

Journal Pre-proofs

Feasibility assessment of employing near-infrared/visible spectroscopy for detecting olive pomace oil adulteration using LDA and SVM identification methods

Shirin Asadian¹, Ahmad Banakar^{2*}, Bahareh Jamshidi³

DOI: <https://doi.org/10.22104/ift.2024.6519.2155>

*ah_banakar@modares.ac.ir

To appear in: Innovative Food Technologies (IFT)

Received Date: 19 October 2023

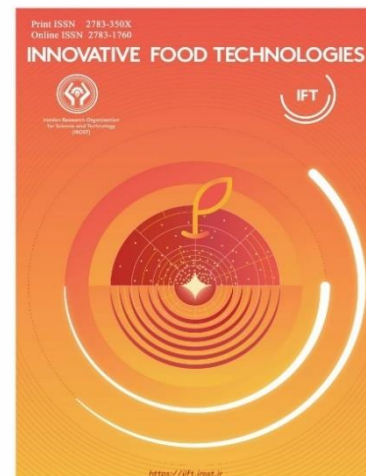
Revised Date: 14 May 2024

Accepted Date: 27 May 2024

Please cite this article as: Shirin Asadian¹, Ahmad Banakar^{2*}, Bahareh Jamshidi³, Feasibility assessment of employing near-infrared/visible spectroscopy for detecting olive pomace oil adulteration using LDA and SVM identification methods, *Innovative Food Technologies* (2024), doi: <https://doi.org/10.22104/ift.2024.6519.2155>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2024 The Author(s). Published by irost.org.



امکان‌سنجی به‌کارگیری طیف‌سنجی مرئی/فروسرخ نزدیک برای تشخیص تقلب روغن تفاله زیتون با روش‌های شناسایی LDA و SVM

شیرین اسدیان^۱، احمد بناکار^{۲*}، بهاره جمشیدی^۳

۱. شیرین اسدیان، دانشجوی کارشناسی ارشد، گروه مهندسی بیوسیستم، دانشکده کشاورزی، دانشگاه تربیت مدرس، تهران، ایران.

۲. احمد بناکار^{*}، دانشیار گروه مهندسی بیوسیستم، دانشکده کشاورزی، دانشگاه تربیت مدرس، تهران، ایران.

۳. بهاره جمشیدی، دانشیار مؤسسه تحقیقات فنی و مهندسی کشاورزی، سازمان تحقیقات، آموزش و ترویج کشاورزی، کرج، ایران.

* نویسنده مسئول: ah_banakar@modares.ac.ir

چکیده

زیتون یک میوه مدیترانه‌ای است که عمدتاً به دلیل روغن آن کشت می‌شود. روغن‌زیتون فرابکر همیشه مورد توجه و خواست استفاده‌کنندگان بوده است. از این رو، تقلب در روغن‌های زیتون بکر و فرابکر، با افزودن روغن‌هایی مثل کانولا، آفتابگردان، تفاله زیتون و غیره مشاهده می‌شود. تشخیص تقلب در روغن‌های گیاهی با ترکیبات مشابه و گونه‌های متفاوت با روش‌های مرسوم از جمله کروماتوگرافی گازی مشکل و زمان‌بر است و نیاز به آماده‌سازی نمونه و اپراتور دارد. به همین دلیل کاربرد فناوری‌های غیرمخرب برای تشخیص تقلب اهمیت دارد. در این پژوهش، با استفاده از فناوری طیف‌سنجی مرئی/فروسرخ نزدیک (Vis/NIR)، تشخیص تقلب روغن پومیس (روغن تفاله زیتون) بررسی شد. از سه نوع روغن‌زیتون فرابکر، بکر و روغن‌زیتون تصفیه شده برای نمونه‌گیری استفاده شد؛ همچنین نمونه‌ها در شش دسته خالص، ۱۱، ۲۰، ۳۳، ۵۰ و ۱۰۰ درصد تقلب ساخته شدند. هر تیمار در ده نمونه تهیه و آزمایش‌ها انجام شد. در ادامه به منظور تحلیل ویژگی‌های کیفی و طبقه‌بندی داده‌های مستخرج از طیف‌سنج، از روش‌های بازشناسی الگو شامل تحلیل تفکیک خطی (LDA) و ماشین بردار پشتیبان (SVM) استفاده شد. نتایج به دست آمده نشان داد که طیف‌سنجی مرئی/فروسرخ نزدیک (Vis/NIR) قادر به تفکیک نمونه‌های روغن‌زیتون بر اساس درصدهای مختلف تقلب پومیس است. هر چند روش LDA توانست با صحت قابل قبولی نمونه‌های روغن‌زیتون را با توجه به نرخ تقلب دسته‌بندی کند، اما روش SVM با صحت آموزش ۹۶/۶۹ درصد و اعتبارسنجی ۹۵/۸۶ درصد از صحت و برازش مطلوب‌تری برخوردار بود. طبق نتایج، تابع خطی، به عنوان بهترین تابع برای ساخت مدل‌های دسته‌بندی به روش SVM پیشنهاد شد.

واژه‌های کلیدی: تحلیل تفکیک خطی، تقلب، روغن پومیس، ماشین بردار پشتیبان.

۱. مقدمه

درخت زیتون درختی مقاوم، بردبار و همیشه سبز بوده است و میوه آن برای تغذیه به صورت کنسرو^۱ و روغن‌کشی^۲ مورد استفاده قرار می‌گیرد [1]. در حال حاضر بیش از ۹۵ درصد درختان زیتون جهان در مناطق مدیترانه‌ای رشد می‌کنند. نزدیک به ۸۱ درصد کل تولید زیتون از اتحادیه اروپا (اسپانیا، ایتالیا، یونان، پرتغال و فرانسه)، حدود ۷ درصد از خاور نزدیک و حدود ۱۱ درصد از آفریقای شمالی تامین می‌شود. ۱ درصد باقیمانده به مناطق آمریکایی مثل آرژانتین، مکزیک، پرو و ایالات متحده تعلق دارد [2]. لیبی‌ها نقش کلیدی در سلامت انسان دارند [3]. روغن زیتون از زمان‌های قدیم از نظر خوراکی و حتی درمانی بسیار مورد توجه بوده است؛ بطوریکه روغن زیتون می‌تواند در درمان بیماری‌های قلبی عروقی به کار گرفته شود [4]. روغن زیتون از نظر تجاری محصولی ارزشمند است بنابراین اغلب در معرض تقلب با سایر روغن‌های خوراکی مانند آفتابگردان، روغن فندق، روغن پومیس و دیگر روغن‌ها با ترکیب اسیدهای چرب مشابه مانند کانولا و کنجد است [5]. تقلب در مواد غذایی از جنبه‌های زیان اقتصادی به مصرف‌کننده، زیان آسیب به سلامتی و یا عدم بهره‌مندی از فواید تغذیه‌ای و جنبه‌ی روانی آن یعنی عدم اعتماد به محصول معرفی شده، نیازمند تحقیق و بررسی و نیز تصویب و اجرایی نمودن الزامات، قوانین و استانداردهای مربوط می‌باشند. تقلب در مواد غذایی مخصوصاً در روغن زیتون، با افزودن مواد با ارزش غذایی و قیمت پایین‌تر به روغن اصل ایجاد می‌گردد و این افزودن باعث تغییر در بعضی ویژگی‌ها می‌گردد. تشخیص روغن اضافه‌شده به روغن زیتون معمولاً به آسانی انجام نمی‌شود؛ چرا که گاهی روغن اضافه را طوری انتخاب می‌کنند که

^۱Visible Near Infrared

^۲Canned

^۳Oil Extraction

هیچ تغییری در شاخص‌ها (عدد یدی، عدد صابونی، تغییر ضریب شکست، نور و غیره) به وجود نیاید [6]. روغن زیتون بکر^۱ با کیفیت بالا را می‌توان تنها از میوه‌های سالم و تازه با درجه رسیدگی مناسب تولید کرد. نوع و میزان ترکیبات شیمیایی موجود در روغن زیتون به عواملی نظیر رقم، شرایط اقلیمی و روش استخراج وابسته است [5]. پارامترهای فنی مانند روش‌های نگهداری و مدت زمان فرایند هم می‌تواند بر کیفیت روغن دخالت داشته باشد. پایداری اکسایشی و ارزش تغذیه‌ای روغن زیتون نسبت به روغن‌های گیاهی دیگر به دلیل وجود ترکیبات آنتی‌اکسیدانی طبیعی، بیشتر است [7].

باتوجه به اینکه روغن زیتون یک نوع روغن مصرفی به‌ویژه برای پخت‌وپز است، روی سلامتی عموم مردم، به‌ویژه بیماری‌های قلبی عروقی، دیابت و سرطان تأثیرگذار است [8]. بر این اساس، کنترل کیفیت این روغن یکی از مسائل پراهمیت در بهداشت و ایمنی مواد غذایی است [9]. همچنین به دلیل اینکه روغن زیتون، یک روغن گران‌قیمت است می‌تواند در معرض تقلب افراد سودجو قرار گیرد. از این رو، وجود استاندارد ها در سطوح مختلف یک امر ضروری و مهم است. از طرفی تنوع کشورهای تولیدکننده آن نیز لزوم توجه به حفظ کیفیت واردات این روغن را افزایش می‌دهد. پس از به دست آوردن روغن زیتون با تکنیک‌های فیزیکی، تفاله‌ای ایجاد می‌شود که محتوی بیش از ۸ درصد روغن است و می‌تواند توسط حلال (معمولاً هگزان) استخراج شود. این روغن، روغن پومیس^۲ نام دارد. به علت خصوصیات حلال، روغن استخراج شده با حلال، مقادیر بیشتری از اجزاء کمیاب را دارا است [10]. روغن پومیس به علت استخراج با حلال‌های شیمیایی و در نتیجه آلوده بودن به بنزوالفاپیرن^۳، مونوکلوروپروپان دی‌ال‌ها^۴ و هیدروکربن‌های حلقوی^۵ سرطان‌زا است و روی اعصاب و سایر ارگان‌های بدن تأثیر نامطلوب دارد [11]. محتوای میزان اسیدهای چرب^۶ یک پارامتر مهم برای ارزیابی کیفیت روغن محسوب می‌شود. از ترکیبات اصلی روغن زیتون، اسیدهای چرب هستند. روغن زیتون دارای ۱۵-۱۰ درصد اسید چرب اشباع^۷ و ۷۰-۸۰ درصد اسید چرب غیراشباع^۸ است [12]. فراوان‌ترین اسید چرب موجود در روغن زیتون اسید اولئیک است [13]. با بهره‌گیری از کروماتوگرافی گازی^۹ (GC) می‌توان درصد اسیدهای چرب را با دقت بالا به دست آورد و از این طریق نوع روغن زیتون را تشخیص داد. کروماتوگرافی دقیق‌ترین روش تحلیلی است [14]. در حال حاضر از روش کروماتوگرافی گازی برای شناسایی تقلب در روغن زیتون استفاده می‌شود. همچنین در **ارزیابی انطباق** برای کنترل کیفیت مواد غذایی از جمله روغن زیتون **برای اصالت سنجی، تشخیص، تعیین تقلبات و کمیت اسیدهای چرب**، توسط آزمایش‌های شیمیایی و تست کروماتوگرافی گازی صورت می‌پذیرد که که نسبت به روش‌های غیرمخرب مانند طیف‌سنجی مرئی-فروسرخ نزدیک روش گران‌قیمت و زمان‌بری هستند و نیازمند آزمایشگاه‌های تخصصی کنترل کیفیت و نیروی انسانی متخصص **می‌باشند**. قابل ذکر است که وجود روغن پومیس در روغن زیتون بکر و روغن زیتون تصفیه‌شده به‌روش کروماتوگرافی گازی قابل تشخیص نیست و تشخیص تقلب در این موارد با روش کروماتوگرافی بدون تجهیزات پیشرفته از جمله محفظه تزریق قابل برنامه‌ریزی دما کاربرد ندارد ولی در صورت بهره‌گیری از این

