

## تشخیص به موقع پرنده‌های هدایت پذیر از دور چند بال چرخان با استفاده از الگوریتم YOLOv5 بهینه‌سازی شده

مجید امیرزاده<sup>۱</sup>، سید علی حسینی مرادی<sup>۲\*</sup>، نادر قبادی<sup>۳</sup>

۱- دانشجوی دکتری، دانشگاه همدان، همدان، ایران ۲- پژوهشگر دانشگاه پدافند هوایی خاتم الانبیاء (ص)، تهران، ایران ۳- دانشیار، دانشگاه ملایر، ملایر، ایران

(دریافت: ۱۴۰۱/۱۲/۱۸، بازنگری: ۱۴۰۲/۰۱/۱۵، پذیرش: ۱۴۰۲/۰۱/۳۱، انتشار: ۱۴۰۲/۰۳/۰۱)

DOR: <https://dor.isc.ac/dor/20.1001.1.26762935.1402.14.1.2.2>

### چکیده

در سال‌های اخیر پیشرفت بسیار سریع فناوری در حوزه پهپادها (پرنده‌های هدایت‌پذیر از دور)، در کنار مزایای خود، تهدیدات جدی را در سطوح مختلف اجتماعی و امنیتی به همراه داشته است. از جمله این مشکلات، می‌توان به بحث پروازهای غیرمجاز در مناطق حفاظت‌شده و امنیتی اشاره کرد. لذا تشخیص به موقع این دستگاه‌ها در جهت انجام سریع اقدامات مربوطه، ضروری است. در همین راستا، در این پژوهش با بهره‌گیری از الگوریتم YOLOv5 که جزء جدیدترین نسخه الگوریتم‌های یک‌مرحله‌ای بینایی رایانه‌ای است، دو مدل با بهینه‌سازهای SGD و Adam جهت تشخیص به موقع پهپادها توسعه داده شده است. برای توسعه مدل‌های حاضر در این پژوهش، از یک مجموعه داده شامل ۱۰۰۴۶ عدد عکس از انواع و حالات مختلف پهپادها استفاده شده است. پردازش مدل‌ها به کمک بستر گوگل کولب انجام شده است که به صورت رایگان یک سیستم پردازشی قدرتمند را در اختیار توسعه‌دهندگان قرار می‌دهد. ارزیابی مدل‌ها بر روی چهار مجموعه‌آزمون ۱۰۰۰ عددی شامل مجموعه‌آزمون معمولی، کم‌حجم، حالت شب، خاکستری مقیاس و همچنین یک مجموعه‌آزمون شامل ۱۰۰ عدد عکس از چندین پهپاد صورت گرفته است. طبق نتایج، مدل توسعه داده شده با بهینه‌ساز Adam نسبت به مدل توسعه داده شده با بهینه‌ساز SGD عملکرد بهتری داشته است.

**کلیدواژه‌ها:** پهپاد (پرنده هدایت‌پذیر از دور)، تشخیص به موقع، پهپاد چند بال چرخان، YOLOv5

## Real Time Detection of Multi-Rotor Unmanned Aerial Vehicle Using YOLOv5 Optimized Algorithm

M. A. Amirzadeh, S. A. Hosseini Moradi<sup>\*</sup>, N. Ghobadi

Khatam Al Anbia University, Thran, Iran

(Received: 2023/03/09 ; revised: 2023/04/04 ; Accepted: 2023/04/20 ; published: 2023/05/22)

### Abstract

In recent years, the very rapid development of technology in the field of UAVs (Unmanned Aerial Vehicles), along with its advantages, has brought serious threats at various social and security levels. Among these problems, the issue of unauthorized flights in protected and security area can be mentioned. Therefore, real time detection of these devices is necessary in order to quickly carry out relevant measures. In this regard, in this research, using the YOLOv5 algorithm, which is part of the latest version of one-stage computer vision algorithms, two models with SGD and Adam optimizers have been developed for the real time detection of UAVs. To develop the models in this research, a dataset containing 10046 images of different types and states of UAVs has been used. The processing of the models has been done with the Google Colab platform, which provides a free powerful processing system for developers. The models have been evaluated on four test sets of 1000 images including normal, low volume, night mode, and gray scale and also a test set including 100 images from several UAVs. According to the results, the developed model with The Adam optimizer performed better than the model developed with the SGD optimizer.

**Keywords:** UAV (Unmanned Aerial Vehicle), Real Time Detection, Multi-Rotor UAV, YOLOv5.

## ۱. مقدمه

ناحیه را طبقه‌بندی و در صورت وجود شیء مورد نظر در آن ناحیه، جعبه‌های مرزی<sup>۷</sup> را در آن رسم می‌کنند [۳]. از جمله این الگوریتم‌ها می‌توان به R-CNN، FasterR-CNN و MaskR-CNN اشاره داشت. این الگوریتم‌های دومرحله‌ای، مدت‌زمان بیشتری جهت پردازش عکس‌ها نیاز دارند. اما در دسته‌ی دوم، الگوریتم‌های<sup>۸</sup> YOLO و<sup>۹</sup> SSD قرار دارند که تمام کارهای پردازشی توضیح داده‌شده را تنها در یک مرحله انجام می‌دهند و به نسبت دسته‌ی اول از سرعت بالاتری برخوردار هستند [۴].

از الگوریتم‌های مختلف بینایی رایانه‌ای، برای تشخیص اشیا و اهداف خاص به وسیله‌ی کامپیوتر، در زمینه‌های مختلفی استفاده می‌شود. از جمله مهم‌ترین کاربردهای این الگوریتم‌ها در حوزه‌ی پدافندی، می‌توان به شناسایی پهپادها به دلیل جلوگیری از بروز تهدیدات امنیتی اشاره کرد. در این زمینه، محققین مطالعات مختلفی را انجام داده‌اند. ثاقب و همکاران [۵] سه مدل را به کمک الگوریتم‌ها ZF، VGG16 و VGG16 برای تشخیص پهپاد توسعه دادند که در نهایت مدل VGG16 با میانگین صحت متوسط<sup>۱۰</sup> به مقدار ۶۶٪ بهترین عملکرد را نشان داد. آکر و همکاران [۶] در تحقیق خود به کمک الگوریتم مدلی را برای تشخیص محل پهپادها در فایل‌های ویدیویی توسعه دادند و سپس برای حل مشکل کمبود داده و عملکرد بهتر مدل از الگوریتم YOLOv2 برای آموزش مجدد مدل بهره بردند. لی و همکاران [۷] در تحقیق خود برای تشخیص پهپاد، از یک الگوریتم DeepCNN استفاده کردند که با توجه به تعداد کم مجموعه داده (۲۰۸۸ عکس) به صحت<sup>۱۱</sup> ۸۹٪ در مدل رسیدند که نشان از عملکرد مناسب این سری الگوریتم‌ها در مجموعه‌داده‌های کوچک دارد. شومانو و همکاران [۸] به کمک الگوریتم‌های CNN، مدلی را جهت تشخیص پهپاد و سایر پرندگان به کمک فیلم‌های دریافتی از دوربین توسعه دادند. هو و همکاران [۳] در تحقیق خود، نسخه بهبودیافته‌ای از الگوریتم YOLOv3 را جهت تشخیص پهپاد استفاده کردند و عملکرد آن را مورد ارزیابی قرار دادند. اونلو و همکاران [۹] نیز یک مدل سبک‌تر از YOLOv3 جهت تشخیص و رهگیری پهپادها، با هدف چرخش کمتر دوربین توسعه دادند. ویژون و همکاران [۱۰] با در نظرگیری استفاده‌های نادرست احتمالی از پهپادها، همانند حادثه‌ی فرودگاه گتویک<sup>۱۲</sup> که حدود ۱۴۰۰۰۰ مسافر را تحت تأثیر خود قرار داد، وجود یک مدل جهت تشخیص به موقع پهپادها را ضروری دانستند. آن‌ها با اعمال تقویت داده، مجموعه‌داده‌ی ۱۴۳۵ عددی خود را به ۷۱۷۵ عدد افزایش دادند و به کمک الگوریتم YOLOv3 مدل خود را توسعه دادند

در سال‌های اخیر توسعه پیوسته فناوری در بخش پهپادی<sup>۱</sup> (پرنده‌های هدایت‌پذیر از دور)، منجر به تولید پهپادهایی با اندازه‌های کوچک و وزن مناسب شده است. این امر مزایای بسیاری را برای استفاده از این دستگاه‌ها در عرصه‌های مختلف مانند بازرسی از بخش‌هایی که دسترسی انسان به آن سخت است، ارسال مرسوله توسط شرکت‌ها، عکس‌برداری و فیلم‌برداری از مناظر و... به همراه داشته است؛ اما در کنار تمام این مزایا، این دستگاه‌ها به نوبه خود دارای تهدیدات جدی در سطوح مختلف اجتماعی و امنیتی نیز هستند. از جمله مهم‌ترین تهدیدات این پهپادها در بخش امنیتی، می‌توان به پروازهای بدون مجوز در مناطق حفاظت‌شده و امنیتی جهت ثبت تصاویر و ضبط فیلم‌های غیرمجاز از این مناطق اشاره داشت. لذا تشخیص به موقع این دستگاه‌ها جهت انجام اقدامات سریع و مربوطه، ضروری است. در این راستا استفاده از سامانه‌های ضد هوایی سنگین و لیزری، جهت تشخیص به موقع این وسایل، می‌تواند برای سازمان‌های مربوطه هزینه‌زا باشد. لذا حرکت به سمت استفاده از فناوری‌های ارزان‌قیمت در این بخش توجیه‌پذیر است. از جمله این فناوری‌ها می‌توان هوش مصنوعی اشاره داشت.

