

طبقه بندی زعفران با استفاده از ویژگی های رنگی استخراج شده از تصویر

مرتضی محمدزاده مقدم^۱، مسعود تقی زاده^{۲*}، حسن صدرنیا^۳ و حمیدرضا پوررضا^۴

تاریخ پذیرش: ۲۱ اردیبهشت ۱۳۹۹

تاریخ دریافت: ۱ آبان ۱۳۹۸

محمدزاده مقدم، م.، تقی زاده، م.، صدرنیا، ح.، و پوررضا، ح. ر. ۱۳۹۹. طبقه بندی زعفران با استفاده از ویژگی های رنگی استخراج شده از تصویر. زراعت و فناوری زعفران، ۸(۳): ۳۱۹-۳۹۹.

چکیده

طبقه بندی زعفران به عنوان گران ترین ادویه از اهمیت بالایی برای مشتریان و تجار برخوردار است. به طور کلی، در حال حاضر دو روش برای درجه بندی زعفران استفاده می شود. روش اول براساس تجربیات فرد خبره و با مشاهده نمونه ها انجام می شود. روش دوم تخریبی بوده و با استفاده از متدهای آزمایشگاهی انجام می گیرد. طبق نظر متخصصان، استفاده از تکنیک های یادگیری ماشین برای طبقه بندی زعفران به دلیل داشتن ماهیت غیر مخرب و خصوصیات بهنگام، یک هدف است. این روش همچنین می تواند باعث افزایش دقت فرآیند درجه بندی در مقیاس صنعتی شود. در این مقاله، یک روش مبتنی بر ماشین بینایی ارائه شده است. با توجه به عدم تحقیقات مستند در مورد این موضوع، جستجوی مشروح جامع در این کار ارائه می شود. تقریباً تمام ویژگی های رنگ استخراج و در تعداد زیادی از طبقه بندی کننده ها استفاده شد. افراد خبره در ایران زعفران را بر اساس خصوصیات ظاهری به سه طبقه اصلی یعنی پوشال، نگین و سرگل طبقه بندی می کنند. در این مقاله، یک بانک اطلاعاتی متشکل از ۴۴۰ تصویر از زعفران برای سه کلاس مختلف با استفاده از دوربین تلفن همراه جمع آوری شد. پس از اعمال تعدادی از مراحل پیش پردازش مانند حذف پس زمینه، بریدن و حذف مناطق ناخواسته تصاویر و غیره، ۲۱ ویژگی رنگی با استفاده از روش های مختلف تحلیل تصویر استخراج شد. برای طبقه بندی از ۲۲ طبقه بندیگر استفاده شدند. مقایسه نتایج طبقه بندی کننده های مختلف نشان داد که Linear Discriminant، Linear SVM، Bagged Trees و RUSBoost Trees می توانند در هنگام استفاده از ویژگی های رنگی، درجه بندی دقیق تری را نسبت به سایر طبقه بندی کننده ها ایجاد کنند. به طور خاص، در این کار، میانگین دقت ۸۲/۲۳ درصد با استفاده از طبقه بندی کننده خطی SVM بدست آمد.

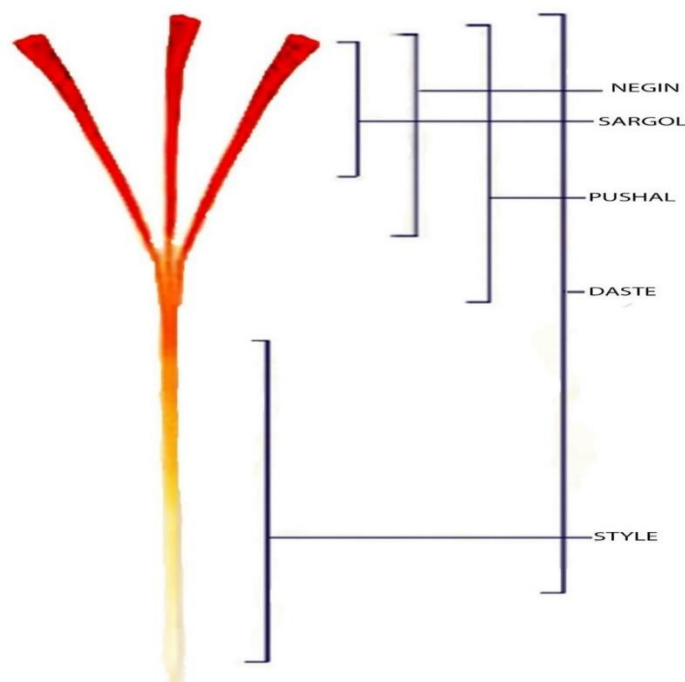
کلمات کلیدی: استخراج رنگ، سرگل، درجه بندی، پردازش تصویر.

۱- دانشجوی دکتری، مهندسی علوم و صنایع غذایی، گروه علوم و صنایع غذایی، دانشکده کشاورزی، دانشگاه فردوسی مشهد
۲- استادیار، گروه علوم و صنایع غذایی، دانشکده کشاورزی، دانشگاه فردوسی مشهد
۳- دانشیار، گروه مکانیک بیوسیستم، دانشکده کشاورزی، دانشگاه فردوسی مشهد
۴- استاد، گروه کامپیوتر، دانشکده مهندسی، دانشگاه فردوسی مشهد
(* نویسنده مسئول: mtaghizadeh@um.ac.ir)

مقدمه

(Shahdadi et al., 2016). یک نوع دیگر زعفران در ایران وجود دارد که به دسته یا دختر پیچ مشهور است و کمتر در بازار تجاری وجود دارد (Bonyadi et al., 2014). برای تهیه زعفران پوشال، کلاله‌های قرمز رنگ همراه مقدار کمی خامه‌ی سفید رنگ جدا شده و سپس خشک می‌شوند. زعفران نگین کاملاً قرمز بوده و از همان ابتدا در هنگام پرکنی گلهای زعفران، کلاله‌های قرمز بصورت تکی جدا می‌شوند و خشک می‌شوند. تفاوت نگین و سرگل در این است که سرگل حاوی رشته‌های کلاله شکسته شده است ولی در نگین عمده رشته‌های کلاله سالم هستند (Atefi et al., 2013) و (Kafi et al., 2006). زعفران سرگل فقط از قسمت قرمز کلاله تشکیل شده است. این نوع از زعفران قدرت رنگی قوی دارد (Azarabadi & Özdemir, 2018). این تعاریف در شکل ۱ نشان داده شده است.

زعفران گرانترین ادویه در جهان است. این محصول در کشورهای مختلف جهان از جمله ایران، هند، اسپانیا، یونان، ایتالیا و مراکش کشت می‌شود (Fernández, 2004). ایران بزرگترین تولیدکننده زعفران در جهان است. هم‌اکنون ۹۴٪ زعفران جهان در ایران تولید می‌شود (Masi et al., 2016). درجه‌بندی کیفیت زعفران همواره خواسته مشتریان و همچنین تجار است. به طور سنتی، طبقه‌بندی زعفران توسط کارشناسان بر اساس پارامترهای کیفی ظاهری انجام می‌شود (Kiani & Minaei, 2016). در کشورهای مختلفی مانند ایران، اسپانیا و هند، زعفران خالص با توجه به خصوصیات ظاهری زعفران درجه‌بندی می‌شود. در بازار محلی ایران زعفران به سه نوع تقسیم می‌شود: سرگل، نگین و پوشال (Peter, 2012) و



شکل ۱- انواع مختلف زعفران
Figure 1- The types of saffron.

به کشور دیگر تغییر می‌کند. همچنین به دلیل اختلاف نظرهای کارشناسی در درجه‌بندی زعفران، ممکن است برخی از خطاها

لازم به ذکر است که هیچ استاندارد برای درجه‌بندی زعفران بر اساس ظاهر آن وجود ندارد و تعاریف آن از یک کشور

تصویر جدید پیش بینی می کند (Faucitano et al., 2005). اولین قدم در الگوریتم های طبقه بندی، استخراج تعداد کافی از ویژگی ها از تصاویر است، به طوری که آنها توانایی جدا کردن کلاس ها به درستی را دارند و با آموزش یک سیستم هوشمند می توانند به صورت خودکار یک تصویر جدید را طبقه بندی کنند. طبقه بندی توسط الگوریتم های آماری و خوشه بندی های مختلف با اختصاص هر تصویر به کلاس مربوطه انجام می شود (Donis-González & Guyer, 2016). هدف از این مطالعه، ارائه یک روش یادگیری ماشین بینایی برای تشخیص انواع مختلف زعفران (سرگل، نگین و پوشال) است. این مقاله به شرح زیر برگزار می شود. انواع مختلف زعفران در بخش ۲ توضیح داده شده است و یک بانک اطلاعاتی با گرفتن تصاویر با استفاده از تلفن همراه از نمونه های فله زعفران جمع آوری می شود. پس از اعمال مراحل پیش پردازش، خصوصیات رنگ و درصد ماده خارجی (براساس رنگ) زعفران بدست آمد. در بخش ۳ نتایج آموزش کلاسیفایرهای مختلف با ویژگی های استخراج شده، ارائه شده است. بحث و مقایسه بین کلاسیفایرها نیز در بخش ۳ انجام شده و بهترین روش با توجه به تجربه نویسندگان ارائه شده است. بخش ۴ به نتیجه گیری و کارهای آینده اختصاص یافته است.

