

## تشخیص هوشمند تاسل در تصاویر پهپادی با استفاده از یادگیری عمیق برای تعیین تاریخ گل دهی

سیده فروه موسوی<sup>۱</sup> و اعظم کرمی<sup>۲</sup>

### چکیده

تخمین عملکرد و بررسی روند رشد در گونه‌های مختلف از یک محصول در کشاورزی دقیق برای محققین و کارشناسان حوزه کشاورزی بسیار حائز اهمیت است. در این مقاله روشی نوین مبتنی بر یادگیری عمیق تک مرحله‌ای به نام GP-YOLOv5 برای شناسایی خودکار تاسل در تصاویر پهپادی از یک مزرعه بزرگ ذرت در تاریخ‌های مختلف رشد و تخمین زمان گل‌دهی ارائه شده است. در این راستا ابتدا به دلیل رشد تعداد کمی از تاسل‌ها در مراحل اولیه رشد برای داده‌افزایی از شبکه مولد متخاصم GP-GAN استفاده شد. سپس برای شمارش و تشخیص تاسل‌ها ساختار و پارامترهای آشکارساز YOLOv5 برای افزایش دقت مطابق با پایگاه داده اصلاح شد. در ادامه شمارش گیاهان در مراحل اولیه کاشت به عنوان یک پارامتر مهم در تعیین تاریخ گل‌دهی در نظر گرفته شد. شمارش گیاهان با استفاده از آشکارساز CenterNet انجام شده است و از الگوریتم‌های درون‌یابی و پیش‌بینی برای تعیین تاریخ گل‌دهی استفاده شد. روش پیشنهادی با دو روش معتبر مبتنی بر تشخیص CenterNet و روش مبتنی بر رگرسیون TasselNetv2+ برای شمارش تاسل‌ها مقایسه شد. دقت میانگین در تشخیص صحیح تاسل‌ها در روش پیشنهادی ۹۶/۸۱ و در روش CenterNet، ۸۱/۷۸ درصد است که نشان می‌دهد دقت روش پیشنهادی بالاتر از روش CenterNet است.

### کلیدواژه‌ها

شبکه‌های مولد متخاصم، یادگیری عمیق، YOLOv5، شناسایی تاسل، تخمین تاریخ گل‌دهی

### ۱ مقدمه

ذرت جز پنج محصول با ارزش دنیا به حساب می‌آید که به جز استفاده غذایی در زمینه تولید سوخت نیز از آن استفاده می‌شود. هدف این مقاله تشخیص گونه‌های مختلف ذرت در تصاویر رنگی اخذ شده توسط پهپاد و همچنین تخمین زمان گل‌دهی است. زمانی که بیشتر از نصف ذرت‌ها گل داده باشند را زمان گل‌دهی در نظر می‌گیرند. در حال حاضر زمان گل‌دهی بصورت بصری تعیین می‌شود. روش‌های سنتی شناسایی، شمارش و بررسی گیاهان یک مزرعه کاری زمانبر، پرهزینه و نیازمند تعداد زیادی نیروی انسانی است [۵]. بنابراین تصاویر پهپادی جایگزین مناسبی برای روش سنتی هستند اما تجزیه و تحلیل تصاویر پهپادی به دلیل تعداد بالای تصاویر اخذ شده از یک مزرعه نیز بسیار زمانبر است. امروزه، محققان در حال بررسی و جایگزین کردن رویکردهای مبتنی بر پردازش تصویر هستند. این امر سبب کاهش هزینه و

یکی از نوآوری‌های اخیر در کشاورزی دقیق، استفاده از پهپاد است. با استفاده از تصاویر اخذ شده توسط پهپاد می‌توان تخمین میزان محصول، تحلیل میزان رشد گیاه، تشخیص آفات، شمارش گیاهان [۱] و... را با دقت بالاتری در مقایسه با روش‌های سنتی انجام داد [۲-۴].

این مقاله در مردادماه سال ۱۴۰۰ دریافت، در مهرماه بازنگری و در آبان‌ماه همان سال پذیرفته شد.

<sup>۱</sup> آزمایشگاه پردازش تصویر و رباتیک، دانشکده فیزیک، دانشگاه شهید باهنر کرمان

رایانامه: farvehmusavi@gmail.com

<sup>۲</sup> دانشکده فیزیک، دانشگاه شهید باهنر کرمان

رایانامه: akarami@uk.ac.ir

نویسنده مسئول: اعظم کرمی

تابع هدف طبقه بندی کرد. از جمله شبکه‌های مولد متخاصم می‌توان به SRGAN<sup>۷</sup> [۱۷]، CycleGAN [۱۸]، InfoGAN [۱۹] و... اشاره نمود. از شبکه‌های مولد متخاصم در زمینه‌های مختلف از جمله شبیه سازی داده [۲۰]، افزایش داده برای مجموعه داده‌های کوچک [۲۱]، تغییر شکل [۱۸]، تولید تصاویر ساختگی پزشکی [۲۲، ۲۳]، انتقال تصویر، افزایش وضوح و ترکیب تصویر و تولید فیلم و... استفاده می‌شود.

انتخاب آشکارساز مناسب نیز برای تشخیص شی در پایگاه داده‌های کوچک امری بسیار مهم محسوب می‌شود. در این راستا روش‌هایی، مانند یادگیری انتقالی<sup>۸</sup> [۲۴] و افزایش تعداد داده<sup>۹</sup> [۲۵] معرفی شده‌اند. بطور کلی آشکارسازهای مبتنی بر یادگیری عمیق به دو دسته مبتنی بر کادر [۲۶] و بدون کادر [۲۶] دسته بندی می‌گردند. شناسه گذاری تصاویر در آشکارسازهای بدون کادر بسیار ساده‌تر از شناسه گذاری آشکارسازهای مبتنی بر کادر است. به همین دلیل اخیراً آشکارسازهای بدون کادر مورد توجه قرار گرفته‌اند. از جنبه‌ی دیگر آشکارسازها به دو دسته تک مرحله‌ای و دو مرحله‌ای تقسیم می‌شوند. در روش دو مرحله‌ای بررسی منطقه-ای انجام می‌شود و مختصات کادرهایی که احتمال وجود شی را دارد استخراج می‌کند و سپس به طبقه‌بندی اشیاء درون کادر می‌پردازد. از جمله آشکارسازهای دو مرحله‌ای می‌توان به Faster R-CNN [۲۷] و RetinaNet [۲۸] اشاره کرد. روش تک مرحله‌ای آشکارساز در مکان‌ها و اندازه‌های مشخص دنبال اجسام می‌گردد و استخراج مختصات کادرها و طبقه‌بندی در یک مرحله انجام می‌شود. به همین دلیل سرعت بالاتری نسبت به آشکارسازهای دو مرحله‌ای دارد. YOLO<sup>۱۰</sup> [۲۹] و SSD<sup>۱۱</sup> [۳۰] از جمله آشکارسازهای تک مرحله‌ای می‌باشند. YOLO قابلیت استفاده به صورت بلادرنگ را دارد.

در این تحقیق برای داده افزایی از GAN-<sup>۱۲</sup> GP به دلیل قابلیت ادغام تصاویر با رزولوشن بالا استفاده شده و متناسب با پایگاه داده اصلاح شده است. برای تشخیص تاسل نیز با توجه به دقت و سرعت بالا در تشخیص اشیا YOLOv5 از این آشکارساز استفاده شد. برای تعیین تاریخ گل‌دهی دو روش پیش بینی و درون‌یابی استفاده شد. در ادامه‌ی این مقاله، در بخش دوم به مروری بر تحقیقات پیشین و در بخش سوم به معرفی روش پیشنهادی پرداخته می‌شود. در گام بعدی در بخش چهارم نتایج عملی ارائه شده و در نهایت نتیجه گیری و معرفی کارهای آینده در بخش پنجم شرح داده شده است.

بالا بردن دقت در مقایسه با نیروی انسانی شده است [۶]. برای محصول ذرت، تشخیص تاسل<sup>۱</sup> ضروری است زیرا تاسل‌ها در مورد کرده افشانی، مراحل تولید مثل و عملکرد رشد تاثیر گذار هستند [۷]. زمان گل‌دهی اصلی‌ترین عامل تعیین کننده میزان بازدهی محصول است [۸]. شناسایی تاسل‌ها به دلیل تنوع در اندازه و شکل، پس‌زمینه‌ها، میزان روشنایی و سایر موارد در تصاویر اخذ شده توسط پهپاد، چالش برانگیز است. از آنجا که مدل‌های مختلف شمارش دارای ویژگی‌های متمایز هستند، انتخاب یک مدل مناسب بسیار مهم است. یکی از روش‌های پیشنهاد شده استفاده از تصاویر چند طیفی است [۹]. این روش نیازمند دوربین‌های مخصوص است و هزینه تصویربرداری آن بالا است. اکثر سیستم‌های مبتنی بر تصویر برای فنوتیپ گیاهان در محیط آزمایشگاهی کنترل شده مستقر می‌شوند. هنگام انتقال سناریوی برنامه به شرایط غیرمحدودی در میدان، تغییرات ذاتی و برونی در طبیعت چالش‌های بزرگی برای شمارش دقیق تاسل‌های ذرت به وجود می‌آورند، که فراتر از توانایی روش‌های کلاسیک پردازش تصویر مانند روش‌های مبتنی بر ماشین بردار پشتیبان<sup>۲</sup> [۱۰] و عملگرهای مورفولوژی<sup>۳</sup> [۱۱] است.

در سال‌های اخیر الگوریتم‌های متعددی مبتنی بر یادگیری عمیق برای شناسایی شی معرفی شده است، اما برای دستیابی به دقت مطلوب ابتدا باید پایگاه داده بزرگ و دقیق تهیه شود که بتوان داده کافی برای آموزش تهیه نمود [۱۲]. تهیه چنین پایگاه داده‌ای بسیار وقت‌گیر و هزینه‌بر است. یکی از چالش‌ها در حوزه کشاورزی نیز کمبود داده در روزهای ابتدایی کاشت به دلیل رشد تعداد کمی از گیاهان است. راهکارهای متعددی برای افزایش تعداد نمونه‌ها معرفی شده است. به عنوان مثال تبدیل‌های گوناگونی از قبیل چرخش، کوچک‌نمایی، بزرگ‌نمایی، تغییر شدت روشنایی و ایجاد نویز بر روی تصاویر اصلی اعمال می‌شوند تا تعداد تصاویر آموزشی افزایش پیدا کند [۱۳]. بطور کلی مدل‌های تولیدی را می‌توان به سه دسته تقسیم کرد: شبکه‌های مولد متخاصم<sup>۴</sup> [۱۴]، رمزگذار خودکار متنوع<sup>۵</sup> [۱۵] و شبکه‌های خود رگرسیون<sup>۶</sup> [۱۶].

