



Semnan University



Research Article

Accurate and real-time detection of bucket teeth in copper mine shovels based on improved yolo model

Mohaddeseh Ghiasi^a, Masoudreza Aghabozorgi^{b,*}

^a Ph.D. Student, Faculty of Electrical Engineering, Yazd University, Yazd, Iran

^b Associate Professor, Faculty of Electrical Engineering, Yazd University, Yazd, Iran

PAPER INFO

Paper history:

Received: 2022-03-13

Revised: 2022-05-12

Accepted: 2022-10-19

Keywords:

Yolo;

Object detection;

Shovel;

Bucket tooth.

ABSTRACT

Shovel is a type of mechanical excavator set that is used in open pit mines. The shovel bucket has a number of teeth, which increase the efficiency of the bucket. Prolonged direct impact of bucket teeth on ore during loading causes unexpected teeth breakage. One of the factors that stop the crusher is the separation of this tooth from the shovel bucket and loading and transfer to the crusher due to the lack of sufficient visibility of the operator on the teeth. The entry of this teeth into the crusher causes the crusher to jam and stop the production cycle. Therefore, it is necessary to propose a bucket shovel teeth detection algorithm with high accuracy and in real time. To solve this problem, we improved the accuracy of the model by making changes in the basic yolov5 model structure. The proposed method was evaluated on a new data set from shovel under real working conditions. The results obtained with an average accuracy of 93.5% and a complexity of 16.1 indicate the improvement of the detection accuracy and the reduction of the complexity in the detection process, which meets the requirements of accurate and real-time detection of bucket shovel teeth.

DOI: <https://doi.org/>

© 2024 Published by Semnan University Press.

This is an open access article under the CC-BY 4.0 license. (<https://creativecommons.org/licenses/by/4.0/>)

* Corresponding author.

E-mail address: aghabozorgi@yazd.ac.ir

How to cite this article:

تشخیص دقیق و بلادرنگ دندانهای باکت در شاولهای معدن مس بر اساس مدل بهبود یافته

YOLO

محدثه قیاسی^۱، مسعودرضا آقابزرگی^{۲*}

اطلاعات مقاله	چکیده
دریافت مقاله: ۱۳۹۵/۰۳/۲۰	شاول نوعی از مجموعه بیل‌های مکانیکی است که در معادن روباز استفاده می‌شود. باکت شاول دارای تعدادی دندان می‌باشد، که باعث افزایش بازدهی باکت می‌شود. تاثیر مستقیم طولانی مدت دندانهای باکت بر روی سنگ معدن در حین بارگیری باعث شکستگی غیرمنتظره دندانها می‌شود. یکی از عوامل توقف سنگ‌شکن، جدا شدن این دندان از باکت شاول و بارگیری و انتقال به سنگ‌شکن به دلیل عدم دید کافی متصدی روی دندانها می‌باشد. ورود این دندان به سنگ‌شکن باعث گیر کردن سنگ‌شکن و توقف چرخه تولید می‌شود. بنابراین، لازم است یک الگوریتم تشخیص دندانهای باکت شاول با دقت بالا و در زمان واقعی پیشنهاد شود. برای حل این مسئله با ایجاد تغییراتی در ساختار مدل پایه YOLOv5، دقت مدل را بهبود بخشیدیم. روش پیشنهادی بر روی مجموعه داده‌ای جدید از شاول تحت شرایط کاری واقعی، ارزیابی شد. نتایج به دست آمده با میانگین دقت ۹۳.۵ درصد و پیچیدگی ۱۶.۱ بیانگر بهبود دقت تشخیص و کاهش پیچیدگی در فرآیند تشخیص انجام شده می‌باشد که الزامات تشخیص دقیق و بلادرنگ دندانهای باکت شاول را برآورده می‌کند.
بازنگری مقاله: ۱۳۹۵/۰۶/۲۰	
پذیرش مقاله: ۱۳۹۵/۰۷/۱۵	
واژگان کلیدی: Yolo، تشخیص اشیا، شاول، دندان باکت.	

DOI: <https://doi.org/>

© 2024 Published by Semnan University Press.

This is an open access article under the CC-BY 4.0 license. (<https://creativecommons.org/licenses/by/4.0/>)

۱-مقدمه

امروزه معدن کاری روباز شامل روش‌های حمل و نقل و استخراج متفاوتی در طیف سیستم ناپیوسته، نیمه پیوسته و پیوسته است [۱]. تمام خطوط تولید در معادن، وابسته به خردایش اولیه یا همان سنگ‌شکن هستند. در فرآیند سنگ‌شکنی بحث توقف‌ها و تعمیرات، همچنین نحوه ارسال

بار از معدن مقوله‌های مهمی می‌باشد. در تمام معادن سعی بر کاهش زمان توقف‌های عملیاتی، تعمیراتی و باردهی متناسب با ظرفیت سنگ‌شکن است. در فرآیند تولید، شاول نیز نقش به‌سزایی در کاهش توقف‌ها و افزایش تولید سنگ‌شکن تا ظرفیت نامی آن دارد. شاول کابلی یک تجهیز

* پست الکترونیک نویسنده مسئول: aghabozorgi@yazd.ac.ir

۱. دانشجوی دکتری، دانشکده مهندسی برق، دانشگاه یزد

۲. دانشیار، دانشکده مهندسی برق، دانشگاه یزد

دسته خاص مانند حیوانات، انسان‌ها یا ماشین‌ها در تصاویر دیجیتال می‌شود [۵]. تاکنون الگوریتم‌های مختلفی جهت تشخیص دندان‌ها باکت مطرح شده‌اند. به عنوان نمونه، آقای لو و همکارش برای تشخیص افتادن دندان‌ها از یک فاصله‌یاب لیزری برای اسکن و بازسازی مدل سه‌بعدی دندان‌های باکت در حال کار استفاده کرده و سپس آن را با مدل اصلی مقایسه می‌نمایند تا تشخیص دهند دندان‌ها باکت در حال افتادن است [۶]. آقای هی و همکارانش یک الگوریتم تشخیصی شی با ترکیبی از ویژگی‌های ظاهری، ساختاری و شکل پیشنهاد داده‌اند [۷].

آقای دووان و همکارانش یک الگوریتم بینایی ماشین برای تشخیص افتادن دندان‌های باکت شاول الکتریکی را پیشنهاد کرده‌اند، که در آن تصاویر به وسیله یک تصویرگر حرارتی مادون قرمز جمع آوری می‌شود. در روش مذکور یک حد آستانه تطبیقی برای اینکه تشخیص افتادن دندان‌های باکت با دقت بالایی انجام شود، استفاده می‌شود [۸]. این الگوریتم‌ها اغلب بر اساس اطلاعات لایه سطحی تصاویر هستند و برای به دست آوردن ویژگی‌های بهینه نیاز به برورسانی مداوم پارامترها و تطبیق الگو دارند. بنابراین از آنجایی که تشخیص دندان‌های باکت شاول در مقایسه با سایر تصاویر چالش‌های بیشتری دارد، روش‌های سنتی تشخیص اشیاء به سختی الزامات این زمینه را برآورده می‌کنند.

با توسعه مستمر یادگیری عمیق، پیشرفت‌هایی در زمینه بینایی کامپیوتر ایجاد شده و شبکه عصبی کانولوشن به ابزار مهمی برای پردازش تصویر تبدیل شده است و در زمینه تشخیص هدف مورد استفاده قرار گرفته است. یادگیری عمیق زیر شاخه‌ای از هوش مصنوعی است که مهندسی ویژگی و طبقه‌بندی را در یک روش ترکیب می‌کند. این یک تکنیک داده محور است که یک مدل پیش بینی را از طریق یادگیری از یک مجموعه داده بزرگ بهینه می‌کند. دیجیتالی شدن در صنعت شامل کسب و ذخیره انواع مجموعه داده‌های بزرگ برای تفسیر و تصمیم‌گیری است. این امر منجر به پذیرش یادگیری عمیق در صنایع مختلف مانند حمل و نقل، تولید، پزشکی و کشاورزی شده است. با این حال، در صنعت معدن، پذیرش و توسعه فناوری‌های جدید، از جمله روش‌های یادگیری عمیق، مانند سایر صنایع پیشرفت نکرده است. با این وجود، در ۵ سال گذشته، کاربردهای یادگیری عمیق در فضای تحقیقات معدنی

مهم برای معدن‌کاری نیمه‌پایسته است که در معدن‌کاری روباز به طور وسیعی استفاده می‌شود [۲]. به دلیل این‌که شاول‌های کابلی سازگاری زیادی با شرایط محیطی و همچنین کاهش میزان مواد منفجره جهت استخراج مواد دارند، لذا استفاده از این دستگاه علاوه بر استخراج مواد معدنی جهت باطله‌برداری همراه با تراک‌های معدنی کاربرد زیادی دارد [۳].

