

Research Paper

Development of a deep learning model based on VGG16 network for classifying saffron plants and weeds by color images

Seyed Iman Saedi ¹, Hassan Makarian ^{2*}

¹Department of Water and Soil, Faculty of Agricultural Engineering, Shahrood University of Technology, Iran

²Department of Agronomy and Plant Breeding, College of Agriculture, Shahrood University of Technology, Iran

Article history

Submitted: 2024/12/02

Revised: 2024/12/22

Accepted: 2025/02/24

Published
online: 2025/06/21

Keywords:

Color images

Saffron

Classification Weeding

Deep Learning

*Corresponding author
email:

isaedi@shahroodut.ac.ir

ORCID: 

0000-0003-2020-4142



Abstract

The use of technologies based on site-specific management that enable effective weed control while causing minimal damage to the main crop is of great importance. In this study, color images and a deep learning strategy based on convolutional neural networks were employed to distinguish saffron from two of its common weeds, namely Azmak and Khakshir. A total of 291 images from the three classes (saffron, Azmak, and Khakshir) were captured under completely natural field conditions. The images were resized to 150×150 pixels, and data augmentation was then applied to virtually increase the number of images. The proposed deep learning architecture consisted of an improved VGG16 model, with initial weights applied using the transfer learning approach. Based on the obtained results, the accuracy of the proposed model on the test data was 91%, with an error rate of 0.3759. For the three target classes—saffron, Azmak, and Khakshir—the F1-scores were 85%, 100%, and 86%, respectively. The results also showed that the model achieved a precision of 100% in distinguishing saffron and Azmak from the other classes. This study can be considered a foundation for the development of robotic systems aimed at site-specific weed management in saffron fields.

How to cite this paper:

Saedi, S.I., Makarian, H. (2025). Development of a deep learning model based on VGG16 network for classifying saffron plants and weeds by color images. *Journal of Researches in Mechanics of Agricultural Machinery*. 35: 69-82. <https://dx.doi.org/10.22034/jrmam.2025.15034.738>. (In Persian)



Authors retain the copyright and full publishing rights. Published by [Shahrood University](https://www.shahroodut.ac.ir). This article is an open access article licensed under the [Creative Commons Attribution 4.0 International \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

<https://dx.doi.org/10.22034/jrmam.2025.15034.738>

EXTENDED ABSTRACT

Introduction

Precision agriculture technologies play a crucial role in optimizing crop management by enabling site-specific interventions, particularly in weed control. Traditional weed management methods often result in excessive herbicide use, environmental damage, and crop damage. Automated weed detection using deep learning offers a promising solution, accurately distinguishing between crops and weeds, thereby facilitating targeted removal. Saffron, a high-value crop, faces competition from invasive weeds such as Flixweed and Hoary Cress, which reduce yield and quality. This study leverages computer vision and deep learning to classify saffron and these two common weeds under natural field conditions. Convolutional Neural Networks (CNNs), were employed due to their proven effectiveness in image classification tasks. Transfer learning was applied to enhance model performance by utilizing pre-trained weights from ImageNet. The research aims to develop a robust classification model that can support precision agriculture tools, such as robotic weeders, by accurately identifying weeds while preserving the main crop. The success of this approach could significantly reduce herbicide use, lower production costs, and improve saffron yield through automated, site-specific weed management.

Material and Methods

A dataset of 291 field images of saffron, Flixweed, and Hoary Cress was collected under natural lighting and environmental conditions. Each image was resized to 150×150 pixels, and data augmentation techniques (e.g., rotation, flipping, scaling) were applied to expand the dataset and improve model generalization artificially. The study utilized the VGG16 CNN architecture, fine-tuned via transfer learning with ImageNet weights. The model was trained to classify the three plant categories, and its performance was evaluated using test data. Key metrics included accuracy, F1-

score, and precision to assess classification effectiveness.

Results and Discussion

The proposed model achieved an overall accuracy of 91% on unseen test data, with a loss of 0.3759. The F1-scores for Saffron, Hoary Cress, and Flixweed were 85%, 100%, and 86%, respectively. Notably, the model demonstrated perfect precision (100%) in distinguishing saffron from Hoary Cress, indicating no false positives for these classes. Flixweed recognition was slightly less precise but still highly effective (86% F1-score). The high classification accuracy suggests that deep learning, combined with transfer learning, is a viable approach for weed detection in precision agriculture. The model's ability to differentiate saffron from invasive weeds under real-world conditions supports its potential integration into automated weeding systems. These results suggest that robotic weeders equipped with such AI models can selectively target weeds while minimizing crop damage, thereby reducing reliance on broad-spectrum herbicides.

Conclusions

This study demonstrates the effectiveness of deep learning in distinguishing saffron from Flixweed and Hoary Cress under natural field conditions. The improved VGG16 model achieved high accuracy (91%) and near-perfect precision for certain weed classes, validating its potential for real-world agricultural applications. The findings provide a foundation for developing AI-driven weed removal robots, which could enhance precision farming by enabling targeted, sustainable weed management. Future research should focus on optimizing models for real-time processing and integration with robotic systems. Expanding the dataset to include more weed species and varying environmental conditions could further improve robustness. Overall, this work contributes to advancing precision agriculture technologies, offering a scalable

solution for automated weed control in saffron fields and similar high-value crops.

Author Contributions

S.I. Saedi: Software, Methodology, Modeling, Writing of the initial draft, Final review, and editing.

H.Makarian: Review and editing of the text, Research, Data collection, Validation.

Data Availability Statement

Not applicable.

Acknowledgements

Not applicable.

Ethical Considerations

This section states ethical approval details (e.g., Ethics Committee, ethical code) and confirms adherence to ethical standards, including avoidance of data fabrication, falsification, plagiarism, and misconduct.

Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper

Funding Statement

The author(s) received no specific funding for this research



توسعه مدل یادگیری عمیق مبتنی بر شبکه VGG16 به منظور طبقه‌بندی گیاه زعفران و علف‌های هرز از روی تصاویر رنگی