شبکه‌های مولد متخاصم (GANها) از تئوری بازی الهام گرفته شده‌اند، که شامل مولد و متمایز کننده است. وظیفه مولد تولید داده‌های جعلی تا حد امکان متناسب با توزیع داده‌های واقعی است، در حالیکه وظیفه متمایز کننده تشخیص صحیح داده‌های واقعی از داده‌های جعلی است. مدل‌های مختلف GAN را می‌توان به دو گروه GANهای مبتنی بر بهینه سازی معماری و بهینه سازی

<sup>7</sup> Super Resolution Generative Adversarial Network

<sup>8</sup> Transfer Learning

<sup>9</sup> Data augmentation

<sup>10</sup> You Only Look Once

<sup>11</sup> Single Shot Detector

<sup>12</sup> Gaussian Poisson

<sup>1</sup> Tassel

<sup>2</sup> Support Vector Machine (SVM)

<sup>3</sup> Morphology

<sup>4</sup> Generative Adversarial Network (GAN)

<sup>5</sup> Variational Auto Encoder

<sup>6</sup> Autoregressive

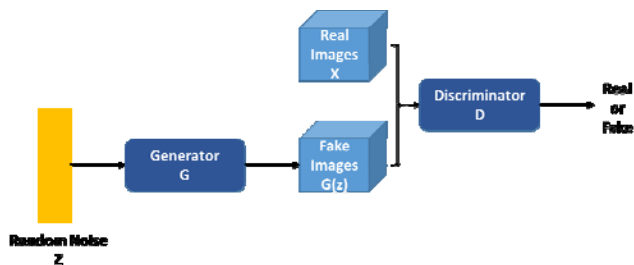
یادگیری ماشین، آموزش مدل ماشین بردار پشتیبان<sup>۲</sup> استفاده نمودند و خوشه‌های گل‌دار را در یک بازه زمانی تشخیص دادند. این مدل می‌تواند تغییرات روزانه را در طول مدت گل‌دهی مشخص کند [۳۷]. در سال ۲۰۱۹ Su و همکاران جهت تشخیص ارتفاع و سطح برگ گیاه، آنالیز تصاویر لایدار را پیشنهاد دادند که توسط آن‌ها می‌توان زمان گل‌دهی را تعیین نمود. این تحقیق بر روی ۲۰ گونه ذرت انجام شد و به نتایج قابل قبولی دست یافت [۳۸]. در سال ۲۰۲۱ Ji و همکاران برای تشخیص خودکار تاسل یک روش مبتنی بر پردازش تصویر و یادگیری ماشین را پیشنهاد دادند. در این روش از یک الگوریتم تشخیص Saliency برای تعیین موقعیت تاسل و به دست آوردن نقشه Saliency در یک تصویر استفاده شده است. سپس الگوریتم LS-SVM جهت طبقه‌بندی اعمال می‌گردد [۳۹].

### ۳ روش پیشنهادی

معماری روش پیشنهادی GP-YOLOv5 در ادامه شرح داده می‌شود.

#### ۳-۱ داده افزایی با استفاده از شبکه مولد متخاصم GP-GAN

معماری کلی GANها در شکل ۱ نشان داده شده است. ورودی مولد یک بردار نویز تصادفی  $z$  است. نویز از طریق مولد  $G$  به یک فضای داده جدید ترسیم می‌شود تا یک نمونه جعلی،  $G(z)$ ، بدست آید. متمایزکننده  $D$  یک طبقه‌بند باینری است که نمونه واقعی را از مجموعه داده و نمونه جعلی تولید شده توسط مولد  $G$  را به عنوان ورودی دریافت می‌کند. خروجی تفکیک کننده  $D$  تعیین کننده واقعی یا جعلی بودن نمونه است. زمانی که تفکیک کننده  $D$  نتواند تعیین کند که داده از مجموعه داده واقعی است یا مولد، حالت مطلوب حاصل می‌شود. در این مرحله، یک مولد  $G$  بدست می‌آید که توزیع داده‌های واقعی را یاد گرفته است.



شکل ۱: معماری شبکه مولد متخاصم

در این مقاله تاسل‌های برچسب گذاری شده در روزهایی که تاسل‌ها رشد کرده‌اند، به تصاویر ارتوفوتو روزهای اولیه از طریق شبکه GP-GAN اضافه می‌شوند. همانطور که در شکل ۲ نشان داده شده، این روش شامل دو مرحله است: در مرحله اول یک

### ۲ مروری بر تحقیقات پیشین

روش‌های مبتنی بر یادگیری عمیق متعددی که در سال‌های اخیر در زمینه تشخیص تاسل استفاده شده‌اند که به دو دسته اصلی مبتنی بر شمارش و تشخیص شی دسته بندی می‌شوند. در ذیل به چند مورد از آن‌ها اشاره شده است.

روش‌های مبتنی بر شمارش: در سال ۲۰۱۷ Lu و همکاران شبکه بر پایه رگرسیون و مبتنی بر شبکه عصبی TasselNet را پیشنهاد کردند. این شبکه با یادگیری نگاشت از ویژگی‌های تصویر شمارش تاسل را انجام می‌دهد [۳۱]. در سال ۲۰۱۹ Xiong جهت شمارش تاسل شبکه TasselNetv2 را با شبکه رگرسیون تقویت شده معرفی نمود که دقت و سرعت بالاتری نسبت به نسخه قبلی دارد. در این نسخه لایه‌های پیچشی جایگزین لایه‌های تمام متصل شد [۳۲]. در سال ۲۰۲۰ Lu و همکاران Cao شبکه TasselNetv2+ را پیشنهاد کردند. در این روش نرمالایز کننده استفاده شده در نسخه قبلی را تغییر دادند و رمزگذار آن را بهبود بخشیدند [۳۳]. نسخه اول و دوم TasselNet در نرم افزار متلب توسعه داده شده بودند که کاربردی نبودند و در نسخه TasselNetv2+ بر مبنای پایتون توسعه داده شد که به یک الگوریتم کاربردی در این زمینه تبدیل شد.

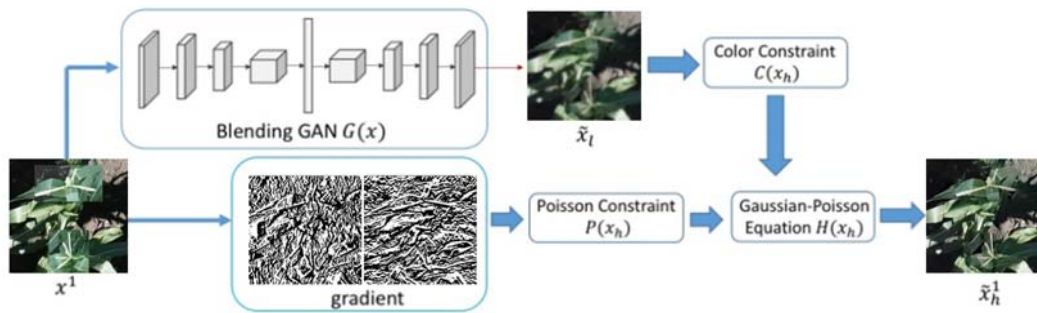
روش‌های مبتنی بر تشخیص شی: در سال ۲۰۲۰ Zou و همکاران عملکرد آشکارساز دو مرحله‌ای Faster R-CNN، آشکارساز تک مرحله‌ای RetinaNet و TasselNetv2+ را با یکدیگر مقایسه کردند. طبق نتایج بدست آمده شبکه RetinaNet عملکرد بهتری نسبت به دو آشکارساز دیگر داشته است. مدت زمان آموزش روش TasselNetv2+ کوتاه است اما این روش برخلاف روش‌های مبتنی بر تشخیص مانند Faster R-CNN، RetinaNet قادر به تشخیص دقیق موقعیت مکانی و اندازه تاسل نمی‌باشد [۳۴]. در سال ۲۰۲۱ Mirnezami و همکاران برای کاهش تعداد دوره‌های آموزش از یادگیری انتقالی در آشکارساز RetinaNet استفاده کردند و با تغییر پارامترها و داده‌های آموزش نتیجه گرفتند که بالاترین دقت مربوط به آموزش با استفاده از تصاویری است که قبلاً آن‌ها را ندیده است، زیرا در این حالت آشکارساز به جای حفظ الگوهای دیده شده ویژگی‌های جدید را یاد می‌گیرد. آن‌ها همچنین با استفاده از روش تقسیم بندی<sup>۱</sup> و به دست آوردن مشخصات فیزیکی ذرت زمان گل‌دهی را برای هر ذرت بدست آوردند [۳۵].

زمانی که بیشتر از نصف ذرت‌ها گل داده باشند را زمان گل‌دهی در نظر می‌گیرند. یکی از روش‌های تعیین زمان گل‌دهی استفاده از الگوی دمایی که این روش برای مناطق گرمسیر مناسب است [۳۶]. در سال ۲۰۱۵ Gue و همکاران جهت تشخیص خودکار گل‌دهی شلتوک برنج در تصاویر رنگی از یک روش

<sup>2</sup> Support Vector Machine (SVM)

<sup>1</sup> Segmentation

با وضوح بالا و واقع‌بینانه تولید می‌شود [۴۰]. هرم لاپلاسی در سطح اول از طریق تفاضل تصویر اولیه و تصویر گوسی سطح دوم که در فیلتر گوسی کانولوشن می‌شود بدست می‌آید و سطوح بعدی نیز به همین ترتیب بدست می‌آیند.

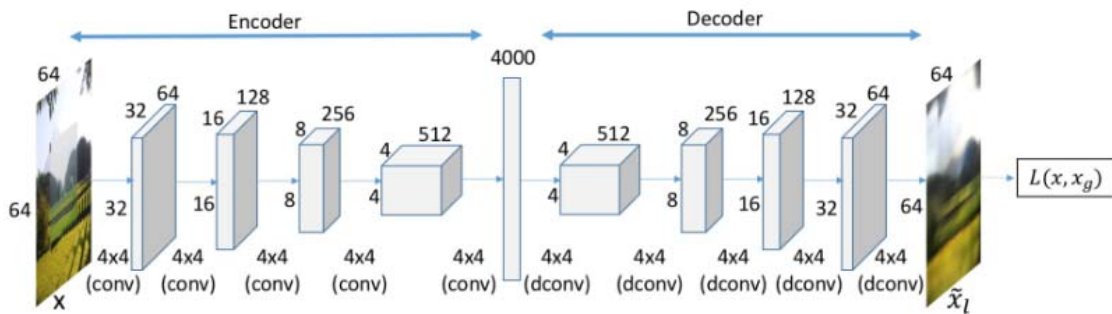


شکل ۲: مراحل تولید تصویر در روش GP-GAN [۴۰]

شود. چنین الزاماتی سبب در نظر گرفتن دو محدودیت رنگ و محدودیت شیب می‌شوند. محدودیت رنگ  $\tilde{x}_h$  را مجبور می‌کند که رنگی مشابه  $\tilde{x}_l$  داشته باشد که با تولید تصویری با سیگنال‌های فرکانس پایین می‌توان به آن دست یافت. ساده‌ترین راه برای استخراج سیگنال‌های با فرکانس پایین استفاده از فیلتر گوسی است. محدودیت شیب باعث بازگرداندن جزئیات با وضوح بالا می‌شود. این مرحله می‌تواند با استفاده از اپراتور واگرایی اجرا شود.

