

تشخیص آفاتوکسین پسته با استفاده از تکنیک طیف‌سنجی رامان و شبکه عصبی

رضامحمدی گل^{۱*}، محمد هادی خوش تقاضا^۲، رسول ملک فر^۳، منصوره میرابوالفتحی^۴، علی محمد نیکبخت^۵

۱- دانشجوی دکتری مکانیک بیوسیستم، دانشگاه تربیت مدرس، r.mohammadigol@modares.ac.ir

۲و۳- دانشیار گروه مکانیک بیوسیستم و استاد بخش فیزیک، دانشگاه تربیت مدرس

۳- دانشیار موسسه تحقیقات گیاه پزشکی کشور

۴- استادیار گروه مکانیک بیوسیستم دانشگاه ارومیه

چکیده

آلودگی ناشی از زهرابه قارچی آفاتوکسین به عنوان یک معضل اساسی برای صادرات پسته محسوب می‌شود. با توجه به استقبال روزافزون استفاده از تکنیک طیف‌سنجی رامان در تشخیص و تفکیک مواد مختلف و همچنین مسائل پیش‌روی روش‌های آزمایشگاهی سنجش سم مذکور (مانند هزینه بالا و زمان‌بر بودن)، هدف از این پژوهش بررسی امکان تشخیص و سنجش آفاتوکسین پسته با تکنیک طیف‌سنجی رامان و استفاده از شبکه‌های عصبی بوده است. نمونه‌های مورد تحقیق در ۳ سطح بدون آلودگی (سالم)، آلودگی ۲۰ و ۱۰۰ نانوگرم در گرم (ppb) از مجموع آفاتوکسین‌های (B1+B2+G1+G2) آماده شدند. بعد از طیف برداری، با توجه به نتایج، هنجارسازی داده‌های طیفی به عنوان روش پیش‌پردازش مناسب، انتخاب شد و به دنبال آن برای کاهش ابعاد داده‌های طیفی استخراج مؤلفه‌های اصلی صورت پذیرفت. برای طبقه‌بندی نمونه‌ها، شبکه پرسپترون یک لایه با قانون یادگیری پس انتشار خطا (با ۴ مولفه اصلی موثر به عنوان ورودی و ۳ نرون در لایه پنهان) مورد استفاده قرار گرفت. متوسط دقت طبقه‌بندی شبکه ۹۸ درصد به دست آمد و بنابراین، مدل‌سازی غیرخطی داده‌های طیف رامان توسط شبکه عصبی پرسپترون در طبقه بندی نمونه‌ها موفقیت‌آمیز ارزیابی شد.

واژه‌های کلیدی: آفاتوکسین، آنالیز مؤلفه اصلی، پسته، شبکه عصبی، طیف‌سنجی، رامان

مقدمه

ایران با داشتن حدود ۴۴۰ هزار هکتار سطح زیر کشت باغات پسته در حدود ۵۷ درصد تولید جهانی و متجاوز از ۶۰ درصد صادرات جهانی این محصول مهم و اقتصادی را به خود اختصاص می‌دهد و به عنوان بزرگ‌ترین و مهم‌ترین کشور تولیدکننده و صادرکننده پسته جهان در بین کشورهای (ایران، آمریکا، ترکیه، سوریه، یونان و...) به شمار می‌رود (Moghadam, 2007).

مسئله آلودگی پسته به آفاتوکسین (سمی که برای انسان و حیوانات سرطان‌زاست) یکی از مسائل مهم تحقیقاتی و یکی از معضلات ملی و منطقه‌ای این محصول با ارزش است و همواره کشورهای صادرکننده پسته در راستای رقابت در بازار بین‌المللی از این حربه علیه پسته ایران استفاده نموده‌اند (Sherafati, 2008). آفاتوکسین به عنوان عضو خانواده میکوتوکسین‌ها، از متابولیت-



های ثانویه کپک‌هایی چون *آسپرژیلوس فلاوس*^۱ و *آ. پارازیتیکوس*^۲ است. قارچ‌های مذکور هم روی درخت و هم طی مراحل فرآوری و انبارمانی می‌توانند محصول را آلوده کنند و در صورت آلوده شدن محصول به هاگ آن‌ها و فراهم شدن شرایط دمایی و رطوبتی مطلوب، هاگ رشد کرده و قارچ را تولید می‌کند و قارچ‌ها نیز سم تولید می‌کنند. اگر چه تعداد معدودی از پسته‌ها، آلوده به آفلاتوکسین هستند ولی سرعت انتشار آلودگی زیاد بوده و توده بزرگی از پسته آلوده خواهد شد. (Boutrif, 1998). آفلاتوکسین‌ها به دلیل فراوانی قابل‌ملاحظه آن‌ها در طبیعت و خاصیت سمی و سرطان‌زا بودن، به عنوان سردسته میکوتوکسین‌ها شناخته شده‌اند. تا کنون چندین نوع آفلاتوکسین شناسایی شده است که از بین آن‌ها آفلاتوکسین‌های B1، B2، G1 و G2 دارای بیش‌ترین اهمیت هستند (Airus *et al.*, 2005). بیش‌ترین حد قابل‌قبول آفلاتوکسین B1 پسته در اتحادیه اروپا 12 ppb - میکروگرم بر کیلوگرم) است و این مقدار در کشورهای دیگر ممکن است متفاوت باشد (۲۰-۲ ppb) البته در برخی موارد مجموع آفلاتوکسین‌های B1، B2، G1 و G2 به عنوان محدودیت میکوتوکسین‌های واردات مواد غذایی کشورها لحاظ می‌شود (Anon, 2012).