سید ایمان ساعدی^{۱*}، حسن مکاریان^۲

۱- گروه آب و خاک، دانشکده کشاورزی، دانشگاه صنعتی شاهرود، ایران

۲- گروه زراعت و اصلاح نباتات، دانشکده کشاورزی، دانشگاه صنعتی شاهرود، ایران

تاریخچه مقاله	چکیده
دریافت: ۱۴۰۳/۰۹/۱۲	استفاده از فناوری‌های مبتنی بر مدیریت خاص مکانی که منجر به مبارزه مؤثر با علف
بازنگری: ۱۴۰۳/۱۰/۰۲	هرز گردد و کمترین آسیب را به محصول اصلی وارد آورد، بسیار حائز اهمیت است. در
پذیرش: ۱۴۰۳/۱۲/۰۶	این پژوهش، به منظور تمایز زعفران از دو علف هرز رایج آن یعنی ازمک و خاکشیر از
انتشار: ۱۴۰۴/۰۳/۳۱	تصاویر رنگی و راهبرد یادگیری عمیق مبتنی بر شبکه‌های عصبی پیچشی استفاده
واژه‌های کلیدی:	شد. ۲۹۱ عکس از سه طبقه مذکور (زعفران، ازمک و خاکشیر) در شرایط کاملاً
تصاویر رنگی	طبیعی مزرعه گرفته شد. عکس‌ها به اندازه ۱۵۰ در ۱۵۰ تصادله تغییر ابعاد داده
زعفران	شدند و سپس، از داده‌افزایی برای افزایش مجازی تعداد تصاویر استفاده شد. ساختار
طبقه‌بندی	پیشنهادی برای مدل یادگیری عمیق مشتمل بر مدل بهبودیافته VGG16 و اعمال
علف هرز	وزن‌های اولیه با استفاده از روش انتقال یادگیری بود. با توجه به نتایج به‌دست‌آمده،
یادگیری عمیق	دقت مدل پیشنهادی در داده‌های آزمون ۹۱٪ و میزان خطا ۳۷۵۹/۰ بود. در سه طبقه
*پست الکترونیکی نویسنده	مدنظر، زعفران، ازمک و خاکشیر به ترتیب با امتیاز $F1$ معادل ۸۵٪، ۱۰۰٪ و ۸۶٪
مسئول:	طبقه‌بندی شدند. نتایج همچنین نشان داد مدل در تمایز زعفران و ازمک از دو طبقه
isaedi@shahroodut.ac.ir	دیگر دارای دقت (Precision) معادل با ۱۰۰٪ بود. این پژوهش را می‌توان به عنوان
ORCID: 	بستری در توسعه ربات‌ها به منظور مدیریت خاص مکانی علف هرز در مزارع زعفران
۰۰۰۰-۰۰۰۳-۲۰۲۰-۴۱۴۲	مدنظر قرار داد.



نحوه استناد به این مقاله:

ساعدی، س.ا.، مکاریان، ح. (۱۴۰۴). توسعه مدل یادگیری عمیق مبتنی بر شبکه VGG16 به منظور طبقه‌بندی گیاه زعفران و علف‌های هرز از روی تصاویر رنگی. نشریه پژوهش‌های مکانیک ماشین‌های کشاورزی، ۳۵: ۸۲-۶۹. شناسه دیجیتال:

<https://dx.doi.org/10.22034/jrmam.2025.15034.738>

مقدمه

نیاز روزافزون جهان به محصولات کشاورزی ضرورت استفاده از رویکردهای نوآورانه برای تولید را نمایان می‌کند. در این راستا، اتخاذ راهبردهایی که استفاده حداقلی نهاده‌ها و تولید حداکثر عملکرد را به دنبال داشته باشد امری حیاتی است (Gebbers and Adamchuk, 2010; Saedi *et al.* 2019). زعفران به عنوان یک محصول صادراتی مهم در اقتصاد ایران و جهان به شمار می‌آید. کشور ایران یکی از مهم‌ترین کشورهای تولید کننده این محصول است. مبارزه مناسب با علف‌های هرز این محصول می‌تواند نقش کلیدی در توسعه کمی و کیفی آن ایفا نماید. در این راستا، استفاده از فناوری‌های مبتنی بر کشاورزی دقیق که منجر به مبارزه انتخابی با علف هرز گردد و کم‌ترین آسیب را به محصول اصلی وارد آورد حائز اهمیت است.

رویکردهای مختلفی برای شناسایی و طبقه‌بندی محصول و علف هرز وجود دارد. این رویکردها بیش‌تر به ویژگی‌های بصری بستگی دارند که با استفاده از روش‌های پردازش تصویر، مسئول تشخیص محصول اصلی از علف هرز هستند. این روش‌ها اساساً چالش برانگیزند، زیرا اثرات محیطی مانند پس‌زمینه پیچیده، نور متغیر (Feng *et al.* 2019)، همپوشانی‌ها و انسدادها با سایر بخش‌های گیاه (Wang *et al.* 2017) تشخیص دقیق را دشوارتر می‌کند. چالش‌های دیگر در مورد ویژگی تعمیم و هم‌چنین زمان اجرای یک روش است، به این معنی که الگوریتم پیشنهادی باید بتواند انواع مختلف محصولات را در زمان واقعی برای کاربردهای غیر ایستا (مثل ربات مبارزه با علف هرز) شناسایی نماید. هم‌چنین، به منظور مبارزه رباتی با علف هرز، می‌توان به راهکارهایی برای تشخیص زمان مناسب اعمال تیمار مبارزه توجه نمود. محققان موفق به ارائه الگوریتم‌های مناسب تشخیص محصول برای انواع کاربردها در کشاورزی دقیق شده‌اند که هرکدام دارای دقت متفاوتی هستند. تصاویر RGB یا رنگی، که اطلاعات دوبعدی را ارائه می‌کنند، در این زمینه بیش‌تر مورد استفاده قرار گرفته‌اند (Gené-Mola *et al.* 2019; Juwono *et al.*, 2023).

استفاده از یادگیری عمیق برای نیل به اهداف فوق می‌تواند نتایج امیدبخشی به همراه داشته باشد. یادگیری عمیق یک زمینه یادگیری ماشین خاص بر اساس نمایش پی در پی لایه‌های داده‌است. این رویکردها که معمولاً مبتنی بر شبکه‌های عصبی پیچشی هستند دارای خاصیت تعمیم قوی

و هم‌چنین قابلیت بالا برای استخراج غیرخطی ویژگی‌های تصویر برای طبقه‌بندی اند (Nasiri *et al.* 2019; Peng *et al.* 2018). شبکه‌های عصبی پیچشی^۱ زیرمجموعه‌ای از یادگیری عمیق هستند که در زمره شبکه‌های عصبی با نظارت طبقه‌بندی می‌شوند (Chollet, 2017). به طور کلی، یک شبکه عصبی پیچشی عمیق که دارای دو مرحله استخراج ویژگی و طبقه‌بندی است مجموعه‌ای از لایه‌های مختلف است که مهم‌ترین آنها لایه‌های پیچش و Max Pooling هستند. در کنار آنها لایه‌هایی از قبیل Batch, Global Average Pooling, Normalization, Dropout و Dense نیز ممکن است استفاده شود. شبکه ابتدا عکس‌هایی با اندازه‌های مختلف را دریافت می‌کند. این عکس با انجام پیش‌پردازش‌هایی که بعضاً مورد نیاز است وارد مرحله بعدی شبکه یعنی بخش استخراج ویژگی یا یادگیری ویژگی می‌شود. در این قسمت با استفاده از لایه‌های مختلف عمل استخراج ویژگی‌ها صورت می‌گیرد. با افزایش لایه‌ها در این بخش، ویژگی‌های استخراج شده به تدریج پیچیده‌تر می‌شوند. بعد از انجام فرایند استخراج ویژگی فرایند طبقه‌بندی صورت می‌گیرد. نحوه قرارگیری لایه‌های مختلف (در کنار ویژگی هر یک) موجب تولید شبکه‌های مختلف با ویژگی‌ها و توانایی‌های مختلف می‌گردد. اما با این حال، شبکه‌های استاندارد وجود دارد که توسط شرکت‌ها و واحدهای تحقیقاتی مختلف تولید شده‌اند و عملکرد خود را به اثبات رسانیده‌اند.

