

پیاده‌سازی و ارزیابی الگوریتم یادگیری عمیق Detecto در تشخیص میوه سیب روی درخت

ایمان احمدی^{۱*} و فاطمه توسلی^۲

۱- دانشیار، گروه مهندسی مکانیک ماشین‌های کشاورزی، واحد خوراسگان، دانشگاه آزاد اسلامی، اصفهان، ایران

۲- دانش آموخته کارشناسی ارشد ریاضی کاربردی، دانشگاه یزد، یزد، ایران

* ایمیل نویسنده مسئول: imanahmadi2013@gmail.com

(تاریخ دریافت: ۱۴۰۳/۱۱/۱۳ - تاریخ پذیرش: ۱۴۰۳/۱۲/۲۶)

چکیده

یکی از کاربردهای یادگیری عمیق که در سالهای اخیر پیشرفت زیادی در کشاورزی داشته، تشخیص اشیاء با استفاده از روشهای یادگیری عمیق است. هدف از پژوهش حاضر پیاده‌سازی یک مدل یادگیری عمیق برای تشخیص میوه سیب روی درخت به کمک تصاویر گرفته شده از درخت دارای میوه است. ارزیابی عملکرد مدل با محاسبه نسبت مساحت اشتراکی به اجتماعی کادرهای واقعی و پیش‌بینی شده محیط بر سیب‌های موجود در تصویر انجام شد. این کار روی تعدادی از تصاویر نزدیک دارای یک یا دو سیب صورت گرفت و سپس معیارهای حساسیت، دقت و نمره F1 با توجه به نسبت مساحتی محاسبه شده و مقایسه آن با مقدار آستانه تشخیص به‌دست آمدند. به منظور استفاده از کل تصاویر پوشه اعتبارسنجی در ارزیابی عملکرد الگوریتم، از معیار متوسط صحت در تشخیص سیب در تمام تصاویر با لحاظ کردن سطح آستانه تشخیص پنجاه درصدی استفاده شد. بر اساس نتایج به‌دست آمده فرآیند آموزش مدل باعث کاهش مقدار تابع هزینه شد و در شماره دوره‌های آموزشی بیش از ۱۰ نمودار تقریباً افقی شد. همچنین گونه سیب و تغییر پس زمینه عکس‌ها باعث تأثیرگذاری بر نتایج مدل نشد. از نظر کمی معیارهای حساسیت، دقت و نمره F1 مربوط به مدل روی تعدادی از تصاویر نزدیک دارای یک و یا دو سیب بترتیب برابر با ۰/۹۲، ۰/۹۴ و به‌دست آمد. این معیارها در سطح آستانه نسبت مساحت اشتراکی به اجتماعی برابر با ۰/۸ به دست آمدند. از سوی دیگر متوسط صحت به‌دست آمده روی کل تصاویر پوشه اعتبارسنجی با لحاظ کردن سطح آستانه تشخیص پنجاه درصدی برابر با ۰/۸۸ به‌دست آمد. در نتیجه کارایی مدل پیاده‌سازی شده در این پژوهش در تشخیص میوه سیب روی درخت تایید شد.

واژه‌های کلیدی: الگوریتم detecto، تشخیص سیب، نسبت مساحت اشتراکی به اجتماعی، یادگیری عمیق

مقدمه

پیشرفتهای فناوریانه در سالهای اخیر باعث ایجاد ابزارهایی شده است که می‌توانند در زمینه کشاورزی نیز کاربرد داشته باشند. این فناوریها عبارتند از: سنجش از دور ماهواره‌ای (V'elez et al., 2022) و پهپادی (Sishodia et al., 2020; Tsouros et al., 2019)، پردازش داده و دید رایانه‌ای. به کمک این فناوریها، ابزارهای جدیدی در اختیار کشاورزان قرار گرفته که باعث تطبیق کارهای کشاورزی با شرایط و نیازهای موجود در مزرعه به صورت خاص مکانی شده، روشهای پردازش تصویر سنتی بهبود یافته و تسهیل انجام کارهای کشاورزی به صورت دستی را به همراه داشته است. برخی از کاربردهای این فناوریها عبارتند از: تشخیص مناسب آفات و بیماریها که کاربرد دقیق نهادهای شیمیایی را به همراه خواهند داشت (V'elez et al., 2023)، ارزیابی جامع محصولات زراعی (Chang et al., 2020)، ارزیابی تنش آبی گیاهان (Gutiérrez et al., 2018) و ارزیابی پارامترهای بیوفیزیکی گیاهان (Martínez-Peña et al., 2023). همچنین با به خدمت درآمدن اتوماسیون و رباتیک در کارهای کشاورزی، امور نیازمند به نیروی کارگری کاهش یافته است و پیشه کشاورزی کارآمدتر شده است (Marinoudi et al., 2019).

از سوی دیگر امور مهم و متنوعی مانند تشخیص نوع گیاه، تخمین مقدار باردهی، دسته بندی محصول و تشخیص بیماریها با پیاده‌سازی روشهای یادگیری عمیق، پیشرفتهای چشمگیری داشته است (Zhou et al., 2019). یکی از کاربردهای ویژه یادگیری عمیق تشخیص اشیاء موجود در تصاویر است که با استفاده از الگوریتم شبکه عصبی پیچشی قابل پیاده‌سازی

است (Xu et al., 2022). به علت متن باز بودن این الگوریتم، استفاده از آن در تشخیص اشیاء در پروژه‌های مربوط به کشاورزی دقیق رو به افزایش است (Magalhães et al., 2023).

به طور کلی سه نوع هدف گذاری در تشخیص اشیاء در کاربردهای کشاورزی مطرح می‌باشد که عبارتند از: مدیریت رباتها، شمارش و ارتباط دهی ویژگیهای ظاهری به واکنشهای فیزیولوژیکی. تشخیص میوه (Santos et al., 2020) و سبزی (Vitzrabin & Edan, 2016) آماده برای برداشت از وظایف مهمی است که برای خودکار کردن فرآیند برداشت با استفاده از رباتها انجام آن الزامی است. دیگر امور رباتی که به تشخیص اشیاء احتیاج دارند شامل سم‌پاشی گیاه (Berenstein et al., 2010) و تشخیص و مدیریت آفات و بیماریها (Fuentes et al., 2017) می‌شوند. مفهوم شمارش شامل تخمین مقدار محصول (Xiong et al., 2019) و یا تخمین تراکم شکوفه‌های درختان (Farjon et al., 2019) می‌شود که حداقل در مواردی به طور صریح از روشهای تشخیص اشیاء برای انجام آن استفاده می‌گردد. ارتباط دهی ویژگیهای ظاهری به واکنشهای فیزیولوژیکی شامل بررسی صفات و ویژگیهای گیاه است که به کمک آن رشد، مقاومت و شرایط فیزیولوژیکی گیاه ارزیابی شود (Costa et al., 2019). برای دستیابی به عملکرد مناسب در هر یک از وظایف مورد اشاره در بالا وجود یک آشکارساز شیئی با کیفیت ضروری است.

