

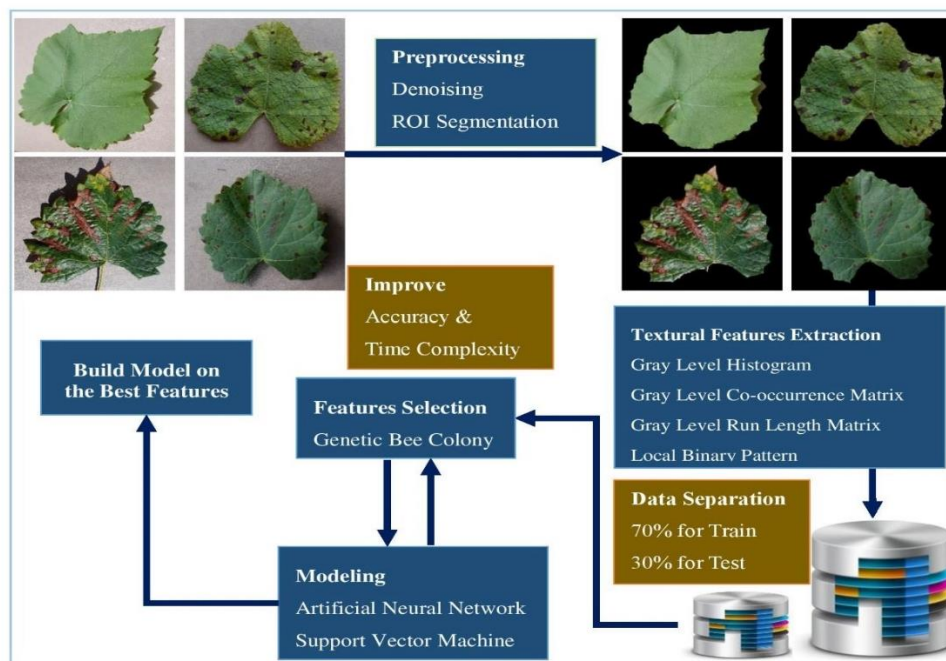
Image Processing Based Method for Automatic Detection of Grape leaf Diseases

Sajjad Nasiri¹, Mostafa Khojastehnazhand^{1*}

1. Computer Engineering Department, Faculty of Engineering, University of Bonab, Bonab, Iran

2. Mechanical Engineering Department, Faculty of Engineering, University of Bonab, Bonab, Iran

(Received: Oct. 15, 2021- Revised: Feb. 4, 2022- Accepted: Feb. 9, 2022)



ABSTRACT: Rapid detection and prevention of disease spread in agricultural products can significantly reduce losses and costs of disease control. In this study, an intelligent system based on image processing method has been presented for detection of grape (*Sultana - Vitis vinifera*) leaf diseases. For this purpose, different image texture features were extracted from the Gray Level histogram (GLH), Gray Level Co-occurrence Matrix (GLCM), Gray Level Run Length Matrix (GLRM) and Local Binary Pattern (LBP) algorithms. Two models of Artificial Neural Network (ANN) and Support Vector Machine (SVM) were used to model the features. The dataset consists of 4062 images including healthy leaves, Black Rot, Esca and Isariopsis leaves. The results showed that the SVM model based on GLRM features with an average accuracy of 89.70% showed the best performance. The results also showed that the use of all extracted features as a single feature vector increases the accuracy of classification. The accuracy of the SVM and ANN models using all of the features for training data were 91.10%, 95.04%, and for the test data were 89.93% and 91.75%, respectively. Finally, using Genetic Bee Colony (GBC) algorithm and reducing the number of features to 34 and 46 for ANN and SVM models, respectively, the average accuracy of 97.20% and 94.10% for training and testing of ANN model and 93.01% and 92.33% for training and testing of SVM model were obtained, which shows the improvement of results by GBC algorithm. The proposed method was evaluated as efficient in diagnosing grape leaf diseases.

Keywords: Machine learning, Image texture analysis, Black rot, Esca and Isariopsis

* Corresponding Author's Email: khajasteh@ubonab.ac.ir

روش مبتنی بر پردازش تصویر به منظور تشخیص خودکار بیماری برگ درخت انگور

سجاد نصیری^۱، مصطفی خجسته‌نژاد^{۱*}

۱. گروه مهندسی کامپیوتر، دانشکده فنی و مهندسی، دانشگاه بناب، بناب، ایران

۲. گروه مهندسی مکانیک، دانشکده فنی و مهندسی، دانشگاه بناب، بناب، ایران

(تاریخ دریافت: ۱۴۰۰/۷/۲۳ - تاریخ بازنگری: ۱۴۰۰/۱۱/۱۵ - تاریخ تصویب:

۱۴۰۰/۱۱/۱۹)

چکیده: تشخیص سریع و پیشگیری از گسترش بیماری محصولات کشاورزی، می‌تواند تلفات مقابله با بیماری را به میزان قابل توجهی کاهش دهد. در این پژوهش، سامانه‌ای هوشمند بر مبنای پردازش تصویر برای تشخیص بیماری‌های برگ درخت انگور (*Sultana - Vitis vinifera*) ارائه گردیده است. بدین منظور، ویژگی‌های مختلف بافت تصویر از هیستوگرام سطح خاکستری (GLH)، ماتریس هم-رخداد سطح خاکستری (GLCM)، ماتریس طول بردار سطح خاکستری (GLRM) و الگوی دودویی محلی (LBP) استخراج شد. برای مدل‌سازی ویژگی‌ها، از دو مدل شبکه عصبی (ANN) و ماشین بردار پشتیبان (SVM) استفاده شد. پایگاه داده مورد استفاده، متشکل از ۴۰۶۲ تصویر، شامل برگ سالم، مبتلا به پوسیدگی سیاه، اسکا و لکه ایزاریوپسیس است. نتایج نشان دادند که مدل SVM با استفاده از ویژگی‌های GLRM با متوسط دقت ۸۹/۷۰٪ بهترین عملکرد را از خود نشان داد. همچنین نتایج نشان دادند، استفاده از تمام ویژگی‌های استخراج یافته به صورت بردار ویژگی واحد، افزایش دقت دسته‌بندی را به دنبال دارد. مدل SVM و ANN با استفاده از تمام ویژگی‌ها بترتیب برای داده‌های آموزشی دقت ۹۱/۱۰٪، ۹۵/۰۴٪ و برای داده‌های آزمون میزان دقت ۸۹/۹۳٪ و ۹۱/۷۵٪ را نتیجه دادند. در نهایت، با استفاده از الگوریتم کلونی زنبور ژنتیکی (GBC) و کاهش تعداد ویژگی‌ها به ۳۴ و ۴۶ به ترتیب برای مدل‌های ANN و SVM میانگین دقت ۹۷/۲۰٪ و ۹۴/۱۰٪ برای آموزش و آزمون مدل ANN و ۹۳/۰۱٪ و ۹۲/۳۳٪ برای آموزش و آزمون مدل SVM به دست آمد که نشان دهنده بهبود نتایج توسط الگوریتم GBC می‌باشد. روش پیشنهادی در تشخیص بیماری‌های برگ انگور کارآمد ارزیابی شد.

واژه‌های کلیدی: یادگیری ماشین، تحلیل بافت تصویر، پوسیدگی سیاه انگور، اسکای انگور، لکه ایزاریوپسیس.

مقدمه

انگور به دلیل عرضه به صورت خام و سایر فرآورده‌های تبدیلی، از جمله کشمش، غوره، نوشیدنی و شیر، به یکی از محصولات مهم و تجاری کشاورزی تبدیل شده است. این محصول به دلیل مواد قندی و درصد رطوبت بالا، مستعد ابتلا به بیماری‌ها، آفات و فسادهای میکروبی می‌باشد که برخی از این بیماری‌ها مربوط به دوران رشد درخت انگور می‌باشد. در این دوران، تغییرات بافتی و رنگی در برگ درختان به دلیل وجود بیماری و آفت ایجاد می‌شود (Barbedo, 2016; 2019). از جمله بیماری‌های درختان انگور می‌توان به پوسیدگی سیاه^۱، اسکا^۲ و لکه ایزاریوپسیس^۳ اشاره نمود. این بیماری‌ها به دلیل سرعت گسترش بالا، باعث آسیب‌های فراوانی می‌شوند. از طرفی مقابله و سرکوب بیماری نیز مستلزم هزینه‌های بسیاری است (Scheck et al., 1998a; 1998b; Ye et al., 2021).