در یک تعریف کلی، هوش مصنوعی برای توصیف رایانه‌هایی به کار می‌رود که برخی از فعالیت‌های شناختی وابسته به ذهن انسان مانند قدرت یادگیری و حل مسئله را به خوبی انجام می‌دهند [۱]. یکی از حوزه‌های هوش مصنوعی که در این زمینه می‌تواند مورد استفاده قرار گیرد، بینایی رایانه‌ای است که در یک تعریف ساده، به دنبال خودکارسازی وظایف سیستم چشم انسان است [۲]. با استقرار دوربین‌های اپتیکی<sup>۲</sup> و یا الکترونیکی<sup>۳</sup> در مناطق حفاظت‌شده و پردازش عکس‌ها و فریم‌های<sup>۴</sup> ورودی آن‌ها به کمک مدل‌های توسعه داده شده با الگوریتم‌های بینایی رایانه‌ای، می‌توان به صورت به موقع حضور پهپاد را در محل تشخیص داد و اقدامات مربوطه را در راستای ساقط سازی آن انجام داد.

الگوریتم‌های بینایی رایانه‌ای در جهت تشخیص اشیا، مورد استفاده قرار می‌گیرند. این الگوریتم‌ها به دودسته‌ی کلی تقسیم می‌شوند: دومرحله‌ای و یک‌مرحله‌ای. در دسته‌ی اول، شبکه‌های عصبی پیچشی منطقه محور<sup>۵</sup> وجود دارند که ابتدا در آنها یک جست‌وجوی انتخابی در تصویر با هدف ساخت تعداد زیادی ناحیه‌ی پیشنهادی<sup>۶</sup> انجام می‌شود. سپس شبکه عصبی پیچشی، ویژگی‌های هر ناحیه را استخراج می‌کند و در مرحله‌ی بعدی آن

<sup>۷</sup> Bounding Box

<sup>۸</sup> You Only Look Once

<sup>۹</sup> Single Shot Multibox Detector

<sup>۱۰</sup> mAP (Mean Average Precision)

<sup>۱۱</sup> Precision

<sup>۱۲</sup> Gatwick

<sup>۱</sup> UAV (Unmanned Aerial Vehicle)

<sup>۲</sup> Optical Viewfinder

<sup>۳</sup> Electronical Viewfinder

<sup>۴</sup> Frame

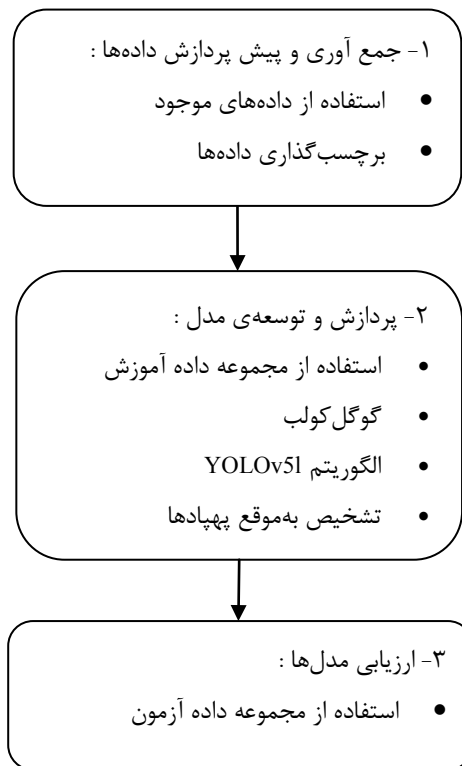
<sup>۵</sup> R-CNNs (Region-based Convolutional Neural Networks)

<sup>۶</sup> Region Proposal

ارائه‌شده است، جهت تشخیص به‌موقع پهپادهای چند بال چرخان استفاده شده است. الگوریتم YOLOv5<sup>1</sup> از سمت توسعه‌دهندگان اصلی، با بهینه‌ساز SGD ارائه شده است. در این پژوهش با اصلاح بهینه‌ساز این الگوریتم به Adam، یک مدل جهت تشخیص به‌موقع پهپادهای چند بال چرخان توسعه پیدا کرده است و عملکرد آن نیز با الگوریتم شامل بهینه‌ساز SGD و سایر الگوریتم‌های استفاده شده در پژوهش‌های پیشین مقایسه شده است. در این پژوهش از یک مجموعه داده ۱۰۰۴۶ عددی برای آموزش مدل‌ها استفاده شده است و عملکرد مدل‌ها توسط چهار مجموعه داده آزمون شامل: مجموعه آزمون معمولی، کم‌کیفیت، حالت شب و خاکستری مقیاس مورد ارزیابی قرار گرفته است. همچنین در این پژوهش جهت پردازش مدل‌ها از بستر گوگل کولب<sup>۵</sup> استفاده شده است.

## ۲. روش تحقیق

تحقیق انجام‌شده دارای سه مرحله کلی است. در ادامه، شکل (۱) مراحل انجام تحقیق را نمایش می‌دهد.



شکل ۱. نقشه راه انجام تحقیق.

در ادامه به‌صورت مختصر به شرح هر یک از مراحل پرداخته شده است.

به‌گونه‌ای که قادر به تشخیص پهپاد از فاصله‌ی ۴۰ متری بوده و ضرب اطمینان مدل نیز به‌صورت متوسط ۸۸/۹٪ به‌دست آمده است. در پژوهش دیگری، جمیل و همکاران [۱۱] با توصیف پهپادهای مخرب به‌عنوان دسته‌ای از پهپادها که بسته‌ها و اشیای خطرناک و مشکوک را حمل می‌کنند، یک ترانسفورماتور بینایی<sup>۱</sup> جهت تشخیص پهپادهای مخرب از غیر مخرب ارائه کردند و همچنین عملکرد مدل خود را نیز با چندین مدل یادگیری ماشین و مدل شبکه عصبی پیچشی عمیق مقایسه کردند که مدل مذکور با میزان صحت ۹۸/۳٪ بهترین عملکرد را دارا بود.

یکی از چالش‌های موجود در زمینه‌ی شناسایی پهپادها، تمایز آنها از پرندگان است. در این راستا، آپدهایی و همکاران [۱۲] در پژوهش خود جهت شناسایی به‌موقع سه نوع خاص از پهپادهای کوچک و همچنین پرندگان، مدلی را با یکپارچه‌سازی فن‌های پردازش سیگنال راداری و شبکه‌های عصبی پیچشی توسعه دادند. در این مدل، مجموعه داده از طریق سیگنال‌های راداری به‌دست‌آمده از پرندگان و پهپادها جمع‌آوری شده که همین امر، صحت مدل را به ۸۷/۵٪ رساند. صمدزادگان و همکاران [۱۳] نیز مدلی را به کمک الگوریتم YOLOv4 جهت شناسایی دو نوع از پهپادها و همچنین پرندگان توسعه دادند که برای این امر از یک مجموعه داده ۱۰۰۰۰ عددی استفاده شده است که در نهایت مدل قادر به تفکیک این دودسته پهپاد از پرندگان با میانگین صحت متوسط به مقدار ۸۴٪ بود. وجود مدل سبک‌تر جهت تشخیص هر چه سریع‌تر پهپادها نیز امری ضروری است در همین راستا شارما و همکاران [۱۴] با استفاده از شبیه‌ساز AirSim و یک رابط برنامه‌نویسی کاربردی<sup>۲</sup> به کمک پایتون<sup>۳</sup>، مجموعه داده موردنیاز خود را جمع‌آوری کردند و به کمک الگوریتم tiny-YOLOv4 مدل خود را توسعه دادند که دارای میانگین صحت متوسط ۷۷٪ است. در پژوهش دیگری نیز سون و همکاران [۱۵] شبکه عصبی جدیدی را به نام TIB-Net توسعه دادند. این شبکه جدید با بهره‌گیری از ساختار خاص به خود، قابلیت استخراج ویژگی مؤثرتری را از عکس‌های دارای اشیای کوچک برخوردار است و حجم مدل را نیز در یک سطح قابل قبولی نگه می‌دارد. عملکرد این مدل با چندین مدل دیگر مقایسه شد که در نهایت مدل مذکور به بیشترین معیار صحت به مقدار ۸۹/۲٪ دست پیدا کرد.

