

A Vector Machine (R-SVM) Method for Mold Detection of Tobacco Leaves in Silk Processing Based on Improved Radial Basis Functions

Yinxiang Huang

Beijing Yuanshan Intelligent Technology Co., Ltd., Beijing, 100010, China

Abstract

Tobacco leaves are prone to mold during storage and processing, which can seriously affect the quality and taste of tobacco and may also pose a threat to human health. This paper conducts an in-depth study on the tobacco mold detection method based on improved radial basis function vector machine (R-SVM). During this period, the mold detection method for shredded tobacco leaves is introduced in detail. By optimizing the RBF kernel function parameters, the accuracy and efficiency of identifying moldy tobacco leaves are comprehensively enhanced. And with the help of experimental analysis, it was clarified that the improved R-SVM performed better in detecting moldy tobacco leaves, providing effective technical support for quality control in the tobacco industry and helping related industries to improve economic benefits.

Keywords

radial basis function; vector machine; silk tobacco leaves; mold detection

基于改进径向基函数的向量机 (R-SVM) 制丝烟叶霉变检测方法

黄银祥

北京远舫智能科技有限公司, 中国·北京 100010

摘要

烟叶在储存和处理过程中容易发生霉变, 这种情况会严重影响烟草的品质和口感, 还可能对人体健康造成威胁。论文对基于改进径向基函数的向量机 (R-SVM) 的烟叶霉变检测方法进行深入研究, 在此期间详细介绍制丝烟叶霉变检测方法, 通过优化RBF核函数参数, 全面增强霉变烟叶识别的精确度和效率。并借助实验分析加以了解, 明确改进后的R-SVM在霉变烟叶检测中表现更为出色, 为烟草行业质量控制提供了有效的技术支持, 以此帮助相关产业提供帮助, 提高经济效益。

关键词

径向基函数; 向量机; 制丝烟叶; 霉变检测

1 引言

烟草行业需要及时准确地检测烟叶霉变。而传统的霉变检测主要依赖于人工观察和经验判断, 存在效率低, 以及准确性易受主观因素影响等问题。此方面应借助机器学习和图像处理技术, 实现自动化、智能化的霉变检测方法。

2 支持向量机 (SVM)

支持向量机 (SVM) 核心理念在于寻找最优的超平面。其目标是最大化不同类别之间的间隔, 从而实现高效的分类效果。在实际应用中, 会遇到线性不可分的问题, 对此

SVM引入核函数的概念。核函数的作用是通过非线性变换, 将原本线性不可分的数据映射到更高维的空间中, 使得数据在新的空间里变得可分。这种变换能够有效增强 SVM 的分类能力, 还可以极大地扩展其应用范围。在众多核函数中, 径向基函数 (RBF) 核无疑是最常用一种。RBF 核通过复杂的非线性变换, 将样本从原始空间映射到高维空间。在此空间中, 原本纠缠不清、难以划分的数据点, 变得井然有序, 清晰可分。通过这种方式, SVM 在各种复杂的情况下, 都能展现出卓越的分类效果。当面对需要高效、准确分类的任务时, 支持向量机 (SVM) 中的径向基函数 (RBF) 核引入, 让 SVM 在处理复杂数据时变得简单, 使得 SVM 在机器学习领域中脱颖而出, 成为众多研究者和工程师的首选工具。值得一提的是, SVM 的应用并不仅限于分类问题。随着技术的不断发展和研究的深入, SVM 还在回归分析、异常检