روش کروماتوگرافی مایع با عملکرد بالا^۴ (HPLC) و ستون‌های ایمونوآفینیتی^۵ به طور گسترده در تشخیص آفلاتوکسین پسته و دیگر محصولات کشاورزی در عرصه‌های تحقیقاتی و تجاری استفاده می‌شود و مقالات متعددی نیز در تأیید دقت روش ذکر شده به چاپ رسیده‌اند (Dragacci *et al.*, 2001; Pearson *et al.*, 1999; Stroka *et al.*, 2000). در این روش مراحل جداسازی و استخراج آفلاتوکسین پیچیده است به علاوه زمان‌بر بودن و نیاز به تجربه زیاد آزمایش‌کننده از خصوصیات دیگر به‌کارگیری این تکنیک عنوان شده است (Sinha, 1999).

طیف‌سنجی (لیزری) رامان، شدت و طول موج نور پراکنده شده لیزری که با یک نمونه در وضعیت گاز، مایع، جامد و یا پلاسما تعامل می‌نماید را اندازه‌گیری می‌کند. به عبارتی دیگر سیگنال رامان از تعامل نور (فوتون‌ها) با نوسانات مولکولی سرچشمه می‌گیرد. در سال‌های اخیر تحقیقات وسیعی در حوزه‌های گوناگون علم پزشکی و داروسازی به منظور به‌کارگیری طیف‌سنجی رامان و به خصوص تکنیک ارتقا سیگنال‌های رامان انجام شده است. عمدتاً تمرکز این تحقیقات روی شناسایی و تشخیص سریع و زود هنگام بافت‌های سرطانی، باکتری‌ها، قارچ‌ها و مواد شیمیایی- دارویی بوده است. در تحقیقات مذکور عموماً از مدل‌های طبقه‌بندی یا رگرسیونی چند متغیره برای نگاشت و تفسیر اطلاعات طیفی استفاده شده است (Huang *et al.*, 2010; Teh *et al.*, 2008; 2010).

¹ *Aspergillus flavus*

² *A. parasiticus*

³ Particle per billion

⁴ High-Performance Liquid Chromatography

⁵ Immunoaffinity



مهم‌ترین مزیت استفاده از روش طیف‌سنجی رامان، در تحلیل زیستی^۱ غنای اطلاعات موجود در هر طیف است (DasAgrawal, 2011). طیف رامان یک سلول را می‌توان به عنوان « اثر انگشت» ترکیبات بیوشیمیایی آن به حساب آورد. بنابراین اگر یک عامل سمی باعث تغییرات بیوشیمیایی شود طیف رامان حاصله دست‌خوش تغییر خواهد شد (Notingher, 2007). رویکرد رو به رشد در اغلب محصولات کشاورزی، افزایش استفاده از روش‌های سریع، قابل اطمینان و غیر مخرب ارزیابی کیفی است. با توجه به جایگاه ممتاز پسته در صادرات غیر نفتی ایران و به منظور جلوگیری از به مخاطره افتادن صادرات این محصول با ارزش و همچنین برای حضور در تجارت جهانی و رقابت با دیگر کشورها، اولویت تحقیقات بر روی روش‌های نوین تشخیص و سنجش آلودگی پسته به آفلاتوکسین انکار ناپذیر بوده و جایگاه ویژه‌ای دارد. پتانسیل بالای تشخیصی طیف‌سنجی رامان و پیشرفت‌های صورت گرفته در عرصه فناوری ساخت تجهیزات اپتیکی و نانو باعث توسعه و ارتقاء چشم‌گیر طیف‌سنج‌های مزبور شده است (تا آنجا که انواع قابل انتقال آن‌ها به بازار آمده است) از طرف دیگر، افزایش مطالبه جهانی برای استفاده از روش‌های سریع و دقیق تشخیص آلودگی در مواد غذایی، زمینه‌ساز انجام تحقیق حاضر با هدف امکان سنجی تشخیص و سنجش کیفی آفلاتوکسین پسته با تکنیک طیف‌سنجی رامان و استفاده از شبکه عصبی بوده است.

مواد و روش‌ها

سامانه طیف برداری



شکل ۱- سامانه طیف‌سنج رامان مورد استفاده

سامانه طیف‌سنج رامان مدل (2009) SENTERRA از نوع پاشنده ساخت شرکت BRUKER برای به دست آوردن طیف‌های رامان مورد استفاده قرار گرفت (شکل ۱). این سامانه برداشت طیف رامان از مایعات و جامدات را فراهم کرده و سیستم کاوش میکروسکوپی آن امکان بررسی میکروسکوپی مواد و بلورها را میسر می‌کند. لیزر دستگاه طول موج ۷۸۵ نانومتر بوده و حداکثر توان لیزر ۱۰۰ میلی وات است

¹ Bio analysis



استفاده از لیزر با طول موج ۷۸۵ نانومتر به عنوان منبع تهییج در طی فرآیند طیف برداری رامان علاوه بر داشتن حساسیت مناسب به طور معنی‌داری سبب کاهش پدیده فلورسانس در طیف حاصله می‌گردد. البته در هر صورت وجود مقداری فلورسانس در طیف رامان اجتناب‌ناپذیر است که معمولاً توسط قابلیت نرم‌افزاری قابل حذف است. در تحقیقات زیادی در رابطه با طیف‌سنجی رامان نمونه‌های زیستی از لیزر با طول موج ۷۸۵ نانومتر به عنوان منبع تهییج استفاده شده است (Dennis, 2013a; 2013b).