^۱Virgin Olive Oil

^۲Pomace Olive Oil

^۳Benze Alphaperenn

^۴Monochloropropane DL

^۵Cyclic hydrocarbons

^۶Fatty acids

^۷Saturated fatty acid

^۸Unsaturated fatty acids

^۹ Gas Chromatography

تجهیزات، تقلب روغن تفاله زیتون با مقادیر بسیار کمتر از ۱ درصد هم قابل تشخیص است [15]. امروزه روش‌های متنوعی برای کنترل کیفیت میوه‌ها در دنیا وجود دارد که در این بین روش‌های غیرمخرب به دلیل توانایی در تشخیص ویژگی‌های مختلف محصول بدون آسیب‌رساندن به آن‌ها از اهمیت ویژه‌ای برخوردار هستند. برخی از این روش‌ها شامل طیف‌سنجی مرئی - فروسرخ (Vis/NIR) ، تصویربرداری چند طیفی و فراطیفی، تصویربرداری رزونانس مغناطیسی هسته،^۳ توموگرافی محاسباتی با استفاده از اشعه ایکس و طیف‌سنجی دی‌الکترونیک هستند [16]. روش‌های طیف‌سنجی مانند مرئی، فروسرخ نزدیک، فروسرخ میانه و طیف‌سنجی فراطیفی غیرمخرب‌تر از روش‌های مکانیکی و شیمیایی هستند و داشتن ابزارهای ساده، سریع، قابل حمل و فراگیر، از دیگر فواید استفاده از این روش‌ها هستند. بنا بر همین دلایل امروزه از این فناوری به شکل گسترده‌ای در صنایع کشاورزی، غذا، دارو و پتروشیمی استفاده می‌شود [17]. یکی از کاربردهای مناسب روش طیف‌سنجی، آشکارسازی افزودنی‌های تقلبی در محصولات به‌ویژه مواد غذایی پرکاربرد و گران‌قیمت است [18]. در پژوهشی از طیف‌سنجی FTIR^۴ برای تجزیه و تحلیل و طبقه‌بندی روغن‌های زیتون فرابکر در مناطق مختلف مراکش استفاده شد. در این تحقیق از خوشه‌بندی^۵ به همراه آنالیز تفکیک حداقل مربعات جزئی^۶ استفاده شده است. براساس آنالیز خوشه‌بندی و آنالیز مؤلفه‌های اصلی^۷ چهار خوشه متمایز شناسایی شدند. این روش نتایج قابل اعتمادی را در طبقه‌بندی روغن زیتون فرابکر از مناطق مختلف مراکش با مزایای سریع، ارزان بودن و بدون نیاز به روش جداسازی قبلی ارائه می‌دهد [19]. از ترکیب روش‌های بینی الکترونیک^۸ و زبان الکترونیک^۹ نیز برای توصیف پنج روغن زیتون بکر (VOO) از مناطق مختلف جغرافیایی مراکش استفاده شد. آنالیز مؤلفه‌های اصلی نشان داد که وقتی از هرکدام از دو روش به تنهایی استفاده می‌شود نمی‌توان آن‌ها را به‌طور مشخص از یکدیگر تفکیک کرد. آنالیز کلاس‌بندی به روش ماشین بردار پشتیبان (SVM) انجام گرفت و نشان داد نمونه‌ها با توجه به منشأ جغرافیایی‌شان به درستی طبقه‌بندی شدند. نتایج به‌دست آمده از داده‌ها به‌صورت ترکیبی (استفاده همزمان از بینی و زبان الکترونیک) از نتایج طبقه‌بندی به صورت جداگانه بهتر عمل کردند [20]. در تحقیق دیگری تقلب روغن زیتون با استفاده از پردازش تصویر و طیف‌سنجی دی‌الکترونیک از لحاظ ترکیب آن با روغن‌های ذرت، کلزا، آفتابگردان و سویا مورد بررسی قرار گرفت. مقادیر ضریب همبستگی و میانگین مربعات خطا برای نمونه‌های مخلوط شده زیتون و آفتابگردان برابر ۰/۹۶۲ و ۰/۰۰۸ و برای نمونه‌های مخلوط شده زیتون و کلزا مقادیر ۰/۹۶۱ و ۰/۰۱۳ حاصل شد. نتایج ارزیابی‌ها نشان داد که سامانه طراحی شده، توانایی تشخیص دقیق موردنظر را دارا است [1]. در آزمایش دیگری از طیف‌سنجی فروسرخ برای تجزیه و تحلیل کیفیت غیرمخرب دانه‌های روغنی و دانه‌های خوراکی استفاده شد. نتایج نشان داد که این فناوری سریع، غیرمخرب و کم‌هزینه است و یک گزینه امیدوارکننده برای تشخیص پارامترهای کیفیت دانه‌های روغنی و همچنین برای تشخیص اصالت و قابلیت ردیابی منشأ دانه‌های روغنی و روغن‌های خوراکی با صحت ۸۹ درصد است [21]. مطالعات گذشته بیانگر اهمیت بالای روش طیف‌سنجی به‌عنوان یک روش سریع، ساده،

¹ Visible-Near infrared

^۲Multispectral/Hyperspectral imaging

^۳Nuclear magnetic resonance

^۴ Fourier transform infrared

^۵ Cluster analysis

^۶ Discriminant analysis-Partial least squares

^۷ Principal component analysis

^۸ Electronic nose

^۹ Electronic tongue

^۱ Virgin olive oil

^{۱۱}Support Vector Machine

کم هزینه با دقت قابل قبول در سنجش کیفیت محصولات کشاورزی است. به همین دلیل، استفاده از آن به صورت های مختلف و برای محصولات گوناگون روز به روز در حال گسترش است. اما پژوهشی که در آن تشخیص تقلب پومیس در انواع روغن زیتون (فرا بکر، بکر و روغن زیتون تصفیه شده) مبتنی بر روش طیفسنجی و با استفاده از روش های طبقه بندی پرداخته شده باشد، یافت نشد. بنابراین، در این پژوهش، شناسایی تقلب روغن پومیس، با استفاده از روش سریع و غیرمخرب طیفسنجی مرئی - فروسرخ نزدیک بدون سنجش و اندازه گیری خواص کیفی با استفاده از روش های شناسایی الگو نظیر LDA و SVM انجام پذیرفت. در واقع تجزیه و تحلیل داده های چندمتغیره از جمله روش های آماری است که بر مشاهده همزمان، بهره برداری و پردازش چندین متغیر آماری به منظور استخراج اطلاعات سینتیک مربوطه استوار است. از این رو، اهداف خاص این پژوهش عبارتند از: ۱- مقایسه عملکرد روش های پیش پردازش طیفی مختلف ۲- آنالیز توزیع متغیرهای انتخاب شده به وسیله الگوریتم های مختلف انتخاب متغیر ۳- مقایسه صحت تشخیص مدل ساخته شده به وسیله ترکیب پیش پردازش های طیفی مختلف و الگوریتم های انتخاب متغیر.

۲. مواد و روش ها

۱.۲. نمونه گیری

روغن های زیتون فرا بکر، بکر و روغن زیتون تصفیه شده مورد استفاده در این پژوهش، از کارخانه تولید روغن زیتون اتکا در شهرستان ورامین استان تهران تهیه شد. نمونه روغن پومیس خام و غیرخوراکی نیز از شرکت دارویی، آرایشی و بهداشتی تهیه شد. ابتدا روغن کشتی به صورت پرس سرد صورت گرفت و سپس برای اطمینان از اصالت آن ها، مشخصات و درصد ترکیبات سازنده آن با سامانه کروماتوگرافی گازی مدل Unicam 4600 (با دتکتور FID در دمای ۳۰۰ درجه سلسیوس و ستون Capillary, BP*70) در آزمایشگاه گروه علوم دام و طیور دانشکده کشاورزی دانشگاه تربیت مدرس در شهر تهران استخراج شد. کروماتوگرافی گازی، روشی برای جداسازی و بررسی مواد فرار یک ماده (معمولاً به صورت مایع یا گاز) است. روغن زیتون فرا بکر، بکر و روغن زیتون تصفیه شده با روغن پومیس با درصدهای وزنی ۱۱، ۲۰، ۳۳، ۵۰ و ۱۰۰ درصد مخلوط شدند. این درصدها با توجه به تقلب های موجود در بازار انتخاب شدند. در مجموع ۱۶۰ نمونه که به ازای هر تیمار ده تکرار وجود داشت، تهیه شد. شکل ۱، صفحه ۹، دسته های مختلف تقلب را نشان می دهد. همان گونه که پیداست از لحاظ ظاهری نمونه ها با هم هیچ تفاوتی ندارند.

۲.۲. اندازه گیری با دستگاه طیفسنج مرئی/ فروسرخ نزدیک

به منظور طیف گیری از نمونه ها نیازی به هیچ گونه آماده سازی روی نمونه نیست که این از مزایای طیفسنجی مرئی/ فروسرخ نزدیک است. طیف مرئی/ فروسرخ نزدیک با استفاده از فیبر دومسیره که قسمت مرکزی برای انتقال نور از نمونه به آشکارساز و فیبرهای بیرونی برای انتقال نور از منبع تابش به نمونه است، استفاده شد، و در وضعیت عبوری و در ناحیه ۸۵۰-۲۰۰ nm با استفاده از طیفسنج مدل USB2000 ساخت شرکت Ocean Optic و با تفکیک پذیری ۱ نانومتر توسط نرم افزار OceanView ثبت شد. این طیفسنج بسیار کوچک، سبک و قابل حمل است مجموعه این دستگاه به وزن ۱۹۰ گرم شامل یک طیفسنج با ابعاد ۳/۳×۶۳ mm ×۸۹/۱ mm ×۳۴/۴mm و آشکارساز Sony ILX51B CCD است. سرعت برداشت هر منحنی طیفی ۰/۰۰۱ میلی ثانیه است. از یک کاوشگر فیبر نوری دومسیره استفاده شد که قابلیت انتقال اطلاعات به کامپیوتر را به منظور نمایش و ذخیره سازی طیف های به دست آمده در نرم افزار Ocean View، از طریق پورت USB دارد. از یک کامپیوتر رومیزی برای نمایش و ذخیره داده های طیفی به کمک نرم افزار مخصوص داده های طیفسنجی استفاده شد. به منظور تولید نور مناسب، از منبع نور هالوژن - تنگستن مدل High Intensity Contact Probe A122300 (ASD, Inc. USA) با قابلیت اتصال به فیبر نوری بهره گیری شد. این منبع نور دارای توان خروجی ۶/۵ وات و ولتاژ ۱۸-۱۲ DC ولت است. طیف نمونه های روغن با ۱۰ اسکن متوالی گرفته و میانگین ۱۰ طیف حاصله، برای هر نمونه گزارش شد.