شاول‌های الکتریکی به دلیل راندمان تولید بالا، نرخ عملیاتی بالا و هزینه عملیاتی پایین یکی از تجهیزات ضروری معادن روباز می‌باشند. باکت شاول بسته به نوع آن هیدرولیک یا کابلی دارای تعدادی دندان‌ها می‌باشد که دندان‌ها باعث افزایش بازدهی باکت شاول و کندن مواد انفجار شده از بلوک می‌شود. ضربه مستقیم طولانی مدت دندان‌های باکت بر روی سنگ معدن در حین حفاری می‌تواند باعث شود که دندان‌ها در معرض ضربه، اصطکاک و نیروی خمشی قوی قرار گیرند که باعث شکستن غیرمنتظره دندان‌ها می‌شود و در نتیجه منجر به زیان‌های اقتصادی و اتلاف منابع انسانی و مادی می‌شود. از آنجایی که دندان‌ها از فولاد منگنز با سختی بسیار بالاتر از سنگ معدن و سایر مواد است، هنگامی که دندان‌های شکسته همراه با سنگ معدن داخل سنگ شکن تخلیه می‌شوند، آسیب‌های جدی به چرخ دنده‌های سنگ شکن و سایر تجهیزات وارد شده و باعث خرابی سنگ شکن و توقف آن شده و کل خط تولید سنگ شکن معدن را تحت تاثیر قرار داده و ضرر اقتصادی فراوانی را به همراه دارد. در ضمن فرآیند تخلیه مواد از داخل سنگ‌شکن و بیرون آوردن دندان‌ها به دلیل عدم فضای کافی، خطرناک و وقت گیر بوده و باعث کاهش تولید می‌گردد [۴]. بنابراین اگر دندان‌های جدا شده باکت شاول به موقع و دقیق تشخیص داده نشود خطرات ایمنی جدی را در پی خواهد داشت. لذا ضمانت عدم انسداد سنگ‌شکن را می‌توان با تشخیص دقیق دندان‌های جدا شده قبل از ورود به سنگ‌شکن از طریق سیستم‌هایی همچون تشخیص دندان‌های باکت شاول، انجام داد. برای جلوگیری از وقوع چنین اتفاقی، می‌توان با نصب دوربینی بر روی بازوهای شاول جدا شدن دندان‌ها را تشخیص داده و با دادن هشدار به متصدی شاول فرایند بارگیری را متوقف کرده و با بیرون کشیدن دندان‌ها جلوی توقف احتمالی سنگ‌شکن را گرفت. تشخیص اشیاء یک کار مهم در بینایی کامپیوتری است که شامل شناسایی نمونه‌هایی از اشیاء بصری متعلق به یک

مختصات کادرها و طبقه‌بندی در یک مرحله انجام می‌شود. به همین دلیل سرعت بالاتری نسبت به آشکارسازهای دو مرحله‌ای دارد. [۱۴] YOLO، [۱۵] RetinaNet، [۱۶] CnterNet و [۱۷] از جمله آشکارسازهای تک مرحله‌ای می‌باشند. یکی از کارهای پیشگام در آشکارسازهای شی یک مرحله‌ای، YOLO می‌باشد.

YOLO از یک CNN^۵ استفاده می‌کند تا به طور همزمان چندین جعبه محدود کننده و دسته را برای آن‌ها در کل تصویر با تقسیم تصویر به یک شبکه $S \times S$ پیش‌بینی کند و هر سلول شبکه جعبه‌های محدود و امتیازات اطمینان اشیا را پیش‌بینی کند. نویسندگان YOLO مدل خود را با نسخه‌های جدید بهبود بخشیدند، در YOLO9000 [۱۸] آموزش چند مقیاسی و در YOLOv3 [۱۹] از ستون فقرات بهبود یافته‌ای به نام Darknet53 استفاده شد و تشخیص چند مقیاسی برای تشخیص بهتر اجسام در اندازه‌های مختلف معرفی شد. در حالی که YOLOv3 آخرین نسخه رسمی YOLO بود که توسط نویسندگان اصلی YOLO توسعه و معرفی شد، محققان دیگر به تحقیق در مورد این مدل تشخیص اشیا ادامه دادند و مدل‌های جدید بهبود یافته مبتنی بر YOLO را معرفی کردند. به عنوان مثال، در سال ۲۰۲۰ با معرفی YOLOv4 [۲۰] که شامل چندین پیشرفت جدید مانند اتصالات جزئی متقاطع، فعال‌سازی اشتباه،^۵ CIoU و روش‌های جدید تقویت تصویر بود، مدل YOLOv3 بهبود یافت.

یکی از پیشرفت‌های عمده در آشکارسازهای شی YOLO، YOLOv5 [۲۱] است که ساختار نسبتاً مشابهی با YOLOv4 دارد، اما چندین پیشرفت از جمله تولید جعبه‌های لنگر دقیق‌تر با استفاده از الگوریتم ژنتیک و پیاده‌سازی را با استفاده از زبان برنامه‌نویسی پایتون و چارچوب pytorch [۲۲] انجام می‌دهد که آن را نسبت به سایر مدل‌های مبتنی بر YOLO کمتر پیچیده و سریع‌تر می‌کند و آن را به گزینه‌ای خوب برای کارهای تشخیص اشیا در زمان واقعی تبدیل می‌کند. YOLOv5 یک پروژه در حال انجام و محبوب است و از زمان انتشار اولیه با معرفی ماژول‌های C3 و SPPF^۶ و عملکرد فعال‌سازی SiLU^۷ [۲۳] توسط نویسندگان بهبود بیشتری یافته است.

افزایش یافته است. یادگیری عمیق برای حل انواع مشکلات مربوط به فرآیندهای اکتشاف معدن، استخراج سنگ معدن و فلزات و فرآیندهای احیا اجرا شده است. یادگیری عمیق یک روش مبتنی بر داده است که در آن مدل تولید شده برای مجموعه داده خاص مورد استفاده، در طول توسعه بهینه می‌شود. یک مدل باید در محیط عملیاتی نسبتاً خوب عمل کند. آشکارسازهای تشخیص اشیا به دو دسته تک مرحله‌ای و دو مرحله‌ای تقسیم می‌شوند. در روش دو مرحله‌ای بررسی منطقه‌ای انجام می‌شود و مختصات کادرهایی که احتمال وجود شی را دارد استخراج می‌کند و سپس به طبقه‌بندی اشیا درون کادر می‌پردازد. از جمله آشکارسازهای دو مرحله‌ای می‌توان به Faster R-CNN [۹] اشاره کرد. با وجود پیشرفت در آشکارسازهای شی دو مرحله‌ای، سرعت آن‌ها هنوز توسط فرآیند چند مرحله‌ای آن‌ها محدود می‌باشد [۱۰]. در مرجع [۱۱] یک روش تشخیص شی مبتنی بر یادگیری عمیق برای تشخیص دندان‌های باکت در شاول پیشنهاد شده است. در مرجع [۱۲] الگوریتم تشخیص دندان‌های باکت شاول با استفاده از یک مجموعه داده بر اساس تصاویر دندان‌های باکت شاول واقعی و مدل‌های پرینت سه بعدی با استفاده از آشکارساز Faster R-CNN انجام شده است. در مرجع [۱۳] یک روش تشخیص شی برجسته مبتنی بر یادگیری عمیق برای حذف تصاویر بی‌فایده که باکت شاول در آن مشاهده نمی‌شود، پیشنهاد شده است تا فشار محاسباتی مراحل بعد را کم‌تر کند. سپس با استفاده از بخش‌بندی معنایی موقعیت نسبی دندان‌های باکت تشخیص داده می‌شود. اکثر این روش‌ها در محیط آزمایشگاهی تا حدی موثر بوده و می‌توانند دندان‌های باکت شاول را شناسایی کنند، اما راه‌حل‌های پیشنهادی در محیطی به سختی سطح معدن هشدارهای کاذب زیادی را تولید می‌کنند. نرخ مثبت کاذب بالا و تعداد بالای هشدارهای کاذب باعث می‌شود متصدی‌های تجهیزات به طور کامل هشدارها را نادیده بگیرند. بنابراین، لازم است یک الگوریتم تشخیص دندان‌های باکت شاول با دقت بالا و در زمان واقعی پیشنهاد شود.

در روش‌های تک مرحله‌ای آشکارساز در مکان‌ها و اندازه‌های مشخص دنبال اجسام می‌گردد و استخراج

⁵ Complete IoU

⁶ Spatial Pyramid Pooling Fast

⁷ Sigmoid Linear Unit

² Single Shot Detector

³ You Only Look Once

⁴ Convolutional Neural Network

علاوه بر این، در [۳۰] یک مدل بهبود یافته YOLOv5 برای تشخیص نقص اپیدرم مرکبات معرفی شد. ابتدا یک سیستم تصویربرداری با لامپ دوگانه برای ثبت تصاویر مرکبات با نقص نامرئی طراحی و مورد استفاده قرار گرفت. سپس مدل YOLOv5 با ادغام سازوکار توجه CBAM و اصلاح تابع تلفات به صورت DIoU بهینه شد و در نهایت اثربخشی مدل بهبودیافته از طریق آزمایش‌های مقایسه مورد ارزیابی قرار گرفت.

در حالی که عملکرد مدل های YOLO در طول زمان بهبود یافته است، شایان ذکر است که آنها اغلب به جای تمرکز بر دقت، سرعت و دقت را در اولویت قرار می دهند. این مبادله ضروری است و امکان تشخیص اشیاء به صورت بلادرنگ در اپلیکیشن های مختلف را فراهم می کند. بنابراین از آنجایی که بسیاری از برنامه‌های کاربردی در دنیای واقعی وجود دارد که نیاز به تشخیص شی در زمان واقعی دارد، ایجاد تعادل بین سرعت و دقت چالش برانگیز است.

برای مقابله با این چالش، کار حاضر بر روی یکی از برجسته‌ترین اعضای آشکارساز شی YOLO [۲۱] تمرکز کرده است. به منظور دستیابی به تشخیص واقعی و دقیق دندان‌های باکت شاول با اعمال تغییراتی بر روی آشکارساز شی YOLOv5 قابلیت‌های افزایش دقت تشخیص آن بهبود می‌یابد. نوآوری‌های اصلی این مقاله به شرح زیر است:

- بهبود ساختاری در واحدهای مدل پایه YOLOv5
- اضافه کردن سازوکار توجه
- استفاده از تابع زیان رگرسیون جعبه مرزی تطبیقی
- استفاده از Ghost Convolution برای کاهش تعداد پارامترها، FLOPS و اندازه مدل

برای تشخیص دندان‌های باکت مجموعه داده شاول در معدن مس تحت شرایط کاری واقعی برای آموزش مدل‌های یادگیری عمیق جمع‌آوری شد. مقایسه چندین مجموعه آزمایش و ارزیابی نتایج مدل نشان داد که مدل پیشنهادی بر روی مجموعه داده مورد نظر عملکرد بهتری از مدل پایه دارد و میانگین دقت متوسط الگوریتم پیشنهادی می‌تواند به ۹۳ درصد برسد.

