



Research Paper

Performance Evaluation of a Hybrid Transfer Learning-Based Classifier for Distinguishing Healthy Apples from Apples with Three Types of Diseases

Iman Ahmadi

Department of Genetics and Plant Production Engineering, Isf.C., Islamic Azad University, Isfahan, Iran

Article history	Abstract
Submitted: 2025/11/03	The aim of this research is the development of hybrid models based on transfer learning concept to classify apple images into four classes named: healthy apples, apples affected by bacterial disease, apples affected by fungal disease, and apples with tissue level disease propagation. A single-staged four-class model as well as six hybrid models were utilized for classification. Each hybrid model consisted of a three-class model followed by a two-class model. To train them, 850 images of the Kaggle dataset was used. The backbone of all transfer learning models was the EfficientNet algorithm. The models were implemented in the PyCharm environment, using the Python programming language. Evaluation of the models were carried out using the confusion matrix, as well as the calculation of the precision, accuracy, sensitivity, specificity, and F1-score criteria. According to the results of this study, the value of all performance evaluation criteria exceeded 0.94, indicating that the developed models performed robustly in classifying apple images; however, some hybrid models performed worse than the single-staged classifier. Therefore, the optimal configuration of hybrid classifiers cannot be assumed a priori; instead a systematic evaluation of all possible combinations is necessary to identify the most effective approach.
Revised: 2025/11/14	
Accepted: 2025/12/05	
Published online: 2025/12/21	
Keywords: Apple grading, Hybrid classifier, Machine vision, Transfer learning	
*Corresponding author: ImanAhmadi1358@iaau.ac.ir ORCID:  0000-0003-3492-9443 	

How to cite this paper:

Ahmadi, I. (2025). Performance Evaluation of a Hybrid Transfer Learning-Based Classifier for Distinguishing Healthy Apples from Apples with Three Types of Diseases. *Journal of Researches in Mechanics of Agricultural Machinery*. 37: 75-94. [https://dx.doi.org/ 10.22034/jrmam.2026.14805.742](https://dx.doi.org/10.22034/jrmam.2026.14805.742) (In Persian)



Authors retain the copyright and full publishing rights. Published by [Shahrood University](#). This article is an open access article licensed under the [Creative Commons Attribution 4.0 International \(CC BY 4.0\)](#)

<https://dx.doi.org/10.22034/jrmam.2026.14805.742>

EXTENDED ABSTRACT

Introduction

Recent advances in precision agriculture highlight the pivotal role of machine vision and deep learning in developing automated systems for fruit classification and grading. A wide range of studies have been conducted with the primary aim of replacing manual methods and improving accuracy and efficiency. These studies can be broadly categorized into three main approaches.

The first involves hybrid frameworks, where fast detection models such as YOLO are integrated with advanced classification architectures, achieving accuracies of approximately 92% across multiple fruit types. The second focuses on multi-stage systems, designed for specific fruits, in which the classification process is structured into sequential steps such as fruit type identification, freshness assessment, and ripeness evaluation. The third approach emphasizes explainable models, exemplified by frameworks such as XAI-FruitNet, which not only achieve very high accuracies (exceeding 97%) across diverse datasets but also provide interpretability in model decision-making. Overall, while these studies share a common objective—enhancing accuracy and overcoming the limitations of conventional approaches—they differ considerably in their technical methodologies and scope of application. Some prioritize feature integration, others employ multi-stage analysis, and yet others focus on explainability.

Building on these advancements, the present study proposes the use of sequential hybrid classifiers based on transfer learning to categorize apple images into four distinct classes: healthy apples, apples affected by bacterial disease, apples affected by fungal disease, and apples with tissue-level disease progression.

Material and Methods

The dataset used in this study was obtained from the Kaggle repository and consists of four categories of apple images:

- Bacterial Disease (Blotch): 234 images, divided into 146 for training and 88 for testing.
- Healthy Apples (Normal): 222 images, divided into 130 for training and 92 for testing.
- Advanced Tissue Rot (Rot): 226 images, divided into 134 for training and 92 for testing.
- Fungal Disease (Scab): 168 images, divided into 100 for training and 68 for testing.

The objective of this study was to classify apples into one of these four categories. Two classification strategies were investigated: a single unified four-class classifier and sequential hybrid classifiers. The hybrid classifier first classifies images into three classes, with one class combining two of the four original categories. Subsequently, a binary classifier further separates the images within this composite class into the two original groups.

Given the four-class problem, six possible hybrid configurations were constructed. The performance of each hybrid classifier was evaluated and compared with the unified four-class classifier. Using the abbreviations B, N, R, and S to represent the four categories (Blotch, Normal, Rot, and Scab, respectively), the classifiers investigated were as follows:

Unified four-class classifier: B–N–R–S

- Hybrid classifier BN–R–S followed by a binary classifier B–N
- Hybrid classifier BR–N–S followed by a binary classifier B–R
- Hybrid classifier BS–N–R followed by a binary classifier B–S
- Hybrid classifier B–NR–S followed by a binary classifier N–R
- Hybrid classifier B–NS–R followed by a binary classifier N–S
- Hybrid classifier B–N–RS followed by a binary classifier R–S

All classifiers employed the pre-trained EfficientNet model as the backbone, and the implementations were carried out in Python using the PyCharm environment. Performance evaluation was conducted using

confusion matrices and standard metrics including accuracy, precision, sensitivity, specificity, and F1-score.

Results and Discussion

By employing the proposed model instead of a conventional convolutional neural network, the number of trainable parameters was drastically reduced. This substantial reduction not only decreased the training time of the transfer learning-based model but also enabled its deployment on computers without GPU support.

All performance evaluation metrics exceeded 0.94, indicating that the developed model performs robustly in classifying apple images. These results confirm the suitability of the transfer learning approach using the pre-trained EfficientNet architecture for categorizing apples into four classes: Bacterial Disease (Blotch), Healthy (Normal), Advanced Tissue Rot (Rot), and Fungal Disease (Scab).

Regarding hybrid classifiers, the B–NR–S and BS–N–R configurations outperformed the unified four-class classifier, while the BR–N–S hybrid achieved comparable accuracy. In contrast, the BN–R–S, B–NS–R, and B–N–RS hybrid classifiers performed worse than the unified model. This variation can be explained by the fact that combining classes acts like a form of feature engineering, and since all possible hybrid combinations are considered, this feature engineering occurs stochastically. Consequently, some combinations may lead to reduced overall accuracy. Therefore, when developing hybrid classifiers, it is essential to account for this property and discard configurations that underperform relative to the unified classifier.

Conclusions

This study demonstrates that all developed models achieved robust, reliable performance, with all evaluation metrics exceeding 0.94, which can be attributed to the use of a sufficiently large image dataset and the incorporation of transfer learning algorithms. The consistent high performance across models underscores the effectiveness

of transfer learning in enhancing the accuracy and generalizability of automated apple classification systems.

A comparative analysis revealed that some hybrid models outperformed the unified four-class classifier, while others performed comparably worse. These findings indicate that the optimal configuration of hybrid classifiers cannot be assumed a priori; instead, a comprehensive development and systematic evaluation of all possible hybrid combinations are necessary, with each being benchmarked against the unified model to identify the most effective approach.

