

# 骨关节炎软骨的近红外光谱学研究及分期诊断

符娟娟, 马丹英, 唐金兰, 包一麟, 赵远, 尚林伟, 尹建华\*

南京航空航天大学自动化学院生物医学工程系, 江苏 南京 211106

**摘要** 骨关节炎是一种威胁中老年人公共健康和生活质量的重大医学疾病。骨关节炎的早期病变主要表现在细胞外基质成分含量的变化, 患者自身很难发现, 现有的临床方法和实验方法也不能较准确地识别骨关节炎的早期病变。近年来, 傅里叶变换近红外(FTNIR)光谱技术因为其分析速度快、成本低、易于穿透组织获得样本的光谱信息等特点已被用于手术导航、无损检测和疾病诊断等各个领域。基于以上优势, 采用FTNIR技术对不同深度分区(表层区、过渡区、深层区)的健康和骨关节炎的关节软骨进行NIR光谱采集和预处理, 结合主成分分析(PCA)和Fisher判别(FDA)分别研究不同的预处理方式对判别结果的影响、不同深度下基质成分含量的变化以及骨关节炎分期识别。比较其他2种(基线校正、二阶导数3次多项式25点Savitzky-Golay平滑)预处理方式, 同分区中一阶导数2次多项式21点Savitzky-Golay平滑预处理的判别结果最优, 其中表层区的识别率高达95%(初始案例)和90%(交互验证案例); 表层区的判别结果优于过渡区, 更优于深层区, 恰可证明骨关节炎的早期病变主要发生在表层区。在骨关节炎分期识别中, 经数据优化后模型的初始案例识别率为100%, 交互验证识别率为93.3%, 预测集识别率为87.5%。结果表明: NIR光谱的一阶导数预处理结合PCA-FDA方法能有效地鉴别关节软骨病变与否并进行骨关节炎的分期诊断, 对骨关节炎监测和早期诊断研究具有重要意义, 并有潜力应用于骨关节炎的原位分期和早期临床诊断。

**关键词** 骨关节炎; NIR光谱; 导数预处理; 主成分分析; Fisher判别; 分期诊断

**中图分类号:** O433 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2021)09-2770-06

## 引言

关节软骨是覆盖于骨关节表面的半透明、光滑的结缔组织, 具有弹性和韧性, 能减少关节面间的摩擦、承受高负荷和缓冲震动<sup>[1]</sup>。关节软骨主要由软骨细胞(占总体积的1%~2%)和软骨基质(占总体积的98%~99%)组成<sup>[2]</sup>。软骨基质从外表面到软骨下骨之间具有明显的板层状结构, 一般分为3层: 表层区(SZ, 软骨总厚度的10%)、过渡区(TZ, 软骨总厚度的10%)、深层区(DZ, 软骨总厚度的80%)<sup>[3]</sup>。基质中主要含有水、胶原蛋白(Collagen)、蛋白多糖(PG)和少量无机盐离子<sup>[4]</sup>。其中胶原蛋白排列成网架结构, 维持软骨的结构和形状, 并能有效固定PG<sup>[5]</sup>。而PG可以保证关节软骨的弹性和耐压性。在3个分区中, 深度不同其基质主成分结构和含量也不同<sup>[3]</sup>。肥胖、增龄、负伤和过度使用等因素都可能会导致关节软骨整体结构和成分含量的变化, 从而导致骨关节炎(OA)<sup>[6]</sup>。OA的初级病变主要表现在细胞外基质成分

含量和软骨细胞形态的变化, 在临床上的表现并不明显, 且现有的临床方法和实验方法并不能较准确地识别OA的早期病变<sup>[7-10]</sup>, 对OA的早期诊断造成极大困难。

傅里叶变换近红外(FTNIR)光谱技术近年来因为其分析速度快、成本低、易于穿透组织并含有对应组织成分信息等特点而发展应用较为迅速, 已被用于手术导航、无损检测和疾病诊断等各个领域。NIR光谱范围内的光谱吸收主要来自O—H, C—H, N—H和S—H键泛频的伸缩振动, 不同的基团或是同一基团在不同的化学环境中NIR吸收的波长和强度存在差异<sup>[11]</sup>。关节软骨基质成分中的胶原蛋白和蛋白多糖中含有近红外光可探测的大分子键(O—H, C—H, N—H和S—H), 使得NIR光谱学成为探测关节软骨微观和宏观变化合适的光学技术<sup>[12-13]</sup>。然而, 样品的状态、光的散射、仪器等因素会使NIR光谱产生一些与待测样品无关的干扰, 导致NIR光谱的基线漂移或重复性不强等问题, 因此对原始光谱进行预处理非常必要。同样, 光谱预处理选择会很大程度的影响预测的准确性, 故选择合适的预处理方法至关