مواد و روش ها

نمونه های زعفران

در مجموع ۴۴۰ نمونه زعفران مختلف موجود در بازار که فاقد هرگونه افزودنی و تقلب بودند و میانگین درصد رطوبت آنها بین ۵ تا ۷ درصد بود از شهرهای مختلف استان خراسان رضوی شامل: گناباد، بجستان، رشتخوار، سبزوار، مشهد، تربت حیدریه و کاخک تهیه شدند و سپس نمونه ها کدگذاری شدند. چهار

رخ دهد که با استفاده از یک رویکرد عینی مانند پردازش تصویر می توان از بروز آن جلوگیری کرد (Pourreza et al., 2012). پیشرفت در فناوری بینایی ماشین، الگوریتم های دقیق، قوی و کم هزینه ای را ارائه می دهد که می توان از آنها برای تشخیص کیفیت مواد غذایی و همچنین درجه بندی کیفیت زعفران استفاده کرد (Kiani & Minaei, 2016). کیانی و همکاران (Kiani et al., 2018) پیشنهاد یک سیستم بویایی-چشایی و سیستم بینایی ماشین را برای ارزیابی کیفیت زعفران و جایگزینی آن برای تشخیص حسی توسط افراد خبره ارائه کردند. مینایی و همکاران (Minaei et al., 2017) نشان دادند که ترکیب سیستم بینایی کامپیوتری^۱ و پرسپترون چند لایه^۲ ابزاری برای ارزیابی کیفیت زعفران بر اساس قدرت رنگی آنها است. آنها نشان دادند که عملکرد مدل پرسپترون چند لایه برای تشخیص قدرت رنگی زعفران بهتر از حداقل مربعات جزئی^۳ و MLR^۴ بود و مقدار ۵CSR ۹۶/۶ درصد بدست آمد وقتی که بانک داده خودشان استفاده کردند. امروزه از سیستم های بینایی رایانه ای رنگی در صنایع غذایی مختلف و سیستم های درجه بندی محصولات کشاورزی استفاده می شود زیرا آنها قابل اعتماد، سریع و ارزان هستند (Donis-González & Guyer, 2016). از این فناوری برای طبقه بندی یا تشخیص کیفیت محصولات کشاورزی و انواع غذاها از جمله خرما (Muhammad, 2015)، پسته (Omid et al., 2017)، سیب (Paulus & Schrevens, 1999) گندم (Pourreza et al., 2012) و پیتزا (Sun, 2016) استفاده می شود. سیستم بینایی کامپیوتر بر اساس الگوهای خاص استخراج شده از مجموعه تصاویر رنگی تهیه شده با کلاس های مختلف آموزش داده می شود مانند: بافت، خصوصیات هندسی و خصوصیات رنگ. سپس سیستم دسته خاص را از یک

۴- Machine-learned Ranking

۵- Success rate of classification

۱-Computer vision system

۲- Multilayer perceptron

۳-Partial least squares regression



شکل ۲- محفظه تصویر برداری
Figure 2- Imaging chamber.

پردازش تصاویر

تصویر اولیه در شکل ۳a ارائه شده است. در مرحله اول از یک فیلتر پایین گذر جهت حذف نویزها و صاف کردن تصویر استفاده شد و تصویر برای فرایند بعدی مطابق شکل ۳b آماده شد. سپس پیش زمینه تصویر با شدت بیشتر از ۲۰ پیکسل انتخاب شد که در شکل ۳c نشان داده شده است. اجزای کوچک از پیش زمینه باینری حذف شدند (شکل ۳d).

ویژگی‌های رنگی

رنگ یکی از مهم‌ترین فاکتورهای کیفی زعفران می باشد. رنگ جسم می تواند در چندین فضای رنگی نمایش داده شود. در بررسی محصولات کشاورزی از فضاهای رنگی مختلف می توان استفاده کرد. در این پژوهش از فضاهای رنگی RGB، HSr، L^*a^*b ، CIE، YCbCr، YIQ و برای توصیف اطلاعات نهفته در پیکسل‌های تشکیل دهنده تصاویر استفاده شد. از توابع مختلف پردازش تصویر در نرم‌افزار متلب برای تبدیل تصاویر دیجیتال به مدل‌های مختلف رنگی استفاده شد. اطلاعات مولفه‌های مختلف رنگی تشکیل دهنده فضاهای رنگی شامل R، G، B، H، S، L، a، b، c، E، I، Y، Cr، Cb، Y، I، Q و از هر تصویر با استفاده از نرم افزار متلب استخراج شد (Mohammadzadeh et al., 2016).

متخصص که سابقه طولانی در تجارت زعفران داشتند نیز در این کار برای ارزیابی نمونه‌ها همکاری داشتند. کارشناسان نمونه‌ها را به سه کلاس سرگل، نگین و پوشال طبقه‌بندی کردند. یک پایگاه داده برای این نمونه‌های مرجع، همانطور که در مقالات (Zheng & Lu, 2012) و (Zhang et al., 2017) و (Donis- González & Guyer, 2016) شرح داده شده، ایجاد شد.

تهیه تصاویر

برای تهیه تصاویر از یک گوشی موبایل (Samsung Galaxy S7 Edge SM-G935FD Dual SIM 32GB Mobile Phone) استفاده شد که بر روی یک محفظه تصویربرداری به فاصله ۹ سانتی متر از نمونه قرار داده شد. برای سیستم روشنایی محفظه تصویربرداری از لامپ‌های ال ای دی نواری (۴۰۱۴ SMD LED Module) استفاده شد که برای یکنواختی نور لامپ‌ها، یک دیفیوزر در زیر لامپ‌ها نصب شد. محل قرار گرفتن نمونه بصورت کشویی تعبیه شد. برای اینکه رنگ پس زمینه تصاویر مشکی باشد در کف آن از یک پارچه مخمل مشکی مات استفاده شد (شکل ۲). سرعت شاتر دوربین ۱/۵۰۰ ثانیه، بدون استفاده از فلش، فاصله کانونی لنز ۴/۲ میلی متر، مقدار دیافراگم F1/7 و ایزو ۱۰۰ تنظیم شد. تصاویر با حداکثر رزولوشن (۴۰۳۲×۳۰۲۴ pixels) گرفته و با فرمت "JPG" ذخیره شدند. بر اساس میانگین نظر افراد خبره با مشاهده تصاویر ۴۴۰ نمونه مختلف، آن‌ها به سه طبقه شامل: ۱۹۵ نمونه پوشال، ۱۲۹ نمونه نگین و ۱۱۶ نمونه سرگل تقسیم شدند. در این حالت میانگین نظرات افراد خبره به عنوان معیار برجسب گذاری نمونه‌ها انتخاب شد.