چندین مدل غیر رسمی دیگر مبتنی بر YOLO در سال‌های ۲۰۲۲ و ۲۰۲۳ توسط نویسندگان مختلف معرفی شدند. مدل YOLOv6 [۲۴] شامل یک شبکه ستون فقرات جدید به نام EfficientRep و یک شبکه گردن جدید به نام Rep-PAN است. همچنین سرهای محلی‌سازی و طبقه‌بندی جعبه جدا شده‌اند. چند ماه بعد، وانگ و همکاران مدل YOLOv7 را معرفی کردند [۲۵]. این مدل یک ستون فقرات جدید به نام ELAN^۸ معرفی می‌کند و سرهای تشخیص کمکی را به شبکه اضافه می‌کند تا دقت را افزایش دهد و آن را به طور قابل توجهی متفاوت از سایر مدل های مبتنی بر YOLO می‌کند. علاوه بر این، اولین نسخه YOLOv8 در ژانویه ۲۰۲۳ توسط نویسندگان YOLOv5 منتشر شد. این مدل محلی‌سازی شی بدون لنگر، بلوک‌های C2F و تکنیک‌های افزایش تصویر آنلاین را معرفی کرد. با این حال، YOLOv8 هنوز در حال توسعه فعال است [۲۶].

در مقایسه با مدل‌های قبلی مانند YOLOv5 که در حال حاضر به طور گسترده مورد استفاده قرار می‌گیرند، مدل‌های جایگزین جدید YOLO ذکر شده به ندرت در عمل استفاده می‌شوند و نیاز به بهبود و ارزیابی مداوم برای کاربردهای دنیای واقعی دارند [۲۷]. در نتیجه، بسیاری از محققان در سال‌های اخیر مدل‌های جدید بهبود یافته مبتنی بر YOLOv3، YOLOv4، YOLOv5 را معرفی کرده‌اند.

در [۲۸] یک مدل تشخیص عیب آج چرخ ریلی بر اساس YOLOv3 بهبود یافته معرفی شد. در این روش، مدل YOLOv3 با ادغام نقشه‌های ویژگی کم‌عمق‌تر برای بهبود دقت تشخیص اهداف کوچک و متوسط و بهبود عملکرد تلفات، بهبود یافت.

در [۲۹] یک مدل بهبودیافته YOLOv4-miny برای تشخیص درخت کاج پیشنهاد شد. این مدل شبکه پیش‌فرض ستون فقرات را با ShuffleNet سبک و پیشرفته جایگزین کرد، از یک شبکه هرمی ویژگی فشرده و تحریک برای ترکیب اطلاعات چند مقیاسی استفاده کرد و از تلفات کانونی استفاده کرد. این پیشرفت‌ها این مدل را دقیق‌تر از مدل‌های پایه مقایسه می‌کند.

⁸ Extended Efficient Layer Aggregation

YOLOv5 در بلوک دیاگرام شکل ۱ نشان داده شده است. این مدل تشخیص شی شامل سه واحد اصلی ستون فقرات، گردن و سر است که به ترتیب توضیح داده خواهد شد.

۲-۱- ستون فقرات

ستون فقرات وظیفه استخراج ویژگی را بر عهده دارد. در ابتدا YOLOv5، ورودی را با مقیاس‌بندی و برش تصادفی تصاویر برای تعمیم بهتر، افزایش می‌دهد. سپس ورودی از میان لایه‌های پیچشی و C3 عبور می‌کند. ماژول‌های پیچشی از لایه کانولوشن، لایه نرمال‌سازی دسته‌ای و تابع فعال‌سازی واحد خطی سیگموئید تشکیل شده است. همچنین هر ماژول C3 شامل سه لایه کانولوشن و دنباله‌ای از چندین ماژول تنگنا است که تعداد پارامترها و محاسبات را کاهش می‌دهد.

در انتهای واحد ستون فقرات، ماژول SPPF وجود دارد، که ویژگی‌های رزولوشن‌های مختلف را ترکیب می‌کند. این ماژول شامل سه لایه حداکثر ادغام و لایه کانولوشن است که فرآیند استخراج ویژگی را تکمیل می‌کند و اطلاعات بیشتری را جمع‌آوری می‌کند.

۲-۲- گردن

بین واحدهای ستون فقرات و سر یک رابط به نام گردن وجود دارد. این واحد ویژگی‌های استخراج شده توسط ستون فقرات را از مراحل مختلف جمع می‌کند. در مراحل اولیه ستون فقرات، نقشه‌های ویژگی با وضوح بالا وجود دارد، اما اطلاعات معنایی ضعیفی دارند. با این حال، در مراحل بعدی، نقشه‌های ویژگی دارای اطلاعات معنایی غنی اما اطلاعات مکانی کم هستند. شبکه تجمع مسیر^۹ به منظور دستیابی به نقشه‌های ویژگی با اطلاعات معنایی و مکانی قوی، نقشه‌های ویژگی مراحل مختلف را در دو جهت (پایین به بالا و بالا به پایین) به هم متصل می‌کند.

۲-۳- سر

در نهایت، واحد سر نقشه‌های ویژگی را در سه مقیاس مختلف پیش‌بینی می‌کند که به طور پیش‌فرض 20×20 ، 40×40 و 80×80 هستند. تشخیص اشیا در مقیاس چندگانه برای تشخیص اجسام کوچک، متوسط و بزرگ در تصاویر کمک می‌کند. در نهایت، دقیق‌ترین جعبه مرزبندی

در ادامه، در بخش دوم ساختار کلی شبکه YOLOv5 توضیح داده شده است. سپس در بخش سوم، جزئیات شبکه پیشنهادی بر اساس معماری YOLOv5 و تغییرات پیشنهادی انجام شده در مدل پایه به منظور بهبود عملکرد مدل در تشخیص دندان‌های باکت شاول که کوچک هستند، آورده شده است و پس از آن معیارهای ارزیابی و تحلیل نتایج در بخش چهارم ذکر شده است و در نهایت، بخش آخر به نتیجه‌گیری و پیشنهادهای آینده اختصاص داده شده است.

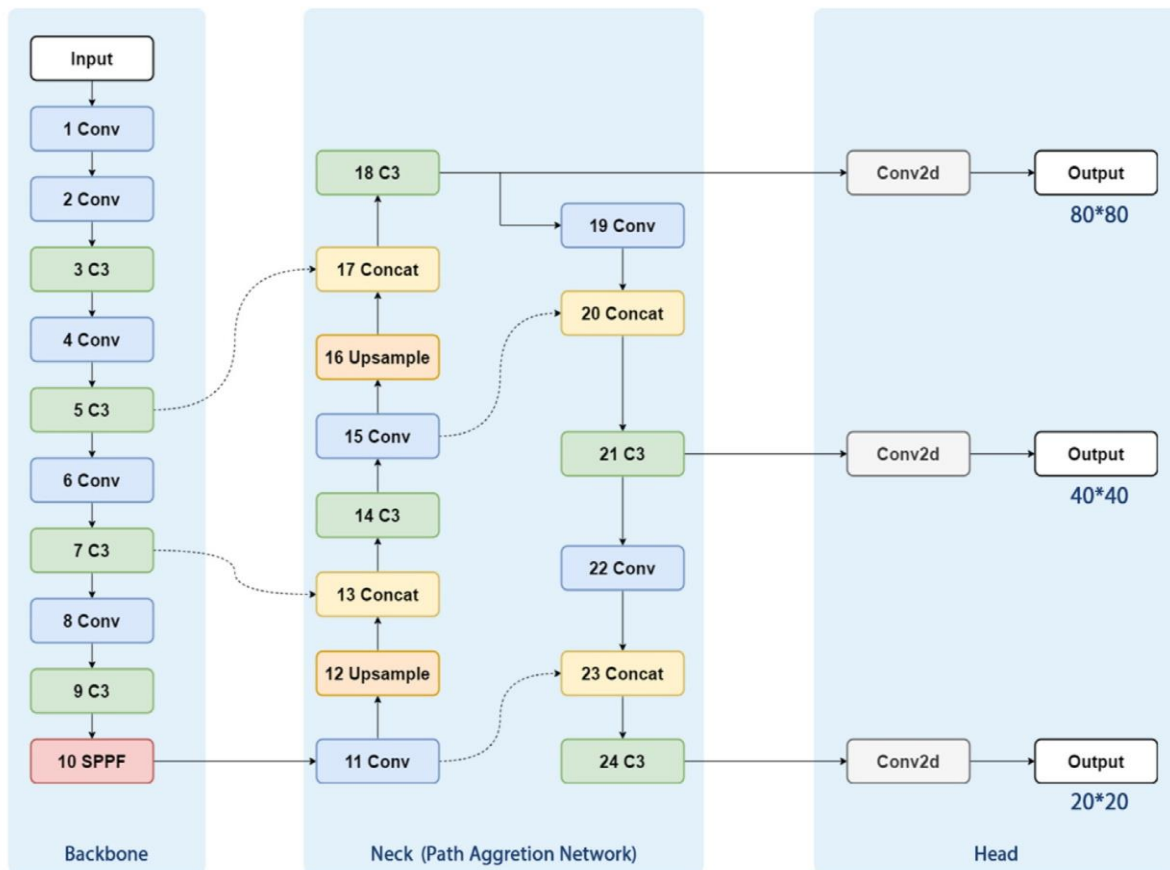
۲- معماری شبکه YOLOv5

همانطور که قبلاً ذکر شد، آشکارسازهای YOLO عمدتاً شبکه‌های عمیق تازه توسعه یافته برای هدف تشخیص اشیا هستند. آنها سرعت استنتاج بالاتری در مقایسه با سایر مدل‌ها دارند که باعث می‌شود برای برآورده کردن الزامات تشخیص اشیا در زمان واقعی مناسب‌تر باشند. شبکه‌های YOLO در نسخه‌های متوالی بهبود یافته‌اند. از آنجایی که YOLOv5 سریعتر از نسخه‌های قبلی خود است و عملکرد بهتری دارد، ما مدل پیشنهادی خود را بر اساس YOLOv5 v6.1 که در سال ۲۰۲۲ منتشر شد توسعه دادیم.