收稿日期: 2020-07-19, 修订日期: 2020-10-27

基金项目: 国家自然科学基金项目(61378087)和江苏省自然科学基金项目(BK20151478)资助

作者简介: 符娟娟, 女, 1994年生, 南京航空航天大学自动化学院硕士研究生 e-mail: 15951993619@163.com

\* 通讯作者 e-mail: yin@nuaa.edu.cn

重要。关节软骨主成分胶原蛋白、蛋白多糖和水的特征谱带在  $5\ 300\sim 4\ 000\ \text{cm}^{-1}$  范围内<sup>[14]</sup>, 将此波段的光谱进行分析并与化学计量学算法相结合有利于实现快速的样本分类和识别<sup>[15]</sup>。

主成分分析(PCA)是一种通过降维技术从多个原始数据提取重要信息的分析方法。为了尽量减少变量和研究对象数据信息的损失, 选取累计方差贡献率达 85% 以上的主因子作主成分分析<sup>[16]</sup>。Fisher 判别(FDA)通过方差分析建立判别函数, FDA 的原则是使类别间的分散尽可能大, 类别内的分散尽可能小。通过比较分类中心和样本之间的距离, 快速对数据集进行分类识别<sup>[17]</sup>。

本研究采用 FTNIR 技术对不同分区(SZ, TZ, DZ)健康和病变的关节软骨进行分析, 结合 PCA-FDA 方法挑选出最佳光谱预处理方法和识别分区, 基于该最佳预处理方法将 PCA-FDA 应用于 OA 不同时期的有效识别。

## 1 实验部分

### 2.1 样本制备

用于 NIR 光谱分析的所有比格犬的膝关节样本, 均由江苏南京亚东实验动物研究中心提供, 并经伦理审查机构批准(SCXK 苏 2016-0009)。选择 9 只健康比格犬, 对其中 3 只(Y7, Y8, Y9)的左后腿膝关节胫骨平台进行原位提取, 从而获取健康样本; 其余 6 只进行任一单后腿的膝关节前交叉韧带横切(ACL)手术后<sup>[18]</sup>, 分别培养 3 个月(OA-3M; Y1, Y2, Y3)和 7 个月(OA-7M; Y4, Y5, Y6), 再从相应手术侧的后腿膝关节胫骨平台提取样本。从实验犬获得关节软骨样本后, 用生理盐水冲洗, 然后将关节软骨切割成大小为  $2\ \text{mm}\times 2\ \text{mm}\times 2\ \text{mm}$  的小块样本, 再经去离子水包埋及液氮速冻后, 用低温切片机(Leica CM 1950, Germany)沿着平行于软骨表面的方向将软骨样本以  $50\ \mu\text{m}$  的厚度进行连续切片, 并将切片移取到直径为 13 mm, 厚 2 mm 的硒化锌(ZnSe)晶片上(江苏南晶红外光学仪器有限公司), 然后将样本置于空气中风干 2 h 再进行 NIR 光谱数据采集。

软骨切片样本信息如表 1 所示, 实验所选用的比格犬关节软骨切块厚度低于  $1\ 050\ \mu\text{m}$ 。因实验中 3, 4 和 11 号切块深度为  $50\ \mu\text{m}$ (最表层)的切片样本缺失, 未进行光谱检测。

### 1.2 NIR 光谱数据采集及数据分析

采用傅里叶变换红外光谱仪(Vertex 70, Bruker 公司)及配套的 OPUS 7.0 软件对 ZnSe 晶片上样本以透射模式进行 NIR 光谱数据采集, 光谱采集范围  $10\ 000\sim 4\ 000\ \text{cm}^{-1}$ , 光谱分辨率  $8\ \text{cm}^{-1}$ , 扫描 16 次。在 OPUS 软件中采用计算相对积分面积表示基线校正后的特征谱带的吸光度。OA-3M 组中切块编号为 1—3 和 5—6; OA-7M 组中切块编号为 7—10; 健康组切块编号为 12—15 和 18 的切片(214 组)光谱数据用于健康和 OA 的分类分析, 研究不同预处理方法对健康和病变样本识别率的影响。在导数预处理不同参数结合 PCA-FDA 的结果中, 发现一阶导数中 2 次多项式 21 点 Savitzky-Golay 平滑和二阶导数中 3 次多项式 25 点 Savitzky-