در حال حاضر، تشخیص این بیماری‌ها به روش سنتی و توسط افراد خبره انجام می‌گیرد که این روش مستعد خطا، پرهزینه و زمان‌بر است. علاوه بر این، وجود فرد خبره در تمام مناطق زیر کشت انگور ناممکن است. از طرف دیگر، شناسایی سریع و به موقع و البته دقیق بیماری، در کاهش هزینه‌ها و تلفات بسیار مهم و حیاتی است. بنابراین ایجاد روشی برای تشخیص خودکار درختان مبتلا به بیماری از اولویت زیادی برخوردار است. در سال‌های اخیر، روش‌های هوشمند و غیرمخرب بسیاری بر مبنای پردازش تصاویر دیجیتال برای حل این مسأله و مسائل مشابه رواج پیدا کرده‌اند. برای طبقه‌بندی بیماری‌های سیب از ماشین بردار پشتیبان^۴ (SVM) و ویژگی‌های رنگ، بافت و شکل، شامل هیستوگرام گابور^۵، بردار وابستگی رنگی^۶، الگوی دودویی

محلی^۷ (LBP) استفاده شده است (Dubey & Jalal, 2016). نتایج طبقه‌بندی نشان داد که استفاده از ترکیب ویژگی‌های یاد شده، دقت تشخیص را بهبود می‌بخشد. در تحقیقات دیگری نیز با استفاده از ویژگی‌های گرادیان رنگ و ویژگی‌های بافت استخراج شده از ماتریس‌های هم-رخداد سطح خاکستری^۸ (GLCM)، ماتریس طول بردار پیکسل^۹ (PRL) و الگوی دودویی محلی محصول انار و انبه، مدل طبقه‌بندی کننده ماشین بردار پشتیبان ارائه گردیده است که نشان دهنده عملکرد قابل قبول ویژگی‌های استخراج شده است (Gurubelli et al., 2020). روشی بر مبنای تجزیه و تحلیل ویژگی‌های رنگی و بافتی حاصل از ماتریس GLCM برای شناسایی و قطع‌بندی سیب در تصاویر گرفته شده از درخت، ایجاد شده است (Zhang et al., 2020). مدل‌سازی با نه الگوریتم مختلف انجام و این روش با دقت ۹۴ درصدی مؤثر ارزیابی شده است.

تحقیقات مختلفی در زمینه تشخیص و طبقه‌بندی بیماری‌های گیاهی با الگوریتم‌های مختلف انجام پذیرفته است. به‌عنوان مثال، الگوریتم ژنتیک برای تشخیص بیماری گیاهی (Dubey & Jalal, 2016)، شبکه عصبی مصنوعی^{۱۰} (ANN) به منظور تشخیص بیماری گیاه پنبه (Revathi & Hemalatha, 2012) و مدل ماشین بردار پشتیبان و شبکه عصبی جهت تشخیص بیماری گیاه درخت مو (Zhang et al., 2020) استفاده گردیده‌اند که در تمامی موارد نتایج قابل قبولی گزارش گردیده است. برگ گیاه یکی از اندام‌هایی است که در صورت ابتلای درخت به بیماری به سرعت دچار تغییرات ظاهری می‌شود و در صورت تشخیص به هنگام، می‌توان از آسیب به محصول جلوگیری کرد. بنابراین، این تحقیقات از ویژگی‌های مختلف رنگ و بافت برگ درخت

6 Color Coherence Vector

7 Local Binary Pattern

8 Gray Level Co-occurrence Matrix

9 Pixel Run Length Matrix

10 Artificial Neural Network

1 Black rot

2 Esca

3 Isariopsis

4 Support Vector Machine

5 Gabor Histogram

بافت تصویر از هیستوگرام سطح خاکستری^۱ (GLH)، ماتریس طول بردار سطح خاکستری^۲ (GLCM)، و LBP استخراج گردید. دسته‌بندی تصاویر با یک شبکه عصبی پیش‌خور و یک ماشین بردار پشتیبان چند کلاسی انجام شد. نتایج به دست آمده از هر یک از گروه‌های ویژگی، مورد مقایسه قرار گرفته و برای معرفی مدل بهینه بر اساس دقت دسته‌بندی، از ادغام تمام ویژگی‌ها در یک بردار ویژگی بهره گرفته شد. در نهایت، به منظور دستیابی به روشی سریع و دقیق، توسط الگوریتم کلونی زنبور ژنتیکی^۳ (GBC) کمترین تعداد ویژگی‌هایی که بیشترین دقت را ارائه دهد، انتخاب شد. شکل (۱) نمای کلی روش پیشنهادی را نمایش می‌دهد.

پایگاه داده

با توجه به اینکه برای ارزیابی بهتر سامانه پیشنهاد شده، نیازمند نمونه تصاویر زیاد و مورد تایید بود، از پایگاه‌های اطلاعاتی موجود که توسط محققین دیگری ارائه شده است استفاده گردید. بدین منظور، پایگاه داده آماده‌ای با تعداد ۴۰۶۲ تصویر برگ درخت انگور رقم کشمش بی‌دانه سفید (*Sultana - Vitis vinifera*) در اندازه استاندارد ۲۵۶*۲۵۶ پیکسل با چگالی ۹۶dpi و استاندارد JPG در فرمت RGB که در تحقیقات قبلی نیز مورد استفاده قرار گرفته‌اند، استفاده شد (Mohammed *et al.*, 2021). تصاویر شامل ۴۲۳ برگ سالم (H)، ۱۱۸۰ برگ مبتلا به لکه سیاه (BR)، ۱۳۸۳ برگ مبتلا به اسکا (E) و ۱۰۷۶ برگ مبتلا به لکه ایزاریوپسیس (I) می‌باشند. شکل (۲) نمونه‌ای از هر یک از کلاس‌های تصویر موجود در پایگاه داده را به نمایش می‌گذارد (Anon, 2021, Retrieved July 4).

پیش‌پردازش

به منظور بهبود نتایج روش پیشنهادی در تشخیص و

انگور در فضاها رنگی RGB یا HSI جهت تهیه پایگاه داده استفاده نموده‌اند که در تحقیقی که بر روی بیماری درخت انگور گزارش گردیده است، دقت بالای ۹۶٪ به دست آمده است (Zhang *et al.*, 2020). تحقیقی بر روی ۴۰۰ تصویر مختلف در ۴ کلاس، شامل ۱۰۰ برگ سالم، ۱۰۰ برگ با پوسیدگی سیاه، ۱۰۰ برگ با اسکا و ۱۰۰ برگ با لکه ایزاریوپسیس صورت گرفته است (Chakraborty *et al.*, 2017). نتایج تحقیقات گذشته نشان دهنده عملکرد و کارایی روش‌های پردازش تصویر در تلفیق با برخی روش‌های مدل‌سازی است که می‌تواند در بخش باغداری انگور و مرحله داشت آن مورد استفاده قرار گیرد.

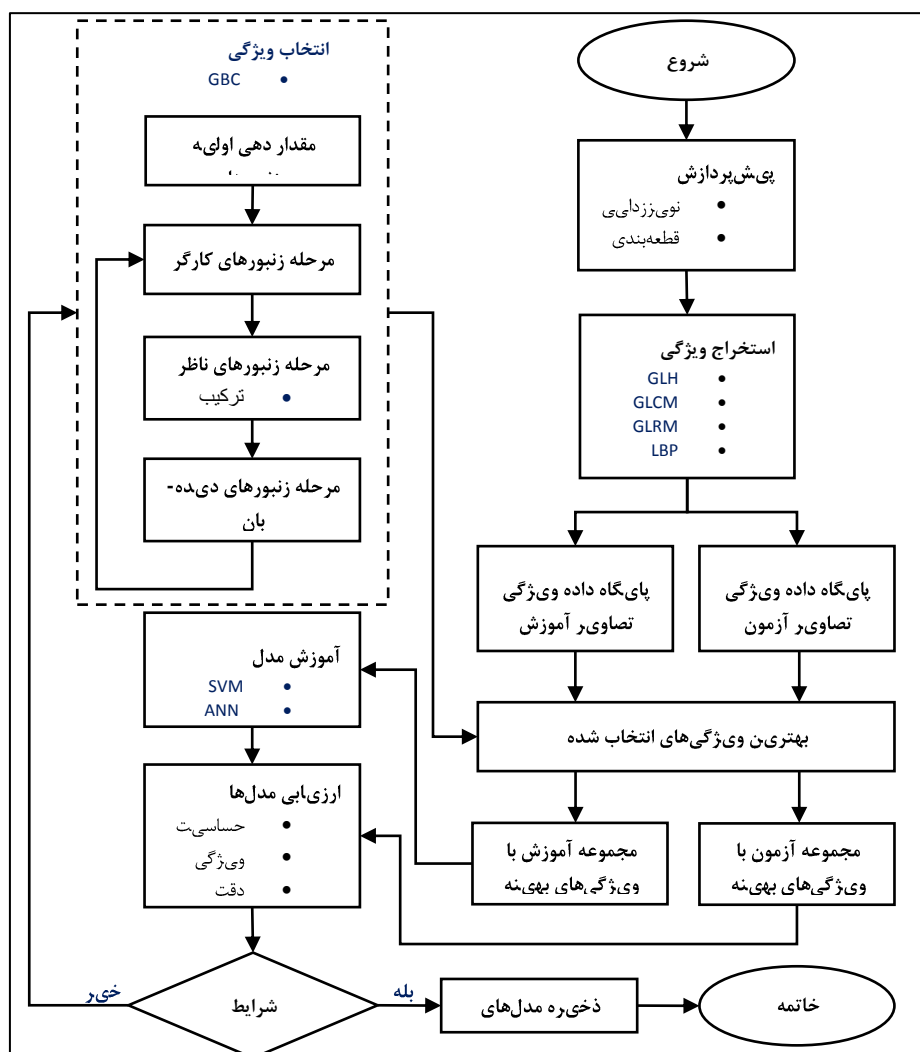
علی‌رغم کاربردهای فراوان روش بینایی ماشین و پردازش تصویر در تشخیص انواع بیماری گیاهان و همین‌طور نتایج مطلوبی که گزارش گردیده است، همچنان بررسی جامع‌تر و وسیع‌تر انواع الگوریتم‌های استخراج ویژگی‌های بافت، به‌عنوان یکی از نیازهای اساسی این حوزه می‌باشد. در این پژوهش سعی شد با استفاده از الگوریتم‌های مختلف استخراج ویژگی‌های بافت تصویر و بهره‌گیری از یادگیری ماشین، بهترین مدل و بهترین الگوریتم برای تشخیص بیماری و آفت درخت انگور پیشنهاد گردد.