به دلیل اهمیت بالای تشخیص به‌موقع پهپادها در مناطق پدافندی و حفاظت‌شده در راستای تأمین امنیت این مناطق، با در نظرگیری پژوهش‌های پیشین، در این مقاله از جدیدترین نسخه الگوریتم YOLO با اندازه‌بزرگ، یعنی الگوریتم YOLOv5<sup>۱</sup> که با هدف عملکرد و سرعت پردازشی بهتر نسبت به سایر الگوریتم‌ها

<sup>1</sup> Vision Transformer

<sup>2</sup> API

<sup>3</sup> Python

<sup>4</sup> YOLO Version 5 Large

تمام داده‌های این پژوهش به کمک جستجو در سطح اینترنت، از طریق عکس‌های موجود و استخراج فریم از ویدئوها، جمع‌آوری شده است. تعداد داده‌های جمع‌آوری شده برای هر نوع در جدول (۲) مشخص شده است.

**جدول ۲.** تعداد داده‌های جمع‌آوری شده از هر دسته پهپاد بال چرخان.

نوع پهپاد	Quadrotor	Hexarotor	Octorotor	Octo Coax Wide
تعداد داده	۴۵۴۶	۲۰۰۰	۲۰۰۰	۱۵۰۰

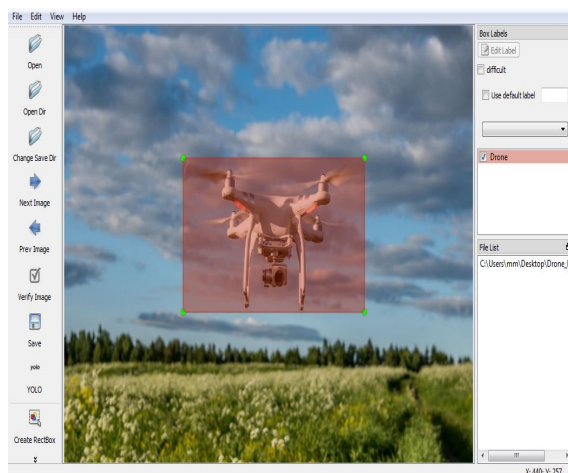
در ادامه جدول (۳) دسته‌بندی مجموعه داده‌های جمع‌آوری شده را نشان می‌دهد.

**جدول ۳.** تقسیم‌بندی داده‌های جمع‌آوری شده [۱۸].

تعداد	درصد	مجموعه داده
۸۰۳۶	۸۰	آموزش
۲۰۱۰	۲۰	اعتبارسنجی

داده‌های آموزش، جهت آموزش مستقیم مدل و داده‌های اعتبارسنجی، به‌منظور ارزیابی مدل پس از هر مرحله آموزش، استفاده می‌شوند.

آموزش مدل در این پژوهش به‌صورت نظارت شده<sup>۳</sup> انجام می‌شود، به این معنی که داده‌ها همراه با برچسب خود، جهت آموزش مدل به آن وارد می‌شوند. لذا نیاز است که تمام عکس‌های جمع‌آوری شده، برچسب‌گذاری<sup>۴</sup> شوند. جهت برچسب‌گذاری عکس‌ها از نرم‌افزار LabelIMG استفاده شده است. دلیل استفاده از این نرم‌افزار، محیط بصری زیبا و سهولت در استفاده از آن است. تمام عکس‌ها با نام Drone برچسب‌گذاری شده‌اند. در ادامه، شکل (۳) یک نمونه عکس که توسط این نرم‌افزار برچسب‌گذاری شده است را نمایش می‌دهد.



شکل ۳. عکس برچسب‌گذاری شده در نرم‌افزار LabelIMG [۱۹].

## ۱-۲. جمع‌آوری و پیش‌پردازش داده‌ها

پرنده‌های هدایت‌پذیر از دور، نوعی هواپیما هستند که بدون خلبان و از راه دور قابلیت پرواز و هدایت‌پذیری دارند. اکثر پهپادها به دودسته<sup>۱</sup> بال ثابت<sup>۲</sup> و چند بال چرخان<sup>۳</sup> تقسیم می‌شوند [۱۶].

در یک تعریف ساده پهپادهای بال ثابت مانند هواپیماهای معمولی از یک بدنه که بال‌ها به طور ساکن به آن متصل شده‌اند، تشکیل شده است و به کمک سرعت هوای روبه‌جلو، عملیات پرواز در آنها انجام می‌شود. در ادامه شکل (۲) یک نمونه از این پهپادها را نمایش می‌دهد.



شکل ۲. نمونه پهپاد بال ثابت [۱۷].

پهپادهای چند بال چرخان نیز از چندین بال چرخان جهت پرواز استفاده می‌کند [۱۳]. در ادامه جدول (۱) نمونه‌هایی از انواع پهپادهای چند بال چرخان را نمایش می‌دهد.

**جدول ۱.** دسته‌بندی پهپادهای بال چرخان [۱۳].

تعداد بال چرخان	نمونه	پهپاد
۴		Quadrotor
۶		Hexarotor
۸		Octorotor
۸ (۴ بال بالا و ۴ بال پایین)		Octo Coax Wide

به دلیل استفاده بیشتر از پهپادهای چند بال چرخان نسبت به پهپادهای بال ثابت، در این پژوهش جهت آموزش مدل یک مجموعه داده<sup>۴</sup> ۱۰۰۴۶ عددی عکس از چهار نوع پهپاد چند بال چرخان نمایش داده شده در جدول (۱) جمع‌آوری شده است.

<sup>۳</sup> Supervised

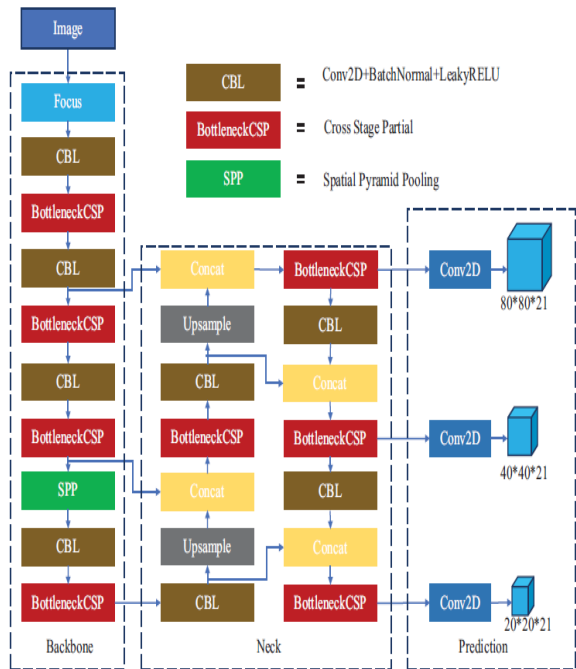
<sup>۴</sup> Labeling

<sup>۱</sup> Fixed Wing

<sup>۲</sup> Multi Rotor

می‌شوند و تا زمانی که نتیجه‌ای نزدیک‌تر به خروجی واقعی و با خطای کم‌تر به دست آید و یا تعداد معینی تکرار سپری شود، مراحل قبل تکرار می‌شوند [۲۵].

در ادامه شکل (۴) ساختار شبکه YOLOv5 را نمایش می‌دهد.



شکل ۴. ساختار کلی شبکه YOLOv5 [۲۶].

همان‌گونه که در شکل (۴) نشان داده شده است، شبکه YOLOv5 از سه بخش پشتیبان<sup>۶</sup>، گردن<sup>۷</sup> و خروجی تشکیل شده است. در قسمت پشتیبان، عکس ورودی با وضوح  $3^8 \times 640 \times 640$  وارد لایه Focus می‌شود. سپس با عملیات برش<sup>۹</sup> در ابتدا به یک نقشه ویژگی<sup>۱۰</sup> با ابعاد  $12 \times 320 \times 320$  تبدیل می‌شود، سپس بعد از عملیات کانولوشنی<sup>۱۱</sup> هسته  $3 \times 3$  تبدیل به یک نقشه ویژگی با ابعاد  $32 \times 320 \times 320$  می‌شود. ماژول CBL یک ماژول کانولوشنی پایه است. این ماژول بیانگر Conv2D + BatchNormal + LeakyRELU است.