هدف از پژوهش حاضر پیاده‌سازی یک مدل یادگیری عمیق برای تشخیص میوه سیب روی درخت به کمک تصاویر گرفته شده از درخت دارای میوه است. این تصاویر به عنوان ورودی به مدل داده

کرد، یعنی اگر الگوریتمی دارای صحت بالا روی تصاویر اعتبارسنجی باشد، احتیاج به افزودن تعداد تصاویر آموزشی ندارد، چون الگوریتم مورد اشاره با همین تعداد کم تصاویر آموزشی هم کارکرد درستی دارد. همچنین عامل دیگری که بر تعداد تصاویر آموزشی اثرگذار است، تعداد کلاسهای اشیایی است که تشخیص آن به وسیله الگوریتم هدف گذاری شده است. در این پژوهش تشخیص میوه سیب روی درخت هدف گذاری شده است تا در پژوهشهای آتی از نتایج آن برای توسعه ربات برداشت کننده سیب استفاده شود. در نتیجه هدف این پژوهش تشخیص میوه سیب بتهایی است و تشخیص چند دسته میوه مختلف هدف گذاری نشده است تا به تصاویر با تعداد بیشتر در مرحله آموزش مدل احتیاج باشد.

۲- کادربندی و برچسب زنی میوه سیب در تصاویر گروه‌های آموزش و آزمون:

در حال حاضر تعدادی نرم افزار و وب سایت اینترنتی برای کادربندی و برچسب زنی تصاویر مورد استفاده در تشخیص اشیاء با هوش مصنوعی وجود دارد. در این پژوهش از وب سایت makesense.ai برای این کار استفاده شد و تمام تصاویر سیب قابل مشاهده در تک تک عکسهای موجود در پوشه‌های آموزش و آزمون کادر بندی و برچسب زنی شدند. کادردهی به گونه‌ای انجام شد که مرز کادرها مماس بر لبه‌های میوه‌های سیب موجود در تصاویر باشد. تمام فایل‌های خروجی مربوط به این مرحله به پوشه‌های آموزش و آزمون مربوطه منتقل شدند و در نهایت هر سه پوشه آموزش، آزمون و اعتبارسنجی به پوشه دیگری به نام [apple](https://apple.com) منتقل شدند. این پوشه برای استفاده در فرآیند مدل‌سازی به فرمت فشرده

می‌شوند تا به کمک آنها مدل آموزش ببیند و پس از آن قادر به تشخیص میوه سیب در تصاویر دیگری که تا به حال ندیده، باشد. ارزیابی عملکرد مدل هدف دیگری است که در این پژوهش دنبال خواهد شد. کاربرد آتی این آشکارساز می‌تواند در طراحی ربات برداشت کننده سیب باشد.

مواد و روش‌ها

به منظور توسعه مدل تشخیص دهنده میوه سیب روی درخت و ارزیابی کارکرد آن مراحل زیر اجرا شد (Padilla et al., 2020):

۱- جمع آوری تصاویر خام برای آموزش، آزمون و اعتبارسنجی مدل:

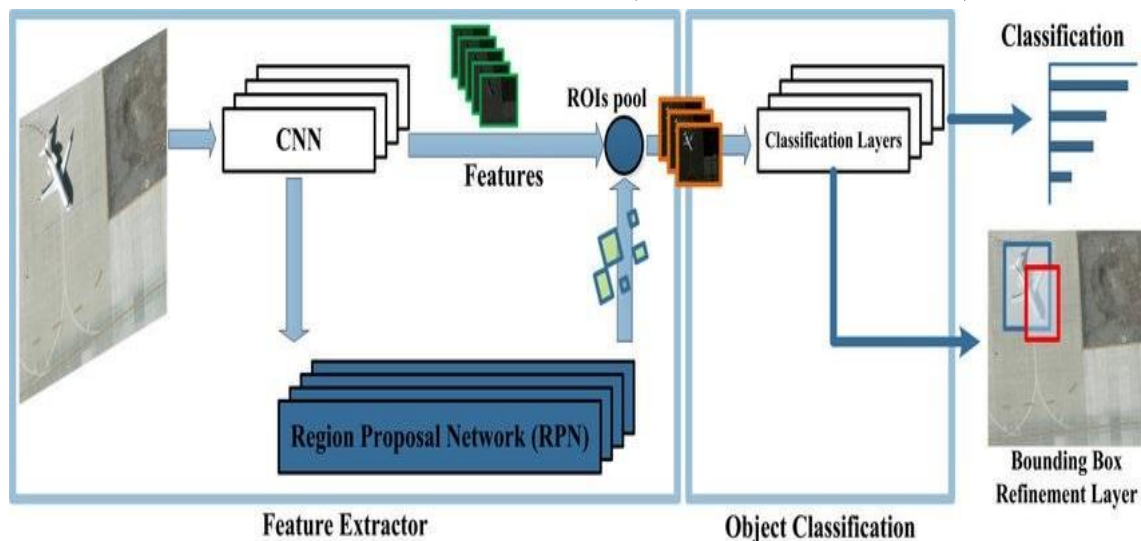
تعداد ۱۹۹ عکس دارای پسوند jpg. مربوط به درخت سیب دارای محصول که شامل گونه‌های مختلف سیب بودند به عنوان داده‌های خام در نظر گرفته شد. این تصاویر به سه گروه آموزش، آزمون و اعتبارسنجی تقسیم شدند که مشتمل بر بترتیب ۲۲، ۱۱ و ۱۶۶ عکس بودند. از تصاویر گروه آموزش برای آموزش مدل استفاده شد. کیفیت مدل توسعه یافته توسط تصاویر گروه آزمون و با محاسبه مقدار تابع هزینه مدل در دوره‌های آموزشی مختلف امتحان شد. پس از استقرار مدل از تصاویر گروه اعتبارسنجی برای ارزیابی نهایی کارکرد مدل استفاده شد. لازم به ذکر است که در الگوریتم‌های یادگیری عمیق، که استخراج ویژگیهای اشیاء در آنها به صورت خودکار انجام می‌شود، نسبت به روشهای یادگیری ماشین، که در آنها استخراج و معرفی ویژگیها به صورت دستی صورت می‌گیرد، می‌توان از تعداد تصاویر مرحله آموزش و آزمون مدل کم کرد. از طرف دیگر تعداد تصاویر لازم را می‌توان با توجه به کارکرد مدل تعیین