مواد و روش‌ها

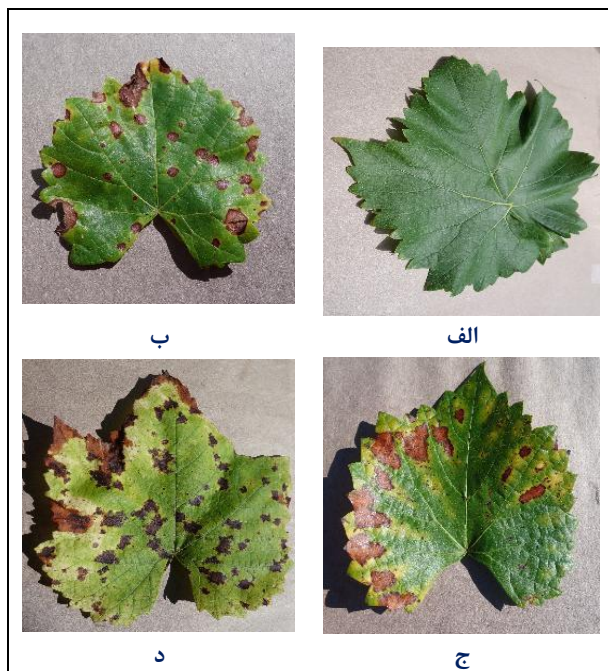
در این تحقیق، سامانه هوشمندی طراحی و ارائه می‌شود که توانایی تشخیص برگ سالم و ناسالم درخت انگور را داشته و علاوه بر این، سه بیماری پوسیدگی سیاه، اسکا و لکه ایزاریوپسیس موجود در نمونه را طبقه‌بندی می‌نماید. به منظور ایجاد روش پیشنهادی، ابتدا در مرحله پیش‌پردازش، تصاویر نوپزدایی و برگ به عنوان ناحیه مورد نظر از پس‌زمینه جدا می‌شوند. ویژگی‌های رنگ و

عنوان لبه تعریف می‌شوند که بزرگی آن‌ها در راستای گرادیان بیشینه باشد، بنابراین گرادیان محلی و جهت لبه هر پیکسل محاسبه شده و لبه‌های تصویر ایجاد می‌شود. ناحیه مورد نظر نیز با یک مرحله عملگر بستن بوسیله یک ساختار دیسک با شعاع ۵، بر روی تصویر لبه ایجاد شد. در نهایت، به منظور اطمینان از پیوستگی کامل سطح برگ، حفره‌های احتمالی تصویر نیز پر می‌شوند. شکل (۳) نتایج گام‌های مختلف پیش‌پردازش را نمایش می‌دهد.

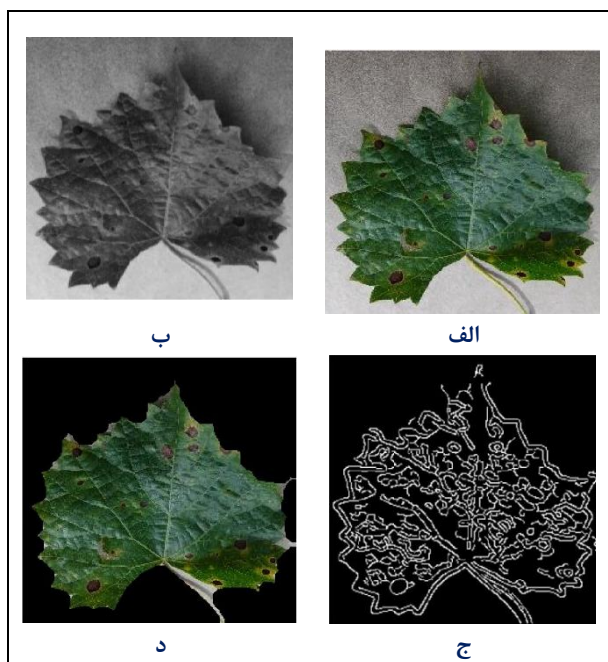
دسته‌بندی بیماری برگ انگور، از یک مرحله پیش‌پردازش استفاده شد. این مرحله شامل دو فرآیند نویززدایی و استخراج برگ از پس‌زمینه می‌باشد. به این منظور ابتدا تصاویر رنگی به سطح خاکستری تبدیل و با استفاده از فیلتر گوسی $\sigma = 3$ و 3×3 نویززدایی انجام شد. در گام بعدی، به منظور استخراج ناحیه مورد نظر (برگ)، از قطعه‌بندی مبتنی بر لبه بهره گرفته شد. برای این کار، ابتدا لبه‌های تصویر نویززدایی شده به کمک روش کنی^۱ استخراج گردید. در روش کنی نقاطی به



شکل ۱- روند نامی روش پیشنهادی برای تشخیص خودکار بیماری برگ درخت انگور
 Fig. 1. The flow chart of proposed method for automatic detection of grape leaf diseases



شکل ۲- نمونه‌ای از تصاویر برگ انگور به منظور استخراج ویژگی و تشکیل پایگاه داده
 الف: برگ سالم. ب: برگ دارای لکه سیاه. ج: برگ دارای لکه اسکا. د: برگ دارای لکه ایزاریوپسیس
Fig. 2. Sample grape leaf images for feature extraction and formation of the database
 a. Healthy leaf, b. leaf with black rot, c. leaf with Esca, d. leaf with Isariopsis



شکل ۳- خروجی گام‌های پیش‌پردازش برای حذف زمینه
 الف: تصویر اصلی برگ درخت انگور ب: تصویر سطح خاکستری نویزدایی شده. ج: لبه‌یابی با استفاده از روش کنی. د: نتیجه نهایی
Fig.3. Output of preprocessing methods: steps for removing the background
 a. the main image of grape leaf, b. gray scale noise removal image, c. thresholding by Canny method, d. final output

ظاهری اصلی‌ترین پارامترها برای اندازه‌گیری می‌باشند که اغلب افراد خبره با توجه به همین ویژگی‌ها، برگ درختان را از لحاظ بیماری مورد ارزیابی قرار می‌دهند.

استخراج ویژگی
 به منظور ایجاد روشی غیرمخرب برای دسته‌بندی برگ-های سالم و ناسالم و تشخیص انواع بیماری، ویژگی‌های

بافت تصویر را مشخص می‌کند، همواره یکی از ساده-ترین و پرکاربردترین روش‌های تحلیل بافت می‌باشد (Gonzalez et al., 2004). در این تحقیق، بر اساس احتمال وقوع شدت‌های مختلف برای هر یک از تصاویر، شش ویژگی میانگین، انحراف معیار، همواری، گشتاور سوم، یکنواختی و آنتروپی محاسبه گردید. جدول (۱) روابط مورد استفاده برای محاسبه این ویژگی‌ها را نمایش می‌دهد. این ویژگی‌ها به ترتیب میانگین شدت بافت، میانگین کنتراست بافت، همواری نسبی شدت نواحی، چولگی هیستوگرام، میزان یکنواختی سطوح شدت و میزان تصادفی بودن را به صورت کمی نمایش می‌دهند. به طوری که i مقادیر ممکن سطوح شدت از 0 تا $L-1$ ، z مقدار احتمال شدت و $p(z_i)$ هیستوگرام سطوح شدت در یک ناحیه را نمایش می‌دهند.