در این بخش ساختار پیش فرض این نسخه از YOLOv5 را توضیح می‌دهیم. تفاوت اصلی بین YOLOv5 و سایر مدل‌های YOLO بهبود معماری، پیاده‌سازی و ویژگی‌های پیشرفته آن است. YOLOv5 بر اساس معماری YOLOv4 ساخته شده و آن را با تمرکز بر دقت و کارایی بهبود می‌بخشد. YOLOv5 ویژگی‌های معماری بهبود یافته‌ای مانند فعال‌سازی SiLU و ادغام هرم فضایی سریع و همچنین پارامترهای کمتری نسبت به سایر نسخه‌ها دارد. YOLOv5 با استفاده از Python و PyTorch پیاده‌سازی می‌شود و آن را سریع‌تر کرده و پیچیدگی کمتری نسبت به نسخه‌های قبلی دارد و شامل ویژگی‌های پیشرفته‌ای مانند تولید جعبه لنگر مبتنی بر الگوریتم ژنتیک خودکار، تکنیک‌های مختلف تقویت تصویر و صادرات به فرمت‌های مختلف است.

این پیشرفت‌ها منجر به بهبود قابل توجهی نسبت به نسخه‌های قبلی شده است. ساختار کلی شبکه

⁹ Path Aggregation Network



شکل ۱- بلوک دیاگرام شبکه YOLOv5

اجسامی کوچک هستند، محاسبات مدل را افزایش داده و سرعت استنتاج را کاهش می‌دهد. بنابراین علاوه بر تغییرات پیشنهادی انجام شده در مدل پایه به منظور بهبود دقت مدل در تشخیص دندانها، تغییراتی نیز جهت افزایش سرعت استنتاج شبکه ایجاد کردیم.

شکل ۲ بلوک دیاگرام شبکه بهبود یافته YOLOv5 را نشان می‌دهد که اصلاحات پیشنهادی را برجسته می‌کند و بخش‌های فرعی زیر به توضیح جزئیات تغییرات پیشنهادی اختصاص داده شده است.

۳-۱- بهبودهای ساختاری

از آنجایی که دندانهای باکت شاول در تصاویر کوچک به نظر می‌رسند در ساختار شبکه YOLOv5 برای بهبود تشخیص اشیاء کوچک تغییراتی انجام دادیم. در لایه‌های اولیه شبکه در ستون فقرات، وضوح و اطلاعات مکانی نقشه‌های ویژگی بالا است.

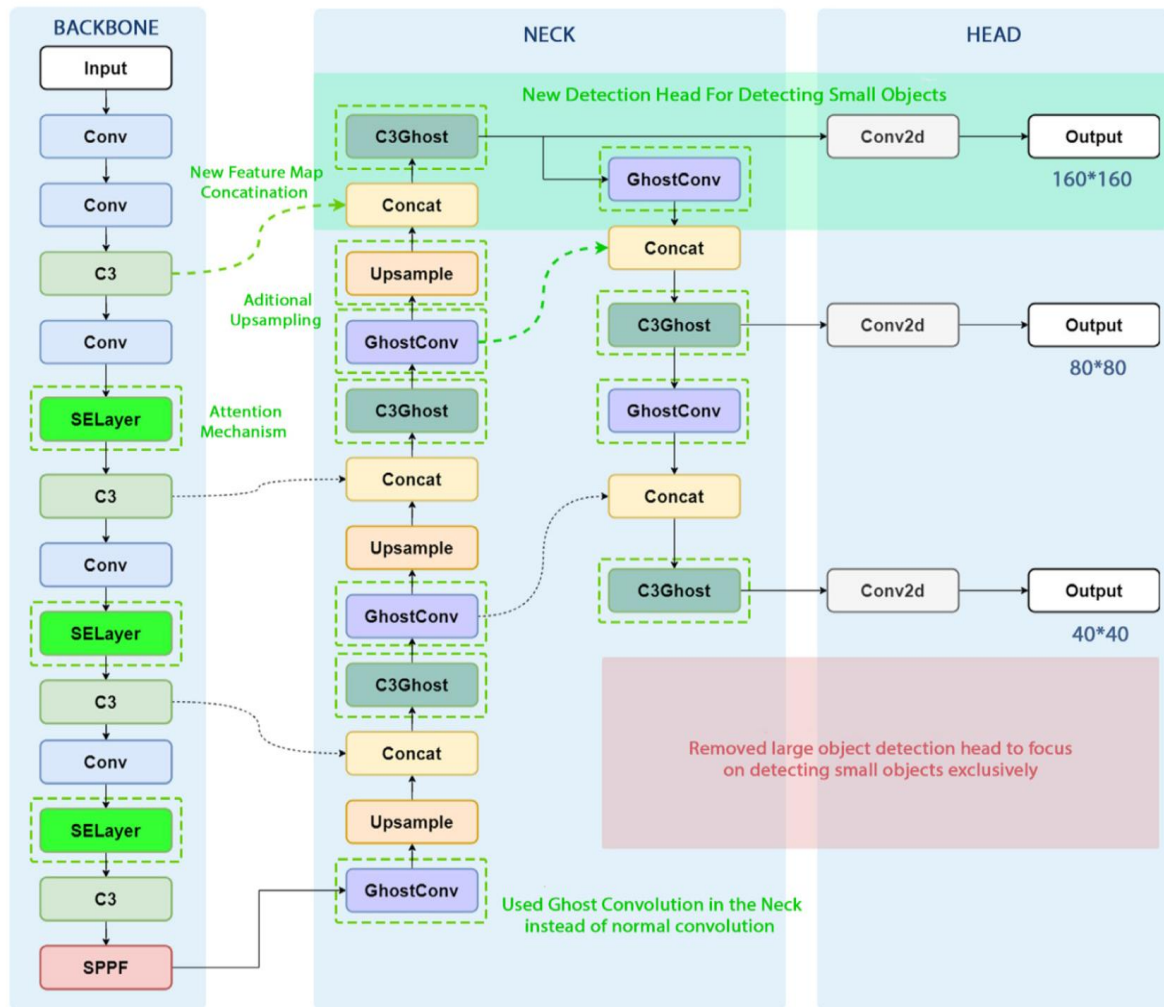
برای هر شی با استفاده از الگوریتم سرکوب غیر حداکثری^{۱۰} محاسبه می‌شود. با دانستن ساختار کلی شبکه YOLOv5، بخش بعدی نشان می‌دهد که چگونه عملکرد این شبکه را به صورت بلادرنگ به منظور تشخیص دقیق دندانها که کوچک هستند، تغییر داده و بهبود بخشیده‌ایم.

۳-۲- روش پیشنهادی

در این بخش، شبکه پیشنهادی بر اساس معماری YOLOv5 توسعه یافته به تفصیل توضیح داده شده است. در این تحقیق هدف اصلی توسعه یک مدل دقیق برای تشخیص بلادرنگ دندانهای باکت در شاولهای معدن مس می‌باشد. بدیهی است که دستاوردهای شبکه پیشنهادی می‌تواند الهام‌بخشی جهت تشخیص دقیق و بلادرنگ شی مشابه باشد.

تغییرات پیشنهادی انجام شده در مدل پایه به منظور بهبود عملکرد مدل در تشخیص بلادرنگ دندانهای باکت شاول می‌باشد. از سوی دیگر تمرکز بر روی تشخیص دندانها که

¹⁰ Non-maximum suppression



شکل ۲- بلوک دیاگرام شبکه YOLOv5 پیشنهادی

لایه جدید `concat` به گردن اضافه می‌شود که نقشه‌های ویژگی با وضوح بالا را از اولین لایه `C3` به گردن الحاق می‌کند. این کار با اضافه کردن یک دنباله جدید از چهار لایه کانولوشن، `upsamp`، `concat` و `C3` به گردن انجام می‌شود که به طور موثر اطلاعات مکانی بیشتری را حفظ می‌کند. با افزودن یک لایه `upsample` جدید به گردن، تعداد این لایه‌ها از دو به سه افزایش می‌یابد که از دست دادن اطلاعات مکانی را کاهش می‌دهد. همانطور که قبلاً ذکر شد، YOLOv5 دارای چندین سر تشخیص یا لایه‌های خروجی است که برای تشخیص اجسام در مقیاس‌های مختلف است. به طور خاص، مدل پایه دارای سه لایه خروجی با ۳۲، ۱۶ و ۸ گام برای اجسام بزرگ، متوسط و کوچک است. با توجه به تغییرات پیشنهادی در لایه‌های ستون فقرات و گردن، یک لایه خروجی با گام پایین ۴ نیز به عنوان سر جدید اضافه می‌شود تا توجه مدل پیشنهادی

پس از اعمال عملیات ادغام متوالی، وضوح و اطلاعات مکانی نقشه‌های ویژگی به تدریج کاهش می‌یابد، در حالی که در مقابل، اطلاعات معنایی و ویژگی‌های آموخته شده از دسته‌ها بهبود می‌یابند. برای بهره‌وری حداکثری از اطلاعات، YOLOv5 نقشه‌های ویژگی از مراحل مختلف لایه ستون فقرات را با نقشه‌های ویژگی در لایه گردن به هم متصل می‌کند. بنابراین شبکه قادر به تشخیص اشیاء در مقیاس‌های مختلف است. با این حال، با توجه به کاربرد عمومی شبکه YOLOv5 در تشخیص اشیاء، نقشه‌های ویژگی از اولین لایه‌های کانولوشن نادیده گرفته می‌شوند. زیرا آنها هنوز وضوح بالایی دارند که برای اهداف عمومی تشخیص اشیاء ضروری نیست.