The relative advantage of the unified classifier stems primarily from its single-model architecture, in contrast to the paired structure of hybrid models. While some hybrid configurations offer potential improvements in classification accuracy, their deployment in an operational apple grading line requires additional hardware, which may pose practical challenges. This highlights a critical trade-off between achieving marginal performance gains and maintaining the feasibility of real-world implementation.

Overall, the findings of this study not only confirm the efficacy of transfer learning-based approaches for multi-class fruit classification but also provide important insights into the design considerations, performance trade-offs, and practical constraints associated with implementing hybrid versus unified classifier architectures in industrial agricultural settings.

Data Availability Statement

Data are available on request from the author.

Ethical Considerations

The author has utilized the Microsoft Copilot to summarize articles for writing the introduction section of the paper, and to translate Persian to English in writing the English abstract.

Conflict of Interest

The author declare that he has no known competing financial interest or personal

relationship that could have appeared to influence the work reported in this paper.

Funding Statement

The author received no specific funding for this research.



مقاله پژوهشی

ارزیابی کارکرد دسته‌بند ترکیبی بر مبنای یادگیری انتقالی در جداسازی تصاویر سیب سالم از سیب‌های دارای سه نوع بیماری

ایمان احمدی

گروه مهندسی تولید و ژنتیک گیاهی، واحد اصفهان (خوراسگان)، دانشگاه آزاد اسلامی، اصفهان، ایران

تاریخچه مقاله	چکیده
دریافت: ۱۴۰۴/۰۸/۱۱	هدف این پژوهش، استفاده از دسته‌بندهای ترکیبی مبتنی بر یادگیری انتقالی برای طبقه‌بندی تصاویر سیب در چهار گروه شامل سیب سالم، سیب مبتلا به بیماری باکتریایی، سیب مبتلا به بیماری قارچی و سیب دارای بیماری گسترش‌یافته به بافت است. در این روش، یک دسته‌بند سه طبقه به عنوان گام نخست و سپس یک دسته‌بند دوطبقه به کار گرفته شد. همچنین، مقایسه عملکرد این دسته‌بندهای ترکیبی با یک دسته‌بند چهار طبقه مستقل، یکی دیگر از اهداف پژوهش بود. برای توسعه مدل‌ها، تعداد ۸۵۰ تصویر از پایگاه داده Kaggle استفاده شد. در تمام دسته‌بندها از مدل یادگیری انتقالی بر مبنای الگوریتم EfficientNet، بهره گرفته شد. پیاده‌سازی مدل‌ها با زبان Python در محیط PyCharm انجام شد. به منظور ارزیابی عملکرد، از ماتریس درهم‌ریختگی و معیارهای دقت، صحت، حساسیت، ویژگی (Specificity) و نمره F1 استفاده گردید. نتایج نشان داد که به دلیل حجم مناسب داده‌ها و استفاده از الگوریتم یادگیری انتقالی، همه مدل‌ها عملکرد مطلوبی داشتند و مقدار همه شاخص‌های ارزیابی بالاتر از ۰/۹۴ به دست آمد. مقایسه دسته‌بندها نشان داد که برخی از مدل‌های ترکیبی بهتر از مدل یکپارچه چهار طبقه عمل کردند و برخی دیگر عملکرد ضعیف‌تری داشتند. علت این تفاوت می‌تواند به نحوه ترکیب گروه‌ها بازگردد؛ چراکه این ترکیب‌ها مشابه نوعی مهندسی ویژگی عمل کرده و به صورت تصادفی ممکن است منجر به بهبود یا افت صحت کلی مدل شوند.
واژه‌های کلیدی: بینایی ماشین، درجه‌بندی سیب، دسته‌بند ترکیبی، یادگیری انتقالی	
*پست الکترونیکی نویسنده مسئول: ImanAhmadi1358@iaui.ac.ir ORCID:  ۰۰۰۰-۰۰۰۳-۳۴۹۲-۹۴۴۳	



نحوه استناد به این مقاله:

احمدی، ا. (۱۴۰۴). ارزیابی کارکرد دسته‌بند ترکیبی بر مبنای یادگیری انتقالی در جداسازی تصاویر سیب سالم از سیب‌های دارای سه نوع بیماری. نشریه پژوهش‌های مکانیک ماشین‌های کشاورزی، ۳۷: ۷۵-۹۴. شناسه دیجیتال:

<https://dx.doi.org/10.22034/jrmam.2026.14805.742>

مقدمه

در سال‌های اخیر استفاده از یادگیری عمیق و شبکه‌های عصبی کانولوشنی (CNN) به عنوان رویکردی نوین در پردازش تصویر برای طبقه‌بندی و ارزیابی کیفی محصولات کشاورزی مورد توجه گسترده قرار گرفته است. Golzar et al. (2024) یک الگوریتم یادگیری عمیق برای دسته‌بندی خیارها به سه گروه سالم، معیوب و خمیده ارائه کردند. آن‌ها با معرفی یک شبکه کانولوشنی اصلاح‌شده (MCNN) مجهز به لایه *global average pooling (GAP)* و مقایسه آن با روش *InceptionResNetV2 transfer learning* نشان دادند که مدل پیشنهادی با دقت ۹۴/۱۴٪ بالاترین عملکرد را داشته و استفاده از لایه‌های GAP برتری محسوسی نسبت به لایه‌های fully connected در این مسئله دارد.

در همین راستا، Vidyarthi et al. (2020) با به کارگیری شبکه‌های عصبی کانولوشنی عمیق همراه با پردازش تصویر، به طبقه‌بندی بادام‌هندی در پنج دسته بر اساس میزان ناخالصی پرداختند. چهار مدل *DCNN* شامل *Inception-V3*، *ResNet50*، *VGG-16* و یک مدل سفارشی در این پژوهش مورد استفاده قرار گرفت که همگی دقت بالای ۹۵/۱٪ را کسب کردند. در میان آن‌ها، *Inception-V3* و *ResNet50* با دقت ۹۸/۴٪ و امتیاز *F1-score* برابر با ۹۶٪ بهترین نتایج را ارائه دادند.

در مطالعه‌ای دیگر، Jahanbakhshi et al. (2020) توسعه یک الگوریتم CNN بهبودیافته مجهز به *stochastic pooling* موفق به شناسایی عیوب ظاهری لیموی ترش و طبقه‌بندی آن‌ها به دو گروه سالم و آسیب‌دیده شدند. مقایسه این روش با تکنیک‌های سنتی استخراج ویژگی و طبقه‌بندی همچون *HOG*، *LBP*، *KNN*، *ANN*، *Fuzzy*، و *SVM* و درخت تصمیم نشان داد که CNN پیشنهادی با دقت ۱۰۰٪ توانسته است عملکردی برتر در ارزیابی غیرمخرب این محصول ارائه دهد.

به همین ترتیب، Vidyarthi et al. (2021) چهار مدل *DCNN* شامل *Inception-V3*، *ResNet50*، *VGG-16* و یک مدل سفارشی را برای طبقه‌بندی بادام‌ها بر اساس سطح ناخالصی ناشی از بادام‌های شکسته (۰٪ تا ۳۵٪) به کار بردند. نتایج نشان داد که *ResNet50* با دقت ۹۹٪ بهترین عملکرد را داشته است، در حالی که مدل

سفارشی با دقت ۹۲/۲٪ ضعیف‌ترین نتیجه را کسب کرده و اندکی *underfitting* در منحنی خطای آن مشاهده شد. در پژوهشی جدیدتر، Farahani and Bagherpour (2025) یک سیستم خودکار مبتنی بر یادگیری عمیق برای شناسایی آلودگی قارچی در فندق‌ها توسعه دادند. این سیستم با تحلیل تغییرات جزئی رنگ پوسته و بهینه‌سازی عواملی همچون اندازه ورودی تصویر، روش‌های تخت‌سازی (*flattening*) و معماری شبکه، به این نتیجه رسید که مدلی با چهار لایه کانولوشنی، لایه GAP و نرخ *dropout* برابر با ۰.۵ بهترین عملکرد را دارد. این مدل CNN توانست با دقتی بالاتر از روش‌های دستی (بهبود ۰/۸٪) در تفکیک فندق‌های سالم از آلوده عمل کند.