Golay 平滑的结果最优。故本研究先采用 Unscrambler X 软件(CAMO Software, Inc., Woodbridge, NJ)对获得的光谱数据分别进行基线校正(BC)、一阶导数 2 次多项式 21 点 Savitzky-Golay 平滑(1<sup>st</sup>-2-21SG)和二阶导数 3 次多项式 25 点 Savitzky-Golay 平滑(2<sup>nd</sup>-3-25SG)三种预处理。再使用 IBM SPSS Statistics 20 软件对光谱数据进行主成分分析(PCA), 并选择适当主成分因子数进行 Fisher 判别(FDA)。

表 1 关节软骨水平切片的样本信息  
Table 1 Sample information of horizontal section of articular cartilage

分组	实验犬编号	软骨切块编号	切片深度/ $\mu\text{m}$	切片数量/个
OA-3M	Y1	1	50~800	16
		2	50~700	14
	Y2	3	100~1 050	20
		4	100	1
	Y3	5	50~550	11
		6	50~700	14
OA-7M	Y4	7	50~950	19
		8	50~950	19
	Y5	9	50~900	18
		10	50~700	14
	Y6	11	100	1
		12	50~850	17
健康	Y7	13	50~1 050	21
		14	50~600	12
		15	50~450	9
	Y8	16	50	1
		17	50	1
	Y9	18	50~500	10

为更进一步对 OA 样本进行早期准确识别, 在预测时增加样本的数量, 仅对 SZ 进行样本切片。OA-3M 组中 1—6 号切块; OA-7M 组中 7—11 号切块和健康组 12—18 号切块的光谱数据(24 组数据)用于 OA 的分期分析, 其中包括 9 组健康和 15 组 OA 样本的光谱数据。在 PCA-FDA 分类处理过程中, 在健康、OA-3M 和 OA-7M 光谱数据中随机选择 6, 5 和 5 组(共 16 组)作为 FDA 的训练集, 其余 8 组光谱数据作为预测集进行分类识别。

## 2 结果与讨论

### 2.1 光谱分析

图 1(a), (b) 和 (c) 分别是经 BC 后的健康和 OA 样本(3M, 7M)在不同分区(SZ, TZ, DZ)的近红外吸收平均光谱。可发现, 健康和病变的关节软骨在不同深度下的光谱谱形基本相似且具有相同的特征峰。其中  $5\ 150\ \text{cm}^{-1}$  谱带的主要贡献来自结合水和自由水;  $4\ 130\sim 4\ 460\ \text{cm}^{-1}$  特征谱带主要贡献来自胶原和蛋白多糖综合成分;  $4\ 050, 4\ 610$  和  $4\ 898\ \text{cm}^{-1}$  谱带主要来自胶原蛋白<sup>[14]</sup>。

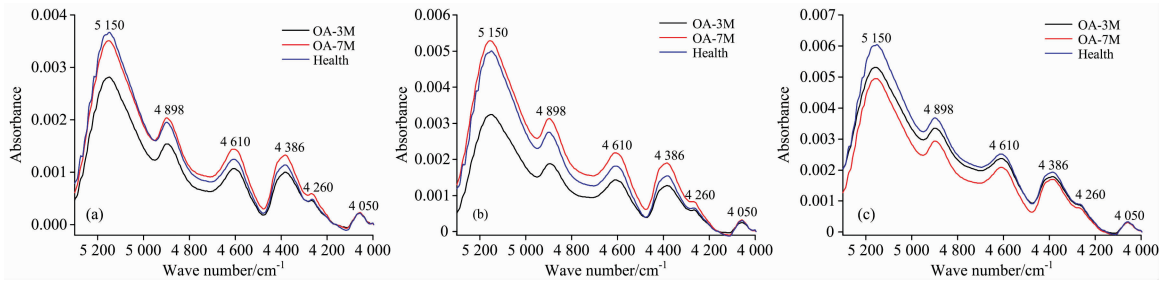


图 1 健康和多期病变样本的 SZ(a), TZ(b), DZ(c) 的 NIR 平均光谱图

Fig. 1 The near-infrared average spectra of healthy and multi-stage OA tissues at SZ(a), TZ(b), DZ(c), respectively