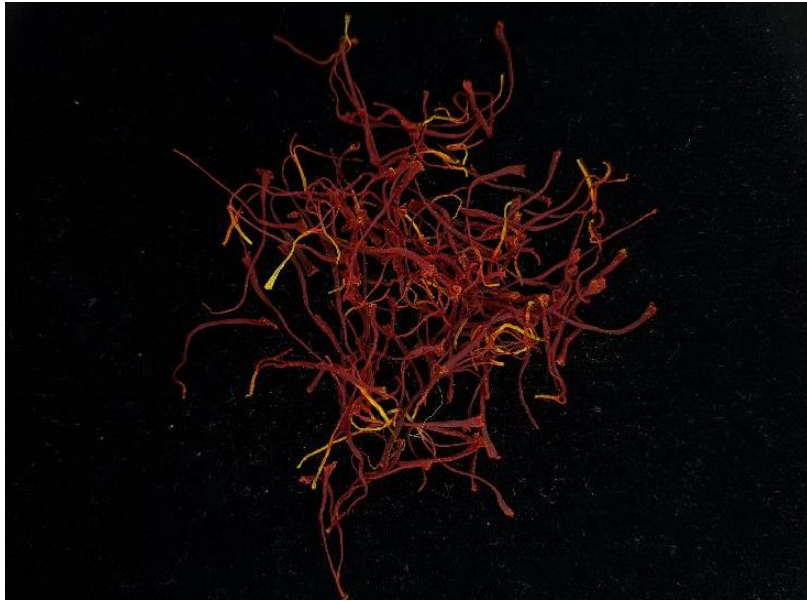


Figure- 3a



Figure- 3b

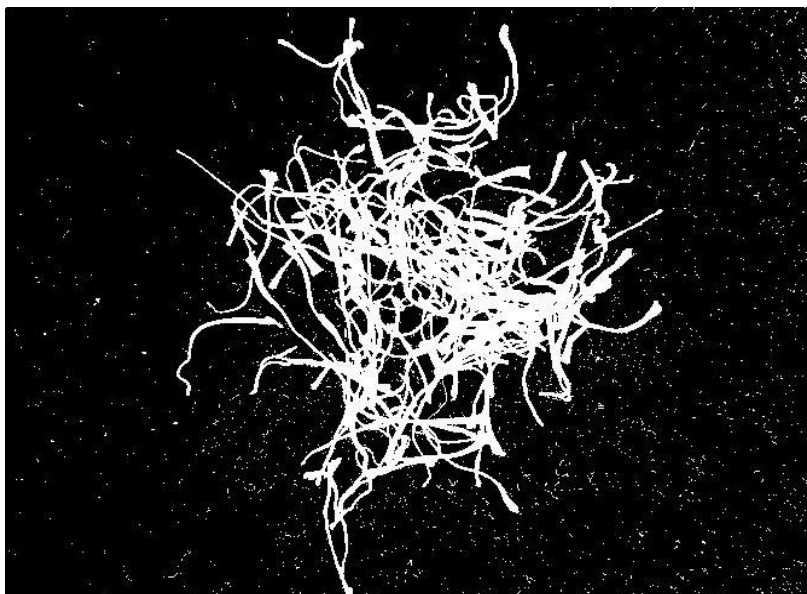


Figure- 3c

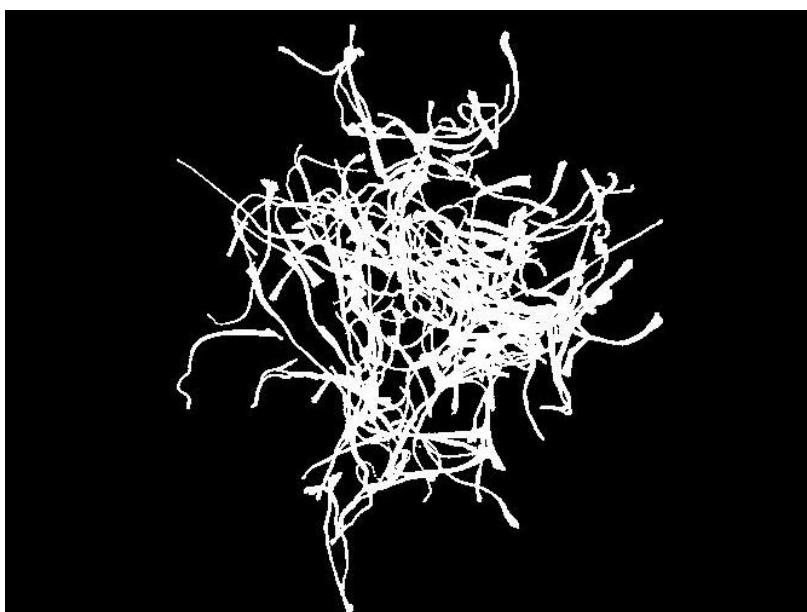


Figure- 3d



Figure- 3e

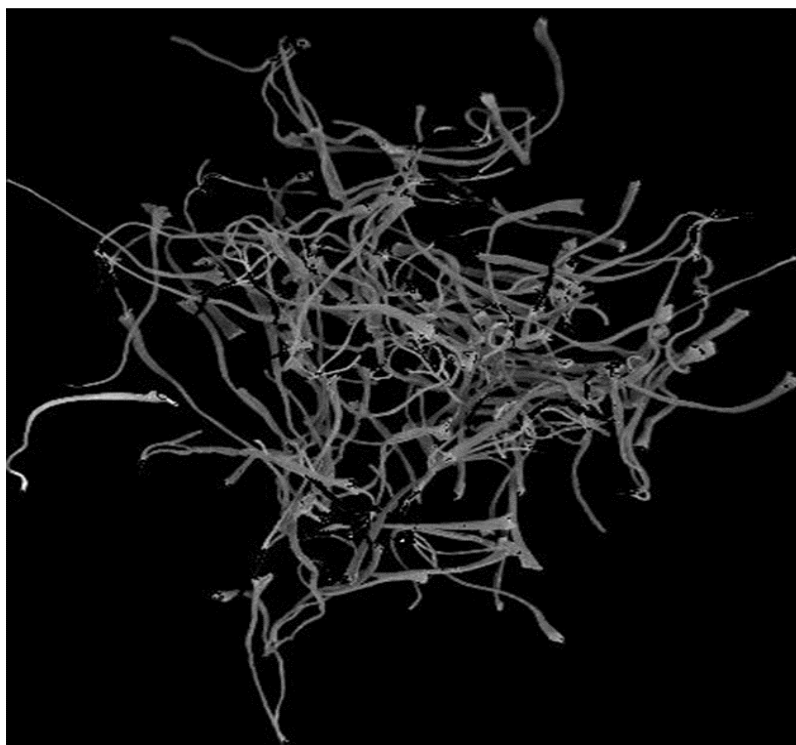


Figure- 3f



Figure- 3g

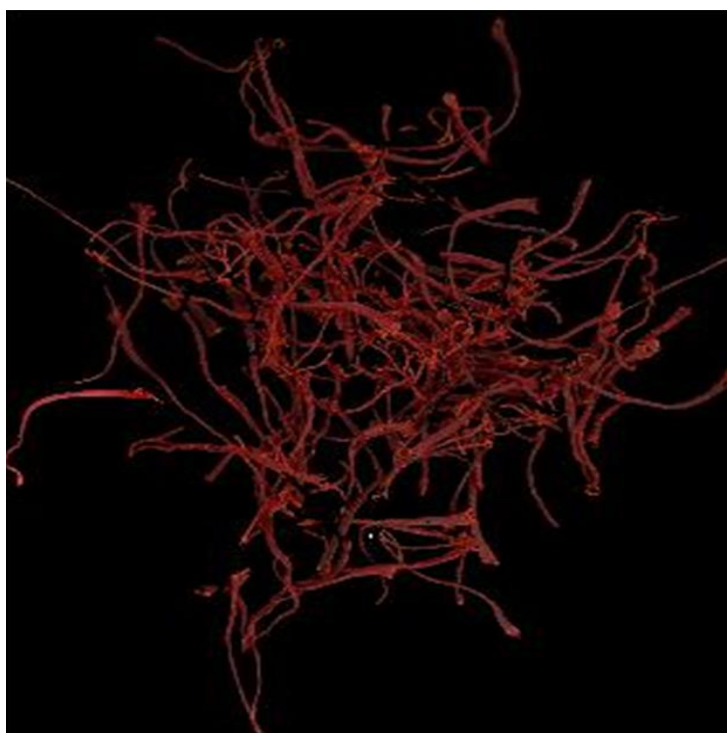


Figure- 3h

شکل ۳- پیش پردازش تصویر: تصویر ابتدایی یک نمونه زعفران پوشال (a)، اعمال فیلتر میانگین (b)، مشخص کردن پس زمینه تصویر (c)، حذف نویزهای پس زمینه تصویر (d) جداکردن اجزای سفید و زرد رنگ از کلاله زعفران (e)، تفکیک کلاله‌های قرمز رنگ از سایر اجزای زعفران (f)، قسمت‌های زرد و سفید زعفران (g)، قسمت‌های خالص کلاله زعفران (h)

Figure 3- Image pre-processing: (a) Original sample image, (b) filtered image, (c) foreground of the image, (d) foreground of the image after opening, (e) Separate the white and yellow components from the saffron stigma, (f) Separation of red stigmas from other components of saffron, (g) Yellow and white parts, (h) Pure saffron parts.

استخراج ویژگی های رنگی از تصویر

که بتوان از این مدل برای پیش بینی رده رکوردهایی که برچسب رده آن ها (متغیر هدف) ناشناخته می باشد، استفاده نمود (Azizi et al., 2016). از ویژگی های استخراج شده از تصاویر برای طبقه بندی استفاده شد. ۲۲ کلاسیفایر مختلف استفاده شد که پارامترهای هر کلاسیفایر مطابق جدول ۱ می باشد. کلاسیفایرها به شرح زیر می باشند:

درخت تصمیم

درخت تصمیم^۲ یک الگوریتم یادگیری ماشینی است که داده های آموزش را بطور بازگشتی توسط هر گره به منظور به حداکثر رساندن جداسازی داده ها طبقه بندی می کند. تصمیم گیری در درخت از گره ریشه تا گره برگ شروع می شود تا یک پاسخ را پیش بینی کنید. گره برگ حاوی پاسخ است (Kamiński et al., 2018). انواع مدل هایی که در این گروه استفاده می شوند عبارتند از: Medium Tree، Fine Tree و Coarse Tree.

آنالیزهای افتراقی

آنالیزهای افتراقی یک روش طبقه بندی چند متغیره است. این روش برای مجموعه داده های وسیعی مفید می باشد. در این روش فرض می شود که طبقات مختلف تولید داده بر اساس توزیع های گوسی مختلف ایجاد می شود. برای آموزش یک طبقه بندی کننده، تابع فیت شده، پارامترهای توزیع گوسی را برای هر کلاس برآورد می کند (Riveiro-Valiño et al., 2009). مدل های استفاده شده در این گروه شامل: آنالیز تشخیص خطی^۳ و آنالیز تشخیص درجه دوم^۴ می باشند. آنالیز تشخیص خطی یک روش آماری می باشد که از جمله در یادگیری ماشین و بازساخت الگو برای پیدا کردن ترکیب خطی خصوصیتی که به بهترین صورت دو یا چند کلاس از اشیا را از

برای آستانه گیری از ابزار آستانه گیری رنگ^۱ در محیط متلب استفاده شد. تابعی در قالب یک ماسک ایجاد شد که بتواند قسمت های زائد تصویر را حذف کند و کلاله ها (قسمت های قرمز رنگ) و مواد خارجی از قبیل اجزاء خامه و پرچم زعفران را از همدیگر تفکیک کند. شکل ۳e، ۳f، ۳g و ۳h نتیجه استفاده از این آستانه را نشان می دهد. در نهایت یک کد متلب با استفاده از نرم افزار MATLAB 2017b تهیه شد که قادر بود ۲۱ ویژگی را از هر تصویر توده استخراج کند. این ویژگی ها شامل: درصد مواد خارجی به کل توده، درصد کلاله به کل توده، درصد مواد خارجی به کلاله، اجزای فضاهای رنگی مختلف شامل R، G، B، H، S، L، a، b، C، I، E، Y، Cr، Cb، Y، I، Y، Cr، Cb، Y، E، I، C، b، a، L، r، S، H، B، G و Q برای هر تصویر از توده زعفران بود.

الگوریتم های طبقه بندی

یادگیری ماشین یکی از شاخه های پرکاربرد هوش مصنوعی است. در روش های یادگیری ماشینی با استفاده از اطلاعات مجموعه آموزشی، بر پایه الگو و قواعد یادگیری، پیش بینی های آینده انجام می گیرد. یادگیری تحت نظارت، یک روش عمومی در یادگیری ماشین است که در آن به یک سیستم، مجموعه ای از جفت های ورودی - خروجی ارائه شده و سیستم تلاش می کند تا تابعی از ورودی به خروجی را فرا گیرد. یادگیری تحت نظارت نیازمند تعدادی داده ورودی به منظور آموزش سیستم است. یادگیری تحت نظارت خود به دو دسته رگرسیون و طبقه بندی تقسیم می شود. در یادگیری ماشین، طبقه بندی نوعی یادگیری با نظارت است، که مجموعه ای از ویژگی ها برای آموزش موجودند. طبقه بندی فرایندی برای پیدا کردن مدلی است که رده های موجود در داده ها را تعریف می نماید و متمایز می کند، با این هدف

۳-Linear Discriminant Analysis

۴-Quadratic Discriminant Analysis

۱- Color Thresholder

۲- Decision Tree

به یک نمونه شناخته شده مشخص می‌شوند که در مجموعه آموزش قرار می‌گیرند و سپس طبقه‌بندی بر اساس مجموعه آموزش انجام می‌گیرد (Xie et al., 2017). انواع طبقه‌بندی کننده‌های این گروه عبارتند از: Medium، Fine KNN، Coarse KNN، Cubic KNN و KNN، Cosine KNN، Weighted KNN.

