

به کارگیری تبدیل موجک تصویر برای درجه بندی کیفی چای سبز به کمک الگوریتم های فراابتکاری

عادل بخشی پور زیارتگاهی^۱، حماد ذرعی فروش^{۱*}، ایرج باقری^۱

۱- استادیار، گروه مهندسی مکانیزاسیون کشاورزی، دانشگاه گیلان، رشت، ایران

* نویسنده مسئول (hemad.zareiforoush@guilan.ac.ir)

تاریخ دریافت: ۱۳۹۷/۰۵/۱۰

تاریخ پذیرش: ۱۳۹۷/۱۰/۱۱

واژه های کلیدی

الگوریتم های فراابتکاری

بافت تصویر

پردازش تصویر

تبدیل موجک

چکیده

این مطالعه با هدف بررسی بهترین مشخصه های استخراج شده از تصاویر برای تعیین بهترین روش طبقه بندی کیفی چای سبز توسط الگوریتم های فراابتکاری انجام شد. ۵ طبقه مختلف چای سبز مطابق با استاندارد سازمان ملی استاندارد ایران ارزیابی شدند. پس از دریافت تصاویر گروه های مختلف چای سبز در رایانه، تعداد ۶ بلوک تصویر مربعی از هر کدام از تصاویر رنگی اولیه جدا شدند. این بلوک های تصویر از حالت RGB به تصاویر سطح خاکستری تبدیل شدند. فیلتر موجک گسسته هار سطح اول روی تصاویر خاکستری اعمال شد و ۴ زیر تصویر موجکی استخراج شدند. ماتریس های هم رخداد برای هر کدام از تصاویر زیرباند موجک محاسبه شدند و ۱۷ ویژگی بافتی پرکاربرد در مطالعه های بافتی تصویر، از تصاویر زیرباند استخراج شدند (مجموعاً ۶۸ ویژگی بافتی برای هر بلوک تصویر). با استفاده از آنالیز مؤلفه های اصلی، تعداد ۸ ترکیب ویژگی از ویژگی های اولیه تولید شدند و برای جداسازی ۵ گروه چای سبز استفاده شدند. نتایج نشان داد که الگوریتم هایی از شبکه های عصبی مصنوعی، ماشین بردار پشتیبان و درخت تصمیم، قادر به طبقه بندی کیفی چای سبز با دقت بالایی بودند. در حالی که شبکه بیزین عملکرد قابل قبولی نداشت. با توجه به آماره های ارزیابی، شبکه های عصبی مصنوعی پرسپترون چندلایه (با مقادیر آماره کاپا، ریشه میانگین مربعات خطا و دقت طبقه بندی به ترتیب برابر با ۰/۹۹۰۱، ۰/۴۲۰ و ۹۹/۱۷ درصد) به عنوان بهترین طبقه بندی انتخاب شد. براساس نتایج این پژوهش، استفاده از ماشین بینایی و ویژگی های بافتی مستخرج از زیرباندهای موجک تصویر، روش مناسبی برای طبقه بندی کیفی چای سبز می باشد.

مقدمه

چای یکی از نوشیدنی های محبوب در سرتاسر جهان است که پس از فراوری برگ سبز گیاه چای^۱ به دست می آید. چای در ایران به عنوان یکی از محصولات استراتژیک منطقه شمال کشور شناخته شده است که دارای تاریخچه ای ۱۲۰ ساله می باشد. سطح زیر کشت چای

حدود ۳۲۰ هزار هکتار است که عمدتاً در استان گیلان متمرکز بوده و بخش کمی نیز در استان مازندران قرار گرفته است. براساس گزارش وزارت جهاد کشاورزی- معاونت برنامه ریزی و اقتصادی (۱۳۹۶)، میزان تولید چای کشور در سال ۱۳۹۵ برابر با ۳۱/۳ هزار تن بوده است. با توجه به نوع عملیات فراوری که پس از برداشت برگ های سبز چای انجام می شود، انواع مختلف چای به دست می آید

¹ *Camellia Sinensis*

پردازش تصویر در ارزیابی کیفی محصولات غذایی توسعه فراوانی یافته و در پژوهش‌های مختلفی مورد استفاده قرار گرفته است (Al Ohali, 2011; Benalia *et al.*, 2016; Liming & Yanchao, 2010; Mery, Pedreschi, & Soto, 2013; Payman, Bakhshipour, & Zareiforoush, 2018; Sabanci, Toktas, & Kayabasi, 2017; Semary, 2015; Tharwat, Elhariri, & Hassanien, 2015).

یکی از بخش‌های مهم هر سیستم ماشین بینایی، بخش تصمیم‌گیری سیستم می‌باشد که براساس پارامترهای استخراج شده از تصاویر و باتوجه به اهداف سیستم ماشین بینایی، به تصمیم‌گیری می‌پردازد. یکی از روش‌های مورد استفاده برای تصمیم‌گیری در ماشین بینایی، تکنیک داده‌کاوی^۳ است. داده‌کاوی با استفاده از ترکیبی از روش‌های تجزیه و تحلیل سنتی داده‌ها مانند روش‌های آماری و الگوریتم‌های پیچیده جدید مانند هوش مصنوعی، به استخراج الگوها و ارتباط‌های معنی‌دار جدید، از میان حجم زیادی از داده‌ها می‌پردازد (Farooqi, 2012). روش‌ها و الگوریتم‌های مختلفی برای داده‌کاوی در سامانه‌های ماشین بینایی برای طبقه‌بندی کیفی مواد غذایی مورد استفاده قرار گرفته‌اند. از جمله این روش‌ها می‌توان به شبکه‌های عصبی مصنوعی^۴ (ANN) (Sanaeifar, Mohtasebi, Ghasemi-Varnamkhashti, & Wang & Liu, 2014) شبکه بیزین^۵ (BN) (Siadat, 2014) ماشین بردار پشتیبان^۶ (SVM) (Nouri, 2015) و درخت تصمیم^۷ (DT) (Velásquez, Cruz-Tirado, 2017) اشاره کرد. (Siche, & Quevedo, 2017)

در پژوهشی، طبقه‌بندی کیفی برنج با استفاده از ماشین بینایی و به‌کارگیری^۴ روش طبقه‌بندی فراابتکاری^۸ مختلف شامل شبکه عصبی مصنوعی، شبکه بیزین، ماشین بردار پشتیبان و درخت تصمیم انجام شد. تصاویری از ۴ طبقه کیفی متفاوت برنج توسط یک سامانه ماشین بینایی تهیه شد و پس از انجام پیش‌پردازش و بخش‌بندی تصویر، ویژگی‌های رنگی و ابعادی از تصاویر

که مهم‌ترین انواع آن، چای سیاه، چای سبز، و چای اولانگ^۱ می‌باشند (Hayat, Iqbal, Malik, Bilal, & Mushtaq, 2015).

یکی از ابتدایی‌ترین و مهم‌ترین عملیات‌های فراوری محصولات کشاورزی، درجه‌بندی کیفی محصولات است. این مرحله از آن جهت اهمیت می‌یابد که درجه‌بندی کیفی این محصولات باعث افزایش ارزش افزوده محصول گردیده و برای خریداران امکان انتخاب محصولی منطبق بر نیاز خود را فراهم می‌کند.

