

Winter 2024, 5 (4), 169-186

DOR:

Received: 29 Sep 2024

Accepted: 9 Nov 2024

مقاله پژوهشی

Fusion of Electronic Nose and Computer Vision for Detecting Lampante Oil Adulteration in Extra Virgin Olive Oil Using Machine Learning Algorithms

Seyyede Mahsa Mirhoseini Moghaddam¹, Mohammad Reza Yamaghani^{2*}, Adel Bakhshipour³

1. Ph.D. Student, Department of Computer Engineering, Lahijan Branch, Islamic Azad University, Lahijan, Iran.

2. Assistant Professor, Department of Computer Engineering, Lahijan Branch, Islamic Azad University, Lahijan, Iran. * Corresponding Author, o_yamaghani@iau.ac.ir

3. Assistant Professor, Department of Biosystems Engineering, Faculty of Agricultural Sciences, University of Guilan, Rasht, Iran.

Abstract

Introduction: Olive (*Olea europaea* L.), a key economic fruit species primarily cultivated in the Mediterranean region, is widely consumed as table olives and olive oil. Olive oil adulteration, a significant food fraud issue, typically involves mixing lower-quality oils with extra virgin olive oil (EVOO), compromising its quality and nutritional value. This study explores the use of electronic nose (e-nose) and computer vision (CV) systems to detect adulteration in EVOO.

Method: Four levels of adulteration (5%, 10%, 15%, and 20%) were simulated by blending lampante oil with pure EVOO. The aroma and appearance of the pure and adulterated samples were captured using an e-nose system equipped with 13 gas sensors and a CV apparatus. The collected signals and images were processed in MATLAB (R2021a, The MathWorks, USA) for feature extraction, resulting in 78 odor-related features (13 sensors $\times$ 6 features) and 9 color-related features. These were then used to train machine learning (ML) algorithms.

Results: Using e-nose data, the quadratic discriminant analysis (QDA) algorithm, combined with principal components analysis (PCA), we achieved 100% accuracy in distinguishing pure from adulterated EVOO. With CV data, the PCA-QDA model provided 98.50% accuracy. Furthermore, the partial least squares (PLS) algorithm predicted the percentage of lampante oil adulteration with determination coefficients (R^2) of 0.8565 in the training phase and 0.7858 in the evaluation phase. These values dropped to 0.6983 and 0.4936, respectively, when only color features were used. However, by fusing aroma and color data, the prediction accuracy significantly improved, with R^2 values of 0.9266 for training and 0.9184 for evaluation.

Discussion: The findings demonstrate that combining e-nose and CV data with appropriate ML algorithms allows highly accurate detection of EVOO adulteration. Robustness tests, performed by intentionally altering the e-nose and CV data, confirmed the models' reliability. Given the e-nose and CV systems' capabilities and the achieved performance, this approach offers a fast, accurate, cost-effective, and non-destructive method for detecting adulteration in EVOO.

Keywords: Adulteration, chemometrics, electronic nose, image processing, data fusion.

ترکیب بینی الکترونیک و بینایی کامپیوتر برای تشخیص تقلب روغن لامپانت در روغن زیتون فرابکر با استفاده از الگوریتم‌های یادگیری ماشین

دوره پنجم، زمستان ۱۴۰۳
شماره چهارم، صص: ۱۶۹-۱۸۶

تاریخ دریافت: ۱۴۰۳/۰۷/۰۸
تاریخ پذیرش: ۱۴۰۳/۰۸/۱۹

سیده مهسا میرحسینی مقدم^۱، محمدرضا یمقانی^{۲*}، عادل بخشی پور^۳

۱. دانشجوی دکتری، گروه مهندسی کامپیوتر، واحد لاهیجان، دانشگاه آزاد اسلامی، لاهیجان، ایران.

mahsa.moghaddam1368@yahoo.com

۲. استادیار، گروه مهندسی کامپیوتر، واحد لاهیجان، دانشگاه آزاد اسلامی، لاهیجان، ایران. (نویسنده مسئول). o_yamaghani@iau.ac.ir

۳. استادیار، گروه مهندسی بیوسیستم، دانشکده علوم کشاورزی دانشگاه گیلان، رشت، ایران. abakhshipour@guilan.ac.ir

چکیده: روغن زیتون یک محصول با ارزش غذایی محسوب می‌شود که تقلب در آن با روغن‌های کم کیفیت رو به افزایش است. روش‌های مختلفی برای ارزیابی کیفیت روغن زیتون خوراکی وجود دارد، اما یکی از روش‌های غیرمخرب، دقیق و سریع، استفاده از سامانه‌های ماشین بویایی و بینایی است. در این پژوهش چهار سطح تقلب (۵، ۱۰، ۱۵ و ۲۰ درصد ناخالصی) با اختلاط روغن لامپانت در روغن زیتون فرابکر خالص ایجاد شد و سپس نمونه‌های خالص و تقلبی توسط یک سامانه بینی الکترونیکی با ۱۳ حسگر گازی و یک سامانه بینایی رایانه‌ای مورد تجزیه و تحلیل قرار گرفتند. نتایج حاصل از تجزیه و تحلیل مؤلفه‌های اصلی (PCA) نشان داد که ویژگی‌های استخراج‌شده از سامانه‌های بینی الکترونیکی و بینایی رایانه‌ای توانایی بالایی در تفکیک گروه‌های مختلف روغن‌های زیتون خالص و ناخالص دارند. نتایج طبقه‌بندی نشان داد که الگوریتم تحلیل تشخیصی مرتبه دوم (QDA) که با استفاده از مؤلفه‌های اصلی (PCs) تولید شده از ویژگی‌های بینی الکترونیک توسعه یافته است، قادر به تشخیص ۱۰۰ درصدی روغن‌های خالص و ناخالص است. این صحت در هنگام استفاده از مؤلفه‌های اصلی استخراج‌شده از ویژگی‌های تصویر برای تشخیص ناخالصی‌ها با روغن لامپانت برابر با ۱۰۰ درصد بود. نتایج نشان داد که الگوریتم حداقل مربعات جزئی (PLS) قادر به پیش‌بینی میزان ناخالصی روغن لامپانت در روغن زیتون فرابکر خالص با ضریب تبیین (R^2) آموزش و اعتبارسنجی به ترتیب برابر با ۰/۸۵۶۵ و ۰/۷۸۵۸ بود. این مقادیر در هنگام استفاده از ویژگی‌های رنگی به ترتیب برابر با ۰/۶۹۸۳ و ۰/۴۹۳۶ بودند که نشان داد استفاده از داده‌های حاصل از بوی روغن روش مناسب‌تری نسبت به رنگ روغن برای تخمین نرخ ناخالصی لامپانت در روغن زیتون فرابکر است. از سوی دیگر، ادغام داده‌های استخراج‌شده از عطر و رنگ برای تعیین میزان ناخالصی، دقت پیش‌بینی را به‌طور قابل توجهی در مقایسه با استفاده از داده‌های رنگ و رایحه به‌صورت مجزا افزایش داد. مقدار R^2 آموزش و اعتبارسنجی مدل PLS با داده‌های ترکیبی برای پیش‌بینی اختلاط روغن لامپانت برابر با ۰/۹۲۶۶ و ۰/۹۱۸۴ به دست آمد. نتایج این تحقیق نشان داد که با استفاده از داده‌های بینی الکترونیک و بینایی الکترونیک در کنار الگوریتم یادگیری ماشین مناسب، می‌توان وجود اختلاط در روغن زیتون فرابکر خالص را با دقت بالایی تعیین کرد. براساس نتایج این پژوهش می‌توان از سامانه ترکیبی بینی و بینایی الکترونیک برای توسعه یک سامانه غیر مخرب و دقیق برای پایش تقلب در روغن زیتون استفاده کرد.

واژه‌های کلیدی: تقلب، کمومتریک، بینی الکترونیک، پردازش تصویر، همجوشی داده‌ها.

۱. مقدمه

زیتون (*Olea europaea* L.) یک گونه میوه‌ای است که به‌طور گسترده و عمدتاً در مناطق مدیترانه‌ای کشت می‌شود و از نظر اقتصادی مهم است که هم به عنوان روغن زیتون و هم به عنوان مصرف روی سفره استفاده می‌شود [۱]. بر اساس آخرین آمار سازمان کشاورزی و مواد غذایی ملل متحد (FAO)^۱، تولید جهانی انواع در حدود ۲۱/۴۵ میلیون تن در سال ۲۰۲۲ گزارش شده است که کشور اسپانیا با بیش از ۳/۹۴ میلیون تن اولین تولیدکننده این محصول است [۲]. روغن زیتون بکر که عصاره میوه زیتون است که از محبوب‌ترین و شناخته‌شده‌ترین روغن‌های گیاهی است. این ماده نه تنها طعم خوبی دارد، بلکه سرشار از مواد مغذی و خواص دارویی است. این روغن دارای محتوای بالای اسید اولئیک، ترکیبات آنتی‌اکسیدانی مانند ترکیبات فنلی و اسیدهای چرب غیراشباع است که در پیشگیری از بیماری‌های مختلف نقش دارند [۳]. روغن زیتون بکر به روغنی اطلاق می‌شود که صرفاً با فشار و سایر فرآیندهای مکانیکی استخراج شده و هیچ گونه فرایند دیگری روی آن انجام نشده است [۴]. بر اساس آمار FAO، میزان تولید روغن زیتون در سال ۲۰۲۲ در حدود ۳/۳۵ میلیون تن بوده است [۲]. اتحادیه اروپا بزرگترین تولیدکننده روغن زیتون در سراسر جهان است (با حدود ۷۳ درصد از تولید جهانی) که از این مقدار حدود ۶۰ درصد آن توسط اسپانیا (بزرگترین تولیدکننده) و پس از آن ایتالیا و یونان تولید می‌شود [۲].

تقلب در مواد غذایی مخصوصاً در روغن زیتون با افزودن موادی با ارزش غذایی و قیمت پایین‌تر به روغن اصل ایجاد می‌شود و این افزودن باعث تغییر در بعضی ویژگی‌های روغن می‌شود. روش‌های معمول تعیین تقلب روغن زیتون، مانند کروماتوگرافی اسپکترو سکوپ گران و وقت‌گیر بوده و نیاز به ابزارهای پیچیده آزمایشگاه‌های مجهز دارند. روش ارزیابی حسی نیز نیاز به شخص متخصص دارد و تحت تأثیر شرایط فیزیکی، سلامتی، و صداقت شخص ارزیاب است. بنابراین دسترسی به روش‌های جدید و دقیق برای تشخیص تقلب روغن خوراکی لازم و ضروری است. در چند دهه اخیر این توجه به‌سوی استفاده از روش‌های بینایی، بویایی، چشایی رایانه‌ای معطوف شده است.

بینی الکترونیک از جمله روش‌های غیرمخرب آسان، ارزان قیمت و نسبتاً دقیق است که در سال‌های اخیر در زمینه‌های مختلف از جمله صنایع غذایی با موفقیت استفاده شده است. بینی الکترونیک که با عبارات دیگری مانند ماشین بویایی، سامانه تشخیص بو یا بوسنجی الکترونیکی نیز نامیده می‌شود، درواقع آرایه‌ای از حسگرهای شیمیایی نیمه‌انتخاب‌گر، با حساسیت به یک یا چند گاز خاص در ترکیب با یک سامانه تشخیص الگوی (PARC)^۲ مناسب جهت شناسایی بوهای ساده یا پیچیده است [۵]. در این سامانه اجزای آلی فرار (VOCs)^۳ مواد معطر به صورت یک رد اثر یا الگو در فضای بالای نمونه شناسایی

می‌شوند. در پژوهشی از آرایه حسگرهای گاز در یک سامانه بینی الکترونیک به منظور تشخیص تقلب حاصل از اختلاط موادی مانند کدو تنبل، سیب زمینی و نشاسته در درصد‌های مختلف در رب گوجه‌فرنگی استفاده شد. نتایج ایشان نشان داد که استفاده از حسگرهای TGS26 و MQ3 به همراه طبقه‌بندهای تحلیل مؤلفه‌های اصلی (PCA)^۴ و ماشین بردار پشتیبان (SVM)^۵ در تشخیص تقلب‌ها در رب گوجه‌فرنگی بسیار کارآمد است [۶]. در پژوهش دیگری، از یک سامانه بینی الکترونیک بر پایه شش حسگر نیمه‌هادی اکسید فلزی به منظور نظارت بر اکسیداسیون روغن زیتون فرابرک در طول نگهداری استفاده شد. مدل تحلیل تشخیصی خطی (LDA)^۶ توانست نمونه‌ها را بر اساس سطح اکسیداسیون به خوبی تفکیک کند. همچنین همبستگی خوبی با ارزیابی حسی حاصل شد [۷]. در پژوهش‌های دیگری توانایی سامانه بینی الکترونیک در تشخیص اصالت روغن زیتون بکر با افزودن روغن‌های گیاهی کم‌ارزش‌تر مانند روغن‌های آفتاب‌گردان، تفال‌ه زیتون، کلزا و فندق مورد بررسی قرار گرفت [۸]. پژوهش‌های دیگری از بینی الکترونیک به منظور تشخیص کیفیت روغن زیتون [۹]، تقلب در شیر [۱۰] و تقلب مارگارین در کره حیوانی [۱۱] استفاده شده است.

بینایی رایانه یا ماشین یکی از شاخه‌های بسیار مهم از هوش مصنوعی است که کاربردهای فراوانی در صنایع غذایی خصوصاً در آزمایشگاه‌های کنترل کیفیت دارد. دستگاه‌های بینایی ماشین از جمله دستگاه‌هایی هستند که در ارزیابی کیفی محصولات کشاورزی و دارویی به کار می‌روند. به دلیل سرعت و دقت بالای دستگاه‌های بینایی ماشین، صنعت کشاورزی و صنایع غذایی در سطح جهان تمایل فراوان به استفاده از آن به جای دستگاه‌های نظارت دستی پیدا کرده‌اند. در مطالعه‌ای امکان تشخیص تقلب مربوط به اختلاط روغن سویا در روغن زیتون فرابرک با استفاده از پردازش اطلاعات در فضای رنگی RGB بررسی شد. صحت طبقه‌بندی ۱۰۰ درصد در غلظت‌های تقلب ۰/۵ تا ۱۰ درصد حجمی حاصل شد [۱۲]. تحقیقات دیگری که مربوط به پردازش رنگی تصاویر می‌شوند، شامل شناسایی منشأ روغن زیتون [۱۳]، پیش‌بینی محتوای روغن در زیتون به منظور بهینه‌سازی زمان برداشت [۱۴]، و تشخیص شاخص‌های کلروفیل و کاروتنوئیدها در ارقام مختلف روغن زیتون بکر [۱۵] هستند. در پژوهشی دیگر یک سامانه ماشین بینایی به منظور تشخیص ناخالصی‌های روغن زیتون در سه سطح ارائه شد. به این منظور پارامترهای مربوط به فضاهای رنگی RGB، HSV و L^*a^*b از روغن زیتون استخراج شده و با استفاده از روش‌های شبکه عصبی مصنوعی و ماشین بردار پشتیبان نتایج طبقه‌بندی مقایسه شدند [۱۶].