طبقه‌بندی گروهی

بوستینگ یک فرا الگوریتم ترکیبی در حوزه یادگیری ماشین است که برای کاهش عدم توازن و همچنین واریانس به کار می‌رود. این روش در یادگیری با نظارت مورد استفاده قرار گرفته و از خانواده الگوریتم‌های یادگیری ماشین به شمار می‌رود. این تکنیک، روشی برای تبدیل سیستم‌های یادگیری ضعیف به قوی بر اساس ترکیب نتایج طبقه‌بندی‌های مختلف است (Dutta et al., 2015). انواع مدل‌هایی که در این گروه استفاده می‌شوند شامل: Bagged Trees، Boosted Trees، Subspace، Discriminant و RUSBoost Trees.

شاخص‌های اعتبارسنجی و ارزیابی عملکرد

از روش اعتبارسنجی متقابل ۵ لایه برای اعتبارسنجی طبقه‌بندی استفاده شد. در اعتبارسنجی متقابل K لایه^۱، داده‌ها به K زیرمجموعه افراز می‌شوند. از این K زیرمجموعه، هر بار یکی برای اعتبارسنجی و $K-1$ تای دیگر برای آموزش بکار می‌روند. این روال K بار تکرار می‌شود و همه داده‌ها دقیقاً یک بار برای آموزش و یک بار برای اعتبارسنجی بکار می‌روند. در نهایت میانگین نتیجه این K بار اعتبارسنجی به‌عنوان یک تخمین نهایی برگزیده می‌شود. به‌طور معمول از روش اعتبارسنجی پنج لایه یا ده لایه در پژوهش‌های مدل‌سازی و پیش‌بینی استفاده می‌شود.

هم جدا می‌کند، استفاده می‌شود. آنالیز تشخیصی خطی بسیار به تحلیل واریانس و تحلیل رگرسیون نزدیک است. در هر سه این روش‌های آماری متغیر وابسته به صورت یک ترکیب خطی از متغیرهای دیگر مدل‌سازی می‌شود (Martínez & Kak, 2001).

ماشین بردار پشتیبان^۱

ماشین بردار پشتیبان یک ابزار مدل‌سازی مؤثر برای طبقه‌بندی است و برای رگرسیون، طبقه‌بندی الگو، پیش‌بینی و تشخیص مشکل مورد استفاده قرار گرفته است (Nasirahmadi et al., 2019). در واقع ماشین بردار پشتیبان برای اینکه داده‌های غیر خطی را از هم تفکیک کند باید از کرنل‌های مختلف استفاده کند. برای این کار دیگر در فضای دو بعدی کار نمی‌کند بلکه داده‌ها به فضایی با ابعاد بیشتر نگاشت داده می‌شوند تا بتوان آن‌ها را در این فضای جدید بصورت خطی تفکیک نمود. در واقع ایده اصلی ماشین بردار پشتیبان این است که ابر صفحه‌هایی در فضا ترسیم کند که عمل تمایز نمونه‌های مختلف داده را بطور بهینه انجام دهند. ابرصفحه‌ای که بیشترین حاشیه جداسازی را دارد پیدا می‌کند و نزدیک‌ترین داده‌های آموزشی به ابر صفحه جداکننده، بردارهای پشتیبان نامیده می‌شوند (Huang et al., 2016). انواع مدل‌هایی که در این گروه استفاده می‌شوند شامل: Quadratic SVM، Linear SVM، Medium، Fine Gaussian SVM، Cubic SVM، Gaussian SVM و Coarse Gaussian SVM.

نزدیک‌ترین همسایه

طبقه‌بندی نزدیک‌ترین همسایه در ابعاد کم دقت پیش‌بینی خوبی دارند. اما ممکن است در ابعاد زیاد این قابلیت را نداشته باشند. حافظه بالایی را استفاده می‌کنند و برای درک و تفسیر آسان نیستند. در این طبقه‌بندی نمونه‌هایی از همسایگی یا شبیه

^۲ -k-fold cross-validation

^۱ - Support vector machine

جدول ۱- پارامترهای تنظیم شده برای طبقه بندی کننده ها در متلب

Table 1- Parameters adjusted for classifiers in Matlab

پارامترهای تنظیم شده Parameters adjusted					
طبقه بندی Classifier	حداکثر تعداد اسپلیت ها Maximum number of splits	معیار اسپلیت Split criterion	تصمیم جانشین اسپلیت ها Surrogate decision splits	حداکثر گره Maximum surrogates per node	
درخت تصمیم Decision Tree	درخت ظریف Fine Tree	100	شخص تنوع جینی Ginis diversity index	خاموش off	10
	درخت متوسط Medium Tree	20	شخص تنوع جینی Ginis diversity index	خاموش off	10
	درخت درشت Coarse Tree	4	شخص تنوع جینی Ginis diversity index	خاموش off	10
تشخیص خطی Linear discriminant	ساختار کواریانس Covariance structure				
	قطری Diagonal				
تشخیص افتراقی Quadratic discriminant	قطری Diagonal				
	تابع کرنل Kernel function	سطح محدودیت جعبه Box constraint level	حالت مقیاس کرنل Kernel scale mode	روش چند کلاسی Multiclass method	
ماشین بردار پشتیبان SVM	خطی Linear	خطی Linear	1	خودکار Auto	One-Vs-One
	درجه دوم Quadratic	Quadratic	1	خودکار Auto	One-Vs-One
	مکعبی Cubic	مکعبی Cubic	1	خودکار Auto	One-Vs-One
	گاوسی ظریف Fine Gaussian	گاوسی Gaussian	1	دستی manual:2.7	One-Vs-One
	گاوسی متوسط Medium Gaussian	گاوسی Gaussian	1	دستی manual:11	One-Vs-One
	گاوسی درشت Coarse Gaussian	گاوسی Gaussian	1	دستی manual:44	One-Vs-One
	تعداد همسایگان Number of neighbors	متریک فاصله Distance metric	وزن فاصله Distance weight		
نزدیک ترین همسایه KNN	ظریف Fine	1	Euclidean	برابر Equal	
	متوسط Medium	10	Euclidean	برابر Equal	
	درشت Coarse	100	Euclidean	برابر Equal	
	زیبا Cosine	10	Cosine	برابر Equal	
	مکعبی Cubic	10	Minkoski	برابر Equal	
	وزن داشت Weighted	10	Euclidean	معکوس مربع Squared inverse	

	روش گروهی Ensemble method	نوع یادگیرنده Learner type	حداکثر تعداد اسپلیتها Maximum number of splits	تعداد یادگیرنده Number of learners	نرخ یادگیری Learning rate	
	درختان تقویت شده Boosted Trees	آدابوست Ada Boost	درخت تصمیم Decision Tree	20	30	0.1
	درختان بسته بندی شده Bagged Trees	کیسه ای Bag	درخت تصمیم Decision Tree	439	30	-
طبقه بندی گروهی Ensemble	تشخیص فضای فرعی Subspace Discriminant	فضای فرعی Subspace	تشخیص Discriminant	-	30	ابعاد فضای فرعی Subspace dimension: 60
	نزدیک ترین همسایه فضای فرعی Subspace KNN	فضای فرعی Subspace	نزدیکترین همسایگان Nearest Neighbors	-	30	ابعاد فضای فرعی Subspace dimension: 60
	درختان راس بوست شده RUS Boosted Trees	راس بوست RUS Boost	درخت تصمیم Decision Tree	20	30	0.1

کننده‌ها استفاده شد. تجزیه و تحلیل آماری با استفاده از نرم افزار SPSS انجام شد.

نتایج و بحث

در این مطالعه از ۴۴۰ عکس رنگی از نمونه‌های مختلف زعفران شامل ۱۹۵ نمونه از پوشال، ۱۲۹ نمونه نگیں و ۱۱۶ نمونه سرگل استفاده شد. واژه‌نامه تعریف شده برای کلاسیفایرها، شامل ۲۱ ویژگی رنگی، استخراج شده از ۴۴۰ نمونه بود. سپس طبقه‌بندی با استفاده از اعتبار سنجی متقاطع ۵ لایه ارزیابی شد. در اعتبارسنجی متقابل، نمونه‌های اصلی به طور تصادفی به پنج گروه تقسیم شدند. از چهار گروه به عنوان داده‌های آموزشی برای توسعه مدل استفاده شد و گروه باقی مانده نیز به عنوان داده‌های اعتبارسنجی برای آزمون طبقه‌بندی نگهداری شدند.

مزیت این روش نسبت به اعتبارسنجی تصادفی زیر نمونه تکراری^۱ این است که از همه مشاهدات برای آموزش و اعتبار سنجی استفاده می‌شود و از هر مشاهده دقیقاً یک بار برای اعتبارسنجی استفاده می‌شود (Siedliska et al., 2014). همچنین درصد دقت^۲، ماتریس سردرگمی^۳، نرخ مثبت صحیح^۴، نرخ منفی کاذب^۵، ارزش بیش بینی مثبت^۶ و نرخ کشف کاذب^۷ محاسبه و گزارش شد (Xie et al., 2017). همچنین منحنی مشخصه عملکرد سیستم^۸ در MATLAB براساس نرخ مثبت صحیح و نرخ منفی کاذب محاسبه شد. سطح زیر منحنی مشخصه عملکرد سیستم که از ۰/۵ (بدون توانایی تمایز) تا ۱ (بهترین توانایی تمایز) متغیر است نیز محاسبه شد (Nasirahmadi et al., 2019). از آنالیز واریانس یک طرفه و آزمون دانکن برای تعیین تفاوت معنی داری بین دقت طبقه‌بندی

۵-False negative rate

۶-Positive predictive rate

۷-False discovery rate

۸-Receiver operating characteristic

۱-Repeated random sub-sampling validation

۲-accuracy

۳-confusion matrix

۴-True positive rate

از Linear SVM, Bagged Trees و RUSBoost Trees سایر کلاسیفایرها بیشتر بود. میانگین دقت این ۴ کلاسیفایر با هم اختلاف معنی داری نداشتند ($p < 0.05$) و دقت طبقه بندی برای کلاسیفایر Linear SVM مقدار ۸۲/۲۳ درصد ($\pm 0/66$ درصد) بدست آمد.