Detecto بر مبنای مدل Faster R-CNN استوار است که معماری آن در شکل ۱ نشان داده شده است، بنابراین می‌توان گفت که الگوریتم Detecto الگوریتمی است بر مبنای شبکه عصبی عمیق پیچشی Convolutional Neural Network (CNN). همانطور که در شکل مشاهده می‌شود، ساختار این الگوریتم از دو بخش استخراج ویژگی (Feature Extractor) و دسته‌بندی شیء (Object Classification) تشکیل شده است. در بخش اول علاوه بر استخراج ویژگیها، کادرهایی هم به منظور استفاده در فرآیند کادربندی اشیاء توسط شبکه پیشنهاد کادر (RPN) ارائه می‌شود. در بخش دسته‌بندی اشیاء از تمام این اطلاعات برای دسته‌بندی و اعمال کوچکترین کادرهای محیط بر اشیاء استفاده می‌شود.

درآمد و در مسیر کاری زبان برنامه نویسی Python قرار داده شد.

۳- به کارگیری کتابخانه Detecto در فرآیند آموزش و آزمون مدل (Bi, 2022):

کد مربوط به فرآیند آموزش و آزمون به زبان برنامه‌نویسی Python و با استفاده از توابع کتابخانه Detecto آماده و اجرا شد. در پیوست مقاله کد مورد نظر به همراه توضیحاتی راجع به کارکرد دستورات مختلف آن آورده شده است. خروجی حاصل از فرآیند آموزش و آزمون، نمودارهایی است که مقدار تابع هزینه ایجاد شده در دوره‌های مختلف آموزش و آزمون مدل را نشان می‌دهد. توضیح اینکه هر چقدر مقدار تابع هزینه کمتر باشد به معنی این است که مدل در بخشهای آموزش و آزمون بهتر عمل کرده است و نتیجه پیش‌بینی‌ها با واقعیت همخوانی بیشتری دارد. لازم به ذکر است که الگوریتم



شکل ۱- معماری مدل Faster R-CNN که در الگوریتم Detecto به کار رفته است (<https://viso.ai/deep-learning/faster-r-cnn-2>)

۴- ارزیابی مدل با استفاده از تصاویر پوشه اعتبارسنجی:

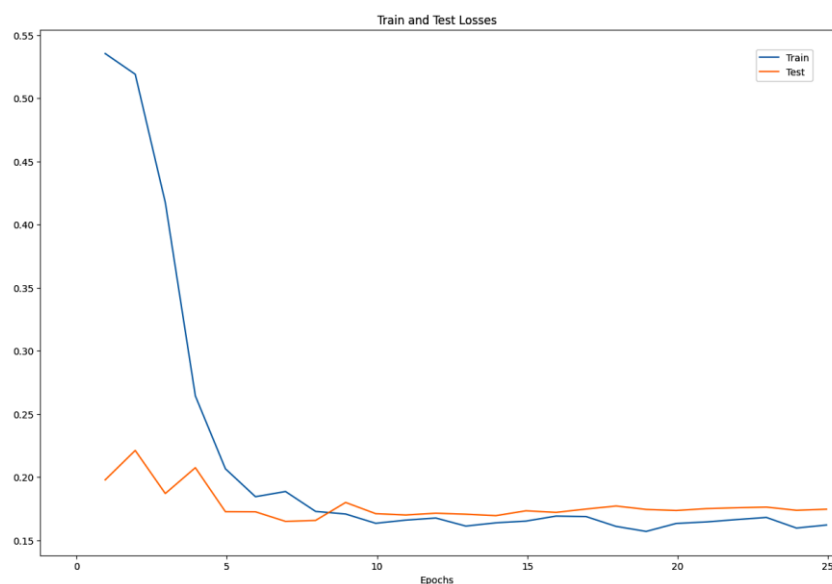
برای ارزیابی، از تعدادی تصویر نزدیک دارای یک یا دو سیب از پوشه اعتبارسنجی که شامل کادرهای

عکس می‌توان معیارهای حساسیت، دقت و نمره F1 را به ترتیب با استفاده از فرمولهای $Sensitivity = \frac{TP}{TP+FN}$ و $Precision = \frac{TP}{TP+FP}$ و $F1 - score = \frac{2 \times Sensitivity \times Precision}{(Sensitivity + Precision)}$ برای هر تصویر محاسبه کرد و از متوسط این معیارهای سه گانه که برای همه تصاویر مورد نظر محاسبه شده‌اند به عنوان حساسیت، دقت و نمره F1 مدل یاد کرد (Padilla et al., 2020). در این پژوهش اندازه‌گیری مساحت‌های اشتراکی و اجتماعی دو کادر واقعی و پیش‌بینی شده توسط نرم‌افزار KLOONK Image Measurement صورت گرفت. از سوی دیگر برای توسعه معیار ارزیابی کننده‌ای که شامل تمام ۱۶۶ تصویر پوشه اعتبارسنجی باشد، از متوسط صحت تشخیص سیب با سطح آستانه ۵۰٪ استفاده شد. به این صورت که تعداد تصاویر دارای بیش از ۵۰٪ تشخیص سیب شمارش شد و عدد حاصل بر عدد ۱۶۶ (تعداد کل تصاویر پوشه اعتبارسنجی) تقسیم شد.

نتایج و بحث

شکل ۲ تغییرات مقدار تابع هزینه محاسبه شده برای مراحل آموزش و آزمون در برابر شماره دوره آموزشی را نشان می‌دهد.