تصویر واقع‌گرایانه با رزولوشن پایین با استفاده از Blending GAN تولید می‌شود. در مرحله دوم معادله گوسین-پواسون، بر اساس بردار شیب تصویر تولید شده در مرحله اول که توسط هرم لاپلاسی بدست آمده، حل می‌شود. با استفاده از این روش تصاویر

در شکل ۳ معماری Blending GAN،  $G(x)$  نشان داده شده است. Blending GAN پیشنهادی فقط می‌تواند تصاویر با وضوح پایین ایجاد کند. برای تولید تصاویر با وضوح بالا از تصاویر واقع‌گرایانه تولید شده توسط Blending GAN، معادله گوسین-پواسون را که توسط هرم معروف لاپلاس ساخته شده استفاده می‌شود. اگر چه Blending GAN نمی‌تواند تصاویر با وضوح بالا تولید کند اما تصاویر تولید شده  $\tilde{x}_l$  دارای وضوح پایین، طبیعی و واقع‌گرایانه است. بنابراین می‌توان با گرفتن جزئیات غنی مانند بافت‌ها و لبه‌ها در تصویر اصلی با وضوح بالا  $x$ ، با تقریب رنگ  $\tilde{x}_l$  تصویری با وضوح بالا و واقعی  $\tilde{x}_h$  تولید



شکل ۳: معماری شبکه Blending GAN [۴۰]

$$C(x_h) = \int_T \left\| g(x_h) - \tilde{x}_h \right\|_2^2 dt \quad (۳)$$

$\beta$  پارامتر حفظ رنگ،  $T$  نشان دهنده کل منطقه تصویر است، عملگر  $\text{div}$ ، واگرایی را تفکیک می‌کند و  $\Delta$  نماینده عملگر لاپلاسی است.  $v$  نیز به صورت زیر تعریف می‌شود:

$$V^i = \begin{cases} \nabla x_{mask}^i & \text{if } x_{mask}^i = 1 \\ \nabla x_{dst}^i & \text{if } x_{mask}^i = 0 \end{cases} \quad (۴)$$

در روش پیشنهادی تابع هدف باید به صورت زیر بهینه شود:

$$H(x_h) = P(x_h) + \beta C(x_h) \quad (۱)$$

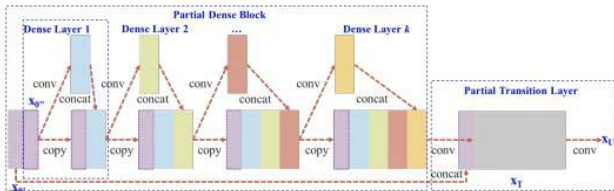
$P(x_h)$  از معادله معروف پواسون الهام گرفته شده است و به صورت زیر تعریف می‌شود:

$$P(x_h) = \int_T \left\| \text{div}(v) - \Delta x_h \right\|_2^2 dt \quad (۲)$$

$C(x_h)$  به صورت زیر تعریف می‌شود:

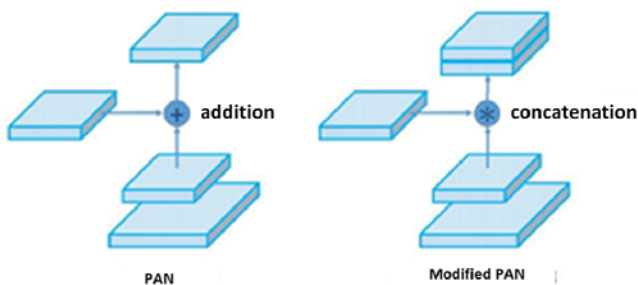


بر اساس ایده DensNet، با کپی کردن نقشه ویژگی لایه پایه و ارسال نسخه از طریق بلوک متراکم به مرحله بعدی، نقشه ویژگی لایه پایه را جدا می‌کند. شکل ۴ معماری CSPNet در نسخه چهارم و پنجم YOLO را نشان می‌دهد. این روش به طور موثر مشکل از بین رفتن شیب را کاهش می‌دهد و تغییرات گرادینان را از ابتدا تا انتها در نقشه ویژگی ادغام می‌کند، در نتیجه تعداد پارامترهای شبکه را کاهش می‌دهد.



شکل ۴: معماری CSPNet در نسخه چهارم و پنجم YOLO [۴۴]

طراحان YOLO در بخش گردن از تجمیع هرم فضایی (SPP) و شبکه متراکم‌سازی مسیر (PAN) استفاده کرده‌اند. البته شبکه PAN استفاده‌شده تفاوت‌هایی با شبکه اصلی دارد و نسخه اصلاح شده آن در نسخه چهارم و پنجم بکار گرفته شده که بجای افزودن لایه‌ها، آن‌ها را با هم تلفیق می‌کند.



شکل ۵: شبکه PAN اصلی و اصلاح شده در نسخه چهارم و پنجم YOLO [۴۵]

پیشرفت و تغییرهای عمده این نسخه، عبارتند از داده افزایشی موزاییکی<sup>۹</sup> و یادگیری خودکار کادرهای محصور کننده البته داده افزایشی موزاییکی در نسخه چهارم هم استفاده شده است. از آنجایی که معرفی نسخه چهارم و پنجم بسیار نزدیک به یکدیگر بود این امر باعث شد که معماری این دو نسخه بسیار شبیه به یکدیگر باشد اما در نسخه پنجم، زمانی که با اندازه دسته بزرگتری آزمایش شود، سرعت استنباط بالاتری نسبت به بسیاری از مدل‌های تشخیص دهنده دارد [۴۶]. وزن‌های بدست آمده از نسخه پنجم بسیار حجم کمتری نسبت به وزن‌های بدست آمده از نسخه چهارم دارد.

در این مقاله از YOLOv5s که کوچکترین مدل YOLO است استفاده شد. مدل‌های دیگری از YOLOv5 موجود است که از مدل استفاده شده بزرگتر هستند. با افزایش اندازه شبکه، ممکن است عملکرد آن بهبود یابد، اما زمان پردازش نیز افزایش می‌یابد.

### ۲-۳ تاریخچه نسخه اول، دوم، سوم و چهارم YOLO

نسخه اول YOLOv1 بصورت یکپارچه و بلادرنگ<sup>۱</sup> در سال ۲۰۱۶ معرفی شد [۲۹]. YOLO معماری الگوریتم‌های تشخیص اشیا را دست‌خوش تغییراتی کرده و برای تشخیص اشیا، مستقیم از پیکسل‌های تصویر به مختصات کادر و احتمال کلاس‌ها می‌رسد. YOLO بسیار ساده است و فقط یک شبکه کانولوشنی دارد که تصویر تغییر اندازه داده شده ورودی را دریافت و سپس به صورت همزمان چندین کادر را به همراه احتمال کلاس‌ها پیش‌بینی می‌کند. کادرهای با ضریب اطمینان بالا نگه داشته می‌شوند و بقیه کادرها حذف می‌گردند. ضریب اطمینان در واقع حاصل ضرب احتمال تشخیص صحیح شی در  $IoU^2$  است.

نسخه دوم با عنوان بهتر، سریعتر و قوی‌تر در سال ۲۰۱۷ منتشر شد [۴۱]. YOLOv2 نسبت به نسخه اولیه دارای دقت بیشتری است. این شبکه، از ۲۳ لایه کانولوشن تشکیل شده است؛ هر کدام از این لایه‌ها، واحد نرمال سازی دسته‌ای<sup>۲</sup>، تابع فعال‌سازی Leaky RELU و لایه‌های نمونه بردار حداکثری<sup>۴</sup> مختص به خود را دارد تا فرابرازش<sup>۵</sup> کم‌تر اتفاق بیافتد.

در سال ۲۰۱۸ نسخه سوم YOLO با عنوان یک پیشرفت روزافزون ارائه شد [۴۲]. در این نسخه از یک استخراج کننده‌ی ویژگی عمیق‌تر و قوی‌تر به نام DarkNet-53 استفاده شد که از ResNet الهام گرفته شده بود.

در سال ۲۰۲۰ نسخه چهارم با عنوان سرعت و دقت بهینه تشخیص اشیا را معرفی کردند [۴۳]. این آشکارساز به داده آموزش کمتری نیاز دارد. YOLOv4 از  $BoF^6$  منطبق بر آخرین پیشرفت‌های علمی و چندین  $BoS^7$  تاثیر می‌گیرد.  $BoF$  بدون افزایش مدت زمان استنباط، دقت مدل تشخیص دهنده را بهبود می‌بخشد و فقط هزینه آموزش را افزایش می‌دهد. از طرف دیگر،  $BoS$  با وجود اینکه هزینه استنباط را اندکی افزایش می‌دهد، دقت تشخیص اشیا را به طور قابل توجهی بالا می‌برد.

### ۳-۳ نسخه پنجم YOLO

دو ماه پس از منتشر شدن نسخه چهارم YOLO، نسخه پنجم آن ارائه شد. این نسخه با تمامی نسخه‌های قبل از خود تفاوت دارد؛ زیرا مبتنی بر کتابخانه PyTorch پیاده سازی شده است. در معماری YOLOv5 نیز همانند YOLOv4، قسمت پشتیبان از نوع CSP<sup>۸</sup> و قسمت گردن از نوع PAN<sup>۹</sup> است. CSPNet در واقع بر

<sup>1</sup> Real Time

<sup>2</sup> Intersection over Union

<sup>3</sup> Batch Normalization

<sup>4</sup> Max Pooling

<sup>5</sup> Overfitting

<sup>6</sup> Bag of Freebies

<sup>7</sup> Bag of Specials

<sup>8</sup> Cross Stage Partial DenseNet

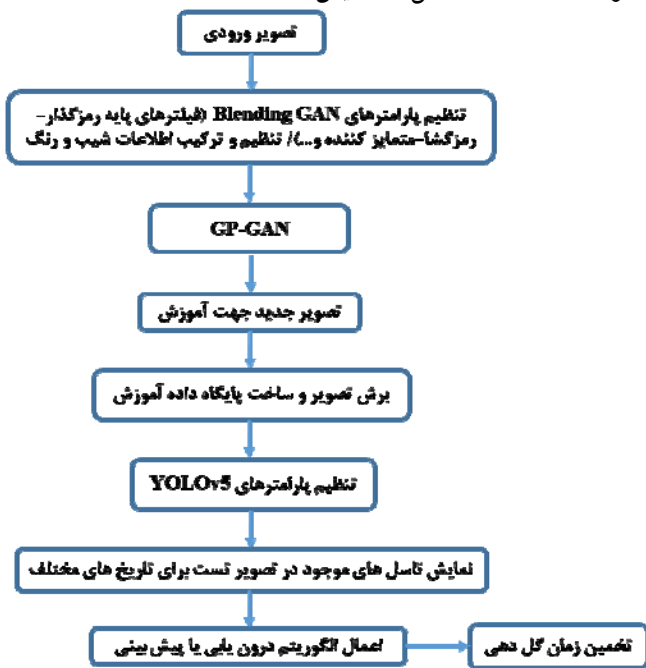
<sup>9</sup> Path Aggregation Network

<sup>10</sup> Mosaic Data Augmentation

$$b = \frac{S_{xy}}{S_{xx}} \quad (۸)$$

در این روابط  $x_i$  تعداد تاسل در تاریخ‌های قبل از گل دهی که توسط GP-YOLOv5 شمارش شده‌اند، می‌باشد.  $\bar{x}$  تعداد میانگین تاسل‌هایی است که در آن بازه زمانی رشد کرده‌اند.  $a$  عرض از مبدا و  $b$  شیب معادله خطی در نظر گرفته شده می‌باشد.  $N$  تعداد تصاویر تست است. پارامتر  $x$  تعداد نصف ذرت‌های شمارش شده در روزهای اولیه کاشت است و در نهایت  $y(x)$  تاریخ پیش‌بینی شده را مشخص می‌کند. برای درون‌یابی از روش خطی، یک خط از بین نقاط عبور داده می‌شود و طبق معادله خط در نظر گرفته شده مقدار نقطه مورد نظر بدست می‌آید.