图 2 表示健康和多期病变样本在不同深度 (SZ, TZ, DZ) 特征带吸光度 (积分面积) 比值: 4 610  $\text{cm}^{-1}$ /4 898  $\text{cm}^{-1}$  (a)、复合带/4 898  $\text{cm}^{-1}$  (b)、复合带/4 610  $\text{cm}^{-1}$  (c)。从图 2 (a,b) 知: SZ(TZ) 中各积分带比值随着病变时间 (健康、OA-3M、OA-7M) 的延长呈下降趋势; 从图 2(a) 和 (b) 得知, 健康样本中 SZ 和 TZ 的吸光度比值高于 OA-3M 和 OA-7M 样

本, 分析认为病变初期软骨中胶原含量几乎不变, PG 含量减少且主要发生在 SZ 和 TZ<sup>[19]</sup>。OA-7M 中 DZ 比值突然升高, 可能是胶原含量不变, DZ 中 PG 修复合成含量增加所致。图 2(a) 和 (b) 相似的变化趋势证明 4 610  $\text{cm}^{-1}$  带可能亦是复合带。

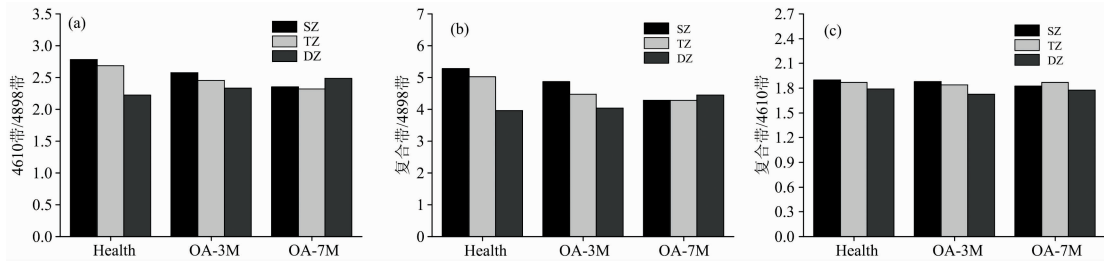


图 2 健康和多期病变样本的 4 610  $\text{cm}^{-1}$  带和 4 898  $\text{cm}^{-1}$  带 (a)、复合带和 4 898  $\text{cm}^{-1}$  带 (b)、复合带和 4 610  $\text{cm}^{-1}$  带 (c) 在不同深度 (SZ, TZ, DZ) 的吸光度比

Fig. 2 The absorbance ratios of the 4 610 to 4 898  $\text{cm}^{-1}$  band (a), the complex to 4 898  $\text{cm}^{-1}$  band (b), the complex to 4 610  $\text{cm}^{-1}$  band (c) of healthy and multi-stage OA tissues at different depths (SZ, TZ, DZ), respectively

进一步比较 SZ, TZ 和 DZ 的吸光度比值变化趋势, 发现图 2(b) 中同一病期 (健康) 的吸光度比值梯度下降比较显著, OA-3M 略有缓和, OA-7M 病期的梯度变化则趋于平缓甚至相反, 表明手术诱发 OA 病变后 PG 含量的变化。图 2 (c) 中复合带/4 610  $\text{cm}^{-1}$  带的吸光度比值随分区深度的变化在健康、OA-3M 和 OA-7M 中趋于一个非常接近的斜率, 进一步证明 4 610  $\text{cm}^{-1}$  带是包含 PG 成分和胶原蛋白的复合带。

上述特征带吸光度的比值研究, 证实了 4 610  $\text{cm}^{-1}$  带是

综合成分的整体贡献, 而不似文献[14]所报道的 4 610  $\text{cm}^{-1}$  带仅来源于胶原蛋白的这一说法。因为 NIR 光谱的吸收带多是 X-H 基团振动的重叠, 所以比较复杂, 不像 MIR 特征带分配定义较为清晰, 易于定量分析。因此, 仍需将 NIR 光谱结合化学计量学方法进行进一步的定量研究或分期诊断。