معمولاً ارزیابی کیفی چای به صورت آزمون حسی توسط افراد خبره انجام می‌شود که نمونه‌های چای را براساس عطر، رنگ، بافت و خواص مورفولوژیکی ارزیابی می‌کنند (Liang, Lu, Zhang, Wu, & Wu, 2005). این نوع بازرسی اگرچه مؤثر می‌باشد اما زمان‌بر، گران و خسته‌کننده است. ازسوی دیگر، نتایج ارزیابی حسی می‌تواند تحت تأثیر شرایط جسمانی و روانی شخص آزمونگر قرار گیرد (Yu, Wang, Yao, Zhang, & Yu, 2008). بنابراین استفاده از روش‌های سریع، قابل اطمینان و با صحت و دقت زیاد در ارزیابی و کنترل کیفیت چای بسیار حائز اهمیت است.

در پژوهش‌های مختلف، برخی از روش‌های جدید مانند کروماتوگرافی مایع با کارایی بالا^۲ (HPLC) (He *et al.*, 2015)، تجزیه و تحلیل شیمیایی (Diniz, Pistonesi, 2015) Alvarez, Band, & de Araújo, 2015) طیف‌سنجی (Diniz, Barbosa, de Melo Milanez, Pistonesi, & de Chen, Zhao, 2016) و بینی و زبان الکترونیکی (Chen, Lin, & Zhao, 2011; Roy, Chattopadhyay, Tudu, Bhattacharyya, & Bandyopadhyay, 2014) به‌طور موفقیت‌آمیزی برای طبقه‌بندی چای استفاده شده‌اند. اما این روش‌ها مخرب، گران و تقریباً پیچیده هستند.

تکنولوژی ماشین بینایی با تقلید از سامانه بینایی موجودات زنده، امکان تشخیص و شناسایی اشیاء را فراهم می‌آورد و هدف از کاربرد آن، جایگزینی نظارت ماشینی به‌جای نظارت انسانی می‌باشد (ایزدی، ۱۳۹۲). در سال‌های اخیر، کاربرد ماشین بینایی و تکنیک‌های

³ Data mining

⁴ Artificial Neural Networks

⁵ Bayesian Network

⁶ Support Vector Machine

⁷ Decision Tree

⁸ Meta-heuristic

¹ Oolong tea

² High Performance Liquid Chromatography

بافتی مستخرج از ماتریس هم‌رخداد سطح خاکستری (GLCM^۷) مورد بررسی قرار گرفتند و بهترین دقت طبقه‌بندی ماشین‌بردار پشتیبان برابر با ۹۶/۳۳ درصد گزارش شد.

بررسی‌های انجام‌شده در پژوهش‌های پیشین، هیچ مطالعه‌ای درباره‌ی مقایسه‌ی روش‌های طبقه‌بندی و استراتژی‌های انتخاب ویژگی و داده‌کاوی برای طبقه‌بندی کیفی چای سبز براساس ویژگی‌های مستخرج از ماشین بینایی وجود ندارد. بنابراین این مطالعه با هدف بررسی بهترین مشخصه‌های تصویر استخراج‌شده توسط روش‌های داده‌کاوی مختلف برای بررسی بهترین روش طبقه‌بندی نمونه‌های چای سبز انجام شده است.

مواد و روش‌ها

تهیه‌ی نمونه‌ها

نمونه‌های چای سبز مورد استفاده در این مطالعه از یکی از کارخانه‌های فراوری چای شهرستان لاهیجان در استان گیلان دریافت شدند. ۵ طبقه‌ی کیفی مختلف از چای سبز مطابق با استاندارد ارائه‌شده توسط استاندارد ملی ایران به شماره ۵۳۶۰ (سازمان ملی استاندارد ایران، ۱۳۸۰) مورد بررسی قرار گرفتند. جزئیات بیشتر این گروه‌ها در جدول (۱) ارائه شده است. تصاویر نمونه‌های چای سبز مورد بررسی در این پژوهش، در شکل (۱) نمایش داده شده‌اند.

دانه‌های برنج استخراج گردید. نتایج نشان داد که شبکه‌ی عصبی مصنوعی بهترین عملکرد (۹۸/۷۲ درصد) را در طبقه‌بندی کیفی دانه‌های برنج براساس خصوصیات استخراج‌شده از تصاویر داشت (Zareiforoush, Minaei, Alizadeh, & Banakar, 2016). در پژوهش دیگری، از روش‌های داده‌کاوی به‌منظور طبقه‌بندی کیفی ۴ درجه‌ی مختلف چای سیاه براساس خصوصیات مستخرج از تصاویر استفاده شد. مقادیر مرتبط با ویژگی‌های رنگ و بافت تصویر و نیز ویژگی‌های آماری زیرباندهای تبدیل موجک از تصاویر نمونه‌های چای استخراج شدند. از دو روش انتخاب ویژگی^۱ مختلف شامل روش انتخاب ویژگی مبتنی بر همبستگی (CFD^۲) و آنالیز مؤلفه‌ی اصلی (PCA^۳) برای کاهش تعداد ویژگی‌های ورودی به طبقه‌بندها استفاده شد. سپس ویژگی‌های انتخاب‌شده به طبقه‌بندهای مختلف داده شد. در این پژوهش نیز شبکه‌ی عصبی مصنوعی بهترین دقت طبقه‌بندی را داشت (Bakhshipour, Sanaeifar, Payman, & de la Guardia, 2018).

Wu و He، Liu، Yu (۲۰۱۱) به طبقه‌بندی ۶ گروه مختلف از چای سبز براساس ویژگی‌های رنگی و بافتی و با استفاده از طبقه‌بندی ماشین‌بردار پشتیبان بر پایه‌ی حداقل مربعات^۴ (LSSVM) پرداختند. ترکیب‌های مختلف از مؤلفه‌های رنگی در فضای قرمز، سبز و آبی (RGB^۵) و فضای رنگی رنگ، اشباع و شدت (HSI^۶) و نیز ویژگی‌های

جدول ۱- مشخصات گروه‌های کیفی چای مورد بررسی در این پژوهش

نام فارسی	نام انگلیسی	علامت اختصاری
چای سبز قلم	Green Orange Pekoe	GOP
چای سبز شکسته ممتاز	Green Flowery Broken Orange Pekoe	GFBOP
چای سبز شکسته	Green Broken Orange Pekoe	GBOP
چای سبز باروتی	Green Broken Orange Pekoe Fannings	GBOPF
چای سبز خاکه	Green Dust	GD