【作者简介】黄银祥 (1984-), 男, 中国江苏泰州人, 硕士, 高级工程师, 从事深度学习研究。

测等领域展现出了巨大的潜力^[1]。

而 R-SVM 算法是一种基于支持向量机 (SVM) 的回归算法,其核心是找到最优超平面,使得所有样本点到该平面的距离之和最小。此距离被称为间隔,而 R-SVM 算法的目标是最大化这个间隔。为了实现这一目标, R-SVM 算法引入一个损失函数,用于衡量预测值与实际值之间的误差。在 R-SVM 算法中,每个样本点都有对应的松弛变量,用于表示该样本点到最优超平面的距离。如果样本点位于最优超平面的上方或下方,则其松弛变量为正值;如果样本点位于最优超平面上,则其松弛变量为零。R-SVM 算法的目标是最小化所有样本点的松弛变量之和,从而最大化间隔。为了求解 R-SVM 算法,通常采用迭代优化的方法,如梯度下降法或牛顿法。此类方法通过不断调整最优超平面的位置和松弛变量的值,逐步逼近最优解。R-SVM 算法的一个重要参数是正则化参数 C,其用于平衡模型的复杂度和拟合误差。较大的 C 值会使模型更复杂,但可能导致过拟合;较小的 C 值会使模型更简单,但可能导致欠拟合。除了正则化参数 C 外, R-SVM 算法还需要调整核函数的选择和核函数的参数。核函数用于将原始特征空间映射到高维空间,以便更好地拟合非线性关系。常见的核函数包括线性核、多项式核和径向基函数 (RBF) 核等。选择合适的核函数和参数提高 R-SVM 算法的预测性能。在使用 R-SVM 算法时,还需要注意数据的预处理和特征选择。数据预处理可以包括标准化、去噪和缺失值处理等操作,以减少数据中的噪声和异常值对模型的影响。特征选择则是从原始特征中选择最具有代表性的特征,以提高模型的泛化能力。

3 径向基函数 (RBF)

对采集的光谱数据 X 进行预处理,以获得更加准确的特征波长数据。令 $M = \{M_1, M_2, \dots, M_i, \dots, M_n\}$, M 表示第 i 个烟叶样本的光谱的初始权重, $1 \leq i \leq n$, 并有 $M_i = \{M_{i1}, M_{i2}, \dots, M_{i1}, M_{i2}, \dots, M_{im}\}$, $1 \leq j \leq m$, M_j 表示第 i 个烟叶样本第 j 个光谱特征的初始权重数值。

求解特征波长权重,利用如下公式构建最优化目标函数:

$$\min_M \left\{ \sum_{i=1}^n M_i \left(\frac{1}{2} \prod X_i - \mu_i \prod^2 + \lambda \cdot \text{Regularizer}(M) \right) \right\}$$

其中, M 为波长权重矩阵; M_i 为第 i 个烟叶特征波长的权重; X_i 为第 i 个样本的光谱数据; Π 为 x 的均值; λ 为正则化参数; Regularizer(W) 为对 M 的正则化项。

RBF 核的局部性特点在高维空间中,如果两个数据点在原始空间中距离较近,在映射到高维空间后也会保持相近。这种特性使得 RBF 核能够有效地处理非线性关系和复杂模式的数据。

4 烟叶霉变的相关研究

在烟叶质量检测领域,化学分析和图像处理技术是两

种主要的检测方法。化学分析方法虽然准确度高,但其操作繁琐且耗时长,难以实现实时检测。相比之下,图像处理技术因其快速和非破坏性的特点而逐渐受到重视。然而,传统的图像处理方法主要依赖于手工提取的特征和专家知识,这种方式限制了其在复杂多变环境下的应用效果。

为了解决上述问题,不断探索新的图像处理技术。深度学习技术因其强大的特征学习能力而被广泛应用于图像处理领域。在烟叶霉变的检测中,深度学习模型能够自动从大量图像数据中学习霉变的特征,并实现端到端的检测。这种方法提高了检测的自动化程度,并提升了在复杂环境下的应用效果。此外,还尝试将多种传感器数据与图像数据相结合,以进一步提高烟叶霉变检测的准确性和鲁棒性。通过融合红外、紫外等多光谱图像数据,更全面地捕捉到烟叶表面的细微变化,加强霉变检测的灵敏度和特异性。在此期间,结合温度、湿度等环境传感器数据,可以更准确地判断烟叶霉变的程度和发展趋势。除了改进检测方法外,研究人员还致力于开发新的预防和控制烟叶霉变的策略。通过研究霉变微生物的生长特性和影响因素,可以制定出更有效的烟叶储存和运输条件,降低霉变的风险。利用生物防治技术,如引入天敌微生物或使用抗菌剂,也可以在一定程度上抑制霉变微生物的生长,保护烟叶的质量^[2]。