در سال‌های اخیر کاربرد یادگیری عمیق در زمینه‌های مختلف کشاورزی شتاب گرفته است (Kamilaris and Prenafeta-Boldú 2018). به عنوان مثال می‌توان به طبقه‌بندی گونه‌های گیاهی (Dyrmann *et al.*, 2016)، تشخیص علف‌های هرز (Yu *et al.*, 2019)، تشخیص میوه روی شاخه (Saedi & Khosravi, 2020)، تعیین زمان برداشت میوه سیب (Saedi, 2023)، و پیش‌بینی عملکرد محصول (Nevavuori *et al.*, 2019) اشاره کرد. در یک مطالعه، یک رویکرد شبکه عصبی عمیق پیشرفته به نام R-CNN Fast توسط محققان برای توسعه یک سامانه کارآمد تشخیص میوه استفاده شد (Ren *et al.* 2015). هم‌چنین، با استفاده از روش انتقال یادگیری^۲، محققین یک الگوریتم تشخیص میوه را توسعه دادند که در آن دو نوع تصویر RGB و NIR با روش هم‌جوشی اولیه و سریع ترکیب شدند (Sa *et al.* 2016). در پژوهش دیگری محققین از سه روش شبکه عصبی پیچشی به

(GWN) در مطالعه‌ای مشابه برای شناسایی علف‌های هرز در تصاویر RGB از مراجع استفاده شد (Hu et al., 2020). در نهایت، از مدل یادگیری عمیق برای طبقه‌بندی گونه‌های علف هرز و محصول در شرایط گلخانه توسط محققین استفاده شد. آنها از الگوریتم انتخاب ویژگی ReliefF برای انتخاب مهم‌ترین ویژگی‌ها برای مدل‌های پیش‌بینی استفاده کردند. سپس تصاویر پردازش شده به مدل یادگیری عمیق که شبکه VGG16 بود وارد شد. میانگین امتیازات F1 طبقه‌بندی‌کننده مدل VGG16 بین ۹۳٪ و ۹۷٪ به دست آمد (G C et al., 2022).

مبارزه مؤثر با علف‌های هرز زعفران نقش مهمی در توسعه کمی و کیفی این محصول به همراه داشت. از این رو، در پژوهش حاضر با هدف ایجاد بستری در توسعه سامانه‌های خودکار کنترل علف‌های هرز این محصول، استفاده از فناوری پیشرفته یادگیری عمیق مدنظر قرار گرفت. در اینجا، به منظور تشخیص و طبقه‌بندی گیاه زعفران و تمایز آن از دو علف هرز متداول یعنی خاکشیر و از مک، یک الگوریتم یادگیری عمیق معرفی گردید که بتواند با پردازش تصاویر رنگی گرفته‌شده، هر یک از سه طبقه مدنظر (زعفران، خاکشیر، و از مک) را شناسایی کرده و زمینه‌ای برای مدیریت دقیق علف‌های هرز ایجاد نماید.

مواد و روش‌ها

آماده سازی و پیش پردازش داده ها

مهمترین دلیل استفاده از راهبرد یادگیری عمیق در این مطالعه توانایی بسیار بالای آن (نسبت به روش‌های سنتی پردازش تصویر) در پردازش عکس‌های رنگی گرفته شده در شرایط کاملاً طبیعی و بدون ساختارهای خاص عکس‌برداری یا تنظیمات خاص دوربین است.

به منظور توسعه مدل شبکه عصبی عمیق در این پژوهش، ابتدا عکس‌هایی از سه طبقه مدنظر یعنی زعفران، از مک، و خاکشیر در طول دوره رشد گرفته شد. در شکل ۱ نمونه‌ای از این عکس‌ها نمایش داده شده است. عکس‌های تهیه شده مورد بازبینی قرار گرفت و نمونه‌های نامناسب (مثلاً مواردی که تصویر شفافیت لازم را به همراه نداشت) حذف شدند. در نهایت تعداد عکس‌های مورد استفاده در توسعه مدل برای هر کدام از طبقه‌های زعفران، خاکشیر و از مک به ترتیب ۶۹، ۱۱۷ و ۱۰۵ بود. لذا، در مجموع ۲۹۱ عکس از طبقه‌های مختلف تهیه شد که از آنها برای توسعه مدل و انجام فرایند آموزش شبکه عصبی عمیق استفاده شد.

نام‌های ResNet50 و ResNet152 برای تشخیص مرکبات استفاده کردند تا در یک ربات برداشت میوه آزمایش شوند. برگ‌ها، شاخه‌ها و میوه‌ها نیز با سامانه بینایی پیشنهادی متمایز شدند و بالاترین دقت تشخیص توسط ResNet152 به دست آمد (Liu et al. 2019). در یک مطالعه کاربردی دیگر، شبکه استاندارد Xception به منظور شناسایی میوه استفاده شد (Salim et al., 2023). هم‌چنین، در پژوهشی یک الگوریتم یادگیری عمیق SSD با استفاده از مدل ResNet101 برای تشخیص سیب و چند نوع میوه دیگر توسعه داده شد. در این مطالعه، یک الگوریتم بهینه سازی SGD و یک روش افزایش داده نیز برای بهبود دقت تشخیص میوه و تعمیم مدل استفاده شد (Peng et al. 2018). یکی دیگر از رویکردهای یادگیری عمیق در مورد ترکیب اطلاعات RGB-D و موج سنجی برای تشخیص سیب فوجی بود. شدت سیگنال تصحیح دامنه و ثبت تصویر در کنار مدل R-CNN Faster به عنوان فرآیند تشخیص لحاظ گردید (Gené-Mola et al. 2019). یک مدل VGG16 در یک مطالعه توسعه داده شد تا میوه‌های خرما را بر اساس بلوغ و وضعیت سلامت آنها مرتب کند که به دقت ۹۶/۹۸٪ در طبقه‌بندی دست‌یافت (Nasiri et al. 2019). در تحقیقی مشابه، یک مدل Inception-ResNetV2 برای طبقه‌بندی هفت نوع میوه زیتون، با دستیابی به حداکثر دقت ۹۵/۹۱٪ استفاده شد (Ponce et al., 2019). در پژوهشی دیگر، یک رویکرد شبکه عصبی عمیق برای طبقه‌بندی دو رقم میوه زیتون روی درخت در مراحل رشدی مختلف استفاده گردید. مدل طراحی شده توانست با دقت بیش از ۹۱٪ هر یک از این مراحل را تشخیص دهد (Khosravi et al. 2021). این روش هم‌چنین در جداسازی پس از برداشت میوه زیتون براساس رقم و میزان بلوغ استفاده گردید (Saedi et al. 2024).