¹ Feature selection

² Correlation-based Feature Selection

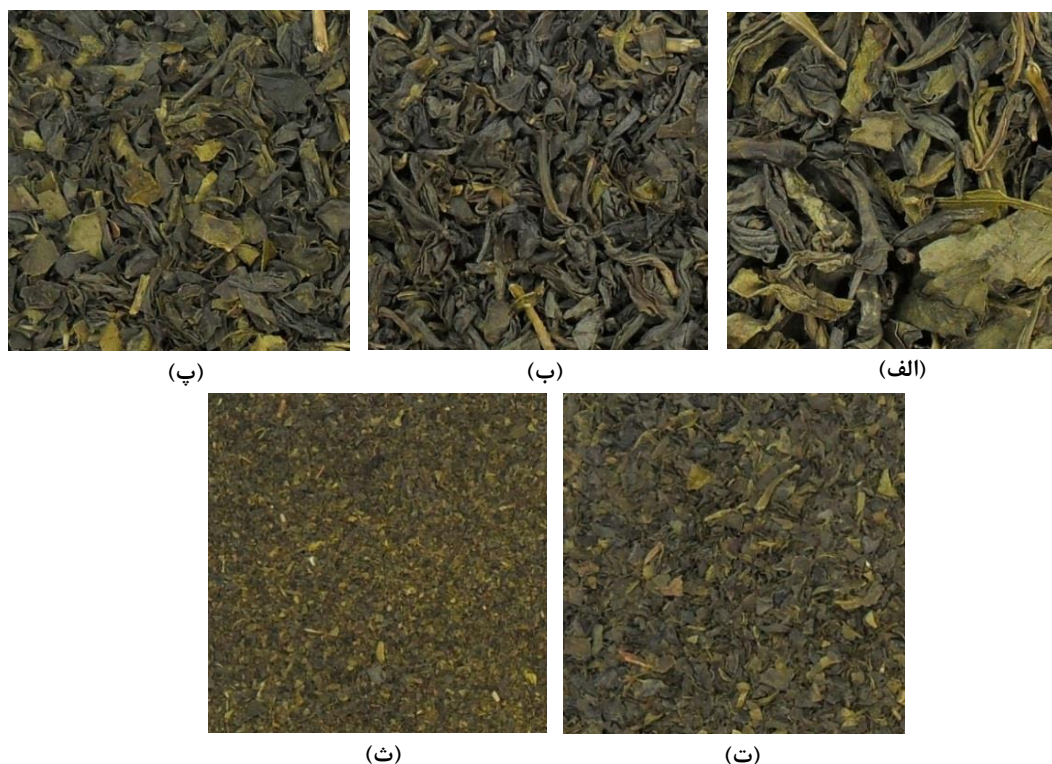
³ Principal Component Analysis

⁴ Least Squares Support Vector Machine

⁵ Red, Green, Blue (RGB)

⁶ Hue, Saturation, Intensity (HIS)

⁷ Grey-Level Co-occurrence Matrix



شکل ۱- تصاویر نمونه‌های چای سبز مورد بررسی در این پژوهش؛ الف) چای سبز قلم، ب) چای سبز شکسته ممتاز، پ) چای سبز شکسته، ت) چای سبز باروتی، ث) چای سبز خاکه

قرار گرفت و تصاویر در حالت فلاش خاموش^۳ و ایزو^۴ ۱۰۰ تهیه شدند. نمونه‌های چای سبز روی یک سطح سفید و مات (مقوای سفید غیرروغنی) قرار داده شدند و دوربین عمود بر سطح نمونه‌ها و در فاصله ۳۰ سانتی‌متری آنها نصب شده و عملیات عکس‌برداری انجام شد. ۸ تصویر رنگی (RGB) با بزرگ‌نمایی یکسان برای هر درجه کیفیت چای سبز تهیه شد.

پردازش تصاویر و استخراج ویژگی‌ها

پس از دریافت عکس‌های گروه‌های مختلف چای سبز در رایانه، تعداد ۶ بلوک تصویر مربعی با ابعاد ۶۰۰×۶۰۰ پیکسل از هرکدام از تصاویر رنگی اولیه جدا شدند (مجموعاً تعداد ۴۸ بلوک تصویر برای هر درجه کیفیت چای سبز) و برای عملیات پردازشی بعدی مورد استفاده قرار گرفتند.

برای هر گروه کیفی چای تعداد ۸ نمونه 150 ± 5 گرمی تهیه شد. به‌منظور پیش‌گیری از ایجاد هرگونه تغییر ظاهری و کیفی در نمونه‌های چای، نمونه‌ها تا زمان داده‌برداری در محفظه‌های پلاستیکی در بسته و در دمای ۲۵-۲۸ درجه سانتی‌گراد نگهداری شدند. همچنین برای جلوگیری از خرد شدن نمونه‌های چای، بسته‌ها از اعمال فشارهای فیزیکی خارجی محافظت شدند.

تهیه تصاویر نمونه‌های چای سبز

بخش‌های اصلی سامانه تصویربرداری استفاده‌شده در این پژوهش شامل اتاقک نورپردازی مناسب برای کنترل شرایط نور محیط در هنگام عکس‌برداری، یک دوربین دیجیتال^۱ CCD (مدل Canon IXUS 960IS، ساخت ژاپن) با وضوح ۱۲ مگاپیکسل برای تهیه تصاویر رنگی (RGB)، و یک رایانه شخصی (PC^۲) به‌منظور دریافت و نمایش تصاویر بود. تنظیمات دوربین روی حالت دستی

^۳ Flash off

^۴ ISO

^۱ Charge-Coupled Device

^۲ Personal Computer

(زیرتصویر تقریب^۵، زیرتصویر بیانگر جزئیات عمودی^۶، زیرتصویر بیانگر جزئیات افقی^۷ و زیرتصویر بیانگر جزئیات قطری^۸) استخراج شدند (شکل ۲-ج). در مرحله بعد، ماتریس‌های هم‌رخداد^۹ برای هرکدام از تصاویر زیرباند موجک محاسبه شدند و ۱۷ ویژگی آماری^{۱۰} پرکاربرد در مطالعه‌های بافتی تصویر، برای هرکدام از تصاویر زیرباند با استفاده از اطلاعات ماتریس‌های هم‌رخداد استخراج شدند.

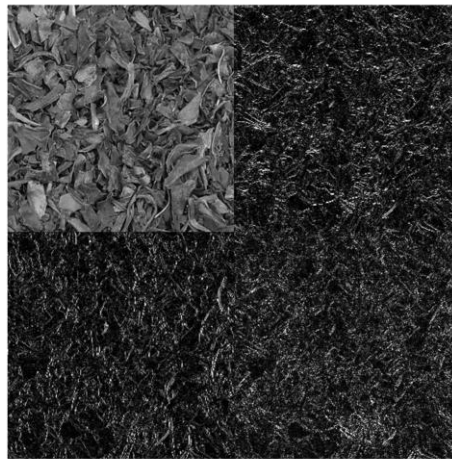
به‌منظور استخراج ویژگی‌های آماری از زیرباندهای تبدیل موجک در تصاویر نمونه‌های چای، از جعبه‌ابزار پردازش تصویر نرم‌افزار MATLAB (نسخه ۲۰۱۷a) استفاده شد. در ابتدا این تصاویر از حالت RGB به تصاویر مقیاس خاکستری^۱ تبدیل شدند (شکل ۲-ب). سپس به‌منظور انتقال تصاویر خاکستری به حوزه فرکانس و استخراج تصاویر موجکی، فیلتر موجک^۲ هار^۲ به‌عنوان ساده‌ترین تبدیل موجک روی تصاویر خاکستری اعمال شد و ۴ زیرتصویر^۳ حاصل از تبدیل موجک مرتبه اول^۴



(ب)



(الف)



(پ)

شکل ۲- الف) تصویر رنگی چای سبز، ب) تصویر مقیاس خاکستری چای سبز، پ) زیرباندهای تبدیل موجک مرتبه اول

⁵ Approximation image

⁶ Vertical coefficients sub-image

⁷ Horizontal coefficients sub-image

⁸ Diagonal coefficients sub-image

⁹ Co-occurrence matrices

¹⁰ Statistical texture features

¹ Grey-scale images

² Haar Wavelet

³ Sub-image

⁴ First level wavelet transform

از ۴ نوع طبقه‌بندی مختلف پرکاربرد به‌منظور طبقه‌بندی چای سبز براساس بردارهای ویژگی استخراج‌شده آنالیز مؤلفه‌های اصلی استفاده شدند که عبارت بودند از درخت تصمیم، شبکه بیزین، شبکه‌های عصبی مصنوعی و ماشین بردار پشتیبان (Mollazade, Omid, & Arefi, 2012; Vlontzos & Pardalos, 2017; Zareiforouh *et al.*, 2016).