۳,۲. آماده‌سازی و پیش‌پردازش طیف‌ها

داده‌های طیفی به‌دست‌آمده از طیف‌سنج، افزون بر اطلاعات نمونه شامل اطلاعات ناخواسته پس‌زمینه و نویزها نیز هستند. از این رو، برای دستیابی به مدل‌های واسنجی پایدار، دقیق و قابل‌اعتماد، نیاز به پیش‌پردازش داده‌های طیفی پیش از توسعه مدل‌های دسته‌بندی متناسب با داده است. ساخت مدل دسته‌بندی برای توصیف رابطه بین داده‌های طیفی و آزمایشگاهی یکی از مهم‌ترین و سخت‌ترین مراحل طیف‌سنجی به شمار می‌رود و فرایندی زمان‌بر است که معمولاً با آزمون و خطا انجام می‌گیرد. بعد از ثبت و ذخیره طیف‌های به‌دست‌آمده برای تحلیل طیف‌های دریافتی، پیش‌پردازش داده‌ها انجام شد. بخشی از داده‌ها از ناحیه ۵۵۰ نانومتر به قبل دارای نویز زیاد بودند و به روش‌های پیش‌پردازش معمول قابل هموارسازی نبودند؛ بنابراین حذف شدند. به‌منظور اجرای مراحل پیش‌پردازش داده‌های طیفی نیاز به تبدیل داده‌های مد عبور به مد جذبی است که بدین منظور از رابطه ۱ استفاده شد که به قانون بیر-لمبرت مشهور است [22].

$$A = \log 1/Tr_{\lambda} \quad (1)$$

که در آن A : میزان جذب و Tr_{λ} : بازتاب یا عبور در طول موج λ است.

پس از این که طیف‌های عبوری به‌دست‌آمده از همه نمونه‌ها به‌منظور ایجاد یک رابطه خطی با غلظت مولکولی نمونه به طیف‌های جذبی تبدیل شدند، برای آماده‌سازی و پیش‌پردازش طیف‌ها از برخی روش‌های هموارسازی و کاهش نویز مانند میانگین‌گیری متحرک^۱ (MA)، ساویتزکی گولای^۲ (SG)، توزیع نرمال استاندارد^۳ (SNV) و مشتق اول^۴ (D₁) استفاده شد.

روش میانگین‌گیری متحرک یکی از معمول‌ترین روش‌های هموارسازی و حذف نویزهای تصادفی موجود در طیف است. به‌گونه‌ای که با استفاده از یک فیلتر میانگین‌گیری متحرک، هر نقطه از طیف با میانگین یک همسایگی (نقطه‌های مجاور شامل خود نقطه) از آن جایگزین می‌شود. در این تحقیق، پس از تبدیل طیف‌ها به جذب، هموارسازی آن‌ها با فیلتر میانگین‌گیری متحرک و عرض پنجره ۷ اجرا شد. هموارسازی ساویتزکی گولای شیوه‌ای دیگر از هموارسازی است که می‌تواند به طور مؤثری نویزهای تیز را بدون از بین رفتن اطلاعات طیفی برطرف کند. روش SG در این پژوهش، با در نظر گرفتن تعداد نقطه‌های همسایگی و درجه چندجمله‌ای به ترتیب برابر ۳ و ۵، برای هموارسازی طیف‌ها همچنین مقایسه با روش هموارسازی MA به کار گرفته شد. روش توزیع نرمال استاندارد برای حذف اثرهای جمعی و افزایش پخش نور، اندازه نمونه و تغییر فاصله نوری است. از این روش در این پژوهش برای نرمال‌سازی طیف‌ها، پس از هموارسازی و حذف نویز طیف‌ها استفاده شد. همچنین از روش مشتق اول برای افزایش قدرت تفکیک طیفی، حذف پس‌زمینه و جابه‌جایی خط پایه استفاده شد.

۴,۲. توسعه مدل کالیبراسیون

در آنالیز کمی و کیفی محصولات غذایی، ساخت و توسعه مدل واسنجی^۵ قابل‌اعتماد بسیار ضروری است. به کمک مدل می‌توان مشخصات نامعلوم نمونه‌ها را تعیین و پیش‌بینی کرد. در واقع هدف از توسعه مدل در این پژوهش ارزیابی نرخ تقلب روغن‌های زیتون و طبقه‌بندی آن‌ها است. نمونه‌ها به دو دسته آموزش (۸۰ درصد) برای توسعه مدل‌های واسنجی، و دسته آزمون (۲۰ درصد) برای اعتبارسنجی بیرونی مدل‌های واسنجی توسعه یافته به منظور بررسی توانمندی و دقت مدل‌ها در کلاس‌بندی درصد‌های مختلف تقلب نمونه‌های ناشناخته تقسیم شدند. در این پژوهش از روش تحلیل تفکیک خطی^۶ (LDA) و ماشین بردار پشتیبان^۷ (SVM)

^۱Moving Average

^۲Savitzky-Golay

^۳Standard Normal Variate

^۴First derivative

^۵Calibration

^۶Linear Discriminant Analysis

^۷Support Vector Machine

استفاده شد. ارزیابی مدل ساخته شده با روش اعتبارسنجی متقاطع کامل^۱ برای انتخاب مدلی با بیشترین توانایی دسته‌بندی، در بین مدل‌های آماری مورد استفاده قرار می‌گیرد. به عبارت دیگر پس از ساخت مدل در مرحله واسنجی با ارزیابی مجدد مدل، به بررسی اعتبار و پایداری مدل نیز پرداخته شد.

تحلیل تفکیک خطی (LDA)

تحلیل تفکیک خطی در حوزه استخراج ویژگی‌ها تکنیک شناخته شده‌ای است که در بسیاری از مسائل بازشناسی الگو با نظارت با نتیجه موفقیت‌آمیزی استفاده شده است. هدف اصلی LDA نگاشت داده‌ها به فضایی است که در آن نمونه‌ها را در گروه‌های مجزا، تفکیک کند. LDA برای دستیابی به این هدف به دنبال افزایش جدایی بین کلاس‌ها و درعین حال کاهش پراکندگی داخل کلاس‌ها است [23]. در این پژوهش برای ایجاد مدل دسته‌بندی به روش LDA از سه نوع تابع Linear، Quadratic و Mahalanobis و از ۴ نوع ضرایب وزنی مختلف (Downweight، Block weighting، Constant، 1/SDev) استفاده شد.

ماشین بردار پشتیبان (SVM)

ماشین بردار پشتیبان از جمله روش‌های غیر احتمالی است که برای انجام طبقه‌بندی از ابر صفحه‌ها برای تعریف مرز بین کلاس‌ها استفاده می‌کند. در واقع هدف اصلی این روش، یافتن بهترین ابر صفحه‌ای است که به‌عنوان یک سطح تصمیم‌گیری، حاشیه بین دو کلاس را بیشینه و داده‌های مربوط به دو کلاس را از هم تفکیک کند. SVM همراه با توابع کرنل مختلف از جمله توابع خطی، چند جمله‌ای، پایه شعاعی و سیگموئید، می‌تواند با انتقال فضای ورودی داده‌ها به فضایی با بعد بالاتر در طبقه‌بندی غیرخطی داده‌ها استفاده شود [24]. در بیشتر موارد طبقه‌بندی داده‌ها با استفاده از دو نوع ماشین پشتیبان C-SVM و Nu-SVM انجام می‌شود. در طبقه‌بندی به روش C-SVM پارامتر C براساس اطلاع از نویز موجود در داده‌ها انتخاب می‌شود. پارامتر Nu در روش Nu-SVM به‌عنوان مرز بالایی از خطاها و حد پایین برای بردار پشتیبان عمل می‌کند. با افزایش آن، در زمان افزایش حاشیه جداسازی کلاس‌ها، خطاها افزایش می‌یابند [25].

۳. نتایج و بحث

شکل ۲، صفحه ۱۰، طیف جذبی از چهار نمونه روغن را نشان می‌دهد که در آن طیف‌ها به‌خوبی از یکدیگر متمایز بودند و نشان می‌دهد امکان تبعیض در آن‌ها وجود دارد. تفاوت‌ها ممکن است مطابق با ویژگی‌های مختلف نمونه‌های روغن‌زیتون، مانند ترکیبات اسیدهای چرب، تری گلیسرول، گلیسرول، استرول، توکوفرول و فنل و غیره باشد [26].

یافته‌های ما نشان می‌دهد که روغن زیتون فرابکر تقریباً در تمام طیف‌ها جذب بالاتری نسبت به روغن‌های زیتون بکر و تصفیه شده و پومیس دارد و این امر با پژوهش‌های گذشته نیز مطابقت دارد [27]. از شکل ۲ می‌توان دریافت که دو پیک اصلی در ناحیه ۶۵۵-۶۵۰ نانومتر و ۷۵۸-۷۵۵ نانومتر وجود دارد که می‌تواند مربوط به اورتون سوم CH و CH₃ در طیف پومیس باشد [28]. هنگامی که این روغن به سایر روغن‌ها افزوده شود، در طیف اصلی این نوع روغن‌ها تغییراتی در این نواحی مشاهده می‌شود. در شکل ۲ پیک‌های جذبی موجود روغن فرابکر و بکر در ناحیه ۶۲۰-۶۱۰ نانومتر، ۶۸۰-۶۵۰ و ۷۷۰-۷۴۰ نانومتر مشاهده می‌شود که می‌توانند ناشی از اورتون سوم H₂O یا اورتون چهارم CH باشند [29]. از شکل ۲ می‌توان دریافت در روغن بکر در ناحیه ۸۵۰-۸۳۰ نانومتر پیک دیگری وجود دارد که می‌تواند مربوط به اورتون چهارم Arch باشد [30]. پیک‌های موجود در روغن‌زیتون تصفیه شده در ناحیه ۶۷۰-۶۵۵ نانومتر، ۷۶۵-۷۴۵ نانومتر ناشی از اورتون سوم H₂O یا اورتون چهارم CH و ۸۴۰-۸۱۰ نانومتر هستند که می‌تواند مربوط به اورتون چهارم Arch یا NH₂ باشد [28].

۱.۳. طبقه‌بندی نمونه‌ها به روش LDA

نتایج حاصل از روش LDA در تفکیک تیمارهای مختلف تحقیق به همراه صحت هر کدام در جدول ۱، صفحه ۱۳ نشان داده شده است. نتایج به‌دست آمده نشان داد که پیش‌پردازش MA، بر پایه تابع Quadratic بالاترین صحت را ارائه می‌دهد. در مرحله بعد پیش‌پردازش

SG بر پایه تابع Quadratic مدلی باصحت بهتر نسبت به سایر روش‌های پیش‌پردازش ارائه داد؛ بنابراین در این روش استفاده از پیش‌پردازش MA و SG و تابع Quadratic تأثیر بیش‌تری در ساخت مدل‌ها باصحت بالاتر داشتند. قابل‌ذکر است که ضرایب وزنی متفاوت بر صحت روش LDA بی‌تأثیر بود. انتخاب ضرایب وزنی مناسب به نوع داده‌ها بستگی دارد و معمولاً با آزمون‌وخطا می‌توان به روش مناسب اعمال ضریب وزنی دست‌یافت.

شکل ۳، صفحه ۱۱، عملکرد روش LDA با استفاده از پیش‌پردازش MA و تابع Quadratic در تفکیک تیمارهای تحقیق را نشان می‌دهد.