از این رو، برای ایجاد روش پیشنهادی از استخراج و تحلیل ویژگی‌های بافت و رنگ استفاده شده است. اغلب الگوهای تصاویر بر اساس شباهت‌های موجود در پیکسل‌های همسایگی ایجاد می‌شوند. در روش‌های تجزیه و تحلیل بافت مهم‌ترین موضوع، انتخاب ویژگی‌هایی می‌باشد که اطلاعات مؤثری از تصویر را در بر گیرند. با استفاده از چند روش مختلف، ویژگی‌های تصاویر استخراج و مورد تجزیه و تحلیل قرار گرفت که در ادامه به شرح هر یک پرداخته می‌شود.

ویژگی‌های استخراج شده از (GLH) علی‌رغم اینکه تحلیل بافت به وسیله مشخصات آماری هیستوگرام تصویر، نحوه قرارگیری پیکسل‌ها نسبت به پیکسل‌های همسایگی را مورد تجزیه و تحلیل قرار نمی‌دهد، اما در عین حال، به دلیل اینکه دید کلی از توزیع

جدول ۱- ویژگی‌های مستخرج از هیستوگرام سطح خاکستری (GLH) به منظور تشکیل پایگاه داده
Table 1. Extracted features using Gray Level Histogram (GLH) for creation of data set

ویژگی	رابطه
میانگین	$m = \sum_{i=0}^{L-1} z_i p(z_i)$
انحراف معیار	$\sigma = \sqrt{\sum_{i=0}^{L-1} (z_i - m)^2 p(z_i)}$
همواری	$R = 1 - 1/(1 + \sigma^2)$
گشتاور سوم	$\mu_3 = \sum_{i=0}^{L-1} (z_i - m)^3 p(z_i)$
یکنواختی	$U = \sum_{i=0}^{L-1} p^2(z_i)$
آنتروپی	$e = - \sum_{i=0}^{L-1} p(z_i) \log_2 p(z_i)$

i - مقادیر ممکن سطوح شدت، z - مقدار احتمال شدت، $p(z_i)$ - هیستوگرام سطوح شدت در ناحیه L .

خاکستری، ماتریس هم-رخدادی، سطح خاکستری میزان احتمال مشترک وقوع دو سطح رنگ مختلف را در یک فاصله و زاویه مشخص نشان می‌دهد. زمانی که فاصله انتخابی برابر با یک باشد، صرفاً الگوهای موجود در پیکسل‌های مجاور مورد بررسی قرار خواهد گرفت.

ویژگی‌های استخراج شده از (GLCM) یکی از روش‌های آماری برای توصیف بافت تصویر استفاده از ماتریس هم-رخدادی می‌باشد (Haralick et al., 1973). این ماتریس برای یافتن الگوهای خاص موجود در سطح تصاویر به کار می‌رود. برای یک تصویر

می‌شوند، برابر با تعداد پیکسل‌های هم‌جواری می‌باشد که در یک راستای مشخص، شدت خاکستری یکسانی دارند. در این تحقیق، برای هر تصویر ۱۱ ویژگی در چهار راستای صفر، ۴۵، ۹۰ و ۱۳۵ درجه از ماتریس طول بردار اجرای سطح خاکستری استخراج شد. روابط ویژگی‌های استخراج شده در جدول (۳) مشخص شده‌اند، که i, j ؛ $Q(i, j)$ و S به ترتیب معرف میزان سطح خاکستری، میزان طول اجرا، درایه‌ها و مجموع ماتریس طول بردار اجرا می‌باشند.

ویژگی‌های استخراج شده از (LBP)

ماتریس الگوی دودویی محلی، ساختار محلی همسایگی یک پیکسل را توصیف می‌کند (Ojala et al., 2002). به عنوان مثال، در یک همسایگی 3×3 ، ابتدا پیکسل‌های همسایه با استفاده از آستانه‌گذاری با مقدار پیکسل مرکزی، به صفر یا ۱ تبدیل می‌شوند و سپس با تبدیل کد دودویی متناظر با همسایگی‌ها به عدد صحیح، مقدار الگوی دودویی محلی به دست می‌آید. با توجه به سادگی عملکرد و قدرت متمایزکنندگی این روش، انواع مختلفی از آن گسترش یافته و مورد استفاده قرار گرفته است.

برای استخراج ویژگی‌های انرژی، آنروپی، بیشترین احتمال، کنتراست، همبستگی و همگنی از ماتریس هم‌رخدادی برای پیکسل‌های مجاور و در چهار راستای صفر، ۴۵، ۹۰ و ۱۳۵ درجه استفاده شد که روابط آن‌ها در جدول (۲) مشخص شده است. در این جدول، مقدار p_{ij} -آمین مؤلفه ماتریس هم-رخداد سطح خاکستری نرمالیزه شده می‌باشد، m_c ، m_r به ترتیب میانگین سطری و ستونی ماتریس هم-رخدادی و $\sigma_r \neq 0$ و $\sigma_c \neq 0$ نیز به همین ترتیب انحراف معیار استاندارد را برای این ماتریس مشخص می‌کنند. بدین ترتیب، ۲۴ ویژگی از ماتریس هم-رخدادی در بانک اطلاعات ویژگی ایجاد گردید.

ویژگی‌های استخراج شده از (GLRM)

روش دیگری که قابلیت استخراج الگوهای یک تصویر را برای تحلیل بافت دارد، استخراج ویژگی از ماتریس طول بردار اجرای سطح خاکستری می‌باشد (Karimi et al., 2017). ماتریس طول بردار اجرا برای یک تصویر سطح خاکستری دو بعدی، هر یک از درایه‌های این ماتریس که با عنوان طول اجرا (Run Length) نیز شناخته

جدول ۲- ویژگی‌های مستخرج از ماتریس هم‌رخدادی سطح خاکستری (GLCM) به منظور تشکیل پایگاه داده

Table 2. Extracted features using Gray Level Co-Occurrence Matrix (GLCM) for creation of data set

ویژگی	رابطه
انرژی	$\sum_{i=1}^k \sum_{j=1}^k p_{ij}$
آنروپی	$-\sum_{i,j=1}^k p_{ij} \log_2 p_{ij}$
بیشترین احتمال	$MAX(p_{ij})$
کنتراست	$\sum_{i=1}^k \sum_{j=1}^k (i-j)^2 p_{ij}$
همبستگی	$\sum_{i=1}^k \sum_{j=1}^k \frac{(i-m_r)(j-m_c)p_{ij}}{\sigma_r \sigma_c}$
همگنی	$\sum_{i=1}^k \sum_{j=1}^k \frac{p_{ij}}{1+ i-j }$

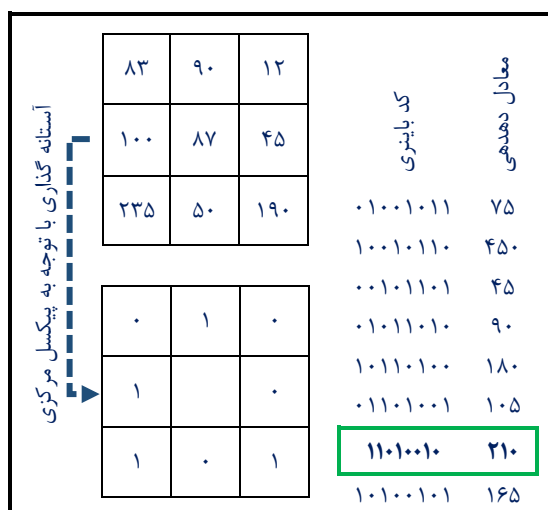
p_{ij} - مقدار ij -آمین مؤلفه ماتریس GLCM و m_c و m_r به ترتیب میانگین سطری و ستونی ماتریس GLCM. $\sigma_r \neq 0$ و $\sigma_c \neq 0$ به ترتیب انحراف معیار سطری و ستونی ماتریس GLCM.