ماژول BottleneckCSP بیشتر استخراج ویژگی بر روی نقشه ویژگی را انجام می‌دهد، این استخراج، اطلاعات از عکس را حاصل می‌کند. در مقایسه با سایر شبکه‌های عصبی کانولوشنی بزرگ‌مقیاس، ساختار BottleneckCSP می‌تواند گرادپان تکثیر اطلاعات در فرایند بهینه‌سازی شبکه عصبی کانولوشنی را کاهش

## ۲-۲. پردازش و توسعه مدل YOLOv5

در این پژوهش از الگوریتم YOLOv5 جهت آموزش مدل‌ها استفاده شده است. YOLO از جمله شبکه‌های پیچشی و الگوریتمی یک‌مرحله‌ای و یکپارچه برای تشخیص اشیا است که از طریق رسم کادرهای محصورکننده با امتیاز معین که بیان‌کننده احتمال وجود شیء در کادر است، به شناسایی آن شیء و برچسب مربوط به آن می‌پردازد. اولین نسخه این الگوریتم در سال ۲۰۱۶ منتشر شد و پس از آن YOLOv2 [۲۰]، YOLOv3 [۲۱]، YOLOv4 [۲۲] منتشر شدند و نسخه پنجم این الگوریتم نیز اولین بار ماه می سال ۲۰۲۰ در Github منتشر شد [۲۳]. YOLOv5 در چهار نسخه شامل YOLOv5m، YOLOv5s، YOLOv5x و YOLOv5l با عمق‌های مختلف عرضه شده است. هرچه شبکه‌ای عمیق‌تر باشد، تعداد لایه‌های بیشتری جهت آموزش در شبکه آن به کار رفته است و در نتیجه، آن شبکه ویژگی‌های بیشتری را می‌آموزد و البته کندتر نیز عمل می‌کند.

از بین الگوریتم‌های نسخه پنجم YOLO در این پژوهش از الگوریتم YOLOv5l استفاده شده است. دلیل انتخاب این الگوریتم را می‌توان عمق زیاد و در نتیجه آن، دقت بالا در پردازش تصاویر دانست. به صورت کلی از جمله مهم‌ترین دلایل استفاده از سری YOLO را می‌توان در زبان مورد توسعه آن که پایتون است جویا شد، چرا که زبان پایتون به دلیل دارا بودن کتابخانه‌های بسیار زیاد، قابلیت توسعه و یکپارچه‌سازی این مدل برای استفاده در سایر بخش‌ها و دستگاه‌ها را به راحتی میسر می‌سازد. همچنین این نسخه، از فریم‌ورک<sup>۱</sup> پای‌تورچ<sup>۲</sup> بهره می‌برد که در مقایسه با فریم‌ورک دارک نت<sup>۳</sup> به کار گرفته شده در نسخه‌های قبلی، قوی‌تر و دارای پشتیبانی به‌شدت بهتری است که توسعه و پیشرفت روزافزون را برای این الگوریتم به همراه دارد.

آموزش در سری الگوریتم‌های YOLO مانند سایر شبکه‌های عصبی پیچشی از طریق روش پس انتشار خطا<sup>۴</sup> انجام می‌گیرد [۲۴]. در این روش، ابتدا داده‌های آموزشی به شبکه داده می‌شوند و سپس ویژگی‌های آن‌ها در حرکتی روبه‌جلو<sup>۵</sup>، لایه به لایه شبکه عمیق را طی می‌کنند تا جهت شناسایی، آموزش داده شوند. آموزش به‌گونه‌ای انجام می‌پذیرد که یک تابع خطا حداقل گردد. پس از آن جهت ارزیابی عملکرد شبکه آموزش‌یافته، شبکه با داده‌های اعتبارسنجی بررسی می‌شود و در صورتی که خطای آموزش زیاد باشد، مجدد داده‌ها از طریق لایه‌ها برگشت داده

<sup>6</sup> Backbone

<sup>7</sup> Neck

<sup>8</sup> Resolution

<sup>9</sup> Slicing operation

<sup>10</sup> Feature Map

<sup>11</sup> Kernel

<sup>1</sup> Framework

<sup>2</sup> Pytorch

<sup>3</sup> DarkNet

<sup>4</sup> Back Propagation of Error

<sup>5</sup> Feed Forward

سخت‌افزار ارائه شده توسط گوگل کولب جهت پردازش داده‌ها در این پژوهش شامل: ۱۳ گیگابایت رم<sup>۱۰</sup> و واحد پردازش گرافیکی از نوع TeslaK80 با ۱۲ گیگابایت حافظه و ۴۰ گیگابایت فضای هارددیسک جهت ذخیره‌سازی است.

پیش از شروع پردازش هر مدل، لازم است که ابر پارامتر<sup>۱۵</sup>‌های مدل تنظیم شود. در ادامه جدول (۴) ابر پارامترهای تنظیم شده جهت پردازش مدل‌ها را نشان می‌دهد. این ابر پارامترها در هر دو مدل یکسان در نظر گرفته شده است.

**جدول ۴.** ابر پارامترهای تنظیم شده جهت پردازش مدل‌ها [۲۵].

تعداد دوره	اندازه دسته	اندازه عکس	مدل
۳۰۰	۱۶	۶۴۰ × ۶۴۰	YOLOv5l

اندازه عکس<sup>۱۶</sup>، بیانگر اندازه عکس ورودی به مدل است که مدل با استفاده از روش‌های خاصی آن را مقیاس‌بندی و به صورت مربعی تغییر داده و آموزش را شروع می‌کند، باید توجه داشت که هر چه اندازه عکس بزرگ‌تر باشد، توانایی مدل جهت تشخیص اشیا کوچک‌تر ارتقا می‌یابد. لذا در این مدل بیشترین مقدار ممکن برای اندازه عکس در نظر گرفته شده است. اندازه دسته<sup>۱۷</sup>، اشاره به تعداد داده‌های موجود در هر دسته جهت ورود به مدل دارد که به دلیل بزرگ بودن مجموعه داده آموزش، وارد کردن به یک‌باره تمام داده‌ها به مدل سخت می‌شود، لذا داده‌ها به دسته‌های کوچک‌تر با تعداد مشخص داده تقسیم می‌شوند. جهت پردازش بهتر مدل، مقدار پیشنهادی اندازه دسته توسط مدل، در نظر گرفته شده است. دوره<sup>۱۸</sup> نیز اشاره به تعداد دفعاتی دارد که تمام داده‌ها به صورت کامل در شبکه پردازش شده‌اند.

**توسعه مدل:** پس از کامل شدن آموزش مدل، فایل وزن‌های آموزش دیده مدل‌ها دریافت شده و سپس توسعه مدل‌ها در قالب زبان برنامه‌نویسی پایتون، جهت تشخیص به موقع پهناده‌ها توسط فیلم‌های ورودی از طریق دوربین‌های متصل به سیستم، انجام شده است. برای این امر از کتابخانه‌های OpenCV و Pytorch در بستر پایتون کمک گرفته شده است. در ادامه، شکل (۵) ساختار عملکردی مدل‌ها را پس از توسعه نشان می‌دهد. به طور کلی عکس و یا فیلم وارد مدل می‌شود و بلافاصله پردازش انجام می‌شود و خروجی آن، همان عکس یا فیلم همراه با اطلاعات شامل موقعیت پهناده تشخیص داده شده و عدد ضریب اطمینان تشخیص صورت گرفته شده است.

دهد. با تنظیم عرض و عمق ماژول BottleneckCSP در ساختار شبکه YOLOv5 می‌توان به چهار مدل YOLOv5m، YOLOv5s، YOLOv5l و YOLOv5x دست پیدا کرد. ماژول SPP نیز عمدتاً زمینه‌های پذیرشی مدل را افزایش داده و ویژگی‌های با مقیاس‌های مختلف را حصول می‌کند.

YOLOv5 همچنین یک ساختار هرمی<sup>۱</sup> پایین - بالا<sup>۲</sup> بر مبنای ساختار FPN تولید می‌کند. با این عملیات ترکیب، لایه FPN ویژگی‌های معنایی قوی را از بالا به پایین انتقال می‌دهد. ترکیب ویژگی تجمع از لایه‌های ویژگی مختلف، برای بهبود توانایی شبکه در تشخیص اشیا با مقیاس‌های مختلف است. در قسمت آخر نیز، خروجی شامل نتایج طبقه‌بندی و مختصات اشیا تشخیص داده شده است [۲۶].