همچونشی داده‌ها یک راهبرد جدید در یادگیر ماشین برای ترکیب داده‌های به دست آمده از فناوری‌های مختلف برای دستیابی به یک طبقه مدل سازی کامل‌تر و دقیق‌تر است [۱۷]. بسته به اینکه داده‌های

نرخ اختلاط با روغن لامپانت نیز تعداد ۱۹ نفر از ۲۸ نفر، نرخ اختلاط ۱۵ تا ۲۰ درصد را تأیید کردند. بر این اساس، آزمایش‌ها در پنج کلاس مختلف شامل اختلاط صفر درصد (روغن زیتون فرابرکر خالص)، و ۵ درصد، ۱۰ درصد، ۱۵ درصد و ۲۰ درصد اختلاط با روغن لامپانت انجام شد. تعداد ۱۰ نمونه با وزن خالص ۱۰۰ گرم برای هر کلاس بررسی شدند. روغن زیتون فرابرکر و لامپانت مورد نیاز این آزمایش از ایستگاه تحقیقات زیتون رودبار تهیه شدند و تا زمان آزمایش در دمای یخچال (۴ درجه سلیسیوس) نگهداری شد.

جدول ۱: خلاصه برخی از پژوهش‌های مرتبط اخیر

محبول	سامانه استفاده شده	مدل و نتایج	منبع
روغن زیتون	بینایی ماشین و ویژگی‌های رنگی	تشخیص تقلب در روغن زیتون فرابرکر با روغن سویا با حداکثر خطای ۵ درصد توسط رگرسیون خطی	[۲۲]
روغن زیتون	بینایی ماشین و ویژگی‌های رنگی	طبقه‌بندی روغن‌های زیتون فرابرکر دارای اختلاط با روغن سویا با نرخ طبقه‌بندی حداکثر ۹۴ درصد توسط مدل PLS-DA ^v	[۱۲]
روغن زیتون	بینی الکترونیک تجاری	مدل PCA با صحت ۸۱ درصد و مدل LDA با صحت ۹۲ درصدی طبقه‌بندی	[۲۳]
روغن دانه گیاه پائونیا	بینی الکترونیک مدل تجاری PEN3 با ۱۰ حسگر گازی	طبقه‌بند LDA به خوبی توانستند روغن آفتابگردان و روغن سویا را از روغن پائونیا تشخیص دهند	[۲۴]
روغن سویا	بینی الکترونیک آزمایشگاهی با ۸ حسگر	مدل PLS با ضریب همبستگی ۰/۸۴۳ توانست نرخ تقلب روغن سرخ شده در روغن سویا را تشخیص دهد	[۲۵]
روغن آرگان و روغن آفتابگردان	بینی الکترونیک آزمایشگاهی با ۵ حسگر	مدل PCA توانست تقلب را با صحت ۹۸/۸۱ درصد در روغن آرگان خوراکی و با صحت ۹۸/۰۹ درصد روغن آرگان آرایشی تشخیص دهد	[۲۶]
روغن نارگیل بکر	بینی الکترونیک تجاری مدل zNose TM Model 7100	مدل PLS توانست با ضریب همبستگی ۰/۹۱ تقلب روغن پالم در روغن نارگیل بکر را تشخیص دهد	[۲۷]

۲.۲. کلیات روش تحقیق

شکل ۱ راهبرد کلی اجرا شده در این پژوهش پس از تهیه و آماده‌سازی نمونه‌ها را نشان می‌دهد. به‌طور کلی در این پژوهش از ترکیب داده‌های بینی الکترونیک و بینایی ماشین برای تشخیص سطوح مختلف تقلب‌های مختلف در روغن زیتون فرابرکر استفاده شد. در بخش بینی الکترونیک سیگنال‌های حاصل از حسگرها پیش‌پردازش شدند و سپس

خام از منابع مختلف به عنوان ورودی مدل‌ها استفاده شوند، داده‌های بهینه از منابع مختلف انتخاب شوند و با هم ترکیب شوند، یا نتایج حاصل از مدل‌های توسعه داده‌شده از منابع مختلف برای به دست آوردن نتیجه تصمیم نهایی استفاده شوند، راهبرد همجوشی داده‌ها به سه سطح پایین، متوسط و بالا تقسیم بندی می‌شوند [۱۸]. در سال‌های اخیر از همجوشی داده‌ها برای بهبود عملکرد مدل‌سازی ویژگی‌های کیفی محصولات غذایی و کشاورزی استفاده شده‌است. در پژوهشی برای تعیین تقلب گلرنگ رنگ‌شده و کنج زعفران از روش ترکیبی بینایی ماشین و ماشین بویایی استفاده شد. تفسیر و بررسی مقادیر حاصل از برآورد سامانه ترکیبی در مقایسه با مقادیر اندازه‌گیری‌شده به روش شیمی‌سنجی نشان داد که این سامانه توانایی تشخیص تقلب در زعفران با صحت ۸۹ درصد توسط بینایی ماشین و صحت ۱۰۰ درصد توسط ماشین بویایی را دارد [۱۹]. در پژوهش دیگری از روش‌های ترکیبی پردازش تصویر و بینی الکترونیکی برای تعیین کیفیت گوشت مرغ استفاده شد. نتایج حاصل از بررسی‌های سامانه ترکیبی نشان داد که اثر اسانس زردچوبه بر روی نمونه‌ها به میزان درصد ۳۰ موجب حفظ کیفیت مناسب و افزایش ماندگاری آن‌ها در طول مدت نگهداری در شرایط سرما می‌شود [۲۰]. مطالعه‌ای بر روی تفکیک گیاهان دارویی به صورت برخط با استفاده از سامانه ترکیبی بینایی ماشین و ماشین بویایی انجام شد. نتایج حاصل از بررسی چندین گیاه مختلف دارویی مثل گلگاوزبان و یا رازیانه نشان داد که گیاهان دارویی اغلب با فیلترهای تصویری و بویایی قابل تفکیک هستند و این کار می‌تواند به صورت برخط با صحت میانگین ۹۳/۵ درصد انجام پذیرد [۲۱].

در جدول ۱ سابقه برخی از پژوهش‌ها در زمینه تشخیص تقلب در روغن زیتون ارائه شده است. با توجه به مرور منابع تاکنون پژوهشی در زمینه ترکیب داده‌های رایحه و ظاهر که پارامترهای بسیار مهمی در ارزیابی محصولات هستند، برای تشخیص تقلب در روغن زیتون انجام نگرفته است. بنابراین هدف این پژوهش ترکیب داده‌های بینایی ماشین و بینی الکترونیک به منظور تشخیص اختلاط روغن لامپانت در روغن زیتون فرابرکر در نظر گرفته شد.

۲. مواد و روش‌ها

۱.۲. آماده سازی نمونه

به منظور تعیین نوع و میزان اختلاط در روغن زیتون، یک پرسشنامه ساده شامل نوع و میزان اختلاط در روغن زیتون تهیه شد. با بازدید میدانی از تولیدکنندگان و فروشندگان روغن زیتون، تعداد ۵۰ مورد پرسشنامه در بین پاسخ‌دهندگان توزیع شد و خواسته شد که به صورت ناشناس پاسخ دهند. تعداد ۳۹ مورد از پاسخ‌دهندگان تأیید کردند که تقلب در روغن زیتون اتفاق می‌افتد. براساس نتایج پرسشنامه، یکی از مرسوم‌ترین تقلب‌های منطقه، تقلب با روغن لامپانت به دست آمد که در ۲۸ مورد از پرسشنامه‌های پاسخ داده شده تأیید شده بود. در مورد

استخراج ویژگی انجام شد. در بخش بینایی ماشین نیز پس از پیش پردازش و پردازش تصویر، ویژگی های مورد نظر از تصاویر استخراج شدند. سپس الگوریتم های مختلف طبقه بندی برای تشخیص روغن های خالص و دارای اختلاط توسعه یافتند. الگوریتم های پیش بینی نیز به منظور تخمین نرخ اختلاط در روغن زیتون فرابکر مورد ارزیابی قرار گرفتند. این مراحل به تفصیل در ادامه توضیح داده خواهند شد. نکته مهم اینکه در تمامی مدت زمان پژوهش، نمونه ها در دمای اتاق (۲۰ درجه سلسیوس) نگهداری شدند. لازم به ذکر است آزمایش های بین الکترونیک در اتاقی در آزمایشگاه عاری از هرگونه بو انجام شد. تهیه تصاویر نیز در شرایط نورپردازی کنترل شده، انجام شد.

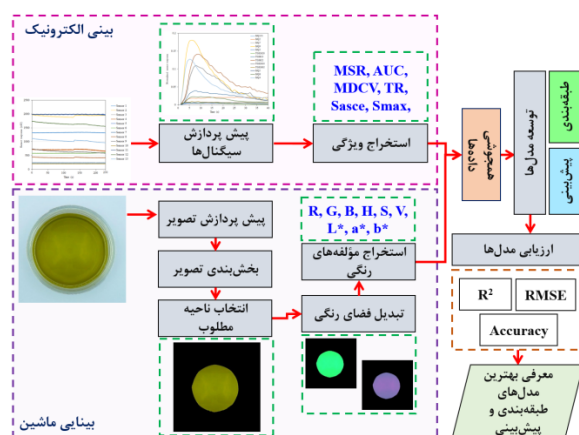
جدول ۲: مشخصات و کاربرد حسگرهای به کار رفته در آرایه

حسگری

نام حسگر	محدوده تشخیص (ppm)	کاربرد
MQ135	10-300 NH3 10-1000	CO ₂ , دود، بزن، الکل، NH ₃ , NOx
MQ2	300-5000-LPG and propane	LPG, i-butane, propane, methane, alcohol, Hydrogen, smoke
MQ7	100-300	CO
MQ9	20-2000 CO	CO and CH ₄ , LPG
TGS813	500-10000	Domestic gas leak detectors and alarms
TGS822	50-5000	Breath alcohol detectors, Gas leak detectors
TGS2602	1-30	Air cleaners, Ventilation control
MQ5	200-10000	LPG, natural gas, town gas
TGS2620	50-5000	Alcohol, toluene, xylene, other volatile organic vapors
MQ-3	0.05-10	Alcohol
MQ8	100-1000	H ₂
TGS2610	500-1000	Butane, liquid petroleum gas
MQ4	300-10000	Natural gas/ Methane

۴.۲. داده برداری بین الکترونیک

در این بخش ابتدا، یک مرحله پیش پردازش بر داده ها اعمال شد. پس از آن طی یک مرحله استخراج ویژگی، ابعاد داده ها کاهش داده شد. با دسته بندی داده ها در فضای ویژگی می توان نوع آنالیت را تشخیص داد. با توجه به اینکه تعداد داده های استخراجی در هر ثانیه برابر با ۱۰ داده بود، به منظور حذف نویزهای احتمالی در داده های ضبط شده در هر ثانیه، در ابتدا میانگین داده های ضبط شده در هر ثانیه محاسبه شد. سپس با حذف داده احتمالی دارای تفاوت بالا (در این پژوهش ۱۰ میلی و لت در نظر گرفته شد) با میانگین داده های آن ثانیه، میانگین گیری مجدد از سایر داده های مربوط به هر زمان انجام شد و به عنوان پاسخ حسگر در آن ثانیه در نظر گرفته شد. از سوی دیگر، به منظور حذف نویز در سیگنال حاصل از میانگین گیری، در این مطالعه از روش هموارسازی پنجره متحرک (MWS)^۸ برای پیش پردازش سیگنال های دریافت شده از حسگرها استفاده شد. روش MWS پنجره ای را با اندازه از پیش تعریف شده، انتخاب می کند که روی سیگنال حرکت می کند و مقدار میانگین سیگنال در داخل پنجره متحرک را محاسبه و با مقدار اولیه مربوط به مرکز پنجره جایگزین می کند [۲۸]. مرحله دیگر پیش پردازش شامل تصحیح خط مبنا بود. به این منظور از روش کسری که معمولاً برای حسگرهای MOS به کار می رود، استفاده شد [۲۹]. در این روش، مقدار مبنای کمینه از پاسخ حسگرها از کم شده و نتیجه بر تفاوت مقادیر مبنای بیشینه و کمینه (دامنه داده ها) تقسیم می شود. مقدار مبنای کمینه با میانگین گرفتن از پاسخ



شکل ۱: راهبرد کلی تعیین تقلب در روغن زیتون با استفاده از داده های بین الکترونیک و بینایی ماشین در این پژوهش

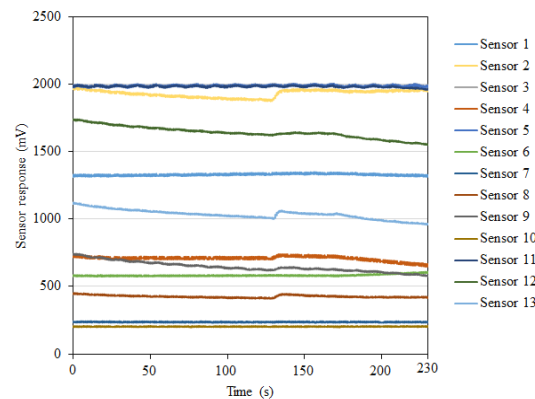
۳.۲. آرایه حسگری بین الکترونیک

یک سامانه آرایه حسگری (ماشین بویایی) مبتنی بر ۱۳ حسگر گازی جهت تشخیص تقلب روغن خوراکی در آزمایشگاه گروه مهندسی بیوسیستم دانشگاه گیلان طراحی و ساخته شد. سامانه شامل کپسول هوا، محفظه حسگرها، محفظه نمونه گیری، سامانه تحویل و گردآوری داده و شیرهای قطع و وصل جریان هوا بود. خصوصیات حسگرهای استفاده شده در آرایه حسگری در جدول ۲ آورده شده است. مراحل کاری سامانه ماشین بویایی شامل تصحیح خط مبنا، اندازه گیری و پاکسازی محفظه حسگرها است. مرحله تصحیح خط مبنا، به منظور رساندن پاسخ آرایه حسگرها به حالت پایدار انجام شد. در این مرحله گاز حامل با پمپ به مدت ۶۰ ثانیه به محفظه حسگرها تزریق شد. در مرحله اندازه گیری، گاز حامل با ورود به محفظه نمونه، گاز فضای هد روغن را به محفظه حسگرها منتقل می کرد. مدت زمان ۴۰ ثانیه برای رسیدن پاسخ حسگرها به حالت ماکزیمم در نظر گرفته شد. مرحله پاکسازی برای رساندن پاسخ حسگرها به خط مبنا و آماده سازی سامانه برای آزمایش های بعدی، انجام می شود. گاز حامل به مدت ۶۰ ثانیه به محفظه حسگرها تزری شد.