درخت تصمیم یک روش داده‌کاوی است که نتایج طبقه‌بندی را در یک ساختار مشابه درخت ارائه می‌دهد. برگ‌های میانی درخت تصمیم نشان‌دهنده متغیرهای مؤثر در طبقه‌بندی و برگ‌های انتهایی نشان‌دهنده گروه‌ها هستند، شاخه‌های درخت تصمیم نیز قوانین را نشان می‌دهند (Maheshwari, Jain, & Khanna, 2015). بسته به نحوه ایجاد ساختار درخت و کاهش خطا، الگوریتم‌های مختلف درخت تصمیم وجود دارند که برخی از این الگوریتم‌ها در این پژوهش به‌منظور طبقه‌بندی نمونه‌های چای استفاده شدند. این الگوریتم‌ها عبارت بودند از کاهش خطای هرس^{۲۲} (REP)، درخت مدل لجستیک^{۲۳} (LMT)، درخت تصادفی^{۲۴} (RT)، جنگل تصادفی^{۲۵} (RF) و J48 ارزیابی شدند. این الگوریتم‌ها در منابع علمی مختلف تشریح شده‌اند (Chandra & Maheshkar, 2017; Granitto *et al.*, 2007; Landwehr, Hall, & Frank, 2017; Rajesh & Karthikeyan, 2005).

ماشین بردار پشتیبان^{۲۶} (SVM) یکی از روش‌های طبقه‌بندی خطی است. این طبقه‌بند بهترین ابرسطحی^{۲۷} را پیدا می‌کند که داده‌های مربوط به کلاس‌های مختلف را با حداکثر فاصله^{۲۸}، از هم تفکیک کند (Wang, 2005). تابع هسته یک بخش کلیدی ماشین بردار پشتیبان است. توابع هسته الگوریتم‌هایی هستند که امکان تعیین روابط کلی بین متغیرها به‌منظور تعیین الگو برای طبقه‌بندی را فراهم می‌کنند (Bordoni *et al.*, 2018). در این پژوهش، الگوریتم‌های خطی، چندجمله‌ای و تابع پایه شعاعی^{۲۹} (RBF) از ماشین بردار پشتیبان برای طبقه‌بندی چای سبز استفاده شدند. اصول و روابط مربوط به این توابع در

این ویژگی‌های آماری عبارت بودند از همبستگی خودکار^۱، کنتراست^۲، آنتروپی^۳، انرژی^۴، همبستگی^۵، همگنی^۶، بی‌شبهتی^۷، گشتاور تفاضل معکوس^۸، حداکثر احتمال^۹، سایه خوشه^{۱۰}، برتری خوشه^{۱۱}، مجموع مربعات^{۱۲}، مجموع میانگین‌ها^{۱۳}، مجموع واریانس‌ها^{۱۴}، مجموع آنتروپی^{۱۵}، تفاضل آنتروپی^{۱۶} و تفاضل واریانس^{۱۷} (Bakhshipour *et al.*, 2018; Gonzalez & Woods, 2002; Haralick & Shanmugam, 1973; Majumdar & Jayas, 1999; Park & Chen, 2001). با توجه به استخراج ۱۷ ویژگی آماری از هر کدام از ۴ زیربند موجک، مجموعاً ۶۸ ویژگی آماری از هر نمونه چای سبز استخراج شدند.

آنالیز داده‌ها

به‌منظور بررسی ارتباط داده‌های بافتی استخراج‌شده از زیرباندهای موجک با درجه کیفی نمونه‌های چای سبز، این داده‌ها در نرم‌افزار یادگیری ماشین وکا^{۱۸} فراخوانی شده و مورد تحلیل قرار گرفتند. از آنجاکه تعداد ویژگی‌های ورودی بسیار زیاد بود، از روش آنالیز مؤلفه‌های اصلی^{۱۹} (PCA) برای کاهش ویژگی^{۲۰} استفاده شد. آنالیز مؤلفه‌های اصلی، ترکیب‌هایی خطی (بردارهای ویژگی^{۲۱}) از ویژگی‌های اولیه را ایجاد می‌کند به‌طوری‌که این بردارهای ویژگی نسبت به هم عمود بوده و حداکثر مقدار تغییر داده‌ها را دربرمی‌گیرند (Jolliffe, 2011). در این پژوهش، با اعمال PCA، تعداد ۸ ترکیب خطی ویژگی تولید شد که برای انجام عملیات طبقه‌بندی توسط طبقه‌بندها مورد استفاده قرار گرفتند.

¹ Autocorrelation

² Contrast

³ Entropy

⁴ Energy

⁵ Correlation

⁶ Homogeneity

⁷ Dissimilarity

⁸ Inverse difference moment

⁹ Maximum Probability

¹⁰ Cluster shade

¹¹ Cluster prominence

¹² Sum of squares

¹³ Sum of averages

¹⁴ Sum of variances

¹⁵ Sum of entropy

¹⁶ Difference of entropy

¹⁷ Difference of variance

¹⁸ WEKA machine learning software

¹⁹ Principal Components Analysis

²⁰ Feature reduction

²¹ Feature vectors

²² Reduced-Error Pruning

²³ Logistic Model Tree

²⁴ Random Tree

²⁵ Random Forest

²⁶ Support Vector Machine

²⁷ Hyper plane

²⁸ Maximum margin

²⁹ Radial Basis Function

لجستیک (LMT)، درخت تصادفی (RT)، جنگل تصادفی (RF) و J48 ارزیابی شدند. باتوجه‌به جدول (۲) الگوریتم درخت تصمیم LMT با مقدار RMSE برابر با ۰/۰۵۸۱ و آماره کاپا برابر با ۰/۹۸۹۶ دارای کمترین میزان خطای طبقه‌بندی و بیشترین میزان انطباق نتایج طبقه‌بند با مشاهده‌های واقعی بود. درخت تصمیم با الگوریتم LMT توانست ۵ گروه چای سبز را با دقت کلی معادل ۹۹/۱۷ درصد طبقه‌بندی نماید که نشان‌دهنده قابلیت بالای درخت تصمیم در طبقه‌بندی کیفی چای براساس خصوصیات بافتی زیرباندهای موجک تصاویر می‌باشد.