شمای کلی الگوریتم پیشنهادی GP-YOLOv5 و مراحل آن به صورت خلاصه در شکل ۷ نمایش داده شده است.



شکل ۷: شمای کلی الگوریتم پیشنهادی GP-YOLOv5

## ۴ نتایج عملی

### ۴-۱ مشخصات سیستم پردازشی

سیستم عامل استفاده شده، ابونتو ورژن ۱۸،۰۴ است. آموزش آشکارسازها به وسیله یک سیستم با مشخصات زیر انجام شد.

جدول ۱: مشخصات سیستم کامپیوتری

CPU	Core i7-6850K 3.60GHz
RAM	128 Gb RAM
GPU	1x Nvidia Geforce GTX 1080 Ti (11Gb)

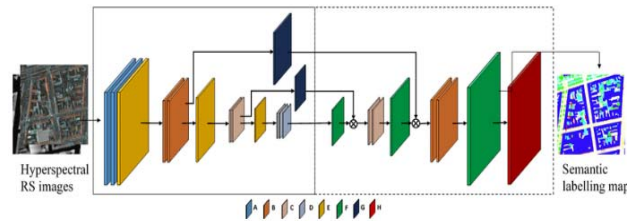
### ۴-۲ پایگاه داده

کاشت ذرت در مرکز تحقیقات و آموزش زراعت (ACRE) دانشگاه پردو انجام شده است. موقعیت پرواز عبارت است از: شهر وست لافایت، ایندیانا، آمریکا ("40°16'49.6" شمال، "86°35'24.8" غرب). شکل ۸- الف نمای کلی مزرعه را نشان می‌دهد.

بنابراین در اکثر موارد مدل‌های بزرگتر برای پایگاه داده‌های بزرگ و پیچیده مفید هستند [۴۷].

## ۳-۴ CenterNet

در آشکارسازها فرآیند تشخیص معمولاً به تعیین محل کادرهای اطراف شی در یک تصویر اطلاق می‌شود. بیشتر آشکارسازهای شی موفق تقریباً یک لیست جامع از مناطق احتمالی و طبقه‌بندی بالقوه آن اشیا را ارائه می‌دهند که در این امر پردازش‌های اضافی مورد نیاز است. در این آشکارساز، یک شی به عنوان تک نقطه مدل می‌شود که این تک نقطه همان مرکز کادر خواهد بود. آشکارساز CenterNet از روش تخمین نقاط کلیدی، برای یافتن نقاط مرکز و اعمال رگرسیون به تمام ویژگی‌های دیگر شی استفاده می‌کند [۴۸]. در این آشکارساز تصویر ورودی با استفاده از یک شبکه تمام کانولوشنی بررسی و از آن یک نقشه حرارتی تشکیل می‌شود که قله‌ی این نقشه تخمین مرکز شی را نشان می‌دهد. برای آموزش شبکه از مدل Hourglass [104] استفاده شده است که در اینجا برای افزایش دقت از دو مدل پشت به پشت استفاده شده است. این مدل یک شبکه متقارن کانولوشنی است که از ۵ لایه بالا و ۵ لایه پایین تشکیل شده که معماری آن در شکل ۶ نشان داده شده است.



شکل ۶: معماری Hourglass [۴۹]

A, B: لایه‌های کانولوشنی، C, D: ماژول‌های دریافتی، E: لایه max pooling، F: لایه کانولوشنی انتقال دهنده، G: ماژول باقیمانده، H: ماژول اتلاف

برچسب زنی آشکارساز CenterNet بسیار ساده و برچسب هر شی بصورت نقطه است. دقت این آشکارساز در تشخیص اشیا کوچک مناسب است. به همین دلیل از این آشکارساز برای تشخیص و شمارش گیاه ذرت در مراحل اولیه رشد قبل از رشد تاسل استفاده شده است.

### ۳-۵ درون‌یابی و پیش‌بینی

برای پیش‌بینی تاریخ گل دهی از مدل رگرسیون خطی استفاده شد. ضرایب  $a$  و  $b$  در معادله خطی  $y(x) = a + bx$  بصورت زیر محاسبه می‌شود.

$$S_{xx} = \sum_{i=0}^N (x_i - \bar{x})^2 \quad (۵)$$

$$S_{xy} = \sum_{i=0}^N (x_i - \bar{x})(y_i - \bar{y}) \quad (۶)$$

$$a = \bar{y} - b\bar{x} \quad (۷)$$

خروجی نرم افزار کلاسها به تفکیک و با یک فایل CSV شامل نام خطا و مختصات  $(x_s, y_s, x_e, y_e)$  در اختیار کاربر قرار می‌دهد. شکل ۹ شماتیک کلی نرم‌افزار Vira Label را نشان می‌دهد.



شکل ۹: شماتیک کلی نرم‌افزار Vira Label

مشخصات داده آموزش برای دو روش CenterNet و YOLOv5 در جدول ۲ نشان داده شده است.

جدول ۲: مشخصات داده آموزش

نام روش	نوع برچسب	تصاویر اندازه/ آموزش (تعداد)	تصاویر اندازه/ تست (تعداد)
CenterNet	نقطه	۱۸۰۰/(۵۱۲×۵۱۲)	۱۵/(۶۲۰×۲۱۰۰)
YOLOv5	چهارضلعی	۴۲۰/(۶۲۰×۲۱۰۰)	۱۵/(۶۲۰×۲۱۰۰)
TasselNet v2+	چهارضلعی	۴۲۰/(۶۲۰×۲۱۰۰)	۱۵/(۶۲۰×۲۱۰۰)

اندازه تصاویر تست  $۶۲۰ \times ۲۱۰۰$  است بنابراین اندازه تصاویر آموزشی برای روش YOLOv5 نیز همین مقدار در نظر گرفته شد اما در روش CenterNet به دلیل محدودیت در اندازه تصویر ورودی، ناحیه آموزش به صورت تصادفی با اندازه  $۵۱۲ \times ۵۱۲$  با همپوشانی برش داده شد.

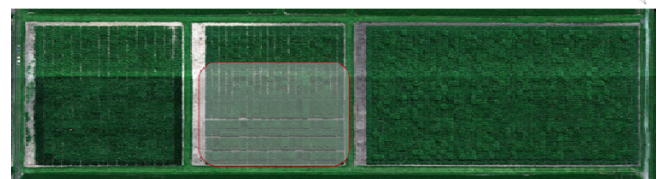
### ۳-۴ تنظیم پارامترهای بهینه

پارامترهای شبکه GP-GAN برای تولید تصاویر با رزولوشن بالا مطابق جدول ۳ تنظیم شد:

جدول ۳: پارامترهای بهینه GP-GAN

پارامتر	GP-GAN
اندازه فیلترهای پایه در انکدر	۶۴
اندازه فیلترهای پایه در دیکدر	۶۴
اندازه فیلترهای پایه در متمایز کننده	۶۴
تعداد کانال‌های خروجی در رمزگشا	۳
گرادیان	نرمال
تعداد تکرار	۱۲۰۰
وزن رنگ	۰٫۱

دهد. ذرت‌ها در تاریخ ۳۰ آوریل سال ۲۰۲۰ کاشته شدند. سه هفته بعد از کاشت گیاهان رشد کردند و دو ماه بعد (جولای) رشد تاسل‌ها بسته به نوع گونه ذرت شروع شد. در تاریخ ۸ جولای تعداد بسیار کمی از ذرت‌ها تاسل داشتند و بعد از این تاریخ تعداد بیشتری از تاسل‌ها شروع به رشد کردند. تصویر برداری با فواصل ۳ الی ۴ روزه انجام شده تا روند رشد تاسل‌ها تا زمان تاسل‌دهی همگی ذرت‌ها (۲۴ جولای) برای ایجاد آرشیو و درون‌یابی ثبت گردد. تصاویر در طول فصل رشد ذرت در سال ۲۰۲۰ در تاریخ‌های ۸، ۱۱، ۱۳، ۱۷، ۲۰ و ۲۴ جولای با استفاده از دوربین رنگی DJI Sony Alpha 7R-III نصب شده بر روی پهپاد Matrice M600 Pro در ارتفاع ۲۰ متری اخذ شدند. این دوربین دارای رزولوشن  $۴۲/۴$  مگاپیکسل است. تصویر برداری از ساعت ۱۰:۰۰ الی ۱۲:۰۰ صبح انجام شد، در این ساعات نور مناسب برای تصویر برداری وجود داشته است. تصاویر خام رنگی<sup>۱</sup> پردازش پردازش شدند و به ارتوفوتو<sup>۲</sup> با رزولوشن مکانی  $۰/۲۵$  سانتی متر تبدیل شدند که این تصاویر بسیار با کیفیت هستند. هر تاریخ شامل ۱۵ تصویر رنگی با اندازه  $۳۶۷۰ \times ۲۱۵۰$  است. اطلاعات فنوتیپی از جمله شمارش و تاریخ گل‌دهی با بازرسی‌های بصری در مزرعه در طول مدت رشد نیز ثبت شده است. بخش آموزش و تست مطابق شکل ۸-ب از تاریخ‌های اولیه رشد ۸، ۱۱، ۱۳ و ۱۷ جولای انتخاب شده است.



(الف)



(ب)

شکل ۸: الف) موقعیت پرواز ب) ناحیه آموزش (قرمز)، ناحیه تست (آبی)

برای ساخت پایگاه داده و برچسب زنی تصاویر از نرم‌افزار تخصصی Vira Label [۵۰] استفاده شد. در این نرم‌افزار می‌توان کادر مورد نظر را به صورت دلخواه رسم نمود و در نهایت

<sup>۱</sup> RGB

<sup>۲</sup> Orthophotos



از پیش آموزش دیده پایگاه داده‌های استاندارد ImageNet [۵۱] و MS-COCO [۵۲] استفاده می‌شود اما این وزن‌ها شباهتی به داده‌های تاسل موجود نداشتند به همین دلیل از Scratch آموزش انجام شده است. در حال حاضر وزن‌های خوبی برای تشخیص تاسل بدست آمده که می‌توان از این وزن‌ها برای تصاویر سال‌های آینده استفاده نمود و شبکه از پیش آموزش دیده شده حاصل می‌تواند در زمان کوتاه‌تری با دقت بالاتری تاسل‌ها را تشخیص دهد.

جدول ۵: مدت زمان آموزش

آشکارساز	مدت زمان آموزش
CenterNet	۷ ساعت و ۱۵ دقیقه
CenterNet+GP	۸ ساعت و ۲۰ دقیقه
YOLOv5	۲۰ دقیقه
YOLOv5+GP	۲۵ دقیقه
TasselNetv2+	۲۲ دقیقه

همانگونه که در جدول ۵ مشاهده می‌گردد بعد از افزایش داده تعداد تاسل‌ها در پایگاه داده افزایش یافته و این امر باعث افزایش مدت زمان آموزش شد.