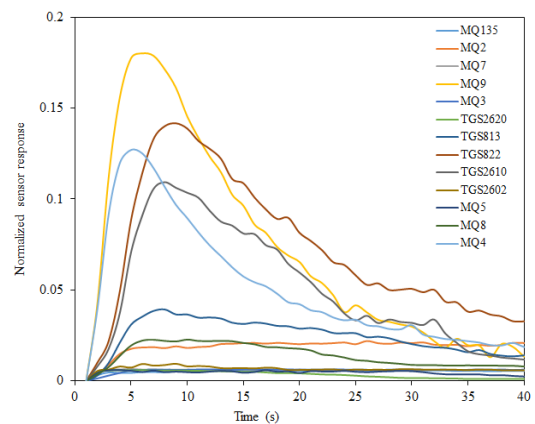
حسگرها در طی ۱۰ ثانیه منتهی به شروع مرحله تزریق گاز فضای فوقانی نمونه‌های روغن زیتون محاسبه‌شد. از رابطه ۱ به منظور نرمال‌سازی داده‌ها استفاده‌شد [۳۰].

$$y_s(t) = \frac{x_s(t) - x_s(0)}{x_s(0)} \quad (1)$$

که در آن، $y_s(t)$ داده نرمال شده از حسگر s در زمان t و $x_s(0)$ و $x_s(t)$ به ترتیب نماینده پاسخ حسگر s در لحظه اول و هر زمان t است. یک نمونه از سینگنال‌های نرمال شده از حسگرها در شکل ۲ آمده‌است.



(الف)



(ب)

شکل ۲: یک نمونه از سینگنال ضبط شده در سامانه بینی الکترونیک (الف) و سینگنال پیش پردازش شده حسگرها (ب)

پس از نرمال سازی، شش ویژگی از پاسخ‌های از پیش پردازش شده استخراج شدند که عبارتند از [۳۱-۳۴]:

- (۱) حداکثر پاسخ حسگر (MSR)^۹
- (۲) مقدار مساحت زیر نمودار (AUC)^{۱۰} که مساحت ناحیه بین منحنی پاسخ و خط پایه است، که مجموع پاسخ کلی یک حسگر به یک نمونه کیوی در یک بازه زمانی را نشان می‌دهد.
- (۳) زمان اشباع (Tr)^{۱۱} که به عنوان زمان رسیدن به حداکثر پاسخ حسگر تعریف می‌شود.

(۴) شیب متوسط بخش صعودی (S_{asce})^{۱۲}

(۵) مقدار حداکثر شیب لحظه‌ای (S_{max})^{۱۳}

(۶) مقدار میانگین ضریب تفاضلی (MDCV)^{۱۴} که نشان دهنده

سرعت متوسط پاسخ حسگرها در بازه انتخاب شده است.

در مجموع ۷۸ ویژگی (۱۳ حسگر × ۶ ویژگی = ۷۸) از رایحه نمونه‌های روغن زیتون خالص و غیرخالص استخراج شدند و برای توسعه الگوریتم‌های یادگیری ماشین استفاده شدند.

۵.۲. دریافت و پردازش تصویر

تصاویر اصلی نمونه‌های روغن از نمای بالا با استفاده از یک سامانه بینایی ماشین متشکل از یک دوربین دیجیتال (Samsung CCD (ES95 digital color camera, Korea, 10 MP) و یک محفظه نورپردازی چوبی مجهز به نورهای نوری LED ۱۲ ولت به عنوان سامانه روشنایی استفاده شد. دوربین در فاصله عمودی ۰/۲ متری بالای نمونه‌های روغن که بر روی مقوای سفید مات قرارداشتند، قرارگرفت. اندازه تصاویر گرفته شده ۲۷۳۶×۳۶۴۸ پیکسل بود. تصاویر رنگی به رایانه منتقل شده و در جعبه ابزار پردازش تصویر نرم افزار برنامه نویسی متلب (Version 2021a, the MathWorks, USA) برای پردازش و عملیات استخراج ویژگی فراخوانده شدند. ابتدا، یک بلوک ۱۲۰۰×۱۲۰۰ پیکسل در اطراف مرکز تصویر برش داده شد که حاوی نمونه روغن بود (شکل ۳الف) تا برخی از پیکسل‌های غیرلازم در نواحی لبه تصویر حذف شوند. قرار دادن نمونه‌های روغن در داخل ظروفی با مقطع دایره‌ای و برش بلوک تصویر به شکل مربعی به این دلیل انجام شد که دوران دوربین تاثیری در نتایج نداشته باشد.

از فیلتر میانه با بلوک ۳×۳ برای حذف نویزهای احتمالی در تصاویر استفاده شد که یک روش معمول برای کاهش نویز در تصویر است. فیلتر میانه برای حذف نویزهای ضربه‌ای مانند نویز نمک و فلفلی^{۱۵} که ممکن است به دلایل سخت افزاری در تصویر ایجاد شوند، استفاده می‌شود. فیلتر میانه تا حدودی مانند فیلتر میانگین عمل می‌کند، اما اغلب عملکرد بهتری نسبت به فیلتر میانگین در حفظ جزئیات مفید در تصویر دارد. در این روش پنجره‌ای بر روی تصویر حرکت می‌کند و شامل تعداد فرد از پیکسل (۹ پیکسل در این پژوهش) است. پیکسل مرکزی پنجره که همان پیکسل مبنای فیلتر است با مقدار میانه آماری تمامی پیکسل‌های درون پنجره در تصویر جایگزین می‌شود [۳۵]. از آنجاکه فیلتر میانه بر روی تصویر دوبعدی عمل می‌کند، در ابتدا مؤلفه‌های رنگی قرمز، سبز و آبی در تصویر RGB جدا شدند. سپس اعمال فیلتر میانه با استفاده از تابع $medfilt2(I, [3 \ 3])$ در نرم افزار MATLAB انجام شد (که در آن، I همان مؤلفه تصویر است). پس از اعمال فیلتر میانه بر روی مؤلفه‌های تصویر، به منظور بخش بندی تصویر و تشخیص ناحیه مربوط به روغن از پس زمینه تصویر، مؤلفه قرمز مزاد (EXR)^{۱۶} (شکل ۳ب) با استفاده از رابطه ۲ محاسبه شد.

(۲)

$$EXR=2 \times R-G-B$$

که در آن R، G، و B به ترتیب مؤلفه‌های رنگی قرمز، سبز و آبی از بلوک تصویر RGB برش خورده بعد از اعمال فیلتر میانه هستند.

با آستانه گذاری بهینه روی تصویر EXR، تصویر باینری اولیه روغن به دست آمد. پس از حذف نواحی مرزی تصویر قطعه‌بندی شده، و پر کردن حفره‌های احتمالی داخل ناحیه مورد نظر، تصویر دودویی نهایی روغن به دست آمد (شکل ۳ج). در نهایت با اعمال عملیات AND منطقی بین تصویر باینری و تصویر رنگی، تصویر RGB روغن با مقدار صفر در مناطق پس زمینه به دست آمد (شکل ۳د). سپس این تصویر از فضای رنگی RGB به فضاهای رنگی HSV (شکل ۳ه) و L^*a^*b (شکل ۳و) تبدیل شده و مقادیر متوسط مؤلفه‌های رنگی قرمز، سبز، آبی، فام^{۱۷}، اشباع^{۱۸}، مقدار^{۱۹}، روشنایی^{۲۰} و نیز اجزای رنگی a^* و b^* استخراج شدند [۳۶].

۶.۲. تحلیل داده‌ها

۶.۲.۱. تحلیل مؤلفه‌های اصلی (PCA)

با اعمال روش PCA، متغیرهای اولیه به مؤلفه‌های جدید و مستقل از یکدیگر (با ضریب همبستگی صفر برای هر دو مؤلفه اصلی) تبدیل می‌شوند. مؤلفه‌های جدید ایجاد شده ترکیبی خطی از متغیرهای اولیه هستند [۳۷]. با استفاده از این روش، ترکیباتی از P متغیر اولیه، $X_1, X_2, \dots, X_p$ برای ایجاد P مؤلفه مستقل (معادل با تعداد متغیرهای اولیه مورد استفاده) $Z_1, Z_2, \dots, Z_p$ ایجاد می‌شود [۳۸]. با توجه به تعداد بالای متغیرهای ورودی، در این پژوهش از روش PCA به منظور کاهش متغیرهای ورودی به مدل‌ها با حفظ حداکثری واریانس اطلاعات ورودی استفاده شد. ثابت شده است که PCA علاوه بر کاهش ابعاد ویژگی‌های ورودی به مدل، با حذف اطلاعات غیرمرتبط یا اضافی باعث افزایش عملکرد مدل‌های یادگیری ماشین می‌شود [۳۹].

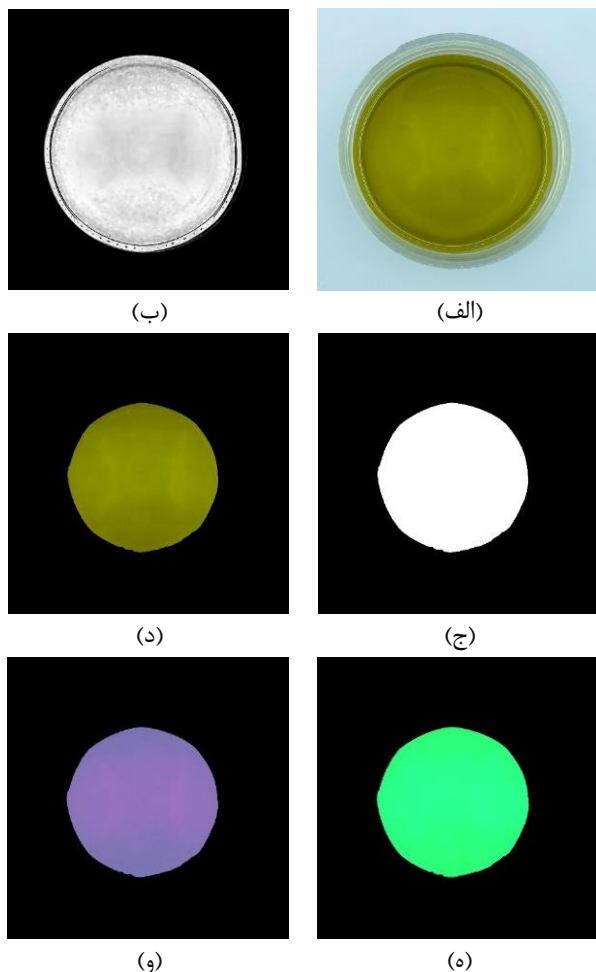
۶.۲.۲. حداقل مربعات جزئی (PLS)

PLS توانایی آنالیز داده‌های که دارای مشکلات هم‌خطی، خطای زیاد و یا تعداد زیادی از متغیرهای x و y هستند، را داراست. PLS ابعاد داده‌های متغیر مستقل را به استخراج متغیرهای پنهان کاهش می‌دهد [۴۰] و درواقع PLS یک ماکزیمم کوواریانس بین متغیرهای پنهان x و متغیرهای وابسته y به وجود می‌آورد [۴۱]. در این پژوهش از PLS به منظور تخمین میزان ناخالصی در روغن زیتون با استفاده از داده‌های بینی الکترونیک و بینایی ماشین استفاده شد.

۶.۲.۳. رگرسیون خطی چند متغیره (MLR)^{۲۱}

مدل MLR یک الگوریتم یادگیری نظارت شده است که به عنوان بسط رگرسیون خطی ساده برای زمانی است که دو یا چند متغیر مستقل وجود دارد، تعریف می‌شود [۴۲]. در این پژوهش، علاوه بر روش PLS، از

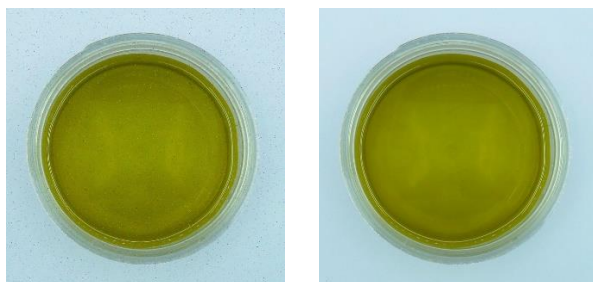
الگوریتم MLR برای یافتن روابط احتمالی بین اطلاعات استخراج شده از سامانه‌های بینایی و بویایی الکترونیک و درصد تقلب استفاده شد.