جدول ۲ - نتایج طبقه‌بندی الگوریتم‌های مختلف درخت تصمیم برای طبقه‌بندی چای سبز

الگوریتم درخت تصمیم	آماره کاپا	RMSE	دقت (درصد)
REP	۰/۷۸	۰/۲۴	۸۲/۰۸
LMT	۰/۹۹	۰/۰۶	۹۹/۱۷
RT	۰/۷۴	۰/۲۹	۷۹/۱۷
RF	۰/۹۰	۰/۱۹	۹۲/۰۸
J48	۰/۸۱	۰/۲۴	۸۴/۵۸

ماتریس اغتشاش طبقه‌بند LMT در جدول (۳) ارائه شده است. همان‌طوری‌که از جدول (۳) مشخص است، ۴ گروه کیفی چای سبز قلم، شکسته ممتاز، باروتی و خاکه، کاملاً از سایر گروه‌ها جداسازی شدند. درحالی‌که تنها در حدود ۴ درصد از نمونه‌های چای سبز شکسته به اشتباه در گروه‌های چای سبز شکسته ممتاز و چای سبز قلم طبقه‌بندی شدند.

جدول ۳ - ماتریس اغتشاش طبقه‌بند درخت تصمیم با الگوریتم LMT در طبقه‌بندی ۵ گروه چای سبز (داده‌ها به درصد)

درجه کیفی	GD	GBOPF	GBOP	GFBOP	GOP
GOP	۰	۰	۰	۰	۱۰۰
GFBOP	۰	۰	۰	۱۰۰	۰
GBOP	۰	۰	۹۵/۸۴	۲/۰۸	۲/۰۸
GBOPF	۰	۱۰۰	۰	۰	۰
GD	۱۰۰	۰	۰	۰	۰

مقاله‌ها توضیح داده شده‌اند (Ge et al., 2014; Yao et al., 2009).

طبقه‌بند شبکه بیزین نیز یک طبقه‌بند با ناظر است که از قانون شرط بیز برای جداسازی گروه‌های مختلف استفاده می‌کند. توضیحات جامع در مورد این الگوریتم طبقه‌بند در منابع علمی ارائه شده است (Garg, Pavlovic, & Huang, 2002; Ouali, Cherif, & Krebs, 2006).

برای ارزیابی طبقه‌بندها از اعتبارسنجی متقابل ۱۰ برابری^۱ استفاده گردید (Silva, Koga, Cugnasca, & Costa, 2013) و به‌منظور مقایسه روش‌های فوق از دو معیار ارزیابی، آماره کاپا^۲ و مجذور میانگین مربعات خطا^۳ (RMSE) استفاده شدند. آماره کاپا میزان توافق بین نتایج طبقه‌بند و داده‌های تجربی را نشان می‌دهد (Dissing et al., 2013). مقدار RMSE نیز یک معیار پرکاربرد است که میزان تفاوت بین نتایج یک مدل پیش‌بینی با مقادیر واقعی مشاهده‌شده را نمایش می‌دهد (Stojanova, Panov, Gjorgjioski, Kobler, & Džeroski, 2010). همچنین میزان دقت کلی طبقه‌بندی^۴ برای طبقه‌بندها براساس داده‌های ماتریس اغتشاش^۵ و با استفاده از رابطه (۱) محاسبه شدند:

رابطه (۱)

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

در رابطه (۱)؛ TP ، FP ، TN و FN به‌ترتیب عبارت بودند از مقادیر مثبت صحیح^۶، مثبت کاذب^۷، منفی صحیح^۸ و منفی کاذب^۹ در ماتریس اغتشاش می‌باشد.

نتایج و بحث

در جدول (۲) مقادیر آماره کاپا، RMSE و نیز دقت طبقه‌بندی ۵ گروه کیفی چای سبز توسط طبقه‌بند درخت تصمیم ارائه شده‌اند. ۵ الگوریتم مختلف از درخت تصمیم با نام‌های کاهش خطای هرس (REP)، درخت مدل

¹ 10-fold cross validation

² Kappa statistics

³ Root Mean Squared Error

⁴ Accuracy

⁵ Confusion matrix

⁶ True Positive

⁷ False Positive

⁸ True Negative

⁹ False Negative

باتوجه به آماره‌های کاپا و RMSE، شبکه عصبی پرسپترون چندلایه عملکرد بهتری را نسبت به شبکه تابع با پایه شعاعی داشت. دقت شبکه عصبی پرسپترون چندلایه برای جداسازی ۵ گروه چای سبز برابر با ۹۹/۱۷ درصد مشاهده شد. این دقت برابر با بیشترین دقت طبقه‌بندی درخت تصمیم LMT بود. هرچند مقادیر آماره‌های کاپا و RMSE مربوط به شبکه عصبی MLP بهتر از درخت تصمیم LMT بودند. ماتریس اغتشاش طبقه‌بند MLP در جدول (۶) ارائه شده است. در جدول (۶) مشاهده می‌شود که طبقه‌بند شبکه عصبی پرسپترون چندلایه در حدود ۲ درصد از نمونه‌های چای سبز قلم را به اشتباه به‌عنوان چای سبز شکسته طبقه‌بندی کرده و در حدود ۲ درصد از نمونه‌های چای سبز شکسته را نیز به‌عنوان چای سبز قلم طبقه‌بندی نموده است. سایر گروه‌های کیفی با دقت ۱۰۰ درصد جداسازی شده‌اند.

جدول ۶- ماتریس اغتشاش طبقه‌بند شبکه عصبی مصنوعی MLP در طبقه‌بندی ۵ گروه چای سبز (داده‌ها به درصد)

درجه کیفی	GD	GBOPF	GBOP	GFBOP	GOP
GD	۱۰۰	۰	۰	۰	۰
GFBOP	۰	۰	۰	۱۰۰	۰
GBOP	۰	۰	۹۷/۹۲	۰	۲/۰۸
GBOPF	۰	۱۰۰	۰	۰	۰
GOP	۰	۰	۲/۰۸	۰	۹۷/۹۲

در ادامه، نتایج حاصل از طبقه‌بندی ۵ گروه چای سبز با استفاده از الگوریتم‌های خطی، چندجمله‌ای و تابع پایه شعاعی در ماشین‌بردار پشتیبان، در جدول (۷) ارائه شده‌اند (Schölkopf, Smola, & Bach, 2002). اصول تئوری این روش‌ها در منابع تشریح شده‌اند. باتوجه به مقادیر آماره‌های کاپا و RMSE، بهترین عملکرد در بین ۳ الگوریتم ماشین‌بردار پشتیبان متعلق به حالت خطی می‌باشد (RMSE=۰/۰۷ و کاپا=۰/۹۸). این طبقه‌بند توانست ۵ گروه چای سبز را با دقتی برابر با ۹۸/۷۵ درصد طبقه‌بندی نماید.