5 烟叶霉变的快速识别方式

5.1 近红外光谱

5.1.1 数据采集

不同红外光谱仪的数据采集范围不同,其中 BRUKER 公司的 MPA 型傅立叶变换近红外光谱仪主要设定光谱采集的范围为 $4000 \sim 12000 \text{cm}^{-1}$ 。此范围内的光谱信息能够提供关于烟叶样本的重要特征。为保证光谱数据的精确性,应选择分辨率为 8cm^{-1} ,因此相邻两个数据点之间的波长差为 8cm^{-1} 。这样的分辨率能够保证获得足够的细节信息,同时又不会过于烦琐。在此实验中,选择扫描次数为 64 次。通过多次扫描,增强光谱数据的信噪比,更准确地反映烟叶样本的真实情况。每次扫描都会记录下烟叶样本在特定波长下的吸收强度,这些数据将用于后续的分析。在准备好光谱采集参数后,将不同霉变程度的烟叶样本装入样品杯中。每个样本都要经过仔细的处理和准备,以确保其代表性。然后使用近红外光谱仪对每个样本进行光谱采集。此步骤将记录下每个烟叶样本在近红外波段内的吸收特性,作为其基础光谱信息。为了提高数据的可靠性,需要对每个样品进行两次重复装样测定。此过程要将同一个烟叶样本重新装入样品杯中,并再次进行光谱采集。通过比较两次测定的结果,评估数据的一致性,并排除可能的误差。再计算每个样品的平均结果作为最终光谱。此步骤涉及将两次测定的数据进行平均处理,以获得更为准确和稳定的光谱曲线。最终光谱将代表每个烟叶样本在近红外波段内的特征吸收情况,为后续的分析和研究

做出保障。

5.1.2 光谱预处理

在分析近红外光谱数据时，为了提高信号的质量和准确性，需要对数据进行预处理以减少系统噪音。其中离散小波变换（DWT）是一种有效的预处理方法，其通过将信号分解为不同尺度和频率的小波子空间实现。小波变换的核心是将信号分解为多个层次的小波子空间，每个子空间对应不同的尺度和频率。这种分解使得小波变换能够灵活地提取信号中的特征，相较于其他信号预处理方法具有极大优势。在实际应用中，根据波形或长度选择合适的母小波。

通过小波变换，原始信号被分解为两部分：低频信号和高频信号。低频信号代表信号的主要趋势和近似信息，而高频信号则包含信号的细节信息。这两个部分分别对应于小波变换中的近似系数和细节系数。近似系数描述信号的整体趋势和大致形状，主要通过低通滤波获得。这些系数反映信号的主要特征，如峰值、谷值和趋势变化等。通过对近似系数的分析，可以了解信号的基本特性，并为后续的信号处理提供基础。细节系数则描述信号中的高频成分，即信号的细节信息。这些系数通常通过高通滤波得到，其反映信号中的快速变化和局部特征。细节系数对于信号的精确分析和特征提取包含信号中的微小变化和瞬时现象。在近红外光谱数据的预处理过程中，离散小波变换的应用可以提高信号的信噪比，减少噪声干扰，并增强信号的特征。通过对近似系数和细节系数的分析，可以更准确地识别和提取光谱数据中的关键信息，从而提高后续分析的准确性和可靠性^[3]。