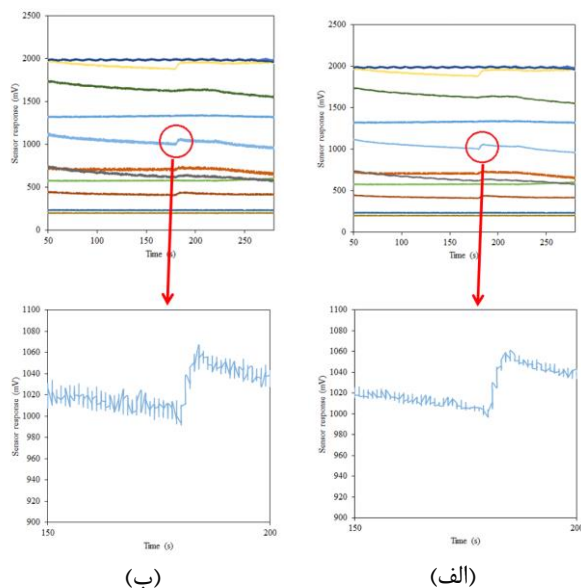
شکل ۳: گالری تصاویر حاصل از مراحل مختلف پردازش تصویر، (الف) تصویر RGB برش خورده، (ب) تصویر EXR، (ج) تصویر باینری پس از حذف نواحی مرزی و پر کردن حفره‌های ناحیه مورد نظر، (د) تصویر RGB روغن با مقدار صفر برای مناطق پس زمینه، (ه) تصویر HSV از روغن با مقدار صفر برای مناطق پس زمینه، (و) تصویر L^*a^*b از روغن با مقدار صفر برای مناطق پس زمینه

۶.۲.۴. تحلیل تشخیصی (DA)

تحلیل تشخیصی یکی از گسترده‌ترین روش‌های تحلیل چند متغیره است که برای تمایز اشیاء به گروه‌های از پیش تعریف شده، استفاده شده است [۴۳]. الگوریتم DA یک الگوریتم ساده و در عین حال دقیق است که ترکیب موفقیت‌آمیز آن با PCA برای در تشخیص اصالت مواد غذایی با داده‌های استخراج شده توسط سامانه‌های بینایی الکترونیک به اثبات رسیده است [۴۴]. بنابراین در این پژوهش، اجزای اصلی (PCs) تولید شده از ویژگی‌های استخراج شده از بینی الکترونیک و بینایی ماشین به دو الگوریتم تجزیه و تحلیل متمایز متداول تحلیل تشخیصی خطی (LDA) و تحلیل تشخیصی درجه دوم (QDA)^{۲۲} با



شکل ۴: تصویر دریافتی توسط سامانه بینایی الکترونیک (الف) و تصویر دارای نویز مصنوعی (ب)



شکل ۵: سیگنال دریافتی توسط سامانه بینایی الکترونیک (الف) و سیگنال دارای نویز مصنوعی (ب)

نکته قابل ذکر دیگر در این پژوهش، در مورد مقایسه پذیری داده‌ها است. مقایسه پذیری یک اصطلاح کلی است که مجموعه‌ای از ویژگی‌های نمونه‌گیری و تجزیه و تحلیل را در بر می‌گیرد. بر این اساس، به طور کلی دو دسته از عوامل بر روی مقایسه پذیری داده‌ها مؤثر هستند که عبارتند از عوامل مربوط به نمونه و داده‌برداری، که عبارت است از نحوه دریافت نمونه‌ها، روش‌های اکتساب داده و شرایط محیطی در زمان نمونه برداری؛ و عوامل مربوط به روش‌های تحلیلی.

در این پژوهش تلاش شد که فرآیند تهیه نمونه‌ها، شرایط نگهداری تا زمان آزمایش مطابق روش‌های موجود در پژوهش‌های مشابه باشد. همچنین، همان‌طور که پیش از این ذکر شد، آزمایش‌های بینایی الکترونیک و تصویربرداری در شرایط کنترل شده و مشابه با پژوهش‌های مشابه انجام شد. چگونگی تهیه تصاویر و ضبط سیگنال‌های بینایی الکترونیک نیز مشابه با عموم پژوهش‌های مرتبط صورت گرفت.

در بخش تحلیل، اصول استخراج اطلاعات از سیگنال‌های بینایی الکترونیک و تصاویر نمونه‌های روغن نیز مطابق با روش‌های مورد

راهبرد اعتبارسنجی متقابل ۱۰ برابری^{۲۳} برای طبقه‌بندی روغن زیتون استفاده شد [۴۵].

تمامی تحلیل‌های آماری با استفاده از نرم‌افزار (CAMO AS, Trondheim, Norway) Unscrambler x10.4 انجام شد.

۷.۲. معیارهای ارزیابی

به منظور مقایسه مدل‌های توسعه یافته و انتخاب بهترین مدل‌های طبقه‌بندی و پیش‌بینی از معیارهای ارزیابی آماری استفاده شد. در مدل‌های طبقه‌بندی، آماره صحت^{۲۴} از ماتریس درهم شدگی^{۲۵} استخراج شد. تعاریف و روابط محاسبه این آماره در پژوهش‌های پیشین موجود است [۴۶، ۴۷]. برای مقایسه مدل‌های پیش‌بینی از آماره‌های ضریب تبیین^{۲۶} (R^2) و مجذور مربعات میانگین خطا (RMSE)^{۲۷} استفاده شد. تعاریف و روابط محاسبه این پارامترهای عملکردی نیز در پژوهش‌های اخیر ارائه شده‌اند [۴۸]. مقادیر R^2 بیشتر و RMSE کمتر نشان دهنده عملکرد بهتر مدل هستند.

۸.۲. بررسی استحکام مدل‌ها

به منظور بررسی استحکام مدل‌ها در این پژوهش، از روش اعمال نویز بر روی داده‌ها استفاده شد. اعمال نویز بر روی سیگنال‌ها و تصاویر به صورت زیر انجام شد. در ابتدا به تصاویر تهیه شده در بخش ۳-۵ نویز از نوع نمک و فلفل با تراکم نویز ۱۰/۰۱ اضافه شد. به این معنی که تقریباً ۱ درصد از پیکسل‌ها تحت تأثیر نویز قرار گرفتند (شکل ۴). از سوی دیگر، به داده‌های اولیه بینایی الکترونیک نیز نویز اضافه شد. به این صورت که مقدار نویز بر اساس رابطه ۳ برای هر حسگر ایجاد شد (شکل ۵).

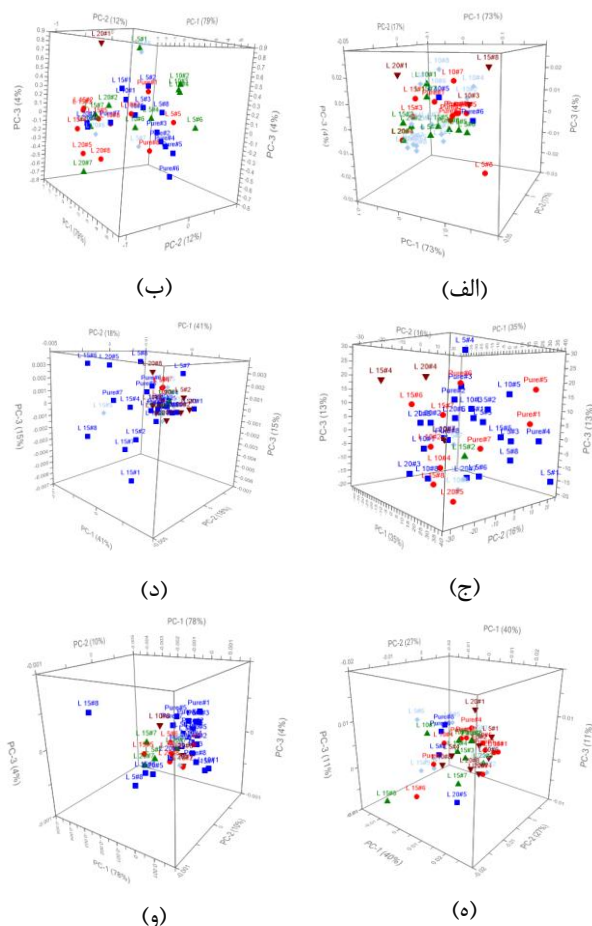
$$(۳) \quad \text{دامنه داده هر حسگر} \pm 0.3 \leq \text{نویز} \leq \text{دامنه داده هر حسگر} - 0.3$$

که در آن دامنه داده هر حسگر عبارت بود از:

$$(۴) \quad \text{دامنه داده حسگر} = \text{کمینه پاسخ حسگر} - \text{بیشینه پاسخ حسگر}$$

پس از ایجاد بردار نویز، نویزها به‌طور تصادفی به هر یک از داده‌ها در نمودارهای اولیه سامانه بینایی الکترونیک اضافه شدند.

بعد از اعمال نویز به داده‌های اولیه تصویر و بینایی الکترونیک، راهبرد ذکر شده در پردازش تصویر و سیگنال بر روی آن‌ها اعمال شد و از ویژگی‌های استخراج شده به منظور ارزیابی استحکام مدل‌ها استفاده شد.



شکل ۶: نمودارهای سه بعدی مؤلفه‌های اصلی برای طبقه‌بندی روغن زیتون فراگیر خالص از نمونه‌های ناخالص دارای روغن لامپانت، با استفاده از ویژگی‌های (الف) MSR، (ب) AUC، (ج) TR، (د) Smax، (ه) MDCV

۲.۱.۳. طبقه‌بندی روغن زیتون خالص و ناخالص با داده‌های بینی الکترونیک و الگوریتم PCA-DA:

نتایج الگوریتم‌های DA برای طبقه‌بندی روغن زیتون خالص با روغن زیتون دارای ناخالصی روغن لامپانت در جدول ۳ ارائه شده‌اند. در این حالت نیز مدل‌های PCA-QDA دقیق‌تر از الگوریتم‌های PCA-LDA با تعداد مؤلفه‌های اصلی ورودی یکسان بودند. بر سی مطالعات پیشین نشان می‌دهد که روش QDA با صحت ۹۵/۸ درصد، عملکرد بهتری نسبت به روش LDA با صحت ۹۴/۴ درصد برای تشخیص تقلب افزودن آب در آبمیوه طبیعی داشته‌است [۴۹]. علاوه بر این، نتایج ارزیابی مدل‌ها نشان داد که صحت طبقه‌بندی با افزایش تعداد مؤلفه‌های اصلی عنوان پارامترهای ورودی افزایش یافت. بالاترین صحت طبقه‌بندی ۱۰۰ درصد توسط PCA-QDA با تعداد در تعدادی از مدل‌ها با تعداد مؤلفه‌های اصلی مختلف تولید شده از ویژگی‌های MSR، AUC به عنوان پارامتر ورودی به دست آمد. این نشان‌دهنده پتانسیل بالای ویژگی‌های بینی الکترونیک برای جداسازی نمونه‌های خالص روغن زیتون فراگیر از نمونه‌های دارای ناخالصی لامپانت است. از بین

استفاده در حوزه‌های مرتبط و براساس منابع معتبر بودند. تصحیح آفست پایه و نرمال سازی داده‌ها، که از روش‌های پیش پردازش رایج در بینی‌های الکترونیکی هستند و باعث افزایش قابلیت مقایسه پذیری داده‌ها، و از بین بردن تفاوت در مقیاس بین داده‌ها می‌شوند، نیز در این پژوهش استفاده شدند. در نهایت، مدل‌های تحلیلی و فرآیند آموزش (اعتبار سنجی متقابل) و آماره‌های عملکردی مورد استفاده نیز براساس اصول حاکم بر مطالعات مشابه انجام شد تا امکان مقایسه نتایج این پژوهش با پژوهش‌های مشابه وجود داشته باشد.

۳. نتایج و بحث

۱.۳.۱. نتایج بینی الکترونیک

۱.۳.۱.۱. نتایج تحلیل مؤلفه‌های اصلی داده‌های بینی الکترونیک:

نمودار امتیازهای حاصل از تحلیل PCA، برای تشخیص روغن زیتون خالص و تقلبی با درصدهای مختلف روغن لامپانت در شکل ۶ ارائه شده‌است. نمودار امتیازهای PCA به منظور گروه‌بندی نمونه‌ها بر اساس شباهت‌ها یا عدم شباهت‌های بر مبنای هر کدام از مؤلفه‌های اصلی استفاده می‌شود. اعداد کنار محورها نشان می‌دهد که هر کدام از مؤلفه‌های اصلی (PCs) ایجاد شده توسط PCA چند درصد از تغییرات داده‌های اولیه را پوشش می‌دهند.

بر اساس نتایج حاصل، سه مؤلفه اول اکثر ویژگی‌های استخراج شده قادر به پوشش درصد بالایی واریانس داده‌ها بودند به طوری که سه مؤلفه اول در ویژگی‌های MDCV، Smax، Sase، TR، AUC، MSR به ترتیب ۹۴ درصد (PC1=۷۳ درصد، PC2=۱۷ درصد و PC3=۴ درصد)، ۹۵ درصد (PC1=۷۹ درصد، PC2=۱۲ درصد و PC3=۴ درصد)، ۶۴ درصد (PC1=۳۵ درصد، PC2=۱۶ درصد و PC3=۱۳ درصد)، ۷۴ درصد (PC1=۴۱ درصد، PC2=۱۸ درصد و PC3=۱۵ درصد)، ۷۸ درصد (PC1=۴۰ درصد، PC2=۲۷ درصد و PC3=۱۱ درصد)، و ۹۲ درصد (PC1=۷۸ درصد، PC2=۱۰ درصد و PC3=۴ درصد) از کل واریانس داده‌ها را پوشش دهند. بیشترین پوشش داده‌ها به ترتیب مربوط به سه مؤلفه اول ویژگی‌های AUC با ۹۵ درصد، MSR با ۹۴ درصد، و MDCV با ۹۲ درصد بود. از سوی دیگر، در نمودارهای شکل ۶ مشاهده می‌شود که توزیع داده‌های گروه‌های مختلف در نمودار سه بعدی به گونه‌ای است که با نگاه کردن از هر کدام از نماهای دوبعدی (حاصل از یک جفت مؤلفه اصلی) نمی‌توان تمامی گروه‌ها را از هم جدا کرد. بنابراین می‌توان گفت که استفاده از جفت مؤلفه‌ها در جدا سازی نمونه‌های خالص از درصدهای مختلف ناخالصی منتج به دقت بالایی نخواهد شد. در این صورت به نظر می‌رسد استفاده از ترکیبی از چند مؤلفه اصلی می‌تواند راهگشا باشد.