**بهینه‌سازی:** یک تابع است که ویژگی‌های شبکه عصبی مانند وزن‌ها و نرخ خطا را اصلاح می‌کند. الگوریتم YOLOv5l به صورت پیش فرض از سمت توسعه دهندگان با بهینه‌سازی گرادینت کاهشی تصادفی<sup>۳</sup> (SGD) ارائه شده است. این بهینه‌سازی، الگوریتم تکراری مرتبه اول برای یافتن کمینه‌های موضعی<sup>۴</sup> روی یک تابع مشتق پذیر است. این بهینه‌سازی از یک نرخ یادگیری واحد برای به روزرسانی تمام وزن‌ها استفاده می‌کند و در طول فرایند آموزش مدل دچار تغییر نمی‌شود [۲۷]. در این پژوهش در کنار مدل توسعه داده شده با این بهینه‌سازی، یک مدل دیگر با بهینه‌سازی Adam که یک تعمیم یافته از بهینه‌سازی گرادینت تصادفی کاهشی است، توسعه داده شده است. بهینه‌سازی Adam پس از محاسبه اولیه نرخ یادگیری برای هر یک از وزن‌ها، برخلاف SGD با شروع فرایند یادگیری نرخ یادگیری را به صورت جداگانه برای هر کدام از وزن‌ها تطبیق می‌دهد.

**پردازش مدل:** معمولاً پردازش داده‌ها در مسائل یادگیری ماشین، به سخت‌افزارهای قدرتمند و مجهز به واحد پردازش گرافیکی<sup>۷</sup> با سرعت بالا نیاز دارد، به دلیل محدودیت سخت‌افزاری موجود، جهت پردازش داده‌ها از بستر گوگل کولب استفاده شده است. گوگل کولب یک محیط توسعه یکپارچه<sup>۸</sup> رایگان از شرکت گوگل است که توسعه و آموزش مدل‌های مبتنی بر هوش مصنوعی را به دلیل استفاده از واحدهای پردازش گرافیکی، پشتیبانی می‌کند. این بستر آنلاین است و از زبان برنامه‌نویسی پایتون تحت محیط ژوپیتر نوت‌بوک<sup>۹</sup> استفاده می‌کند.

<sup>1</sup> Pyramid

<sup>2</sup> Bottom-Up

<sup>3</sup> Stochastic gradient descent

<sup>4</sup> Local Minimum

<sup>7</sup> GPU (Graphics Processing Unit)

<sup>8</sup> Integrated Development Environment

<sup>9</sup> Jupyter Notebook

<sup>10</sup> Ram

<sup>15</sup> Hyper parameter

<sup>16</sup> Image Size

<sup>17</sup> Batch Size

<sup>18</sup> Epoch



بیشتر، به معنای تشخیص درست بیشتر توسط مدل و معیار یادآوری بیشتر، به معنای از دست رفتن کمتر تشخیص تصاویر حاوی شیء مورد نظر است.

**معیار F1 Score:** این معیار میزان تعادل بین دو معیار صحت و یادآوری را نشان می‌دهد که طبق معادله (۳)، از میانگین همساز این دو معیار ذکر شده، محاسبه می‌شود [۲۸].

$$F1 = 2 \times \frac{P \times R}{P + R} \quad (3)$$

کمینه مقدار این معیار، عدد صفر و بیشینه مقدار آن عدد یک است که افزایش آن به معنی بالابودن دو معیار صحت و پوشش و عملکرد مناسب مدل است و کمبود آن معنی برعکس می‌دهد.

**معیار میانگین صحت متوسط:** این معیار از محبوب‌ترین شاخص‌ها برای اندازه‌گیری میزان دقت و عملکرد مدل در تشخیص اشیا است. در یک تعریف کلی، این معیار به دنبال محاسبه مقدار میانگین معیار صحت برای مقدار معیار یادآوری از مقدار صفر تا به مقدار یک است که از طریق محاسبه سطح زیر منحنی صحت - یادآوری، طبق معادله (۴) محاسبه می‌شود [۲۸].

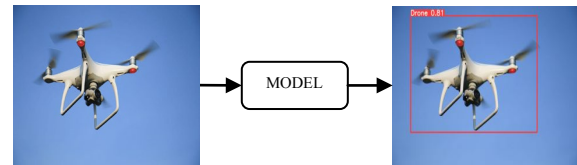
$$AP = \int_0^1 P(R) dR \quad (4)$$

معیار میانگین صحت متوسط، نیز از تقسیم تعداد کلاس‌های موجود بر معیار میانگین صحت، طبق معادله (۵) به دست می‌آید [۲۸].

$$mAP = \frac{1}{n} \int_0^1 P(R) dR \quad (5)$$

به دلیل این که مدل فقط یک کلاس از اشیا را تشخیص می‌دهد، پس این معیار برابر با معیار میانگین صحت است.

باید دقت داشت که در الگوریتم YOLOv5 دو نوع معیار میانگین صحت متوسط گزارش می‌شود. معیار اول با مقدار مؤلفه<sup>۸</sup> IOU برابر ۰/۵ و مؤلفه<sup>۹</sup> دوم با مقدار IOU از ۰/۵ تا ۰/۹۵ با گام ۰/۰۵، مقدار میانگین صحت برای هر کلاس از اشیا مورد تشخیص را محاسبه کرده و در نهایت برای تمام کلاس‌ها میانگین می‌گیرد. لازم به ذکر است که IOU یک مؤلفه ارزیابی است که از تقسیم اشتراک سطح تشخیص داده شده توسط مدل از شیء مورد نظر در تصویر با سطح واقعی شیء مورد نظر در تصویر بر اجتماع آن‌ها به دست می‌آید. در نهایت باید عنوان کرد که مقدار بیشتر معیار میانگین صحت متوسط بیانگر دقت بیشتر مدل در شناسایی است [۲۸].



شکل ۵. ساختار عملکردی مدل پس از توسعه.

### ۳. نتایج و بحث

#### ۱-۳. معیارهای ارزیابی مدل

مؤلفه‌هایی که جهت ارزیابی عملکرد مدل به کاررفته است به شرح زیر است: معیار صحت<sup>۱</sup>، معیار یادآوری<sup>۲</sup>، معیار F1 Score و معیار میانگین صحت متوسط<sup>۳</sup> در ادامه هر کدام به صورت مختصر شرح داده خواهند شد:

**معیارهای صحت و یادآوری:** این دو معیار از مهم‌ترین شاخص‌های پایه و متداول در ارزیابی مدل‌های تشخیص اشیا هستند. در ادامه معادلات (۱) و (۲) نحوه محاسبه این دو معیار را نشان می‌دهند [۲۸].

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

در معادلات بالا مؤلفه<sup>۴</sup> TP اشاره به تعداد مثبت صحیح<sup>۴</sup> دارد که به صورت درست، وجود شیء مورد نظر را پیش‌بینی کرده است. همچنین مؤلفه<sup>۵</sup> TN نیز اشاره به تعداد منفی صحیح<sup>۵</sup> دارد که به صورت درست، عدم وجود شیء مورد نظر را پیش‌بینی کرده است. مؤلفه<sup>۶</sup> FP اشاره به تعداد مثبت کاذب<sup>۶</sup> دارد که به صورت نادرست، وجود شیء مورد نظر را در تصویر پیش‌بینی کرده است و مؤلفه<sup>۷</sup> FN نیز اشاره به تعداد منفی کاذب<sup>۷</sup> دارد که به صورت نادرست، عدم وجود شیء مورد نظر را در تصویر پیش‌بینی کرده است.

به صورت کلی معیار صحت به این سؤال پاسخ می‌دهد که در میان تمام تصاویری که وجود شیء مورد نظر در آن‌ها (چه به صورت درست و چه نادرست) توسط مدل تشخیص داده شده است، در چه تعداد تصویر، به درستی وجود شیء پیش‌بینی شده است. معیار یادآوری نیز به دنبال پاسخ به این سؤال است که چه تعداد از تصاویری که باید در آن‌ها شیء مورد نظر تشخیص داده شود، توسط مدل، تشخیص داده شده است؛ بنابراین معیار صحت

<sup>1</sup> Precision

<sup>2</sup> Recall

<sup>3</sup> mAP (Mean Average Precision)

<sup>4</sup> True Positive

<sup>5</sup> True Negative

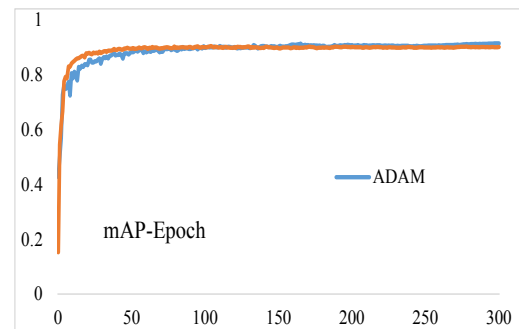
<sup>6</sup> False Positive

<sup>7</sup> False Negative

<sup>8</sup> Intersection Over Union

## ۳-۲. ارزیابی مدل‌ها و تحلیل نتایج

به طور معمول، جهت ارزیابی آموزش مدل‌ها از نتایج به دست آمده در طول آموزش مدل استفاده می‌شود. در این پژوهش برای درک بهتر نتایج، از نمودار استفاده شده است. نمودار همگرایی میانگین صحت متوسط برحسب تعداد دوره، از جمله نمودارهای مورد توجه برای ارزیابی فرایند آموزش مدل‌ها است که از طریق داده‌های اعتبارسنجی در طی فرایند آموزش مدل‌ها حاصل می‌شود. در ادامه شکل (۶)، این نمودار را ارائه می‌کند.