通过小波变换，可以得到近似系数和细节系数，其中包含原始光谱的主要信息和一些噪声或冗余信息。通过适当的算法，去除或压缩这些噪声和冗余信息，得到一组新的近似系数和细节系数。运用新的系数来重构光谱，得到更加清晰、准确的光谱信号。在重构的光谱和独立变量之间，建立关联模型。与此同时，也可以直接使用小波分解得到的小波系数作为变量建立模型。这种方法的基本思路是，通过小波变换，得到一组小波系数，以此直接反映原始光谱的特征。再直接用这些小波系数作为变量，建立起关联模型。这种方法的优点是方便省时，主要由于其避免重构光谱的步骤，直接使用小波系数。因此，这种方法在实际中应用得较多。

5.2 R-SVM 制丝烟叶霉变检测性能评估

经过上述的数据收集与预处理，将提取到的特征输入到 R-SVM 模型中，可得到烟叶是否霉变的预测结果。在此基础上，为了全面评价 R-SVM 模型在此任务中的有效性，在检测过程中，主要采取将其与其他机器学习算法进行比较的方法。通过一系列对比实验，发现 R-SVM 模型展现出其独特的优势。特别是在处理非线性数据方面，R-SVM 模型

表现出较强的能力，这种情况表现出其能够更好地捕捉数据中的复杂关系。当面对含有噪声的数据时，R-SVM 模型也显示出较好的鲁棒性，此种情况相比许多机器学习算法进一步做出难以克服的挑战。在此基础上，为了提升烟叶霉变检测的准确性，将 R-SVM 模型与深度学习算法相结合的检测方式。在此过程中，卷积神经网络（CNN）在图像处理领域具有卓越表现，因此将其选用为烟叶图像的特征提取检测方式。卷积神经网络（CNN）能够有效地从烟叶图像中提取出关键特征，这些特征直接关系到后续的分类任务。接下来，将 CNN 提取到的特征输入到 R-SVM 模型中进行分类处理。这种结合检测策略不仅充分利用深度学习算法在图像特征提取方面的强大能力，还有效结合 R-SVM 模型在分类任务中的高效性能。通过这种结合方法，观察到烟叶霉变检测的效果得到大幅度显著增强。这种情况说明通过合理结合不同的算法，可以充分发挥各自算法的优势，从而达到提高整体检测准确性的目的。这种发现为烟叶霉变检测技术的研究提供了全新的思路和方法。

R-SVM 模型在烟叶霉变检测任务中展现出其独特的性能优势，特别是在处理非线性数据和噪声数据方面。同时，通过与深度学习算法的结合，尤其是利用卷积神经网络进行特征提取，再将提取到的特征用于 R-SVM 模型进行分类，能够进一步提高烟叶霉变检测的准确性。这种跨算法的结合策略，不仅证明了不同算法之间的互补性，也为未来的研究和应用提供宝贵的参考。因此，继续探索和优化这种结合策略，将对烟叶霉变检测技术的发展产生深远的影响，不断加强烟叶霉变检测的效率与速度，促使其相关工作能够快速得到烟叶霉变的相关数据，以此获得更大的经济效益。

6 结论

综上所述，基于改进径向基函数的向量机（R-SVM）为烟叶霉变检测提供一种全新的方式。此过程通过优化 RBF 核函数的参数，R-SVM 在实验中显示出较传统方法更优的霉变检测性能。此方法可以有效提高识别准确率，而且大幅度增强模型对复杂数据的处理能力。将此技术广泛应用，能够帮助烟草及其他农产品领域整体加工质量获得提升。

参考文献

- [1] 付慧,史明明,李双伟,等.基于变异系数和支持向量机的含DG台区线损智能诊断研究[J].电气自动化,2024,46(3):100-103.
- [2] 尚宇辉,孟伟,房健,等.改进径向基函数插值法的多聚焦图像滤波融合[J].计算机仿真,2024,41(2):222-226.
- [3] 陈兆麟,张奕,王诗羽,等.基于电子鼻技术的烟叶霉变自动检测设备的研发与应用[J].信息与电脑(理论版),2022,34(22):118-122.