و دو مؤلفه اصلی اول در مدل های AUC-PCA-QDA و TR-PCA-QDA به ترتیب منتج به صحت طبقه بندی ۶۰ و ۷۲/۵ درصدی شده اند. از سوی دیگر، مشاهده شد که اگرچه با استفاده از سه مؤلفه اول PCA (شکل ۶)، ویژگی هایی مانند MSR جداسازی بهتری بین کلاس های مختلف انجام داده اند، اما استفاده از سایر مؤلفه ها باعث افزایش صحت طبقه بندی DA بر اساس ویژگی TR می شود (جدول ۳). این امر اهمیت مؤلفه های سوم تا ششم را در ویژگی TR نشان می دهد. اینکه سه مؤلفه اول ویژگی TR درصد واریانس کمتری از داده ها را

جدول ۳: مقادیر صحت طبقه بندی (بر حسب درصد) ساختارهای DA برای تمایز روغن زیتون فرابرگر خالص از نمونه های ناخالص دارای درصدهای مختلف روغن لامپانت با استفاده از داده های بینی الکترونیک

طبقه بند	تعداد مؤلفه ورودی							
	۱	۲	۳	۴	۵	۶	۷	۸
MSR-PCA-LDA	۲۷/۵	۶۵	۸۰	۸۰	۸۲/۵	۸۷/۵	۸۷/۵	۹۲/۵
MSR-PCA-QDA	۳۵	۷۰	۹۰	۹۵	۹۵	۹۷/۵	۱۰۰	۱۰۰
AUC-PCA-LDA	۲۷/۵	۵۵	۵۵	۶۷/۵	۷۰	۷۲/۵	۸۰	۸۰
AUC-PCA-QDA	۳۲/۵	۶۰	۸۲/۵	۸۵	۹۵	۱۰۰	۱۰۰	۱۰۰
TR-PCA-LDA	۲۲/۵	۷۰	۷۵	۷۷/۵	۷۷/۵	۷۵	۷۵	۸۰
TR-PCA-QDA	۳۰	۷۲/۵	۸۲/۵	۹۲/۵	۹۷/۵	۱۰۰	۱۰۰	۱۰۰
Sasce-PCA-LDA	۳۰	۶۲/۵	۷۰	۶۰	۷۵	۸۵	۸۷/۵	۹۲/۵
Sasce-PCA-QDA	۳۵	۵۷/۵	۶۷/۵	۷۷/۵	۹۲/۵	۹۷/۵	۱۰۰	۱۰۰
Smax-PCA-LDA	۳۵	۶۰	۶۲/۵	۶۰	۷۰	۶۷/۵	۷۷/۵	۷۷/۵
Smax-PCA-QDA	۴۷/۵	۶۷/۵	۷۵	۸۵	۹۵	۹۲/۵	۹۷/۵	۱۰۰
MDCV-PCA-LDA	۲۵	۴۷/۵	۵۲/۵	۶۲/۵	۶۵	۶۷/۵	۶۷/۵	۷۰
MDCV-PCA-QDA	۳۵	۵۲/۵	۷۲/۵	۹۲/۵	۹۷/۵	۹۷/۵	۱۰۰	۱۰۰

پوشش می دهند نیز به خاطر همین موضوع است.

۳.۱.۳. نتایج مدل سازی میزان ناخالصی روغن زیتون با داده های بینی الکترونیک و الگوریتم MLR:

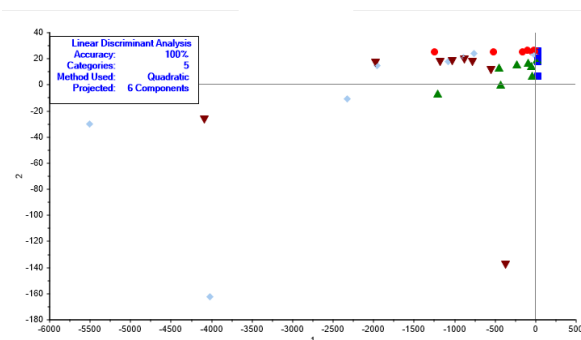
نتایج پیش بینی درصد ناخالصی روغن لامپانت در روغن زیتون فرابرگر با استفاده از مدل های رگرسیون خطی در جدول ۴ ارائه شده است. مشاهده شد که مدل رگرسیون خطی توسعه یافته بر اساس ویژگی TR موفق ترین مدل در پیش بینی میزان تقلب بود.

جدول ۴: مقادیر عملکردی مدل های MLR برای پیش بینی میزان تقلب در روغن زیتون خالص با روغن لامپانت با استفاده از داده های بینی الکترونیک

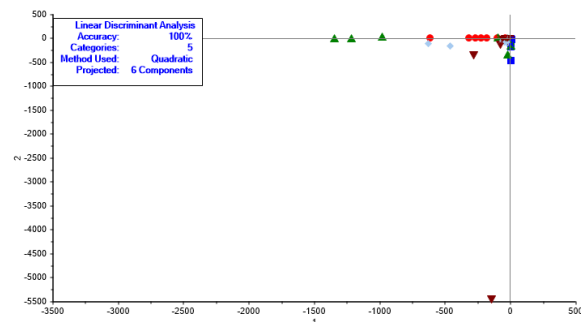
ویژگی ورودی	کالیبراسیون		اعتبارسنجی	
	R ²	RMSE	R ²	RMSE
MSR	۰/۸۰۷۰	۳/۸۵۳۱	۰/۴۳۱۰	۵/۴۰۲۲
AUC	۰/۸۴۸۲	۳/۴۱۷۴	۰/۶۲۹۰	۴/۳۶۲۰
TR	۰/۸۷۰۷	۳/۱۵۴۰	۰/۶۳۰۵	۴/۳۵۲۸
Sasce	۰/۷۵۱۳	۴/۳۷۳۶	۰/۴۰۶۶	۵/۵۱۶۳
Smax	۰/۸۴۷۷	۳/۴۲۳۳	۰/۵۰۸۱	۵/۰۲۲۷
MDCV	۰/۷۶۴۸	۴/۲۵۳۷	۰/۳۱۲۳	۶/۳۶۶۷

مدل های با صحت طبقه بندی ۱۰۰ در صد، ساختارهای AUC-PCA-QDA و TR-PCA-QDA با شش مؤلفه ورودی (طبقه بندهای AUC-PCA6-QDA و TR-PCA6-QDA) با توجه به کمتر بودن تعداد مؤلفه های اصلی که به عنوان داده های ورودی آن ها استفاده شده بودند، به عنوان مناسب ترین طبقه بندی کننده انتخاب شدند. این عملکرد بالاتر از پژوهشی است که در آن طبقه بندی DA بدون عملیات استخراج ویژگی بر روی داده های بینی الکترونیک توانست تقلب در رب انار را با صحت ۹۷/۶۵ درصد تشخیص دهد [۵۰].

نمایش گرافیکی نتایج این دو طبقه بندی در شکل ۷ ارائه شده است. نمونه های مربوط به کلاس های مختلف با شکل و رنگ متفاوتی نشان داده شده اند. نکته اینکه در این شکل، نمونه ها در صفحه مؤلفه اصلی دوم در مقابل مؤلفه اصلی اول ارائه شده اند، در حالیکه تعداد مؤلفه های اصلی مورد نیاز برای نیل به صحت ۱۰۰ درصد در هر دو مدل برابر با ۵ مؤلفه اصلی بوده است.



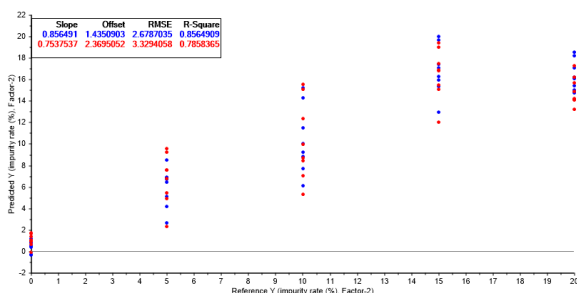
(الف)



(ب)

شکل ۷: خروجی گرافیکی دقیق ترین طبقه بندهای AUC-PCA6-QDA (الف) و TR-PCA6-QDA (ب) در صفحه مؤلفه های اصلی اول و دوم برای طبقه بندی روغن زیتون فرابرگر خالص از روغن های دارای درصدهای مختلف ناخالصی با روغن لامپانت با استفاده از داده های بینی الکترونیک

به عبارت دیگر، شکل ۷ نشان می دهد که استفاده از تنها دو مؤلفه اصلی اول و دوم منتج به جداسازی کامل نمونه های مربوط به کلاس های مختلف نمی شود. نکته ای که در جدول ۳ نیز مشاهده می شود



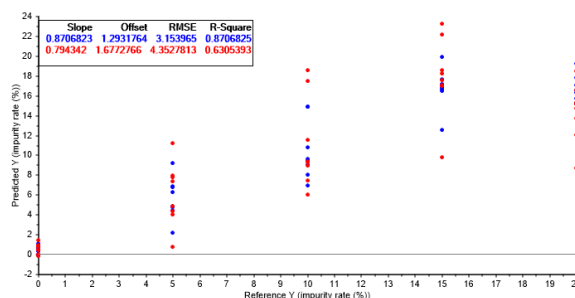
شکل ۹: نمودار گرافیکی نتیجه مدل PLS برای تخمین میزان ناخالصی لامپانت در روغن زیتون فرابکر خالص براساس ویژگی TR در مراحل کالیبراسیون (آبی) و اعتبارسنجی (قرمز)

در مطالعه‌ای که بر روی بررسی ماندگاری روغن‌های خوراکی توسط بینی الکترونیک انجام شد از یک سامانه بینی الکترونیک با چهار حسگر گاز نیمه‌هادی اکسید فلزی، شامل TGS2600، TGS2611، TGS2620 و MQ138 استفاده شد. در نهایت، سه روش تشخیص الگو، شامل LDA، نزدیک‌ترین همسایه (KNN) و SVM برای ایجاد یک مدل شناسایی کیفی دوره ذخیره سازی روغن خوراکی مقایسه شدند. نتایج نشان داد که روش‌های کمومتریک در ترکیب با بینی الکترونیک دقت بالایی در طبقه بندی و تشخیص ماندگاری روغن خوراکی دارند [۵۱]. در پژوهش دیگری از یک ماشین بینی الکترونیک جهت بررسی میزان اکسیداسیون روغن‌های خوراکی استفاده شد. در این تحقیق از آرایه حسگری گازی استفاده شد و حسگرهای TGS2602 و MQ136 بالاترین مقادیر ضریب حساسیت و حسگرهای MQ9، TGS822، TGS813 و TGS2620 کمترین مقادیر حساسیت را داشتند. بر اساس نتایج، صحت دو روش تحلیل PCA و PLS به ترتیب ۹۸/۳۳ درصد و ۱۰۰ درصد بود [۵۲]. همچنین از ترکیب بینی الکترونیک، زبان الکترونیک و بینایی ماشین جهت ارزیابی کیفیت روغن زیتون در طول نگهداری در دماهای مختلف استفاده شد. نتایج نشان داد که با ترکیب داده‌های حواسی الکترونیک و کاربرد الگوریتم KNN می‌توان برای طبقه‌بندی نمونه‌ها در دو کلاس تازگی با صحت ۹۴ درصد استفاده کرد [۵۳].

۲.۲. نتایج بینایی ماشین

۲.۲.۱. نتایج تحلیل مؤلفه‌های اصلی داده‌های رنگی:

نتایج آنالیز مؤلفه‌های اصلی ویژگی‌های رنگی استخراج شده از تصاویر نمونه‌های خالص زیتون فرابکر و نمونه‌های دارای سطوح مختلف انواع روغن ناخالص در شکل ۱۰ آمده است. سه مؤلفه اول ویژگی‌های تصویر قادر به پوشش درصد بالایی از واریانس داده‌ها در انواع تقلب بودند، به‌طوری‌که میزان پوشش واریانس سه مؤلفه اول برای تشخیص تقلب با روغن لامپانت برابر با ۹۲ درصد (PC1=۵۷ درصد، PC2=۲۴ درصد و PC3=۱۱ درصد) بود. این درصدهای بالا در پوشش واریانس داده‌ها نشان می‌دهد، سه مؤلفه اصلی استخراج شده از ویژگی‌های تصاویر



شکل ۸: نمودار گرافیکی نتیجه مدل MLR برای تخمین میزان ناخالصی لامپانت در روغن زیتون فرابکر خالص براساس ویژگی TR در مراحل کالیبراسیون (آبی) و اعتبارسنجی (قرمز)

نتایج ارزیابی نشان داد که مقدار R^2 مدل TR-MLR در مرحله کالیبراسیون و اعتبارسنجی برای به ترتیب برابر با ۰/۸۷۰۷ و ۰/۶۳۰۵ بود که نشان می‌دهد دقت سامانه بینی الکترونیک در پیش‌بینی میزان تقلب حاصل از اختلاط روغن لامپانت در روغن زیتون فرابکر، بالا بود. این موضوع از آن جهت قابل توجه است که روغن لامپانت نیز از جنس روغن زیتون اما با کیفیت بسیار پایین است (شکل ۸).

۴/۱/۳. نتایج مدل‌سازی میزان ناخالصی روغن زیتون با داده‌های

بینی الکترونیک و الگوریتم PLS:

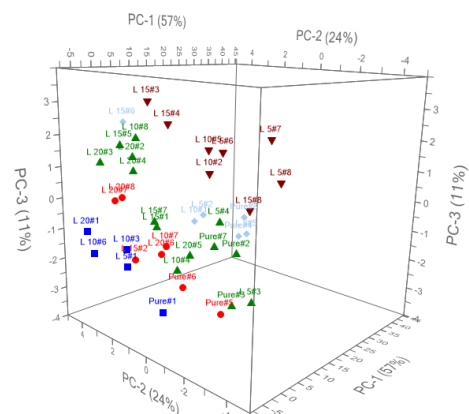
نتایج مدل‌های PLS برای تخمین تقلب روغن لامپانت در روغن زیتون فرابکر با استفاده از ویژگی‌های استخراج شده از رایحه نمونه‌ها در جدول ۵ آمده است. مشاهده شد که مدل TR-PLS به ترتیب با مقادیر R^2 و RMSE برابر با ۰/۸۵۶۵ و ۲/۶۷۸۷ در مرحله کالیبراسیون و مقادیر ۰/۷۸۵۸ و ۳/۳۲۹۴ در مرحله اعتبارسنجی، دقیق‌ترین مدل پیش‌بینی کننده درصد تقلب روغن لامپانت بوده است. مدل PLS از مدل رگرسیون خطی دارای عملکرد بهتری بود اما همچنان نتایج پیش‌بینی درصد تقلب با روغن لامپانت در هنگام استفاده از داده‌های بینی الکترونیک مناسب نبود. شکل ۹ مقادیر پیش‌بینی شده در صد تقلب با روغن لامپانت را در مقابل مقادیر مرجع نشان می‌دهد.