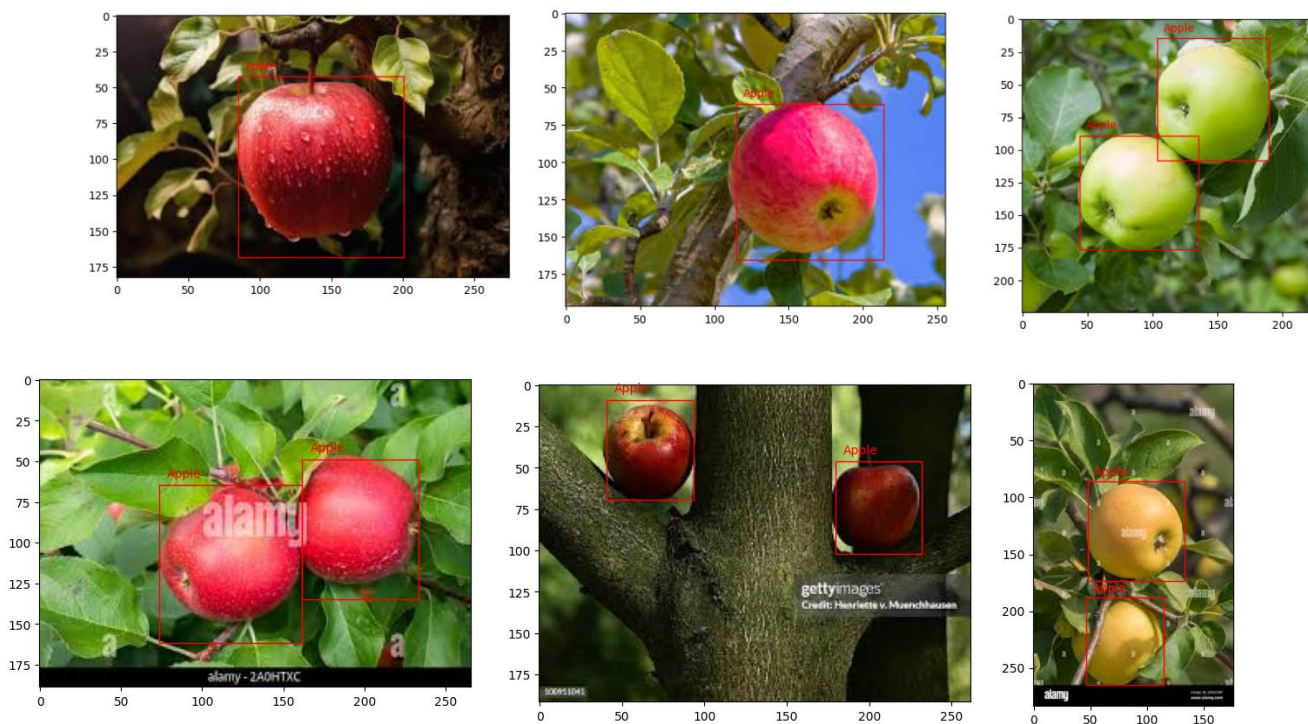
دارای برچسب روی سیب‌های تشخیص داده شده بودند، استفاده شد. روش مورد استفاده در ارزیابی مدل‌های بر مبنای تشخیص اشیاء، استفاده از مفهوم نسبت مساحت‌های اشتراکی به اجتماعی است. توضیح اینکه پس از کادردهی به میوه‌های سیب موجود در تصاویر توسط مدل، این کادرها با کادرهایی که در مرحله کادردهی توسط مدل‌ساز ایجاد شده‌اند مقایسه شده و نسبت مساحت اشتراکی دو کادر به مساحت اجتماعی آنها محاسبه می‌شود. اگر این نسبت از یک مقدار آستانه که توسط کاربر تنظیم می‌شود، بیشتر باشد، آن سیب به عنوان مثبت واقعی (TP) شناخته می‌شود. اگر نسبت محاسبه شده از مقدار آستانه کمتر باشد، آن سیب به عنوان مثبت کاذب (FP) در نظر گرفته می‌شود و اگر کادر ایجاد شده توسط مدل در محلی قرار گرفته باشد که هیچ همپوشانی با کادر اصلی نداشته باشد، آن سیب به عنوان منفی کاذب (FN) لحاظ می‌شود. بنابراین در مدل‌های تشخیص اشیاء برخلاف مدل‌های دسته‌بندی تصاویر، شیئی منفی واقعی (TN) نداریم چون در صورت عدم وجود یک شیئی در تصویر می‌توان بینهایت کادر تصور کرد که به درستی شامل آن شیئی نباشند. پس از محاسبه مقادیر TP، FP و FN مربوط به هر سیب و سپس هر



شکل ۲- نمودار تابع هزینه محاسبه شده در مراحل آموزش و آزمون بر حسب شماره دوره آموزشی

مشاهده می‌شود که فرآیند آموزش مدل باعث کاهش مقدار تابع هزینه در فرآیندهای آموزش و آزمون شده و در شماره دوره‌های آموزشی بیش از ۱۰ نمودارها تقریباً افقی شده‌اند، در نتیجه می‌توان گفت آموزش مدل کامل بوده و افزایش بیشتر شماره دوره باعث بهبود کارکرد مدل نخواهد شد.

در شکل ۳ برخی از نتایج به دست آمده در تشخیص میوه سیب روی درخت نشان داده شده است.