طبق جدول ۲ تعداد تصاویر تست ۱۵ است که با نماد S1 تا S15 نشان داده شده است. نتایج بدست آمده از آشکارسازهای YOLOv5 و CenterNet قبل و بعد از افزایش داده و روش TasselNetv2+ برای تصاویر ۷ و ۹ که از دو گونه‌ی متفاوت هیبرید<sup>۲</sup> و اینبرد<sup>۳</sup> ذرت می‌باشند در شکل‌های ۱۲ تا ۱۴ نشان داده شده است. طبق نتایج بدست آمده روش پیشنهادی بیشترین تعداد تشخیص صحیح را داشته است. روش TasselNetv2+ نیز در تصاویری که تعداد کمی تاسل وجود داشته عملکرد مناسبی نداشته و دقت آن بشدت افت کرده است.

با استفاده از روش آزمون و خطا پارامترهای بهینه سه الگوریتم برای افزایش دقت در تشخیص تاسل در تاریخ‌های مختلف تعیین و تنظیم شد. مقادیر در جدول ۴ نشان داده شده است.

جدول ۴: پارامترهای بهینه TasselNetv2+, YOLOv5, CenterNet

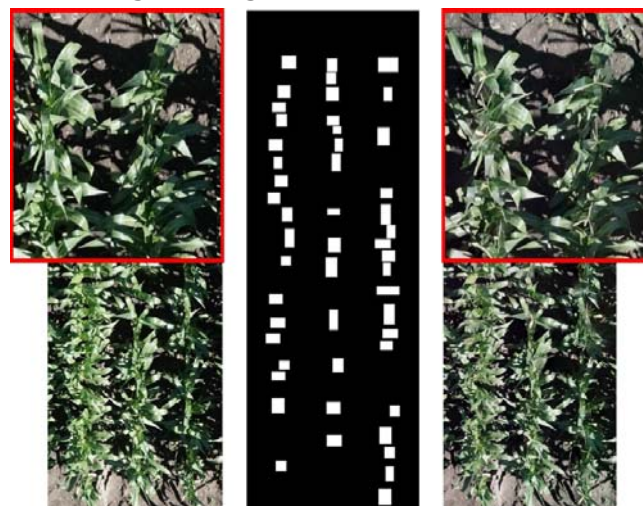
الگوریتم	CenterNet	YOLOv5	TasselNetv2+
مدل	ExtremeNet	YOLOv5s	MTC
شبکه پشتیبان	Hourglass104	CSPNet	CNN
اندازه هر دسته	۱۱	۱۵	۹
مجموع مراحل <sup>۱</sup>	۲۴۰	۱۵۰	۵۰۰
نرخ یادگیری	۰,۰۰۰۱۲۵	۰,۰۰۰۰۰۱	۰,۰۰۰۰۰۱
بهینه ساز	Adam	Adam	SGD
عمق	۱۰۱	۱۰	-

#### ۴-۴ نتایج

به دلیل رشد تعداد کمی از تاسل‌ها در تاریخ ۱۳ جولای، افزایش داده برای تصاویر این تاریخ انجام شد. به این صورت که از تاسل‌های تاریخ ۱۷ ماسک ساخته شد و با استفاده از GP-GAN به تاریخ ۱۳ جولای اضافه شد. نتیجه بدست آمده در شکل‌های ۱۰ و ۱۱ نشان داده شده است.



شکل ۱۰: اضافه کردن تاسل از تاریخ ۱۷ به تاریخ ۱۳



(الف) (ب) (ج)

شکل ۱۱: الف) تصویر ۱۳ جولای ب) تصویر ماسک ۱۷ جولای ج) تصویر اضافه شدن تاسل به ۱۳ جولای

زمان آموزش آشکارسازهای استفاده شده قبل و بعد از افزایش داده در جدول ۵ آورده شده است. برای آموزش معمولاً از وزن‌های

<sup>۲</sup> Hybrid<sup>۳</sup> Inbred<sup>۱</sup> Epoch



$$Pr = \frac{TP}{TP+FP} \quad (9)$$

$$Re = \frac{TP}{TP+FN} \quad (10)$$

$$Sc = 2 \times \frac{Re \times Pr}{Re + Pr} \quad (11)$$

$$MSE = \frac{1}{N} \sum_i (Gt_i - Nt_i)^2 \quad (12)$$

$$MAE = \frac{1}{N} \sum_i |Gt_i - Nt_i| \quad (13)$$

$TP^6$ : تعداد تاسل‌هایی که در تصاویر تست صحیح تشخیص داده است.

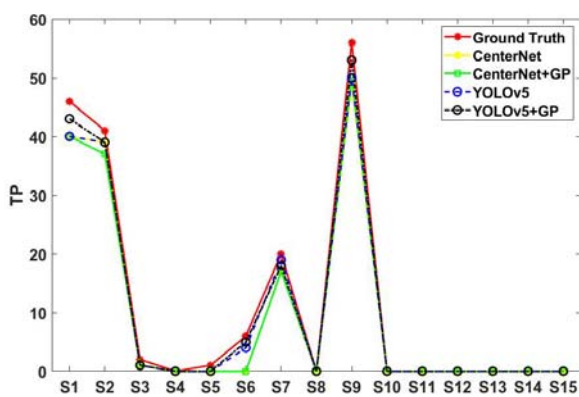
$FN^7$ : تعداد تاسل‌هایی که آشکارساز تشخیص نداده است.

$FP^8$ : تعداد تاسل‌هایی که آشکارساز اشتباه و اضافه تشخیص داده است.

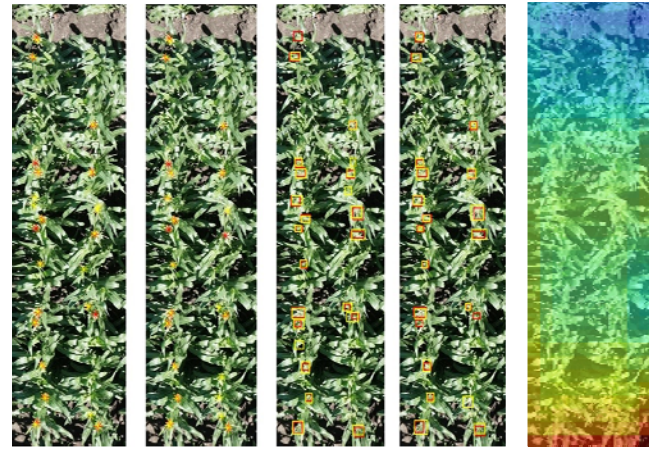
$Gt^9$ : تعداد تاسل گراندتروث در تصاویر استفاده شده برای تست آشکارساز است.

$Nt^{10}$ : تعداد کل تاسل‌هایی که آشکارساز در تصاویر تست تشخیص داده است.

در شکل ۱۵، مقادیر TP مربوط به ۱۵ تصویر S1 تا S15، تاریخ ۱۳ جولای نشان داده شده است. در این تاریخ در اغلب تصاویر تست تاسل رشد نکرده یا تعداد کمی تاسل وجود دارد. در تصاویر S1، S2، S6، S7 و S9 تعداد قابل توجهی تاسل رشد کرده است. عملکرد روش پیشنهادی GP-YOLOv5 در مقایسه با سایر روش‌ها بهتر است و تعداد تشخیص صحیح به تعداد گراندتروث نزدیکتر است.

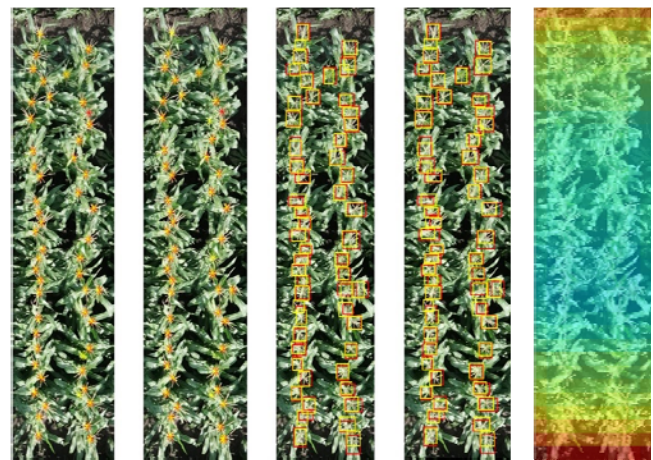


شکل ۱۵: تعداد تشخیص صحیح TP در تاریخ ۱۳ جولای



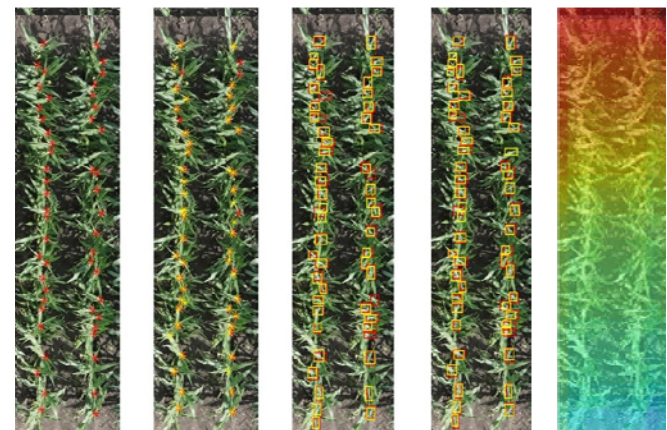
CenterNet CenterNet+GP YOLOv5 YOLOv5+GP TasselNetv2+

شکل ۱۲: نتیجه تصویر ۷ (S7) برای تاریخ ۱۳ جولای



CenterNet CenterNet+GP YOLOv5 YOLOv5+GP TasselNetv2+

شکل ۱۳: نتیجه تصویر ۷ (S7) برای تاریخ ۱۷ جولای



CenterNet CenterNet+GP YOLOv5 YOLOv5+GP TasselNetv2+

شکل ۱۴: نتیجه تصویر ۹ (S9) تاریخ ۱۷ جولای

#### ۴-۵ معیارهای ارزیابی

برای ارزیابی عملکرد آشکارساز و مقایسه روش‌ها از معیارهای دقت<sup>۱</sup>، یادآوری<sup>۲</sup>، نمره ارزیابی<sup>۳</sup>، خطاهای میانگین مربعات<sup>۴</sup> و میانگین مطلق<sup>۵</sup> استفاده می‌شود [۵۳]:

<sup>3</sup> Score (Sc)

<sup>4</sup> Mean Square Error (MSE)

<sup>5</sup> Mean Absolute Error (MAE)