محققین توانستند از یک شبکه عصبی پیش عمیق مبتنی بر معماری YOLOv3 برای تشخیص و شناسایی علف هرز *Convolvulus sepium* در مزرعه چغندر قند استفاده کنند. آنها در تشخیص علف هرز و چغندر قند به ترتیب به دقت ۷۶٪ و ۸۲٪ دست یافتند (Gao et al., 2020). در پژوهشی مشابه، با استفاده از تصاویر گرفته شده توسط هواپیماهای بدون سرنشین، پژوهشگران تصادف‌ها را به عنوان علف‌های هرز یا محصولات زراعی طبقه‌بندی کردند. آنها مدلی از CNN ارائه کردند که می‌توانست با موفقیت بین علف‌های هرز و محصولات تمایز قائل شود (Bah et al., 2018). از معماری یادگیری عمیق مبتنی بر نمودار به نام Graph Weeds Net



شکل ۱ - تصاویری از طبقه‌های مختلف مورد مطالعه

را به‌گونه‌ای به کار می‌گیرد که تصاویر تولید شده همچنان در ظاهر باور پذیر باشند. این امر تعمیم مدل را بهبود می‌بخشد، زیرا همان تصویر بیش از یک بار در معرض دید قرار نمی‌گیرد (Chollet, 2017). در این مطالعه، از یک روش داده‌افزایی خودکار شامل چرخش عکس، جابجایی عرضی، جابجایی ارتفاع، آیینه کردن افقی و آیینه کردن عمودی استفاده شد. پس از بازبینی تصاویر تولید شده و حذف موارد نامعتبر، از تصاویر حاصل شده برای آموزش مدل استفاده شد. ویژگی‌های داده‌افزایی در جدول ۱ نمایش داده شده‌اند. فرآیند بعدی پیش‌پردازش ایجاد اندازه یکسان برای تمامی تصاویر بود. در این راستا تمام تصاویر به اندازه ۱۵۰ در ۱۵۰ تصادف تغییر داده شدند که از آنها برای توسعه مدل و آموزش شبکه استفاده شد.

به عنوان یک عملیات مهم، بایستی تصاویر ورودی پیش‌پردازش شوند. اجرای درست این فرایند به بهبود دقت مدل کمک قابل توجهی خواهد کرد، از بیش‌برازش^۱ جلوگیری کرده و قابلیت تعمیم بالاتری نیز به مدل خواهد داد. یکی از اشکالات عمده شبکه‌های عصبی عمیق این است که برای ایجاد دقت خوب به مقدار زیادی از داده‌های برچسب‌گذاری شده وابسته هستند. این مجموعه داده‌های بزرگ، مرحله آموزش را برای یادگیری تمام ویژگی‌های استفاده شده و همچنین به حداقل رساندن خطر بیش‌آموزش شبکه تسهیل می‌کند. تهیه این تعداد زیاد عکس بسیار پر زحمت، پرهزینه و وقت‌گیر است. داده‌افزایی که در بینایی کامپیوتر مدنظر قرار می‌گیرد یک پاسخ قدرتمند شناخته شده برای کاهش مؤثر بیش‌برازش با ایجاد داده‌های آموزشی بیش‌تر از نمونه‌های فعلی است. این فرآیند برخی از تبدیل‌های تصادفی

جدول ۱ - ویژگی‌های داده‌افزایی به منظور افزایش مجازی داده‌ها در آموزش مدل شبکه عمیق.

نام ویژگی	چرخش تصادفی	آیینه شدن عمودی	آیینه شدن افقی	تغییر ارتفاع	جابجایی عرضی	حالت پرشدگی
مقادیر لحاظ شده	0.2	True	True	0.2	0.2	nearest

داده‌های آزمون در حین آموزش استفاده نمی‌شوند، بلکه از آنها برای بررسی عملکرد نهایی مدل و آزمودن دقت مدل بر روی داده‌های دیده نشده استفاده می‌گردد.

ساختار شبکه پیشنهادی

مبنای توسعه مدل در این پژوهش، یک شبکه استاندارد تحت عنوان VGG16 بود که اصلاحاتی در آن به منظور افزایش دقت و بهبود عملکرد صورت گرفت. VGG16 شامل ۱۳ لایه پیش‌پس هسته ۳×۳ و پنج لایه Max-Pooling با هسته‌های ۲×۲ است. پس از هر لایه پیش‌پس، تابع فعال‌سازی استفاده می‌شود. تابع فعال‌سازی

به منظور توسعه مدل شبکه عصبی عمیق داده‌های تصویری موجود به سه بخش آموزش، اعتبارسنجی و آزمون تبدیل شدند. به این منظور ۸۰ درصد داده‌ها به عنوان آموزش و ۲۰ درصد باقی مانده به عنوان داده‌های آزمون انتخاب شدند. از بین داده‌های آموزش ۲۰ درصد آن‌ها به عنوان داده‌های اعتبارسنجی انتخاب شدند. داده‌های اعتبارسنجی مجموعه‌ای از داده‌های جدا از مجموعه آموزشی است که برای اعتبارسنجی عملکرد مدل در طول آموزش استفاده می‌شود. این فرآیند اعتبارسنجی اطلاعاتی را به ما می‌دهد که کمک می‌کند ابرویژگی‌ها و پیکربندی‌های مدل را مطابق با آن تنظیم کنیم. ذکر این نکته حائز اهمیت است که

¹ Overfitting

Propagation یا پس انتشار. در مرحله Feed-Forward، خطای شبکه با مقایسه تفاوت بین خروجی شبکه بر اساس تصویر داده شده با شبکه به عنوان ورودی، و خروجی برچسب‌گذاری شده واقعی محاسبه می‌شود. در مرحله پس انتشار، گرادیان ویژگی‌ها بسته به خطای شبکه محاسبه می‌شود و سپس هر ماتریس وزنی بر اساس گرادیان محاسبه شده به روز می‌شود، یعنی ورودی وارد شبکه می‌شود و از لایه‌های مختلف عبور می‌کند و نتیجه حاصل می‌شود. این نتیجه با نتیجه مطلوب مقایسه می‌شود و اختلاف این دو خطا را به ما می‌دهد. با استفاده از این خطا ویژگی‌های شبکه اصلاح شده و مجدداً داده به شبکه وارد می‌شود و نتایج و خطای جدید محاسبه شده و سپس ویژگی‌ها اصلاح می‌شود.

این چرخه به دفعات تکرار می‌شود. هر دفعه یک تکرار نامیده می‌شود. برای محاسبه خطای شبکه روابط و توابع مختلفی وجود دارد. پس از آنکه خطا محاسبه شد بایستی توسط این خطا ویژگی‌ها را اصلاح نمود. به این فرایند بهینه‌سازی گفته می‌شود. این فرایند حرکت قدم به قدم به سمت کم‌ترین خطا است، یعنی اصلاح و به روز رسانی وزن‌ها برای رسیدن به حداقل خطا. به منظور آموزش مدل توسعه یافته، از تابع خطای categorical crossentropy، و به منظور بهینه‌سازی ویژگی‌ها از تابع Nadam استفاده گردید. ویژگی‌های این بهینه‌ساز در جدول ۳ معرفی شده‌است.

با انتخاب اندازه بسته (Batch Size) برابر با ۳۲ و نیز انتخاب ۱۵۰ تکرار، فرایند آموزش مدل صورت پذیرفت. برای توسعه مدل از محیط ابری گوگل تحت عنوان Google Colaboratory استفاده شد. همچنین، زبان برنامه نویسی Python 3.10.12 و کتابخانه‌های keras و tensorflow و سایر کتابخانه‌های مفید دیگر استفاده شد.