شکل ۶. نمودار همگرایی میانگین صحت متوسط برحسب تعداد دوره [۲۹].

با توجه به همگرایی که میانگین صحت متوسط در هر دو مدل در تعداد ۳۰۰ دوره انجام داده‌اند، می‌توان انتخاب این عدد را برای تعداد دوره و همچنین عملکرد سخت‌افزاری و به صورت کلی آموزش انجام شده بر روی مدل‌ها را مناسب دانست. مقدار تقریبی میانگین صحت متوسط در هر دو مدل در آخرین دوره ۰/۹۱ است.

در ادامه جهت ارزیابی نهایی مدل‌ها، یک مجموعه داده آزمون<sup>۱</sup> شامل ۱۰۰۰ عدد عکس جدید از پهپادها جمع‌آوری شده است. از آنجایی که داده‌های پردازشی مدل‌ها عکس است، هر عاملی که باعث تغییر عکس‌ها شود، می‌تواند در عملکرد مدل‌ها تأثیرگذار باشد. از جمله این عوامل می‌توان کیفیت، رنگ‌بندی و میزان روشنایی عکس را نام برد. لذا در راستای ارزیابی منصفانه و شبیه‌سازی شرایط مختلف، سه مجموعه داده جدید از مجموعه داده آزمون ساخته شده است. این مجموعه داده‌ها شامل: مجموعه آزمون کم‌حجم، مجموعه آزمون حالت شب و مجموعه آزمون خاکستری مقیاس است.

در ادامه جدول (۵) نمونه داده‌های مجموعه آزمون معمولی و کم‌حجم را نمایش می‌دهد. در مجموعه آزمون کم‌حجم، کیفیت و حجم عکس‌ها به طور قابل توجهی به کمک نرم‌افزار سزیم<sup>۲</sup> کاهش پیدا کرده است.

## جدول ۵. نمونه داده برای مجموعه آزمون معمولی و کم‌حجم.

	مجموعه آزمون معمولی	مجموعه آزمون کم‌حجم
نمونه داده		

در ادامه نیز جدول (۶) نمونه داده‌هایی از مجموعه آزمون حالت شب و خاکستری مقیاس را نمایش می‌دهد. در مجموعه آزمون حالت شب سطح رنگ آبی عکس‌ها به میزان چشمگیری افزایش پیدا کرده است و در مجموعه آزمون خاکستری مقیاس رنگ عکس‌ها به سیاه‌وسفید تغییر پیدا کرده است.

## جدول ۶. نمونه داده برای مجموعه آزمون حالت شب و خاکستری مقیاس.

	مجموعه آزمون حالت شب	مجموعه آزمون خاکستری مقیاس
نمونه داده		

همچنین جهت ارزیابی عملکرد مدل در تشخیص هم‌زمان چندین پهپاد، یک مجموعه داده<sup>۱۰۰</sup> عددی متشکل از عکس‌هایی که شامل چندین پهپاد می‌شوند نیز جمع‌آوری گردیده است. در ادامه جدول (۷) نتایج حاصل شده از ارزیابی مدل‌ها بر روی مجموعه آزمون‌های مختلف را نمایش می‌دهد.

## جدول ۷. نتایج حاصل از ارزیابی مدل‌ها بر روی مجموعه آزمون‌ها [۱۳].

YOLOv5l-Adam				
مجموعه آزمون	R	P	mAP@0.5	mAP@0.5:0.95
معمولی	۰/۹۹۴	۰/۹۹۳	۰/۹۹۴	۰/۸۴۶
کم‌حجم	۰/۹۹۴	۰/۹۹۳	۰/۹۹۴	۰/۸۰۹
حالت شب	۰/۹۹۳	۰/۹۹	۰/۹۹۴	۰/۸
خاکستری مقیاس	۰/۹۹۴	۰/۹۹۲	۰/۹۹۴	۰/۸۱۱
چندین پهپاد	۰/۹۹۳	۰/۹۱۱	۰/۹۶۵	۰/۸۲۸
YOLOv5l-SGD				
معمولی	۰/۹۹۴	۰/۹۸۳	۰/۹۹۲	۰/۸۲۱
کم‌حجم	۰/۹۹۳	۰/۹۸۵	۰/۹۹۲	۰/۷۹۱
حالت شب	۰/۹۸۷	۰/۹۷۸	۰/۹۹۲	۰/۷۸۳
خاکستری مقیاس	۰/۹۹۲	۰/۹۸۳	۰/۹۹۲	۰/۷۸۸
چندین پهپاد	۰/۹۹۲	۰/۹۰۶	۰/۹۵۹	۰/۸۱۶

همان گونه که از نتایج ارائه شده در جدول (۷) مشخص است، می‌توان عملکرد الگوریتم YOLOv5l را در برابر چالش‌های اصلی پردازش تصویر، یعنی تغییر رنگ، تغییر روشنایی و کاهش کیفیت

<sup>1</sup> Test set

<sup>2</sup> Caesium



یکی از معیارهایی که برای ارزیابی سرعت عملکردی مدل وجود دارد، معیار زمان استنتاج<sup>۱</sup> است. زمان استنتاج مدت‌زمانی است که یک عکس یا یک فریم از ویدئو، توسط مدل مورد پردازش قرار می‌گیرد. به کمک رابطه (۶) می‌توان نیز می‌توان مقدار نرخ فریم پردازشی<sup>۲</sup> مدل را به دست آورد [۳۰].

$$FPS = \frac{1}{\text{Inference Time}} \quad (6)$$

در این پژوهش جهت ارزیابی سرعت عملکردی مدل‌ها، از یک ویدئو که در آن یک پهپاد در حال پرواز وجود دارد، استفاده شده است. ویدئوهای استفاده‌شده دارای کیفیت‌های HD (با رزولوشن ۱۲۸۰×۷۲۰) و FULL HD (با رزولوشن ۱۹۲۰×۱۰۸۰) هستند. جدول (۹) نتایج حاصل‌شده از ارزیابی سرعت پردازش مدل‌ها را نمایش می‌دهد.

جدول ۹. نتایج ارزیابی سرعت عملکردی مدل‌ها [۳۱].

HD Video		
YOLOv5l	Inference Time(ms)	FPS
Adam	۱۶/۳	۶۱
SGD	۱۶/۷	۶۰
Full HD Video		
Adam	۱۵/۱	۶۶
SGD	۱۶/۳	۶۱

نتایج جدول (۹) نشان از عملکرد سریع‌تر الگوریتم YOLOv5l با بهینه‌ساز Adam نسبت به SGD دارد.

با توجه به اینکه اکثر پژوهش‌ها در حوزه پردازش تصاویر نرخ پردازشی بالای ۵ فریم بر ثانیه را پردازش به‌موقع می‌دانند. عملکرد پردازشی هر دو مدل را نیز می‌توان به‌موقع دانست [۳۲].

در ادامه جدول (۱۰) نمونه‌ای از تشخیص انجام شده بر روی هر ویدئو را نمایش می‌دهد. جهت مقایسه عملکرد مدل‌های پژوهش حاضر با سایر روش‌ها، در جدول (۱۱) نتایج چندین پژوهش اخیر انجام‌شده به کمک نسخه‌های قبلی الگوریتم YOLO و همچنین سایر الگوریتم‌های بینایی رایانه‌ای، آورده شده است.

همان‌گونه که از مقایسه نتایج جداول (۷) و (۱۱) پیداست، به‌طور کلی میانگین صحت متوسط که به‌عنوان معیار ارزیابی استاندارد، هر دو عامل صحت و یادآوری را هم‌زمان در نظر می‌گیرد، در مدل‌های توسعه داده شده در این پژوهش (که از الگوریتم YOLOv5l بهره‌برده‌اند) نسبت سایر الگوریتم‌ها بیشتر

محسوس عکس‌ها، مناسب دانست. همچنین باید توجه داشت که تغییر بهینه‌ساز الگوریتم YOLOv5l از SGD به Adam در تمام مجموعه‌آزمون‌ها، باعث بهبود در میانگین صحت متوسط، صحت و یادآوری در تشخیص پهپادها شده است که در مجموع بیانگر افزایش تعداد تشخیص‌های درست انجام‌شده توسط مدل پیشنهادی و مؤثر بودن نوآوری در نظر گرفته شده است. در ادامه جدول (۸) نمونه‌هایی از تشخیص صورت‌گرفته بر روی داده‌های مختلف را توسط هر دو مدل نمایش می‌دهد.