نقاط نشان‌داده‌شده در شکل ۳ میزان پراکندگی دسته‌های مختلف روغن‌زیتون را نشان می‌دهد

۲،۳. طبقه‌بندی نمونه‌ها به روش SVM

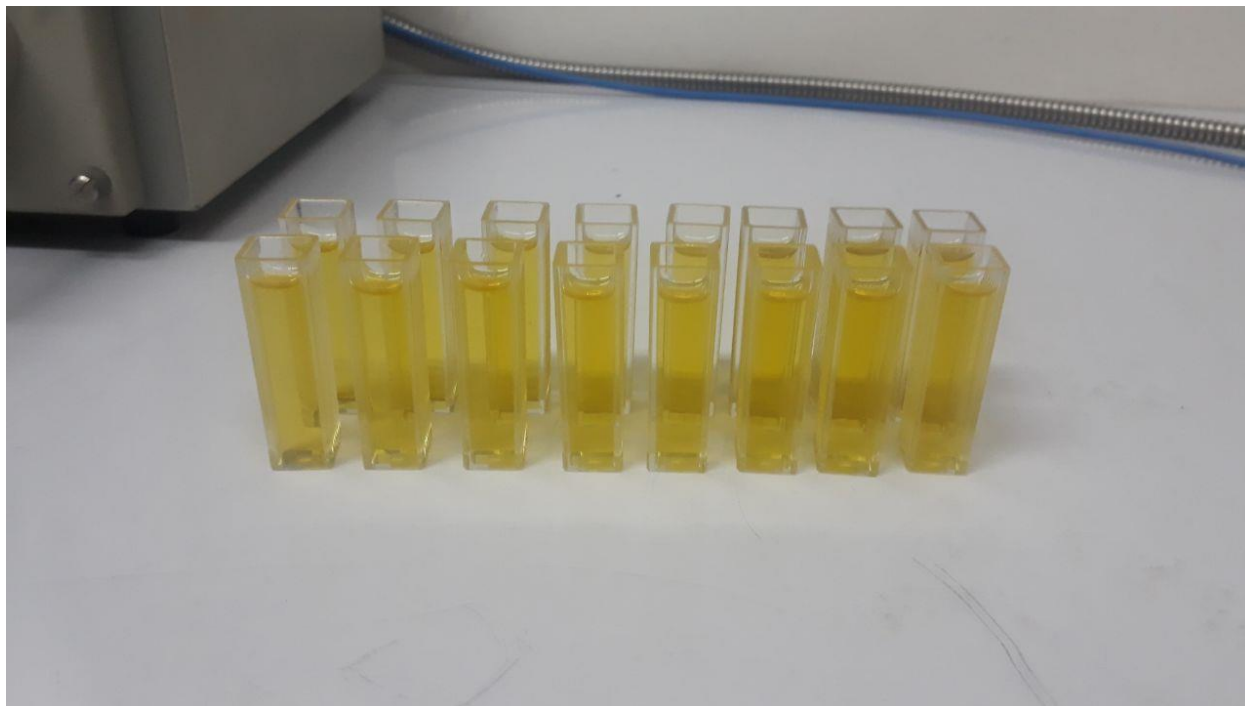
برای طبقه‌بندی نمونه‌ها در روش SVM از دو نوع C-SVM و Nu-SVM استفاده شد. پارامترهای C، Nu و γ با سعی و خطا و از طریق به‌حداقل‌رساندن خطای مدل در بخش واسنجی و اعتبارسنجی انتخاب شدند. زمانی که از روش Nu-SVM استفاده می‌شد صحت مدل در بخش آموزش و اعتبارسنجی به طرز چشمگیری کاهش پیدا می‌کرد بنابراین نتایج مربوط به این روش در ادامه ذکر نشده است. از چهار نوع تابع کرنل شامل خطی، درجه دوم، تابع پایه شعاعی و سیگموئید استفاده شد. تابع خطی با پیش‌پردازش MA باصحت آموزش ۱۰۰ و صحت اعتبارسنجی ۹۷/۵ بیش‌ترین صحت اعتبارسنجی را در طبقه‌بندی نمونه‌های روغن‌زیتون به خود اختصاص داد. (شکل ۴، صفحه ۱۲) نتایج حاصل از روش SVM در جدول ۲، صفحه ۱۴ نشان‌داده‌شده است. پیش‌پردازش SG نیز صحت آموزش بالایی را ارائه داد؛ ولی به دلیل صحت کم‌تر اعتبارسنجی نسبت به روش MA، نمی‌توان آن‌ها را بهترین مدل دانست. قابل‌ذکر است که تابع Linear و روش C-CVM تأثیر زیادی در مدل‌های باصحت بیش‌تر با پیش‌پردازش‌های مختلف داشتند. بنابراین هر دو روش مدل‌سازی SVM و LDA در تفکیک و دسته‌بندی تیمارهای تحقیق عملکرد بسیار قابل‌قبولی داشتند. روش SVM به دلیل صحت بیش‌تر در بخش واسنجی و اعتبارسنجی ارجحیت دارد. هم‌چنین جدول ۳، صفحه ۱۵، کلاس‌های پیش‌بینی‌شده توسط نمونه‌های دسته‌آموزش را نشان می‌دهد. طبق این جدول از ۳۲ نمونه دسته‌آموزش تنها ۲ نمونه اشتباه کلاس‌بندی شده‌اند. این ۲ نمونه مربوط به تقلب ۱۱ درصد بوده است که به اشتباه در دسته ۲۰ درصد تقلب کلاس‌بندی شده‌اند. بنابراین می‌توان نتیجه گرفت برای دقت بالای ۲۰ درصد کلاس بند با دقت ۱۰۰ درصد قادر به تفکیک نمونه‌های روغن‌زیتون براساس درصد‌های تقلب پومیس بوده است. با پیش‌بینی بسیار خوب توسط نمونه‌های دسته‌آموزش می‌توان مطمئن شد که بیش‌برازش و کم‌برازش در مدل اتفاق نیفتاده است و مدل از اعتبار بالایی برخوردار است. تاکنون پژوهشی که بتواند روغن پومیس را از انواع دیگر روغن‌زیتون متمایز کند گزارش نشده است و وجه اشتراک تنها در مقایسه با یک نوع از انواع روغن‌زیتون (به‌جز روغن‌زیتون تصفیه‌شده) و سایر روغن‌های خوراکی بوده است. تحقیقات گذشته نیز نشان داد از بین روش‌های مختلف طبقه‌بندی هم‌چون ماشین بردار پشتیبان، شبکه عصبی مصنوعی و آدابوست، روش ماشین بردار پشتیبان به‌عنوان بهترین روش با دقت ۸۶،۵۲٪ قادر به تشخیص نمونه‌های تقلبی روغن زیتون با استفاده از ماشین بویایی بوده است [31]. در سایر پژوهش‌ها اگرچه تشخیص تقلب روغن‌زیتون فرابکر با پومیس به روش طیف‌سنجی اعم از رامان، تبدیل فوریه و غیره باصحت قابل‌قبولی ارائه شده است؛ ولی در پژوهش حاضر نخستین‌بار است که محدوده پایین فرسرخ نزدیک یعنی بازه ۷۸۰-۸۵۰ نانومتر در کنار ناحیه مرئی به‌منظور شناسایی دقیق تقلب پومیس در انواع مختلف روغن‌زیتون بهره‌گیری شده است. این محدوده طیفی در اکثر طیف‌سنج‌های مرئی/فروسرخ نزدیک وجود دارد و نیازی به محدوده بالای طیفی NIR نیست. روش‌های متداول دیگر مانند کروماتوگرافی گازی می‌تواند مقادیر پایین روغن پومیس را با اندازه‌گیری تری‌ترین دی‌الکل‌ها و میزان موم در روغن تشخیص دهد اما در برابر روش طیف‌سنجی فرسرخ نزدیک نیاز به آماده‌سازی، حلال‌های شیمیایی، هزینه بالا استهلاک دستگاه و اپراتور متخصص دارد.

۴. نتیجه‌گیری

در این پژوهش توانمندی فناوری طیف‌سنجی مرئی/فروسرخ نزدیک (Vis/NIR) به‌منظور طبقه‌بندی نمونه‌های روغن‌زیتون با درصدهای مختلف تقلب با روش‌های بازشناسی الگو شامل LDA و SVM مورد بررسی قرار گرفت. نتایج به‌دست‌آمده نشان داد که طیف‌سنجی (Vis/NIR) قادر به تفکیک نمونه‌های روغن‌زیتون بر اساس نرخ تقلب پومیس است. هرچند روش LDA توانست باصحت خوبی نمونه‌های روغن‌زیتون را دسته‌بندی کند، اما روش SVM باصحت آموزش ۹۶/۶۹ و صحت اعتبارسنجی ۹۵/۸۶ بهترین روش برای دسته‌بندی نمونه‌ها بود. همچنین، بر اساس نتایج حاصله، تابع خطی، بهترین تابع برای ساخت مدل دسته‌بندی به روش SVM تعیین شد. همچنین پیشنهاد می‌شود از این محدوده طیفی به منظور شناسایی حد شناسایی پومیس و اختلاط سایر روغن‌ها مانند آفتابگردان، کلزا، سویا و غیره در روغن زیتون و تشخیص تقلب استفاده شود.

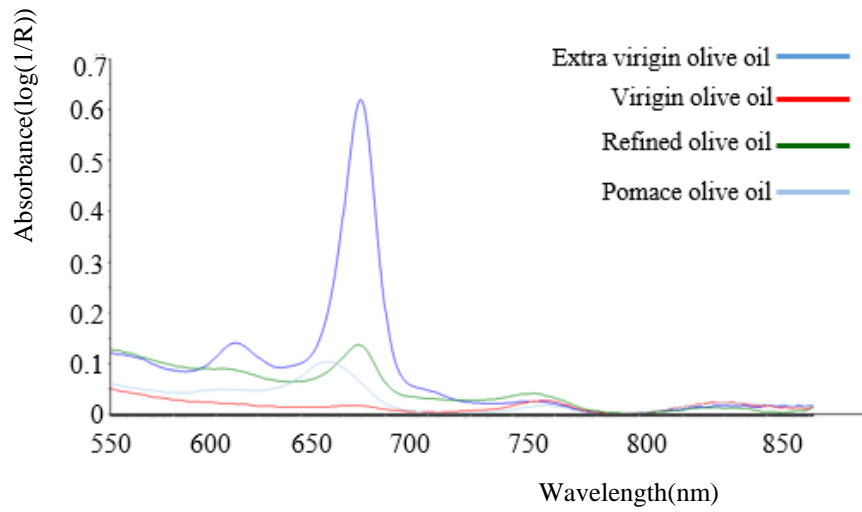
۵. تعارض منافع

هیچ‌گونه تعارض منافع توسط نویسندگان بیان نشده است.