## 2.2 光谱预处理后样本分类的 PCA-FDA

图 3 为健康和 OA 样本 SZ 位置的 NIR 光谱数据分别经 BC, 1<sup>st</sup>-2-21SG, 2<sup>nd</sup>-3-25SG 预处理后的前 3 个主因子的散点图。从图 3(a) 得知, 经 BC 预处理和 PCA 后健康和 OA 样

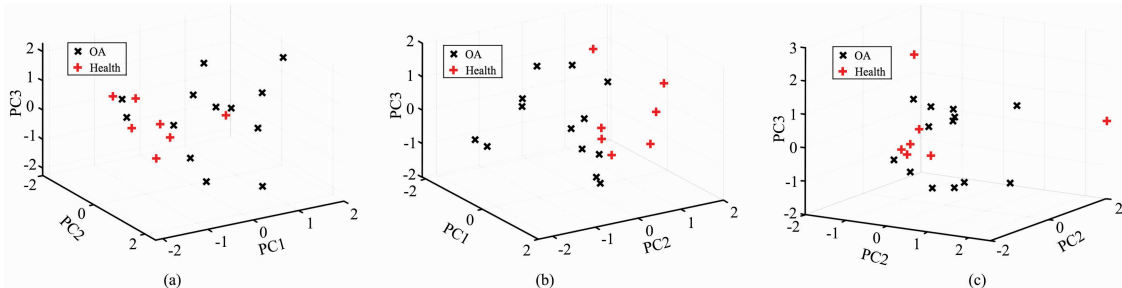


图 3 健康和 OA 样本 SZ 的 NIR 光谱数据分别经 BC(a), 1<sup>st</sup>-2-21SG(b) 和 2<sup>nd</sup>-3-25SG(c) 预处理后的前 3 个主因子散点图

Fig. 3 The scatter plots of the first 3 principal factors on the SZ NIR spectra of healthy and OA tissues with preprocess of BC(a), 1<sup>st</sup>-2-21SG (b) and 2<sup>nd</sup>-3-25SG (c), respectively

本的 3 个主因子混合在一起,无法较好的区分;比较图 3(b)和(c)发现,经 2<sup>nd</sup>-3-25SG 预处理和 PCA 后的样本主因子混合程度较 1<sup>st</sup>-2-21SG 严重。TZ 和 DZ 的光谱数据经与 SZ 相同的预处理后,前 3 个主因子图的变化趋势与 SZ 大抵相同。说明导数预处理可以较好地地区分不同类间的差异。预处理(BC, 1<sup>st</sup>-2-21SG, 2<sup>nd</sup>-3-25SG)后的 NIR 光谱的 SZ, TZ 及 DZ 前 3 个主因子的累计贡献率均超过 85%,但随着因子数的增加,样本的正确识别率上升,所以本研究选择固定的前 5 个主成分因子(累计贡献率均超过 90%)进行 FDA 处理。

表 2 为健康和病变样本的 NIR 光谱(SZ 20 片、TZ 21 片、DZ 173 片)经 BC, 1<sup>st</sup>-2-21SG, 2<sup>nd</sup>-3-25SG 等 3 种不同预处理后 PCA-FDA 的判别结果,可分析不同的预处理方法对关节软骨分类的影响。从表 2 知, SZ 的 NIR 光谱经 BC 预处理后的初始案例和交叉案例的正确识别率分别为 80% 和 50%; 经 1<sup>st</sup>-2-21SG 预处理后的初始案例和交叉案例的正确识别率分别为 95% 和 90%; 经 2<sup>nd</sup>-3-25SG 预处理后的初始案例和交叉案例的正确识别率分别是 85% 和 70%。在样本 TZ 中的 NIR 光谱经 1<sup>st</sup>-2-21SG 预处理后正确识别率为 85.7%(初始案例)和 85.7%(交互验证案例), 经 BC 预处理后的初始案例正确识别率为 76.2% 和交叉验证的正确识别率为 71.4%。经 2<sup>nd</sup>-3-25SG 预处理后的初始案例和交叉案例的正确识别率分别为 76.2% 和 42.9%。在样本 DZ 中的 NIR 光谱经 1<sup>st</sup>-2-21SG 预处理后的判别结果最优, 正确识别率也仅有 73.8%(初始案例)和 71.8%(交互验证案例), BC, 2<sup>nd</sup>-3-25SG 的识别率更低, 其中初始案例的最高的识别率才达到 71.3%。