جدول ۳- ویژگی‌های مستخرج از ماتریس طول بردار سطح خاکستری (GLRM) به منظور تشکیل پایگاه داده

Table 3. Extracted features using Gray Level Run Length Matrix (GLRM) for creation of data set

رابطه	ویژگی
$\sum_i \sum_j (Q(i,j)/j^2) / S$	اجرای کوتاه
$\sum_i \sum_j (j^2 Q(i,j)) / S$	اجرای بلند
$\sum_i (\sum_j Q(i,j))^2 / S$	ناهمگونی سطح خاکستری
$\sum_j (\sum_i Q(i,j))^2 / S$	ناهمگونی طول اجرا
$\sum_i \sum_j S/jQ(i,j)$	نسبت اجرا
$\sum_i \sum_j Q(i,j) / Si^2$	اجرای سطح خاکستری پایین
$\sum_i \sum_j i^2 Q(i,j) / S$	اجرای سطح خاکستری بالا
$\sum_i \sum_j Q(i,j) / Sj^2 i^2$	اجرای سطح خاکستری پایین کوتاه
$\sum_i \sum_j i^2 Q(i,j) / Sj^2$	اجرای سطح خاکستری بالای کوتاه
$\sum_i \sum_j j^2 Q(i,j) / Si^2$	اجرای سطح خاکستری پایین بلند
$\sum_i \sum_j i^2 j^2 Q(i,j) / S$	اجرای سطح خاکستری بالای بلند

i - مقدار سطح خاکستری، j - مقدار طول اجرا، $Q(i,j)$ - درایه‌های GLRM، S - مجموع GLRM.



شکل ۴- نمونه‌ای از نحوه محاسبه LBP برای یک قطعه ۳*۳ از تصویر به منظور تشکیل پایگاه داده و کد باینری از پیکسل سمت چپ بالا شروع شده و بصورت ساعت‌گرد شیفت داده می‌شود. بیشترین مقدار محاسبه شده به عنوان LBP پیکسل مرکزی انتخاب می‌شود.

Fig. 4. Sample procedure to calculate the LBP for a 3 × 3 window of an image to form a data set. The binary code starts from the top left pixel and shifts clockwise. Maximum calculated value is selected as the LBP of the center pixel.

در این مقاله، از الگوی دودویی محلی پایه و به منظور ایجاد کدهای یکتا برای الگوهای محلی یکسان از حرکت چرخشی در جهت عقربه‌های ساعت معرفی شده در تحقیقات قبلی استفاده شد (Karimi et al., 2017). در این روش، هر بار پس از محاسبه مقدار دودویی همسایگی‌ها، درایه‌های ماتریس دودویی حاصل از آستانه‌گذاری یک واحد در راستای ساعت‌گرد جابجا شده و برای هر یک از حالات مقدار متناظر محاسبه می‌شود. در نهایت، بیشترین مقدار، به عنوان الگوی دودویی محلی پیکسل مورد نظر انتخاب می‌شود (شکل ۴). برای هر تصویر از ماتریس الگوی دودویی محلی شش ویژگی میانگین، انحراف معیار، همواری، گشتاور سوم، یکنواختی و آنتروپی محاسبه گردید.

انتخاب ویژگی

کلونی زنبورهای ژنتیکی با ادغام قابلیت اکتشاف الگوریتم زنبورهای عسل و قابلیت بهره‌برداری الگوریتم ژنتیک (ترکیب و جهش) و تعادل این موارد در زمان کمتری به جواب بهینه می‌رسد (Alshamlan *et al*, 2015). زنبور ژنتیکی نیز به مانند الگوریتم کلونی زنبورهای عسل دارای چهار مرحله اصلی مقداردهی اولیه، زنبورهای کارگر، زنبورهای ناظر و زنبورهای دیده‌بان می‌باشد. با فرض اینکه اندازه جمعیت برابر با SN باشد، تعدادی از ویژگی‌ها به صورت تصادفی انتخاب می‌شوند. اگر نشان دهنده ژن i -ام باشد $(i = 1, 2, 3, \dots, SN)$ ؛ هر ژن یک بردار $D = 1, 2, 3, \dots, D$ بعدی با مقادیر ۰ یا ۱ را شامل می‌شود که نشان دهنده انتخاب یا عدم انتخاب ویژگی می‌باشد. در فاز مقداردهی اولیه، جمعیت اولیه به صورت تصادفی ایجاد و میزان برازندگی هر ژن محاسبه می‌شود که در این تحقیق برای اندازه‌گیری برازندگی از پارامتر دقت محاسبه شده توسط الگوریتم‌های شبکه عصبی و ماشین بردار پشتیبان استفاده شد. در مرحله بعد، زنبورهای کارگر تمام فضای مسأله که در واقع شامل ۸۰ ویژگی استخراج شده بافت برای هر تصویر برگ است را جستجو می‌کنند. زنبورهای کارگر با جستجو در اطراف x_i به دنبال اندیس‌های بهتر در موقعیت جدید v_i هستند که توسط رابطه (۱) مشخص شده است.

$$v_{ij} = x_{ij} + R_{ij}(x_{ij} - x_{kj}) \quad (\text{رابطه ۱})$$

به طوری که $v_i = [v_{i1}, v_{i2}, v_{i3}, \dots, v_{in}]$ نمایش اندیس‌های ژن جدید (بردار موقعیت زنبور)، $x_i = [x_{i1}, x_{i2}, x_{i3}, \dots, x_{in}]$ اندیس‌های ژن فعلی است، $k(k \neq j)$ یک راه‌حل تصادفی انتخاب شده از راه‌حل‌های فعلی است و R_{ij} یک عدد تصادفی در بازه $[-1, 1]$ است.

زنبورهای ناظر در طبیعت اطلاعات منابع جدید را از طریق تماشای رقص چرخشی زنبورهای کارگر

کسب می‌کنند، اما در روش پیشنهادی از عملگر ترکیب یکنواخت برای به اشتراک‌گذاری اطلاعات منابع غذایی (ویژگی‌ها) و انتخاب همسایگی بر اساس رابطه (۱) استفاده شد. به منظور افزایش سرعت همگرایی، زنبورهای ناظر از مکان بهترین راه‌حل یافته شده به عنوان ملکه برای انتخاب همسایگی استفاده می‌کنند. انتخاب همسایگی‌ها به صورت تصادفی و بر اساس احتمال هر منبع انجام می‌شود؛ این احتمال از نرمالیزه کردن برازندگی هر منبع به دست آمده است. اما در بخش زنبور دیده‌بان؛ در صورت عدم بهبود پاسخ‌های ایجاد شده و ارضای شرط محاکمه، دو زنبور کارگر تبدیل به زنبور دیده‌بان می‌شوند (Alshamlan *et al*, 2015) که یکی به صورت تصادفی و زنبور دوم با استفاده از جهش، منابع جدید را کاوش می‌کند، اما در این تحقیق برای هر دو زنبور دیده‌بان، از عملگر جهش با نرخ ۰/۰۱ و رابطه (۲) استفاده گردید که باعث افزایش بهره‌برداری می‌شود.

$$ScoutB_{ij} = QueenB_{ij} + R_{ij}(RandB_{ij} + QueenB_{ij}) \quad (\text{رابطه ۲})$$

در رابطه (۲)، $QueenB$ معرف زنبور ملکه و بهترین راه‌حل، i اندیس راه‌حل i -ام است، به طوری که جهش بر روی تمام اندیس‌های z که $j \in [1, \dots, D]$ اعمال می‌شود. $RandB$ یک راه‌حل تصادفی انتخاب شده است که از طریق رابطه (۱) به دست آمده است و R_{ij} یک عدد تصادفی در بازه $[-1, 1]$ است.

مدل‌سازی

روش‌های یادگیری ماشین بسیاری برای تشخیص و دسته‌بندی الگوها در دو یا چند کلاس وجود دارند. در این مقاله، با توجه به در دسترس بودن برچسب تصاویر، برای طبقه‌بندی تصاویر در کلاس‌های مد نظر از الگوریتم‌های یادگیری ماشین نظارت شده استفاده گردید. از این رو و برای مقایسه‌پذیری، دو الگوریتم شبکه عصبی مصنوعی (ANN) و ماشین بردار پشتیبان (SVM) جهت ارزیابی بهتر کارایی ویژگی‌های استخراج

عصبی آموزش دیده برای تشخیص الگو یا طبقه‌بندی داده، در فرآیند آموزش، میزان وزن هر یک از نورون‌ها را تنظیم می‌کند، به طوری که در صورت ورود داده جدید به این ساختار، برچسب صحیح تشخیص داده شود.

در مقاله حاضر، برای تشخیص و طبقه‌بندی بیماری برگ درخت انگور از یک شبکه عصبی از نوع پیش‌خور با یک لایه مخفی، شامل ۳۰ نورون و ماشین بردار پشتیبان از نوع چند کلاسی با کرنل چند جمله‌ای درجه ۳ استفاده شد. برای آموزش مدل‌های ذکر شده از روش ۱۰-بخشی (10-Fold) استفاده گردیده است. ۷۰ درصد از داده‌ها برای آموزش مدل‌ها و ۳۰ درصد آن‌ها نیز برای آزمون مدل‌ها استفاده شد. لازم به ذکر است انتخاب نمونه‌های آموزشی و آزمون برای هر دو مدل به صورت تصادفی انجام گرفت.