تعیین توان بهینه لیزر

توان سامانه طیف‌سنجی مورد استفاده از ۱۰ میلی وات تا ۱۰۰ میلی وات قابل تنظیم است. نظر به اهمیت کیفیت سیگنال‌های طیف رامان در تجزیه و تحلیل آن‌ها، با انجام چند پیش‌آزمایش مشخص گردید که توان ۱۰۰ میلی وات دارای نسبت سیگنال به نوفه بالاتر بوده و لذا از این توان لیزر برای طیف برداری استفاده شد.

نمونه‌سازی

رقم پسته مورد استفاده، رقم فندقی بوده است. کشت گسترده، امکان آلودگی بالا (به علت خندانی بیشتر نسبت به رقم‌های دیگر) از دلایل انتخاب این رقم بودند (Tajabadipour, 2008). نمونه‌های پسته به صورت خمیری (Stroka et al., 2000) در ۳ سطح: c- بدون آلودگی (۱۵ طیف)، a- آلودگی ۲۰ ppb (B1+B2+G1+G2) هر کدام ۵ ppb (۱۷ طیف) و A- آلودگی ۱۰۰ ppb (B1+B2+G1+G2) هر کدام ۲۵ ppb (۱۸ طیف) آماده شدند. به منظور تهیه غلظت‌های مورد نظر از قانون بیر-لامبرت^۱ (رابطه ۱) استفاده شد. این قانون تجربی ارتباط شدت نور جذب‌شده در اثر عبور از ماده همگن بدون پراکندگی را با خصوصیات مواد بیان می‌کند (ParnisOldham, 2013).

$$\text{Log}(I_0 I^{-1}) = A \quad (1)$$

که در آن، I_0 شدت نور اولیه، I شدت نور عبوری و A مقدار جذب ماده (بدون واحد) است. که به صورت (رابطه ۲) تعریف می‌شود:

$$A = abc \quad (2)$$

که در آن، a ضریب جذب ماده ($L \text{ mol}^{-1} \text{ cm}^{-1}$)، b ضخامت (cm) و c غلظت (mol L^{-1}) آن است. غلظت‌های فوق‌الذکر در آزمایشگاه میکروتوکسین‌های موسسه تحقیقات گیاه‌پزشکی کشور به این روش تهیه شدند. از آنجا که آفاتوکسین‌ها جزو سموم خطرناک هستند در کلیه مراحل تحقیق سعی شد ملاحظات ایمنی آزمایشگاهی رعایت شود همچنین برای پاک‌سازی، ظروف و ابزارهای استفاده‌شده در محلول هیپو کلریت سدیم قرار داده شدند.

¹ Beer-Lambert



پیش‌پردازش طیف‌ها^۱

طیف‌های رامان برداشت شده علاوه بر اطلاعات نمونه، حامل اطلاعات ناخواسته از قبیل نوفه و نور پس زمینه هستند. لذا برای رسیدن به مدل واسنجی دقیق و قابل‌اعتماد پیش‌پردازش داده‌های طیفی قبل از مدل‌سازی ضرورت دارد. روش‌های پیش‌پردازش بسیار متنوع هستند که معمولاً هر یک از آن‌ها برای منظور خاصی ابداع شده‌اند و با توجه به شرایط متفاوت نمونه‌ها، روش‌ها و ابزار آزمایشگاهی ارائه یک روش پیش‌پردازش مشخص امکان ندارد. از این رو معمولاً در تحقیقات طیف‌سنجی روش‌های رایج برتر مورد ارزیابی قرار گرفته و بهترین روش مورد استفاده قرار می‌گیرد (Jamshidi, 2012). در این تحقیق تأثیر روش‌های هموارسازی^۲، هنجار سازی^۳، مشتق اول و مشتق دوم، Detrending، توزیع نرمال استاندارد (SNV) و تصحیح پراکنش افزایشده (MSC) به صورت انفرادی یا ترکیبی (هموارسازی + هنجار سازی، هموارسازی + تصحیح پراکنش افزایشده، هموارسازی + توزیع نرمال استاندارد، مشتق اول + هموارسازی + هنجار سازی) در تفکیک و تمایز طیف‌های اخذ شده مورد بررسی قرار گرفتند (IshikawaGulick, 2013).