جدول ۸. نمونه تشخیص‌های انجام‌شده توسط مدل‌ها.

دادهٔ آزمون	YOLOv5l-Adam	YOLOv5l-SGD
معمولی		
کم حجم		
حالت شب		
خاکستری مقیاس		
چندین پهپاد		
روز برفی		
پس‌زمینه شلوغ		
حضور پرنده		

همان‌طور که در جدول (۸) نشان داده شده است، در کنار تشخیص‌های صورت گرفته‌شده بر روی مجموعه‌آزمون‌های مختلف، سه نمونه عکس دیگر شامل: وجود پهپاد در روز برفی، پهپاد با پس‌زمینه شلوغ و پهپاد در حضور پرنده برای ارزیابی بهتر مدل‌ها آورده شده است که تشخیص درستی برای هر مورد توسط مدل‌ها انجام شده است.

<sup>۱</sup> Inference Time



<sup>۲</sup> FPS (Frame Per Second)

#### ۴. نتیجه‌گیری

پیشرفت فناوری در حوزه پهبادی موجب ساخت پهبادهای کوچک و ارزان قیمت شده است. این دسته پهبادهای که بیشتر در گروه پهبادهای چند بال چرخان قرار می‌گیرند، با انجام پروازهای غیرمجاز در مناطق حفاظتی شده و امنیتی می‌توانند باعث ایجاد تهدید در این مناطق شوند. لذا شناسایی و تشخیص به‌موقع این پهبادهای جهت انجام اقدامات مربوطه، ضروری است. به همین دلیل در این مقاله به کمک نسخه جدید الگوریتم YOLO یعنی YOLOv5 مدل جهت تشخیص به‌موقع این دسته از پهبادهای توسعه داده شده است. این الگوریتم از سمت توسعه‌دهندگان با بهینه‌سازی SGD ارائه شده است. در این مقاله در توسعه مدل، بهینه‌سازی الگوریتم به Adam تغییر یافته است و عملکرد آن با مدل توسعه دیده‌شده با بهینه‌سازی SGD و پژوهش‌های پیشین مورد ارزیابی قرار گرفته است. جهت آموزش مدل‌ها از یک مجموعه داده ۱۰۰۴۶ عددی عکس از پهبادهای چند بال چرخان استفاده شده است. در این مقاله برخلاف پژوهش‌های پیشین، جهت آزمون منصفانه مدل‌ها از چهار مجموعه آزمون ۱۰۰۰ عددی شامل مجموعه آزمون معمولی، کم‌حجم، حالت شب و خاکستری مقیاس که همگی در رنگ‌بندی، روشنایی و کیفیت دچار تغییر محسوس شده‌اند و یک مجموعه آزمون ۱۰۰ عددی از عکس‌های شامل چندین پهباد، استفاده شده است و نتایج حاصل شده با پژوهش‌های پیشین مقایسه شده است که در نهایت، مدل توسعه داده شده با بهینه‌سازی Adam عملکرد و سرعت بهتری نسبت به مدل توسعه داده شده با بهینه‌سازی SGD و همچنین الگوریتم‌های استفاده شده در پژوهش‌های پیشین داشته است. با استقرار دوربین‌های اپتیکی و یا الکترونیکی در مناطق حفاظت‌شده و امنیتی و استفاده از فریم‌های خروجی آنها به‌عنوان ورودی مدل‌های پژوهش حاضر، می‌توان تشخیص به‌موقع پهبادهای چند بال چرخان را به‌صورت خودکار توسط سامانه‌های پدافندی و ضدهوایی در مناطق محافظت‌شده انجام داد. همچنین با توسعه سخت‌افزاری و مکانیکی سامانه‌های ضدهوایی در کنار مدل‌های حاضر، می‌توان به سامانه‌های ضدهوایی کاملاً هوشمند که قابلیت تشخیص به‌موقع و ساقط‌سازی خودکار پهبادهای چند بال چرخان را دارا هستند، دست پیدا کرد. این امر باعث حذف اپراتور و نیروی انسانی از پشت این سامانه‌ها می‌شود که در نتیجه آن، مدت‌زمان و خطا در ساقط‌سازی این پهبادهای نیز کاهش پیدا خواهد کرد. در پژوهش‌های آینده می‌توان مدل جهت تشخیص به‌موقع پهبادهای بال ثابت نیز توسعه داد. همچنین با به‌کارگیری نسخه عمیق‌تر الگوریتم YOLOv5 یعنی نسخه YOLOv5x و جمع‌آوری عکس از پهبادهای که با دوربین‌های دید در شب گرفته شده است، می‌توان مدل را جهت تشخیص بهتر پهبادهای در شب ارتقا داد.

است که بیانگر تشخیص درست تعداد بیشتر پهباد است. از نظر کمی نیز پژوهش‌های پیشین تنها محدود به مجموعه آزمون‌هایی با داده‌های معمولی و بدون تغییر در کیفیت و رنگ‌بندی عکس‌ها شده‌اند؛ اما در پژوهش حاضر چندین مجموعه آزمون مختلف جهت ارزیابی مدل‌ها مورد استفاده قرار گرفته که هر دو مدل عملکرد مناسبی را از خود نشان داده‌اند. در بررسی دقیق‌تر پژوهش [۱۴] که از الگوریتم YOLOv4 بهره می‌برد نسبت به الگوریتم YOLOv5 با بهینه‌سازی SGD در ارزیابی با مجموعه آزمون معمولی عملکرد بهتر داشته است، اما در مقایسه با الگوریتم YOLOv5 با بهینه‌سازی Adam که نوآوری اصلی این پژوهش است، عملکرد به نسبت ضعیف‌تری داشته است.

جدول ۱۰. نمونه تشخیص انجام‌شده بر روی ویدئوها.

نمونه تشخیص انجام‌شده	
	Full HD Video
	HD Video

جدول ۱۱. خلاصه نتایج پژوهش‌های پیشین.

مرجع	تعداد داده آزمون	الگوریتم و ابزار توسعه مدل	p	mAP	FPS
		ZF	-	۰/۶۱	-
[۵]	۲۷۲۷	VGG16	-	۰/۶۶	-
		VGG_CNN_M_1024	-	۰/۶	-
		Faster R-CNN with ResNet-101	-	۰/۴۹	-
[۲۳]	۸۷۷	Faster R-CNN with Inceptionv2	-	۰/۳۵	-
		SSD	-	۰/۱۵	-
[۲۴]	۹۹۴	Faster R-CNN	-	۰/۸۰۶۹	-
[۶]	-	Fine tuning CNN (YOLOv2)	۹۰	-	۱۵
[۳]	۲۸۰	Improved YOLOv3	۸۹	۰/۳۷۴۱	۵۶/۳
[۱۳]	۳۰۰۰	YOLOv4	۷۶-۹۰	۰/۸۴	-
[۱۴]	۳۰۰	Tiny-YOLOv4	-	۰/۷۷	۶-۴

در خصوص سرعت پردازشی، اکثر پژوهش‌های پیشین تحلیلی نسبت به الگوریتم‌های مورد توسعه خود نشان انجام نداده‌اند، اما از مقایسه نتایج جدول (۹) با چند مورد گزارش شده در جدول (۱۱) می‌توان به عملکرد سریع‌تر الگوریتم YOLOv5 با به خصوص با بهینه‌سازی Adam نسبت سایر الگوریتم‌ها پی برد.