ارزیابی مدل‌های ارائه شده بر اساس پارامترهای آماری حساسیت^۱، ویژگی^۲ و دقت^۳ بر روی ماتریس آشفتگی انجام پذیرفت. این پارامترها به ترتیب در معادلات (۳)، (۴) و (۵) نمایش داده شده است، که TP ، TN ، FP و FN به ترتیب نشان دهنده مقادیر نرخ مثبت صحیح، نرخ منفی صحیح، نرخ مثبت کاذب و نرخ منفی کاذب می‌باشند. حساسیت، میزان کارایی مدل و در واقع توانایی شناسایی صحیح کلاس‌ها را می‌سنجد. در طرف مقابل، پارامتر آماری ویژگی، نرخ پاسخ‌های منفی درست را اعلام می‌کند؛ به عبارت دیگر، نسبت صحیح برای پیش‌بینی عدم تعلق به یک کلاس می‌باشد. در نهایت، دقت، میزان کل پیش‌بینی‌های صحیح را نمایش می‌دهد (Karimi et al., 2017).

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (\text{رابطه ۳})$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (\text{رابطه ۴})$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (\text{رابطه ۵})$$

تمامی مراحل استخراج ویژگی‌ها، مدل‌سازی و

شده به کار گرفته شدند.

در یادگیری نظارت شده که هر نمونه یک جفت شامل شیء ورودی و خروجی مد نظر را تشکیل می‌دهد، الگوریتم تحت آموزش با داده‌های آموزشی، تابعی برای نگاشت صحیح داده‌های جدید (آزمون) به خروجی مطلوب را به دست می‌آورد؛ یعنی الگوریتم با استفاده از تابع به دست آمده، خروجی مطلوب را برای داده‌های جدید پیش‌بینی خواهد کرد.

ماشین بردار پشتیبان یک طبقه‌بند دو کلاسی است که برای طبقه‌بندی و رگرسیون مورد استفاده قرار می‌گیرد (Singh & Kaur, 2018). برای یک مجموعه از نمونه‌های آموزشی، ماشین بردار پشتیبان مدلی را برای پیش‌بینی اینکه نمونه جدید درون یک کلاس یا کلاس دیگر می‌باشد، ایجاد می‌کند. این جداسازی به کمک ابر صفحه جداساز اتفاق می‌افتد. اساساً برای دو کلاس که به طور خطی قابل جداسازی می‌باشند، بیش از یک ابر صفحه جداساز وجود دارد. بنابراین صفحه‌ای انتخاب می‌شود که حاشیه بین کلاس‌ها را به بیشترین میزان برساند. اما در غالب مسائل دنیای واقعی، بیش از دو کلاس برای داده‌ها وجود دارد. در این گونه مسائل، دو رویکرد برای ایجاد مدل وجود دارد؛ یک رویکرد به صورت مستقیم چند کلاس را طبقه‌بندی می‌کند و رویکرد دیگر با ترکیب چند طبقه‌بند ماشین بردار پشتیبان دو کلاسی انجام می‌شود.

شبکه عصبی که از ساختار بیولوژیکی مغز انسان الهام گرفته شده است، یک الگوریتم پردازش اطلاعات می‌باشد (Karimi et al., 2017). ساختار یک شبکه عصبی شامل تعدادی از لایه‌های متشکل از نورون‌های مصنوعی می‌باشد. در فرم ابتدایی این ساختار شامل سه لایه ورودی، لایه پنهان و لایه خروجی می‌باشد. هر یک از نورون‌ها بسته به لایه‌ای که در آن هستند، با اتصالات وزنی به لایه‌های قبل و بعد مرتبط می‌شوند. یک شبکه

ارزیابی آن‌ها در نرم‌افزار متلب نسخه ۹ و با استفاده از جعبه‌ابزارهای استاندارد پردازش تصویر و یادگیری ماشین انجام پذیرفت.

نتایج و بحث

آموزش مدل‌های ماشین بردار پشتیبان و شبکه عصبی به وسیله ویژگی‌های استخراج شده از GLH، GLCM، LBP و GLRM انجام پذیرفت و تصاویر برگ در چهار گروه سالم (H)، پوسیدگی سیاه (BR)، اسکا (E) و لکه ایزاریوپسیس (I) دسته‌بندی شدند. برای این منظور، تصاویر پایگاه داده به صورت تصادفی به نسبت ۷۰ به ۳۰ به ترتیب برای بخش آموزش و آزمون تقسیم‌بندی شدند. بدین ترتیب، در بخش اول، برای هر گروه از

ویژگی‌ها مدلی مجزا ایجاد و مورد ارزیابی قرار گرفت. جدول (۴) نتایج این بررسی را بر روی تمام پایگاه داده نمایش می‌دهد. ویژگی‌های مستخرج از هیستوگرام سطح خاکستری، کمترین میزان دقت را در هر دو مدل داشت که می‌تواند به علت عدم دخالت پیکسل‌های همسایگی در استخراج ویژگی در این روش قابل توجیه باشد.

بر اساس جدول (۴)، بیشترین میزان دقت برابر با ۸۹/۷۰٪ متعلق به مدل ماشین بردار پشتیبان مبتنی بر ویژگی‌های GLRM خاکستری می‌باشد. همچنین مشاهده می‌شود، ویژگی‌های مستخرج از GLRM در هر دو مدل با دارا بودن ویژگی و حساسیت بالاتر، کارایی بهتری نسبت به سایر ویژگی‌ها داشتند.

جدول ۴- نتایج مدل‌سازی‌های شبکه عصبی و ماشین بردار پشتیبان با استفاده از ویژگی‌های بافت GLH، GLCM، GLRM و LBP به منظور مقایسه مدل‌ها و ویژگی‌ها

Table 4. Results of ANN and SVM modeling using texture features of GLH, GLCM, GLRM and LBP to compare the models and features

مدل / ویژگی	شبکه عصبی			ماشین بردار پشتیبان		
	دقت	ویژگی	حساسیت	دقت	ویژگی	حساسیت
GLH	۸۰/۸	۹۰/۸	۷۹/۷	۷۸/۶	۸۹/۴	۷۷/۱
GLCM	۸۶/۳	۹۴/۰	۸۵/۷	۸۴/۲	۹۲/۸	۸۳/۴
GLRM	۸۸/۲	۹۴/۹	۸۷/۸۰	۸۹/۷	۹۳/۲	۸۹/۶
LBP	۸۳/۰	۹۱/۰	۸۲/۱	۷۹/۵	۹۲/۱	۷۹/۲

مدل شبکه عصبی داشته است.

در پژوهش حاضر، با بررسی دقت‌های اندازه‌گیری شده در جدول (۴) مشخص می‌شود، به طور کلی شبکه عصبی، دقت بیشتری در دسته‌بندی تصاویر برگ‌های سالم و بیمار داشته است. بر همین اساس، مدل شبکه عصبی مبتنی بر ویژگی‌های GLH، GLCM، LBP و GLH دقت بهتری نسبت به ماشین بردار پشتیبان با همین ویژگی‌ها داشته است. نتایج جدول (۴)، مشخص می‌کند هر دو مدل میزان ویژگی بالاتری نسبت به حساسیت ارائه داده‌اند. در واقع، مدل‌های ایجاد شده با ویژگی‌های بافت و رنگ، در اعلام صحیح عدم تعلق به یک کلاس موفق‌تر از شناسایی صحیح کلاس‌ها بوده‌اند. همچنین دقت خوب و نزدیکی نتایج شبکه عصبی و ماشین بردار

در پژوهشی که برای تشخیص کیفیت کشمش فله انجام یافته است (Khojastehnazhand & Ramezani, 2020)، مدل مبتنی بر ماشین بردار پشتیبان در مقایسه با آنالیز تفکیک خطی با استفاده از ویژگی‌های GLRM، بیشترین کارایی را داشته و بهترین مدل در تشخیص کشمش فله شامل کشمش خوب، بد و خار و خاشاک می‌باشد. اما در پژوهش انجام یافته، به منظور تشخیص درصد ترکیب کشمش خوب و بد، ویژگی‌های GLCM بهترین نتایج را برای آموزش ماشین بردار پشتیبان در پی داشته و نتایج ارائه شده نشان از دقت پایین شبکه عصبی و ماشین بردار پشتیبان مبتنی بر ویژگی‌های GLRM اجرا دارد (Karimi et al., 2017). همچنین بر اساس نتایج ارائه شده در این تحقیق، مدل SVM عملکرد بهتری در مقایسه با