جدول ۳ - ویژگی‌های بهینه ساز Nadam

learning_rate	beta_1	beta_2	نام ویژگی
0.001	0.9	0.999	مقدار

یکی از اجزای مهم شبکه عصبی است که خروجی هر نورون را پایش می‌کند. توابع فعال‌سازی معمولاً غیرخطی هستند و به شبکه عصبی خاصیت غیرخطی می‌دهند. در یک شبکه عصبی در حالت کلی، در یک نورون، ورودی‌ها در وزن‌های متناظرشان ضرب شده و سپس با یکدیگر جمع می‌شوند. نتیجه از تابع فعال‌سازی عبور می‌کند. در این مدل، تابع فعال‌سازی برای هر لایه پیچش تابع ReLU است که عملکرد آن در رابطه ۱ بیان شده‌است. این تابع در حال حاضر تقریباً در تمام شبکه‌های عصبی عمیق استفاده می‌شود.

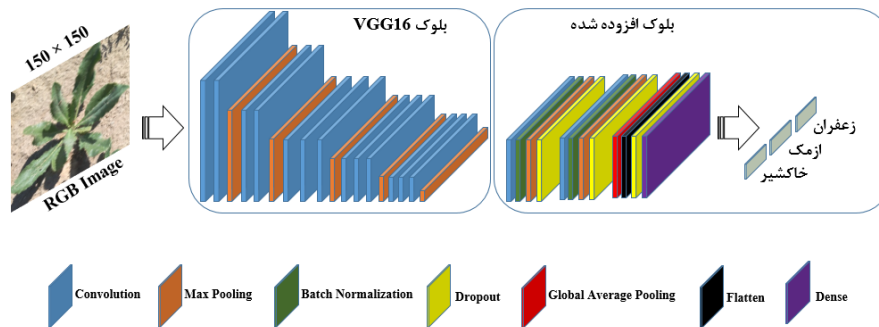
$$f(x) = \begin{cases} 0, & x \leq 0 \\ x, & x > 0 \end{cases} \quad (1)$$

ساختار مدل VGG16 در شکل ۲ نمایش داده شده است. این ساختار یک معماری از پیش آموزش دیده شده‌است که این آموزش از طریق داده‌های ImageNet صورت گرفته است. وزن‌های از پیش آموزش داده شده در مجموعه داده ImageNet در مدل بارگذاری شده‌اند. این وزن‌ها نقطه شروع خوبی برای ویژگی‌های اولیه مدل هستند. لایه‌های کاملاً متصل (لایه‌های انتهایی) مدل VGG16 در مدل اولیه لحاظ نشدند تا بتوان لایه‌های مدنظر را در انتهای مدل پایه برای کار خاص خود سفارشی کرده و اضافه نمود. اجازه داده شد تا وزن این لایه‌ها در طول فرایند آموزش به روز شود تا مدل از پیش آموزش دیده به منظور بهبود عملکرد در مجموعه داده خاص این پژوهش تنظیم شود. با افزودن لایه‌های دیگر به ساختار اصلی VGG16 در نهایت، معماری نشان داده شده در شکل ۲ به دست آمد که به عنوان ساختار پیشنهادی این پژوهش ارائه می‌گردد.

جزئیات مربوط به مدل پیشنهادی و ویژگی‌های دخیل در هر لایه مدل، در جدول ۲ نشان داده شده است. همان‌گونه که در این جدول مشاهده می‌شود این مدل دارای بیش از چهارده میلیون ویژگی است که بجز ۱۲۸ ویژگی سایر ویژگی‌ها قابل آموزش هستند.

آموزش مدل

فرایند آموزش همانند تمام شبکه‌های عصبی سنتی است. آموزش CNN ها شامل دو مرحله است: Feed-Forward و Back-



شکل ۲ - ساختار اصلاح شده VGG16 به منظور شناسایی سه طبقه مختلف مورد مطالعه.

جدول ۲ - جزئیات مدل توسعه یافته و شکل خروجی هر لایه.

تعداد ویژگی‌ها	شکل خروجی	نوع لایه
14714688	(None, 4, 4, 512)	vgg16
147488	(None, 4, 4, 32)	Conv2d-1 + ReLU
128	(None, 4, 4, 32)	BatchNorm2d-1
0	(None, 2, 2, 32)	MaxPool2d-1
0	(None, 2, 2, 32)	dropout_2 (Dropout)
9248	(None, 2, 2, 32)	Conv2d-1 + ReLU
128	(None, 2, 2, 32)	BatchNorm2d-1
0	(None, 1, 1, 32)	MaxPool2d-1
0	(None, 1, 1, 32)	dropout_2 (Dropout)
0	(None, 32)	GlobalAvgPool2d (GAP)
0	(None, 32)	flatten_2 (Flatten)
0	(None, 32)	dropout_4 (Dropout)
99	(None, 3)	dense_3 (Dense)
14,871,779		Total parameters
14,871,651		Trainable parameters
128		Non-trainable parameters

سنجش عملکرد طبقه بندی تصاویر

به منظور سنجش عملکرد مدل در طبقه‌بندی تصاویر استفاده از مفاهیم چهارگانه TP, TN, FP, FN لحاظ شده‌است که تعاریف آنها به این صورت بیان می‌گردد:

True Positive (TP) به مواردی اطلاق می‌شود که طبقه‌بندی کننده به درستی طبقه مثبت را پیش‌بینی کند. به عبارت دیگر، یک نمونه مثبت را به درستی شناسایی می‌کند.

True Negative (TN) به مواردی اطلاق می‌شود که طبقه‌بندی کننده به درستی طبقه منفی را پیش‌بینی کند. به عبارت دیگر، یک مصداق منفی را به درستی شناسایی می‌کند.

False Negative (FN) به مواردی اطلاق می‌شود که طبقه‌بندی کننده طبقه منفی را به اشتباه پیش‌بینی می‌کند در حالی که باید طبقه مثبت را پیش‌بینی می‌کرد. به عبارت دیگر، به اشتباه یک نمونه مثبت را منفی تشخیص می‌دهد.