<sup>6</sup> True Positive

<sup>7</sup> False Negative

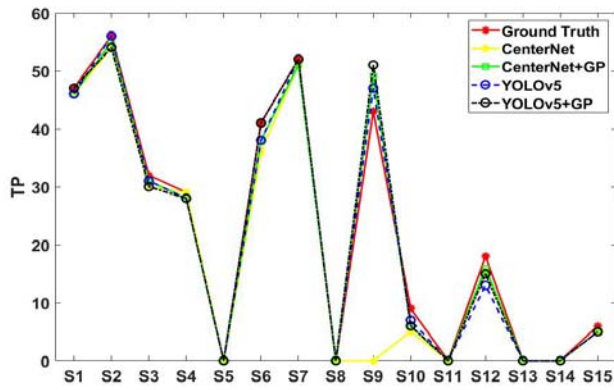
<sup>8</sup> False Positive

<sup>9</sup> Ground Truth

<sup>10</sup> Number Test

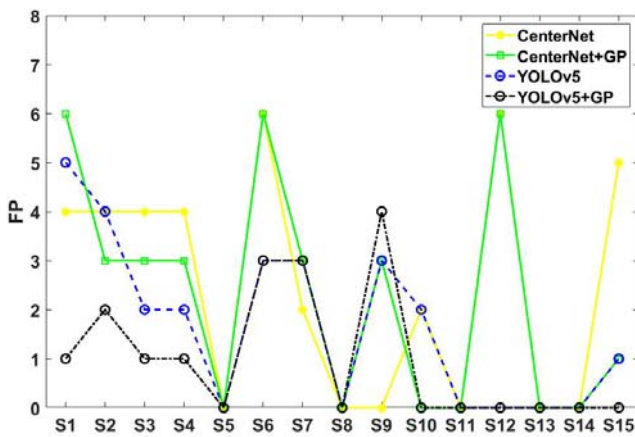
<sup>1</sup> Precision (Pr)

<sup>2</sup> Recall (Re)



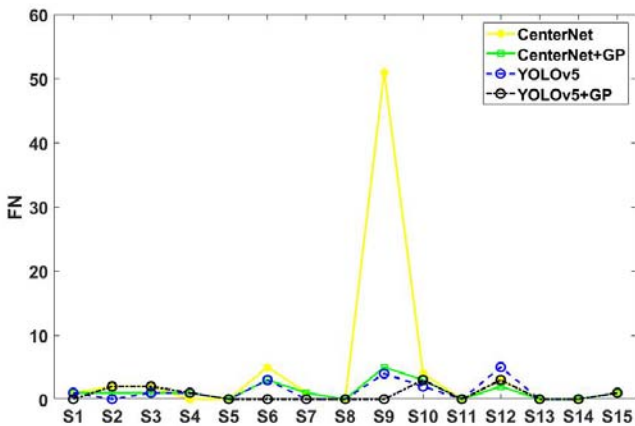
شکل ۱۸: تعداد تشخیص صحیح TP در تاریخ ۱۷ جولای

در شکل ۱۹، مقادیر FP مربوط به ۱۵ تصویر S1 تا S15، تاریخ ۱۷ جولای نشان داده شده است. با توجه به این نمودار، روش پیشنهادی GP-YOLOv5 تعداد تشخیص اضافه در مقایسه با سایر روش‌ها کمتر است و در بعضی موارد مانند تصاویر S5، S8، S10، S11، S12، S13، S14 و S15 تشخیص اضافه یا اشتباه نداشته است.



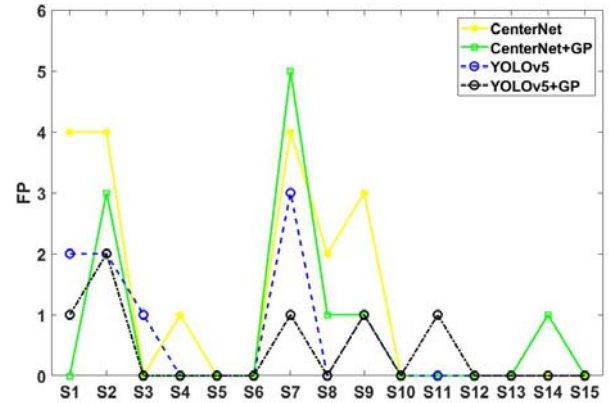
شکل ۱۹: تعداد تشخیص اضافه FP در تاریخ ۱۷ جولای

در شکل ۲۰، مقادیر FN مربوط به ۱۵ تصویر S1 تا S15، تاریخ ۱۷ جولای نشان داده شده است. روشی که تعداد تشخیص نداده کمتری داشته باشد عملکرد بهتری دارد. با توجه به شکل ۲۰ روش پیشنهادی نتیجه بهتری داشته است.



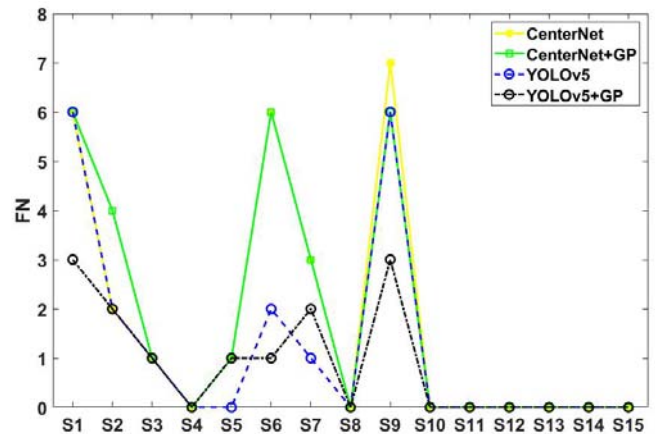
شکل ۲۰: تعداد تشخیص نداده FN در تاریخ ۱۷ جولای

در شکل ۱۶، مقادیر FP مربوط به ۱۵ تصویر S1 تا S15، تاریخ ۱۳ جولای نشان داده شده است. با توجه به این نمودار در اغلب تصاویر تست آشکارساز CenterNet قبل و بعد از داده افزایشی تعداد بیشتری تشخیص اضافه داشته است.



شکل ۱۶: تعداد تشخیص اضافه FP در تاریخ ۱۳ جولای

در شکل ۱۷، مقادیر FN مربوط به ۱۵ تصویر S1 تا S15، تاریخ ۱۳ جولای نشان داده شده است. با توجه به این نمودار، روش پیشنهادی در تشخیص تاسل‌ها موفق‌تر عمل کرده زیرا تعداد تاسل تشخیص داده نشده برای این روش در مقایسه با سایر روش‌ها کمتر است.

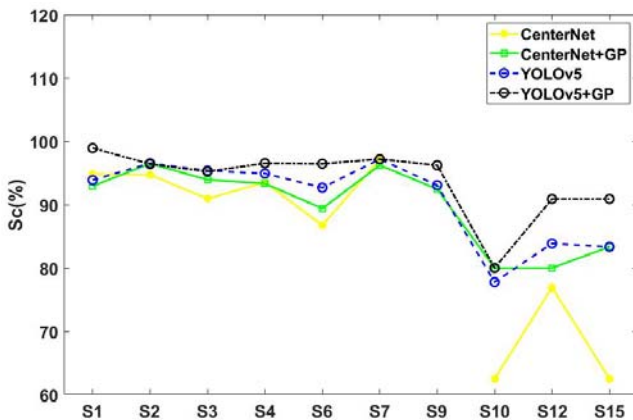


شکل ۱۷: تعداد تشخیص نداده FN در تاریخ ۱۳ جولای

در شکل ۱۸، مقادیر TP مربوط به ۱۵ تصویر S1 تا S15، تاریخ ۱۷ جولای نشان داده شده است. در این تاریخ در اکثر تصاویر تست، تعداد قابل توجهی تاسل رشد کرده است و تقریباً همه روش‌ها عملکرد مناسبی داشته‌اند اما در مجموع تعداد تشخیص صحیح روش پیشنهادی به تعداد گراندتروث نزدیکتر و در برخی تصاویر برابر با گراندتروث است.



شکل ۲۳ مقدار معیار Sc در روش پیشنهادی بهتر از سایر روش‌ها است و در اغلب تصاویر این معیار نزدیک به ۱۰۰ درصد است.



شکل ۲۳: معیار Sc در تاریخ ۱۷ جولای

در روش‌های مبتنی بر شمارش بر خلاف روش‌های مبتنی بر تشخیص امکان محاسبه معیارهای ارزیابی دقت، یادآوری و نمره ارزیابی وجود ندارد به همین دلیل برای ارزیابی، مقدار خطاهای میانگین مربعات و میانگین مطلق برای روش TasselNetv2+، CenterNet و روش پیشنهادی در جدول ۶ محاسبه شد.

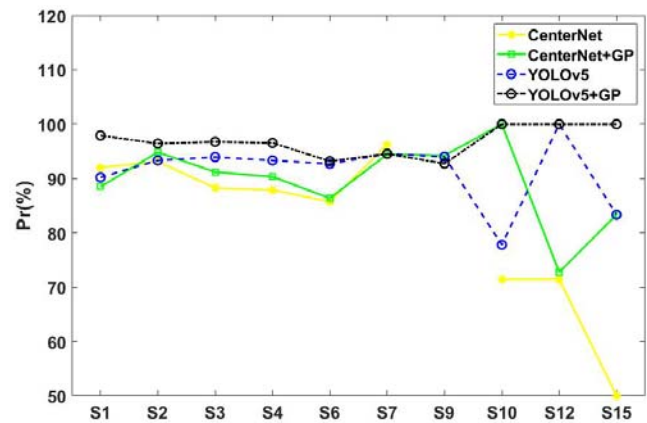
جدول ۶: محاسبه خطای میانگین مربعات و میانگین مطلق

روش	MSE	MAE
CenterNet	۱۶/۷۸	۴/۳۳
GP-CenterNet	۸/۲۶	۲/۱۳
YOLOv5	۴/۹۱	۱/۲۷
GP-YOLOv5	۴/۹۰	۱/۲۶
TasselNetv2+	۱۳/۶۸	۳/۵۳

پایین بودن مقادیر MSE و MAE نشان دهنده عملکرد بهتر در شمارش تاسل است. مطابق نتایج بدست آمده در جدول ۶ روش YOLOv5 و روش پیشنهادی پایین‌ترین مقادیر MSE و MAE را دارند. میزان خطا در روش TasselNetv2+ در مقایسه با GP-CenterNet و روش پیشنهادی بیشتر است. در دو روش YOLOv5 و GP-YOLOv5 تعداد تاسل‌های تشخیص داده شده تقریباً با یکدیگر برابر هستند به همین دلیل خطای میانگین مربعات و میانگین مطلق دو روش نزدیک به هم هستند اما در روش پیشنهادی تعداد تشخیص صحیح بیشتر از روش YOLOv5 است. میزان خطا در روش TasselNetv2+ در مقایسه با GP-CenterNet و روش پیشنهادی بیشتر است.

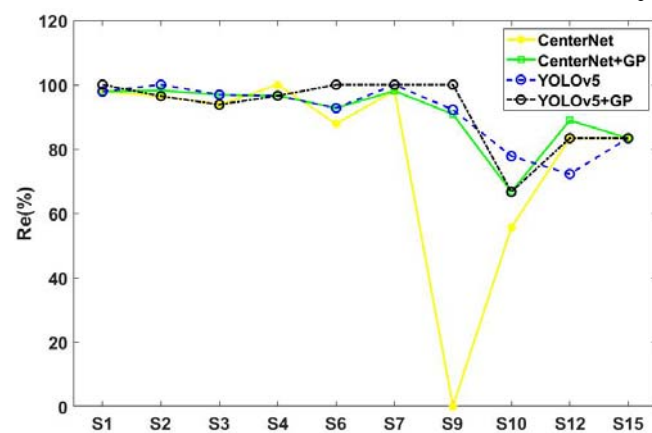
برای تعیین تاریخ گل‌دهی ابتدا تعداد ذرت‌ها در مراحل اولیه کاشت در تاریخ ۳۰ می ۲۰۲۰، قبل از رشد تاسل در قسمت تست با استفاده از CenterNet شمارش شد. نمونه نتایج بدست آمده از آشکارساز CenterNet در شکل ۲۴ نشان داده شده است.