از سوی دیگر، نقشه‌های ویژگی با وضوح بالا و اطلاعات مکانی برای تشخیص اجسام کوچک حیاتی هستند. بنابراین در مدل پیشنهادی برای تشخیص دندانها، یک

که به کانال‌هایی با ویژگی‌های استخراج ارزشمندتر توجه بیشتری می‌کند. کانال‌های نقشه ویژگی که خروجی یک لایه کانولوشن هستند، اندازه $H \times W \times C$ دارند، در حالی که ابعاد به ترتیب نشان‌دهنده ارتفاع و عرض نقشه‌های ویژگی و تعداد کانال‌ها می‌باشد. به منظور تعیین وزن یا اهمیت هر کانال ویژگی از GAP^{12} به عنوان عملیات فشرده‌سازی استفاده می‌شود که میانگین تمام نقاط تصویر در هر نقشه ویژگی را گرفته و منجر به یک تانسور با ابعاد $C \times 1 \times 1$ می‌شود. سپس یک ساختار گلوگاه با استفاده از دو لایه کاملاً متصل¹³ وجود دارد. اولین لایه کاملاً متصل یک لایه پنهان است که به عنوان یک بلوک کاهش استفاده می‌شود که کانال‌های ویژگی C را به C/r کاهش می‌دهد. لازم به ذکر است که بین افزایش پیچیدگی و بهبود عملکرد با کاهش r تعادل وجود دارد.

در روش پیشنهادی از مقدار پیش‌فرض $C/16$ در مرجع [۳۱] استفاده می‌شود. سپس، لایه دوم کاملاً متصل کانال‌های کاهش یافته را به بعد اصلی C گسترش می‌دهد. مزایای استفاده از این لایه‌های کاهش و گسترش در مقایسه با استفاده مستقیم از یک لایه کاملاً متصل، یادگیری بهتر همبستگی پیچیده بین کانال‌ها و کاهش تعداد پارامترها در شبکه می‌باشد. پس از لایه‌های کاملاً متصل، از یک تابع فعال‌سازی سیگموئید (σ) استفاده می‌شود که وزن کانال‌ها را بین ۰ و ۱ نرمال می‌کند.

در نهایت، این وزن‌های نرمال شده برای کانال‌های ویژگی با عملیات مقیاس اعمال می‌شوند. در این مقاله از ماژول SE در ستون فقرات شبکه YOLOv5 استفاده کرده‌ایم. بلوک‌های اصلی تشکیل دهنده ستون فقرات در YOLOv5، بلوک‌های Conv و C3 هستند. همانطور که

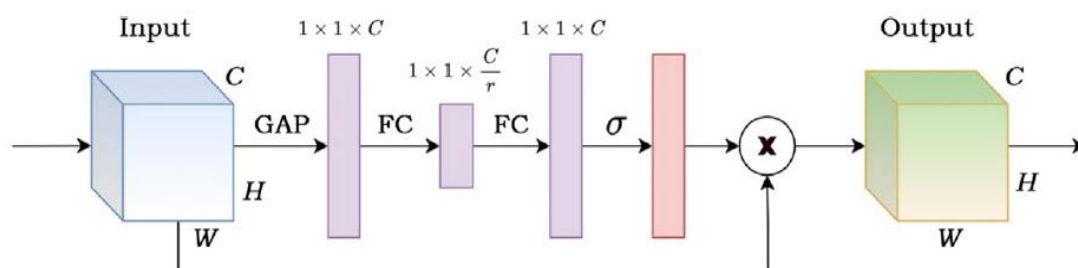
را بیشتر بر روی اجسام کوچک متمرکز کند. همچنین لایه خروجی با گام ۳۲ و لایه‌های متناظر با آن در گردن را حذف کردیم. با استفاده از گام پایین‌تر، سلول‌های بیشتری در شبکه تشخیص با اندازه‌های کوچکتر وجود دارد. به عنوان مثال، اگر ورودی شبکه را 640×640 فرض کنیم ابعاد نقشه ویژگی ۱۶۰ خواهد بود که برای تشخیص اجسام کوچک مناسب است. در شکل ۲ به منظور مقایسه بین شبکه پیشنهادی و مدل پایه لایه‌های اضافه شده با رنگ سبز و لایه‌های حذف شده با رنگ قرمز نشان داده شده است.

۳-۲- افزودن سازوکار توجه SE¹¹ به ستون فقرات

ما از سازوکار توجه کلنال جدید در ستون فقرات برای بهبود استخراج ویژگی و در نتیجه بهبود دقت استفاده کردیم. بنابراین توانایی شبکه برای تشخیص دقیق دندانها در باکت شاول بهبود می‌یابد و در نتیجه تشخیص با دقت خوبی در کاربردهای عملی به دست می‌آید. شکل ۳ ساختار دقیق ماژول افزودن سازوکار توجه SE را نشان می‌دهد.

SE [۳۱] یک بلوک نوری است که سازوکار توجه را [۳۲] به مدل‌ها اضافه می‌کند. این ماژول افزودنی ساده و در عین حال کارآمد را می‌توان به هر معماری پایه اضافه کرد تا بهبود عملکرد را با سربار محاسباتی ناچیز به دست آورد. در CNN‌ها، فیلترهای پیچشی (کلنال‌ها) نقشه‌های ویژگی را بر اساس وزن آن فیلترها می‌سازند. با این حال، ویژگی‌های استخراج شده در برخی کانال‌ها مهم‌تر از سایر کانال‌ها هستند.

بنابراین ایده خوبی است که به کلنال‌های مهم‌تر اهمیت بیشتری داده شود. این پایه و اساس سازوکار توجه می‌باشد



شکل ۳- معماری ماژول SE

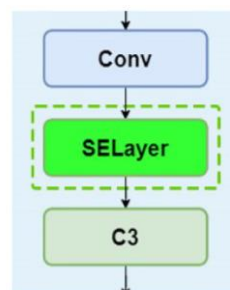
¹³ fully connected

¹¹ Squeeze and Excitation

¹² Global Average Pooling

YOLO

در شکل ۴ نشان داده شده است برای استفاده موثرتر از ویژگی‌های استخراج شده از بلوک Conv، ماژول SE بین این دو بلوک اضافه می‌شود.



شکل ۴- استفاده از ماژول SE بین لایه‌های Conv و C3

با این حال، از آنجایی که نقشه‌های ویژگی از دو لایه اول Conv در ستون فقرات دارای وضوح بالایی هستند، افزودن لایه SE بعد از این لایه‌های Conv تعداد پارامترهای مدل را به شدت افزایش می‌دهد و تاثیر منفی بر سرعت آموزش و استنتاج دارد. به همین دلیل ماژول SE از لایه Conv سوم به بعد بین لایه‌های Conv و C3 اضافه شده است. از آنجایی که بیشتر استخراج ویژگی در ستون فقرات انجام می‌شود، شبکه پیشنهادی فقط از ماژول SE در ستون فقرات استفاده می‌کند.

با این حال، از آنجایی که نقشه‌های ویژگی از دو لایه اول Conv در ستون فقرات دارای وضوح بالایی هستند، افزودن لایه SE بعد از این لایه‌های Conv تعداد پارامترهای مدل را به شدت افزایش می‌دهد و تاثیر منفی بر سرعت آموزش و استنتاج دارد. به همین دلیل ماژول SE از لایه Conv سوم به بعد بین لایه‌های Conv و C3 اضافه شده است. از آنجایی که بیشتر استخراج ویژگی در ستون فقرات انجام می‌شود، شبکه پیشنهادی فقط از ماژول SE در ستون فقرات استفاده می‌کند.

با این حال، از آنجایی که نقشه‌های ویژگی از دو لایه اول Conv در ستون فقرات دارای وضوح بالایی هستند، افزودن لایه SE بعد از این لایه‌های Conv تعداد پارامترهای مدل را به شدت افزایش می‌دهد و تاثیر منفی بر سرعت آموزش و استنتاج دارد. به همین دلیل ماژول SE از لایه Conv سوم به بعد بین لایه‌های Conv و C3 اضافه شده است. از آنجایی که بیشتر استخراج ویژگی در ستون فقرات انجام می‌شود، شبکه پیشنهادی فقط از ماژول SE در ستون فقرات استفاده می‌کند.

۳-۳- استفاده از Ghost Convolution در گردن

ماژول Ghost Convolution یک ماژول خلاقانه است که قادر به بدست آوردن نقشه‌های ویژگی بیشتر از پارامترها و محاسبات کمتر می‌باشد [۳۳]. بنابراین می‌تواند پیچیدگی شبکه را کاهش داده و سرعت آموزش و استنتاج را افزایش دهد که یکی از اهداف شبکه پیشنهادی است. اگر ابعاد نقشه ویژگی ورودی $H \times W \times C$ (H، W و C ارتفاع، عرض و تعداد کانال‌های ورودی) باشد، ابعاد نقشه ویژگی خروجی $H' \times W' \times C'$ (H'، W' و C' ارتفاع، عرض و تعداد کانال‌های خروجی) و کرنل‌ها $k \times k$ هستند. میزان پیچیدگی محاسباتی در یادگیری ماشین معمولاً توسط FLOPS^{۱۴} محاسبه می‌شود. تعداد FLOPS های مورد نیاز در کانولوشن معمولی برابر با مقدار حاصلضرب $C' \times$

۳-۴- استفاده از تابع زبان تطبیقی و وزن مجدد نسبی

رگرسیون جعبه مرزی یک تکنیک محبوب است که برای پیش‌بینی جعبه‌های محلی‌سازی در بسیاری از رویکردهای تشخیص اشیا استفاده می‌شود. جعبه‌های مرزبندی دقیق‌تر که به درستی حاوی شی هستند، می‌توانند دقت مدل را در طبقه‌بندی بهبود بخشند. برای هر جعبه مرزی چند پیش‌بینی وجود دارد: مختصات مرکز، طول و عرض جعبه مرزی، احتمال شی بودن و احتمال دسته‌های مختلف برای شی شناسایی شده در یک جعبه. تابع هزینه مدل‌های

¹⁴ Floating Point Operations Per Second

در رابطه ۲، b و bgt مراکز جعبه مرزی پیش‌بینی شده و جعبه حقیقت زمین هستند. p فاصله اقلیدسی بین این دو مرکز و c طول مورب کوچکترین جعبه مرزی است که دو جعبه را در بر می‌گیرد. β یک پارامتر تنظیم مثبت است و v ثبات نسبت طول و عرض دو جعبه را اندازه‌گیری می‌کند. (w, h, wgt و hgt عرض و ارتفاع جعبه مرزی پیش‌بینی شده و جعبه مرزی حقیقت زمین می‌باشند). این تابع در طول آموزش در مدل پایه ثابت است. برای بهبود بیشتر کارایی توابع زیان مبتنی بر IoU، یک تبدیل Box-Cox [۳۷] به تابع زیان $L_{IoU} = 1 - IoU$ با استفاده از یک ضریب α تطبیقی مطابق با رابطه ۳ اعمال شد.