در نهایت، False Positive (FP) به مواردی اطلاق می‌شود که طبقه‌بندی کننده طبقه مثبت را به اشتباه پیش‌بینی

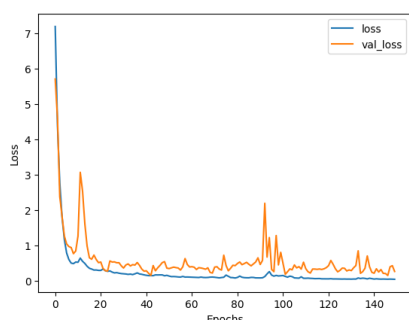
می‌کند در حالی که باید طبقه منفی را پیش‌بینی می‌کرد. به عبارت دیگر، مصداق منفی را به اشتباه مثبت تشخیص می‌دهد. با استفاده از مفاهیم TP, TN, FP, FN، به منظور نمایش نتایج طبقه‌بندی، از گزارش طبقه‌بندی برای هر کدام از طبقه‌های مدنظر (زعفران، ازیمک و خاکشیر) استفاده می‌شود که در بخش مواد و روش‌ها ارائه شده‌است. ویژگی‌های نمایش داده شده در این جدول شامل دقت (Precision)، بازیابی (Recall) و امتیاز F1 (F1-score) هستند که حداقل آنها صفر و حداکثر آنها یک می‌تواند باشد. ویژگی‌های فوق به صورت روابط ۲ تا ۴ تعریف می‌شوند:

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

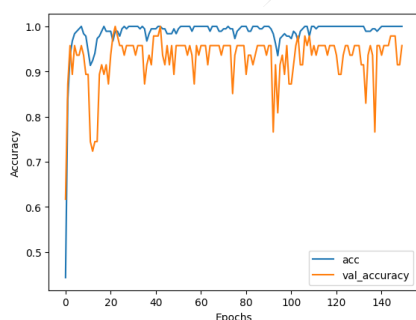
$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F1 - score = \left(2 \times \frac{Precision \times Recall}{Precision + Recall} \right) \quad (4)$$

پیوسته و البته با نوساناتی مشاهده می‌شود. همچنین با توجه به این نمودار کم‌ترین خطا در داده‌های آموزش در تکرار ۱۵۰ و به میزان ۰/۰۴۵۷ و در داده‌های اعتبارسنجی در تکرار ۴۳ و با مقدار ۰/۱۵۶۷ به دست آمد. در شکل ۴ نمودار تغییرات دقت در برابر تکرار برای داده‌های آموزش و اعتبارسنجی آورده شده است. با توجه به این شکل، در مورد داده‌های آموزش بالاترین دقت به میزان ۱۰۰٪ در چندین تکرار به دست آمد و در داده‌های اعتبارسنجی در تکرار ۲۳ بالاترین دقت به میزان ۱۰۰٪ حاصل گردید.



شکل ۳ - نمودار تغییرات خطا در روند آموزش برای داده‌های آموزش و اعتبارسنجی.



شکل ۴ - نمودار تغییرات دقت در روند آموزش برای داده‌های آموزش و اعتبارسنجی.

در جدول ۴ خلاصه‌ای از نتایج حاصل از توسعه مدل ارائه شده است. در این جدول بهترین حالات حاصل شده برای داده‌های آموزش، اعتبارسنجی و آزمون ارائه شده است. بر طبق این جدول، مقدار خطا در داده‌های آزمون برابر با ۰/۳۸ و مقدار دقت برابر با ۰/۹۱ به دست آمد.

گزارش طبقه‌بندی تصاویر زعفران، خاکشیر، و ازمک در جدول ۵ ارائه شده است. با توجه به این جدول دقت شناسایی ازمک و خاکشیر ۱۰۰٪ بود، اما دقت شناسایی زعفران ۷۴٪ به دست آمد. در مورد ازمک، هر سه ویژگی دقت، بازیابی و امتیاز F1 به میزان ۱۰۰٪ به دست آمدند که این نتیجه بیان می‌کند که مدل به بهترین شکل قادر است ازمک را از دو طبقه دیگر

ویژگی دقت نشان می‌دهد چه تعداد از نمونه‌های مثبت پیش‌بینی شده در واقع درست هستند. بالا بودن ویژگی دقت نشان دهنده این است که نسبت تعداد مواردی که مدل اشتباه پیش‌بینی کرده است به تعداد پیش‌بینی‌های درست کمتر است و بالعکس. بازیابی، که همچنین با نام حساسیت یا نرخ درست مثبت نیز شناخته می‌شود، کامل بودن مدل طبقه‌بندی در شناسایی نمونه‌های مثبت را اندازه‌گیری می‌کند. به طور ساده، بازیابی نشان می‌دهد چه تعداد از نمونه‌های واقعی مثبت توسط مدل طبقه‌بندی به درستی شناسایی می‌شوند. این ویژگی هرچقدر بالاتر باشد به معنی این است که نسبت تعداد مواردی که انتظار داشتیم مدل پیش‌بینی کند ولی پیش‌بینی نشده است، به تعداد پیش‌بینی‌های درست کمتر بوده است، و بالعکس. در نهایت، امتیاز F1 میانگین هارمونیک دقت و بازیابی است. این امتیاز با توجه همزمان به دقت و بازیابی، ارزیابی متوازنی از عملکرد مدل طبقه‌بندی را ارائه می‌دهد. امتیاز F1 بین ۰ تا ۱ متغیر است، مقدار ۱ به معنای دقت و بازیابی کامل و مقدار صفر ۰ نشان‌دهنده عملکرد ضعیف است. این ویژگی به عنوان یک معیار تکمیلی، ارزیابی کلی از عملکرد مدل طبقه‌بندی ارائه می‌دهد.

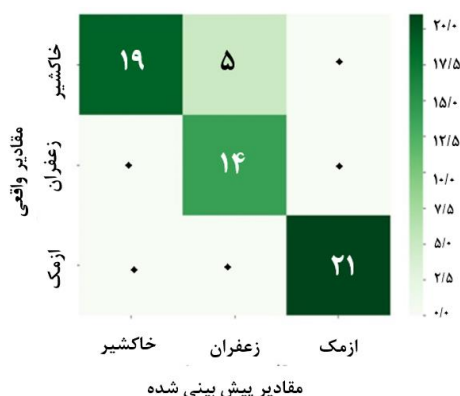
معیار دیگری که با استفاده از آن، عملکرد یک مدل طبقه‌بندی را بررسی می‌کنند، ماتریس آشفتگی نام دارد. این ماتریس پیش‌بینی‌های انجام شده توسط مدل را بر روی مجموعه داده‌های آزمون خلاصه می‌کند و آنها را با برچسب‌های واقعی شناخته شده مقایسه می‌کند. این ماتریس تعداد TP, TN, FP و FN توسط مدل را برای داده‌های آزمون ارائه می‌دهد. ماتریس آشفتگی به ما امکان می‌دهد معیارهای ارزیابی مختلفی مانند دقت، بازیابی، امتیاز F1 و موارد دیگر را ارزیابی کرده و عملکرد مدل را از نظر موارد طبقه‌بندی صحیح و نادرست بررسی نماییم. به این ترتیب، یک نمای کلی روشن از نقاط قوت و ضعف مدل در تمایز بین طبقه‌هایی که برای پیش‌بینی آموزش داده شده است، ارائه می‌شود.

نتایج و بحث

در این بخش نتایج توسعه مدل پیشنهادی برای تمایز بوته زعفران و دو علف هرز آن یعنی خاکشیر و ازمک با استفاده از تصاویر رنگی ارائه می‌شود.