آنالیز مؤلفه‌های اصلی (PCA)^۴

به طور کلی تعداد مشاهدات (طیف‌ها) برای آموزش موفق یک طبقه بند خودآموز ماشینی^۵، با تعداد داده‌ها (متغیرهای طیفی) توسط رابطه‌ی نمایی افزایش می‌یابد (IshikawaGulick, 2013)a. لذا در اغلب تحقیقات مرتبط با طیف‌سنجی به دلیل اینکه تعداد متغیرهای طیفی بسیار زیاد است، قبل از فرآیند مدل‌سازی، استخراج مؤلفه‌های اصلی (توسط الگوریتم کاهش داده) با استفاده از روش PCA انجام شده است (Sigurdsson et al., 2004; Wu et al., 2007). تحلیل مؤلفه‌های اصلی یکی از انواع روش‌های تحلیل داده‌های چند متغیره است که هدف اصلی آن تقلیل بعد مسئله مورد مطالعه است. با استفاده از تحلیل مؤلفه‌های اصلی می‌توان تعداد زیادی متغیر توضیحی (متغیر مستقل) را با تعداد محدودی متغیر توضیحی جدید که با یکدیگر وابستگی خطی ندارند و مؤلفه‌های اصلی نامیده می‌شوند، جایگزین نمود (Nicolai et al., 2007). از آنجا که تمرکز روش یادگیری شبکه‌های عصبی مصنوعی بر معادل‌سازی متغیر وابسته با ورودی‌های شبکه بدون اعمال محدودیت بر پارامترهای مختلف (از جمله وزن‌ها) است بنابراین توجیهاتی وجود دارد که شبکه عصبی در مقایسه با روش‌های PLS و PCR نسبت به Overfitting (وراً موزی) حساس تر است. به عنوان یک مصالحه پیشنهاد می‌شود که از مؤلفه‌های اصلی (PCs) به جای متغیرهای اصلی به عنوان ورودی

¹ Preprocessing

² Smoothing

³ Normalization

⁴ Principal Components Analysis

⁵ Machine-Learning Classifier



شبکه استفاده شود (Næs *et al.*, 1993). با توجه به حجم بالای متغیرهای طیف رامان (عدد موجها) از این روش برای کاهش متغیرها استفاده شد.

مدل‌سازی

یک شبکه عصبی این قابلیت را دارد که تمایزها و تفاوت‌های مرموز را که در شکل کلی و وضعیت نوارهای مختلف رامان وجود دارد و به راحتی توسط بشر درک نمی‌شوند، را فهمیده و بر مبنای آن آموزش ببیند (IshikawaGulick, 2013). در بسیاری از مسائل علمی و پیچیده داده کاوی و شناسایی الگو، از شبکه عصبی پرسپترون برای جداسازی کلاس‌های مختلف بر مبنای الگوهای آموزشی استفاده می‌شود. در این موارد، معیار مناسب بودن طبقه‌بندی کننده عصبی، تجربی بوده و در صورتی که جواب خوبی به دست نیاید، سعی می‌شود با تغییر پارامترهای شبکه عصبی (تعداد نرون و لایه) و استفاده از روش سعی و خطا، جواب مطلوب گرفته شود (MoallemMonadjemi, 2007) در این تحقیق از شبکه پرسپترون یک لایه و قانون یادگیری پس انتشار خطا برای طبقه‌بندی طیف نمونه‌ها استفاده شد. به منظور بهبود عملکرد شبکه داده‌های ورودی در بازه (۱, -۱) هنجار سازی شدند (Gordon *et al.*, 1998).

به منظور انجام مدل‌سازی نرم‌افزارهای متعددی وجود دارد. بسته نرم‌افزاری Unscrambler با سهولت کارکرد و امکان انتخاب عملیات از لیست^۱ و همچنین قابلیت انجام پیش‌پردازش روی داده‌های طیفی از آن جمله است. (Nicolai *et al.*, 2007). در این تحقیق نرم‌افزارهای IBM Modeller و Unscrambler برای مدل‌سازی شبکه عصبی و تجزیه و تحلیل داده‌های طیفی مورد استفاده قرار گرفتند.

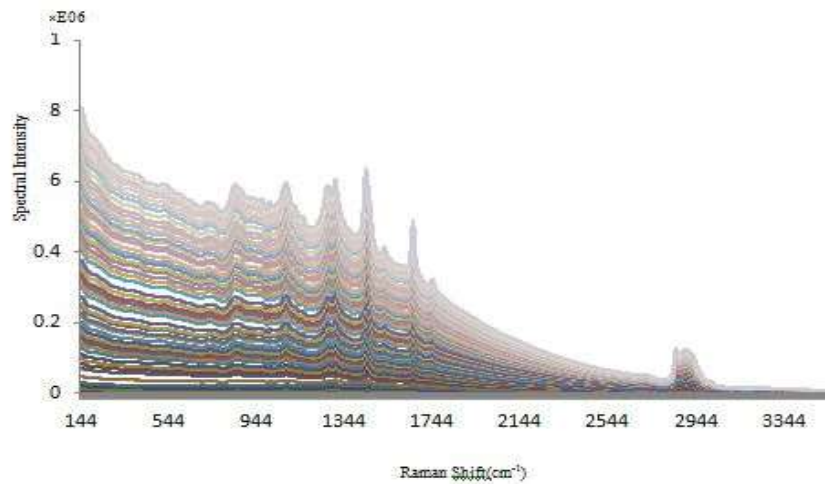
نتایج و بحث

کاوش در طیف‌ها

پس از بررسی طیف‌ها استنباط شد که در نوار $2000 - 848 \text{ cm}^{-1}$ شدت^۲ پراکندگی فوتون‌های نمونه‌های گروه A بیشتر از گروه a و آن هم بیشتر از گروه c است (شکل ۲). بنابراین مطلب و همچنین مقایسه آنالیز مؤلفه‌های اصلی کل محدوده طیف، نوار مذکور در مراحل بعدی مورد تحلیل و بررسی قرار گرفت.