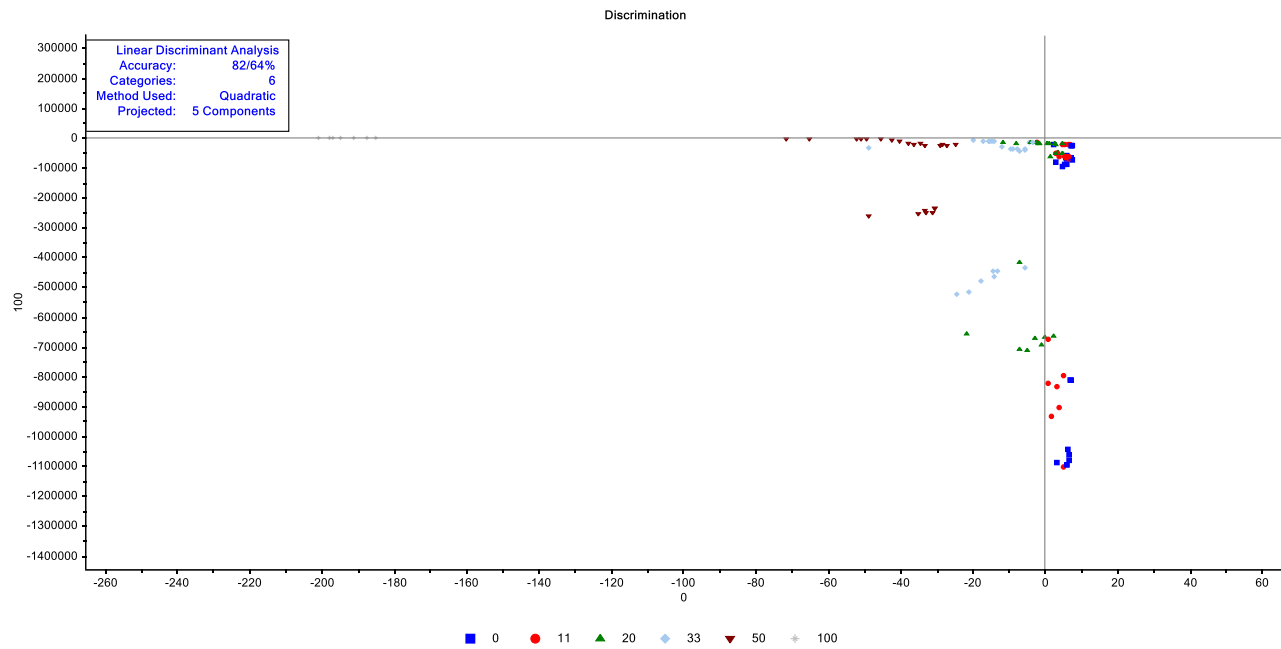


شکل (1) نمونه‌های مورد استفاده در پژوهش با درصد تقلبات مختلف
Figure (1) Samples used in the research with different percentages of fraud

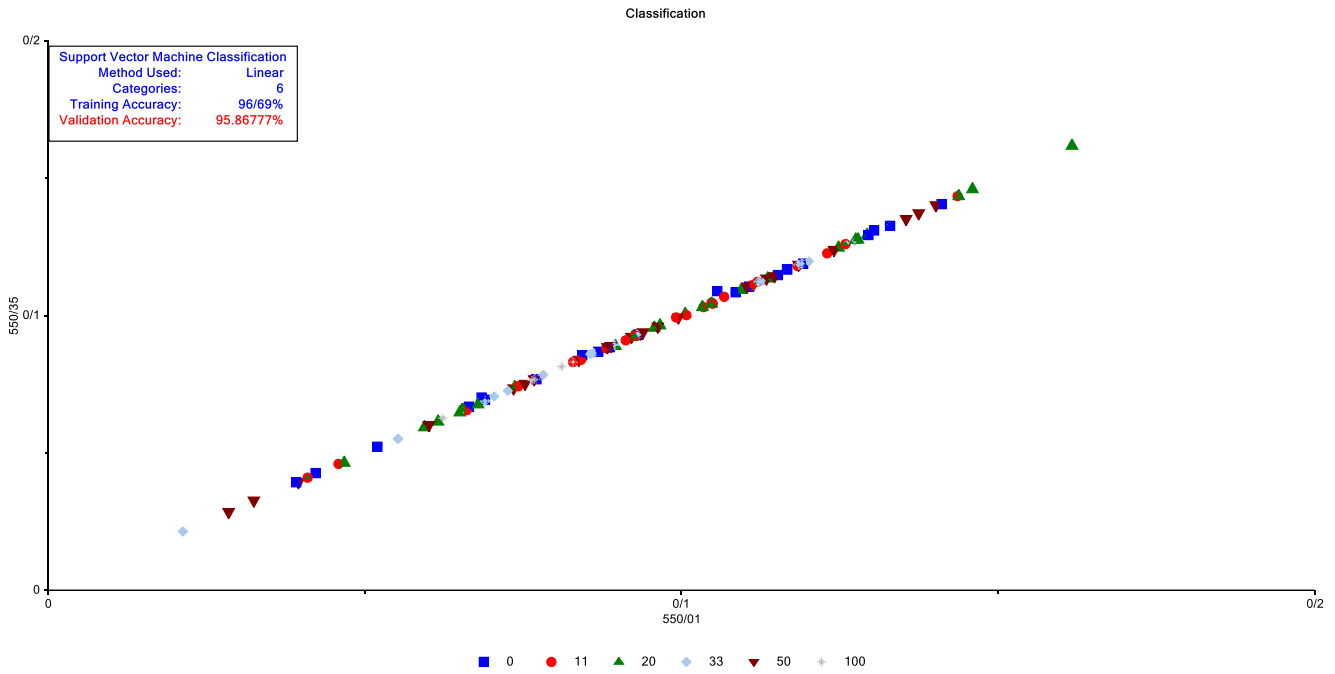
Journal Pre-proofs



شکل (۲) طیف جذبی از چهار نمونه مختلف روغن زیتون
Fig (2) Absorption spectra of four different olive oil samples



شکل (۳) عملکرد روش LDA با استفاده از تابع درجه دوم در جداسازی تیمارهای تحقیقاتی
Fig (3) Performance of LDA method using quadratic function in separating research treatments



شکل (۴) تابع خطی C-SVM با صحت ۹۶/۶۹ درصد
Fig (4) C-SVM linear function with 100% accuracy

جدول (1) نتایج مدل LDA نمونه‌های روغن‌زیتون

Table (1) Results of LDA model of olive oil samples

Accuracy (%)	Function Type	Preprocessing
79.99	Linear	MA
82.64	Quadratic	MA
81.82	Mahalanobis	MA
49.59	Linear	MA+SNV
74.38	Quadratic	MA+SNV
73.55	Mahalanobis	MA+SNV
79.34	Linear	SG
81.56	Quadratic	SG
80.12	Mahalanobis	SG
43.84	Linear	SG+SNV
71.94	Quadratic	SG+SNV
73.06	Mahalanobis	SG+SNV
80.42	Linear	MA+D ₁
78.34	Quadratic	MA+D ₁
72.94	Mahalanobis	MA+D ₁
79.96	Linear	SG+D ₁
72.91	Quadratic	SG+D ₁
74.91	Mahalanobis	SG+D ₁

جدول (۲) نتایج مدل‌های Nu-SVM و C-SVM تحت توابع کرنل

Table (2) Results of Nu-SVM and C-SVM models under kernel function

Validation(%)	Calibration(%)	γ	C	Function Type	Preprocessing
95.86	96.69	1	100	Linear	MA
19	18.8	10	1	Polynomial	MA
29.75	23.97	1	100	Radial	MA
27.27	24.79	0.01	100	Sigmoid	MA
83.47	94.21	1	1	Linear	MA+SNV
48.76	58.68	100	10	Polynomial	MA+SNV
80.99	93.39	0.01	100	Radial	MA+SNV
22.31	18.18	1	0.1	Sigmoid	MA+SNV
85.97	96.86	1	100	Linear	SG
87.5	18.18	10	1	Polynomial	SG
28.73	25.84	1	100	Radial	SG
26.84	26.73	0.01	100	Sigmoid	SG
85.95	100	1	100	Linear	SG+SNV
48.76	58.68	100	10	Polynomial	SG+SNV
80.99	93.39	10	1	Radial	SG+SNV
24.79	18.18	0.1	0.01	Sigmoid	SG+SNV
27.27	23.97	1	100	Linear	MA+D
21.87	24.38	100	10	Polynomial	MA+D
26.44	18.18	100	100	Radial	MA+D
29.37	43.13	1	1	Sigmoid	MA+D
33.05	21.49	1	1	Linear	SG+D
19.45	18.18	100	10	Polynomial	SG+D
24.79	18.18	100	100	Radial	SG+D
15.74	18.18	100	100	Sigmoid	SG+D

جدول (۳) ماتریس اغتشاش حاصل از نمونه‌های دسته آزمون

Table (3) The confusion matrix obtained from the samples of the test group

Predicted amount	Actual amount	Sample number	Predicted amount	Actual amount	Sample number
۰	۰	۵۶	۰	۰	۱
۰	۰	۶۳	۰	۰	۴
۱۱	۱۱	۷۰	۱۰۰	۱۰۰	۱۰
۱۱	۱۱	۷۳	۱۱	۱۱	۱۳
۲۰	۲۰	۷۷	۱۱	۲۰	۱۶
۳۳	۳۳	۸۴	۳۳	۳۳	۱۹
۵۰	۵۰	۹۱	۵۰	۵۰	۲۲
۲۰	۲۰	۹۸	۵۰	۵۰	۲۴
۱۱	۱۱	۱۰۵	۱۱	۱۱	۲۸
۳۳	۳۳	۱۱۲	۳۳	۳۳	۳۱
۵۰	۵۰	۱۱۹	۵۰	۵۰	۳۴
۱۱	۲۰	۱۲۶	۲۰	۲۰	۳۷
۲۰	۲۰	۱۳۳	۳۳	۳۳	۴۰
۳۳	۳۳	۱۴۰	۵۰	۵۰	۴۳
۵۰	۵۰	۱۴۷	۵۰	۵۰	۴۶
۱۰۰	۱۰۰	۱۵۴	۰	۰	۴۹