میزان حساسیت در هر دو مدل مربوط به گروه بیماری لکه سیاه و اسکا می‌باشد. هر چند شبکه عصبی در بخش آموزش با حساسیت ۹۴/۳٪ در کلاس بیماری اسکا عملکرد مناسبی داشته است، اما در فرآیند آزمون با حساسیت زیر ۹۰٪ شاهد کاهش عملکرد هستیم که در نهایت منجر به کاهش دقت کلی گردیده است. اما افزایش حساسیت مدل‌ها که نشان دهنده قابلیت مدل‌ها در تشخیص صحیح کلاس‌ها می‌باشد، در نهایت منجر به افزایش نرخ دقت هر دو مدل شده و در نتیجه دسته-بندی تصاویر بهبود قابل ملاحظه‌ای داشته است. در کار مشابهی که برای دسته‌بندی درصد‌های مختلفی از کشمش فله انجام یافته است نیز، بهترین مدل پس از ادغام ویژگی‌های مختلف به دست آمده است (Karimi et al., 2017)

پشتیبان (جدول ۴)، کارآیی ویژگی‌های استخراج شده در دسته‌بندی تصاویر برگ سالم و مبتلا به بیماری درخت انگور را نمایش می‌دهد. در ادامه برای ارائه مدلی با کارآیی بهتر، تمام گروه-های ویژگی یک تصویر در قالب یک بردار ویژگی واحد برای ساخت مدل مورد تجزیه و تحلیل قرار گرفت. جدول (۵) به ترتیب نتایج دسته‌بندی تصاویر برای بخش‌های آموزش و آزمایش توسط هر دو مدل آموزش یافته بر مبنای ترکیب تمام ویژگی‌های استخراج شده را به نمایش می‌گذارد. با تجمیع ویژگی‌ها، مقدار حساسیت مدل‌ها نیز افزایش قابل قبولی داشت. جدول (۵) به خوبی نشان می‌دهد که میزان حساسیت مدل‌ها نسبت به استفاده از گروه‌های ویژگی مختلف به صورت مجزا که نتایج آن در جدول (۴) درج گردیده است، اختلاف معنی‌دار و قابل ملاحظه‌ای دارد. اما همچنان کمترین

جدول ۵- نتایج مدل‌سازی با استفاده از تمام ویژگی‌ها* به منظور مقایسه نتایج و بررسی ماتریس اغتشاش

Table 5. Modeling results using all features to compare the results and examine the confusion matrix.

مرحله	مدل	شبکه عصبی				ماشین بردار پشتیبان				
		ایزار یوپسیس سالم	اسکا	پوسیدگی	میانگین	ایزار یوپسیس سالم	اسکا	پوسیدگی	میانگین	
آموزش	ماتریس آشفته	پوسیدگی سیاه	۱۷	۵۲	۷۴۱	۳۶	۵	۸۷	۷۰۲	
		اسکا	۵	۸۹۹	۷۳	۴	۰	۸۵۱	۱۲۲	
		سالم	۴	۲۷۴	۳	۱۹	۲۶۸	۰	۴	
	شاخص	ایزار یوپسیس	۷۶۰	۲	۱۲	۶۸۷	۸	۱۰	۴۰	
		حساسیت	۹۶/۷	۹۹/۳	۸۹/۴	۹۴/۹	۹۸/۴	۸۷/۶	۸۱/۸	
		ویژگی	۹۲/۲	۹۹/۷	۹۵/۹	۹۶/۷	۹۹/۴	۹۴/۵	۹۳/۷	
	آزمون	ماتریس آشفته	پوسیدگی سیاه	۱۴	۳۹	۳۱۶	۹	۱	۴۴	۲۹۶
			اسکا	۱	۳۶۴	۴۱	۲	۰	۳۶۲	۴۲
			سالم	۶	۱۳۴	۲	۵	۱۲۵	۰	۲
		شاخص	ایزار یوپسیس	۲۸۶	۳	۶	۳۰۱	۱	۷	۲۲
حساسیت			۹۳/۲	۹۷/۱	۸۶/۶	۹۲/۱	۹۵/۴	۸۹/۸	۸۰/۹	
ویژگی			۹۸/۵	۹۹/۳	۹۳/۷	۹۷/۲	۹۹/۱	۹۳/۳	۹۳/۵	
دقت		دقت	۹۳/۳	۹۷/۱	۸۹/۵	۹۲/۳	۹۵/۵	۸۹/۹	۸۲/۰	

* داده‌های سطر و ستونی به ترتیب بر حسب‌های صحیح و پیش‌بینی مدل‌ها را نمایش می‌دهند.

تصاویر برگ سالم (H) با دقت بیشتری نسبت به سایر کلاس‌ها دسته‌بندی شده‌اند، که با توجه به نزدیکی شدت رنگ‌های مناطق آلوده به بیماری در سایر کلاس‌ها

نتایج حکایت از عملکرد بهتر مدل مبتنی بر شبکه عصبی با متوسط دقت ۹۵/۰٪ و ۹۱/۷٪ به ترتیب برای داده‌های آموزشی و آزمون دارند. بر اساس جدول (۵)،

آمده از این روش برای شبکه عصبی و ماشین بردار پشتیبان به ترتیب برابر ۳۴ و ۴۶ ویژگی بود. جدول (۶) نتایج مدل‌های بررسی شده را برای ویژگی‌های بهینه به دست آمده توسط الگوریتم زنبور ژنتیکی نمایش می‌دهد، همان‌گونه که نتایج ارزیابی نیز نشان می‌دهند، با حذف ویژگی‌های غیربهینه، دقت دسته‌بندی مدل‌ها افزایش یافته است. با مقایسه جداول (۵) و (۶)، در صورت استفاده از ویژگی‌های بهینه، SVM دقت بالاتری از ANN با تمام ویژگی‌های استخراجی داشته است. اما همچنان شبکه عصبی با میانگین دقت ۹۷/۲٪ و ۹۴/۱٪ به ترتیب برای بخش آموزش و آزمون، عملکرد بهتری نسبت به ماشین بردار پشتیبان با میانگین دقت ۹۳/۰٪ و ۹۲/۳٪ داشت. بنابراین می‌توان در بررسی داده‌های جدید، صرفاً با استخراج ویژگی‌های بهینه مشخص شده توسط الگوریتم زنبورهای ژنتیکی، از حجم بسیار زیادی از حجم و محاسبات جلوگیری کرد.

و سطح کاملاً یک‌رنگ برگ‌های سالم، امری بدیهی و قابل قبول می‌باشد. علاوه بر این، با توجه به میزان بالای متوسط ویژگی در هر دو مدل، مشخص است که هر دو مدل در پیش‌بینی عدم تعلق کلاس‌ها نسبت به شناسایی صحیح کلاس‌ها موفق‌تر بوده‌اند.

اما انتخاب کمترین تعداد ویژگی‌هایی که بیشترین دقت دسته‌بندی را نتیجه می‌دهند، سه مزیت اساسی کاهش بیش‌برازش، شامل کاهش داده‌های زائد در اثر کاهش زمان تصمیم‌گیری بر روی داده‌های نویزی، افزایش دقت با کاهش تعداد داده‌های همراه‌کننده و کاهش زمان آموزش را در پی خواهد شد. الگوریتم کلونی زنبور ژنتیکی (GBC) که ترکیبی از الگوریتم‌های زنبورهای عسل و ژنتیک می‌باشد، سعی در یافتن ژن-هایی با بیشترین اطلاعات مفید دارد. بنابراین با انتخاب کمترین تعداد ویژگی که بیشترین میزان دقت دسته‌بندی را دارند، منجر به کاهش پیچیدگی زمانی روش پیشنهادی می‌شود. تعداد ویژگی‌های بهینه به دست

جدول ۶- نتایج مدل‌سازی با استفاده از ویژگی‌های بهینه یافته شده توسط الگوریتم GBC *
Table 6. Modeling results using optimized features found by the GBC algorithm