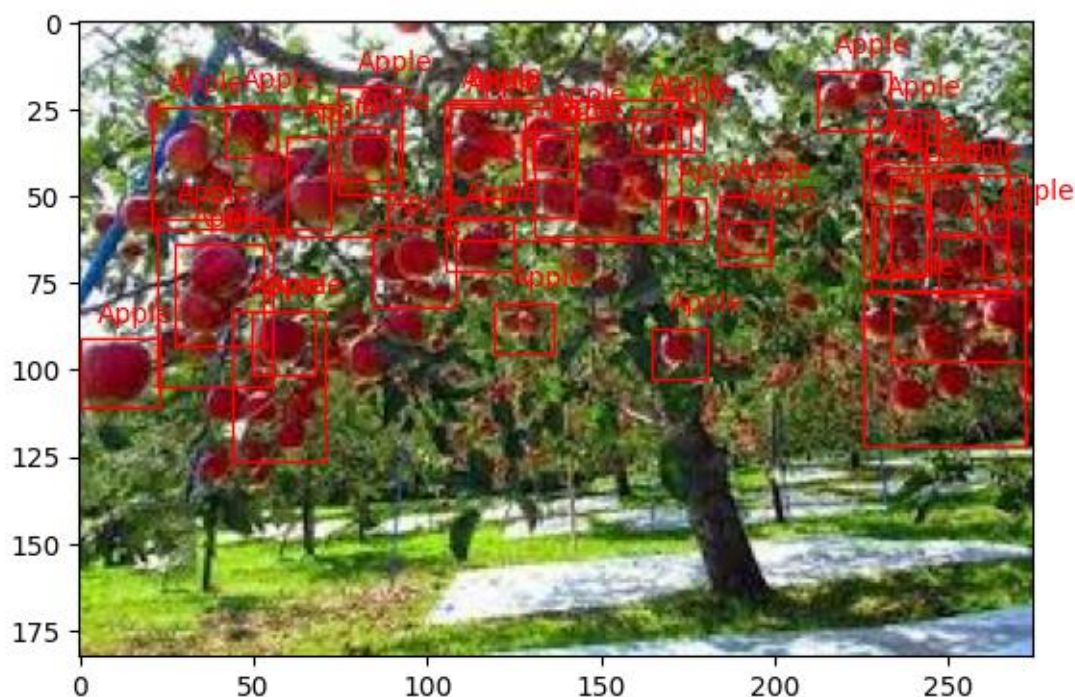


شکل ۳- نتایج به دست آمده از تشخیص سیب روی درخت در تصاویر مختلف پوشه اعتبارسنجی

احمدی و توسلی: پیاده‌سازی و ارزیابی الگوریتم یادگیری عمیق..... ۶۹

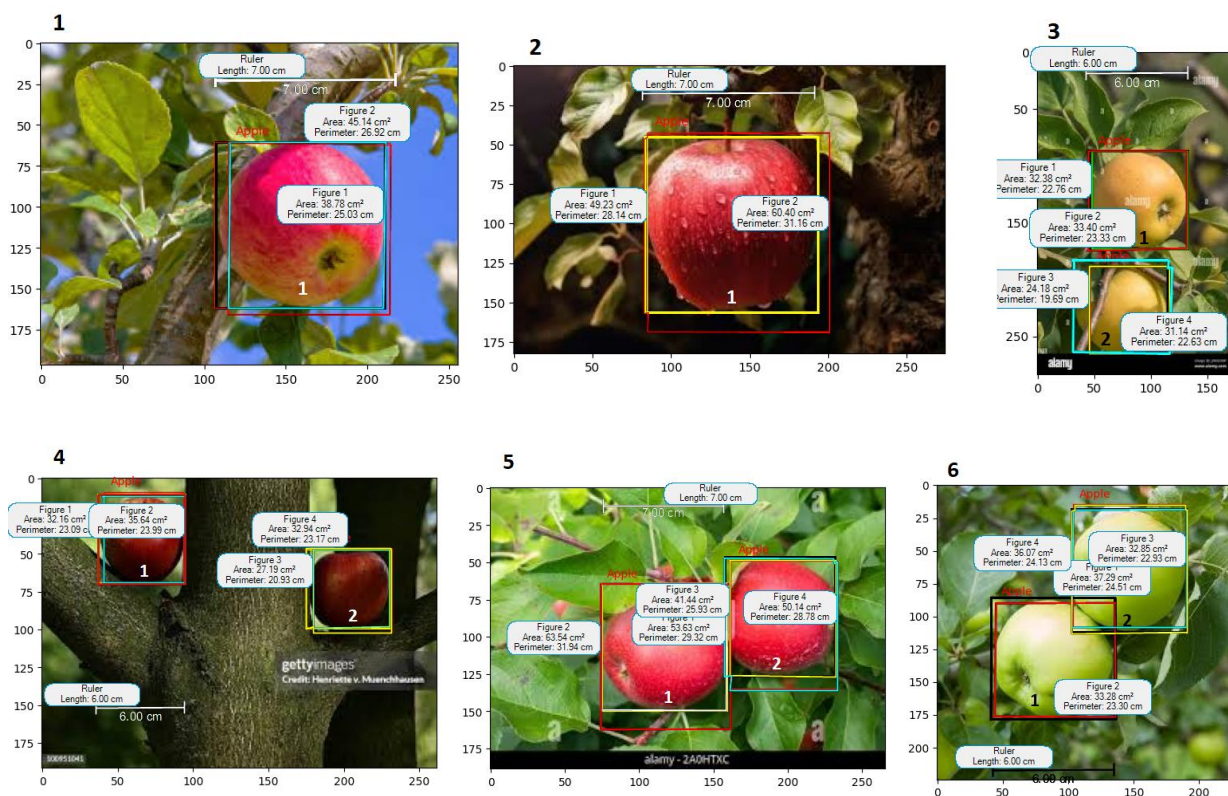


ادامه شکل ۳- نتایج به دست آمده از تشخیص سیب روی درخت در تصاویر مختلف پوشه اعتبارسنجی



ادامه شکل ۳- نتایج به دست آمده از تشخیص سیب روی درخت در تصاویر مختلف پوشه اعتبارسنجی

- با بررسی شکل ۳ نکات زیر به صورت کیفی به دست می‌آید:
- ۱- گونه سیب تأثیری در کارکرد مدل ندارد،
 - ۲- پس زمینه عکس‌ها باعث تأثیرگذاری بر نتایج مدل نشده است،
 - ۳- از نظر کیفی کارکرد مدل در تشخیص سیب روی تصاویر نزدیک، عالی بوده و مدل فقط برخی از میوه‌هایی که به علت پنهان ماندن نسبی پشت شاخ و برگ درخت، دارای تصویر ناقص بوده‌اند را نتوانسته تشخیص بدهد،
 - ۴- با دور شدن دوربین از درخت کیفیت تشخیص سیب به صورت منفرد تحت تأثیر قرار می‌گیرد و مواردی همچون از دست رفتن تشخیص صحیح و یا
- تشخیص چند سیب به عنوان یک سیب منفرد در تصاویر دور دیده می‌شود.
- نتایج به دست آمده در این پژوهش با گزارش پژوهش (Wosner et al., 2021) تطابق دارد، جایی که آنها نیز کارکرد سامانه تشخیص میوه با تصاویر نزدیک را با تشخیص انسان معادل دانستند اما به کاهش کارکرد سامانه در تصاویر دور و یا میوه‌های پوشیده شده با سایر عوارض درخت نیز اشاره داشتند.
- با توجه به نتایج به دست آمده کارکرد مدل به صورت کمی با در نظر گرفتن تعدادی از تصاویر نزدیک پوشه اعتبارسنجی صورت گرفت. شکل ۴ نتایج به دست آمده راجع به مقدار مساحت نواحی اشتراکی و اجتماعی میوه‌های سیب قرار گرفته در تعدادی از تصاویر پوشه اعتبارسنجی را نشان می‌دهد:



شکل ۴- مساحت‌های نواحی اشتراکی و اجتماعی حاصل از کادرهای واقعی و پیش‌بینی شده مربوط به مدل تشخیص سیب