¹ Menu- Driven

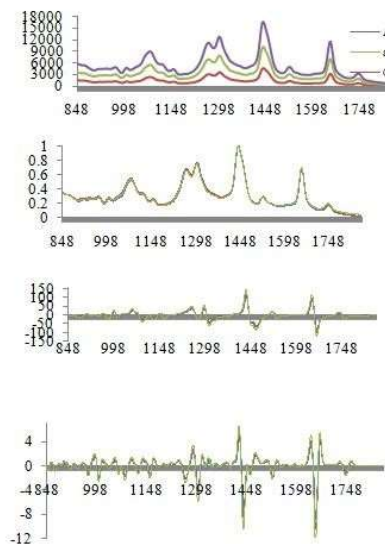
² Intensity



شکل ۲- طیف رامان کامل نمونه‌ها در بازه $144 - 3500 \text{ cm}^{-1}$

ارزیابی روش‌های پیش‌پردازش

در شکل ۳ اثر روش‌های پیش‌پردازش حذف خط زمینه و هموارسازی، هنجار سازی، مشتق اول و مشتق دوم روی طیف‌های میانگین هر گروه نمونه مشاهده می‌شود. در شکل ۳ صرفاً نحوه عملکرد این روش‌ها نشان داده شده است و روش‌های مذکور روی طیف همه نمونه‌ها اعمال شده‌اند.

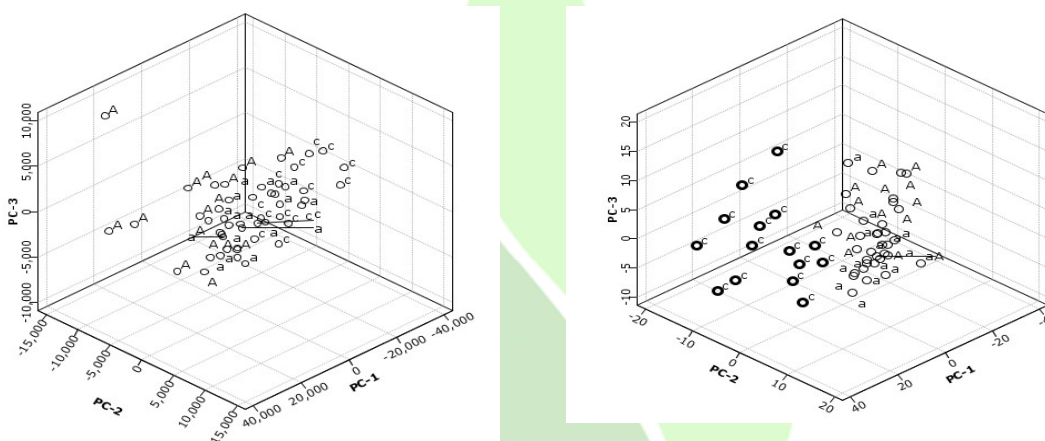


شکل ۳- عملکرد برخی از روش‌های پیش‌پردازش اعمال شده روی میانگین طیف‌های هر گروه نمونه در نوار $848 - 1748 \text{ cm}^{-1}$ ، به ترتیب از بالا به پایین: حذف خط زمینه و هموارسازی، هنجار سازی، اعمال مشتق اول، اعمال مشتق دوم.

آنالیز مؤلفه‌های اصلی PCA



محاسبه مؤلفه‌های اصلی از طیف‌های رامان و به کارگیری آن‌ها، در تحقیقات گوناگون مشاهده می‌شود (Sigurdsson *et al.*, 2004). در شکل ۴ پراکندگی طیف‌ها در مختصات مؤلفه‌های اصلی اول، دوم و سوم قبل و بعد از اعمال مشتق دوم روی داده های طیفی مشاهده می‌شود، همان‌گونه که مشهود است پیش پردازش طیف‌ها می‌تواند در واریانس بین کلاس‌ها و واریانس داده‌ها درون هر کلاس (با توجه به وضعیت قرارگیری طیف‌ها) اثر گذار باشد. در این تحقیق مولفه‌های اصلی از نظر اهمیت در تعیین کلاس طیف نمونه‌ها توسط نرم افزار اولویت بندی شده و سپس به ترتیب اولویت اثرگذاری به عنوان ورودی شبکه لحاظ شدند این شیوه باعث بالا رفتن دقت به خصوص به ازای تعداد کمتر ورودی شبکه شده و توسط محققان دیگر نیز استفاده شده است (Feng *et al.*, 2010).