در مطالعه‌ای توسط Kayaalp (2024) روشی مبتنی بر یادگیری عمیق گروهی (*deep ensemble*) برای طبقه‌بندی هفت گونه مختلف گیلاس ارائه داده شده است. این پژوهش با استفاده از مجموعه داده‌ای شامل ۳۵۷۰ تصویر از منطقه اسپاردا در ترکیه، شش مدل یادگیری عمیق را آموزش داد که در میان آن‌ها مدل *DenseNet169* با دقت ۹۹/۵۷٪ بهترین عملکرد را به صورت مستقل داشت. با ترکیب دو مدل برتر از طریق یادگیری گروهی و رأی‌گیری حداکثری، این سیستم به دقت طبقه‌بندی ۱۰۰٪ دست یافت.

در ادامه این روند Tripathi and Maktedar (2022) یک سامانه غیرمخرب و خودکار برای درجه‌بندی انبه توسعه داده‌اند که بر اساس ویژگی‌های داخلی نظیر محتوای مواد جامد محلول و میزان اسید کل، انبه‌ها را در سه دسته طبقه‌بندی می‌کند. این سامانه با استفاده از ماتریس هم‌وقوع سطح خاکستری، ماتریس طول اجرای سطح خاکستری و طیف‌سنجی نزدیک به مادون قرمز، ویژگی‌ها را استخراج کرده و سپس با بهره‌گیری از تحلیل مؤلفه‌های اصلی مبتنی بر وصله، ابعاد داده‌ها را کاهش می‌دهد. در مرحله طبقه‌بندی، از یک طبقه‌بند گروهی (*ensemble learning*) شامل الگوریتم‌های ماشین بردار پشتیبان، جنگل تصادفی، شبکه‌های عصبی مصنوعی و نزدیک‌ترین همسایه استفاده شده است، که یکی از شبکه‌های عصبی با الگوریتم بهینه‌سازی نهنگ مبتنی بر ماسک ترکیبی Lion-Binary بهینه‌سازی شده است.

Anjali et al. (2023) یک سامانه چندمرحله‌ای برای دسته‌بندی موز معرفی کرده‌اند که با بهره‌گیری از

مدل‌های مختلف در سه مرحله شامل شناسایی نوع میوه، سنجش تازگی و ارزیابی رسیدگی عمل می‌کند. در همین راستا، *Sultana et al. (2024)* مدل XAI-FruitNet را پیشنهاد داده‌اند؛ مدلی تبیینی که با ترکیب روش‌های average pooling و max pooling توانسته در چندین مجموعه داده به دقتی بیش از ۹۷ درصد دست یابد و در عین حال امکان تفسیر نتایج را فراهم کند.

وجه اشتراک این پژوهش‌ها تأکید بر کاربردهای کشاورزی دقیق و استفاده از یادگیری عمیق برای طبقه‌بندی میوه‌هاست. همه آن‌ها تلاش کرده‌اند محدودیت‌های روش‌های دستی را برطرف کنند و دقت و کارایی را بهبود بخشند. تفاوت‌ها بیشتر در رویکردهای فنی دیده می‌شود، به عنوان مثال درحالی‌که *Anjali et al. (2023)* رویکرد چندمرحله‌ای برای میوه خاصی را دنبال کرده‌اند، *Sultana et al. (2024)* بر تبیین‌پذیری مدل تأکید داشته‌اند. همچنین دامنه‌ی مطالعات نیز متفاوت است؛ برخی به طبقه‌بندی کلی انواع میوه پرداخته‌اند و برخی دیگر بر کاربردهای خاص مانند ارزیابی میزان رسیدگی تمرکز داشته‌اند.

در اکثر مقاله‌های بررسی شده منظور از روش‌های ترکیبی، به کارگیری چندین الگوریتم یادگیری ماشین و یا یادگیری عمیق روی یک پایگاه داده‌ای و ادغام نتایج با رأی‌گیری اکثریت (روشی که با عنوان Ensemble learning از آن یاد می‌شود) بوده‌است و به ایده ترکیب چند الگوریتم به صورت سری‌وار به‌ویژه در پژوهش‌های کشاورزی کمتر پرداخته شده‌است. تحقیق حاضر بیشترین شباهت را با پژوهش *Anjali et al. (2023)* دارد که آنها هم از ایده روش ترکیبی سری‌وار در دسته‌بندی موز استفاده کرده‌اند، منتها مدل‌های به کار رفته توسط آنها برای مراحل مختلف دسته‌بندی با یکدیگر متفاوت بود اما در این پژوهش از یک مدل ثابت در مراحل مختلف دسته‌بندی میوه سیب استفاده می‌شود.

بنابراین هدف از پژوهش حاضر استفاده از دسته‌بندهای بر مبنای یادگیری انتقالی به صورت ترکیبی سری‌وار به منظور دسته‌بندی تصاویر سیب به چهار دسته سیب دارای بیماری باکتریایی، سیب سالم، سیب دارای بیماری گسترش یافته به بافت و سیب دارای بیماری قارچی است. دسته‌بندهای ترکیبی موردنظر در این

پژوهش از ترکیب یک دسته‌بند سه طبقه که پس از آن یک دسته‌بند دوطبقه قرارگرفته تشکیل شده‌اند. بنابراین با توجه به چهار گروه سیب مورد مطالعه، شش دسته‌بند ترکیبی قابل‌تصور است که مقایسه کارایی این دسته‌بندهای ترکیبی با دسته‌بند واحد چهارطبقه‌ای هدف دیگری است که در این پژوهش دنبال می‌شود.

مواد و روش‌ها

تصاویر مورد استفاده در این پژوهش از پایگاه داده‌ای Kaggle^۱ به دست آمد (*JayaprakashPondy, 2023*).

تصاویر شامل چهار گروه میوه سیب بود که عبارتند از:

۱- تصاویر سیب دارای بیماری باکتریایی (Blotch)

شامل ۲۳۴ تصویر که به دو گروه آموزش (۱۴۶ تصویر) و آزمون (۸۸ تصویر) تقسیم شده بود.