表 2 健康和病变样本的 NIR 光谱不同预处理后的 Fisher 判别结果

Table 2 Fisher discrimination of healthy and OA tissues with different pretreatment methods of NIR spectroscopy

预处理方式	SZ 识别率/%	TZ 识别率/%	DZ 识别率/%
BC	80	76.2	71.3
初始案例	1 <sup>st</sup> -2-21SG	95	85.7
	2 <sup>nd</sup> -3-25SG	85	76.2
交叉验证	BC	50	71.4
	1 <sup>st</sup> -2-21SG	90	85.7
	2 <sup>nd</sup> -3-25SG	70	42.9

综上所述,同一软骨分区的 NIR 光谱经 3 种不同预处理后, PCA-FDA 的初始案例和交叉验证案例的正确识别率均是经 1<sup>st</sup>-2-21SG 预处理后的判别结果最好, 其中 SZ 的识别率高达 95%(初始案例)和 90%(交互验证案例)。1 号切块 50  $\mu\text{m}$  深度的样本在初始案例和交叉验证中均被误判到健康组, 误判的原因可能是该部位基质成分并未出现明显的变化; 3 号切块 100  $\mu\text{m}$  深度的样本在交叉验证中被误判到健康组, 可能是由于在 OA 过程中该部位并未出现明显的损伤导致误判。本研究结果说明采用一阶导数光谱预处理可以很好的消除基线偏移, 能凸显健康和病变的特征, 从而达到良

好的区分; 基于光谱均作 1<sup>st</sup>-2-21SG 预处理的判别结果, 可得出, 表层区的判别结果优于过渡区, 更优于深层区。

表 3 是 SZ, TZ 和 DZ 的 NIR 光谱数据经 1<sup>st</sup>-2-21SG 预处理后 PCA-FDA 显著性检验结果。Wilk's Lambda 是组内的平方和与总平方的比, 范围从 0 到 1。Wilk's Lambda 值越低, 说明类间差异就越大<sup>[20]</sup>。sig. 为显著性水平假设检验, sig. 值小于 0.05, 说明函数判别显著。从表 3 可知, TZ DZ 中 sig. 的值虽然都小于 0.05, 但 Wilk's lambda 的值偏大(0.366, 0.772), 类间的差异不明显, 故 DZ 的误判率最高。由 SZ 中 Wilk's lambda(0.345) 和 sig.(0.006) 可知: SZ 的 NIR 光谱数据经 1<sup>st</sup>-2-21SG 预处理后对健康和病变组织的分类具有很好的辨别能力。根据相关文献<sup>[19, 21]</sup>, OA 关节软骨早期病变基质成分含量的变化最明显的区域就是 SZ, 也是最易受损的区域, 与本研究的结果分析是一致的。因此进一步选择 SZ 的样本进行不同阶段的 OA 分期判别。

表 3 SZ, TZ 和 DZ 的 NIR 光谱数据经 1<sup>st</sup>-2-21SG 后的 PCA-FDA 显著性检验结果

Table 3 PCA-FDA significant test results of NIR spectral data at SZ, TZ and DZ with preprocess of 1<sup>st</sup>-2-21SG

组别	主因子数 PC/个	Wilk's lambda	sig.
SZ	5	0.345	0.006
TZ	5	0.366	0.005
DZ	5	0.772	0.000

### 2.3 OA 分期的 PCA-FDA

健康和多期 OA 软骨 NIR 光谱的训练集和预测集的 PCA-FDA 结果如表 4 所示。训练集的初始案例识别率为 93.8% 和交叉验证识别率为 87.5%; 预测集的识别率为 87.5%, 其中健康组和 OA-7M 组样本是否病变及病变时期全部被正确识别, OA-3M 组也全被正确识别为病变样本, 只是 OA-3M 组中 Y3 的 5 号切块、50  $\mu\text{m}$  深度的样本切片(预测集)被误判到 OA-7M 组。出现误判的原因可能是: 在 OA 过程中, 由于个体差异, 预测集中 Y3 犬的该部位出现明显的损伤, 导致误判; 在初始案例和交叉案例的识别中, 有同样样本被误判(6 号切块深度为 50  $\mu\text{m}$ ), 该样本的存在, 影响了训练集的模式。

