2012, 45(4): 1342 – 1351.
- [43] Guoqiang Zhong, Cheng-Lin Liu. Error-correcting output codes based ensemble feature extraction[J]. Pattern Recognition, 2013, 46(4): 1091 – 1100.
- [44] P Simeon, C Marrocco, F Tortorella. Design of reject rules for ECOC classification systems[J]. Pattern Recognition, 2012, 45(2): 863 – 875.
- [45] Mikhail Petrovskiy. Probability estimation in error correcting output coding framework using game theory [J]. Lecture Notes in Computer Science, 2005, 3809(5): 186 – 196.
- [46] Omid Dehzangi, Bin Ma, et, al. Discriminative feature extraction for speech recognition using continuous output codes[J]. Pattern Recognition Letters, 2012, 33(13): 1703 – 1709.
- [47] Sergio Escalera, Oriol Pujol, Josepa Mauri. IVUS tissue characterization with subclass error-correcting output codes [J]. Computer Vision and Pattern Recognition, 2008, 34(5): 1 – 8.
- [48] T Windeatt, G Ardeshir. Boosted ECOC ensembles for face recognition[J]. Pattern Recognition, 2003, 35(4): 165 – 168.
- [49] Hongming Zhang, Wen Gao. Robust multi-view face detection using error correcting output codes[J]. Lecture Notes in Computer Science, 2006, 3954(11): 1 – 12.
- [50] Elif Deryableyli. ECG beats classification using multiclass support vector machines with error correcting output codes[J]. Digital Signal Processing, 2007, 45(17): 675 – 684.
- [51] 陶晓燕, 姬红兵. 一种基于 SOM 解码的多类支持向量机 [J]. 系统工程与电子技术, 2006, 24(4): 1147 – 1151.  
Tao Xiao-yan, Ji Hong-bing. Multi-class SVM based on SOM decoding[J]. Systems Engineering and Electronics, 2006, 24(4): 1147 – 1151. (in Chinese)
- [52] 胡莹岑, 王友仁, 崔江. 模拟电路的一种纠错码 SVM 诊断方法 [J]. 计算机辅助设计与图形学学报, 2011, 23(11): 1931 – 1937.  
Hu Yingcen, Wang Youren, Cui Jiang. Method of analog circuit diagnosis based on ECOC and SVM[J]. Journal of Computer-Aided Design & Computer Graphics, 2011, 23(11): 1931 – 1937. (in Chinese)
- [53] 辛轶, 郭躬德, 陈黎飞. 半监督层次纠错输出编码算法 [J]. 小型微型计算机系统, 2010, 31(8): 1660 – 1664.  
Xin Yi, Guo Gongde, Chen Lifei. Semi-supervised Based hierarchical ECOC algorithm [J]. Journal of Chinese Computer Systems, 2010, 31(8): 1660 – 1664. (in Chinese)
- [54] 郭巍, 张平, 朱良. 基于 AdaBoost. ECOC 的合成孔径雷达图像目标识别研究 [J]. 哈尔滨工程大学学报, 2010, 31(2): 232 – 236.  
Guo Wei, Zhang Ping, Zhu Liang. Research on synthetic aperture radar image target recognition based on AdaBoost ECOC [J]. Journal of Harbin Engineering University, 2010, 31(2): 232 – 236. (in Chinese)
- [55] 周进登, 王晓丹. 加权解码在解决纠错输出编码 consistent-diverse 平衡问题的应用 [J]. 电子学报, 2011, 39(7): 1514 – 1522.  
Zhou Jindeng, Wang Xiaodan, et al. Application of weighted decoding for the consistent-diverse balance problem of error correcting output codes [J]. Acta Electronica Sinica, 2011, 39(7): 1514 – 1522. (in Chinese)

#### 作者简介



雷 蕾 女, 1988 年生于四川南充, 博士生, 研究方向为智能信息处理和目标识别。  
E-mail: wendyandpaopao@163.com



王晓丹 女, 1966 年生于陕西汉中, 教授, 博士, 研究方向为模式识别, 机器学习等。  
E-mail: afeu\_w@163.com



周进登 男 1984 年生于江西鹰潭, 博士, 研究方向为模式识别, 智能信息处理和多传感器数据融合  
E-mail: zhousjin198417@163.com