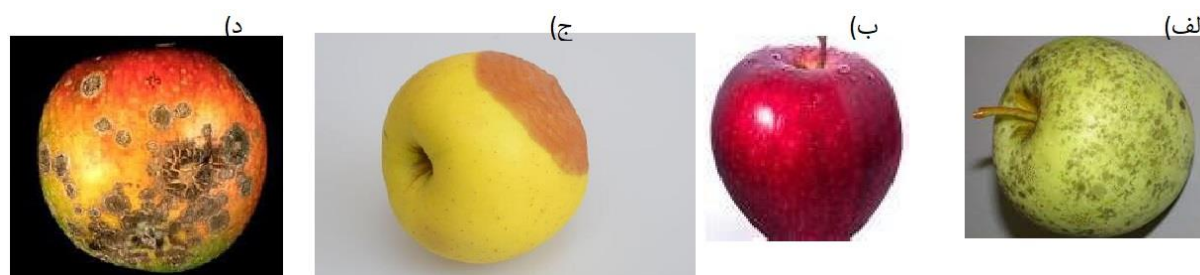
۲- تصاویر سیب سالم (Normal) شامل ۲۲۲ تصویر که به دو گروه آموزش (۱۳۰ تصویر) و آزمون (۹۲ تصویر) تقسیم شده بود.

۳- تصاویر سیب دارای بیماری پیشرفته گسترش یافته به بافت میوه (Rot) شامل ۲۲۶ تصویر که به دو گروه آموزش (۱۳۴ تصویر) و آزمون (۹۲ تصویر) تقسیم شده بود.

۴- تصاویر سیب دارای بیماری قارچی (Scab) شامل ۱۶۸ تصویر که به دو گروه آموزش (۱۰۰ تصویر) و آزمون (۶۸ تصویر) تقسیم شده بود.

در شکل ۱ جهت مقایسه، نمونه‌ای از تصاویر متعلق به هر کدام از چهار گروه تصاویر نشان داده شده است.

هدف این پژوهش دسته‌بندی میوه سیب در یکی از این چهار دسته است. برای انجام این کار می‌توان از یک دسته‌بند یکپارچه چهارطبقه استفاده کرد و یا دسته‌بندهای ترکیبی را به کار برد. دسته‌بند ترکیبی این پژوهش ابتدا تصاویر سیبها را به سه طبقه دسته‌بندی می‌کند به طوری که یکی از طبقات شامل دو گروه از گروه‌های چهارگانه محصول باشد. سپس یک بسته‌بند دوطبقه تصاویر عبور کرده از طبقه مرکب دسته‌بند اول را به دو گروه تجزیه می‌کند. از آنجا که در این پژوهش با چهار گروه از تصاویر مواجه هستیم، شش دسته ترکیبی به وجود می‌آید.



شکل ۱- نمونه‌ای از تصاویر متعلق به گروه الف) سیب دارای بیماری باکتریایی (Blotch)، ب) سیب سالم (Normal)، ج) سیب دارای بیماری گسترش یافته به بافت میوه (Rot) و د) سیب دارای بیماری قارچی (Scab)

دسته‌بند BN با ۲۷۶ تصویر آموزشی مربوط به گروه‌های Blotch و Normal آموزش داده شد و عملکرد آن بر روی ۱۸۰ تصویر آزمونی گروه‌های Blotch و Normal آزمایش شد. دسته‌بند BR با ۲۸۰ تصویر آموزشی مربوط به گروه‌های Blotch و Rot آموزش داده شد و عملکرد آن بر روی ۱۸۰ تصویر آزمونی گروه‌های Blotch و Rot آزمایش شد. دسته‌بند BS با ۲۴۶ تصویر آموزشی مربوط به گروه‌های Blotch و Scab آموزش داده شد و عملکرد آن بر روی ۱۵۶ تصویر آزمونی گروه‌های Blotch و Scab آزمایش شد. دسته‌بند NR با ۲۶۴ تصویر آموزشی مربوط به گروه‌های Blotch و Rot آموزش داده شد و عملکرد آن بر روی ۱۸۴ تصویر آزمونی گروه‌های Blotch و Rot آزمایش شد. دسته‌بند NS با ۲۳۰ تصویر آموزشی مربوط به گروه‌های Scab و Normal آموزش داده شد و عملکرد آن بر روی ۱۶۰ تصویر آزمونی گروه‌های Scab و Normal آزمایش شد و در نهایت، دسته‌بند RS با ۲۳۴ تصویر آموزشی مربوط به گروه‌های Scab و Rot آموزش داده شد و عملکرد آن بر روی ۱۶۰ تصویر آزمونی گروه‌های Scab و Rot آزمایش شد. مدل مورد استفاده در این پژوهش برای همه دسته‌بندها، مدل یادگیری انتقالی بر مبنای الگوریتم از پیش‌آموزش دیده EfficientNet بود. در جدول ۱ به تنظیمات به‌کار برده شده در مدل‌ها اشاره شده است:

جدول ۱- پارامترهای تنظیمی مدل‌های مورد استفاده در این

پژوهش	پارامتر
مقدار تنظیمی	تعداد دوره‌های آموزشی (epoch)
۳۰	تعداد تصاویر مربوط به گروه‌های تصویری
۳۲	استفاده شده در آموزش (batch size)
Adam	تابع بهینه‌ساز (optimizer)
۰/۰۰۱	نرخ یادگیری (learning rate)
CrossEntropy	تابع هزینه (cost function)

اگر حروف B، N، R و S به‌عنوان مخفف گروه‌های چهارگانه Blotch، Normal، Rot و Scab در نظر گرفته شوند، دسته‌بندهای مورد بررسی عبارتند از:

- دسته بند یکپارچه چهارطبقه‌ای B-N-R-S
- دسته‌بند سه‌طبقه‌ای BN-R-S که بعد از آن دسته‌بند دوطبقه‌ای B-N قرار گرفته است.
- دسته‌بند سه‌طبقه‌ای BR-N-S که بعد از آن دسته‌بند دوطبقه‌ای B-R قرار گرفته است.
- دسته‌بند سه‌طبقه‌ای BS-N-R که بعد از آن دسته‌بند دوطبقه‌ای B-S قرار گرفته است.
- دسته‌بند سه‌طبقه‌ای B-NR-S که بعد از آن دسته‌بند دوطبقه‌ای N-R قرار گرفته است.
- دسته‌بند سه‌طبقه‌ای B-NS-R که بعد از آن دسته‌بند دوطبقه‌ای N-S قرار گرفته است.
- دسته‌بند سه‌طبقه‌ای B-N-RS که بعد از آن دسته‌بند دوطبقه‌ای R-S قرار گرفته است.

شایان ذکر است که برای ایجاد امکان مقایسه دسته‌بندهای ترکیبی با دسته‌بند یکپارچه، نتایج به‌دست‌آمده از دو بخش تشکیل دهنده دسته‌بندهای ترکیبی با هم ترکیب شده به‌گونه‌ای که ارزیابی با توجه به خروجی‌های چهارگانه هر دسته‌بند انجام شود و ارزیابی دو بخش تشکیل دهنده دسته‌بندهای ترکیبی به‌صورت مستقل مورد نظر نیست. مدل مورد استفاده برای دسته‌بندهای مطرح در این پژوهش مدل بر مبنای الگوریتم یادگیری انتقالی است (Ahmadi, 2025). برای توسعه مدل‌های بر مبنای دسته‌بند یکپارچه چهارطبقه‌ای و کلیه دسته‌بندهای سه‌طبقه‌ای از تمام ۸۵۰ تصویر پایگاه داده‌ای به شکل زوج مرتب (۳۴۰، ۵۱۰) برای زوج (تعداد تصاویر آزمونی، تعداد تصاویر آزمایشی) استفاده شد. اما تعداد تصاویر مورد استفاده در توسعه دسته‌بندهای دوطبقه‌ای به شکل زیر است:

این معیار نسبت تعداد نمونه‌هایی که به‌درستی «مثبت» پیش‌بینی شده‌اند، به کل نمونه‌های پیش‌بینی شده «مثبت» را اندازه‌گیری می‌کند. دقت با استفاده از رابطه (۳) محاسبه می‌شود:

$$\text{Precision} = \frac{TP}{TP+FP} \times 100 \quad (3)$$

امتیاز (F1-score) F1:

امتیاز F1 میانگین همساز دقت و حساسیت است که از فرمول (۴) محاسبه می‌شود:

$$\text{F1 score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \times 100 \quad (4)$$

صحت (Accuracy):

این معیار نسبت تعداد پیش‌بینی‌های صحیح مدل به کل پیش‌بینی‌ها را نشان می‌دهد. محاسبه صحت از طریق فرمول (۵) صورت می‌گیرد:

$$\text{Accuracy} = \frac{\sum_{i=1}^n (TP_i)}{\sum_{i=1}^n (TP_i + FP_i)} \times 100 \quad (5)$$

که در آن n تعداد طبقه‌هاست.