基于以上结果, 如果将影响训练集模型的该数据移出, 再次做 OA 分期分析, 发现训练集的初始案例正确识别率为 100%、交叉验证正确识别率为 93.3%、预测集的正确识别率为 87.5%, 提高了训练集模型的稳定性。在训练集交叉验证判别结果中 OA-3M 组 Y2 犬 3 号切块、100  $\mu\text{m}$  深度的切片样本被误判到健康组, 且在健康和 OA 分类分析中, 该样本同样出现误判, 出现误判的原因可能是: 由于 ACL 手术 3 个月, 处于 OA 的早期病变, 该部位基质成分并未出现明显的变化。预测集健康组和 OA-7M 组样本是否病变及病变分期也全部被正确识别, OA-3M 组也全被正确识别为病变样本, 同样是 OA-3M 组中 Y3 的 5 号切块 50  $\mu\text{m}$  深度的切片样本被误判到 OA-7M 组。

表 4 健康和多期 OA 样本 NIR 光谱的训练集和预测集的 PCA-FDA(SZ, TZ, DZ)

Table 4 PCA-FDA of NIR spectral training and prediction sets of healthy and multi-stage OA cartilage (SZ, TZ, DZ)

类别	光谱数据	初始判别结果			交叉判别结果			识别率/%	
		健康组	OA-3M	OA-7M	健康组	OA-3M	OA-7M	初始案例	交叉验证
训练集	健康组	6	6		6				
	OA-3M	5	4	1	1	3	1	93.8	87.5
	OA-7M	5		5			5		
预测集	健康组	3	3						
	OA-3M	2	1	1				87.5	
	OA-7M	3		3					

采用 1<sup>st</sup>-2-21SG 的 NIR 光谱预处理方法结合 PCA-FDA 能有效地改善 NIR 光谱的分析效果,并能较准确地对 OA 的不同病期进行识别。仅针对 OA 早期病变进行初步研究,为更深入地对早期 OA 病变进行研究,需扩大样本数量,提高模型的稳定性,提高预测效果。

### 3 结 论

采用 FTNIR 技术结合 PCA-FDA 方法对比格犬健康和不同病期的 OA 进行研究。详细地探讨 BC, 1<sup>st</sup>-2-21SG 和 2<sup>nd</sup>-3-25SG 三种不同预处理方法对 FDA 结果的影响。发现

经 1<sup>st</sup>-2-21SG 预处理后在 SZ 的判别结果最优,初始案例和交叉验证案例的识别率分别达到 95% 和 90%;在 OA 分期识别结果中,模型的初始案例正确识别率为 100%,交叉验证正确识别率为 93.3%,预测集的识别率为 87.5%。采用 NIR 光谱的 1<sup>st</sup>-2-21SG 预处理结合 PCA-FDA 能有效地改善 NIR 光谱的分析效果,较准确地鉴别关节软骨是否发生病变,并能准确地对 OA 的不同病期进行判别,本研究有效实现了 OA 的 NIR 监督分类/分期识别,为 OA 的早期原位诊断提供一种临床可借鉴方法,对骨关节炎 OA 的研究、监测和诊断具有重要意义。