این فرایند ۵ بار تکرار شد و هرکدام از گروهها یک بار به عنوان داده های اعتبار سنجی استفاده شدند (Kuo et al., 2016). جدول ۲ میزان دقت ۲۲ روش مختلف طبقه بندی را با استفاده از ۲۱ ویژگی رنگی نشان می دهد. میانگین دقت طبقه بندی با ۱۰ بار اجرای اعتبارسنجی متقابل ۵ لایه بدست آمد. میانگین دقت کلاسیفایرهای Linear Discriminant

جدول ۲- متوسط دقت طبقه بندی کننده ها برای طبقه بندی زعفران با استفاده از ۲۱ ویژگی رنگی
Table 2- Average accuracy of classifiers for classifying saffron using 21 color features

ردیف NO.	طبقه بندی کننده Classifier	میانگین دقت Average accuracy	انحراف معیار SD	ردیف NO.	طبقه بندی کننده Classifier	میانگین دقت Average accuracy	انحراف معیار SD
1	درخت ظریف Fine Tree	79.65%	1.43%	12	نزدیک ترین همسایه ظریف Fine KNN	77.33%	1.26%
2	درخت متوسط Medium Tree	80.86%	1.68%	13	نزدیک ترین همسایه متوسط Medium KNN	77.71%	0.81%
3	درخت درشت Coarse Tree	79.58%	0.77%	14	نزدیک ترین همسایه درشت Coarse KNN	73.5%	0.4%
4	تشخیص خطی Linear Discriminant	82.23%	0.66%	15	زیبا نزدیک ترین همسایه Cosine KNN	77.69%	1.18%
5	تشخیص افتراقی Quadratic Discriminant	58.17%	0.66%	16	نزدیک ترین همسایه مکعبی Cubic KNN	78.02%	0.87%
6	ماشین بردار پشتیبان خطی Linear SVM	82.27%	0.7%	17	نزدیک ترین همسایه وزن دار Weighted KNN	79.39%	0.85%
7	ماشین بردار پشتیبان درجه دوم Quadratic SVM	80.73%	0.69%	18	درختان تقویت شده Boosted Trees	81.09%	0.97%
8	ماشین بردار پشتیبان مکعبی Cubic SVM	78.89%	0.8%	19	درختان بسته بندی شده Bagged Trees	82.18%	1.04%
9	ماشین بردار پشتیبان گاوسی ظریف Fine Gaussian SVM	79.34%	0.68%	20	تشخیص فضای فرعی Subspace Discriminant	80.65%	0.55%
10	ماشین بردار پشتیبان گاوسی متوسط Medium Gaussian SVM	81.11%	0.58%	21	نزدیک ترین همسایه فضای فرعی Subspace KNN	60.71%	3.12%
11	ماشین بردار پشتیبان گاوسی درشت Coarse Gaussian SVM	76.7%	0.39%	22	درختان راس پوست شده RUSBoost Trees	81.83%	1.19%

کلاسیفایرها برای درجه بندی زعفران هستند. میانگین دقت کلاسیفایرهای Linear SVM, Linear Discriminant

همچنین مشخص شد که کلاسیفایرهای آنالیزهای افتراقی، ماشین بردار پشتیبان و طبقه بندی گروهی بهتر از دیگر

طبقه‌بندی استفاده شود. حساسیت و تشخیص پذیری دو شاخص مهم برای ارزیابی آماری عملکرد نتایج آزمون‌های طبقه‌بندی دو حالت هستند. حساسیت، که در برخی از علوم نرخ مثبت واقعی یا احتمال تشخیص صحیح نیز نامیده می‌شود، نسبتی از موارد مثبت است که کلاسیفایر آن‌ها را به درستی به عنوان مثبت مشخص می‌کند. تشخیص‌پذیری، که نرخ منفی واقعی نیز نامیده می‌شود، به معنی نسبتی از موارد منفی است که کلاسیفایر آن‌ها را به درستی به عنوان منفی تشخیص می‌دهد.

Bagged Trees و RUSBoost Trees به ترتیب مقدار ۸۲/۲۳ درصد، ۸۲/۲۳، ۸۲/۱۸ و ۸۱/۸۳ درصد بدست آمد که با سایر کلاسیفایرها تفاوت معنی داری داشت. شکل ۴ ماتریس درهم‌ریختگی را برای چهار کلاسیفایر ذکر شده در بالا نشان می‌دهد. این شکل نشان می‌دهد که هر کدام از چهار کلاسیفایرهای ذکر شده چند تا از نمونه‌های پوشال، نگین و سرگل را به درستی تشخیص دهند. داده‌هایی که از ماتریس درهم‌ریختگی بدست می‌آید می‌تواند جهت محاسبه سایر شاخص‌های ارزیابی

Types of saffron	Pushal	Negin	Sargol	classifier
Pushal	174, 183, 184, 178	11, 4, 5, 5	10, 8, 6, 12	Linear Discriminant
Negin	8, 9, 9, 10	103, 106, 96, 96	18, 14, 24, 23	Linear SVM Bagged Trees
Sargol	5, 7, 13, 7	28, 35, 20, 24	83, 74, 83, 85	RUSBoost Trees

شکل ۴- ماتریس‌های سردرگمی مدل‌های طبقه‌بندی برای انواع زعفران به عنوان یک متغیر مستقل، هر مدل از نمایه رنگ خاصی برخوردار است و سلول‌های مورب (به رنگ صورتی) طبقه‌بندی‌های صحیح را ارائه می‌دهند

Figure 4- Confusion matrices of the classification models for cultivars as an independent variable. Each model has a specific color representation and the diagonal cells (In pink) present the correct classifications.

تصویر شامل R، G و B (RGB)، S، H، L (HSR) و a، L، R (Lab) استخراج شد. در انتها یک مجموعه تصاویر خاکستری از هر تصویر رنگی بدست آمد، بنابراین ده تصویر مختلف از هر تصویر رنگی مارچوبه بدست آمد. با استفاده از تولباکس Balu نرم افزار متلب ۲۰۱۲ ویژگی‌های مختلف استخراج شد که در مجموع شامل ۱۹۳۱ ویژگی از ده تصویر بود. بعد از نرمالایز کردن، ویژگی‌های مناسب با استفاده از روش‌های انتخاب ویژگی در تولباکس Balu انتخاب شدند تا با استفاده از آن‌ها کلاسیفایرها آموزش داده شوند. از ۵ کلاسیفایر مختلف شامل شبکه عصبی، QDA، LDA، Mahalanobis distance و KNN استفاده شد. اعتبار سنجی کلاسیفایرها با روش اعتبار بخشی متقابل ۴ لایه انجام گرفت. دقت کلاسیفایرها به ترتیب ۹۰/۲، ۸۹/۷، ۸۰/۵، ۷۶/۱ و ۶۵/۱ گزارش شد. بهترین دقت طبقه‌بندی مربوط به شبکه عصبی بود.

نتایج ذکر شده در شکل ۴ نشان می‌دهد که کلاسیفایر Bagged Trees از ۱۹۵ نمونه پوشال ۱۸۴ نمونه را به درستی تشخیص دهد. که به آن تعداد مثبت صحیح گفته می‌شود. از ۱۱ نمونه دیگر ۵ نمونه را به اشتباه نگین و ۶ نمونه را به اشتباه سرگل حدس زده است که به آن مثبت کاذب گفته می‌شود. همچنین از ۱۲۹ نمونه نگین، تعداد ۹ نمونه و از ۱۰۶ نمونه سرگل، تعداد ۱۳ نمونه را به غلط پوشال حدس زده است که به آن منفی کاذب گفته می‌شود. همچنین کلاسیفایر Linear SVM از ۱۲۹ نمونه نگین تعداد ۱۰۶ نمونه و کلاسیفایر RUSBoost Trees از ۱۱۶ نمونه سرگل، تعداد ۸۵ نمونه را به درستی تشخیص دادند. دونیز-گزالس و همکاران (Donis-González & Guyer, 2016) با هدف طبقه‌بندی مارچوبه به سه کلاس نوک، وسط ساقه و انتهای ساقه اقدام به تهیه عکس‌های رنگی با استفاده از اسکنر کردند. اجزای رنگی هر

داد. تصاویر توسط یک دوربین CCD رنگی گرفته شد و در یک کامپیوتر شخصی ویژگی های تصویر آنالیز شد. ابتدا تصاویر از فضای رنگی RGB به فضای رنگی HSV تبدیل شدند. برای سگمنت کردن تصاویر از جزء H در فضای رنگی HSV و روش آستانه گیری Otsu بکار گرفته شد. یک بردار ویژگی شامل ۳۰ ویژگی رنگی از تصاویر گرفته شده استخراج شد. روش انتخاب ویژگی براساس آنالیز حسی برای انتخاب ویژگی های برتر بکار برده شد. مدل های مختلف SVM امتحان شدند. مدل SVM که دارای تابع هسته چند جمله ای مکعبی و ۳۸ بردار تکیه گاه بود، بهترین دقت را داشت (Nouri-Ahmadabadi et al., 2017).