۵ الگوریتم مختلف شبکه بیزین با نام‌های تپه نوردی^۱ (HC)، شبیه‌سازی پروت^۲ (SA)، جستجوی تپه^۳ (TS)، K2 و بیزین ساده با درخت افزوده^۴ (TAN) مورد ارزیابی قرار گرفتند. اطلاعات کامل در مورد این الگوریتم‌ها توسط Bouckaert (۲۰۰۴) ارائه شده است. باتوجه به مقادیر پارامترهای ارزیابی (جدول ۴)، بهترین نتایج شبکه بیزین با الگوریتم‌های SA و TAN به دست آمد (دقت طبقه‌بندی برابر با ۷۹/۱۷ درصد). این نتایج نشان می‌دهد که شبکه بیزین مدل مناسبی برای طبقه‌بندی چای سبز براساس خصوصیات تصویر نمی‌باشد.

جدول ۴- نتایج طبقه‌بندی الگوریتم‌های مختلف شبکه بیزین برای طبقه‌بندی چای سبز

دقت (درصد)	RMSE	آماره کاپا	الگوریتم شبکه بیزین
۷۲/۰۹	۰/۲۷	۰/۶۵	HC
۷۹/۱۷	۰/۲۵	۰/۷۴	SA
۷۲/۰۸	۰/۲۷	۰/۶۵	TS
۷۲/۰۸	۰/۲۷	۰/۶۵	K2
۷۹/۱۷	۰/۲۵	۰/۷۴	TAN

دو ساختار مختلف شبکه عصبی مصنوعی برای طبقه‌بندی چای سبز استفاده شد که عبارت بودند از شبکه‌های عصبی پرسپترون چندلایه^۵ (MLP) و تابع پایه شعاعی (RBF). مفاهیم این دو نوع شبکه منابع مختلف توضیح داده شده‌اند (Alrajeh & Alzohairy, 2012). نتایج طبقه‌بندی این دو شبکه در جدول (۵) نمایش داده شده‌اند.

جدول ۵- نتایج طبقه‌بندی ساختارهای مختلف شبکه‌های عصبی مصنوعی برای طبقه‌بندی چای سبز

دقت (درصد)	RMSE	آماره کاپا	توپولوژی	نوع شبکه عصبی مصنوعی
۹۹/۱۷	۰/۰۴	۰/۹۹	۸-۱۴-۵	MLP
۹۱/۶۷	۰/۱۸	۰/۹۰	۸-۹-۵	RBF

¹ Hill Climbing

² Simulated Annealing

³ Tabu Search

⁴ Tree Augmented Naive Bayes

⁵ Multi-Layer Perceptron

جدول ۷- نتایج طبقه‌بندی الگوریتم‌های مختلف ماشین‌بردار پشتیبان برای طبقه‌بندی چای سبز

دقت (درصد)	RMSE	آماره کاپا	الگوریتم درخت تصمیم
۹۸/۷۵	۰/۰۷	۰/۹۸	خطی
۹۷/۵۰	۰/۱۰	۰/۹۷	چندجمله‌ای
۸۲/۹۲	۰/۲۴	۰/۷۹	RBF

اشتباهی طبقه‌بندی شده‌اند که مقدار خطای طبقه‌بندی کمی را نشان می‌دهد. باتوجه به این نتایج، عملکرد طبقه‌بند ماشین‌بردار پشتیبان نسبت به شبکه عصبی مصنوعی و درخت تصمیم، در جداسازی ۵ گروه چای سبز، ضعیف‌تر می‌باشد. همچنین، باتوجه به جدول‌های (۲) تا (۸)، سه روش درخت تصمیم، شبکه‌های عصبی مصنوعی و ماشین‌بردار پشتیبان توانستند با دقت بالایی گروه‌های مختلف چای سبز را براساس ویژگی‌های بافت زیرباند‌های تصویر، طبقه‌بندی نمایند. طبقه‌بند بیزین روش مناسبی برای طبقه‌بندی کیفی چای سبز نبود. همچنین باتوجه به مقادیر دقت و آماره‌های کاپا و RMSE، روش شبکه عصبی مصنوعی پرسپترون چندلایه به‌عنوان بهترین طبقه‌بند در این مطالعه انتخاب شد.

باتوجه به ماتریس اغتشاش مربوط به طبقه‌بند ماشین‌بردار پشتیبان با الگوریتم خطی که در جدول (۸) ارائه شده است، مشخص است که تنها گروه‌های چای سبز باروتی و خاکه کاملاً جداسازی شده‌اند و در حدود ۲ درصد از مقادیر هرکدام از سایر نمونه‌ها در گروه‌های

جدول ۸- ماتریس اغتشاش طبقه‌بند ماشین‌بردار پشتیبان با الگوریتم خطی در طبقه‌بندی ۵ گروه چای سبز (داده‌ها به درصد)

درجه کیفی	GOP	GFBOP	GBOP	GBOPF	GD
GOP	۹۷/۹۲	۰	۲/۰۸	۰	۰
GFBOP	۰	۹۷/۹۲	۲/۰۸	۰	۰
GBOP	۲/۰۸	۰	۹۷/۹۲	۰	۰
GBOPF	۰	۰	۰	۱۰۰	۰
GD	۰	۰	۰	۰	۱۰۰

نتیجه‌گیری

در این پژوهش، از تکنیک‌های پردازش تصویر به‌منظور طبقه‌بندی درجه‌های کیفی مختلف چای سبز براساس ویژگی‌های بافتی استخراج‌شده از زیرباند‌های تبدیل موجک تصاویر استفاده شد. ۴ نوع مختلف از روش‌های طبقه‌بندی با الگوریتم‌های عملکردی متفاوت موردارزیابی قرار گرفتند. نتایج نشان داد که الگوریتم‌هایی از شبکه‌های عصبی مصنوعی، ماشین‌بردار پشتیبان و درخت تصمیم با دقت بالایی موفق به طبقه‌بندی کیفی چای سبز شدند.

درحالی‌که شبکه بیزین عملکرد قابل‌قبولی نداشت. باتوجه به آماره‌های ارزیابی، شبکه عصبی مصنوعی پرسپترون چندلایه به‌عنوان بهترین طبقه‌بند انتخاب شد (آماره کاپا برابر با ۰/۹۹، مقدار RMSE برابر با ۰/۰۴ و دقت طبقه‌بندی برابر با ۹۹/۱۷ درصد). باتوجه به نتایج این پژوهش، استفاده از ماشین بینایی و به‌کارگیری ویژگی‌های بافتی مستخرج از زیرباند‌های موجک تصاویر چای سبز، معیارهای مناسبی برای طبقه‌بندی کیفی چای سبز می‌باشد.