### References

- [1] Soltz M A, Mauck R L, Wang C C, et al. *Journal of Biomechanical Engineering*, 2000, 122(3): 252.
- [2] ZHANG Li-juan, YAO Wei-wu(张丽娟, 姚伟武). *Chinese Computed Medical Imaging(中国医学计算机成像杂志)*, 2010, 16(1): 87.
- [3] Nagarajan R, Yang Xia, Aruna B, et al. *Applied Spectroscopy*, 2007, 61(12): 1404.
- [4] Kuettner K E. *Clinical Biochemistry*, 1992, 25(3): 155.
- [5] Buckwalter J A, Mankin H J. *Instructional Course Lectures*, 1998, 47(4): 477.
- [6] SONG Ming-qing, DU Min(宋明清, 杜 民). *Chinese Journal of Scientific Instrument(仪器仪表学报)*, 2005, 26(z2): 304.
- [7] Wilson W, Huyghe J M, Donkelaar C C. *Biomechanics & Modeling in Mechanobiology*, 2007, 6(1-2): 43.
- [8] Chen S S, Falcovitz Y, Schneidman R. *Osteoarthritis and Cartilage*, 2001, (9): 561.
- [9] Xia Y, Alhadlaq H, Ramakrishnan N, et al. *Journal of Structural Biology*, 2008, 164(1): 88.
- [10] Tan A, Mitra A K, Chang P C, et al. *Journal of Orthopaedic Surgery*, 2004, 12(2): 199.
- [11] Kennedy J F, Kaczmarek A. *Carbohydrate Polymers*, 2007, 67(4): 648.
- [12] HUANG Jiang-yin, ZHAO Jing, DONG Xiao-wei, et al(黄江茵, 赵 晶, 董晓威, 等). *China Journal of Bioinformatics(生物信息学)*, 2018, 16(1): 57.
- [13] Brown C P, Jayadev C, Glyn-Jones S, et al. *Physics in Medicine & Biology*, 2011, 56(7): 2299.
- [14] Palukuru U P, McGoverin C M, Pleshko N, et al. *Matrix Biology*, 2014, 38: 3.
- [15] Prakash M, Sarin J K, Rieppo L, et al. *Applied Spectroscopy*, 2017, 71(10): 2253.
- [16] YANG Jin-yong, LI Xue-chun, HUANG An-min, et al(杨金勇, 李学春, 黄安民, 等). *Journal of Northeast Forestry University(东北林业大学学报)*, 2013, (12): 132.
- [17] Chiang L H, Russell E L, Braatz R D. *Chemometrics & Intelligent Laboratory Systems*, 2000, 50(2): 243.
- [18] LIU Zhen-long, MAN Zhen-tao, ZHANG Ji-ying, et al(刘振龙, 满振涛, 张继英, 等). *Chinese Journal of Sports Medicine(中国运动医学杂志)*, 2015, 34(2): 103.
- [19] Goldring M B. *Arthritis & Rheumatism*, 2000, 43(9): 1916.
- [20] Patel S, Bhavsar C D. *International Journal of Pharmaceutical Science Invention*, 2013, 2(1): 36.
- [21] Yin Jianhua, Xia Yang. *Spectrochimica Acta Part A: Molecular & Biomolecular Spectroscopy*, 2014, 133: 825.

# NIR Spectroscopic Study and Staging Diagnosis of Osteoarthritic Articular Cartilage

FU Juan-juan, MA Dan-ying, TANG Jin-lan, BAO Yi-lin, ZHAO Yuan, SHANG Lin-wei, YIN Jian-hua\*

Department of Biomedical Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China

**Abstract** Osteoarthritis (OA) is a major medical disease that threatens the middle-aged and aged people's public health and quality of life. The early lesions of OA are mainly shown in the content changes of extracellular matrix components, which are difficult to detect by patients themselves, even the existing clinical and experimental methods. In recent years, Fourier transform near-infrared (FTNIR) spectroscopy has been used for the field of surgical navigation, non-destructive testing and various disease diagnoses due to its fast speed, low cost, ease of penetrating tissue to obtain spectral information, etc. Based on the above advantages, FTNIR technology was used to collect and analyze NIR spectra of healthy and multi-period OA articular cartilage at different depth zones (superficial zone, transitional zone and deep zone) in this paper. Then, principal component analysis (PCA) and Fisher discrimination algorithm (FDA) were combined for studying the influence of different preprocessing methods on discrimination results, the changes of matrix composition at different zones, and the diagnosis of multi-period OA. Compared to other two preprocessing methods (baseline correction, second-derivative cubic polynomials 25-point Savitzky-Golay smoothing) at same zone, the preprocessing of first-derivative quadratic polynomials 21-point Savitzky-Golay smoothing shows the best discrimination results, and the recognition rates at the superficial zone are as high as 95% and 90% for initial case and cross validation case, respectively. The discrimination results at the superficial zone are better than those at the transitional zone, and far better than those at the deep zone, which proves that the early lesions in OA mainly occurs at superficial zone. In the multi-period OA recognition, the recognition rates of initial case, cross-validation and prediction sets through optimized data by the FDA model are 100%, 93.3%, and 87.5%, respectively. The results indicate that the first derivative pretreatment of the NIR spectra combined with the PCA-FDA method can effectively identify whether the articular cartilage is diseased and which period it is, which is of great significance in study of monitoring OA and early diagnosis. NIR technology with the appropriate spectral analysis method can be applied to the in situ staging and early clinical diagnosis of OA.

**Keywords** Osteoarthritis; NIR spectra; Derivative preprocessing; Principal component analysis; Fisher discrimination; Stage diagnosis

(Received Jul. 19, 2020; accepted Oct. 27, 2020)

\* Corresponding author