نتایج و بحث

به‌کارگیری مدل یادگیری انتقالی بر مبنای الگوریتم از پیش آموزش دیده EfficientNet، به‌جای مدل شبکه عصبی پیچشی معمولی، باعث شد تعداد پارامترهای آموزشی کاهش قابل ملاحظه‌ای پیدا کند که در نهایت باعث کاهش زمان آموزش مدل بر مبنای یادگیری انتقالی شد و اجرای مدل را روی رایانه بدون GPU امکانپذیر ساخت. نمودارهای صحت و هزینه مراحل آموزش و آزمون برای دسته‌بند یکپارچه چهار طبقه‌ای بر حسب شماره دوره آموزشی در شکل ۳ نشان داده شده است.

افقی شدن نمودارهای صحت و تابع هزینه در شماره دوره‌های نزدیک به ۳۰، نشان دهنده این است که فرآیند آموزش مدل به‌خوبی انجام شده است. برای سایر دسته‌بندهای مورد مطالعه نیز نمودارهای مشابهی به‌دست آمد. شکل‌های ۴ الف، ب و ج بترتیب ماتریس‌های درهم‌ریختگی مدل یکپارچه چهار طبقه‌ای، مدل سه طبقه‌ای B-RN-S و مدل دوطبقه‌ای R-N را نشان می‌دهد:

کدهای مورد استفاده در محیط برنامه‌نویسی pycharm (2023.3.4) و با استفاده از کتابخانه‌های زبان برنامه‌نویسی python(3.11.0) شامل PIL(10.2.0)، torch(2.2.0) و numpy(1.26.3) ایجاد شد. برای بهتر مشخص شدن معماری روش یادگیری انتقالی مورد استفاده در این پژوهش، در شکل ۲ لایه‌های مدل مربوط به دسته‌بند چهار طبقه‌ای نشان داده شده است.

معیارهای ارزیابی مدل

ماتریس سردرگمی (Confusion Matrix):

برای پی‌بردن به محل پیش‌بینی‌های اشتباه و اینکه گروه مورد نظر با کدام گروه اشتباه شده، از ماتریس سردرگمی استفاده می‌شود. هر چه تمرکز اعداد بزرگ‌تر در خانه‌های قطر اصلی ماتریس بیشتر باشد (و مقدارهای صفر یا اعداد کوچک در سایر خانه‌ها دیده شود)، مدل عملکرد بهتری دارد.

حساسیت (توان تشخیص مثبت‌ها) (Sensitivity):

این معیار نشان می‌دهد که مدل تا چه اندازه توانایی تشخیص نمونه‌های مثبت را دارد. به‌صورت ریاضی، حساسیت از طریق فرمول (۱) محاسبه می‌شود:

$$\text{Sensitivity (Recall)} = \frac{TP}{TP+FN} \times 100 \quad (1)$$

در رابطه (۱)، TP تعداد نمونه‌هایی است که به‌درستی «مثبت» پیش‌بینی شده‌اند، و FN تعداد نمونه‌هایی است که به اشتباه «منفی» پیش‌بینی شده‌اند.

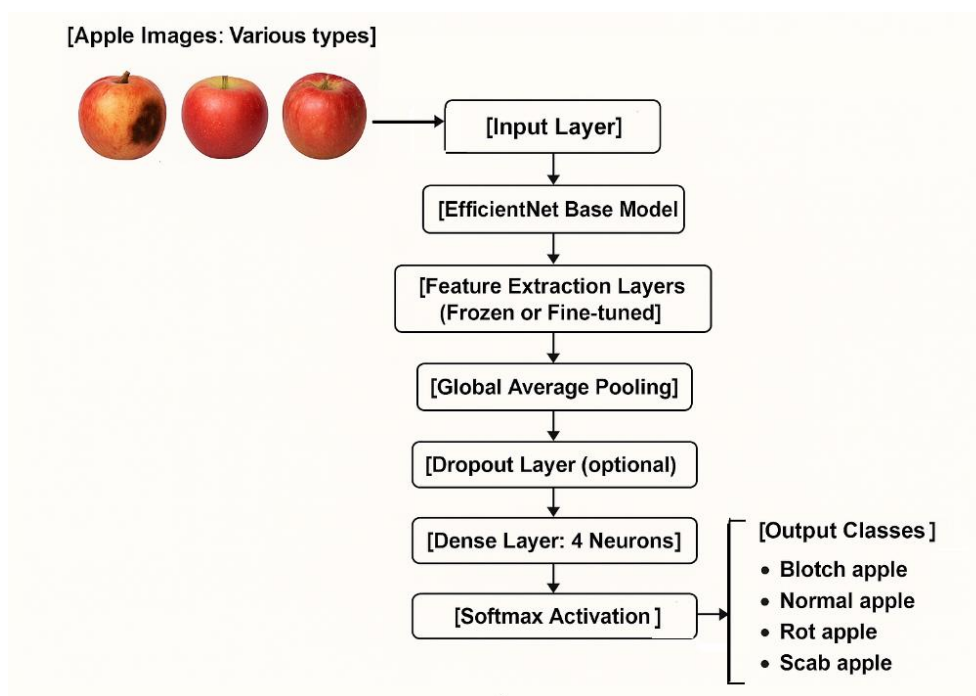
ویژگی (توان تشخیص منفی‌ها) (Specificity):

این معیار توانایی مدل در تشخیص صحیح نمونه‌های منفی هر دسته را نشان می‌دهد. فرمول آن در رابطه (۲) آورده شده است:

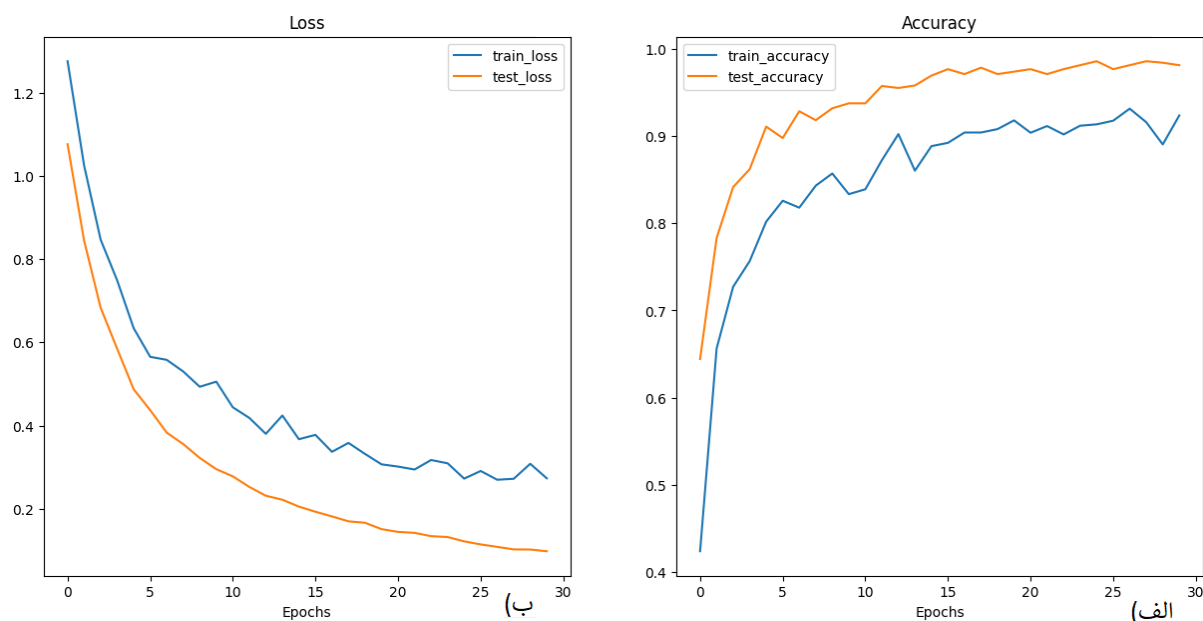
$$\text{Specificity} = \frac{TN}{TN+FP} \times 100 \quad (2)$$

در رابطه (۲)، TN تعداد نمونه‌هایی است که به‌درستی «منفی» پیش‌بینی شده‌اند، و FP تعداد نمونه‌هایی است که به‌اشتباه «مثبت» پیش‌بینی شده‌اند.

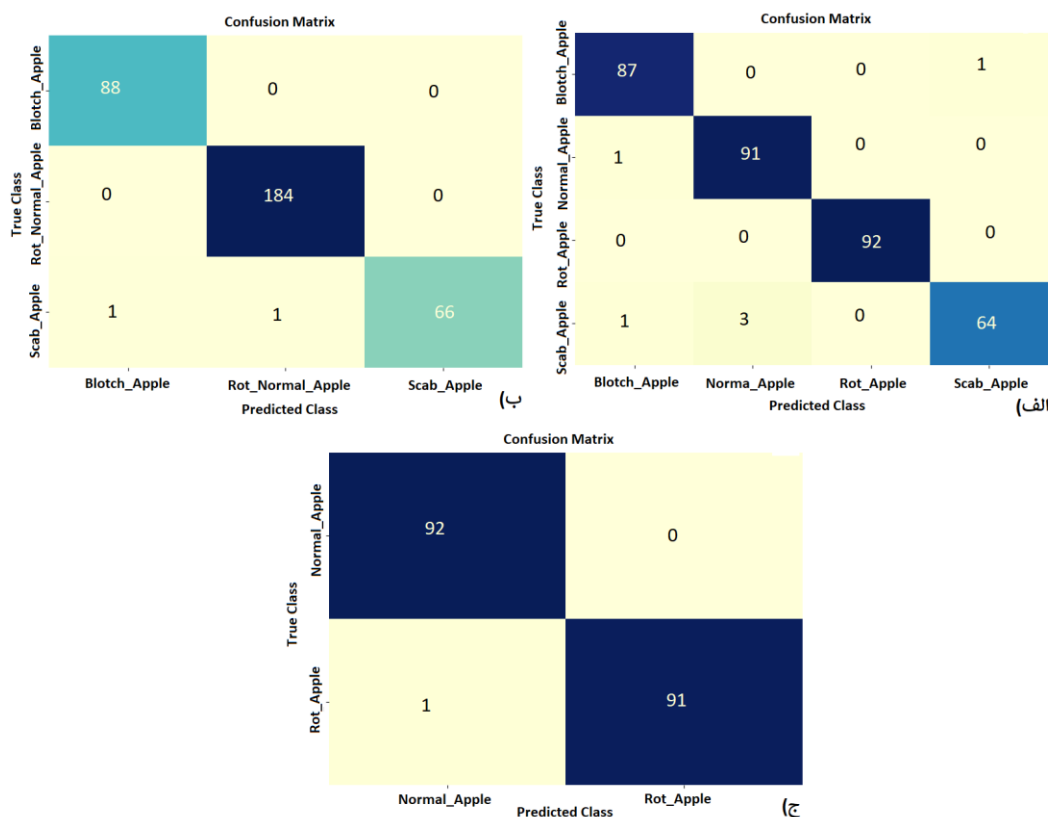
دقت (صحت مثبت‌ها) (Precision):



شکل ۲- معماری مدل یادگیری انتقالی مربوط به دسته‌بند چهارطبقه‌ای



شکل ۳- نمودارهای الف) صحت و ب) تابع هزینه مربوط به دسته‌بند یکپارچه چهارطبقه‌ای بر حسب شماره دوره آموزشی



شکل ۴- ماتریس‌های درهم‌ریختگی حاصل از (الف) مدل یکپارچه چهارطبقه‌ای، (ب) مدل سه طبقه B-RN-S و (ج) مدل دوطبقه R-N

است، بنابراین می‌توان از مناسب بودن مدل یادگیری انتقالی توسعه‌یافته با به‌کارگیری الگوریتم از پیش‌آموزش دیده EfficientNet در دسته‌بندی میوه سیب به چهار دسته دارای بیماری باکتریایی (Blotch)، سالم (Normal)، دارای بیماری گسترش یافته به بافت میوه (Rot) و دارای بیماری قارچی (Scab) اطمینان داشت.

جدول ۲ مقادیر مثبت واقعی، مثبت کاذب، منفی کاذب و منفی واقعی به‌دست آمده از ماتریس درهم‌ریختگی مدل یکپارچه چهارطبقه‌ای، همچنین معیارهای ارزیابی‌کننده حساسیت، ویژگی، دقت، نمره F1 و صحت محاسبه شده را نشان می‌دهد. تمام شاخصهای ارزیابی‌کننده کارکرد مدل مقادیر بالاتر از ۰/۹۴ را دارند که نشان‌دهنده کارکرد مطلوب مدل یکپارچه چهارطبقه‌ای در دسته‌بندی تصاویر میوه سیب

جدول ۲- مقادیر معیارهای ارزیابی‌کننده مدل یکپارچه چهارطبقه‌ای

وضعیت میوه	مثبت واقعی	مثبت کاذب	منفی کاذب	منفی واقعی	حساسیت	ویژگی	دقت	امتیاز F1	صحت
Blotch	87	2	1	250	0.989	0.992	0.978	0.983	0.982
Normal	91	3	1	245	0.989	0.988	0.968	0.978	
Rot	92	0	1	247	0.989	1.000	1.000	0.994	
Scab	64	1	4	271	0.941	0.996	0.985	0.962	