در شکل ۳ نمودار تغییرات خطا در برابر تکرار (تکرار) برای داده‌های آموزش و اعتبارسنجی در روند آموزش ارائه شده است. همانگونه که در این شکل مشاهده می‌شود در هر دو سری داده‌های آموزش و اعتبارسنجی روند کاهش خطا تقریباً به طور

لذا، تکیه بر تصاویر تهیه شده در شرایط آزمایشگاهی که در آنها ویژگی‌های عکس‌برداری ثابت هستند، برای کاربردهای عملی در حذف بلادرنگ علف هرز در مزرعه امری غیرعملی به نظر می‌رسد. از این رو، تحقیقاتی که قابلیت اجرا در شرایط بلادرنگ مزرعه را داشته باشند مطلوب‌تر خواهند بود. البته ذکر این نکته نیز شایان توجه است که یک مدل موفق تمایز علف‌های هرز برای یک محصول، ممکن است برای سایر محصولات عملکرد مطلوبی نداشته باشد. محققین در پژوهش‌های مختلف تلاش کرده‌اند با استفاده از روش‌های یادگیری عمیق پیشرفته با این چالش‌ها مقابله کنند. به عنوان مثال، Shao et al., 2023 با استفاده از روش‌های یادگیری عمیق و تصاویر تهیه شده در شرایط طبیعی مزرعه سعی کردند شش گونه از علف‌های هرز را در شالیزار متمایز کنند. در مقایسه با پژوهش حاضر، این تحقیق بر روی تعداد بیش‌تری از گونه‌های علف‌هرز انجام شد اما دقت کم‌تری در پی داشت.



شکل ۵ - ماتریس آشفتگی به منظور ارزیابی عملکرد مدل طبقه‌بندی پیشنهادی

تمایز دهد و هیچ کدام از دو طبقه زعفران و خاکشیر در عدم تشخیص از مک اشتباه نکردند.

جدول ۴ - نتایج توسعه مدل بر روی داده‌های آموزش، اعتبارسنجی و آزمون.

داده‌های آموزش	داده‌های اعتبارسنجی	داده‌های آزمون
100%	100%	9152%
0.0457	0.1567	0.3759
پایین‌ترین خطا		

زعفران دارای مقدار بازیابی ۱۰۰٪ بوده است، اما این ویژگی برای خاکشیر ۷۵٪ به دست آمد. یعنی، مدل در شناسایی نمونه‌های مثبت خاکشیر بیش از ۷۵٪ دقت نداشته است. مقدار امتیاز F1 برای خاکشیر و زعفران تقریباً مساوی است (۰/۸۶ و ۰/۸۵) که نشان می‌دهد مدل طبقه‌بندی در شناسایی این دو طبقه تقریباً به یک شکل عمل کرده است.

ماتریس آشفتگی برای مدل طبقه‌بندی این پژوهش در شکل ۵ نمایش داده شده است. بر طبق این جدول، پنج نمونه از عکس‌های خاکشیر به اشتباه، زعفران تشخیص داده شد. اما، هیچ کدام از عکس‌های زعفران و از مک با دو طبقه دیگر اشتباه گرفته نشدند و نیز هیچ کدام از عکس‌های دو طبقه دیگر، از مک را به اشتباه در طبقه خود قرار ندادند. به عبارت دیگر، مدل در تشخیص زعفران و از مک دچار هیچ اشتباهی نمی‌شود، اما اگر هدف شناسایی خاکشیر باشد تنها ممکن است آن را با زعفران اشتباه کند و هرگز خاکشیر را با از مک اشتباه نمی‌کند.

توانایی تمایز علف‌های هرز از گیاهان زراعی با استفاده از تصاویر رنگی یک چالش مهم است، به خصوص از این جهت که سامانه‌های مدیریت علف‌های هرز بایستی بتوانند در شرایط طبیعی مزرعه و به‌صورت بلادرنگ علف‌های هرز را شناسایی کنند.

جدول ۵ - گزارش طبقه‌بندی شناسایی زعفران، از مک و خاکشیر از روی تصاویر رنگی.

Support	F1-score	Recall	Precision	
24	0.86	0.75	1.00	خاکشیر
14	0.85	1.00	0.74	زعفران
21	1.00	1.00	1.00	ازمک
59	0.91	0.90	0.91	micro avg.
59	0.90	0.92	0.91	macro avg.
59	0.91	0.90	0.94	weighted avg.
59	0.90	0.90	0.90	samples avg.

گلخانه‌ای با استفاده از یک الگوریتم انتخاب ویژگی پیش‌پردازش کردند و سپس نتایج آن را به منظور انجام عمل

در مطالعه مشابه دیگر G C et al., 2022 تصاویری از چهار گونه علف هرز و شش گونه محصول را در شرایط

حسن مکاریان: بررسی و ویرایش متن، تحقیق، گردآوری داده‌ها، اعتبارسنجی.

اصول اخلاقی

نویسندگان اصول اخلاقی را در انجام و انتشار این اثر علمی رعایت نموده‌اند و این موضوع مورد تایید همه آنها است.

تضاد منافع نویسندگان

در این مقاله هیچ‌گونه تضاد منافی وجود ندارد و این مسئله مورد تایید همه نویسندگان است.

حمایت مالی

نویسنده(ها) هیچ حمایت مالی خاصی برای انجام این پژوهش دریافت نکرده‌اند.

منابع

- Bah, M. D., Hafiane, A., & Canals, R. (2018). Deep Learning with Unsupervised Data Labeling for Weed Detection in Line Crops in UAV Images. In Remote Sensing, 10 (11).. <https://doi.org/10.3390/rs10111690>
- Chollet, F. (2017). Deep learning with Python. Manning Publications Co.
- Dyrmann, M., Karstoft, H., & Midtby, H. S. (2016). Plant species classification using deep convolutional neural network. Biosystems Engineering, 151, 72–80. <https://doi.org/https://doi.org/10.1016/j.biosystemseng.2016.08.024>
- Feng, J., Zeng, L., & He, L. (2019). Apple fruit recognition algorithm based on multi-spectral dynamic image analysis. Sensors (Basel, Switzerland), 19(949). <https://doi.org/10.3390/s19040949>
- G C, S., Zhang, Y., Koparan, C., Ahmed, M. R., Howatt, K., & Sun, X. (2022). Weed and crop species classification using computer vision and deep learning technologies in greenhouse conditions. Journal of Agriculture and Food Research, 9, 100325. <https://doi.org/https://doi.org/10.1016/j.jafr.2022.100325>
- Gao, J., French, A. P., Pound, M. P., He, Y., Pridmore, T. P., & Pieters, J. G. (2020). Deep convolutional neural networks for image-based Convolvulus sepium detection in sugar beet fields. Plant Methods, 16(1), 29. <https://doi.org/10.1186/s13007-020-00570-z>
- Gebbers, R., & Adamchuk, V. I. (2010). Precision agriculture and food security. Science, 327(5968), 828. <https://doi.org/10.1126/science.1183899>

طبقه‌بندی وارد شبکه یادگیری عمیق نمودند. همانند مطالعه این مقاله، شبکه مورد نظر از نوع VGG16 بود، اما به دلیل استفاده از الگوریتم انتخاب ویژگی قبل از ورود داده‌ها به شبکه، دقت بیش‌تری به دست آمد.

دو پیشنهاد به منظور جهت‌دهی به مطالعات آینده مدنظر است: الف) توسعه مدل‌های یادگیری عمیق به منظور شناسایی تمامی گونه‌های ممکن علف هرز در مزارع زعفران به منظور افزایش تعمیم‌پذیری عملکرد مدل، ب) تعمیم مدل‌های یادگیری عمیق برای محصولات مختلف با علف‌هرزهای مشابه، و ب) توسعه مدل‌های یادگیری عمیق جدید برای محصولات مختلف با علف هرزهای متفاوت.