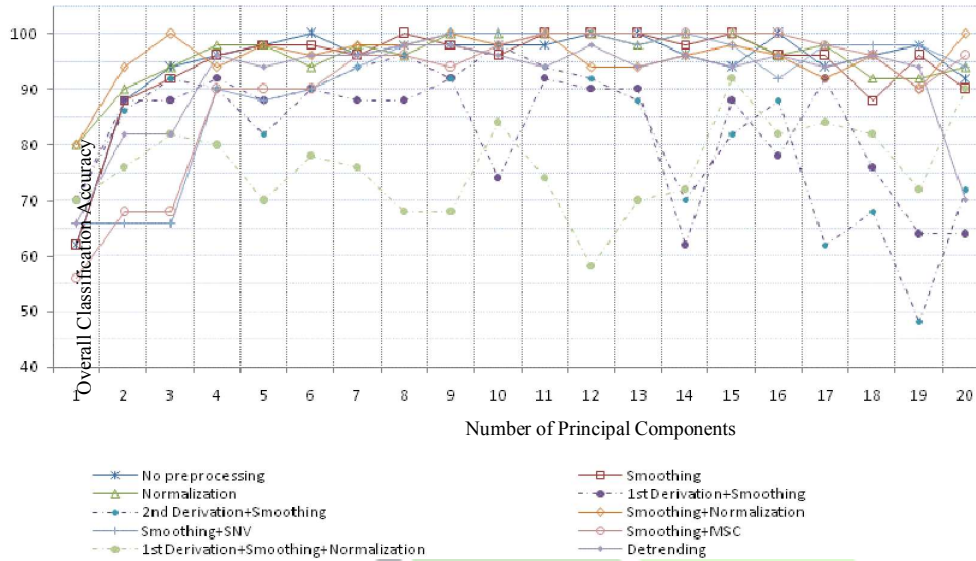
شکل ۴- اثر پیش پردازش مشتق دوم بر پراکندگی طیف‌ها در مختصات مؤلفه اصلی اول، دوم و سوم.

تعیین تعداد بهینه ورودی‌های شبکه

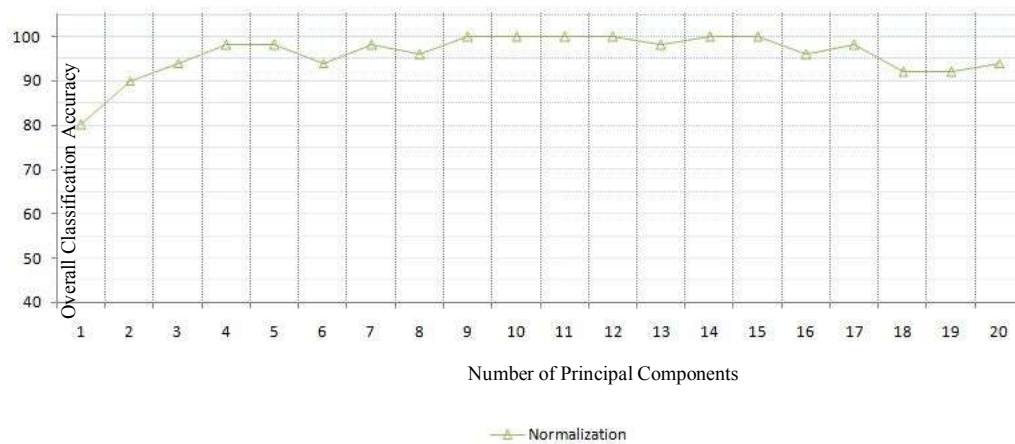
همان‌گونه که قبلاً گفته شد ابعاد داده‌های طیفی به مؤلفه‌های اصلی کاهش داده شدند. به منظور به دست آوردن تعداد بهینه مؤلفه‌های اصلی (و موثر) ورودی شبکه، میانگین دقت طبقه‌بندی شبکه در مقابل تعداد مؤلفه‌های اصلی (۲۰ عدد) به ازای پیش پردازش‌های مختلف ترسیم شد. در شکل ۵ پیش پردازش‌هایی که شامل مشتق گیری (اول یا دوم) هستند بر خلاف روش‌های دیگر، علی‌رغم برتری در تفکیک طیف‌ها در مختصات مولفه‌های اصلی (شکل ۴) دارای روند بسیار ناپایدار بوده و بیشینه دقت متوسط طبقه بندی حاصل از آنها نسبت به پیش پردازش‌های دیگر کمتر است و به طور کلی در این پژوهش مطلوب ظاهر نشده‌اند، نتیجه تحقیق (IshikawaGulick, 2013) نیز این مطلب را تایید می‌نماید. با توجه به شکل ۵ تا تقریباً ۴ مؤلفه به عنوان ورودی، دقت متوسط طبقه بندی در غالب پیش پردازش‌های بررسی شده افزایش و سپس با روند نسبتاً پایداری در بازه بین ۹۰ و ۱۰۰ درصد ادامه می‌یابد. با توجه به تعداد کل طیف‌ها (۵۰ عدد)، تعداد نرون‌های خروجی (۳ کلاس) و تعداد نرون بهینه لایه



پنهان محاسبه شده توسط نرم افزار (۳ عدد)، در نظر گرفتن تعداد ورودی کم از نظر مبانی مدل سازی شبکه عصبی، منطقی تر می باشد، از این رو تعداد ۴ مؤلفه اصلی به عنوان ورودی شبکه و پیش پردازش هنجار سازی (شکل ۶) برای اعمال روی داده های طیفی در نوار $2000 - 848 \text{ cm}^{-1}$ انتخاب شدند.