جدول ۳- معیارهای ارزیابی دسته‌بند ترکیبی B-RN-S

دسته بند	وضعیت میوه	مثبت واقعی	مثبت کاذب	منفی کاذب	منفی واقعی	حساسیت	ویژگی	دقت	امتیاز F1	صحت
سه طبقه‌ای B-RN-S	Blotch	88	1	0	251	1.000	0.996	0.989	0.994	0.994
	Rot-Normal	184	1	0	155	1.000	0.994	0.995	0.997	
	Scab	66	0	2	272	0.971	1.000	1.000	0.985	
دوطبقه‌ای R-N	Normal	1	0	91	1	0.989	0.989	0.994	1.000	0.995
	Rot	0	1	92	0.989	1.000	1.000	0.994	0.000	

جدول ۴- معیارهای ارزیابی نهایی به‌دست آمده از مدل ترکیبی B-RN-S که با نتایج مدل یکپارچه چهارطبقه‌ای قابل مقایسه است

وضعیت میوه	مثبت واقعی	مثبت کاذب	منفی کاذب	منفی واقعی	حساسیت	ویژگی	دقت	امتیاز F1	صحت
Blotch	88	1	0	251	1.000	0.996	0.989	0.994	0.988
Normal	91.502	0.994	0	92.502	1.000	0.989	0.989	0.994	
Rot	90.508	0	0.994	93.497	0.989	1.000	1.000	0.994	
Scab	66	0	2	272	0.971	1.000	1.000	0.985	

$$TP_{converted} = TP_{two\ class} \times HI_{combined\ class} \quad (۶)$$

جایی که $HI_{combined\ class}$ برهمکنش همساز طبقه مرکب از دسته‌بند سه‌طبقه‌ای است که مقدار آن از رابطه (۷) محاسبه می‌شود:

$$HI_{combined\ class} = sensitivity_{combined\ class} \times precision_{combined\ class} \quad (۷)$$

به‌طور مشابه برای تبدیل مقادیر FP و FN وضعیت‌های Rot و Normal از دسته‌بند دوطبقه‌ای به حالت‌های قابل مقایسه با دسته‌بند یکپارچه از روابط (۸) و (۹) استفاده می‌شود:

$$FP_{converted} = FP_{two\ class} \times HI_{combined\ class} \quad (۸)$$

$$FN_{converted} = FN_{two\ class} \times HI_{combined\ class} \quad (۹)$$

و درنهایت $TN_{converted}$ از رابطه (۱۰) محاسبه می‌شود:

$$TN_{converted} = (TP + FP)_{combined\ class} - (TP + FP + FN)_{converted} \quad (۱۰)$$

پس از محاسبه مقادیر TP, FP, FN و TN مربوط به وضعیت‌های Blotch و Scab, معیارهای ارزیابی حساسیت, ویژگی, دقت, نمره F1 و صحت با استفاده از مقادیر TP, FP, FN و TN محاسبه شد.

با توجه به مقادیر عددی جدول ۴، با وجود اینکه فرآیند ترکیب باعث تعدیل معیارهای ارزیابی‌کننده شده‌است، اما برتری دسته‌بند مرکب بر دسته‌بند یکپارچه چهارطبقه‌ای از مقادیر به‌دست آمده مشهود است، به‌عنوان مثال مقایسه مقادیر معیارهای ارزیابی‌کننده جدول‌های ۳ و ۵ نشان می‌دهد که استفاده از مدل ترکیبی کارایی دسته‌بند را بهبود داده است. به‌عنوان مثال شاخص صحت کلی دسته‌بند از ۰/۹۸۲ در دسته‌بند ترکیبی B-RN-S رسیده است.

سؤالی که وجود دارد این است که آیا تمام دسته‌بندهای ترکیبی کارایی بالاتری نسبت به دسته‌بند یکپارچه دارند؟ اگر جواب به سؤال بالا منفی باشد، ترکیب کدام گروه‌ها، دسته‌بند سه‌طبقه‌ای با کارایی بالاتر را به‌وجود می‌آورد؟

در جدول ۳ معیارهای ارزیابی‌کننده مدل‌های سه‌طبقه‌ای B-RN-S و دوطبقه‌ای R-N آورده شده است:

مشاهده می‌شود که مقادیر حساسیت، ویژگی، دقت و امتیاز F1 دسته‌بند سه‌طبقه‌ای B-RN-S برای وضعیت‌های میوه Blotch و Scab از مقادیر مشابه به‌دست آمده از دسته‌بند یکپارچه چهارطبقه‌ای بالاتر است. همچنین مقادیر حساسیت، ویژگی، دقت و امتیاز F1 حاصل از دسته‌بند دوطبقه‌ای برای وضعیت‌های میوه Normal و Rot حداقل با مقادیر مشابه به‌دست آمده از دسته‌بند یکپارچه چهارطبقه‌ای برابر هستند. بعلاوه دو بخش دسته‌بند ترکیبی B-RN-S از نظر معیار صحت از دسته‌بند یکپارچه چهارطبقه‌ای عملکرد بهتری دارند. اما تمام این مقایسه‌ها عملکرد دسته‌بند ترکیبی را در شرایطی نشان می‌دهد که بخش‌های تشکیل دهنده آن به‌صورت مستقل از هم عمل می‌کنند. برای انجام مقایسه‌ای جامع‌تر، دسته‌بندهای سه‌طبقه‌ای B-RN-S و دوطبقه‌ای R-N باهم ترکیب شده و معیارهای ارزیابی‌کننده حاصل که با مدل یکپارچه چهارطبقه‌ای قابل مقایسه است، در جدول ۴ نشان داده شده است.

برای تبدیل دو بخش دسته‌بند ترکیبی B-RN-S به دسته‌بند قابل مقایسه با دسته‌بند یکپارچه چهارطبقه‌ای از عملکرد برهمکنش همساز^۱ استفاده شد (Cullerne Bown, 2024). برهمکنش همساز حاصلضرب حساسیت در دقت است که از آن برای تبدیل مقادیر TP, FP, FN و TN دسته‌بند دوطبقه‌ای به مقادیر مشابه از دسته‌بند قابل مقایسه با دسته‌بند یکپارچه چهارطبقه‌ای استفاده می‌شود. توضیح اینکه مقادیر TP, FP, FN و TN مربوط به وضعیت‌های Blotch و Scab از دسته‌بند ترکیبی سه‌طبقه‌ای عیناً به موقعیت‌های مشابه از حالت قابل مقایسه با دسته‌بند یکپارچه چهارطبقه‌ای منتقل می‌شوند، اما برای تبدیل مقدار TP وضعیت‌های میوه Normal و Rot از دسته‌بند دوطبقه‌ای به حالت قابل مقایسه با دسته‌بند یکپارچه چهارطبقه‌ای از رابطه (۶) استفاده می‌شود:

برای پاسخ به این پرسشها، نتایج مقایسه‌ای معیار صحت دسته‌بندهای مختلف مورد بررسی در این پژوهش در جدول ۵ آورده شده است:

جدول ۵- مقادیر صحت به‌دست آمده از دسته‌بندهای مختلف

مورد بررسی در این پژوهش	
نام مدل	معیار صحت کلی به‌دست آمده
یکپارچه چهارطبقه‌ای	۰/۹۸۲
B-N-RS	۰/۹۷۶
B-NS-R	۰/۹۶۸
BS-N-R	۰/۹۸۵
B-NR-S	۰/۹۸۸
BR-N-S	۰/۹۸۲
BN-R-S	۰/۹۷۷

از داده‌های جدول ۵ مشخص می‌شود که کارایی دسته‌بندهای ترکیبی B-NR-S و BS-N-R از دسته‌بند یکپارچه چهارطبقه‌ای بالاتر است و دسته‌بند ترکیبی BR-N-S صحتی برابر با دسته‌بند یکپارچه چهارطبقه‌ای داشت، درحالی‌که دسته‌بندهای ترکیبی BN-R-S، B-NS-R و B-N-RS کارایی پائین‌تری نسبت به دسته‌بند یکپارچه چهارطبقه‌ای داشتند. دلیل این اتفاق می‌تواند به این صورت قابل توجیه باشد که ترکیب گروه‌ها، مشابه نوعی مهندسی ویژگی‌های مدلها عمل می‌کند و چون با لحاظ کردن تمام حالت‌های ترکیبی، این مهندسی ویژگیها به صورت تصادفی اتفاق می‌افتد، ممکن است حاصل ترکیب در جهت کاهش و یا افزایش صحت کلی مدل باشد. بنابراین در توسعه مدل‌های ترکیبی باید به این خصوصیت توجه داشت و مدل‌های دارای عملکرد پایین‌تر از مدل یکپارچه را کنار گذاشت.