نتیجه‌گیری

در پژوهش حاضر یک مدل شبکه عصبی عمیق مبتنی بر شبکه استاندارد VGG16 که اصلاحاتی بر روی آن انجام شد به منظور تمایز زعفران از دو علف هرز رایج آن، یعنی ازمک و خاکشیر از طریق تصاویر رنگی توسعه داده شد. نتایج حاصل از آزمون مدل با استفاده از داده‌های دیده نشده (داده‌های آزمون) نشان داد که این مدل می‌تواند با دقت مناسبی این سه گیاه را با استفاده از عکس‌برداری متمایز کند. با توجه به نتایج به‌دست‌آمده، از آنجا که هیچ کدام از عکس‌های زعفران با ازمک و خاکشیر اشتباه نشده‌اند، می‌توان سامانه‌ای برای سمپاشی دقیق علف هرز طراحی نمود به گونه‌ای که تنها مبتنی بر شناسایی زعفران باشد. به بیان دیگر، این سامانه قادر خواهد بود سمپاشی مزرعه‌ای را که شامل این سه محصول باشند به شکل پیوسته انجام دهد و هر جا به زعفران برخورد کرد کار سمپاشی را متوقف سازد. نتایج این پژوهش می‌تواند به عنوان بستری در توسعه ربات‌های حذف علف هرز به منظور مبارزه انتخابی با علف هرز را برآورده سازد و به این ترتیب نه تنها صرفه جویی قابل توجهی در علف کش صورت بگیرد بلکه به تولید محصول زعفران با کیفیت تر نیز بیانجامد.

دسترسی به داده‌ها

همه اطلاعات و نتایج در متن مقاله ارائه شده است.

مشارکت نویسندگان

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

سید ایمان ساعدی: نرم‌افزار، مدل سازی، نگارش پیش نویس اولیه، بازبینی و ویرایش نهایی.

- Sa, I., Ge, Z., Dayoub, F., Upcroft, B., Perez, T., & McCool, C. (2016). DeepFruits: A fruit detection system using deep neural networks. *Sensors*, 16(8), E1222. <https://doi.org/10.3390/s16081222>
- Saedi, S. I., (2023). Determining apple fruit harvest time using color images and deep learning. *Iranian Journal of Researches in Mechanics of Agricultural Machinery*. 12.(۳)
- Saedi, S.I., Alimardani, R., Mousazadeh, H., & Salehi, R. (2019). Development and evaluation of an energy and water efficient intensive cropping system. *INMATEH - Agricultural Engineering*, 58(1), 93-104.
- Saedi, S.I., & Khosravi, H. (2020). A deep neural network approach towards real-time on-branch fruit recognition for precision horticulture. *Expert Systems with Applications*, 159, 113594.
- Saedi, S. I., Rezaei, M., & Khosravi, H. (2024). Dual-path lightweight convolutional neural network for automatic sorting of olive fruit based on cultivar and maturity. *Postharvest Biology and Technology*, 216, 113054. <https://doi.org/https://doi.org/10.1016/j.postharvbio.2024.113054>
- Salim, F., Saeed, F., Basurra, S., Qasem, S.N., & Al-Hadhrani, T., (2023). DenseNet-201 and xception pre-trained deep learning models for fruit recognition. *Electronic* 12, 3132. <https://doi.org/10.3390/electronics12143132>.
- Shao, Y., Guan, X., Xuan, G., Gao, F., Feng, W., Gao, G., Wang, Q., Huang, X., Li, J., (2023) . GTCBS-YOLOv5s: A lightweight model for weed species identification in paddy fields, *Computers and Electronics in Agriculture*, 215, 108461. <https://doi.org/10.1016/j.compag.2023.108461>.
- Yu, J., Sharpe, S. M., Schumann, A. W., & Boyd, N. S. (2019). Deep learning for image-based weed detection in turfgrass. *European Journal of Agronomy*, 104, 78–84. <https://doi.org/https://doi.org/10.1016/j.eja.2019.01.004>
- Wang, C., Tang, Y., Zou, X., Luo, L., & Chen, X. (2017). Recognition and matching of clustered mature litchi fruits using binocular charge-coupled device (CCD) color cameras. *Sensors*, 17(5), E1073. <https://doi.org/10.3390/s17051073>
- Gené-Mola, J., Vilaplana, V., Rosell-Polo, J. R., Morros, J.-R., Ruiz-Hidalgo, J., & Gregorio, E. (2019). Multi-modal deep learning for Fuji apple detection using RGB-D cameras and their radiometric capabilities. *Computers and Electronics in Agriculture*, 162, 689-698. <https://doi.org/10.1016/j.compag.2019.05.030>
- Hu, K., Coleman, G., Zeng, S., Wang, Z., & Walsh, M. (2020). Graph weeds net: A graph-based deep learning method for weed recognition. *Computers and Electronics in Agriculture*, 174, 105520. <https://doi.org/https://doi.org/10.1016/j.compag.2020.105520>
- Juwono, F.H., Wong, W.K., Verma, S., Shekhawat, N., Lease, B.A., Apriono, C., (2023). Machine learning for weed-plant discrimination in agriculture 5.0: an in-depth review. *Artif. Intell. Agric.* 10, 13–25. <https://doi.org/10.1016/j.iaia.2023.09.002>.
- Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, 147, 70-90. <https://doi.org/10.1016/j.compag.2018.02.016>
- Khosravi, H., Saedi, S. I., & Rezaei, M. (2021). Real-time recognition of on-branch olive ripening stages by a deep convolutional neural network. *Scientia Horticulturae*, 287, 110252. <https://doi.org/10.1016/j.scienta.2021.110252>
- Liu, Y.-P., Yang, C.-H., Ling, H., Mabu, S., & Kuremoto, T. (2019). A visual system of citrus picking robot using convolutional neural networks. In *Proceedings of the International Conference on Systems Informatics* (pp. 344-349).
- Nasiri, A., Taheri-Garavand, A., & Zhang, Y.-D. (2019). Image-based deep learning automated sorting of date fruit. *Postharvest Biology and Technology*, 153, 133-141. <https://doi.org/10.1016/j.postharvbio.2018.12.001>
- Nevavuori, P., Narra, N., & Lipping, T. (2019). Crop yield prediction with deep convolutional neural networks. *Computers and Electronics in Agriculture*, 163, 104859. <https://doi.org/https://doi.org/10.1016/j.compag.2019.104859>
- Peng, H., Huang, B., Shao, Y., Li, Z., Zhang, C., Chen, Y., & Xiong, J. (2018). General improved SSD model for picking object recognition of multiple fruits in natural environment. In *Transactions of the Chinese Society of Agricultural Engineering*. 155-162.
- Ponce, J. M., Aquino, A., & Andújar, J. M. (2019). Olive-Fruit Variety Classification by Means of Image Processing and Convolutional Neural Networks. *IEEE Access*, 7, 147629–147641. <https://doi.org/10.1109/ACCESS.2019.2947160>
- Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. *arXiv preprint arXiv:1506.01497* .