شکل ۵- دقت متوسط طبقه‌بندی شبکه به ازای تعداد ۲۰ مؤلفه‌های اصلی اول با اعمال پیش پردازش های مختلف



شکل ۶- دقت متوسط طبقه‌بندی شبکه به ازای تعداد ۲۰ مؤلفه‌های اصلی اول با اعمال پیش پردازش هنجار سازی خلاصه عملکرد شبکه عصبی

ماتریس اغتشاش با فراهم نمودن اینکه یک آیتم چقدر به اشتباه طبقه‌بندی شده، دقت الگوریتم‌های طبقه‌بندی را نمایش می‌دهد و خلاصه عملکرد طبقه‌بندی با شبکه‌های عصبی در قالب این ماتریس ارائه می‌شود (Hagan et al., 1996). در شکل ۷

ماتریس اغتشاش عملکرد شبکه عصبی مورد استفاده به ازای اعمال پیش پردازش هنجارسازی روی داده های طیفی و ۴ ورودی (مولفه های اصلی موثرتر) در طبقه بندی طیف نمونه‌ها مشاهده می‌شود.

Classification for class
Overall Percent Correct = 98.0%

Observed	Predicted			Row Percent
	A	a	c	
A	18	0	0	100.00
a	0	17	0	100.00
c	0	1	14	14.29

شکل ۷- ماتریس اغتشاش خلاصه عملکرد شبکه عصبی در طبقه بندی نمونه‌ها

جمع بندی

نوار $2000-848 \text{ cm}^{-1}$ از کل بازه طیفی به دلیل دارا بودن تمایز بین نمونه‌ها مورد تحلیل و بررسی قرار گرفت. روش‌های پیش‌پردازش مورد استفاده در تحلیل داده‌های طیفی اثر داشته و در بین ۱۰ روش پیش‌پردازش بررسی شده، اعمال روش هنجارسازی به علت عملکرد بهتر (از لحاظ دقت متوسط طبقه بندی) و پایداری انتخاب شد. در بین مولفه های اصلی با انتخاب موثر ترین آنها در تعیین کلاس نمونه ها و همچنین اعمال پیش پردازش هنجارسازی، میانگین دقت طبقه بندی شبکه مورد استفاده (با ۴ مولفه اصلی به عنوان ورودی) ۹۸ درصد به دست آمد. بنابراین مدل سازی غیرخطی داده‌های طیف رامان توسط شبکه پرسپترون با الگوی پس انتشار خطا در جداسازی نمونه‌ها موفقیت آمیز ارزیابی شد.

منابع

- 1- Anon. 2012. Mycotoxins Legislation Worldwide European Mycotoxins Awareness Network. <http://services.leatherheadfood.com/eman/FactSheet.aspx?ID=79> (access on September 29, 2013).
- 2- Arrus, K., G. Blank, D. Abramson, R. Clear and R. Holley. 2005. Aflatoxin production by *Aspergillus flavus* in Brazil nuts. *Journal of Stored Products Research* 41: 513-527.
- 3- Boutrif, E. 1998. Prevention of aflatoxin in pistachios. FAO Corporate Document Repository. Food, Nutrition and Agriculture series. Report no. 21.
- 4- Das, R. S. and Y. Agrawal. 2011. Raman spectroscopy: Recent advancements, techniques and applications. *Vibrational Spectroscopy* 57: 163-176.