جدول ۵: مقادیر عملکردی مدل‌های PLS برای پیش‌بینی میزان تقلب در روغن زیتون خالص با روغن لامپانت با استفاده از داده-

های بینی الکترونیک

ویژگی ورودی	متغیرهای پنهان	کالیبراسیون		اعتبارسنجی	
		R^2	RMSE	R^2	RMSE
MSR	۲	۰/۶۳۶۶	۴/۲۶۲۴	۰/۵۷۴۵	۴/۷۷۵۱
AUC	۷	۰/۸۳۱۶	۲/۹۰۱۳	۰/۷۱۲۹	۳/۹۱۹۸
TR	۲	۰/۸۵۶۵	۲/۶۷۸۷	۰/۷۸۵۸	۳/۳۲۹۴
Sasce	۲	۰/۵۶۹۰	۴/۶۴۲۵	۰/۴۶۳۳	۵/۳۰۴۱
Smax	۱۱	۰/۸۴۶۷	۲/۷۶۸۹	۰/۶۲۲۳	۴/۴۴۶۴
MDCV	۲	۰/۵۹۷۶	۴/۴۸۵۷	۰/۵۰۰۸	۵/۰۵۰۸

نمایندۀ مناسبی کل داده‌های استخراج‌شده از تصاویر هستند اما همچنان اطلاعات مهمی در سایر مؤلفه‌های اصلی وجود دارد.



شکل ۱۰: نمودارهای سه بعدی مؤلفه‌های اصلی برای طبقه‌بندی روغن زیتون فرابرک خالص از نمونه‌های ناخالص دارای روغن لامپانت با استفاده از ویژگی‌های استخراج شده از تصاویر

۲/۲/۳. طبقه‌بندی روغن زیتون خالص و ناخالص با داده‌های رنگی و الگوریتم PCA-DA :

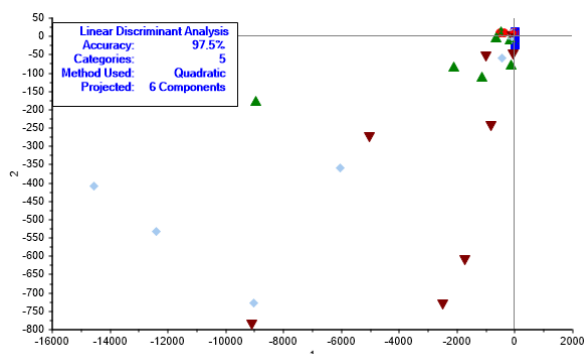
نتایج الگوریتم‌های DA برای طبقه‌بندی روغن زیتون خالص از انواع روغن زیتون دارای ناخالصی، به ازای تعداد مختلف مؤلفه اصلی در جدول ۶ آمده است. به طور کلی مدل‌های QDA از مدل‌های LDA موفق‌تر بودند. بالاترین صحت طبقه‌بندی روغن زیتون خالص از روغن زیتون با درصدهای مختلف روغن لامپانت برابر با ۹۷/۵ درصد بود که توسط مدل PCA-QDA با ۶ مؤلفه اصلی به دست آمد. نتایج نشان می‌دهد که ویژگی‌های تصویر دارای قابلیت بالایی در طبقه‌بندی روغن زیتون فرابرک خالص از انواع روغن زیتون ناخالص ایجاد شده با روغن لامپانت است. بنابراین با ایجاد یک سامانه بینایی ماشین مناسب می‌توان به صورت غیرمخرب و با دقت و سرعت بالایی نمونه‌های روغن زیتون دارای تقلب را تشخیص داد.

جدول ۶: مقادیر صحت طبقه‌بندی (بر حسب درصد)

ساختارهای DA برای تمایز روغن زیتون فرابرک خالص از نمونه‌های ناخالص با استفاده از ویژگی‌های تصاویر

تعداد مؤلفه ورودی	طبقه بند	
	PCA-QDA	PCA-LDA
۱	۶۵	۳۲/۵
۲	۷۵	۴۷/۵
۳	۸۷/۵	۶۰
۴	۹۰	۵۷/۵
۵	۹۲/۵	۶۰
۶	۹۷/۵	۷۰
۷	۹۷/۵	۷۰
۸	۹۷/۵	۷۰

شکل ۱۱: نتایج مربوط به طبقه‌بندی ذکر شده برای تشخیص تقلب در روغن زیتون فرابرک خالص از روغن زیتون ناخالص با سطوح مختلف روغن لامپانت را نمایش می‌دهد.



شکل ۱۱: خروجی گرافیکی دقیق‌ترین طبقه‌بندی‌های PCA-QDA برای طبقه‌بندی روغن زیتون فرابرک خالص از روغن‌های دارای درصدهای مختلف ناخالصی با روغن لامپانت

نکته دیگر اینکه، جدول ۶ نشان می‌دهد که مدل PCA-QDA به مؤلفه اصلی قادر به طبقه‌بندی روغن زیتون فرابرک خالص از روغن زیتون دارای سطوح مختلف ناخالصی لامپانت با صحت ۸۷/۵ درصد است. این عملکرد نشان می‌دهد که اتکا به تنها دو یا سه مؤلفه اصلی در برخی از موارد طبقه‌بندی منتج به نتایج تشخیصی مناسبی نمی‌شود.

۳،۲،۳. نتایج مدل‌سازی میزان ناخالصی روغن زیتون با داده‌های رنگی:

نتایج مدل‌های MLR و PLS برای تخمین تقلب در روغن زیتون فرابرک با استفاده از ویژگی‌های استخراج شده از تصاویر نمونه‌ها در جدول ۷ آمده است. مشاهده شد که دقت مدل‌ها برای پیش‌بینی نرخ تقلب روغن لامپانت در روغن زیتون فرابرک، مقادیر مناسبی نبود. مقادیر عملکردی مدل PLS این مدل با R^2 و RMSE برابر با ۰/۶۹۸۳ و ۴/۴۸۴۷ در مرحله کالیبراسیون و مقادیر ۰/۴۹۳۶ و ۵/۰۹۶۰ در مرحله اعتبارسنجی، نشان داد که استفاده از ویژگی‌های رنگی به‌تنهایی در تشخیص تقلب روغن لامپانت مناسب نیست. این نتیجه با توجه به اینکه روغن لامپانت نیز در طی فرایند روغن‌گیری از زیتون به دست می‌آید و دارای ماهیتی تقریباً مشابه با روغن زیتون فرابرک است، قابل توجه است. نتایج گرافیکی بهترین مدل‌ها جهت پیش‌بینی نرخ انواع تقلب در روغن زیتون فرابرک در شکل ۱۲ ارائه شده است.

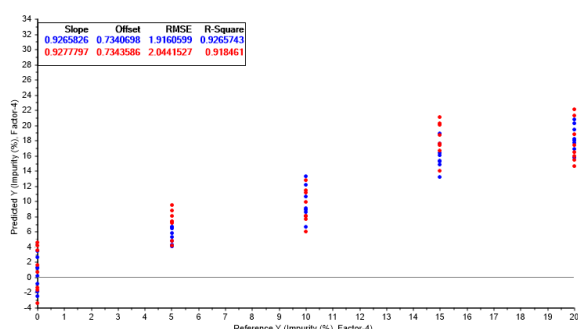
جدول ۷: مقادیر عملکردی مدل‌های PLS و MLR برای پیش‌بینی میزان تقلب در روغن زیتون خالص با روغن لامپانت با استفاده از

داده‌های رنگی

مدل	متغیرهای پنهان	کالیبراسیون		اعتبارسنجی	
		R^2	RMSE	R^2	RMSE
PLS	۵	۰/۶۹۸۳	۴/۴۸۴۷	۰/۴۹۳۶	۵/۰۹۶۰
MLR	-	۰/۶۳۱۳	۴/۴۹۳۶	۰/۴۶۹۰	۵/۳۰۵۴

جدول ۸: معیارهای عملکرد بهترین مدل PLS برای پیش‌بینی نرخ تقلب در روغن زیتون فرابکر خالص با روغن لامپانت با استفاده از ترکیب داده‌های رنگی و بینایی الکترونیک

ترکیب بهینه داده‌ها	متغیرهای پنهان	کالیبراسیون		اعتبارسنجی	
		R ²	RMSE	R ²	RMSE
PLS	۵	۰/۹۲۶۶	۱/۹۱۶۱	۰/۹۱۸۴	۲/۰۴۴۲

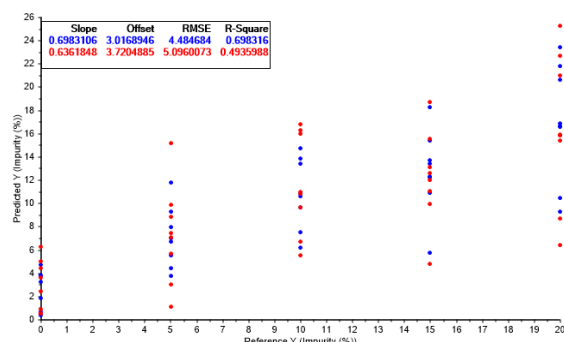


شکل ۱۳: نمودار گرافیکی نتیجه مدل‌سازی برای تخمین میزان ناخالصی روغن لامپانت در روغن زیتون فرابکر خالص براساس ترکیب ویژگی‌های بینایی و بویایی الکترونیک در مراحل کالیبراسیون (آبی) و اعتبارسنجی (قرمز)

مقایسه نتایج این پژوهش با پژوهش‌های مشابه نشان می‌دهد که الگوریتم پیاده‌سازی‌شده در این پژوهش منتج به عملکرد بهتری در تشخیص تقلب شده‌است. به عنوان مثال الگوریتم‌های SPA-LDA و PLS-DA منتج به طبقه‌بندی نمونه‌های EVOO خالص و مخلوط با روغن سویا با نرخ‌های ۸۸ درصد و ۹۴ داده‌های رنگی مستخرج از تصاویر نمونه‌ها شدند [۱۲]. استفاده از ویژگی‌های مستخرج از تصاویر منتج به طبقه‌بندی تقلب در شیر انگور با صحت ۹۱/۶۷ درصد توسط مدل KNN شد [۵۴]. مقایسه الگوریتم‌های مختلف یادگیری ماشین برای درجه‌بندی کیفی روغن زیتون براساس داده‌های بینایی الکترونیک نشان داد که بیشترین صحت طبقه‌بندی مدل‌های SVM، ANN و DT و به ترتیب برابر با ۹۰ درصد، ۹۵ درصد و ۹۸/۳۳ درصد بود [۵۵]. در پژوهش دیگری، صحت مدل LDA برای طبقه‌بندی درجات کیفی مختلف روغن زیتون با استفاده از داده‌های بینایی الکترونیک برابر با ۹۲ درصد گزارش شده‌است [۲۳]. صحت مدل LDA برای طبقه‌بندی تقلب در رب انار مخلوط‌شده با شیر خرمای یک سامانه بینایی الکترونیک برابر با ۹۷/۶۵ درصد گزارش شد [۵۰].

۴.۳. نتایج بررسی استحکام مدل‌ها

به منظور بررسی استحکام بهترین مدل‌های انتخاب‌شده جهت طبقه‌بندی سطوح مختلف تقلب و پیش‌بینی نرخ تقلب، به تصاویر تهیه‌شده از روغن‌ها و سیگنال‌های مستخرج از سامانه بینایی الکترونیک نویز اعمال شد و فرایندهای پردازشی و استخراج ویژگی بر روی تصاویر



شکل ۱۲: نمودار گرافیکی نتیجه مدل‌سازی برای تخمین میزان ناخالصی روغن لامپانت توسط مدل PLS در روغن زیتون فرابکر خالص براساس ویژگی‌های تصویر در مراحل کالیبراسیون (آبی) و اعتبارسنجی (قرمز)

۳.۳. نتایج تشخیص نرخ ناخالصی در روغن زیتون فرابکر با استفاده از ترکیب داده‌های بینایی و بویایی الکترونیک

نتایج بخش‌های قبلی نشان داد که با استفاده از هر کدام از داده‌های بینایی الکترونیک یا بینایی ماشین می‌توان نهایتاً کلاس‌های مختلف روغن زیتون با درصد‌های مختلف تقلب را با صحت بسیار بالایی درصد از همدیگر تشخیص داد. اما در مرحله توسعه مدل‌های رگرسیونی تخمین میزان ناخالصی، همچنان توسعه مدل‌های دارای دقت بالاتر مطلوب است.

بنابراین در بخش دیگری از پژوهش به توسعه مدل‌های پیش‌بینی نرخ تقلب استفاده از ترکیب داده‌های حاصل از رنگ و بوی روغن پرداخته شد. ادغام ویژگی‌ها در سطح متوسط انجام شد و راهبرد انتخاب‌شده به این صورت بود که داده‌های رنگی استخراج‌شده از نمونه‌های روغن با هر یک از ویژگی‌های AUC، MSR، TR، Sase، Smax و MDCV ترکیب شدند. از سوی دیگر با توجه به ارجحیت مدل PLS نسبت به MLR در نتایج بخش‌های پیشین، از مدل PLS برای پیش‌بینی نرخ ناخالصی در روغن زیتون استفاده شد. نتایج بهترین ترکیب‌های داده‌ها در جدول ۸ ارائه شده است. مقایسه پارامترهای عملکردی نشان می‌دهد که ترکیب داده‌های بینایی و بویایی الکترونیک باعث افزایش قابل توجه دقت پیش‌بینی نسبت به حالت‌های استفاده از داده‌های بینایی و بویایی الکترونیک به صورت مجزا می‌شود. از سوی دیگر، با وجود اینکه ویژگی TR در حالت مجزا باعث بهترین دقت پیش‌بینی میزان اختلاط شده بود اما در حالت همجوشی ویژگی‌ها، ترکیب ویژگی AUC از سامانه بینایی الکترونیک با ویژگی‌های رنگی مؤثرترین حالت ترکیب داده برای پیش‌بینی نرخ ناخالصی در روغن زیتون فرابکر شده‌است. نتایج مدل‌های PLS برای پیش‌بینی نرخ ناخالصی در روغن زیتون فرابکر در شکل ۱۳ ارائه شده‌است.