- [1] M. Rashvand, O. Mahmood, H. Mobli, and M. Soltanifiroz, "Evaluation of olive oil adulteration detection system based on image processing and dielectric spectroscopy," *Agric. Mach. Mech. Res. J.*, vol. 6, no. 2, pp. 29–19, 2017.
- [2] B. Vega-Márquez, I. Nepomuceno-Chamorro, N. Jurado-Campos, and C. Rubio-Escudero, "Deep Learning Techniques to Improve the Performance of Olive Oil Classification," *Front. Chem.*, vol. 7, no. January, pp. 1–10, 2020, doi: 10.3389/fchem.2019.00929.
- [3] A. M. Giuffré, M. Caracciolo, C. Zappia, M. Capocasale, and M. Poiana, "Effect of heating on chemical parameters of extra virgin olive oil, pomace olive oil, soybean oil and palm oil," *Ital. J. Food Sci.*, vol. 30, no. 4, pp. 715–739, 2018.
- [4] M. I. Covas, "Olive oil and the cardiovascular system," *Pharmacol. Res.*, vol. 55, no. 3, pp. 175–186, 2007, doi: 10.1016/j.phrs.2007.01.010.
- [5] C. Faulh, F. Reniero, and C. Guillou, "1H NMR as a tool for the analysis of mixtures of virgin olive oil with oils of different botanical origin," *Magn. Reson. Chem.*, vol. 38, no. 6, pp. 436–443, 2000, doi: 10.1002/1097-458X(200006)38:6<436::AID-MRC672>3.0.CO;2-X.
- [6] S. Ok, "Detection of olive oil adulteration by low-field NMR relaxometry and UV-Vis spectroscopy upon mixing olive oil with various edible oils," *Grasas y Aceites*, vol. 68, no. 1, pp. 1–13, 2017, doi: 10.3989/gya.0678161.
- [7] S. A. Mahesar, M. Lucarini, A. Durazzo, A. Santini, A. I. Lampe, and J. Kiefer, "Application of Infrared Spectroscopy for Functional Compounds Evaluation in Olive Oil : A Current Snapshot," vol. 2019, 2019.
- [8] S. Cicerale, X. A. Conlan, A. J. Sinclair, and R. S. J. Keast, "Chemistry and health of olive oil phenolics," *Crit. Rev. Food Sci. Nutr.*, vol. 49, no. 3, pp. 218–236, 2009, doi: 10.1080/10408390701856223.
- [9] P. Vossen, "Olive oil: History, production, and characteristics of the world's classic oils," *HortScience*, vol. 42, no. 5, pp. 1093–1100, 2007, doi: 10.21273/hortsci.42.5.1093.
- [10] R. Rodriguez-Rodriguez, M. D. Herrera, M. A. de Sotomayor, and V. Ruiz-Gutierrez, "Pomace Olive Oil Improves Endothelial Function in Spontaneously Hypertensive Rats by Increasing Endothelial Nitric Oxide Synthase Expression," *Am. J. Hypertens.*, vol. 20, no. 7, pp. 728–734, 2007, doi: 10.1016/j.amjhyper.2007.01.012.
- [11] K. Antonopoulos, N. Valet, D. Spiratos, and G. Siragakis, "Olive oil and pomace olive oil processing," *Grasas y Aceites*, vol. 57, no. 1, pp. 56–67, 2006, doi: 10.3989/gya.2006.v57.i1.22.
- [12] F. Pérez-Jiménez, J. Ruano, P. Perez-Martinez, F. Lopez-Segura, and J. Lopez-Miranda, "The influence of olive oil on human health: Not a question of fat alone," *Mol. Nutr. Food Res.*, vol. 51, no. 10, pp. 1199–1208, 2007, doi: 10.1002/mnfr.200600273.
- [13] L. Di Giovacchino, S. Sestili, and D. Di Vincenzo, "Influence of olive processing on virgin olive oil quality," *Eur. J. Lipid Sci. Technol.*, vol. 104, no. 9–10, pp. 587–601, 2002, doi: 10.1002/1438-9312(200210)104:9/10<587::AID-EJLT587>3.0.CO;2-M.
- [14] V. Maestrello, P. Solovyev, L. Bontempo, L. Mannina, and F. Camin, "Nuclear magnetic resonance spectroscopy in extra virgin olive oil authentication," *Compr. Rev. Food Sci. Food Saf.*, vol. 21, no. 5, pp. 4056–4075, 2022, doi: 10.1111/1541-4337.13005.
- [15] D. Tura, C. Gigliotti, S. Pedò, O. Failla, D. Bassi, and A. Serraiocco, "Influence of cultivar and site of cultivation on levels of lipophilic and hydrophilic antioxidants in virgin olive oils (*Olea Europea L.*) and correlations with oxidative stability," *Sci. Hortic. (Amsterdam)*, vol. 112, no. 1, pp. 108–119, 2007, doi: 10.1016/j.scienta.2006.12.036.
- [16] S. K. Dwivedi, M. Vishwakarma, and P. A. Soni, "Advances and Researches on Non Destructive Testing: A Review," *Mater. Today Proc.*, vol. 5, no. 2, pp. 3690–3698, 2018, doi: 10.1016/j.matpr.2017.11.620.
- [17] H. Xiao, L. Feng, D. Song, K. Tu, J. Peng, and L. Pan, "Grading and sorting of grape berries using visible-near infrared spectroscopy on the basis of multiple inner quality parameters," *Sensors (Switzerland)*, vol. 19, no. 11, 2019, doi: 10.3390/s19112600.
- [18] P. Galvin-King, S. A. Haughey, and C. T. Elliott, "Garlic adulteration detection using NIR and FTIR spectroscopy and chemometrics," *J. Food Compos. Anal.*, vol. 96, p. 103757, 2021, doi: 10.1016/j.jfca.2020.103757.
- [19] M. De Luca *et al.*, "Derivative FTIR spectroscopy for cluster analysis and classification of morocco olive oils," *Food Chem.*, vol. 124, no. 3, pp. 1113–1118, 2011, doi: 10.1016/j.foodchem.2010.07.010.
- [20] R. De Bei *et al.*, "Non-destructive measurement of grapevine water potential using near infrared

- spectroscopy De Bei, R., Cozzolino, D., Sullivan, W., Cynkar, W., Fuentes, S., Damberg, R., Pech, J., & Tyerman, S. D. (2011). Non-destructive measurement of grapevine water potential,” *Aust. J. Grape Wine Res.*, vol. 17, no. 1, pp. 62–71, 2011, doi: 10.1111/j.1755-0238.2010.00117.x.
- [21] X. Li *et al.*, “Review of NIR spectroscopy methods for nondestructive quality analysis of oilseeds and edible oils,” *Trends Food Sci. Technol.*, vol. 101, no. March, pp. 172–181, 2020, doi: 10.1016/j.tifs.2020.05.002.
- [22] S. Bureau and B. Gouble, “rapid and non-destructive analysis of apricot fruit quality using ft-near infrared spectroscopy,” *Mater. Struct.*, no. 2014, pp. 32–35, 2015.
- [23] A. C. de Lima, L. Aceña, M. Mestres, and R. Boqué, “An Overview of the Application of Multivariate Analysis to the Evaluation of Beer Sensory Quality and Shelf-Life Stability,” *Foods*, vol. 11, no. 14, pp. 1–15, 2022, doi: 10.3390/foods11142037.
- [24] A. Amari, N. El Bari, and B. Bouchikhi, “Electronic Nose for Anchovy Freshness Monitoring Based on Sensor Array and Pattern Recognition Methods: Principal Components Analysis, Linear Discriminant Analysis and Support Vector Machine,” *Int. J. Comput.*, vol. 6, no. 3, pp. 61–67, 2014, doi: 10.47839/ijc.6.3.452.
- [25] M. W. Huang, C. W. Chen, W. C. Lin, S. W. Ke, and C. F. Tsai, “SVM and SVM ensembles in breast cancer prediction,” *PLoS One*, vol. 12, no. 1, pp. 1–14, 2017, doi: 10.1371/journal.pone.0161501.
- [26] Z. Liu and J. Tan, “Qualitative and quantitative detection of Sudan I and II adulterated in chili powders by front-face synchronous fluorescence spectroscopy : Aggregation-induced emission in solid food,” pp. 1–21, 2022.
- [27] N. Abu-khalaf and M. Hmidat, “Visible / Near Infrared (VIS / NIR) spectroscopy as an optical sensor for evaluating olive oil quality,” *Comput. Electron. Agric.*, vol. 173, no. December 2019, p. 105445, 2020, doi: 10.1016/j.compag.2020.105445.
- [28] H. Cen, Y. He, and M. Huang, “Combination and comparison of multivariate analysis for the identification of orange varieties using visible and near infrared reflectance spectroscopy,” *Eur. Food Res. Technol.*, vol. 225, no. 5–6, pp. 699–705, 2007, doi: 10.1007/s00217-006-0470-2.
- [29] Y. G. M. Kongbonga *et al.*, “Characterization of Vegetable Oils by Fluorescence Spectroscopy,” *Food Nutr. Sci.*, vol. 02, no. 07, pp. 692–699, 2011, doi: 10.4236/fns.2011.27095.
- [30] T. Wang, Y. Zhang, Y. Liu, Z. Zhang, and T. Yan, “Intelligent Evaluation of Stone Cell Content of Korla Fragrant Pears by Vis/NIR Reflection Spectroscopy,” *Foods*, vol. 11, no. 16, 2022, doi: 10.3390/foods11162391.
- [31] m Zarezade, Mohammad Reza, abonajmi, M. Ghasemi-varnamkhasti, and fazarikia, “Estimation of the Best Classification Algorithm and Fraud Detection of Olive Oil by Olfaction Machine,” vol. 11, no. 2, pp. 371–383, 2021, doi: 10.22067/jam.v11i2.84105.
- [32] M. Bahmaei, R. Nazeri, F. Kalantari, N. T. Branch, and S. B. Company, “Archive of SID Detection of adulteration in canola-virgin olive admixture using instrumental gas liquid chromatography and UV spectroscopy”.

Feasibility assessment of employing near-infrared/visible spectroscopy for detecting olive pomace oil adulteration using LDA and SVM identification methods

Shirin Asadian¹, Ahmad Banakar^{2*}, Bahareh Jamshidi³

¹ Msc Student, Mechanical Engineering of Biosystems Department, Tarbiat Modares University, Tehran, Iran.

² Associate Professor, Mechanical Engineering of Biosystem Department, Tarbiat Modares University, Tehran, Iran.

³ Associate Professor, Agricultural Engineering Research Institute, Agricultural Research Education and Extension Organization (AREEO), Karaj, Iran.

Extended Abstract

Introduction

Olive oil is a commercially valuable product, so it is often adulterated with other edible oils such as sunflower, hazelnut oil, pomace oil and other oils that have a fatty acid composition close to olive oil. Accordingly, quality control of this oil is one of the issues and it is very important in health and food safety. After extracting olive oil by physical techniques, a residue is created that contains more than 8% oil and can be extracted by a solvent (usually hexane). This oil is called pomace olive oil. Due to the characteristics of the solvent, the oil extracted with the solvent has more trace elements. Pomace olive oil is carcinogenic due to extraction with chemical solvents and as a result of being contaminated with benzo(a)pyrene, monochloropropane DLs and cyclic hydrocarbons and has an adverse effect on nerves and other body organs. Currently, gas chromatography method is used to detect adulteration in olive oil. Also, in the health system, to control the quality of food, including olive oil, compared to other frauds and the quantity of fatty acids, it is done by chemical tests and gas chromatography tests, which are expensive and time-consuming methods and require specialized quality control laboratories and specialized human resources. It should be noted that the presence of pomace olive oil in virgin olive oil and refined olive oil cannot be detected by gas chromatography, and detection of adulteration in these cases is not applicable by chromatography. Past studies show the high importance of spectroscopy as a fast, simple, low-cost method with acceptable accuracy in measuring the quality of agricultural products. For this reason, its use in different ways and for different products is expanding day by day. In this research, pomace olive oil adulteration was identified using a fast and non-destructive visible-near-infrared (Vis/NIR) spectroscopy method without measuring qualitative properties using pattern recognition methods such as LDA and SVM.

Materials and Method

The extra-virgin, virgin and refined olive oils used in this research were prepared from the Etkha olive oil factory in Varamin, Tehran province. The sample of raw and non-edible pomace oil was also obtained from a pharmaceutical, cosmetic and health company. Extra virgin olive oil, virgin and refined olive oil were mixed with pomace oil with weight percentages of 11, 20, 33, 50 and 100%. These percentages were chosen according to the frauds in the market. A total of 160 samples were prepared with ten replicates for each treatment. The visible/near infrared spectrum was used using a two-way fiber, which is the central part for light transmission from the sample to the detector and the outer fiber for light transmission from the radiation source to the sample, and in the transition state and in the 200-850 nm region using Spectrometer model usb2000 made by ocean optic company and recorded with 1 nanometer resolution by OceanView software. The spectral data obtained from the spectrometer, in addition to the sample information, also includes unwanted background information and noises. Hence, in order to achieve stable, accurate and reliable calibration models, there is a need to pre-process spectral data before formulating classification models appropriate to the data. After recording and saving the obtained spectra for the analysis of the received spectra, data preprocessing was done. Part of the data from the 550 nm area already had high noise and could not be smoothed by the usual pre-processing methods, so they were removed. After the transmission spectra obtained from all the samples were converted into absorption spectra in order to create a linear relationship with the molecular concentration of the

sample, to prepare and pre-process the spectra some smoothing and noise reduction methods such as moving average (MA), Savitzky Golay (SG), Standard normal variate (SNV) and first derivative (D1) were used. The samples were divided into two groups of training (80 percent) for the development of calibration models, and the test group (20 percent) for the external validation of the developed calibration models in order to check the ability and accuracy of the models in classifying different percentages of fraud of unknown samples. Linear discriminant analysis (LDA) and support vector machine (SVM) were used in this research. Evaluation of the built model with full cross-validation method is used to select the model with the highest classification ability among the statistical models. In other words, after building the model in the re-evaluation phase, the reliability and stability of the model was also investigated. In this research, three types of Linear, Quadratic and Mahalanobis functions and four different weighting coefficients (1/SDev, Constant, Block weighting and Downweight) were used to create a classification model using LDA method.