$$L_{\alpha - IoU} = \frac{1 - IoU^{\alpha}}{\alpha}, \quad \alpha > 0 \quad (3)$$

با استفاده از این ایده، می‌توان عملکرد تابع $CIoU$ استاتیک را با استفاده از تابع از دست دادن تطبیقی α - $CIoU$ پیشنهاد شده در [۳۶] که به صورت رابطه ۴ است، بهبود بخشید.

$$L_{\alpha - CIoU} = 1 - IoU^{\alpha} + \frac{\rho^{2\alpha}(b, bgt)}{c^{2\alpha}} + (\beta v)^{\alpha} \quad (4)$$

با پارامترهایی که قبلاً در رابطه ۲ توضیح داده شد، به صورت تجربی نشان داده شده است که α نسبت به مدل‌ها یا مجموعه داده‌های مختلف بیش از حد حساس نیست و با $\alpha = 3$ در اکثر موارد به طور مداوم خوب عمل می‌کند. لازم به ذکر است که تابع زیان IoU نرمال با IoU رابطه خطی دارد و دارای یک گرادیان ثابت است. یکی از خواص اصلی توابع زیان مبتنی بر α این است که آن‌ها اشیاء را به صورت تطبیقی مجدد وزن می‌کنند (مطابق با به مقادیر IoU آن‌ها). وقتی $\alpha > 1$ باشد وزن از دست رفته و گرادیان اشیاء که IoU بالایی دارند به صورت تطبیقی افزایش می‌یابند و این باعث افزایش دقت جعبه مرزی پیش‌بینی شده می‌شود.

این وزن‌دهی مجدد که بر اساس α و IoU انجام می‌شود، باعث می‌شود که مثال‌های آسان‌تری را با IoU بزرگ‌تر

مبتنی بر YOLO ترکیبی از این پیش‌بینی‌ها به صورت رابطه ۱ می‌باشد.

$$Loss = l_{bx} + l_j + l_s \quad (1)$$

در تابع هزینه l_{bx} نشان‌دهنده تابع رگرسیون جعبه مرزی، l_j کلندید احتمال شیء و l_s ضرری است که دقت طبقه‌بندی را نشان می‌دهد. معمولاً از آنتروپی متقاطع باینری با لاجیت‌ها برای دو عبارت آخر رابطه ۱ استفاده می‌شود، که برای محاسبات باینری و چند دسته‌ای مناسب است.

با این حال، اولین عبارت (l_{bx}) مهم‌ترین عبارت است، زیرا اگر موقعیت جعبه مرزی به درستی پیش‌بینی نشود، دسته آن را نمی‌توان به درستی پیش‌بینی کرد. مدل‌های تشخیص شی مدرن از IoU ^{۱۵} به عنوان هزینه رگرسیون جعبه مرزی پیش‌بینی‌شده در مقابل حقیقت زمین، که یک تابع مستقل از مقیاس است، استفاده می‌کنند. مشکل اصلی با تابع IoU این است که اگر جعبه مرزی پیش‌بینی شده هیچ تقاطعی با جعبه حقیقت‌زمین نداشته باشد ($IoU = 0$)، جهت گرادیان از بین می‌رود و بهینه‌سازی نمی‌تواند اجرا شود [۳۵]. در نتیجه نسخه‌های بهبود یافته این عملکرد مهم به تدریج معرفی شد. نویسندگان در [۳۶] پیشرفت در خانواده از دست دادن IoU را به طور کامل توضیح داده‌اند.

به عنوان مثال در مدل پایه YOLOv5، از تابع $CIoU$ استفاده شده است که نسبت به IoU معمولی پیشرفت قابل توجهی دارد. $CIoU$ تقاطع جعبه‌های مرزی پیش‌بینی‌شده و حقیقت زمین، فاصله مرکزی دو جعبه و نسبت ابعاد دو جعبه را به صورت رابطه ۲ در نظر می‌گیرد [۳۶].

$$L_{CIoU} = 1 - IoU + \frac{\rho^2(b, bgt)}{c^2} + \beta v$$

$$v = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (2)$$

$$\beta = \frac{v}{(1 - IoU) + v}$$

¹⁵ Intersection over Union

YOLO

نمونه از تصاویر مجموعه داده مورد استفاده در شکل ۶ نشان داده شده است.



شکل ۶- نمونه‌هایی از مجموعه داده مورد استفاده

۴-۲- محیط شبیه‌سازی

روش پیشنهادی با استفاده از زبان برنامه‌نویسی پایتون در محیط ویندوز و توسط یک رایانه استاندارد با پردازنده‌ای با مشخصات Intel i7-10750H، کارت گرافیک NVIDIA GTX 1650ti و رم ۱۶ گیگابایت پیاده‌سازی شد.

۴-۳- معیارهای ارزیابی

در این بخش معیارهای ارزیابی مورد استفاده در این مقاله را معرفی می‌کنیم. یکی از مهم‌ترین مراحل پس از طراحی و ساخت یک مدل یا یک الگوریتم ارزیابی کارایی آن است حساسیت (نرخ پاسخ‌های مثبت درست) و اختصاصیت یا تشخیص‌پذیری (نرخ پاسخ‌های منفی درست) دو شاخص مهم برای ارزیابی آماری عملکرد نتایج طبقه‌بندی هستند. زمانی که بتوان داده‌ها را به دو گروه مثبت و منفی تقسیم کرد، عملکرد نتایج یک آزمایش که اطلاعات را به این دو دسته تقسیم می‌کند با استفاده از شاخص‌های حساسیت و تشخیص‌پذیری قابل اندازه‌گیری و توصیف است برای اندازه‌گیری کارایی ۴ پارامتر نیاز است [۳۸]:

بسازید تا به تدریج افزایش سرعت به سمت $CIoU = 1$ یاد بگیرید.

از طرفی سرعت یادگیری و آموزش نمونه‌های سخت در مراحل بعدی که IoU آنها افزایش می‌یابد، بهبود می‌یابد. این راه حل، ضمن بهبود دقت مدل و سرعت همگرایی، بر سرعت استنتاج مدل نهایی تأثیری ندارد.

۴-۴- ارزیابی نتایج

در این بخش ابتدا مجموعه داده مورد استفاده برای ارزیابی الگوریتم پیشنهادی تشریح می‌شود. سپس، معیارهای ارزیابی مورد استفاده برای ارزیابی نتایج معرفی شده‌اند. در ادامه نتایج به دست آمده گزارش و تحلیل می‌شوند.

۴-۱- پایگاه داده

به منظور ارزیابی الگوریتم پیشنهادی جهت تشخیص دندان‌های باکت شاول، یک پایگاه داده شامل ۶۰۰ تصویر در روز و شب جمع‌آوری شده است. مجموعه داده‌های شاول الکتریکی تحت شرایط کاری واقعی برای آموزش مدل‌های یادگیری عمیق جمع‌آوری شده‌اند. دوربین بر روی بازوهای شاول مطابق با شکل ۵ نصب شده و دندان‌های باکت را پایش می‌کند.



شکل ۵- سیستم پایش هوشمند دندان‌های باکت شاول

۷۰ درصد از تصاویر به عنوان مجموعه نمونه آموزشی، ۲۰ درصد به عنوان مجموعه نمونه اعتبارسنجی و ۱۰ درصد به عنوان مجموعه نمونه آزمایشی استفاده می‌شود. چند

معیار اصلی برای ارزیابی دقت مدل‌های تشخیص اشیا میانگین متوسط دقت (mAP) است که بر اساس میانگین متوسط دقت (AP) در دسته‌های مختلف که به صورت رابطه ۸ است، محاسبه می‌شود.

$$mAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (۸)$$

که در آن N تعداد دسته‌ها است. به طور کلی، AP_i نشان‌دهنده دقت متوسط دسته نام بیش از آستانه‌های مختلف IoU است. این آستانه معمولاً بین ۰.۵ تا ۰.۹۵ است و با ۰.۰۵ پله افزایش می‌یابد. ما از معیار mAP50 (۰.۵ = آستانه) برای ارزیابی دقت مدل پیشنهادی استفاده می‌کنیم.

۴-۴-۴-ارائه و تحلیل نتایج

در این بخش ابتدا نتایج روش پیشنهادی با تعدادی از مدل‌های پایه YOLOv5 مقایسه می‌گردد. جدول ۱ نتایج بدست آمده مدل‌های پایه YOLOv5 و مدل پیشنهادی را در مجموعه داده تصاویر ایجاد شده در این پژوهش با استفاده از تصاویر ورودی با اندازه ۶۴۰×۶۴۰ را نشان می‌دهد. نتایج این ارزیابی بر حسب معیارهای مختلف دقت، فراخوانی، F1-score و mAP50 نشان داده شده است.

همان‌طور که در جدول ۱ مشاهده می‌شود، روش پیشنهادی دقت بیشتری نسبت به همه مدل‌های پایه دارد، زیرا مقدار mAP50 آن به ترتیب ۱.۹، ۱.۸ درصد و ۲.۲ درصد بیشتر از YOLOv5M، YOLOv5L و YOLOv5X است.

نتایج به دست آمده با میانگین دقت ۹۳.۵ درصد بیانگر بهبود دقت تشخیص مدل پیشنهادی نسبت به مدل‌های پایه YOLOv5 در فرآیند تشخیص دندان‌های باکت شاول می‌باشد. بنابراین با ایجاد تغییراتی در مدل پایه توانایی شبکه پیشنهادی برای تشخیص دقیق دندان‌ها در باکت شاول بهبود می‌یابد و در نتیجه تشخیص با دقت خوبی در کاربردهای عملی به دست می‌آید.