از نظر مقایسه نتایج به‌دست آمده در این پژوهش با نتایج پژوهش‌های سایر محققین، (Unay et al., 2011) به کمک دوربین سیاه و سفید با رزولوشن بالا به جداسازی میوه سیب دارای دو رنگ از میوه تک رنگ اقدام کرده و با الگوریتم fuzzy k-NN به دقت ۸۳/۵٪ رسیدند.

در پژوهش دیگری با استفاده از ۱۰۰ تصویر به‌دست آمده از دوربین رنگی و با به‌کارگیری الگوریتم Adaptive fuzzy Neural Network که توسط Sabzi et al., (2015) برای دسته‌بندی میوه سیب انجام شد، صحت کلی به‌دست آمده برابر با ۹۹٪ بود.

همچنین Moallem et al., (2017) با به‌کارگیری الگوریتم SVM روی ۱۲۰ تصویر سیب به‌دست آمده از دوربین رنگی

به دسته‌بندی این محصول پرداختند و صحت ۹۲/۵٪ را به‌دست آوردند.

مشارکت نویسندگان

نحوه و میزان مشارکت نویسندگان در انجام این پژوهش به صورت زیر است:

- تمام بخشهای پژوهش شامل گردآوری داده، تحلیل و استخراج نتایج، ترسیم شکل و نگارش مقاله توسط تنها نویسنده مقاله انجام شده است.

دسترسی به داده‌ها

داده‌ها در صورت درخواست، در اختیار قرار خواهند گرفت.

اصول اخلاقی

در این پژوهش برای خلاصه‌سازی مقاله‌ها به‌منظور نگارش بخش مقدمه، همچنین برای ترجمه و نگارش چکیده مبسوط انگلیسی از دستیار هوش مصنوعی copilot استفاده شده است.

تضاد منافع نویسندگان

نویسنده اعلام می‌کند که هیچ گونه منافع مالی رقابتی یا روابط شخصی شناخته‌شده‌ای که ممکن است بر کار گزارش شده در این مقاله تأثیر گذاشته باشد، ندارند.

حمایت مالی

نویسنده هیچ حمایت مالی خاصی برای انجام این پژوهش دریافت نکرده است.

منابع

- Ahmadi, I. (2025). Detection and classification of some diseases of tomato crops using transfer learning. *Journal of Agricultural Machinery*, 15(), 319–335. 3. <https://doi.org/10.22067/jam.2024.88500.1258>
- Anjali, S., Pious, V., Sebastian, J. J., Krishnanunni, J., Johnson, J. K., Mujeeb, A., & Baselios, M. (2023). Multi-Stage Fruit Grading System. In *Lecture Notes in Networks and Systems* (Vol. 672). Springer.
- Cullerne Bown, W. (2024). Sensitivity and specificity versus precision and recall, and related dilemmas. *Journal of Classification*, 41(2), 402–426. <https://doi.org/10.1007/s00357-024-09478-y>
- Farahani, M., & Bagherpour, H. (2025). Using novel optimized deep learning techniques for detecting fungal infections in hazelnuts kernels based on shell

- otomatik sınıflandırılması]. *Tarım Bilimleri Dergisi*, 21(3), 326–336. https://doi.org/10.1501/tarimbil_0000001335
- Sultana, S., Tasir, M. A. M., Nobel, S. M. N., Kabir, M. M., & Mridha, M. F. (2024). XAI-FruitNet: An explainable deep model for accurate fruit classification. *Journal of Agriculture and Food Research*, 18(9), 101474. <https://doi.org/10.1016/j.jafr.2024.101474>
- Tripathi, M. K., & Maktedar, D. D. (2022). Internal quality assessment of mango fruit: An automated grading system with ensemble classifier. *The Imaging Science Journal*, 70:4, 253-272, DOI: [10.1080/13682199.2023.2166657](https://doi.org/10.1080/13682199.2023.2166657)
- Unay, D., Gosselin, B., Kleynen, O., Leemans, V., Destain, M.-F., & Debeir, O. (2011). Automatic grading of bi-colored apples by multispectral machine vision. *Computers and Electronics in Agriculture*, 75(1), 204–212. <https://doi.org/10.1016/j.compag.2010.11.006>
- Vidyarthi, S. K., Singh, S. K., Tiwari, R., Xiao, H.-W., & Rai, R. (2020). Classification of first quality fancy cashew kernels using four deep convolutional neural network models. *Journal of Food Process Engineering*, 43(12), e13552. <https://doi.org/10.1111/jfpe.13552>
- Vidyarthi, S. K., Singh, S. K., Xiao, H.-W., & Tiwari, R. (2021). Deep learnt grading of almond kernels. *Journal of Food Engineering*, 44(4). <https://doi.org/10.1111/jfpe.13662> Process color changes. *Journal of Food Quality*, 2025, Article ID 3350046. <https://doi.org/10.1155/jfq/3350046>
- Golzar, S. H., Bagherpour, H., & Parian, J. A. (2024). A new method to optimize deep CNN model for classification of regular cucumber based on global average pooling. *Journal of Food Processing and Preservation*, 2024, Article ID 5818803. <https://doi.org/10.1155/2024/5818803>
- Jahanbakhshi, A., Momeny, M., Mahmoudi, M., & Zhang, Y.-D. (2020). Classification of sour lemons based on apparent defects using stochastic pooling mechanism in deep convolutional neural networks. *Scientia Horticulturae*, 263, Article 109133. <https://doi.org/10.1016/j.scienta.2019.109133>
- JayaprakashPondy. (2023). Apple Fruit [Dataset]. Kaggle. <https://www.kaggle.com/datasets/jayaprakashpondy/apple-fruit>
- Kayaalp, K. (2024). A deep ensemble learning method for cherry classification. *European Food Research and Technology*, 250(7), 1513–1528. <https://doi.org/10.1007/s00217-024-04490-3>
- Moallem, P., Serajoddin, A., & Pourghassem, H. (2017). Computer vision-based apple grading for Golden Delicious apples based on surface features. *Information Processing in Agriculture*, 4(1), 33–40. <https://doi.org/10.1016/j.inpa.2016.10.003>
- Sabzi, S., Abbaspour-Gilandeh, Y., Javadikia, H., & Havaskhan, H. (2015). Automatic grading of emperor apples based on image processing and ANFIS [Görüntü işleme ve ANFIS ile emperor elmasının