است. در عمل کادرها انطباق کامل ندارند و می‌توان از نسبت مساحت اشتراکی دو کادر به مساحت اجتماعی آنها به عنوان معیار کیفیت انطباق دو کادر استفاده کرد. با توجه به مساحت‌های به‌دست آمده می‌توان جدول زیر را تکمیل کرد:

برای هر کدام از تصاویر سیب موجود در شکل ۴، دو کادر وجود دارد. یکی از آنها خروجی مدل در مورد سیب تشخیص داده شده در تصویر است و دیگری توسط شخص مدل‌ساز در مرحله کادربندی به‌وجود آمده است. اگر دو کادر کاملاً بر هم منطبق باشند به معنی تشخیص صحیح صد در صدی سیب

جدول ۱- روند محاسبه نسبت ناحیه اشتراکی به اجتماعی مربوط به هر کدام از سیب‌های موجود در تصاویر شکل ۴

شماره عکس	1	2	3	4	5	6
شماره سیب	1	1	1	2	1	2
مساحت اشتراکی	38.78	38.78	32.38	24.18	32.16	27.19
مساحت اجتماعی	45.14	45.14	33.4	31.14	35.64	32.94
نسبت مساحت اشتراکی به اجتماعی	0.86	0.86	0.97	0.78	0.9	0.83

صحیح ۰/۷۸ در نظر گرفته شود، دقت مدل ۱۰۰٪ خواهد بود و اگر سطح آستانه عددی بالاتر از ۰/۷۸ در نظر گرفته شود، معیارهای صحت کارکرد مدل با

کمترین مقدار نسبت دو مساحت برابر با ۰/۷۸ به‌دست آمده است. این عدد به این معنی است که اگر سطح آستانه نسبت مساحتی برای تشخیص

کاهش تدریجی مواجه خواهند شد. به عنوان مثال اگر سطح آستانه مساحتی ۰/۸ در نظر گرفته شود، مقادیر TP، FP و FN مربوط به سیبهای موجود در شکل ۴ مطابق آنچه در جدول ۲ نشان داده شده است، خواهد بود:

جدول ۲- مقادیر TP، FP و FN مربوط به سیبهای موجود در شکل ۴ محاسبه شده با سطح آستانه مساحتی برابر با ۰/۸

شماره عکس	1		2		3		4		5		6	
شماره سیب	1	1	1	2	1	2	1	2	1	2	1	2
TP	1	1	1	0	1	1	1	1	1	1	1	1
FP	0	0	0	1	0	0	0	0	0	0	0	0
FN	0	0	0	0	0	0	0	0	0	0	0	0

از سوی دیگر از آنجا که مقادیر معیارهای ارزیابی به ازاء هر عکس محاسبه می شوند، باید مقادیر TP، FP و FN مربوط به سیبهای یک تصویر با هم جمع شود تا بتوان به محاسبه معیارهای ارزیابی مدل اقدام کرد. حاصل کار در جدول ۳ نشان داده شده است:

جدول ۳- مقادیر TP، FP و FN مربوط به عکسهای موجود در شکل ۴ محاسبه شده با سطح آستانه مساحتی برابر با ۰/۸

شماره عکس	1	2	3	4	5	6
TP	1	1	1	2	2	2
FP	0	0	1	0	0	0
FN	0	0	0	0	0	0

با در اختیار بودن مقادیر TP، FP و FN مربوط به عکسها، مقدار معیارهای حساسیت، دقت و نمره F1 متعلق به هر عکس قابل محاسبه است. نتیجه به دست آمده در جدول ۴ نشان داده شده است:

جدول ۴- مقادیر حساسیت، دقت و نمره F1 مربوط به عکسهای موجود در شکل ۴ محاسبه شده با سطح آستانه مساحتی برابر با ۰/۸

شماره عکس	1	2	3	4	5	6
حساسیت	1	1	1	1	1	1
دقت	1	1	0.5	1	1	1
نمره F1	1	1	0.66	1	1	1

۰/۸۳۸ گزارش گردیده است که با نتایج به دست آمده در این پژوهش قابل مقایسه است. Seo et al., (2021) در تشخیص مراحل مختلف رشد محصول گوجه فرنگی در شرایط گلخانه‌ای به کمک پردازش تصویر به صحت ۹۰/۲٪ رسیدند. این نتیجه پتانسیل استفاده از این سامانه در پیش‌بینی زمان برداشت مناسب توسط ربات برداشت کننده محصول را اثبات کرد.

به منظور محاسبه معیارهای ارزیابی برای مدل توسعه یافته از میانگین معیارهای به‌دست آمده از تمام عکسها استفاده شد، بنابراین معیارهای حساسیت، دقت و نمره F1 مربوط به مدل در سطح آستانه نسبت مساحتی اشتراکی به اجتماعی برابر با ۰/۸، برترتیب برابر با ۱، ۰/۹۲ و ۰/۹۴ به‌دست آمد. در پژوهش دیگری که توسط (Sa et al., 2016) با استفاده از الگوریتم Faster R-CNN برای تشخیص فلفل شیرین انجام شده است، مقدار نمره F1 برابر با



از نظر کمی از ۱۶۶ تصویر پوشه اعتبارسنجی، در ۱۴۰ تصویر تشخیص بیش از پنجاه درصدی سیبهای موجود در هر تصویر حاصل شد و در ۶ تصویر هیچ سیبی وجود نداشته که الگوریتم نیز هیچ کادری را به عنوان میوه سیب در آنها تشخیص نداده است و فقط در ۲۰ تصویر درصد تشخیص کمتر از پنجاه درصد بوده است، بنابراین صحت متوسط الگوریتم برای آستانه تشخیص پنجاه درصدی برابر با $\frac{146}{166} \times 100 = 88\%$ به دست آمد که نشان دهنده عملکرد مناسب مدل در تشخیص سیب است.

نتیجه گیری

در صورتی که عکسهای داده شده به مدل پیاده سازی شده در این پژوهش از نوع نزدیک باشند، می توان اطمینان داشت که با دقتی در حد انسان، مدل قادر به تشخیص سیبهای کامل موجود در عکسهاست، هر چند که گونه سیبهای موجود در تصاویر از یک نوع نباشند. همچنین صحت مدل در تشخیص سیب در تصاویر دور نیز قابل قبول بوده، بنابراین از مدل پیاده سازی شده در این پژوهش می توان در پژوهشهای آتی مثل توسعه ربات برداشت کننده سیب استفاده کرد.