منابع

- ایزدی، ه. (۱۳۹۲). توسعه فناوری و ارزیابی سامانه تشخیص عیوب و درجه‌بندی گوجه‌فرنگی با استفاده از فناوری ماشین بینایی و شبکه‌های عصبی- فازی (ANFIS). (پایان نامه کارشناسی ارشد)، انشگاه شیراز، دانشکده کشاورزی، وزارت جهاد کشاورزی- معاونت برنامه‌ریزی و اقتصادی. (۱۳۹۶). گزارش وضعیت صنعت چای کشور (۲۰ صفحه). برگرفته از http://facility.maj.ir/Dorsapax/Data/Sub_46/File/3.pdf
- Al Ohali, Y. (2011). Computer vision based date fruit grading system: Design and implementation. *Journal of King Saud University- Computer and Information Sciences*, 23(1), 29-36. doi:<https://doi.org/10.1016/j.jksuci.2010.03.003>

- Alrajeh, K. M., & Alzohairy, T. A. (2012). Date fruits classification using MLP and RBF neural networks. *International Journal of Computer Applications*, 41(10), 36-41. doi:<https://doi.org/10.5120/5579-7686>
- Bakhshipour, A., Sanaeifar, A., Payman, S. H., & de la Guardia, M. (2018). Evaluation of Data Mining Strategies for Classification of Black Tea Based on Image-Based Features. *Food analytical methods*, 11(4), 1041-1050. doi:<https://doi.org/10.1007/s12161-017-1075-z>
- Benalia, S., Cubero, S., Prats-Montalbán, J. M., Bernardi, B., Zimbalatti, G., & Blasco, J. (2016). Computer vision for automatic quality inspection of dried figs (*Ficus carica* L.) in real-time. *Computers and electronics in agriculture*, 120, 17-25. doi:<https://doi.org/10.1016/j.compag.2015.11.002>
- Bordoni, M., Bittelli, M., Valentino, R., Chersich, S., Persichillo, M., & Meisina, C. (2018). Soil Water Content Estimated by Support Vector Machine for the Assessment of Shallow Landslides Triggering: the Role of Antecedent Meteorological Conditions. *Environmental Modeling & Assessment*, 23(4), 333-352. doi:<https://doi.org/10.1007/s10666-017-9586-y>
- Bouckaert, R. R. (2004). Bayesian network classifiers in weka.
- Chandra, S., & Maheshkar, S. (2017). Verification of static signature pattern based on random subspace, REP tree and bagging. *Multimedia Tools and Applications*, 76(18), 19139-19171. doi:<https://doi.org/10.1007/s11042-017-4531-2>
- Chen, Q., Zhao, J., Chen, Z., Lin, H., & Zhao, D.-A. (2011). Discrimination of green tea quality using the electronic nose technique and the human panel test, comparison of linear and nonlinear classification tools. *Sensors and Actuators B: Chemical*, 159(1), 294-300. doi:<https://doi.org/10.1016/j.snb.2011.07.009>
- Diniz, P. H. G. D., Barbosa, M. F., de Melo Milanez, K. D. T., Pistonesi, M. F., & de Araújo, M. C. U. (2016). Using UV-Vis spectroscopy for simultaneous geographical and varietal classification of tea infusions simulating a home-made tea cup. *Food chemistry*, 192, 374-379. doi:<https://doi.org/10.1016/j.foodchem.2015.07.022>
- Diniz, P. H. G. D., Pistonesi, M. F., Alvarez, M. B., Band, B. S. F., & de Araújo, M. C. U. (2015). Simplified tea classification based on a reduced chemical composition profile via successive projections algorithm linear discriminant analysis (SPA-LDA). *Journal of Food Composition and Analysis*, 39, 103-110. doi:<https://doi.org/10.1016/j.jfca.2014.11.012>
- Dissing, B. S., Papadopoulou, O. S., Tassou, C., Ersbøll, B. K., Carstensen, J. M., Panagou, E. Z., & Nychas, G.-J. (2013). Using multispectral imaging for spoilage detection of pork meat. *Food and Bioprocess Technology*, 6(9), 2268-2279. doi:<https://doi.org/10.1007/s11947-012-0886-6>
- Ezadi, H. (2013). *Development of technology and evaluation of defect detection and tomato grading system using visual machine technology and neuro-fuzzy networks (ANFIS)*. (Unpublished master's degree in Agricultural Machinery Mechanics), Shiraz University, (in Persian)
- Farooqi, S. (2012). Data mining: An overview. *Indian Agricultural Statistics Research Institute (IASRI), Library Avenue, Pusa, New Delhi-110012*.
- Garg, A., Pavlovic, V., & Huang, T. S. (2002). *Bayesian networks as ensemble of classifiers*. Paper presented at the Object recognition supported by user interaction for service robots, 11-15 Aug. 2002.
- Ge, H., Jiang, Y., Xu, Z., Lian, F., Zhang, Y., & Xia, S. (2014). Identification of wheat quality using THz spectrum. *Optics express*, 22(10), 12533-12544. doi:<https://doi.org/10.1364/OE.22.012533>
- Gonzalez, R. C., & Woods, R. E. (2002). Digital Image Processing, Publishing House of Electronics Industry. *Beijing*, 295-300.
- Granitto, P. M., Biasioli, F., Aprea, E., Mott, D., Furlanello, C., Märk, T. D., & Gasperi, F. (2007). Rapid and non-destructive identification of strawberry cultivars by direct PTR-MS headspace analysis and data mining techniques. *Sensors and Actuators B: Chemical*, 121(2), 379-385. doi:<https://doi.org/10.1016/j.snb.2006.03.047>
- Haralick, R. M., & Shanmugam, K. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 3(6), 610-621. doi:<https://doi.org/10.1109/TSMC.1973.4309314>
- Hayat, K., Iqbal, H., Malik, U., Bilal, U., & Mushtaq, S. (2015). Tea and its consumption: benefits and risks. *Critical reviews in food science and nutrition*, 55(7), 939-954. doi:<https://doi.org/10.1080/10408398.2012.678949>
- He, X., Li, J., Zhao, W., Liu, R., Zhang, L., & Kong, X. (2015). Chemical fingerprint analysis for quality control and identification of Ziyang green tea by HPLC. *Food chemistry*, 171, 405-411. doi:<https://doi.org/10.1016/j.foodchem.2014.09.026>
- Jolliffe, I. (2011). *Principal component analysis*. International encyclopedia of statistical science: Springer, Berlin, Heidelberg.
- Landwehr, N., Hall, M., & Frank, E. (2005). Logistic model trees. *Machine learning*, 59(1-2), 161-205. doi:<https://doi.org/10.1007/s10994-005-0466-3>
- Liang, Y., Lu, J., Zhang, L., Wu, S., & Wu, Y. (2005). Estimation of tea quality by infusion colour difference analysis. *Journal of the Science of Food and Agriculture*, 85(2), 286-292. doi:<https://doi.org/10.1002/jsfa.1953>
- Liming, X., & Yanchao, Z. (2010). Automated strawberry grading system based on image processing. *Computers and electronics in agriculture*, 71(1), S32-S39. doi:<https://doi.org/10.1016/j.compag.2009.09.013>
- Maheshwari, C. V., Jain, N. K., & Khanna, S. (2015). Computer Vision Based Classification of Indian Gujarat-17 Rice Using Geometrical Features and Cart. In L. Jain, H. Behera, J. Mandal, & D. Mohapatra (Eds.), *Computational Intelligence in Data Mining-Volume 3. Smart Innovation, Systems and Technologies*, (Vol. 33, pp. 205-216): Springer, New Delhi.
- Majumdar, S., & Jayas, D. (1999). Classification of bulk samples of cereal grains using machine vision. *Journal of Agricultural Engineering Research*, 73(1), 35-47. doi:<https://doi.org/10.1006/jaer.1998.0388>