از آنجایی که تعداد زیادی از تصاویر تست تاریخ ۱۳ تاسل وجود نداشت محاسبه معیارهای ارزیابی برای این تاریخ امکان پذیر نیست اما به دلیل رشد کافی تاسل در تاریخ ۱۷ جولای در برخی تصاویر معیارهای ارزیابی برای این تاریخ در شکل‌های ۲۱ تا ۲۳ نشان داده شده است. شکل ۲۱ معیار Pr در تاریخ ۱۷ جولای مربوط به ۱۰ تصویر که در آن‌ها تاسل رشد کرده را نشان می‌دهد. طبق رابطه ۹ مقدار Pr با تعداد تشخیص اضافه و یا اشتباه رابطه عکس دارد و هر چه تعداد FP بیشتر باشد دقت Pr کاهش می‌یابد. آشکارساز CenterNet در تصویر S9 هیچ تشخیص درست و اشتباهی نداشته است به همین دلیل محاسبه این معیار برای S9 امکان پذیر نبوده است.



شکل ۲۱: معیار Pr در تاریخ ۱۷ جولای

شکل ۲۲ معیار Re در تاریخ ۱۷ جولای مربوط به ۱۰ تصویر که در آن‌ها تاسل رشد کرده است را نشان می‌دهد. طبق رابطه ۱۰ مقدار Re با تعداد تشخیص نداده رابطه عکس دارد و هر چه تعداد FN بیشتر باشد دقت معیار Re کاهش می‌یابد. آشکارساز CenterNet در تصویر S9 هیچ کدام از تاسل‌ها را تشخیص نداده است به همین دلیل مقدار این معیار برای S9 صفر درصد بوده است.



شکل ۲۲: معیار Re در تاریخ ۱۷ جولای

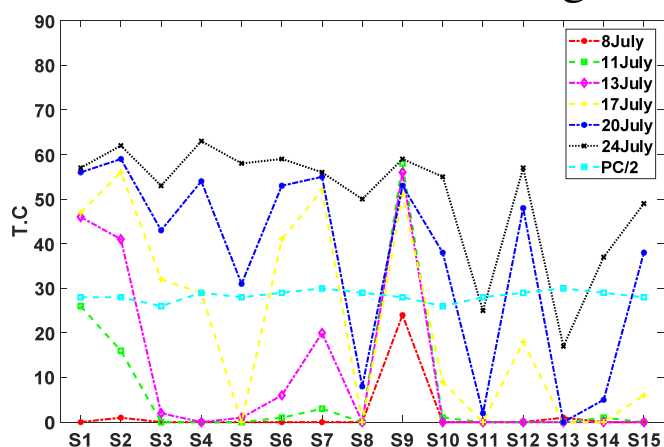
شکل ۲۳ معیار Sc در تاریخ ۱۷ جولای مربوط به ۱۰ تصویر که در آن‌ها تاسل رشد کرده است را نشان می‌دهد. با توجه به

زده شده، بسیار کم است. بنابراین عملکرد روش پیشنهادی قابل قبول بوده است.

جدول ۸: مقایسه آماری نتایج بازدید بصری و تخمین تاریخ گل‌دهی

معیار	میانگین	واریانس	خطای میانگین مربعات	خطای میانگین مطلق
تاریخ حاصل از بازدید بصری (جولای)	۱۸/۱۳	۵/۰۴	۳/۳۶	۰/۸۷
تاریخ تخمین زده شده (جولای)	۱۸/۲	۴/۶۵		

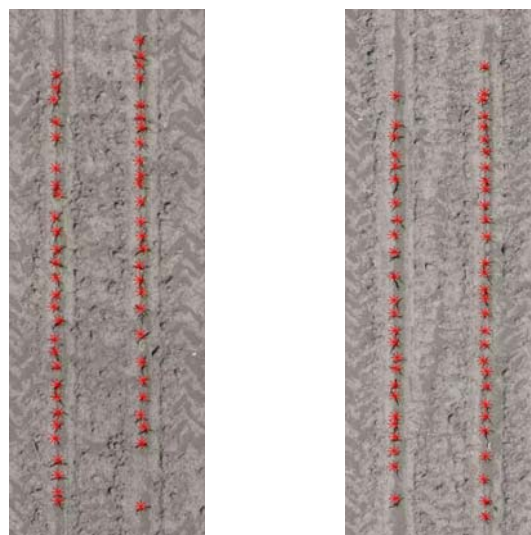
شکل ۲۵ تعداد تاسل‌های موجود در هر تصویر برای تاریخ‌های مختلف و همچنین  $PC/2$  که همان مقدار نصف ذرت‌های شمارش شده در هر تصویر توسط CenterNet است را نشان می‌دهد. هر تاریخی که تعداد تاسل‌ها به تعداد نصف ذرت‌ها نزدیکتر باشد تاریخ گل‌دهی در نظر گرفته شده است.



شکل ۲۵: تعداد تاسل در هر تصویر برای تاریخ‌های مختلف و  $PC/2$

### ۵ نتیجه‌گیری

در این مقاله هدف نهایی تخمین تاریخ گل‌دهی تاسل‌ها در تصاویر رزولوشن بالای اخذ شده توسط پهپاد است. در این راستا در اوایل رشد تاسل‌ها به دلیل کم بودن تعداد آن‌ها از یک شبکه‌ی مولد متخاصم ویژه GP-GAN به علت قابلیت ادغام تصاویر با رزولوشن بالا استفاده شد و تاسل‌ها در تصاویر پهپادی در تاریخی که رشد کافی (۱۷ جولای) داشته‌اند به تصاویر پهپادی از روزهای قبل (۱۳ جولای) که تعداد تاسل‌ها کم هستند اضافه شد. با استفاده از این روش یک پایگاه داده مناسب با تعداد داده آموزشی کافی فراهم می‌شود. در مرحله بعد آشکارساز تک مرحله‌ای YOLOv5 اصلاح شده برای تشخیص تاسل‌ها استفاده شد. نتایج آشکارسازی تاسل‌ها توسط روش پیشنهادی GP-YOLOv5 با دو روش معتبر CenterNet و TasselNetv2+ مقایسه شد که نشان می‌دهد روش CenterNet در تاریخ ۱۳



شکل ۲۴: نمونه نتایج بدست آمده از آشکارساز CenterNet برای شمارش گیاهان ذرت (۳۰ می ۲۰۲۰)

با فرض اینکه هر گیاه یک تاسل داشته باشد، زمانیکه تعداد گل‌ها به نصف بیشتر از تعداد ذرت‌های شمارش شده در CenterNet برسد آن زمان گل‌دهی در نظر گرفته می‌شود. برای تصاویری که در تاریخ‌های ابتدایی تعداد کافی تاسل در آن‌ها وجود داشته است تاریخ گل‌دهی با استفاده از پیش بینی مشخص شد. اما در تعدادی از تصاویر در تاریخ‌های ابتدایی تعداد تاسل‌های رشد کرده صفر است به همین دلیل از تاریخ‌های بعدی ۲۰ و ۲۴ جولای نیز استفاده شد و با استفاده از روش درون‌یابی تاریخ گل‌دهی بدست آمد. مقایسه نتایج بدست آمده از روش درون‌یابی و پیش بینی با گزارش‌های ثبت شده از شمارش روزانه در مزرعه به صورت دستی در جدول ۷ نشان داده شده است.

جدول ۷: نتایج بازدید بصری و تخمین تاریخ گل‌دهی

شماره تصویری	نتیجه بازدید بصری (جولای)	تاریخ تخمین زده شده (جولای)	شماره تصویری	نتیجه بازدید بصری (جولای)	تاریخ تخمین زده شده (جولای)
S1	۱۳	۱۲	S9	۹	۹
S2	۱۳	۱۲	S10	۱۶	۱۹
S3	۱۷	۱۷	S11	۲۶	۲۵
S4	۱۸	۱۷	S12	۲۰	۱۹
S5	۲۰	۲۰	S13	۲۵	۲۷
S6	۱۷	۱۶	S14	۲۲	۲۳
S7	۱۵	۱۴	S15	۲۰	۲۰
S8	۲۲	۲۲			

در جدول ۸ مقایسه آماری نتایج بازدید بصری و تخمین تاریخ گل‌دهی انجام شده است. طبق نتایج بدست آمده تاریخ تخمین زده شده به تاریخ حاصل از بازدید بصری بسیار نزدیک است و خطای تفاوت بین تاریخ حاصل از بازدید بصری و تاریخ تخمین