Results and discussion

Our findings show that extra-virgin olive oil has higher absorption in almost all spectra than virgin and refined olive oils and pomace, and this is also consistent with previous researches. Three types of Linear, Quadratic and Mahalanobis functions were used to create the classification model by LDA method. The results obtained from LDA method showed that MA preprocessing based on Quadratic function provides the highest accuracy. In the next step, SG preprocessing based on the Quadratic function provided a better accurate model than other preprocessing methods; Therefore, in this method, the use of MA and SG preprocessing and the Quadratic function had a greater effect in building models with higher accuracy. It is worth mentioning that different weight coefficients had no effect on the accuracy of LDA method. Choosing the appropriate weighting coefficients depends on the type of data and usually by trial and error, the appropriate method of applying the weighting coefficient can be achieved. Two types of C-SVM and Nu-SVM were used to classify the samples in the SVM method. Four types of kernel functions including linear, quadratic, radial basis function and sigmoid were used. The linear function with MA preprocessing with training accuracy of 100 and validation accuracy of 97.5 had the highest validation accuracy in the classification of olive oil samples. It is worth noting that Linear function and C-CVM method had a great effect on more accurate models with different pre-processing. Therefore, both SVM and LDA modeling methods had a very acceptable performance in separating and categorizing the research treatments. The SVM method is preferred due to its greater accuracy in the calibration and validation section.

Conclusions

In this research, a visible/near infrared (Vis/NIR) spectroscopic system was investigated in order to classify olive oil samples with different adulteration percentages with pattern recognition methods including LDA and SVM. The obtained results showed that spectroscopy (Vis/NIR) is able to distinguish olive oil samples based on pomace olive oil adulteration rate. Although the LDA method was able to classify the olive oil samples with acceptable accuracy, the SVM method with a training accuracy of 96.69 and a validation accuracy of 95.86 was the best method for classifying the samples. Also, based on the results, the linear function, the best function for building the classification model was determined by SVM method.

Keywords: Fraud, Linear Discriminant Analysis, Pomace Oil, Support Vector Machine

جدول (۱) نتایج مدل LDA نمونه‌های روغن‌زیتون

Table (1) Results of LDA model of olive oil samples

Accuracy (%)	Function Type	Preprocessing
79.99	Linear	MA
82.64	Quadratic	MA
81.82	Mahalanobis	MA
49.59	Linear	MA+SNV
74.38	Quadratic	MA+SNV
73.55	Mahalanobis	MA+SNV
79.34	Linear	SG
81.56	Quadratic	SG
80.12	Mahalanobis	SG
43.84	Linear	SG+SNV
71.94	Quadratic	SG+SNV
73.06	Mahalanobis	SG+SNV
80.42	Linear	MA+D ₁
78.34	Quadratic	MA+D ₁
72.94	Mahalanobis	MA+D ₁
79.96	Linear	SG+D ₁
72.91	Quadratic	SG+D ₁
74.91	Mahalanobis	SG+D ₁

جدول (۲) نتایج مدل‌های Nu-SVM و C-SVM تحت توابع کرنل

Table (2) Results of Nu-SVM and C-SVM models under kernel function

Validation(%)	Calibration(%)	γ	C	Function Type	Preprocessing
95.86	96.69	1	100	Linear	MA
19	18.8	10	1	Polynomial	MA
29.75	23.97	1	100	Radial	MA
27.27	24.79	0.01	100	Sigmoid	MA
83.47	94.21	1	1	Linear	MA+SNV
48.76	58.68	100	10	Polynomial	MA+SNV
80.99	93.39	0.01	100	Radial	MA+SNV
22.31	18.18	1	0.1	Sigmoid	MA+SNV
85.97	96.86	1	100	Linear	SG
87.5	18.18	10	1	Polynomial	SG
28.73	25.84	1	100	Radial	SG
26.84	26.73	0.01	100	Sigmoid	SG
85.95	100	1	100	Linear	SG+SNV
48.76	58.68	100	10	Polynomial	SG+SNV
80.99	93.39	10	1	Radial	SG+SNV
24.79	18.18	0.1	0.01	Sigmoid	SG+SNV
27.27	23.97	1	100	Linear	MA+D

21.87	24.38	100	10	Polynomial	MA+D
26.44	18.18	100	100	Radial	MA+D
29.37	43.13	1	1	Sigmoid	MA+D
33.05	21.49	1	1	Linear	SG+D
19.45	18.18	100	10	Polynomial	SG+D
24.79	18.18	100	100	Radial	SG+D
15.74	18.18	100	100	Sigmoid	SG+D

جدول (۳) ماتریس اغتشاش حاصل از نمونه‌های دسته آزمون

Table (3) The confusion matrix obtained from the samples of the test group

Predicted amount	Actual amount	Sample number	Predicted amount	Actual amount	Sample number
۰	۰	۵۶	۰	۰	۱
۰	۰	۶۳	۰	۰	۴
۱۱	۱۱	۷۰	۱۰۰	۱۰۰	۱۰
۱۱	۱۱	۷۳	۱۱	۱۱	۱۳
۲۰	۲۰	۷۷	۱۱	۲۰	۱۶
۳۳	۳۳	۸۴	۳۳	۳۳	۱۹
۵۰	۵۰	۹۱	۵۰	۵۰	۲۲
۲۰	۲۰	۹۸	۵۰	۵۰	۲۴
۱۱	۱۱	۱۰۵	۱۱	۱۱	۲۸
۳۳	۳۳	۱۱۲	۳۳	۳۳	۳۱
۵۰	۵۰	۱۱۹	۵۰	۵۰	۳۴
۱۱	۲۰	۱۲۶	۲۰	۲۰	۳۷
۲۰	۲۰	۱۳۳	۳۳	۳۳	۴۰
۳۳	۳۳	۱۴۰	۵۰	۵۰	۴۳
۵۰	۵۰	۱۴۷	۵۰	۵۰	۴۶
۱۰۰	۱۰۰	۱۵۴	۰	۰	۴۹

- [1] M. Rashvand, O. Mahmood, H. Mobli, and M. Soltanifiroz, "Evaluation of olive oil adulteration detection system based on image processing and dielectric spectroscopy," *Agric. Mach. Mech. Res. J.*, vol. 6, no. 2, pp. 29–19, 2017.
- [2] B. Vega-Márquez, I. Nepomuceno-Chamorro, N. Jurado-Campos, and C. Rubio-Escudero, "Deep Learning Techniques to Improve the Performance of Olive Oil Classification," *Front. Chem.*, vol. 7, no. January, pp. 1–10, 2020, doi: 10.3389/fchem.2019.00929.
- [3] A. M. Giuffré, M. Caracciolo, C. Zappia, M. Capocasale, and M. Poiana, "Effect of heating on chemical parameters of extra virgin olive oil, pomace olive oil, soybean oil and palm oil," *Ital. J. Food Sci.*, vol. 30, no. 4, pp. 715–739, 2018.
- [4] M. I. Covas, "Olive oil and the cardiovascular system," *Pharmacol. Res.*, vol. 55, no. 3, pp. 175–186, 2007, doi: 10.1016/j.phrs.2007.01.010.
- [5] C. Fauhl, F. Reniero, and C. Guillou, "¹H NMR as a tool for the analysis of mixtures of virgin olive oil with oils of different botanical origin," *Magn. Reson. Chem.*, vol. 38, no. 6, pp. 436–443, 2000, doi: 10.1002/1097-458X(200006)38:6<436::AID-MRC672>3.0.CO;2-X.
- [6] S. Ok, "Detection of olive oil adulteration by low-field NMR relaxometry and UV-Vis spectroscopy upon mixing olive oil with various edible oils," *Grasas y Aceites*, vol. 68, no. 1, pp. 1–13, 2017, doi: 10.3989/gya.0678161.
- [7] S. A. Mahesar, M. Lucarini, A. Durazzo, A. Santini, A. I. Lampe, and J. Kiefer, "Application of Infrared Spectroscopy for Functional Compounds Evaluation in Olive Oil : A Current Snapshot," vol. 2019, 2019.
- [8] S. Cicerale, X. A. Conlan, A. J. Sinclair, and R. S. J. Keast, "Chemistry and health of olive oil phenolics," *Crit. Rev. Food Sci. Nutr.*, vol. 49, no. 3, pp. 218–236, 2009, doi: 10.1080/10408390701856223.
- [9] P. Vossen, "Olive oil: History, production, and characteristics of the world's classic oils," *HortScience*, vol. 42, no. 5, pp. 1093–1100, 2007, doi: 10.21273/hortsci.42.5.1093.
- [10] R. Rodriguez-Rodriguez, M. D. Herrera, M. A. de Sotomayor, and V. Ruiz-Gutierrez, "Pomace Olive Oil Improves Endothelial Function in Spontaneously Hypertensive Rats by Increasing Endothelial Nitric Oxide Synthase Expression," *Am. J. Hypertens.*, vol. 20, no. 7, pp. 728–734, 2007, doi: 10.1016/j.amjhyper.2007.01.012.
- [11] K. Antonopoulos, N. Valet, D. Spiratos, and G. Siragakis, "Olive oil and pomace olive oil processing," *Grasas y Aceites*, vol. 57, no. 1, pp. 56–67, 2006, doi: 10.3989/gya.2006.v57.i1.22.
- [12] F. Pérez-Jiménez, J. Ruano, P. Perez-Martinez, F. Lopez-Segura, and J. Lopez-Miranda, "The influence of olive oil on human health: Not a question of fat alone," *Mol. Nutr. Food Res.*, vol. 51, no. 10, pp. 1199–1208, 2007, doi: 10.1002/mnfr.200600273.
- [13] L. Di Giovacchino, S. Sestili, and D. Di Vincenzo, "Influence of olive processing on virgin olive oil quality," *Eur. J. Lipid Sci. Technol.*, vol. 104, no. 9–10, pp. 587–601, 2002, doi: 10.1002/1438-9312(200210)104:9/10<587::AID-EJLT587>3.0.CO;2-M.
- [14] V. Maestrello, P. Solovyev, L. Bontempo, L. Mannina, and F. Camin, "Nuclear magnetic resonance spectroscopy in extra virgin olive oil authentication," *Compr. Rev. Food Sci. Food Saf.*, vol. 21, no. 5, pp. 4056–4075, 2022, doi: 10.1111/1541-4337.13005.
- [15] D. Tura, C. Gigliotti, S. Pedò, O. Failla, D. Bassi, and A. Serraiocco, "Influence of cultivar and site of cultivation on levels of lipophilic and hydrophilic antioxidants in virgin olive oils (*Olea Europea L.*) and correlations with oxidative stability," *Sci. Hortic. (Amsterdam)*, vol. 112, no. 1, pp. 108–119, 2007, doi: 10.1016/j.scienta.2006.12.036.
- [16] S. K. Dwivedi, M. Vishwakarma, and P. A. Soni, "Advances and Researches on Non Destructive Testing: A Review," *Mater. Today Proc.*, vol. 5, no. 2, pp. 3690–3698, 2018, doi: 10.1016/j.matpr.2017.11.620.