به میزان قابل توجهی کاهش یابد و مدل نهایی ارائه شده دارای استحکام مطلوبی باشد.

جدول ۱۰: نتایج ارزیابی استحکام مدل AUC-PCA-QDA برای طبقه‌بندی نمونه‌های روغن زیتون فرابرگر خالص با سطوح مختلف تقلب با روغن لامپانت با استفاده از داده‌های بینایی الکترونیک

صحت (درصد)	مدل
۹۷/۵۰	AUC - PCA - QDA

۳،۴،۳. نتایج بررسی استحکام مدل (Color + AUC)-PLS برای پیش‌بینی نرخ تقلب با استفاده از همجوشی داده‌های بینایی الکترونیک و بینایی ماشین:

جدول ۱۱ نتایج بررسی استحکام مدل حاصل همجوشی داده‌های بینایی ماشین و بینایی الکترونیک در پیش‌بینی نرخ تقلب در روغن زیتون فرابرگر با روغن لامپانت را نشان می‌دهد. مقادیر R^2 و RMSE حاصل از مدل (Color + AUC)-PLS بر روی داده‌های دارای اغتشاش نشان می‌دهد که اعمال نویز به سیگنال‌های بینایی الکترونیک و تصاویر روغن زیتون تاثیر قابل توجهی در صحت مدل برای پیش‌بینی نرخ تقلب در روغن زیتون فرابرگر را نداشته‌است.

با توجه به مقادیر عملکردی به‌دست آمده در بخش‌های ۴-۶ تا ۴-۳ می‌توان گفت که راهبرد به‌کار برده‌شده در این پژوهش به‌منظور استخراج ویژگی‌ها از داده‌های خام و مدل‌های توسعه‌یافته دارای استحکام بالایی نسبت به نویزهای احتمالی است. این امر قابلیت اعتمادپذیری بالایی نتایج این پژوهش را اثبات می‌کند. در پایان مقایسه نتایج به‌دست آمده در این پژوهش با مطالعات خلاصه‌شده در جدول ۱ نشان می‌دهد که راهبرد استفاده شده در این پژوهش جهت همجوشی داده‌های بینایی الکترونیک و بینایی ماشین به‌طور مؤثری باعث افزایش عملکرد طبقه‌بندی روغن زیتون فرابرگر خالص از روغن‌های ناخالص و نیز تعیین نرخ تقلب می‌شود.

جدول ۱۱: نتایج ارزیابی استحکام مدل (Color + AUC)-PLS برای تخمین نرخ تقلب در روغن زیتون فرابرگر خالص با مقادیر مختلف روغن لامپانت با استفاده از ترکیب داده‌های رنگی و بینایی الکترونیک

مدل	R^2	RMSE
(Color + AUC)-PLS	۰/۹۱۰۳	۲/۰۹۹۱

۴. محدودیت‌های سامانه

اگرچه سامانه‌های بینایی الکترونیک در حال حاضر به‌طور گسترده در حوزه‌های تحقیقاتی و حتی توسط برخی از شرکت‌ها به عنوان ابزار کنترل کیفیت استفاده می‌شوند، اما همچنان محدودیت‌هایی در زمینه استفاده از این سامانه‌ها وجود دارد. به عنوان مثال، سامانه‌های بینایی

و سیگنال‌های دارای اغتشاش عمدی اعمال شدند. ویژگی‌های استخراج شده به مدل‌های بهینه داده‌شد و تغییر پارامترهای عملکردی مدل‌ها تحت نویز ارزیابی شد.

همانطور که پیشتر ذکر شد، به دلیل صحت بالای (صحت ۱۰۰ درصد) سامانه‌های بینایی الکترونیک و بینایی تصویر در عملیات طبقه‌بندی نمونه‌های دارای درصد‌های مختلف تقلب، استفاده از داده‌های مجزا به منظور عملیات طبقه‌بندی استفاده‌شد. اما در بخش پیش‌بینی نرخ تقلب، همجوشی داده‌ها باعث افزایش عملکرد مدل‌ها شد. بنابراین به منظور بررسی استحکام مدل‌ها نیز موارد زیر بررسی شد. (۱) مدل PCA-QDA برای طبقه‌بندی نمونه‌های روغن زیتون با استفاده از داده‌های بینایی ماشین.

(۲) مدل AUC-PCA-QDA برای طبقه‌بندی نمونه‌های روغن زیتون با استفاده از داده‌های بینایی الکترونیک.

(۳) مدل (Color + AUC)-PLS برای پیش‌بینی نرخ تقلب با استفاده از همجوشی داده‌های بینایی الکترونیک و بینایی ماشین

۱،۴،۳. نتایج بررسی استحکام مدل PCA-QDA برای طبقه‌بندی نمونه‌های روغن زیتون با استفاده از داده‌های بینایی ماشین:

جدول ۹ نتایج مدل PCA-QDA با استفاده از داده‌های رنگی مستخرج از تصاویر حاوی نویز را در طبقه‌بندی روغن‌های زیتون فرابرگر با درصد‌های مختلف اختلاط با روغن لامپانت را نشان می‌دهد. مشاهده می‌شود اعمال نویز در سطح ذکر شده در بخش ۳-۷ باعث کاهش بسیار اندکی در صحت طبقه‌بندی شده است اما این تاثیر قابل توجه نبوده است و راهبرد مورد استفاده دارای استحکام مناسبی است.

جدول ۹: نتایج ارزیابی استحکام مدل PCA-QDA برای طبقه‌بندی نمونه‌های روغن زیتون فرابرگر خالص با سطوح مختلف تقلب با روغن لامپانت با استفاده از داده‌های تصاویر

صحت (درصد)	مدل
۹۵/۰۰	Color - PCA - QDA

۲،۴،۳. نتایج بررسی استحکام مدل AUC-PCA-QDA برای طبقه‌بندی نمونه‌های روغن زیتون با استفاده از داده‌های بینایی الکترونیک:

جدول ۱۰ نتایج بررسی استحکام مدل AUC-PCA-QDA با استفاده از داده‌های AUC حاصل از سیگنال‌های بینایی الکترونیک حاوی نویز را در طبقه‌بندی نمونه‌های روغن زیتون فرابرگر با درصد‌های مختلف اختلاط با روغن لامپانت را نشان می‌دهد. نتایج بررسی استحکام مدل نشان داد که اعمال نویز به سیگنال‌های بینایی الکترونیک به روش ذکر شده در بخش ۳-۷ تأثیر قابل توجهی در عملکرد طبقه‌بندی نداشته‌است. نرخ داده‌برداری بالای دستگاه (۱۰ داده در هر ثانیه) و عملیات پردازشی صورت‌گرفته بر روی سیگنال‌های خام باعث می‌شود تاثیر احتمالی نویز

الکترونیک مانند برخی ابزارهای آزمایش شیمیایی (به عنوان مثال: کروماتوگرافی گازی) توانایی شناسایی ترکیبات منفرد در مخلوط‌های پیچیده را ندارند و عموماً برای طبقه‌بندی بوها و یا تعیین برخی پارامترهای محصول براساس الگوی بو استفاده می‌شوند. اما در این مورد، برای تعیین نوع یا غلظت یک ماده خاص در ترکیب بوی نمونه، نیاز به انجام آزمایش‌ها و توسعه‌های الگوریتم‌های مدل‌سازی و یافتن ارتباط بین الگوی بو و ویژگی مورد نظر است. از سوی دیگر، محیط آزمایشی در سامانه‌های بینی الکترونیک بسیار مهم است. آرایه حسگرهای بینی الکترونیک ممکن است حساسیت خود را در حضور بخار آب زیاد یا غلظت بالای ماده‌ای مانند الکل از دست بدهند. تغییرات زیاد دما باعث تغییر در پاسخ حسگرهای بینی الکترونیک می‌شود و داده‌برداری در محلی دارای بوهای شدید محیطی می‌تواند باعث ایجاد خطا شود. بنابراین بسیار مهم است که داده‌برداری در محیط کنترل شده انجام شوند. همچنین توصیه می‌شود که حساسیت حسگرها به صورت دوره‌ای کنترل شود.

در مورد سامانه بینی‌های ماشین نیز، از آنجاکه در این پژوهش از داده‌های رنگی استفاده شده‌است، کنترل شرایط نوری محیط و استفاده از سامانه نورپردازی بسیار مهم است. نهایتاً درمورد هدف نهایی پژوهش‌هایی مانند تحقیق حاضر، یعنی توسعه سامانه‌ای که بتواند به‌طور تجاری و قابلیت اعتماد بالا انواع تقلب‌های روغن زیتون را تشخیص دهد، باید گفت که این امر اولاً نیازمند آزمایش‌های بسیار زیاد با انواع تقلب‌ها به منظور ایجاد پایگاه داده جامع بوده همچنین نیازمند توسعه، ارزیابی و انتخاب مدل‌های بهینه یادگیری ماشین و هوش مصنوعی است.

۵. نتیجه‌گیری و مسیر تحقیقات آینده

در این پژوهش از ترکیب روش‌های غیر مخرب ماشین بویایی و بینی‌های جهت شناسایی و طبقه‌بندی روغن زیتون فرابرک خالص از روغن‌های دارای ناخالصی (روغن زیتون خالص به همراه اختلاط روغن لایمانت) استفاده شد. بدین منظور در روغن خوراکی خالص به میزان ۵، ۱۰، ۱۵ و ۲۰ درصد ناخالصی افزوده شد و سپس نمونه‌های روغن خالص و ناخالص (۵ گروه روغن) در مقابل ۱۳ حسگر بینی الکترونیک و سامانه پردازش تصویر قرار گرفت.

سامانه بینی الکترونیک در ترکیب با روش‌های کومتریک نتایج خوبی را برای جداسازی نمونه‌های روغن خالص و ناخالص و همچنین پیش‌بینی نرخ ناخالصی روغن زیتون نشان داد. نتایج نشان داد روش PCA به خوبی توانستند عمل جداسازی نمونه‌های روغن زیتون خالص و ناخالص را انجام دهند. همچنین روش QDA موفق شد با استفاده از داده‌های بینی الکترونیک و با صحت ۱۰۰ درصد روغن‌های خالص و ناخالص دارای درصدهای مختلف اختلاط با روغن‌های لایمانت را طبقه‌بندی کند. درصد طبقه‌بندی الگوریتم QDA برای طبقه‌بندی

روغن‌های زیتون خالص و ناخالص با استفاده از ویژگی‌های تصویر برابر با ۹۷/۵ درصد بود.

نتایج نشان داد که ترکیب داده‌های بینی الکترونیک و بینی‌های ماشین باعث افزایش دقت تخمین نرخ ناخالصی در روغن زیتون فرابرک می‌شود. در این حالت، الگوریتم PLS توانست با استفاده از ترکیب داده‌های رنگی و ویژگی AUC استخراج‌شده از بوی نمونه‌ها، میزان ناخالصی ایجاد شده در روغن زیتون فرابرک با روغن، لایمانت را با میزان R^2 برابر با ۰/۹۲۶۶ در مرحله آموزش، ۰/۹۱۸۴ در مرحله اعتبارسنجی، پیش‌بینی کند. ارزیابی عملکرد مدل‌های توسعه یافته با استفاده از داده‌های اولیه بینی الکترونیک و بینی‌های ماشین دارای نويزهای اعمال شده نشان داد که مدل‌های طبقه‌بندی روغن‌های دارای سطوح مختلف انواع تقلب و نیز پیش‌بینی نرخ تقلب دارای استحکام بالایی هستند. بدین ترتیب می‌توان به کمک مجموعه سامانه‌های بینی الکترونیک و بینی‌های ماشین ناخالصی‌های موجود در روغن زیتون را با دقت بالا شناسایی، طبقه‌بندی و پیش‌بینی کرد. با توجه به ماهیت سامانه‌های بینی و بینی‌های الکترونیک، و با در نظر گرفتن دقت‌های به دست آمده، می‌توان از نتایج این پژوهش جهت توسعه یک سامانه دقیق، سریع، ساده، ارزان و غیرمخرب برای تشخیص تقلب در روغن زیتون استفاده کرد.

در انتها لازم به ذکر است که ایجاد سامانه‌ای که بتواند انواع تقلب‌های موجود در روغن زیتون را تشخیص دهد - که هدف نهایی این پژوهش است - در حله اول نیازمند یک پایگاه داده گسترده از انواع تقلب‌ها است. بنابراین در تحقیقات آینده، انواع دیگر تقلب‌های موجود در روغن زیتون با سایر روغن‌ها و مواد اولیه را مورد توجه قرار خواهیم داد. در همین راستا، نرخ‌های بیشتری از هر نوع تقلب نیز مورد بررسی قرار خواهند گرفت تا بتوان الگوریتم‌هایی با قدرت تفکیک و نرخ پیش‌بینی بالاتری را برای استفاده در یک دستگاه تجاری توسعه داد. افزایش اندازه و تنوع پایگاه داده با تقلب‌های مختلف ممکن است باعث شود که الگوریتم مورد استفاده در این پژوهش به نتیجه یکسانی نرسد. در این صورت نیاز به الگوریتم‌های مناسب را با خود به همراه خواهد داشت. بنابراین سایر الگوریتم‌های یادگیری ماشین و هوش مصنوعی مانند شبکه‌های عصبی مصنوعی، ماشین بردار پشتیبان و یادگیری عمیق را مورد بررسی قرار خواهیم داد. همچنین، علاوه بر عملکرد، پژوهش را در راستای بهبود سرعت سامانه‌ها و الگوریتم‌ها ادامه خواهیم داد.

مراجع

- [1] E. Sakar, H. Unver, M. Bakir, M. Ulas, and Z. M. Sakar, "Genetic relationships among olive (*Olea europaea* L.) cultivars native to Turkey," *Biochemical genetics*, vol. 54, pp. 348-359, 2016.
- [2] FAOSTAT, "Crop statistics. Food and Agriculture Organization of the United Nations (FAO)." 2023.