- 5- Dennis, A. 2013a. Photo-Bleaching and Automatic Baseline Correction for Raman Spectroscopy. www.perkinelmer.com (access on September 29, 2013).
- 6- Dennis, A. 2013b. Automatic Baseline Correction PerkinElmer company. Thechnical note. Available on: www.perkinelmer.com (access on September 29, 2013).
- 7- Dragacci, S. and F. Grosso and J. Gilbert. 2001. Immunoaffinity column cleanup with liquid chromatography for determination of aflatoxin M1 in liquid milk: collaborative study. Journal of AOAC International 84: 437-443.
- 8- Feng, S., R. Chen, J. Lin, J. Pan, G. Chen, Y. Li, M. Cheng, Z. Huang, J. Chen and H. Zeng. 2010. Nasopharyngeal cancer detection based on blood plasma surface-enhanced Raman spectroscopy and multivariate analysis. Biosensors and Bioelectronics 25: 2414-2419.
- 9- Gordon, S. H., B. C. Wheeler, R. B. Schudy, D. T. Wicklow and R. V. Greene. 1998. Neural network pattern recognition of photoacoustic FTIR spectra and knowledge-based techniques for detection of mycotoxigenic fungi in food grains. Journal of Food Protection® 61: 221-230.
- 10- Hagan, M. T. and H. B. Demuth and M. H. Beale. 1996. Neural network design. Pws Pub. Boston.
- 11- Huang, Z., S. K. Teh, W. Zheng, K. Lin, K. Y. Ho, M. Teh and K. G. Yeoh. 2010. In vivo detection of epithelial neoplasia in the stomach using image-guided Raman endoscopy. Biosensors and Bioelectronics 26: 383-389.
- 12- Ishikawa, S. and V. Gulick. 2013. An automated mineral classifier using Raman spectra. Computers & Geosciences 54: 259-268.
- 13- Jamshidi, B. 2012. Non-destructive taste characterization and classification of oranges using Vis/NIR spectroscopy. TMU (in Farsi).
- 14- Moallem, P. and A. Monadjemi. 2007. A Heuristic Criterion for Goodness of Multi Layer Perceptron as a Classifier. in IDMC'07. Tehran, Iran (in Farsi).
- 15- Moghadam, M. M. 2007. Study of pistachio contamination in processing terminals of Semnan province and evaluation of resistance of pistachio cultivars to Aspergillus flavus and aflatoxin B1. IPRI, AREO (in Farsi).
- 16- Næs, T., K. Kvaal, T. Isaksson and C. Miller. 1993. Artificial neural networks in multivariate calibration. Journal of Near Infrared Spectroscopy 1: 1.
- 17- Nicolai, B. M., K. Beullens, E. Bobelyn, A. Peirs, W. Saeys, K. I. Theron and J. Lammertyn. 2007. Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: A review. Postharvest Biology and Technology 46: 99-118.
- 18- Notingher, I. 2007. Raman spectroscopy cell-based biosensors. Sensors 7: 1343-1358.
- 19- Parnis, J. M. and K. B. Oldham. 2013. Beyond the Beer-Lambert law: The dependence of absorbance on time in photochemistry. Journal of Photochemistry and Photobiology A: Chemistry 267: 6-10.
- 20- Pearson, S., A. Candlish, K. Aidoo and J. Smith. 1999. Determination of aflatoxin levels in pistachio and cashew nuts using immunoaffinity column clean-up with HPLC and fluorescence detection. Biotechnology techniques 13: 97-99.
- 21- Sherafati, A. 2008. Pistachio application (challenges and solutions). Sarva Publication (in Farsi).
- 22- Sigurdsson, S., P. A. Philipsen, L. K. Hansen, J. Larsen, M. Gniadecka and H.-C. Wulf. 2004. Detection of skin cancer by classification of Raman spectra. Biomedical Engineering, IEEE Transactions on 51: 1784-1793.
- 23- Sinha, K. K. 1999. Testing methods for aflatoxins in foods. Food & Nutrition Bulletin 20: 458-464.



- 24- Stroka, J., E. Anklam, U. Jorissen and J. Gilbert. 2000. Immunoaffinity column cleanup with liquid chromatography using post-column bromination for determination of aflatoxins in peanut butter, pistachio paste, fig paste, and paprika powder: collaborative study. *Journal of AOAC International* 83: 320-340.
- 25- Tajabadipour, A. 2008. Evaluation of the effectiveness of different devices in the process of separation phase in the garden of nuts to aflatoxin contamination suspected: .IPRI, AREO (in Farsi).
- 26- Teh, S. K., W. Zheng, K. Y. Ho, M. Teh, K. G. Yeoh and Z. Huang. 2008. Diagnosis of gastric cancer using near-infrared Raman spectroscopy and classification and regression tree techniques. *Journal of biomedical optics* 13: 034013-034013-034018.
- 27- Teh, S. K., W. Zheng, K. Y. Ho, M. Teh, K. G. Yeoh and Z. Huang. 2010. Near-infrared Raman spectroscopy for optical diagnosis in the stomach: Identification of *Helicobacter-pylori* infection and intestinal metaplasia. *International Journal of Cancer* 126: 1920-1927.
- 28- Wu, D. and S. Feng and Y. He. 2007. Infrared spectroscopy technique for the nondestructive measurement of fat content in milk powder. *Journal of dairy science* 90: 3613-3619.



Detection of Pistachio Aflatoxin Using Raman Spectroscopy and Artificial Neural Netw

Reza Mohammadigol^{1*} Mohammad Hadi Khoshtaghaza² Rasoul Malekfar³ Mansoureh Mirabolfathi⁴ and Ali Mohammad Nikbakht⁵

1- PhD Student, Department of Biosystem Mechanics Engineering, TMU university, Tehran
r.mohammadigol@modares.ac.ir

2- Associate Professor, Department of Biosystem Mechanics Engineering, TMU university, Tehran

3- Professor, Department of Physics, TMU university, Tehran

4- Associate Professor, Iranian Research institute of plant protection, Tehran

5- Professor Assistant, Department of Biosystem Mechanics Engineering, Uromia university, Uromia

Abstract

Pistachio contamination to aflatoxin has been known as a serious problem for pistachio exportation. Regarding to increasing demand for Raman spectroscopy to detect and classify of different materials and also the current experimental technique problems for measuring toxin (like cost and time consuming), the main objective of this study was to detect aflatoxin contamination in pistachio by using Raman spectroscopy technique and artificial neural networks. Three levels of samples were prepared, non-contaminated (healthy) and contaminated samples with 20 and 100 ppb of the total aflatoxins (B1+B2+G1+G2). After spectrums acquisition, considering to principal components analysis (PCA) results, Normalization preprocessing method was selected and then principal components (PCs) were extracted to reduce the data dimensions. For classification of the samples spectrum, one layer perseptron feed forward back propagation artificial neural networks (4 inputs and 3 neurons on hidden layer) was used. Mean overall accuracy was achieved 98 percent, therefore non-linear Raman spectra data modeling by ANN for samples classification was successful.

Keywords: Aflatoxin, ANN, PCA, Pistachio, Raman Spectroscopy