- [17]H. Xiao, L. Feng, D. Song, K. Tu, J. Peng, and L. Pan, "Grading and sorting of grape berries using visible-near infrared spectroscopy on the basis of multiple inner quality parameters," *Sensors (Switzerland)*, vol. 19, no. 11, 2019, doi: 10.3390/s19112600.
- [18]P. Galvin-King, S. A. Haughey, and C. T. Elliott, "Garlic adulteration detection using NIR and FTIR spectroscopy and chemometrics," *J. Food Compos. Anal.*, vol. 96, p. 103757, 2021, doi: 10.1016/j.jfca.2020.103757.
- [19]M. De Luca et al., "Derivative FTIR spectroscopy for cluster analysis and classification of morocco olive oils," *Food Chem.*, vol. 124, no. 3, pp. 1113–1118, 2011, doi: 10.1016/j.foodchem.2010.07.010.
- [20]R. De Bei et al., "Non-destructive measurement of grapevine water potential using near infrared spectroscopy," *Aust. J. Grape Wine Res.*, vol. 17, no. 1, pp. 62–71, 2011, doi: 10.1111/j.1755-0238.2010.00117.x.
- [21]X. Li et al., "Review of NIR spectroscopy methods for nondestructive quality analysis of oilseeds and edible oils," *Trends Food Sci. Technol.*, vol. 101, no. March, pp. 172–181, 2020, doi: 10.1016/j.tifs.2020.05.002.
- [22]S. Bureau and B. Gouble, "rapid and non-destructive analysis of apricot fruit quality using ft-near infrared spectroscopy," *Mater. Struct.*, no. 2014, pp. 32–35, 2015.
- [23]A. C. de Lima, L. Aceña, M. Mestres, and R. Boqué, "An Overview of the Application of Multivariate Analysis to the Evaluation of Beer Sensory Quality and Shelf-Life Stability," *Foods*, vol. 11, no. 14, pp. 1–15, 2022, doi: 10.3390/foods11142037.
- [24]A. Amari, N. El Bari, and B. Bouchikhi, "Electronic Nose for Anchovy Freshness Monitoring Based on Sensor Array and Pattern Recognition Methods: Principal Components Analysis, Linear Discriminant Analysis and Support Vector Machine," *Int. J. Comput.*, vol. 6, no. 3, pp. 61–67, 2014, doi: 10.47839/ijc.6.3.452.
- [25]M. W. Huang, C. W. Chen, W. C. Lin, S. W. Ke, and C. F. Tsai, "SVM and SVM ensembles in breast cancer prediction," *PLoS One*, vol. 12, no. 1, pp. 1–14, 2017, doi: 10.1371/journal.pone.0161501.
- [26]Z. Liu and J. Tan, "Qualitative and quantitative detection of Sudan I and II adulterated in chili powders by front-face synchronous fluorescence spectroscopy : Aggregation-induced emission in solid food," pp. 1–21, 2022.
- [27]N. Abu-khalaf and M. Hmidat, "Visible / Near Infrared (VIS / NIR) spectroscopy as an optical sensor for evaluating olive oil quality," *Comput. Electron. Agric.*, vol. 173, no. December 2019, p. 105445, 2020, doi: 10.1016/j.compag.2020.105445.
- [28]H. Cen, Y. He, and M. Huang, "Combination and comparison of multivariate analysis for the identification of orange varieties using visible and near infrared reflectance spectroscopy," *Eur. Food Res. Technol.*, vol. 225, no. 5–6, pp. 699–705, 2007, doi: 10.1007/s00217-006-0470-2.
- [29]Y. G. M. Kongbonga et al., "Characterization of Vegetable Oils by Fluorescence Spectroscopy," *Food Nutr. Sci.*, vol. 02, no. 07, pp. 692–699, 2011, doi: 10.4236/fns.2011.27095.
- [30]T. Wang, Y. Zhang, Y. Liu, Z. Zhang, and T. Yan, "Intelligent Evaluation of Stone Cell Content of Korla Fragrant Pears by Vis/NIR Reflection Spectroscopy," *Foods*, vol. 11, no. 16, 2022, doi: 10.3390/foods11162391.
- [31] m Zaregade, Mohammad Reza, abonajmi, M. Ghasemi-varnamkhasti, and f azarikia, "Estimation of the Best Classification Algorithm and Fraud Detection of Olive Oil by Olfaction Machine," vol. 11, no. 2, pp. 371–383, 2021, doi: 10.22067/jam.v11i2.84105.
- [32]M. Bahmaei, R. Nazeri, F. Kalantari, N. T. Branch, and S. B. Company, "Archive of SID Detection of adulteration in canola-virgin olive admixture using instrumental gas liquid chromatography and UV spectroscopy".

Feasibility assessment of employing near-infrared/visible spectroscopy for detecting olive pomace oil adulteration using LDA and SVM identification methods

Shirin Asadian¹, Ahmad Banakar², Bahareh Jamshidi³

1 Msc Student, Mechanical Engineering of Biosystems Department, Tarbiat Modares University, Tehran, Iran.

2 Associate Professor, Mechanical Engineering of Biosystem Department, Tarbiat Modares University, Tehran, Iran.

3 Associate Professor, Agricultural Engineering Research Institute, Agricultural Research Education and Extension Organization (AREEO), Karaj, Iran.

Extended Abstract

Introduction

Olive oil is a commercially valuable product, so it is often adulterated with other edible oils such as sunflower, hazelnut oil, pomace oil and other oils that have a fatty acid composition close to olive oil. Accordingly, quality control of this oil is one of the issues and it is very important in health and food safety. After extracting olive oil by physical techniques, a residue is created that contains more than 8% oil and can be extracted by a solvent (usually hexane). This oil is called pomace olive oil. Due to the characteristics of the solvent, the oil extracted with the solvent has more trace elements. Pomace olive oil is carcinogenic due to extraction with chemical solvents and as a result of being contaminated with benzo(a)pyrene, monochloropropane DLs and cyclic hydrocarbons and has an adverse effect on nerves and other body organs. Currently, gas chromatography method is used to detect adulteration in olive oil. Also, in the health system, to control the quality of food, including olive oil, compared to other frauds and the quantity of fatty acids, it is done by chemical tests and gas chromatography tests, which are expensive and time-consuming methods and require specialized quality control laboratories and specialized human resources. It should be noted that the presence of pomace olive oil in virgin olive oil and refined olive oil cannot be detected by gas chromatography, and detection of adulteration in these cases is not applicable by chromatography. Past studies show the high importance of spectroscopy as a fast, simple, low-cost method with acceptable accuracy in measuring the quality of agricultural products. For this reason, its use in different ways and for different products is expanding day by day. In this research, pomace olive oil adulteration was identified using a fast and non-destructive visible-near-infrared (Vis/NIR) spectroscopy method without measuring qualitative properties using pattern recognition methods such as LDA and SVM.

Materials and Methode

The extra-virgin, virgin and refined olive oils used in this research were prepared from the Etkha olive oil factory in Varamin, Tehran province. The sample of raw and non-edible pomace oil was also obtained from a pharmaceutical, cosmetic and health company. Extra virgin olive oil, virgin and refined olive oil were mixed with pomace oil with weight percentages of 11, 20, 33, 50 and 100%. These percentages were chosen according to the frauds in the market. A total of 160 samples were prepared with ten replicates for each treatment. The visible/near infrared spectrum was used using a two-way fiber, which is the central part for light transmission from the sample to the detector and the outer fiber for light transmission from the radiation source to the sample, and in the transition state and in the 200-850 nm region using Spectrometer model usb2000 made by ocean optic company and recorded with 1 nanometer resolution by OceanView software. The spectral data obtained from the spectrometer, in addition to the sample information, also includes unwanted background information and noises. Hence, in order to achieve stable, accurate and reliable calibration models, there is a need to pre-process spectral data before formulating classification models appropriate to the data. After recording and saving the obtained spectra for the analysis of the received spectra, data preprocessing was done. Part of the data from the 550 nm area already had high noise and could not be smoothed by the usual pre-processing methods, so they were removed. After the transmission spectra obtained from all the samples

were converted into absorption spectra in order to create a linear relationship with the molecular concentration of the sample, to prepare and pre-process the spectra some smoothing and noise reduction methods such as moving average (MA), Savitzky Golay (SG), Standard normal variate (SNV) and first derivative (D1) were used. The samples were divided into two groups of training (80 percent) for the development of calibration models, and the test group (20 percent) for the external validation of the developed calibration models in order to check the ability and accuracy of the models in classifying different percentages of fraud of unknown samples. Linear discriminant analysis (LDA) and support vector machine (SVM) were used in this research. Evaluation of the built model with full cross-validation method is used to select the model with the highest classification ability among the statistical models. In other words, after building the model in the re-evaluation phase, the reliability and stability of the model was also investigated. In this research, three types of Linear, Quadratic and Mahalanobis functions and four different weighting coefficients (1/SDev, Constant, Block weighting and Downweight) were used to create a classification model using LDA method.

Results and discussion

Our findings show that extra-virgin olive oil has higher absorption in almost all spectra than virgin and refined olive oils and pomace, and this is also consistent with previous researches. Three types of Linear, Quadratic and Mahalanobis functions were used to create the classification model by LDA method. The results obtained from LDA method showed that MA preprocessing based on Quadratic function provides the highest accuracy. In the next step, SG preprocessing based on the Quadratic function provided a better accurate model than other preprocessing methods; Therefore, in this method, the use of MA and SG preprocessing and the Quadratic function had a greater effect in building models with higher accuracy. It is worth mentioning that different weight coefficients had no effect on the accuracy of LDA method. Choosing the appropriate weighting coefficients depends on the type of data and usually by trial and error, the appropriate method of applying the weighting coefficient can be achieved. Two types of C-SVM and Nu-SVM were used to classify the samples in the SVM method. Four types of kernel functions including linear, quadratic, radial basis function and sigmoid were used. The linear function with MA preprocessing with training accuracy of 100 and validation accuracy of 97.5 had the highest validation accuracy in the classification of olive oil samples. It is worth noting that Linear function and C-CVM method had a great effect on more accurate models with different pre-processing. Therefore, both SVM and LDA modeling methods had a very acceptable performance in separating and categorizing the research treatments. The SVM method is preferred due to its greater accuracy in the calibration and validation section.

Conclusions

In this research, a visible/near infrared (Vis/NIR) spectroscopic system was investigated in order to classify olive oil samples with different adulteration percentages with pattern recognition methods including LDA and SVM. The obtained results showed that spectroscopy (Vis/NIR) is able to distinguish olive oil samples based on pomace olive oil adulteration rate. Although the LDA method was able to classify the olive oil samples with acceptable accuracy, the SVM method with a training accuracy of 96.69 and a validation accuracy of 95.86 was the best method for classifying the samples. Also, based on the results, the linear function, the best function for building the classification model was determined by SVM method.

Keywords: Fraud, Linear Discriminant Analysis, Pomace Oil, Support Vector Machine