مرحله مدل		شبکه عصبی				ماشین بردار پشتیبان							
آموزش	ماتریس آشفتگی	ایزاربوپسیس سالم اسکا پوسیدگی				ایزاربوپسیس سالم اسکا پوسیدگی							
		پوسیدگی سیاه	۶	۰	۳۱	۷۷۳	پوسیدگی سیاه	۱۸	۷	۶۱	۷۴۴		
		اسکا	۲	۱	۹۳۳	۴۱	اسکا	۳	۱	۸۷۰	۱۰۳		
		سالم	۱	۰	۲۷۸	۲	سالم	۶	۱	۲۸۲	۲		
	میانگین	۷۶۵	۱	۱	۹	میانگین ایزاربوپسیس	۷۱۳	۱	۵	۲۶	میانگین		
	شاخص	حساسیت	۹۸/۸	۹۶/۷	۹۹/۳	۹۳/۷	۹۷/۱	حساسیت	۹۶/۳	۹۲/۸	۸۵/۰	۹۲/۸	
		ویژگی	۹۹/۵	۹۷/۶	۹۹/۹	۹۸/۱	۹۸/۸	ویژگی	۹۸/۴	۹۴/۳	۹۹/۶	۹۵/۱	۹۶/۹
		دقت	۹۸/۸	۹۶/۷	۹۹/۳	۹۳/۸	۹۷/۲	دقت	۹۶/۴	۹۲/۹	۹۶/۹	۸۵/۸	۹۳/۰
	آزمون	ماتریس آشفتگی	ایزاربوپسیس سالم اسکا پوسیدگی				ایزاربوپسیس سالم اسکا پوسیدگی						
			پوسیدگی سیاه	۸	۱	۲۳	۳۳۸	پوسیدگی سیاه	۷	۳	۳۵	۳۰۵	
اسکا			۲	۰	۳۷۵	۲۹	اسکا	۱	۰	۳۶۶	۳۹		
سالم			۵	۰	۱۳۶	۱	سالم	۱	۰	۱۲۸	۳		
میانگین		۲۸۹	۲	۴	۵	میانگین ایزاربوپسیس	۳۱۱	۲	۵	۱۳	میانگین		
شاخص		حساسیت	۹۵/۱	۹۳/۷	۹۶/۴	۹۰/۶	۹۴/۰	حساسیت	۹۷/۲	۹۰/۱	۹۶/۲	۸۴/۷	۹۲/۱
		ویژگی	۹۸/۷	۹۶/۱	۹۹/۴	۹۶/۰	۹۷/۶	ویژگی	۹۷/۸	۹۴/۸	۹۹/۶	۹۶/۲۰	۹۶/۶
		دقت	۹۵/۲	۹۳/۸	۹۶/۵	۹۰/۹	۹۴/۱	دقت	۹۷/۲	۹۰/۴	۹۶/۳	۸۵/۴	۹۲/۳

*داده‌های سطری و ستونی به ترتیب برچسب‌های صحیح و پیش‌بینی مدل‌ها را نمایش می‌دهند.

انگور پایه‌گذاری شد. با استفاده از ۸۰ ویژگی بافت استخراج شده، ماشین بردار پشتیبان و شبکه عصبی به ترتیب متوسط دقتی برابر با ۸۹/۹٪ و ۹۱/۷٪ برای داده‌های آزمون نشان دادند. با استفاده از الگوریتم زنبور ژنتیکی، تعداد ویژگی‌های مورد استفاده برای مدلسازی، به تعداد ۳۴ و ۴۶ ویژگی بهینه به ترتیب برای شبکه عصبی و ماشین بردار پشتیبان انتخاب شد که افزایش دقت هر دو روش مدلسازی را نیز در پی داشت. اما همچنان روش شبکه عصبی با متوسط دقت ۹۴/۱٪ دسته‌بندی قابل اعتمادتری نسبت به ماشین بردار پشتیبان با متوسط دقت ۹۲/۳٪ در دسته‌بندی داده‌های آزمون ارائه داد. قابل ذکر است که کاهش تعداد ویژگی‌ها، کاهش پیچیدگی زمانی روش پیشنهادی را نیز در پی دارد. بنابراین با استفاده از شبکه عصبی و ویژگی‌های بهینه، زمان الگوریتم به حدود یک سوم کاهش یافته و روش پیشنهادی بصورت بلادرنگ نیز قابل استفاده است. بنابراین روش پیشنهادی قابلیت تشخیص و دسته‌بندی بیماری‌های برگ انگور را به طور کارآمد دارا می‌باشد و می‌توان از این روش به عنوان روشی مورد اعتماد در صنعت کشاورزی بهره گرفت.

اگر چه (Sannakki et al. و Chakraborty et al. (2017) (2013) به ترتیب با گزارش میزان دقت ۱۰۰٪ و ۹۵٪/۳۰، دقت بالاتری نسبت به روش پیشنهادی در این مقاله دارند، اما در زمینه پایگاه داده تصاویر مورد استفاده برای مدل‌سازی به ترتیب با ۳۳ و ۴۰۰ تصویر، پایگاه داده به مراتب کوچک‌تری نسبت به این پژوهش با ۴۰۶۲ تصویر، را مورد ارزیابی قرار داده‌اند. همان‌گونه که نتایج نشان دادند، مدل مبتنی بر ویژگی‌های GLRM، دارای عملکرد بهتری نسبت به سایر ویژگی‌های استخراجی بود، در صورتی که (Sannakki et al. (2013) فقط از ویژگی‌های مستخرج از GLCM برای تشخیص دو نوع بیماری بهره برده است.

جمع‌بندی

در این مقاله به منظور کاهش هزینه و ضایعات ناشی از بیماری‌های درخت انگور، روشی خودکار برای شناسایی سریع و دقیق درختان انگور مبتلا به بیماری‌های پوسیدگی سیاه، اسکا و لکه ایزاریوپسیس ایجاد شد. با توجه به تغییرات حاصل از بیماری در برگ درخت؛ ایده اصلی روش پیشنهادی بر اساس استخراج ویژگی‌های بافتی GLH، GLCM، GLRM و LBP تصاویر برگ درخت

REFERENCES

- Alshamlan, H. M., Badr, G. H., & Alohal, Y. A. (2015). Genetic Bee Colony (GBC) algorithm: A new gene selection method for microarray cancer classification. *Computational Biology and Chemistry*, 56, 49-60.
- Anon. n.d. "Grapevine Disease Images | Kaggle." Retrieved July 4, 2021 (<https://www.kaggle.com/piyushmishra1999/plantvillage-grape>).
- Barbedo, J.G.A. (2016). A review on the main challenges in automatic plant disease identification based on visible range images: *Biosystems Engineering*, 144, 52-60.
- Barbedo, J.G.A. (2019). Plant disease identification from individual lesions and spots using deep learning. *Biosystems Engineering*, 180, 96-107.
- Chakraborty, A., Diganta G., & Hassanien A.E. (2017). Studies in computational intelligence artificial intelligence for sustainable development: Theory. *Practice and Future Applications*. Vol. 10.
- Dubey, S.R., & Jalal, A.S. (2016). Apple disease classification using color, texture and shape features from images. *Signal, Image and Video Processing*, 10(5), 819-826.
- Gonzalez, R.C., Woods, R.E., & Eddins, S.L. (2004). Digital image processing using MATLAB, P. *New Jersey-USA*.
- Gurubelli, Y., Malmathanraj, R., & Palanisamy, P. (2020). Texture and colour gradient features for grade analysis of pomegranate and mango fruits using kernel-SVM classifiers. In *2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS)* (pp. 122-126). IEEE.
- Haralick, R.M., Shanmugam, K., & Dinstein, I.H. (1973). Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, 6, 610-621.
- Karimi, N., Kondrood, R.R., & Alizadeh, T. (2017). An intelligent system for quality measurement of Golden Bleached raisins using two comparative machine learning algorithms. *Measurement*, 107, 68-76.
- Khojastehnazhand, M., & Ramezani, H. (2020). Machine vision system for classification of bulk raisins using texture features. *Journal of Food Engineering*, 271, 109864.

- Mohammed, K. K., Darwish, A., & Hassenian, A. E. (2021). Artificial intelligent system for grape leaf diseases classification. In *Artificial Intelligence for Sustainable Development: Theory, Practice and Future Applications* (pp. 19-29). Springer, Cham.
- Ojala, T., Pietikainen, M., & Maenpaa, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on pattern analysis and machine intelligence*, 24(7), 971-987.
- Revathi, P., & Hemalatha, M. (2012). Advance computing enrichment evaluation of cotton leaf spot disease detection using image edge detection. In *2012 Third International Conference on Computing, Communication and Networking Technologies (ICCCNT'12)* (pp. 1-5). IEEE.
- Sannakki, S.S., Rajpurohit, V.S., Nargund, V.B., & Kulkarni, P. (2013). Diagnosis and classification of grape leaf diseases using neural networks. In *2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT)* (pp. 1-5). IEEE.
- Scheck, H.J., Vasquez, S.J., Gubler, W.D., & Fogle, D. (1998a). First report of three *Phaeoacremonium* spp. causing young grapevine decline in California. *Plant Disease*, 82(5), 590-590.
- Scheck, H., Vasquez, S., Fogle, D., & Gubler, W. (1998b). Grape growers report losses to black-foot and grapevine decline. *California Agriculture*, 52(4), 19-23.
- Singh, J., & Kaur, H. (2018). A review on: Various techniques of plant leaf disease detection. In *2018 2nd International Conference on Inventive Systems and Control (ICISC)* (pp. 232-238). IEEE.
- Ye, Q., Jia, J., Manawasinghe, I.S., Li, X., Zhang, W., Mugnai, L., & Yan, J. (2021). *Fomitiporia punicata* and *Phaeoacremonium minimum* associated with Esca complex of grapevine in China. *Phytopathology Research*, 3(1), 1-17.
- Zhang, C., Zou, K., & Pan, Y. (2020). A method of apple image segmentation based on color-texture fusion feature and machine learning. *Agronomy*, 10(7), 972.