مثبت صحیح (True Positive): درست شناسایی شده است.

مثبت کاذب (False Positive): اشتباه شناسایی شده است.

منفی صحیح (True Negative): به درستی رد شده است.

منفی کاذب (False Negative): به اشتباه رد شده است.

یکی از معیارهای ارزیابی روش پیشنهادی معیار دقت^{۱۶} است. این معیار نسبت «تعداد پیش‌بینی‌های صحیح انجام شده» برای نمونه‌های یک دسته خاص، به «تعداد کل پیش‌بینی‌ها» برای نمونه‌های همان دسته خاص را ارائه می‌کند. مقادیر بالا برای معیار دقت، بیانگر تعداد کم داده‌هایی است که به اشتباه در دسته خاص طبقه‌بندی شده‌اند. این معیار به صورت رابطه ۵ محاسبه می‌شود:

$$Precision = \frac{TP}{TP + FP} \quad (۵)$$

معیار بعدی برای ارزیابی فراخوانی^{۱۷} است. این معیار بیان‌کننده نسبت «تعداد داده‌های درست طبقه‌بندی شده» در یک دسته خاص، به «تعداد کل داده‌هایی است که باید در همان دسته خاص طبقه‌بندی شوند» می‌باشد. مقادیر بالا برای این معیار، بیانگر تعداد کم داده‌هایی است که به اشتباه، در آن دسته خاص طبقه‌بندی نشده‌اند و به صورت رابطه ۶ محاسبه می‌شود:

$$Recall = \frac{TP}{TP + FN} \quad (۶)$$

معیار F1 یک معیار مناسب برای ارزیابی دقت یک آزمایش است. این معیار پارامترهای دقت و فراخوانی را با هم ترکیب می‌کند تا مشخص شود یک مدل طبقه‌بند تا چه حد عملکرد خوبی از خود نشان می‌دهد. معیار F1 در بهترین حالت، یک و در بدترین حالت صفر است. چگونگی محاسبه این معیار در رابطه ۷ آورده شده است.

$$F1 = 2 \times \frac{Recall \times Precision}{Recall + Precision} \quad (۷)$$

¹⁷ Recall

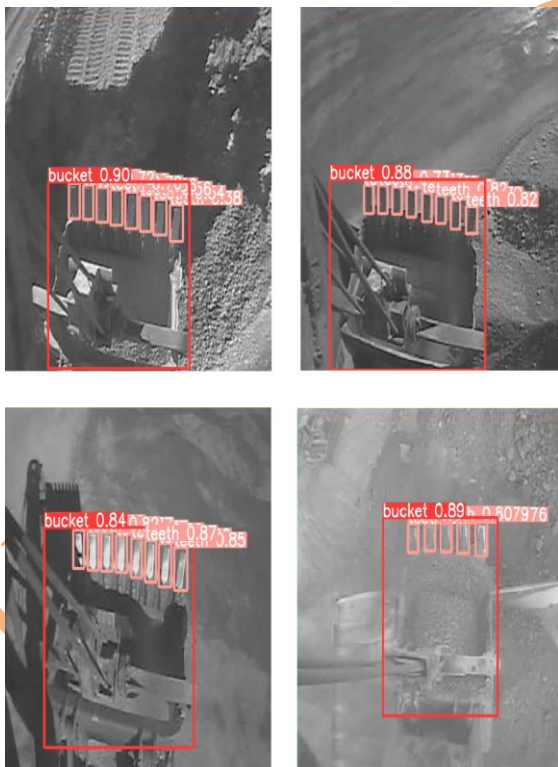
¹⁶ Precision

YOLO

جدول ۱- نتایج حاصل از ارزیابی مدل‌های پایه YOLOv5 و مدل پیشنهادی بر روی مجموعه داده تصاویر باکت شاول

مدل	دقت (%)	فراخوانی (%)	معیار F1 (%)	میانگین دقت متوسط (%)
	Precision	Recall	F1- score	mAP50
YOLOv5M	۸۵.۵	۹۴.۵	۸۹.۷۸	۹۱.۶
YOLOv5L	۸۶.۶	۹۳.۸	۹۰.۰۵	۹۱.۷
YOLOv5X	۸۵.۶	۹۴.۱	۸۹.۶۴	۹۱.۳
روش پیشنهادی	۸۹.۵	۹۵.۵	۹۲.۴۰	۹۳.۵

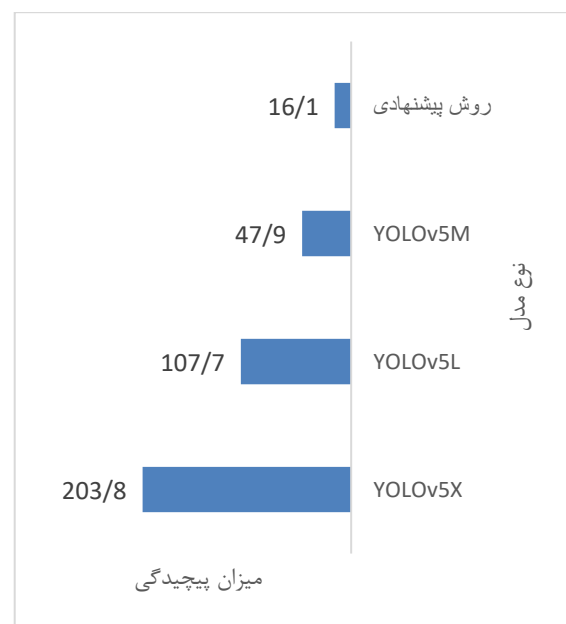
شکل ۷ خروجی مدل پیشنهادی را بر روی تصاویری از باکت شاول نشان می‌دهد.



شکل ۷- نتایج تشخیص مدل پیشنهادی در مجموعه نمونه آزمایشی

جدول ۲ مقایسه‌ای از عملکرد مدل پیشنهادی با تعدادی از مدل‌های پایه YOLOv5 از نظر میزان پیچیدگی را نشان می‌دهد. همان‌طور که در جدول مشاهده می‌شود، نتایج به دست آمده حاصل از مدل پیشنهادی با پیچیدگی ۱۶.۱، بیانگر کاهش پیچیدگی مدل در فرآیند تشخیص می‌باشد.

جدول ۲- مقایسه مدل پیشنهادی با تعدادی از مدل‌های پایه YOLOv5 از نظر میزان پیچیدگی



جدول ۳ مقایسه‌ای از عملکرد مدل پیشنهادی با سایر کارهای انجام شده مربوط به این حوزه در سال‌های اخیر را نشان می‌دهد.

جدول ۳- مقایسه مدل پیشنهادی با سایر کارهای انجام شده در سال‌های اخیر

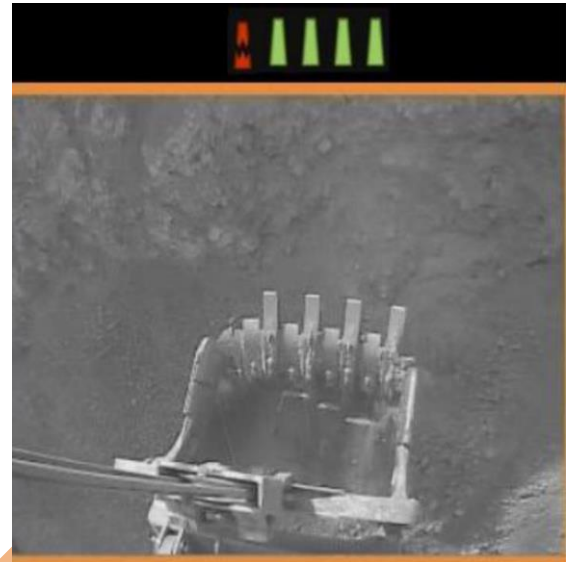
تعداد تصویر پردازش شده در یک ثانیه	پیچیدگی GFLOPS	دقت (%) Precision	روش
۱۰	۸۱	۸۴.۲	[۱۱]
۶	۱۹۰	۶۲.۱	[۱۲]
۱۶	۶۲	۷۹	[۱۳]
۴۹	۱۶.۱	۸۹.۵	روش پیشنهادی

همان‌طور که در جدول مشاهده می‌شود، نتایج به دست آمده حاصل از مدل پیشنهادی با دقت ۸۹.۵ و پیچیدگی ۱۶.۱، بیانگر بهبود در دقت و همچنین کاهش پیچیدگی مدل در فرآیند تشخیص می‌باشد. مدل‌های تشخیص بلادرنگ باید بتوانند تصویرهای ویدئویی را با سرعتی در حدود ۲۰ تا ۳۰ تصویر در ثانیه پردازش کنند. بنابراین سایر روش‌های انجام شده نمی‌تواند به سرعت مورد نیاز برای تشخیص بلادرنگ دندان‌های باکت در شاول‌های معدن دست یابد. در حالیکه روش پیشنهادی با تعداد ۴۹ تصویر پردازش شده در ثانیه می‌تواند به طور کامل نیاز تشخیص بلادرنگ را برآورده نماید.

۵- نتیجه‌گیری

در این مقاله، یک نسخه بهبود یافته از مدل YOLOv5 برای تشخیص دندان‌های باکت شاول پیشنهاد شد و در مجموعه داده‌ای جدید از شاول تحت شرایط کاری واقعی، ارزیابی شد. در مقایسه با مدل YOLOv5، روش

بنابراین با نصب دوربین بر روی بازوهای شاول و تشخیص دقیق و بلادرنگ دندان‌ها در مرحله بعد می‌توان دندان‌های سالم را با رنگ سبز و در صورتی که دندان‌های جدا شده بود آن را با رنگ قرمز مطابق با شکل ۸ نشان داد.



شکل ۸- نتایج نهایی الگوریتم پیشنهادی

بنابراین با مشخص بودن تعداد دندان‌ها و تشخیص دقیق و بلادرنگ دندان‌ها، در صورت جدا شدن دندان از باکت شاول بلافاصله به کارکنان هشدار داده می‌شود که شاول را بررسی نموده و در صورت جدا شدن دندان از فرآیند باربری توسط تراک مطابق با شکل ۹ جلوگیری شده و با بیرون کشیدن دندان جلوی توقف احتمالی سنگ‌شکن به علت گیر کردن دندان گرفته شده و واحد تعمیرگاه جهت قراردادن دندان جدید اقدام خواهد نمود.