همچنین جزئیات آنالیز دقت طبقه بندی چهار کلاسیفایر Bagged Trees, Linear SVM, Linear Discriminant و RUSBoost Trees در جدول ۳ گزارش شده است. ارزش بالای نرخ TP و PP، و ارزش پایین نرخ FN و FD، به معنی مدل طبقه بندی خوب است. این مقادیر برای زعفران پوشال بهتر از سایر طبقات زعفران بود. نرخ TP برای زعفران پوشال از ۸۹ تا ۹۴ درصد برای چهار کلاسیفایر ذکر شده متغیر بود. همچنین نرخ PP بین ۸۹ تا ۹۳ درصد برای پوشال گزارش شد. نرخ FN و نرخ FD نشان داد که خطای طبقه بندی سرگل و نگین بیشتر از پوشال است. این خطاها زمانی اتفاق می افتد که مقادیر نزدیک به یکدیگر باشند و طبقه بندی آن ها مشکل است. از نظر ظاهر، نگین و سرگل بسیار شبیه هستند و تمایز میان آنها دشوار است. در پوشال سه فیلامنت کلاله به هم چسبیده اند ولی در نگین و سرگل این سه فیلامنت از هم جدا هستند. همچنین در پوشال مقداری خامه سفید متمایل به زرد همراه کلاله وجود دارد که باعث می شود تمایز بیشتری با سرگل و نگین داشته باشد. نرخ FN برای سرگل بین ۲۷ تا ۳۶ درصد و نرخ FD بین ۲۳ تا ۲۹ درصد برای سرگل برای هر کدام از چهار مدل طبقه بندی ذکر شده در جدول ۳ گزارش شد. نتایج نشان داد که حساسیت کلاسیفایرها Linear SVM و Bagged Trees نسبت به سایر

استفاده از ویژگی های رنگی تصاویر برای طبقه بندی محصولات کشاورزی توسط سایر محققین بکار گرفته شده است. امید و همکاران (Omid et al., 2017) با استفاده از ماشین بینایی و ویژگی های رنگی، مغزهای پسته پوست شده را طبقه بندی کردند. در این مطالعه از یک الگوریتم بر اساس پردازش تصویر و تکنیک های یادگیری ماشین شامل شبکه عصبی مصنوعی و ماشین بردار پشتیبان برای طبقه بندی مغز پسته به گروه شامل: سبز، سبز مایل به زرد، زرد، رنگ مخلوط شده و مواد ناخواسته بکار گرفته شد. با توجه به نتایج آنالیزهای حسی، جزء H از فضای رنگی HSV و شاخص آماری میانگین بیشترین تاثیر را بر طبقه بندی نمونه ها داشت. توسط آنالیز حسی بردار ویژگی تا ۲۶ ویژگی کاهش یافت. متعاقباً تحلیل مولفه های اصلی برای کاهش بیشتر اندازه بردار ورودی به شبکه عصبی مصنوعی و ماشین بردار پشتیبان استفاده شد. بهترین طبقه بندی ANN یک ساختار ۵-۸-۷ با نرخ طبقه بندی صحیح ۹۹/۴ درصد داشت. بهترین تابع هسته برای الگوریتم SVM بر اساس شعاع با CCR، C، سیگما و تعداد بردارهای پشتیبان به ترتیب ۹۹/۸۸، ۱۰، ۳/۵ و ۲۶۶ بود (Omid et al., 2017). دی الیورا و همکاران (de Oliveira et al., 2016) یک سیستم بینایی کامپیوتری برای طبقه بندی دانه های قهوه بر اساس تکنیک های هوش مصنوعی بکار بردند. ارزیابی رنگ دانه های قهوه سبز یک فرایند مهم در تشخیص کیفیت و تعیین قیمت این محصول می باشد. این ارزیابی بطور معمول توسط بازرسی بصری و ابزارهای سنتی انجام می گیرد که محدودیتهایی دارد. هدف از این مطالعه ایجاد یک سیستم بینایی کامپیوتری بود که قادر به اندازه گیری CIE و Lab دانه های قهوه سبز می باشد و آن ها را مطابق رنگشان طبقه بندی کند (de Oliveira et al., 2016). نوری احمدآبادی و همکاران (Nouri-Ahmadabadi et al., 2017) یک سیستم هوش مصنوعی را براساس ماشین بینایی و ماشین بردار پشتیبان برای سورتینگ مغز پسته و پوست پسته گسترش

برای تشخیص پوشال در بین ۱۹۵ نمونه، به ترتیب مقدار ۹۲ و ۸۹ درصد بود.

کلاسیفایرها برای تشخیص پوشال بیشتر است. یعنی این کلاسیفایرها قادر بودند که از تعداد ۴۴۰ نمونه مقدار ۹۴ درصد آن‌ها را به درستی پوشال تشخیص دهند. همچنین دقت آن‌ها

جدول ۳- تجزیه و تحلیل مشروح دقت چهار کلاسیفایر
Table 3- Detailed accuracy analysis using class of the studied four classifiers

مدل طبقه‌بندی Classification model	نرخ مثبت صحیح (پوشال) TP* rate (Pushal)	نرخ مثبت صحیح (نگین) TP rate (Negin)	نرخ مثبت صحیح (سرگل) TP rate (Sargol)	نرخ منفی کاذب (پوشال) FN* rate (Pushal)	نرخ منفی کاذب (نگین) FN rate (Negin)	نرخ منفی کاذب (سرگل) FN rate (Sargol)	ارزش بیش بینی مثبت (پوشال) PP* rate (Pushal)	ارزش بیش بینی مثبت (نگین) PP rate (Negin)	ارزش بیش بینی مثبت (سرگل) PP rate (Sargol)	نرخ کشف کاذب (پوشال) FD* rate (Pushal)	نرخ کشف کاذب (نگین) FD rate (Negin)	نرخ کشف کاذب (سرگل) FD rate (Sargol)
تشخیص خطی Linear Discriminant	89%	80%	70%	11%	20%	28%	93%	73%	75%	7%	27%	25%
ماشین بردار پشتیبان خطی Linear SVM	94%	82%	64%	6%	18%	36%	92%	73%	77%	8%	27%	23%
بسته بندی شده Bagged Trees	94%	74%	72%	6%	26%	28%	89%	79%	73%	11%	21%	27%
درختان راس بوست شده RUSBoost Trees	91%	74%	73%	9%	26%	27%	91%	77%	71%	9%	23%	29%

* TP= True positive rate, FN= False negative rate, PP= Positive predictive rate and FD= False discovery rate.

آن منحنی مشخصه عملکرد سیستم به یک نزدیک باشد. بدین معنی که هرچه به یک نزدیکتر باشد اندازه گیری دقیق تر و مناسبتر بوده است. همچنین بدترین مدل، مدلی است که در آن منحنی مشخصه عملکرد سیستم به صفر نزدیک باشد. بدین مفهوم که هرچه به صفر نزدیکتر باشد نشان دهنده ضعف در

منحنی مشخصه عملکرد سیستم^۱ یک روش اضافه برای ارزیابی مدل طبقه بندی است. منحنی مشخصه عملکرد سیستم، ارتباط بین مثبت واقعی و کاذب را مشخص می کند. محور X ها نرخ مثبت کاذب و محور Y ها نرخ مثبت واقعی را نشان می دهد (Siedliska et al., 2014). بهترین مدل مدلی است که در

^۱ Receiver operating characteristic (ROC)

است. شکل ۵ منحنی مشخصه عملکرد سیستم همراه با سطح زیر منحنی مشخصه عملکرد سیستم را برای درجه بندی پوشال، نگین و سرگل را مربوط به کلاسیفایر Bagged Trees نشان می دهد. مقادیر AUC به ترتیب برای پوشال و نگین و سرگل، ۰/۹۷، ۰/۹۲ و ۰/۹۲ بدست آمد.

دقت اندازه گیری است. در حقیقت این مدل صفرها را یک و یکها را صفر پیش بینی میکند و این یعنی حداکثر خطا و وقتی که به عدد ۰/۵ برمی خورد، هیچ تفاوتی در ظرفیت لایه ها را نشان نمیدهد. سطح زیر منحنی مشخصه عملکرد سیستم یکی از مهم ترین پارامترهای آماری برای ارزیابی عملکرد طبقه بندی

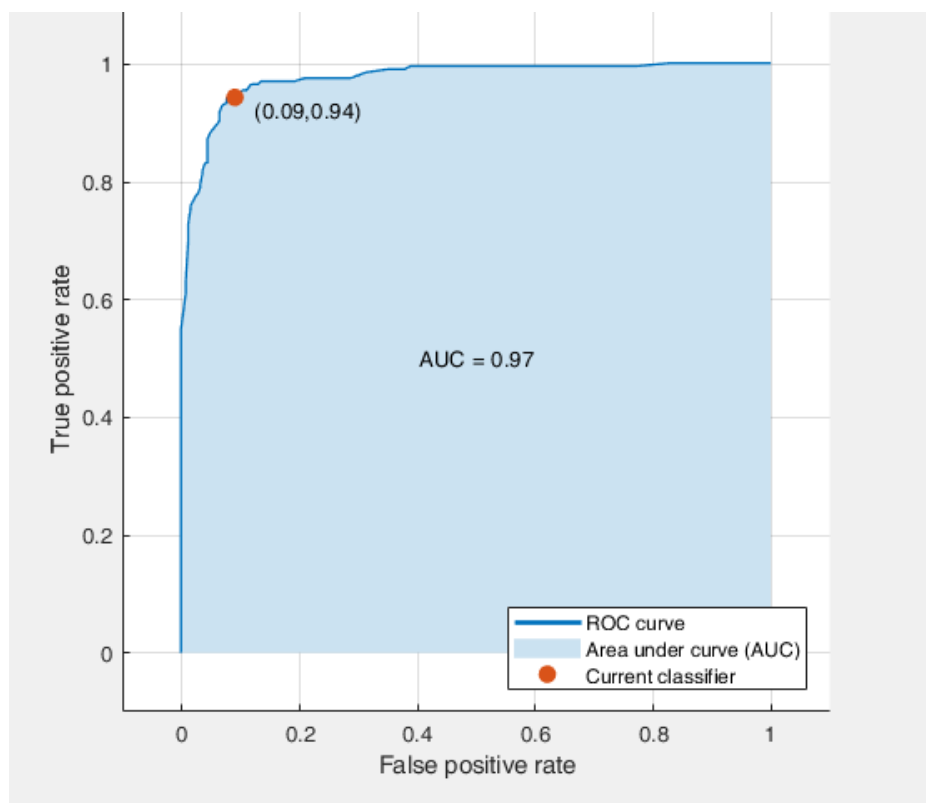


Figure- 5a

\ Area under the curve (AUC)