- Mery, D., Pedreschi, F., & Soto, A. (2013). Automated design of a computer vision system for visual food quality evaluation. *Food and Bioprocess Technology*, 6(8), 2093-2108. doi:<https://doi.org/10.1007/s11947-012-0934-2>
- Ministry of Agriculture-Jahad: Planning & Economic Affairs. (2017). State of the Tea Industry Country Report. pp. 20. Retrieved from http://facility.maj.ir/Dorsapax/Data/Sub_46/File/3.pdf (in Persian)
- Mollazade, K., Omid, M., & Arefi, A. (2012). Comparing data mining classifiers for grading raisins based on visual features. *Computers and electronics in agriculture*, 84, 124-131. doi:<https://doi.org/10.1016/j.compag.2012.03.004>
- Nouri-Ahmadabadi, H., Omid, M., Mohtasebi, S. S., & 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.
- Ouali, A., Cherif, A. R., & Krebs, M.-O. (2006). Data mining based Bayesian networks for best classification. *Computational statistics & data analysis*, 51(2), 1278-1292.
- Park, B., & Chen, Y. (2001). Co-occurrence matrix texture features of multi-spectral images on poultry carcasses. *Journal of Agricultural Engineering Research*, 78(2), 127-140.
- Payman, S., Bakhshipour, A., & Zareiforouh, H. (2018). Development of an expert vision-based system for inspecting rice quality indices. *Quality Assurance and Safety of Crops & Foods*, 10(1), 103-114.
- Rajesh, P., & Karthikeyan, M. (2017). A comparative study of data mining algorithms for decision tree approaches using WEKA tool. *Advances in Natural and Applied Sciences*, 11(9), 230-243.
- Roy, R. B., Chattopadhyay, P., Tudu, B., Bhattacharyya, N., & Bandyopadhyay, R. (2014). Artificial flavor perception of black tea using fusion of electronic nose and tongue response: A Bayesian statistical approach. *Journal of Food Engineering*, 142, 87-93. doi:<https://doi.org/10.1016/j.jfoodeng.2014.06.004>
- Sabanci, K., Toktas, A., & Kayabasi, A. (2017). Grain classifier with computer vision using adaptive neuro-fuzzy inference system. *Journal of the Science of Food and Agriculture*, 97(12), 3994-4000. doi:<https://doi.org/10.1002/jsfa.8264>
- Sanaeifar, A., Mohtasebi, S. S., Ghasemi-Varnamkhasti, M., & Siadat, M. (2014, 3-5 Nov. 2014). *Application of an electronic nose system coupled with artificial neural network for classification of banana samples during shelf-life process*. Paper presented at the 2014 International Conference on Control, Decision and Information Technologies (CoDIT).
- Schölkopf, B., Smola, A. J., & Bach, F. (2002). *Learning with kernels: support vector machines, regularization, optimization, and beyond*: MIT press.
- Semary, N. A., Tharwat, A., Elhariri, E., & Hassanien, A. E. (2015). Fruit-based tomato grading system using features fusion and support vector machine. In F. D. e. al. (Ed.), *Intelligent Systems' 2014. Advances in Intelligent Systems and Computing* (Vol. 323, pp. 401-410): Springer.
- Silva, L., Koga, M., Cugnasca, C., & Costa, A. (2013). Comparative assessment of feature selection and classification techniques for visual inspection of pot plant seedlings. *Computers and electronics in agriculture*, 97, 47-55. doi:<https://doi.org/10.1016/j.compag.2013.07.001>
- Stojanova, D., Panov, P., Gjorgjioski, V., Kobler, A., & Džeroski, S. (2010). Estimating vegetation height and canopy cover from remotely sensed data with machine learning. *Ecological Informatics*, 5(4), 256-266. doi:<https://doi.org/10.1016/j.ecoinf.2010.03.004>
- Velásquez, L., Cruz-Tirado, J., Siche, R., & Quevedo, R. (2017). An application based on the decision tree to classify the marbling of beef by hyperspectral imaging. *Meat science*, 133, 43-50. doi:<https://doi.org/10.1016/j.meatsci.2017.06.002>
- Vlontzos, G., & Pardalos, P. M. (2017). Data mining and optimisation issues in the food industry. *International Journal of Sustainable Agricultural Management and Informatics*, 3(1), 44-64. doi:<https://doi.org/10.1504/IJSAMI.2017.082921>
- Wang, C.-F., & Liu, K. (2015). Learning Bayesian network classifier based on artificial fish swarm algorithm. *IAENG International Journal of Computer Science*, 42(4), 355-360.
- Wang, L. (2005). *Support vector machines: theory and applications* (Vol. 177): Springer Science & Business Media.
- Yao, Q., Guan, Z., Zhou, Y., Tang, J., Hu, Y., & Yang, B. (2009, 2-3 May 2009). *Application of support vector machine for detecting rice diseases using shape and color texture features*. Paper presented at the 2009 international conference on engineering computation.
- Yu, H., Wang, J., Yao, C., Zhang, H., & Yu, Y. (2008). Quality grade identification of green tea using E-nose by CA and ANN. *LWT-Food Science and Technology*, 41(7), 1268-1273. doi:<https://doi.org/10.1016/j.lwt.2007.08.018>
- Yu, X. J., Liu, K. S., He, Y., & Wu, D. (2011). *Color and texture classification of green tea using least squares support vector machine (lssvm)*. Paper presented at the Key Engineering Materials.
- Zareiforouh, H., Minaei, S., Alizadeh, M. R., & Banakar, A. (2016). Qualitative classification of milled rice grains using computer vision and metaheuristic techniques. *Journal of food science and technology*, 53(1), 118-131. doi:<https://doi.org/10.1007/s13197-015-1947-4>

Application of Image Wavelet Transform for Qualitative Classification of Green Tea Using Metaheuristic Algorithms

Adel Bakhshipour¹, Hemad Zareiforoush^{1*}, Iraj Bagheri¹

1- Assistant Professor, Department of Mechanization Engineering, Faculty of Agricultural Sciences, University of Guilan, Rasht, Iran

* Corresponding author (hemad.zareiforoush@guilan.ac.ir)

Abstract

This study was aimed to investigate the best features extracted from images to determine the best technique for qualitative classification of green tea by using meta-heuristic algorithms. Five different classes of green tea were evaluated according to the standards of the Institute of Standards and Industrial Research of Iran. After receiving the images of different green tea classes in the computer, 6 square blocks were extracted from each of the original color images. These image blocks were transformed from RGB to gray scale images. One-level discrete Haar wavelet filter was applied to the gray images and 4 wavelet subimages were obtained. Co-occurrence matrices were calculated for each wavelet subimages and 17 common texture features in the image textural studies, were extracted from subimages (totally 68 texture features for each block image). By using principal component analysis, 8 feature components were produced from the original features and used for the separation of 5 groups of green tea. The results showed that algorithms of artificial neural networks, support vector machine and decision tree were capable of qualitative classification of green tea with high accuracy. However, Bayesian network did not have acceptable performance. According to the evaluation statistics, the multilayer perceptron artificial neural networks (with Kappa statistic, root mean square error and classification accuracy of 0.90, 0.42, and 99.17%, respectively) was the best classifier. Based on the results of this study, the use of machine vision and texture features extracted from image wavelet subimages is a suitable technique for the qualitative classification of green tea.

Keywords: Image Processing, Image texture, Meta-heuristic Algorithms, Wavelet transform