- [17] Q. Li *et al.*, "Machine learning technique combined with data fusion strategies: A tea grade discrimination platform," *Industrial Crops and Products*, vol. 203, p. 117127, 2023.
- [18] M. Guo, K. Wang, H. Lin, L. Wang, L. Cao, and J. Sui, "Spectral data fusion in nondestructive detection of food products: Strategies, recent applications, and future perspectives," *Comprehensive Reviews in Food Science and Food Safety*, vol. 23, no. 1, pp. 1-23, 2024.
- [19] S. Kiani, S. Minaei, and M. Ghasemi-Varnamkhasti, "Application of electronic nose systems for assessing quality of medicinal and aromatic plant products: A review," *Journal of Applied Research on Medicinal and Aromatic Plants*, vol. 3, no. 1, pp. 1-9, 2016.
- [20] U. Khulal, J. Zhao, W. Hu, and Q. Chen, "Nondestructive quantifying total volatile basic nitrogen (TVB-N) content in chicken using hyperspectral imaging (HSI) technique combined with different data dimension reduction algorithms," *Food chemistry*, vol. 197, pp. 1191-1199, 2016.
- [21] W. Wang, J. Xu, H. Fang, Z. Li, and M. Li, "Advances and challenges in medicinal plant breeding," *Plant Science*, vol. 298, p. 110573, 2020.
- [22] N. Karagiorgos *et al.*, "An approach for estimating adulteration of virgin olive oil with soybean oil using image analysis, 2017 6th Int," in *Conf. Mod. Circuits Syst. Technol. MOCAS*, 2017, vol. 2017.
- [23] G. G. Teixeira *et al.*, "Application of a lab-made electronic nose for extra virgin olive oils commercial classification according to the perceived fruitiness intensity," *Talanta*, vol. 226, p. 122122, 2021.
- [24] X. Wei, X. Shao, Y. Wei, L. Cheong, L. Pan, and K. Tu, "Rapid detection of adulterated peony seed oil by electronic nose," *Journal of food science and technology*, vol. 55, pp. 2152-2159, 2018.
- [25] H. Men, D. Chen, X. Zhang, J. Liu, and K. Ning, "Data fusion of electronic nose and electronic tongue for detection of mixed edible-oil," *Journal of Sensors*, vol. 2014, pp. 1-7, 2014.
- [26] M. Bougrini, K. Tahri, Z. Haddi, T. Saidi, N. El Bari, and B. Bouchikhi, "Detection of adulteration in argan oil by using an electronic nose and a voltammetric electronic tongue," *Journal of Sensors*, vol. 2014, 2014.
- [27] A. Marina, Y. Che Man, and I. Amin, "Use of the SAW sensor electronic nose for detecting the adulteration of virgin coconut oil with RBD palm kernel olein," *Journal of the American Oil Chemists' Society*, vol. 87, no. 3, pp. 263-270, 2010.
- [28] X. Yin, "Driven by machine learning to intelligent damage recognition of terminal optical components," *Neural Computing and Applications*, vol. 33, pp. 789-804, 2021.
- [29] T. C. Pearce, S. S. Schiffman, H. T. Nagle, and J. W. Gardner, *Handbook of machine olfaction: electronic nose technology*. John Wiley & Sons, 2006.
- [30] N. Aghilinategh, M. J. Dalvand, and A. Anvar, "Detection of ripeness grades of berries using an electronic nose," *Food Science & Nutrition*, vol. 8, no. 9, pp. 4919-4928, 2020.
- [31] Y. Yin, B. Chu, H. Yu, and Y. Xiao, "A selection method for feature vectors of electronic nose signal based on wilks Λ -statistic," *Journal of Food Measurement and Characterization*, vol. 8, pp. 29-35, 2014.
- [3] A. Bendini *et al.*, "Phenolic molecules in virgin olive oils: a survey of their sensory properties, health effects, antioxidant activity and analytical methods. An overview of the last decade," *Molecules*, vol. 12, no. 8, pp. 1679-1719, 2007.
- [4] A. Bendini, E. Valli, S. Barbieri, and T. G. Toschi, "Sensory analysis of virgin olive oil," *Olive Oil—Constituents, Quality, Health Properties and Bioconversions; InTech: Rijeka, Croatia*, pp. 109-130, 2012.
- [5] M. Roy and B. Yadav, "Electronic nose for detection of food adulteration: A review," *Journal of Food Science and Technology*, pp. 1-13, 2022.
- [6] A. Mohammad-Razdari, M. Ghasemi-Varnamkhasti, S. H. Yoosefian, Z. Izadi, and M. Siadat, "Potential application of electronic nose coupled with chemometric tools for authentication assessment in tomato paste," *Journal of Food Process Engineering*, vol. 42, no. 5, p. e13119, 2019.
- [7] M. Lerma-García, E. Simó-Alfonso, A. Bendini, and L. Cerretani, "Metal oxide semiconductor sensors for monitoring of oxidative status evolution and sensory analysis of virgin olive oils with different phenolic content," *Food Chemistry*, vol. 117, no. 4, pp. 608-614, 2009.
- [8] S. Mildner-Szkudlarz and H. H. Jeleń, "Detection of olive oil adulteration with rapeseed and sunflower oils using mos electronic nose and SMPE-MS," *Journal of food quality*, vol. 33, no. 1, pp. 21-41, 2010.
- [9] M. J. Oates, P. Fox, L. Sanchez-Rodriguez, Á. A. Carbonell-Barrachina, and A. Ruiz-Canales, "DFT based classification of olive oil type using a sinusoidally heated, low cost electronic nose," *Computers and Electronics in Agriculture*, vol. 155, pp. 348-358, 2018.
- [10] C. Li, G. W. Krewer, P. Ji, H. Scherm, and S. J. Kays, "Gas sensor array for blueberry fruit disease detection and classification," *Postharvest Biology and Technology*, vol. 55, no. 3, pp. 144-149, 2010.
- [11] F. Ayari, E. Mirzaee-Ghaleh, H. Rabbani, and K. Heidarbeigi, "Using an E-nose machine for detection the adulteration of margarine in cow ghee," *Journal of Food Process Engineering*, vol. 41, no. 6, p. e12806, 2018.
- [12] K. d. M. Milanez and M. Pontes, "Classification of extra virgin olive oil and verification of adulteration using digital images and discriminant analysis," *Analytical Methods*, vol. 7, no. 20, pp. 8839-8846, 2015.
- [13] P. Kadiroğlu and F. Korel, "Chemometric studies on zNose™ and machine vision technologies for discrimination of commercial extra virgin olive oils," *Journal of the American Oil Chemists' Society*, vol. 92, pp. 1235-1242, 2015.
- [14] T. Ram, Z. Wiesman, I. Parmet, and Y. Edan, "Olive oil content prediction models based on image processing," *Biosystems engineering*, vol. 105, no. 2, pp. 221-232, 2010.
- [15] M. Moyano, A. J. Meléndez-Martínez, J. Alba, and F. J. Heredia, "A comprehensive study on the colour of virgin olive oils and its relationship with their chlorophylls and carotenoids indexes (I): CIEXYZ non-uniform colour space," *Food Research International*, vol. 41, no. 5, pp. 505-512, 2008.
- [16] P. C. Marchal, D. M. Gila, J. G. García, and J. G. Ortega, "Expert system based on computer vision to estimate the content of impurities in olive oil samples,"

- Electronics in Agriculture*, vol. 157, pp. 436-446, 2019.
- [44] A. Gliszczynska-Swiglo and J. Chmielewski, "Electronic nose as a tool for monitoring the authenticity of food. A review," *Food Analytical Methods*, vol. 10, pp. 1800-1816, 2017.
- [45] S. M. Mirhoseini-Moghaddam, M. R. Yamaghani, and A. Bakhshipour, "Application of electronic nose and eye systems for detection of adulteration in olive oil based on chemometrics and optimization approaches," *Journal of Universal Computer Science*, vol. 29, no. 4, p. 300, 2023.
- [46] A. Kaffash, S. R. Kamel, and M. Kheirabadi, "An Effective and Lightweight Intrusion Detection for IoT based on Fog and Cloud using KNN Classification," *Intelligent Multimedia Processing and Communication Systems(IMPCS)*, vol. 5, no. 2, pp. 55-64, 2024.
- [47] S. Shouri, A. Cheshomi, A. Safi, R. Ramezani, and T. E. Salari, "A MLP Neural Network Approach for Validation of Health Insurance Customers in Big Data Environments," *Intelligent multimedia processing and communication systems*, vol. 5, no. 2, pp. 11-21, 2024.
- [48] N. Haghbini, A. Bakhshipour, H. Zareiforush, and S. Mousanejad, "Non-destructive pre-symptomatic detection of gray mold infection in kiwifruit using hyperspectral data and chemometrics," *Plant Methods*, vol. 19, no. 1, p. 53, 2023.
- [49] M. Rasekh and H. Karami, "Application of electronic nose with chemometrics methods to the detection of juices fraud," *Journal of Food Processing and Preservation*, vol. 45, no. 5, p. e15432, 2021.
- [50] A. Sadeghi and H. Hosseini, "Application of Electronic Nose to Detect Pomegranate Paste Adulteration Using Pattern Recognition Methods and Artificial Neural Network," *Innovative Food Technologies*, vol. 10, no. 1, pp. 35-48, 2022.
- [51] H. Jiang, Y. He, and Q. Chen, "Qualitative identification of the edible oil storage period using a homemade portable electronic nose combined with multivariate analysis," *Journal of the Science of Food and Agriculture*, vol. 101, no. 8, pp. 3448-3456, 2021.
- [52] H. Karami, M. Rasekh, and E. Mirzaee-Ghaleh, "Qualitative analysis of edible oil oxidation using an olfactory machine," *Journal of Food Measurement and Characterization*, vol. 14, pp. 2600-2610, 2020.
- [53] S. Buratti, C. Malegori, S. Benedetti, P. Oliveri, and G. Giovanelli, "E-nose, e-tongue and e-eye for edible olive oil characterization and shelf life assessment: A powerful data fusion approach," *Talanta*, vol. 182, pp. 131-141, 2018.
- [54] K. d. M. Milanez and M. Pontes, "Classification of extra virgin olive oil and verification of adulteration using digital images and discriminant analysis," *Analytical Methods*, vol. 7, no. 20, pp. 8839-8846, 2015.
- [55] E. Ordukaya and B. Karlik, "Quality control of olive oils using machine learning and electronic nose," *Journal of Food Quality*, vol. 2017, no. 1, p. 9272404.
- [32] N. Nimsuk, "Improvement of accuracy in beer classification using transient features for electronic nose technology," *Journal of Food Measurement and Characterization*, vol. 13, no. 1, pp. 656-662, 2019.
- [33] A. Sanaeifar, X. Li, Y. He, Z. Huang, and Z. Zhan, "A data fusion approach on confocal Raman microspectroscopy and electronic nose for quantitative evaluation of pesticide residue in tea," *Biosystems Engineering*, vol. 210, pp. 206-222, 2021.
- [34] N. Haghbini, A. Bakhshipour, S. Mousanejad, and H. Zareiforush, "Monitoring Botrytis cinerea infection in kiwifruit using electronic nose and machine learning techniques," *Food and Bioprocess Technology*, vol. 16, no. 4, pp. 749-767, 2023.
- [35] G. Antonov, Y. Podchashynsky, and L. Mohelnyska, "RESEARCHING METHODS OF FILTERING DIGITAL IMAGES THAT CONTAIN MEASURING INFORMATION," 2014.
- [36] A. Amirvaresi, N. Nikouneshad, M. Amirahmadi, B. Daraei, and H. Parastar, "Comparison of near-infrared (NIR) and mid-infrared (MIR) spectroscopy based on chemometrics for saffron authentication and adulteration detection," *Food Chemistry*, vol. 344, p. 128647, 2021.
- [37] Z. Keshavarzi, S. Barzegari Banadkoki, M. Faizi, Y. Zolghadri, and F. H. Shirazi, "Comparison of transmission FTIR and ATR spectra for discrimination between beef and chicken meat and quantification of chicken in beef meat mixture using ATR-FTIR combined with chemometrics," *Journal of food science and technology*, vol. 57, pp. 1430-1438, 2020.
- [38] P. Mishra, A. Biancolillo, J. M. Roger, F. Marini, and D. N. Rutledge, "New data preprocessing trends based on ensemble of multiple preprocessing techniques," *TrAC Trends in Analytical Chemistry*, vol. 132, p. 116045, 2020.
- [39] N. S. Murugan and G. U. Devi, "Feature extraction using LR-PCA hybridization on twitter data and classification accuracy using machine learning algorithms," *Cluster Computing*, vol. 22, no. Suppl 6, pp. 13965-13974, 2019.
- [40] S. Sarkar, J. K. Basak, B. E. Moon, and H. T. Kim, "A comparative study of PLSR and SVM-R with various preprocessing techniques for the quantitative determination of soluble solids content of hardy kiwi fruit by a portable vis/NIR spectrometer," *Foods*, vol. 9, no. 8, p. 1078, 2020.
- [41] I. Muthahharah and A. Juhari, "A Cluster Analysis with Complete Linkage and Ward's Method for Health Service Data in Makassar City," *Jurnal Varian*, vol. 4, no. 2, pp. 109-116, 2021.
- [42] R. Joshi *et al.*, "Evaluation of benzene residue in edible oils using Fourier transform infrared (FTIR) spectroscopy," *Korean Journal of Agricultural Science*, vol. 46, no. 2, pp. 257-271, 2019.
- [43] C. Maione, F. Barbosa Jr, and R. M. Barbosa, "Predicting the botanical and geographical origin of honey with multivariate data analysis and machine learning techniques: A review," *Computers and*

³ Volatile Organic Compound

⁴ Principal Components Analysis (PCA)

⁵ Support Vector Machine (SVM)

⁶ Linear Discriminant Analysis (LDA)

⁷ Partial Least Squares – Discriminant Analysis (PLS-DA)

[^] Moving Window Smoothing (MWS)

-
- ⁹ Maximum Sensor Response (MSR)
¹⁰ Area Under Curve (AUC)
¹¹ Time to Reach MSR (Tim)
¹² Average slope of ascending section (S_{asce})
¹³ Maximum of the instantaneous slopes (S_{max})
¹⁴ Mean Differential Coefficient Value (MDCV)
¹⁵ Salt and pepper
¹⁶ Excessive Red (EXR)
¹⁷ Hue
¹⁸ Saturation
¹⁹ Value

- ²⁰ Luminance
²¹ Multilayer Linear Regression (MLR)
²² Quadratic Discriminant Analysis (QDA)
²³ 10-fold cross validation
²⁴ Accuracy
²⁵ Confusion Matrix
²⁶ Coefficient of determination
²⁷ Root Mean Squared Error (RMSE)