در پژوهش دیگری (Ge et al., 2022) با استفاده از الگوریتم Deep sort به دقت های متوسط (mAP) بترتیب ۹۳/۱٪، ۹۶/۴٪ و ۹۷/۹٪ در تشخیص گل، گوجه فرنگی نارس و گوجه فرنگی رسیده دست یافتند که نشان دهنده بترتیب ۱۷٪، ۲٪ و ۲/۳٪ بهبود در تشخیص در مقایسه با روش برمبنای الگوریتم YOLO-v5 بود.

(Rong et al., 2023) روش بهبود یافته ای را در تشخیص مقدار محصول دهی گلخانه گوجه فرنگی با استفاده از الگوریتم YOLO-4D توسعه دادند و به صحت تشخیص ۹۷/۹٪ و mAP برابر با ۰/۷۴۸ دست یافتند.

در پژوهش دیگری بزرگترین مانع در برداشت رباتی توت فرنگی، ویژگی پنهان ماندن نسبی محصول پشت ساقه و بوته عنوان شد. برای حل این مشکل (Xiong et al., 2021)، الگوریتم های Deep sort و YOLO-v4 را با هم ترکیب کرده و به صحت ۶۲/۴٪ دست یافتند که بیانگر ۳۶/۸٪ بهبود در فرآیند تشخیص توت فرنگی نسبت به پژوهش قبلی انجام شده توسط این پژوهشگران بود.

از سوی دیگر نتایج حاصل از مدل مورد استفاده در این پژوهش در تشخیص میوه سیب روی تمام ۱۶۶ تصویر پوشه اعتبارسنجی در فضای ابری Google قرار داده شد که از طریق QRcode زیر قابل دستیابی است:

REFERENCES

- Berenstein, R., Shahrar, O.B., Shapiro, A., Edan, Y., 2010. Grape clusters and foliage detection algorithms for autonomous selective vineyard sprayer. *Intel. Serv. Robot.* 3, 233–243.

- Bi, A., 2022. Build fully-functioning computer vision models with PyTorch. <https://github.com/alankbi/detecto>.
- Chang, A., Yeom, J., Jung, J., Landivar, J., 2020. Comparison of canopy shape and vegetation indices of citrus trees derived from UAV multispectral images for characterization of citrus greening disease. *Remote Sensing* 12, 4122. <https://doi.org/10.3390/rs12244122>.
- Costa, C., Schurr, U., Loreto, F., Menesatti, P., Carpentier, S., 2019. Plant phenotyping research trends, a science mapping approach. *Front. Plant Sci.* 9, 1933.
- Farjon, G., Krikeb, O., Hillel, A.B., Alchanatis, V., 2019. Detection and counting of flowers on apple trees for better chemical thinning decisions. *Precision Agric.* 1–19.
- Fuentes, A., Yoon, S., Kim, S.C., Park, D.S., 2017. A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. *Sensors* 17, 2022.
- Ge, Y., Lin, S., Zhang, Y., Li, Z., Cheng, H., Dong, J., Shao, S., Zhang, J., Qi, X., Wu, Z., 2022. Tracking and counting of tomato at different growth period using an improving YOLO-deepsort network for inspection robot. *Machines* 10, 489. <https://doi.org/10.3390/machines10060489>.
- Gutiérrez, S., Diago, M.P., Fernández-Navales, J., Tardaguila, J., 2018. Vineyard water status assessment using on-the-go thermal imaging and machine learning. *PLoS One* 13, e0192037.
- Magalhães, S.C., dos Santos, F.N., Machado, P., Moreira, A.P., Dias, J., 2023. Benchmarking edge computing devices for grape bunches and trunks detection using accelerated object detection single shot multibox deep learning models. *Engineering Applications of Artificial Intelligence* 117, 105604. <https://doi.org/10.1016/j.engappai.2022.105604>.
- Marinoudi, V., Sørensen, C.G., Pearson, S., Bochtis, D., 2019. Robotics and labour in agriculture. A context consideration. *Biosystem Engineering* 184, 111–121. <https://doi.org/10.1016/j.biosystemseng.2019.06.013>.
- Martínez-Peña, R., Velez, S., Vacas, R., Martín, H., Álvarez, S., 2023. Remote sensing for sustainable pistachio cultivation and improved quality traits evaluation through thermal and non-thermal UAV vegetation indices. *Applied Sciences* 13, 7716. <https://doi.org/10.3390/app13137716>.
- Padilla, R., Netto S.L., da Silva, E.A.B., 2020. A Survey on Performance Metrics for Object-Detection Algorithms. *International Conference on Systems, Signals and Image Processing (IWSSIP)*, Niteroi, Brazil, 2020, pp. 237-242, doi: 10.1109/IWSSIP48289.2020.9145130.
- Rong, J., Zhou, H., Zhang, F., Yuan, T., Wang, P., 2023. Tomato cluster detection and counting using improved YOLOv5 based on RGB-D fusion. *Computers and Electronics in Agriculture* 207, 107741. <https://doi.org/10.1016/j.compag.2023.107741>.
- Sa, I., Ge, Z., Dayoub, F., Upcroft, B., Perez, T., McCool, C., 2016. Deepfruits: A fruit detection system using deep neural networks. *Sensors* 16, 1222.
- Santos, T.T., de Souza, L.L., dos Santos, A.A., Avila, S., 2020. Grape detection, segmentation, and tracking using deep neural networks and three-dimensional association. *Comput. Electron. Agric.* 170, 105247.
- Seo, D., Cho, B.-H., Kim, K.-C., 2021. Development of monitoring robot system for tomato fruits in hydroponic greenhouses. *Agronomy* 11, 2211. <https://doi.org/10.3390/agronomy11112211>.
- Sishodia, R.P., Ray, R.L., Singh, S.K., 2020. Applications of remote sensing in precision agriculture: A review. *Remote Sensing* 12, 3136. <https://doi.org/10.3390/rs12193136>.