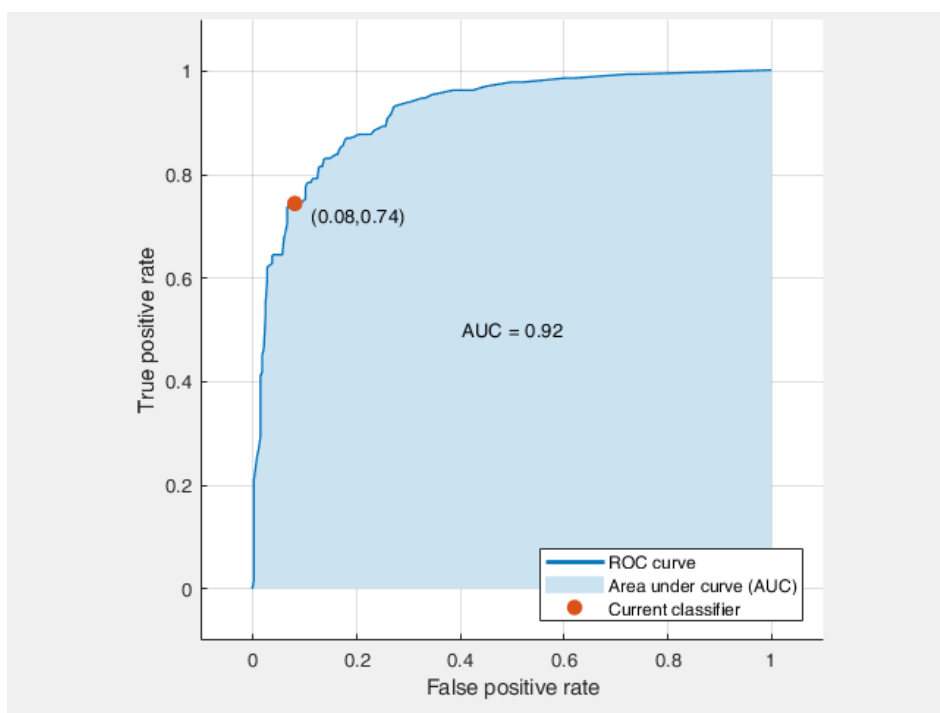


Figure- 5b

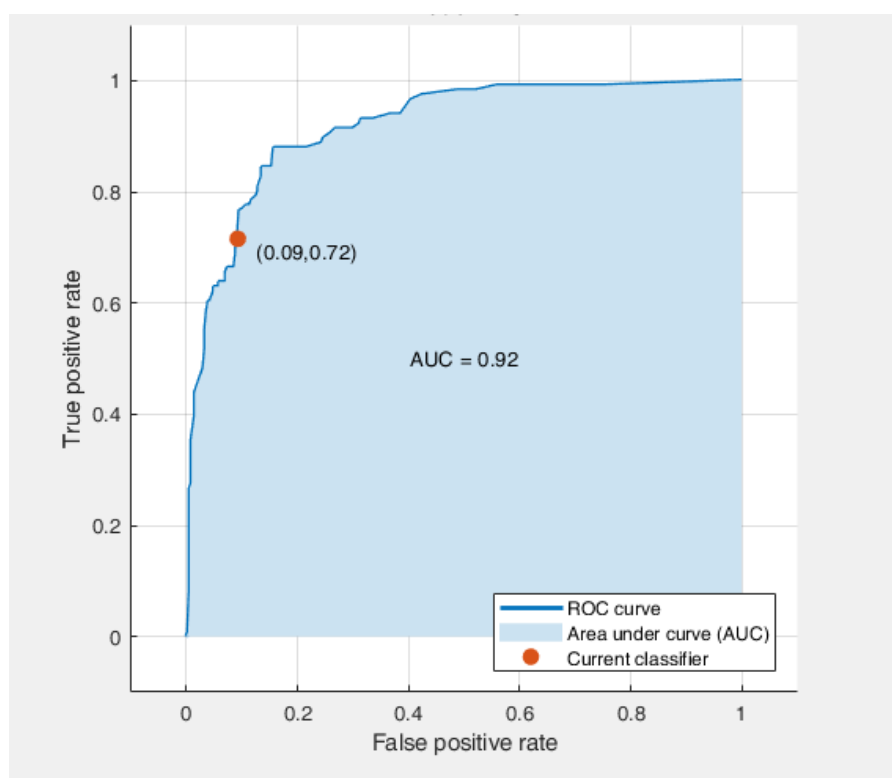


Figure- 5c

شکل ۵- منحنی ROC با مقدار AUC مربوط به کلاسیفایر Bagged Trees: پوشال (a)، نگین (b)، سرگل (c)
 Figure 5- The ROC along with AUC values of the Bagged Trees classifier: (a) Pushal, (b) Negin, (c) Sargol.

شکل ۵a منحنی مشخصه عملکرد سیستم را برای تشخیص پوشال نشان می دهد با توجه به داده های شکل

مقادیر کلی AUC برای کلاسیفایرهای Linear Discriminant، Linear SVM، Bagged Trees و RUSBoost Trees به ترتیب ۰/۹۲۳، ۰/۹۱۳، ۰/۹۳۶ و ۰/۹۱۳ بود. این نتایج نشان داد که این کلاسیفایر در طبقه بندی زعفران بر اساس ویژگی های رنگ در بین ۲۲ کلاسیفایر، دقت بهتری داشتند. مطابق نتایج جداول ۳ و ۴ طبقه بندی زعفران پوشال با استفاده از این مدل ها بهتر از دو کلاس دیگر زعفران، یعنی سرگل و نگین بود. نتایج منحنی مشخصه عملکرد سیستم نشان داد که با توجه به سطح زیر منحنی که برای پوشال نزدیک به یک است و بالا بودن میزان نرخ مثبت صحیح و کم بودن مقدار نرخ مثبت کاذب یعنی حساسیت کلاسیفایرها برای تشخیص صحیح پوشال از دو طبقه دیگر بهتر است و بطور اختصاصی تری می توانند بین پوشال و دو کتگوری دیگر تمایز ایجاد کنند. پس می توان با استخراج ویژگی های رنگ از تصاویر توده و با استفاده از کلاسیفایرهای ذکر شده با دقت نسبتاً خوبی نمونه های پوشال را از نگین و سرگل تشخیص داد.

حساسیت کلاسیفایر برای تشخیص پوشال ۹۴ درصد است و کلاسیفایر قادر است که ۹۱ درصد از نمونه های سرگل و نگین را پوشال تشخیص دهد و ۹ درصد از آن ها را به اشتباه پوشال تشخیص داد. همچنین شکل ۵b نشان می دهد که حساسیت کلاسیفایر برای تشخیص نگین ۷۴ درصد است و کلاسیفایر قادر بود که ۹۲ درصد از نمونه های پوشال و سرگل را نگین تشخیص دهد و ۸ درصد از آن ها را به اشتباه نگین تشخیص داد. همچنین شکل ۵c مشخص می کند که حساسیت کلاسیفایر برای تشخیص سرگل ۷۲ درصد است و کلاسیفایر قادر بود که ۹۱ درصد از نمونه های پوشال و نگین را سرگل تشخیص دهد و ۹ درصد از آن ها را به اشتباه سرگل تشخیص داد.

هرچه مقدار AUC به یک نزدیکتر باشد کارایی تشخیص کلی کلاسیفایر بهتر است (Hu et al., 2016). جدول ۴ نشان می دهد که مقادیر AUC برای پوشال بهتر از سایر کلاس های زعفران است. کلاسیفایر Bagged Trees بیشترین مقدار AUC را برای تشخیص نگین و سرگل و پوشال داشت. علاوه بر این

جدول ۴- مقدار AUC (سطح زیر منحنی ROC) چهار کلاسیفایر برای انواع مختلف زعفران

Table 4- The AUC (area under the ROC curve) of the four classifiers for different saffron classes

مدل طبقه بندی Classification model	پوشال Pushal	نگین Negin	سرگل Sargol	مجموع Overall
تشخیص خطی Linear Discriminant	0.95	0.91	0.91	0.923
ماشین بردار پشتیبان خطی Linear SVM	0.95	0.9	0.89	0.913
درختان بسته بندی شده Bagged Trees	0.97	0.92	0.92	0.936
درختان راس پوست شده RUSBoost Trees	0.96	0.88	0.90	0.913

تجاری در دسترس استفاده شد. تصاویر به افراد خبره داده شد تا نمونه ها را به سه طبقه سرگل، نگین و پوشال طبقه بندی کنند. تعداد ۲۱ ویژگی رنگ از تصاویر زعفران استخراج شد. ویژگی های رنگی به عنوان ورودی به ۲۲ کلاسیفایر مختلف برای طبقه بندی زعفران در نظر گرفته شد. کلاسیفایرهای Linear

نتیجه گیری

این مقاله نشان می دهد که شاخص رنگ شاخص خوبی برای جداسازی زعفران پوشال، نگین و سرگل است. در این تحقیق نمونه های زعفران از شهرهای مختلف استان خراسان جمع آوری شد. برای گرفتن تصاویر زعفران از یک تلفن همراه

مورفولوژی و بافت تصویر و تکنیک‌های یادگیری ماشین (یعنی یادگیری عمیق) می‌تواند دقت شناسایی کلاس زعفران را بهینه کند.