<sup>1</sup> Plant Counting



- International Conference on Smart Systems and Inventive Technology (ICSSIT)*, 2018, pp. 287-292: IEEE.
- [5] M. Arsenovic, M. Karanovic, S. Sladojevic, A. Anderla, and D. Stefanovic, "Solving current limitations of deep learning based approaches for plant disease detection," *Symmetry*, vol. 11, no. 7, p. 939, 2019.
- [6] S. Herwitz *et al.*, "Imaging from an unmanned aerial vehicle: agricultural surveillance and decision support," *Computers and electronics in agriculture*, vol. 44, no. 1, pp. 49-61, 2004.
- [7] H. Lu, Z. Cao, Y. Xiao, B. Zhuang, and C. Shen, "TasselNet: counting maize tassels in the wild via local counts regression network," *Plant methods*, vol. 13, no. 1, pp. 1-17, 2017.
- [8] S. Lacube *et al.*, "Simulating the effect of flowering time on maize individual leaf area in contrasting environmental scenarios," *Journal of Experimental Botany*, vol. 71, no. 18, pp. 5577-5588, 2020.
- [9] A. Kumar, P. Rajalakshmi, W. Guo, B. B. Naik, B. Marathi, and U. Desai, "Detection and Counting of Tassels for Maize Crop Monitoring using Multispectral Images," in *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*, 2020, pp. 789-793: IEEE.
- [10] T. Malisiewicz, A. Gupta, and A. A. Efros, "Ensemble of exemplar-svms for object detection and beyond," in *2011 International conference on computer vision*, 2011, pp. 89-96: IEEE.
- [11] F. Y. Shih, *Image processing and mathematical morphology: fundamentals and applications*. CRC press, 2009.
- [12] Z. Zou, Z. Shi, Y. Guo, and J. Ye, "Object detection in 20 years: A survey," *arXiv preprint arXiv:1905.05055*, 2019.
- [13] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of Big Data*, vol. 6, no. 1, pp. 1-48, 2019.
- [14] A. Antoniou, A. Storkey, and H. Edwards, "Data augmentation generative adversarial networks," *arXiv preprint arXiv:1711.04340*, 2017.
- [15] D. P. Kingma and M. Welling, "Auto-encoding variational bayes," *arXiv preprint arXiv:1312.6114*, 2013.
- [16] H. Akaike, "Fitting Autoregressive for Prediction Models," *Statist Math*, vol. 21, pp. 243-247, 1969.
- [17] C. Ledig *et al.*, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4681-4690.
- [18] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2223-2232.
- جولای به دقت میانگین تشخیص صحیح تاسل ۹۱,۳۶ درصد و روش پیشنهادی به دقت میانگین ۹۷,۶۲ درصد رسید و در تاریخ ۱۷ جولای نیز دقت میانگین تشخیص صحیح تاسل روش CenterNet ۸۱,۷۸ درصد و روش GP-YOLOv5 ۹۶,۸۱ درصد است. اما روش TasselNetv2+ قادر به تشخیص دقیق موقعیت مکانی و اندازه تاسل نیست. ولی دقت بالایی در شمارش تاسل‌های موجود در تصاویر تست زمانی که تعداد تاسل‌ها زیاد باشد، دارد. چنانچه تعداد تاسل‌ها در تصاویر تست کم باشد دقت این الگوریتم به شدت افت می‌کند. تخمین تاریخ گل‌دهی با استفاده از تکنیک‌های پیش‌بینی و درون‌یابی و با کمک اطلاعات شمارش گیاه در روزهای اولیه کاشت انجام شد. با توجه به گزارش‌های حاصل از بازدید روزانه در مزرعه و خطای میانگین مربعات و مطلق که به ترتیب ۳/۳۶ و ۰/۸۷ محاسبه شدند حاکی از دقت بالای تخمین تاریخ گل‌دهی است.
- با توجه به سخت و زمان‌بر بودن افزایش داده با استفاده از روش پیشنهادی، برای کارهای بعدی می‌توان از تکنیک‌های مبتنی بر روش یادگیری انتقالی و یادگیری با تعداد داده کم<sup>۱</sup> استفاده کرد. همچنین از مدل به دست آمده بر روی تاسل‌ها و اعمال وزن‌های اولیه برای تشخیص تاسل در سال‌های آینده و تشخیص گونه‌های متفاوت برای کاهش زمان آموزش و افزایش دقت تشخیص استفاده نمود.
- ## ۶ سپاسگزاری
- این پژوهش مربوط به مرکز تحقیقات و آموزش زراعت (ACRE) دانشگاه پردو، ایندیانا، آمریکا است. بر خود لازم می‌دانیم از شرکت دانش بنیان فنی مهندسی داده کاوی ویرا ژرف کویر که تصاویر پایگاه داده را تهیه نموده و در به ثمر رسیدن این پژوهش نقش بسزایی داشت تقدیر و تشکر نماییم.
- ## مراجع
- [1] س. ف. موسوی، ا. کرمی، "تشخیص گیاه در تصاویر اخذ شده توسط پهباد با استفاده از شبکه YOLOv3"، چهارمین کنفرانس بین‌المللی علوم کشاورزی، محیط زیست، توسعه شهری و روستایی، ۱۳۹۸.
- [2] N. Falco *et al.*, "Influence of soil heterogeneity on soybean plant development and crop yield evaluated using time-series of UAV and ground-based geophysical imagery," *Scientific reports*, vol. 11, no. 1, pp. 1-17, 2021.
- [3] D. I. Patrício and R. Rieder, "Computer vision and artificial intelligence in precision agriculture for grain crops: A systematic review," *Computers and electronics in agriculture*, vol. 153, pp. 69-81, 2018.
- [4] R. H. Shankar, A. Veeraghavan, K. Sivaraman, and S. S. Ramachandran, "Application of UAV for pest, weeds and disease detection using open computer vision," in *2018*

<sup>1</sup> Few-Shot Learning

- context-augmented local regression networks," *Plant Methods*, vol. 15, no. 1, pp. 1-14, 2019.
- [33] H. Lu and Z. Cao, "Tasselnet2+: A fast implementation for high-throughput plant counting from high-resolution RGB imagery," *Frontiers in plant science*, vol. 11, p. 1929, 2020.
- [34] H. Zou, H. Lu, Y. Li, L. Liu, and Z. Cao, "Maize tassels detection: a benchmark of the state of the art," *Plant Methods*, vol. 16, no. 1, pp. 1-15, 2020.
- [35] S. V. Mirnezami, S. Srinivasan, Y. Zhou, P. S. Schnable, and B. Ganapathysubramanian, "Detection of the Progression of Anthesis in Field-Grown Maize Tassels: A Case Study," *Plant Phenomics*, vol. 2021, 2021.
- [36] H. E. Tonnang, D. Makumbi, and P. Craufurd, "Methodological approach for predicting and mapping the phenological adaptation of tropical maize (*Zea mays* L.) using multi-environment trials," *Plant methods*, vol. 14, no. 1, pp. 1-12, 2018.
- [37] W. Guo, T. Fukatsu, and S. Ninomiya, "Automated characterization of flowering dynamics in rice using field-acquired time-series RGB images," *Plant methods*, vol. 11, no. 1, pp. 1-15, 2015.
- [38] Y. Su *et al.*, "Evaluating maize phenotype dynamics under drought stress using terrestrial lidar," *Plant methods*, vol. 15, no. 1, pp. 1-16, 2019.
- [39] M. Ji, Y. Yang, Y. Zheng, Q. Zhu, M. Huang, and Y. Guo, "In-field automatic detection of maize tassels using computer vision," *Information Processing in Agriculture*, vol. 8, no. 1, pp. 87-95, 2021.
- [40] H. Wu, S. Zheng, J. Zhang, and K. Huang, "Gp-gan: Towards realistic high-resolution image blending," in *Proceedings of the 27th ACM international conference on multimedia*, 2019, pp. 2487-2495.
- [41] J. Redmon and A. Farhadi, "YOLO9000: better, faster, stronger," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 7263-7271.
- [42] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *arXiv preprint arXiv:1804.02767*, 2018.
- [43] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. L. YOLO, "Optimal Speed and Accuracy of Object Detection," *arXiv preprint arXiv:2004.10934*, 2020.
- [44] C.-Y. Wang, H.-Y. M. Liao, Y.-H. Wu, P.-Y. Chen, J.-W. Hsieh, and I.-H. Yeh, "CSPNet: A new backbone that can enhance learning capability of CNN," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, 2020, pp. 390-391.
- [45] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "Yolov4: Optimal speed and accuracy of object detection," *arXiv preprint arXiv:2004.10934*, 2020.
- [46] A. Kuznetsova, T. Maleva, and V. Soloviev, "Detecting apples in orchards using YOLOv3 and YOLOv5 in general
- [19] X. Chen, Y. Duan, R. Houthoof, J. Schulman, I. Sutskever, and P. Abbeel, "Infogan: Interpretable representation learning by information maximizing generative adversarial nets," in *Proceedings of the 30th International Conference on Neural Information Processing Systems*, 2016, pp. 2180-2188.
- [20] Z. Zhang, X. Zhou, S. Zhao, and X. Zhang, "Semantic prior guided face inpainting," in *Proceedings of the ACM Multimedia Asia*, 2019, pp. 1-6.
- [21] Y. Hiasa *et al.*, "Cross-modality image synthesis from unpaired data using CycleGAN," in *International workshop on simulation and synthesis in medical imaging*, 2018, pp. 31-41: Springer.
- [22] T. Kong, F. Sun, H. Liu, Y. Jiang, L. Li, and J. Shi, "Foveabox: Beyond anchor-based object detection," *IEEE Transactions on Image Processing*, vol. 29, pp. 7389-7398, 2020.
- [23] P. Welander, S. Karlsson, and A. Eklund, "Generative adversarial networks for image-to-image translation on multi-contrast MR images-A comparison of CycleGAN and UNIT," *arXiv preprint arXiv:1806.07777*, 2018.
- [24] M. Romero, Y. Interian, T. Solberg, and G. Valdes, "Training Deep Learning models with small datasets," *arXiv preprint arXiv:1912.06761*, 2019.
- [25] L. Liu, M. Muelly, J. Deng, T. Pfister, and L.-J. Li, "Generative modeling for small-data object detection," in *Proceedings of the IEEE International Conference on Computer Vision*, 2019, pp. 6073-6081.
- [26] S. Liu, H. Zhou, C. Li, and S. Wang, "Analysis of Anchor-Based and Anchor-Free Object Detection Methods Based on Deep Learning," in *2020 IEEE International Conference on Mechatronics and Automation (ICMA)*, 2020, pp. 1058-1065: IEEE.
- [27] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," in *Advances in neural information processing systems*, 2015, pp. 91-99.
- [28] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2980-2988.
- [29] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779-788.
- [30] W. Liu *et al.*, "Ssd: Single shot multibox detector," in *European conference on computer vision*, 2016, pp. 21-37: Springer.
- [31] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [32] H. Xiong, Z. Cao, H. Lu, S. Madec, L. Liu, and C. Shen, "TasselNetv2: in-field counting of wheat spikes with



سیده فروه موسوی مدرک کارشناسی خود را در رشته مهندسی اپتیک و لیزر از دانشگاه شهید باهنر کرمان دریافت کرد. در سال ۱۳۹۷ تحصیل در مقطع کارشناسی ارشد در رشته فوتونیک را در همان دانشگاه شروع نمود. علاقه‌مندی‌های علمی ایشان یادگیری عمیق و پردازش تصویر است.



اعظم کرمی در سال ۱۳۹۱ مدرک دکترا را در رشته مهندسی برق - مخابرات سیستم از دانشگاه شیراز دریافت نمود و از همان سال به عنوان استادیار به دانشکده فیزیک دانشگاه شهید باهنر کرمان پیوست. دکتر اعظم کرمی نویسنده و هم‌نویسنده ده‌ها مقاله در ژورنال‌ها و کنفرانس‌های معتبر علمی بوده است و زمینه‌های علاقه‌مندی ایشان شامل پردازش تصویر به‌ویژه تصاویر سنجش از دور و یادگیری عمیق است.

- and close-up images," in *International Symposium on Neural Networks*, 2020, pp. 233-243: Springer.
- [47] A. Malta, M. Mendes, and T. Farinha, "Augmented Reality Maintenance Assistant Using YOLOv5," *Applied Sciences*, vol. 11, no. 11, p. 4758, 2021.
- [48] X. Zhou, D. Wang, and P. Krähenbühl, "Objects as points," *arXiv preprint arXiv:1904.07850*, 2019.
- [49] Y. Liu, D. Minh Nguyen, N. Deligiannis, W. Ding, and A. Munteanu, "Hourglass-shapenetwork based semantic segmentation for high resolution aerial imagery," *Remote Sensing*, vol. 9, no. 6, p. 522, 2017.
- [50] Available: [www.virakavir.com](http://www.virakavir.com)
- [51] O. Russakovsky *et al.*, "Imagenet large scale visual recognition challenge," *International journal of computer vision*, vol. 115, no. 3, pp. 211-252, 2015.
- [52] T.-Y. Lin *et al.*, "Microsoft coco: Common objects in context," in *European conference on computer vision*, 2014, pp. 740-755: Springer.
- [53] D. M. Powers, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation," *arXiv preprint arXiv:2010.16061*, 2020.