- Tsouros, D.C., Bibi, S., Sarigiannidis, P.G., 2019. A review on UAV-based applications for precision agriculture. *Information* 10, 349. <https://doi.org/10.3390/info10110349>.
- V´elez, S., Ariza-Sentís, M., Valente, J., 2023. Mapping the spatial variability of Botrytis bunch rot risk in vineyards using UAV multispectral imagery. *European Journal of Agronomy* 142, 126691. <https://doi.org/10.1016/j.eja.2022.126691>.
- V´elez, S., Rançon, F., Barajas, E., Brunel, G., Rubio, J.A., Tisseyre, B., 2022. Potential of functional analysis applied to Sentinel-2 time-series to assess relevant agronomic parameters at the within-field level in viticulture. *Computers and Electronics in Agriculture* 194, 106726. <https://doi.org/10.1016/j.compag.2022.106726>.
- Vitzrabin, E., Edan, Y., 2016. Adaptive thresholding with fusion using a rgb-d sensor for red sweet-pepper detection. *Biosyst. Eng.* 146, 45–56.
- Wosner, O., Farjon, G., Bar-Hillel, A., 2021. Object detection in agricultural contexts: A multiple resolution benchmark and comparison to human. *Computers and Electronics in Agriculture* 189, 106404. <https://doi.org/10.1016/j.compag.2021.106404>.
- Xiong, H., Cao, Z., Lu, H., Madec, S., Liu, L., Shen, C., 2019. Tasselnetv2: in-field counting of wheat spikes with context-augmented local regression networks. *Plant Methods* 15, 150.
- Xiong, Y., Ge, Y., From, P.J., 2021. An improved obstacle separation method using deep learning for object detection and tracking in a hybrid visual control loop for fruit picking in clusters. *Computers and Electronics in Agriculture* 191, 106508. <https://doi.org/10.1016/j.compag.2021.106508>.
- Xu, G., Khan, A.S., Moshayedi, A.J., Zhang, X., Shuxin, Y., 2022. The object detection, perspective and obstacles in robotic: A review: *EAI endorsed trans. AI Robotics* 1, e13–e. <https://doi.org/10.4108/airo.v1i1.2709>.
- Zhou, Y., Rui, T., Li, Y., Zuo, X., 2019. A UAV patrol system using panoramic stitching and object detection. *Computers and Electrical Engineering* 80, 106473. <https://doi.org/10.1016/j.compeleceng.2019.106473>.

پیوست

کد استفاده شده در این پژوهش

توضیح	کد
خارج کردن فایل فشرده محتوی پوشه ها از حالت فشرده، فراخوانی کتابخانه Pytorch و چک کردن دسترسی به cuda	<pre>!unzip apple.zip import torch print(torch.cuda.is_available())</pre>
نصب کتابخانه detecto و انجام سایر فراخوانی های کتابخانه های مورد نیاز برای ایجاد مدل	<pre>!pip install detecto from detecto import core, utils, visualize from detecto.visualize import show_labeled_image, plot_prediction_grid from torchvision import transforms import matplotlib.pyplot as plt import numpy as np</pre>
تبدیل های مورد نیاز که روی تصاویر ورودی صورت می گیرد تا همه تصاویر با استاندارد یکسان در فرآیند مدل سازی وارد شوند	<pre>custom_transforms = transforms.Compose([transforms.ToPILImage(), transforms.Resize(900), transforms.RandomHorizontalFlip(0.5), transforms.ColorJitter(saturation=0.2), transforms.ToTensor(), utils.normalize_transform(),])</pre>
دستورات لازم برای ایجاد مدل	<pre>Train_dataset=core.Dataset("apple/Train/"),transform=custom_transforms) Test_dataset = core.Dataset("apple/Test/") loader=core.DataLoader(Train_dataset, batch_size=2, shuffle=True)</pre>

model = core.Model(["Apple"]) losses = model.fit(loader, Test_dataset, epochs=25, lr_step_size=5, learning_rate=0.001, verbose=True)	
plt.plot(losses) plt.show()	رسم منحنی loss برحسب epoch number
model.save("model_weights.pth") model = core.Model.load("model_weights.pth", ["Apple"])	ذخیره سازی وزنهای به دست آمده در فرآیند آموزش مدل
image = utils.read_image("apple/Validation/08.jpg") predictions = model.predict(image) labels, boxes, scores = predictions show_labeled_image(image, boxes, labels) thresh=0.8 filtered_indices=np.where(scores>thresh) filtered_scores=scores[filtered_indices] filtered_boxes=boxes[filtered_indices] num_list = filtered_indices[0].tolist() filtered_labels = [labels[i] for i in num_list] show_labeled_image(image, filtered_boxes, filtered_labels)	دستورات لازم برای ارزیابی کارکرد مدل روی تصاویر پوشه اعتبارسنجی. برای انجام ارزیابی تک تک تصاویر به صورت منفرد در اختیار مدل قرار داده می‌شود و کارکرد مدل روی آن تصویر با نشان داده شدن کادر پیش بینی کننده میوه به عنوان خروجی ارائه می‌شود.



Application and Assessment of Detecto's Deep Learning Algorithm for Identifying Apple Fruit on Trees

Iman Ahmadi ^{1*} and Fateme Tavassoli ²

¹ Associate Professor, Department of Mechanical Engineering of Agricultural Machinery, Khorasgan Branch, Islamic Azad University, Isfahan, Iran

² Master's Degree in Applied Mathematics, Yazd University, Yazd, Iran

* Corresponding Author's Email: imanahmadi2013@gmail.com,
(Received: February. 1, 2025 – Accepted: March. 16, 2025)

ABSTRACT

One area where deep learning models have shown their potential is the identification of objects within images. The focus of this study is to develop a deep learning model to identify apple fruit on a tree using images captured from trees of apple orchards. The performance of the model was evaluated by calculating the ratio of the intersection to the union areas of the actual and predicted boxes surrounding the apples in the image. This was done on a number of close-up images with one or two apples, and then the sensitivity, accuracy, and F1 score metrics were calculated. In order to use all images of the validation folder in evaluating the performance of the algorithm, the average accuracy criterion for apple detection in all images was used, taking into account a detection threshold level of 50%. According to the results acquired, the model training process decreased the cost function's value, and after more than 10 epochs, the graph appeared nearly flat. Furthermore, the variety of apple and the type of the image backgrounds did not influence the model's results. Additionally, regarding quality, the model performed exceptionally well in identifying apples in close-up photos. Quantitatively, the sensitivity, accuracy, and F1 score measures of the model on a number of close-up images with one or two apples were obtained as 1, 0.92, and 0.94, respectively, considering an 80% detection threshold level. On the other hand, the average accuracy obtained on all images in the validation folder was 88%, considering a 50% detection threshold level. Thus, the effectiveness of the model used in this study was validated in identifying apple fruit on trees.

Keywords: Apple Recognition, Deep Learning, Detecto Algorithm, IOU Criterion