شکل ۹- فرآیند باربری توسط تراک

YOLO

فرآیند تشخیص دندان‌ها می‌باشد. در مجموع نتایج به دست آمده نشان می‌دهد که چگونه تغییرات اعمال شده، شبکه YOLOv5 را به هدف تعیین شده برای آن نزدیکتر کرده است. بنابراین با تشخیص با دقت و بلادرنگ دندان‌ها، در صورت جدا شدن دندان از باکت شاول، بلافاصله به کارکنان هشدار داده می‌شود که باکت شاول را بررسی نمایند. همچنین کارایی مدل را می‌توان با اصلاحات جدید به منظور کاهش بیشتر پیچیدگی آن، حتی بیشتر نیز بهبود بخشید. در معادن روباز ممکن است در بعضی مواقع پدیده مه آشکار شود که بر عملکرد تشخیص دندان‌های باکت شاول تاثیر می‌گذارد. در کارهای آینده می‌توان ماژول مه‌زدایی تصویر را برای بهبود دقت تشخیص استفاده نمود.

پیشنهادی با تغییرات زیر تقویت شد: الف) گردن شبکه برای الحاق ویژگی‌های بهتر با افزودن یک سر تشخیص جدید برای تشخیص دندان‌های باکت که اجسامی کوچک هستند، بهبود یافت. ب) سازوکار توجه SE به ستون فقرات آن اضافه شد. ج) تابع زیان رگرسیون جعبه مرزی با یک تابع تطبیقی جایگزین شد تا دقت مدل را بهبود بخشد. د) برای افزایش سرعت استنتاج مدل از Ghost Convolution استفاده شد. این اصلاحات با هدف افزایش دقت مدل در تشخیص بلادرنگ دندان‌های باکت شاول در محیط روباز معدن انجام شد. مدل پیشنهادی با چندین مدل پایه YOLOv5 مقایسه شد. نتایج به دست آمده با میانگین دقت ۹۳.۵ درصد و پیچیدگی ۱۶.۱ بیانگر بهبود دقت تشخیص و کاهش پیچیدگی مدل پیشنهادی در

مراجع

- [1] C. Zhaoxue and Y. Hong. "Application of open-pit and underground mining technology for residual coal of end slopes." *Mining Science and Technology (China)* 20, no. 2 (2010): 266-270.
- [2] L. Bernardi, M. Kumral, and M. Renaud. "Comparison of fixed and mobile in-pit crushing and conveying and truck-shovel systems used in mineral industries through discrete-event simulation." *Simulation Modelling Practice and Theory* 103 (2020): 102100.
- [3] F. Ng, J. A. Harding, and J. Glass. "Improving hydraulic excavator performance through in line hydraulic oil contamination monitoring." *Mechanical Systems and Signal Processing* 83 (2017): 176-193.
- [4] Z. Deng, L. Wang, W. Liu, Z. Wang, and E. Qiule. "Mathematical modeling and fuzzy approach for disaster analysis on geo-spatial rock mass in open-pit mining." *Computer Communications* 150 (2020): 384-392.
- [5] S. Luo, H. Li, and H. Shen. "Deeply supervised convolutional neural network for shadow detection based on a novel aerial shadow imagery dataset." *ISPRS Journal of Photogrammetry and Remote Sensing* 167 (2020): 443-457.
- [6] X. Luo and H. Zhang. "Missing tooth detection with laser range sensing," In Fifth World Congress on Intelligent Control and Automation, vol. 4, pp. 3607-3610, 2004.
- [7] L. He, H. Wang, and H. Zhang. "Object detection by parts using appearance, structural and shape features," In 2011 IEEE International Conference on Mechatronics and Automation: IEEE, pp. 489-494, 2011.
- [8] Y. X. Duan, W. H. Du and Z. Q. Zeng. "Detection method for missing teeth of electric shovel based on machine vision", *Industrial Mine Automation*, vol. 44, no. 7, pp. 75-79, 2018.
- [9] S. Ren, K. He, R. Girshick, and J. Sun. "Faster r-cnn: towards real-time object detection with region proposal networks." *IEEE transactions on pattern analysis and machine intelligence* 39, no. 6 (2016): 1137-1149.
- [10] Z. Zou, K. Chen, Z. Shi, Y. Guo, and J. Ye. "Object detection in 20 years: a survey." *Proceedings of the IEEE* 111, no. 3 (2023): 257-276.
- [11] H. Shariati, A. Yeraliyev, B. Terai, S. Tafazoli, and M. Ramezani, "Towards autonomous mining via intelligent excavators," In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, pp. 26-32, 2019.
- [12] S. Ji, W. Li, B. Zhang, L. Zhou, and C. Duan. "Bucket teeth detection based on faster region convolutional neural network." *IEEE Access* 9 (2021): 17649-17661.
- [13] X. Liu, X. Qi, and Y. Jiang. "Electric shovel teeth missing detection method based on deep learning." *Computational Intelligence and Neuroscience*, no. 1 (2021): 6503029.

- [14] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg. "Ssd: single shot multibox detector," In Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, Proceedings, Part I 14: Springer, pp. 21-37, 2016.
- [15] J. Redmon, D. Santosh, 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, pp. 779-788, 2016.
- [16] T. Lin, P. Goyal, R. Girshick and K. He and P. Dollár. "Focal loss for dense object detection." In Proceedings of the IEEE international conference on computer vision, pp. 2980–2988, 2017.
- [17] K. Duan, S. Bai, L. Xie, H. Qi, Q. Huang, and Q. Tian. "Centernet: keypoint triplets for object detection." In Proceedings of the IEEE/CVF international conference on computer vision, pp. 6569-6578, 2019.
- [18] J. Redmon and A. Farhadi. "Yolo9000: better, faster, stronger". In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 7263-7271, 2017.
- [19] A. Farhadi and J. Redmon. "Yolov3: an incremental improvement". *arXiv preprint arXiv:1804.02767*, 2018.
- [20] Bochkovskiy, Alexey, Chien-Yao Wang, and Hong-Yuan Mark Liao. "Yolov4: optimal speed and accuracy of object detection." *arXiv preprint arXiv:2004.10934*, 2020.
- [21] G. Jocher, A. Chaurasia, A. Stoken, J. Borovec, Y. Kwon, J. Fang, K. Michael, D. Montes, J. Nadar, and P. Skalski. "Ultralytics/yolov5: v6. 1-tensorrt, tensorflow edge tpu and openvino export and inference." *Zenodo*, 2022.
- [22] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein and L. Antiga. "Pytorch: an imperative style, high-performance deep learning library". *Advances in neural information processing systems*, vol. 32, 2019.
- [23] S. Elfwing, E. Uchibe, and K. Doya. "Sigmoid-weighted linear units for neural network function approximation in reinforcement learning". *Neural networks* 107 (2018): 3-11.
- [24] C. Li, L. Li, H. Jiang, K. Weng, Y. Geng, L. Li, Z. Ke, Q. Li, M. Cheng, and W. Nie. "Yolov6: A single-stage object detection framework for industrial applications". *arXiv preprint arXiv:2209.02976*, 2022.
- [25] C. Wang, A. Bochkovskiy and H. Liao. "Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors." In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 7464-7475. 2023.
- [26] G. Jocher, A. Chaurasia, and J. Qiu. "Ultralytics yolov8 (version 8.0. 0)", 2023.
- [27] S. Zhang, H. Yang, C. Yang, W. Yuan, X. Li, X. Wang, Y. Zhang, X. Cai, Y. Sheng , X. Deng AND W. Huang "Edge device detection of tea leaves with one bud and two leaves based on shufflenetv2-yolov5-lite-e." *Agronomy* 13, no. 2 (2023): 577.
- [28] Z. Xing, Z. Zhang, X. Yao, Y. Qin, and L. Jia. "Rail wheel tread defect detection using improved yolov3." *Measurement* 203 (2022): 111959.
- [29] M. Cui, Y. Lou, Y. Ge, and K. Wang. "Les-yolo: a lightweight pinecone detection algorithm based on improved yolov4-tiny network." *Computers and Electronics in Agriculture* 205 (2023): 107613.
- [30] W. Hu, J. Xiong, J. Liang, Z. Xie, Z. Liu, Q. Huang, and Z. Yang. "A method of citrus epidermis defects detection based on an improved yolov5." *Biosystems Engineering* 227 (2023): 19-35.
- [31] J. Hu, L. Shen, and G. Sun. "Squeeze-and-excitation networks." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 7132-7141. 2018.
- [32] S. E. Petersen and M. I. Posner. "The attention system of the human brain: 20 years after." *Annual review of neuroscience* 35 (2012): 73-89.
- [33] K. Han, Y. Wang, Q. Tian, J. Guo, C. Xu, and C. Xu. "Ghostnet: more features from cheap operations." In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 1580-1589. 2020.
- [34] F. Chollet. "Xception: deep learning with depthwise separable convolutions." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1251-1258. 2017.
- [35] J. Wang, H. Xiao, L. Chen, J. Xing, Z. Pan, R. Luo, and X. Cai. "Integrating weighted feature fusion and the spatial attention module with convolutional neural networks for automatic aircraft detection from sar images." *Remote Sensing* 13, no. 5 (2021): 910.

YOLO

- [36] J. He, S. Erfani, X. Ma, J. Bailey, Y. Chi, and X.-S. Hua. " α -iou: a family of power intersection over union losses for bounding box regression". *Advances in Neural Information Processing Systems* 34 (2021): 20230-20242.
- [37] G. Box and D. Cox. "An analysis of transformations". *Journal of the Royal Statistical Society Series B: Statistical Methodology* 26, no. 2 (1964): 211-243.
- [38] D. Powers. "Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation." *Journal of Machine Learning Technologies* 2,no.1(2011): 37-63.

UNCORRECTED PROOF