و Bagged Trees, Linear SVM, Discriminant RUSBoost Tree بهتر از سایر کلاسیفایرها بودند. دقت طبقه‌بندی برای کلاسیفایر Linear SVM مقدار ۸۲/۲۳ درصد ($\pm ۰/۶۶$ درصد) بدست آمد. مطالعات آینده در مورد ویژگی‌های

منابع

- Atefi, M., Akbari Oghaz, A., and Mehri, A. 2013. Drying effects on chemical and sensorial characteristics of saffron. *Iranian Journal of Nutrition Sciences and Food Technology* 8 (3): 201-208. (In Persian with English Summary).
- Azarabadi, N., and Özdemir, F. 2018. Determination of crocin content and volatile components in different qualities of Iranian saffron. *GIDA/The Journal of FOOD* 43 (3).
- Azizi, Z., Moradi, S.H., Moradi S.M., Rafat, S.A., and Shodja, J. 2016. Genetic classification of Azari and North ecotype Buffalo population using SVM method. *Iranian Journal of Animal Science* 47 (2): 279-290.
- Bonyadi, M.H.J., Yazdani, S., and Saadat, S. 2014. The ocular hypotensive effect of saffron extract in primary open angle glaucoma: a pilot study. *BMC Complementary and Alternative Medicine* 14 (1): 399.
- de Oliveira, E.M., Leme, D.S., Barbosa, B.H.G., Rodarte, M.P., and Pereira, R.G.F.A. 2016. A computer vision system for coffee beans classification based on computational intelligence techniques. *Journal of Food Engineering* 171: 22-27.
- Donis-González, I.R., and Guyer, D.E. 2016. Classification of processing asparagus sections using color images. *Computers and Electronics in Agriculture* 127: 236-241.
- Dutta, R., Dutta, R., Smit, D., Rawnsley, R., Bishop-Hurley, G., Hills, J., Timms, G., and Henry, D. 2015. Dynamic cattle behavioural classification using supervised ensemble classifiers. *Computers and Electronics in Agriculture* 111: 18-28.
- Faucitano, L., Huff, P., Teuscher, F., Garipey, C., and Wegner, J. 2005. Application of computer image analysis to measure pork marbling characteristics. *Meat Science* 69 (3): 537-543.
- Fernández, J.A. 2004. Biology, biotechnology and biomedicine of saffron. *Recent Research Development and Plant Science* 2: 127-159.
- Hu, M.H., Dong, Q.L., and Liu, B.L. 2016. Classification and characterization of blueberry mechanical damage with time evolution using reflectance, transmittance and interactance imaging spectroscopy. *Computers and Electronics in Agriculture* 122: 19-28.
- Huang, M., Tang, J., Yang, B., and Zhu, Q. 2016. Classification of maize seeds of different years based on hyperspectral imaging and model updating. *Computers and Electronics in Agriculture* 122: 139-145.
- Kafi, M., Koocheki, A., and Rashed, M. 2006. *Saffron (Crocus sativus): Production and Processing*. Science Publishers, Enfield, NH, USA, 1-241.
- Kamiński, B., Jakubczyk, M., and Szufel, P. 2018. A framework for sensitivity analysis of decision trees. *Central European Journal of Operations Research* 26 (1): 135-159.
- Kiani, S., and Minaei, S. 2016. Potential application of machine vision technology to saffron (*Crocus sativus* L.) quality characterization. *Food Chemistry* 212: 392-394.
- Kiani, S., Minaei, S., and Ghasemi-Varnamkhasti, M. 2018. Instrumental approaches and innovative systems for saffron quality

- assessment. *Journal of Food Engineering* 216: 1-10.
- Kuo, T.Y., Chung, C. L., Chen, S.Y., Lin, H.A., and Kuo, Y.F. 2016. Identifying rice grains using image analysis and sparse-representation-based classification. *Computers and Electronics in Agriculture* 127: 716-725.
- Martínez, A.M., and Kak, A.C. 2001. Pca versus lda. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2): 228-233.
- Masi, E., Taiti, C., Heimler, D., Vignolini, P., Romani, A., Mancuso, S. 2016. PTR-TOF-MS and HPLC analysis in the characterization of saffron (*Crocus sativus* L.) from Italy and Iran. *Food Chemistry* 192: 75-81.
- Minaei, S., Kiani, S., Ayyari, M., and Ghasemi-Varnamkhasti, M. 2017. A portable computer-vision-based expert system for saffron color quality characterization. *Journal of Applied Research on Medicinal and Aromatic Plants* 7: 124-130.
- Mohammadzadeh, A., Golzarian, M., and Abbaspour, F.M. 2016. Classification of pomegranate arils from image features using linear discriminant analysis. *Iranian Food Science and Technology Research Journal* 12 (1): 182-192. (In Persian with English Summary).
- Muhammad, G. 2015. Date fruits classification using texture descriptors and shape-size features. *Engineering Applications of Artificial Intelligence* 37: 361-367.
- Nasirahmadi, A., Sturm, B., Olsson, A.C., Jeppsson, K.H., Müller, S., Edwards, S., Hensel, O. 2019. Automatic scoring of lateral and sternal lying posture in grouped pigs using image processing and support vector machine. *Computers and Electronics in Agriculture* 156: 475-481.
- Nouri-Ahmadabadi, H., Omid, M., Mohtasebi, S.S., and Firouz, M.S. 2017. Design, development and evaluation of an online grading system for peeled pistachios equipped with machine vision technology and support vector machine. *Information Processing in Agriculture* 4 (4): 333-341.
- Omid, M., Firouz, M.S., Nouri-Ahmadabadi, H., and Mohtasebi, S.S. 2017. Classification of peeled pistachio kernels using computer vision and color features. *Engineering in Agriculture, Environment and Food* 10 (4): 259-265.
- Paulus, I., and Schrevens, E. 1999. Shape characterization of new apple cultivars by Fourier expansion of digitized images. *Journal of Agricultural Engineering Research* 72 (2): 113-118.
- Peter, K.V. 2012. *Handbook of Herbs and Spices*. Elsevier, 1-640.
- Pourreza, A., Pourreza, H., Abbaspour-Fard, M H., and Sadrnia, H. 2012. Identification of nine Iranian wheat seed varieties by textural analysis with image processing. *Computers and Electronics in Agriculture* 83: 102-108.
- Riveiro-Valiño, J., Álvarez-López, C., and Marey-Pérez, M.F. 2009. The use of discriminant analysis to validate a methodology for classifying farms based on a combinatorial algorithm. *Computers and Electronics in Agriculture* 66 (2): 113-120.
- Shahdadi, H., Barati, F., Bahador, R.S., and Eteghadi, A. 2016. Clinical applications of saffron (*Crocus sativus*) and its constituents: A literature review. *Der Pharmacia Lettre* 8 (19): 205-209.
- Siedliska, A., Baranowski, P., and Mazurek, W. 2014. Classification models of bruise and cultivar detection on the basis of hyperspectral imaging data. *Computers and Electronics in Agriculture* 106: 66-74.
- Sun, D.W. 2016. *Computer Vision Technology for Food Quality Evaluation*. Academic Press, 1-583.
- Xie, C., Yang, C., and He, Y. 2017. Hyperspectral imaging for classification of healthy and gray mold diseased tomato leaves with different

infection severities. *Computers and Electronics in Agriculture* 135: 154-162.

Zhang, M., Lee, D.J., Lillywhite, K., and Tippetts, B. 2017. Automatic quality and moisture evaluations using Evolution Constructed Features. *Computers and Electronics in Agriculture* 135: 321-327.

Zheng, H., and Lu, H. 2012. A least-squares support vector machine (LS-SVM) based on fractal analysis and CIELab parameters for the detection of browning degree on mango (*Mangifera indica* L.). *Computers and Electronics in Agriculture* 83: 47-51.

Classification of Saffron Using Color Features Extracted from the Image

Morteza Mohamadzadeh Moghadam¹, Masoud Taghizadeh^{2*}, Hassan Sadrnia³ and Hamid Reza Pourreza⁴

Submitted: 23 October 2019

Accepted: 10 May 2020

Mohamadzadeh Moghadam, M., Taghizadeh, M., Sadrnia, H., Pourreza, H. R. 2020. Classification of Saffron using Color Features Extracted from the Image. Saffron Agronomy & Technology, 8(3): 319-399.

Abstract

The classification of saffron as the most expensive spice is of great importance for customers and traders. In general, two methods are currently used to classify saffron. The first method is based on the experiences of an expert and by observing the samples. The second method is destructive and is performed using laboratory methods. According to experts, the use of machine learning techniques to classify saffron is a goal due to its non-destructive nature and timely characteristics. This method can also increase the accuracy of the industrial scale grading process. In this paper, a vision machine method is presented. Due to lack of documented research on this subject, a comprehensive literature search is presented in this work. Almost all color characteristics were extracted and used in a large number of classifiers. Experts in Iran classify saffron into three main categories based on their appearance: Pushal, Negin and Sargol. In this paper, a database consisting of 440 images from saffron for the three different classes was collected using a mobile phone camera. After applying some preprocessing steps, such as background removal, cropping etc., 21 color features were extracted using different image analysis methods. Twenty-two classifiers were employed for classification. Comparing results of different classifiers showed that the Linear Discriminant, Linear SVM, Bagged Trees and RUSBoost Trees can produce more accurate grading compared to other classifiers when using color features. In particular, mean classification accuracy of 82.23% was achieved in this work using Linear a SVM classifier.

Keywords: Classification, Saffron, Image processing

1- Ph.D Student of Food Science and Technology, Department of Food Science and Technology, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran

2- Assistant Professor, Department of Food Science and Technology, Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran

3- Associate Professor, Biosystems Engineering. Faculty of Agriculture, Ferdowsi University of Mashhad, Mashhad, Iran

4- Department of Computer Engineering. Ferdowsi University of Mashhad IRAN

(*- Corresponding author Email: mtaghizadeh@um.ac.ir)

DOI: 10.22048/JSAT.2020.206278.1362