## ۵. مرجع‌ها

- Swarms for Forest Fire Monitoring”; *Aerosp. Sci. Technol.* 2022,123,107494. <https://doi.org/10.1016/j.ast.2022.107494>.
- [17] Vourtsis, C.; Rochel, V. C.; Serrano, F. R.; Stewart, W.; Floreano, D. “Insect Inspired Self-Righting for Fixed-Wing Drones”; *IEEE Robot. Autom. Lett.* 2021, 6, 6805-6812. <https://doi.org/10.1109/LRA.2021.3096159>.
- [18] Dietterich, T. G. “Approximate Statistical Tests for Comparing Supervised Classification Learning Algorithms”; *Neural Comput.* 1998, 10, 1895-1923 <https://doi.org/10.1162/089976698300017197>.
- [19] Vostrikov, A.; Chernyshev, S. “Training Sample Generation Software”; *Intell. Decis. Technol.* 2019, 145-151. [https://doi.org/10.1007/978-981-13-8303-8\\_13](https://doi.org/10.1007/978-981-13-8303-8_13).
- [20] Redmon, J.; Farhadi, A. “YOLO9000: Better, Faster, Stronger”; *IEEE Conf. CVPR.* 2017, 7263-7271 <https://doi.org/10.48550/arXiv.1804.02767>.
- [21] Redmon, J.; Farhadi, A. “Yolov3: An Incremental Improvement”; *ArXiv.org.* 2018, 1804-2767. <https://doi.org/10.48550/arXiv.180402768>.
- [22] Bochkovskiy, A.; Wang, C.-Y.; Liao, H.-Y. M. “Yolov4: Optimal Speed and Accuracy of Object Detection”; *ArXiv.org.* 2020, 10934. <https://doi.org/10.45550/arXiv.200410934>.
- [23] Jocher, G. et al. “NanoCode012, C. A., “Ultralytics/yolov5: v5. YOLOv5 - 1280 Models, AWS, Supervise.ly and YouTube integrations”; *Zenodo* 2021, 0-6. <https://doi.org/10.1109/78.134446>.
- [24] Leung, H.; Haykin, S. “The Complex Backpropagation Algorithm”; *IEEE Trans. Signal Process.* 1991, 39, 2101-2104. <https://doi.org/10.1109/78.134446>.
- [25] Falahat, S.; Karami, A. “Maize Tassel Detection and Counting Using Deep Learning Techniques”; *J. Agric. Mach.* 2022, 2228-6829. <https://doi.org/10.1109/78.134446>.
- [26] Zhou, F.; Zhao, H.; Nie, Z. “Safety Helmet Detection Based on YOLOv5”; *IEEE Int. Conf. Power Electron. Control. Autom.* 2021, 6-11. <https://doi.org/10.1109/ICPECA51329.2021.9362711>.
- [27] Bottou, L.; Bousquet, O. “The Tradeoffs of Large Scale Learning”; *Adv. Neural Inf. Process. Syst.* 2007, 20.
- [28] Padilla, R.; Netto, S. L.; Da Silva, E. A. “A Survey on Performance Metrics for Object-Detection Algorithms”; *IEEE Int. Conf. Syst. Signal. Image Process.* 2020, 237-242. <https://doi.org/10.1109/IWSSIP48289.2020.9145130>.
- [29] Reddy, A.; Indragandhi, V.; Ravi, L.; Subramaniaswamy, V. “Detection of Cracks and Damage in Wind Turbine Blades Using Artificial Intelligence-Based Image Analytics”; *Meas.* 2019, 147, 106823. <https://doi.org/10.1016/j.measurement.2019.07.051>.
- [30] Bai, L.; Lyu, Y.; Huang, X. “Roadnet-rt: High Throughput CNN Architecture and SOC Design for Real-Time Road Segmentation”; *IEEE Trans. Circuits Syst. I, Reg. Papers.* 2020, 68, 704-714. <https://doi.org/10.1109/TCSI.2020.3038139>.
- [31] Deng, W.; Mou, Y.; Kashiwa, T.; Escalera, S.; Nagai, K.; Nakayama, K.; Matsuo, Y.; Prendinger, H. “Vision Based Pixel-Level Bridge Structural Damage Detection Using a Link ASPP Network”; *Autom. Constr.* 2020, 110, 102973. <https://doi.org/10.1016/j.autcon.2019.102973>.
- [32] Nath, N. D.; Behzadan, A. H.; Paal, S. G. “Deep Learning for Site Safety: Real-Time Detection of Personal Protective Equipment”; *Autom. Constr.* 2020, 112, 103085. <https://doi.org/10.1016/j.autcon.2020.103085>.
- [1] Russell, S.; Norvig, P. “A Modern Approach”; *Artif. Intell.* 1995, 25. <https://doi.org/10.1142/S0219843620500164>.
- [2] Sonka, M.; Hlavac, V.; Boyle, R. “Image Processing, Analysis, and Machine Vision”; *Cengage Learning*, 2014.
- [3] Hu, Y.; Wu, X.; Zheng, G.; Liu, X. “Object Detection of UAV for Anti-UAV Based on Improved YOLO v3”; *Chin. Control Conf.* 2019, 8386-8390. <https://doi.org/10.23919/ChiCC.2019.8865525>.
- [4] Kharchenko, V.; Chyrka, I. “Detection of Airplanes on The Ground Using YOLO Neural Network”; *17<sup>th</sup> IEEE Int. Conf. Math. Methods. Electromagn. Theory.* 2018, 294-297. <https://doi.org/10.1109/MMET.2018.8460392>.
- [5] Saqib, M.; Khan, S. D.; Sharma, N.; Blumenstein, M. “A Study on Detecting Drones Using Deep Convolutional Neural Networks”; *14<sup>th</sup> IEEE Int. Conf. Adv. Video. Signal Based Surveill.* 2017, 1-5. <https://doi.org/10.1109/AVSS.20178078541>.
- [6] Aker, C.; Kalkan, S. “Using Deep Networks for Drone Detection”; *14<sup>th</sup> IEEE Int. Conf. Adv. Video. Signal Based Surveill.* 2017, 1-6. <https://doi.org/10.1109/AVSS.2017.8078539>.
- [7] Lee, D.; La, W. G.; Kim, H. “Drone Detection and Identification System Using Artificial Intelligence”; *Int. Conf. Inf. Commun. Syst.* 2018, 1131-1133. <https://doi.org/10.1109/ICTC.2018.8539442>.
- [8] Schumann, A.; Sommer, L.; Klatte, .; Schuchert, T.; Beyerer, . “Deep Cross-Domain Flying Object Classification for Robust UAV Detection”; *14<sup>th</sup> IEEE Int. Conf. Adv. Video. Signal Based Surveill.* 2017, 1-6. <https://doi.org/10.1109/AVSS.2017.8078558>.
- [9] Unlu, E.; Zenou, E.; Riviere, N.; Dupouy, P.-E. “Deep Learning-Based Strategies for the Detection and Tracking of Drones Using Several Cameras”; *IPS Trans. Comput. Vis.* 2019,11,1-13. <https://doi.org/10.1186/s41074-019-0059>.
- [10] Xun, D. T. W.; Lim, Y. L.; Srigrarom, S. “Drone Detection Using YOLOv3 with Transfer Learning on NVIDIA etson TX2”; *2<sup>th</sup> Int. Symp. Instrum. Control. Artif. Intell. Robot.* 2021, 1-6. <https://doi.org/10.1109/ICA-SYMP50206.2021.9358449>.
- [11] Amil, S.; Abbas, M. S.; Roy, A. M. “Distinguishing Malicious Drones Using Vision Transformer”; *Artif. Intell.* 2022, 3, 260-273. <https://doi.org/10.3390/ai3020016>.
- [12] Upadhyay, M.; Murthy, S. K.; Ra, A. B. “Intelligent System for Real Time Detection and Classification of Aerial Targets Using CNN”; *5<sup>th</sup> Int. Conf. Intell. Comput. Inf. Control. Syst.* 2021, 1676-1681. <https://doi.org/10.1109/ICICCS51141.2021.9432136>.
- [13] Samadzadegan, F.; Dadrass avan, F.; Ashtari Mahini, F.; Gholamshahi, M. “Detection and Recognition of Drones Based on a Deep Convolutional Neural Network Using Visible Imagery”; *IEEE Trans. Aerosp. Electron. Syst.* 2022, 9, 31. <https://doi.org/10.3390/aerospace9010031>.
- [14] Sharma, A.; Ain, N.; Kothari, M. “Lightweight Multi-Drone Detection and 3D-Localization via YOLO”; *ArXiv.org.* 2022, 09097. <https://doi.org/10.48550/arXiv.2202.09097>.
- [15] Sun, H.; Yang, j.; Shen, j.; Liang, D.; Ning-Zhong, L.; Zhou, H. “TIB-Net: Drone Detection Network with Tiny Iterative Backbone”; *IEEE Access.* 2020, 8, 130697-130707. <https://doi.org/10.1109/ACCESS.2020.3009518>.
- [16] Hu, J.; Niu, H.; Carrasco, J.; Lennox, B.; Arvin, F. “Fault-Tolerant Cooperative Navigation of Networked UAV

- [33] Nalamati, M.; Kapoor, A.; Saqib, M.; Sharma, N.; Blumenstein, M. "Drone Detection in Long-Range Surveillance Videos"; 16<sup>th</sup> IEEE Int. Conf. Adv. Video. Signal Based Surveill. 2019, 1-6. <https://doi.org/10.1109/AVSS.2019.8909830>.
- [34] Peng, J.; Zheng, C.; Lv, P.; Cui, T.; Cheng, Y.; Lingyu, S. "Using Images Rendered by PBRT to Train Faster R-CNN for UAV Detection"; *Comput. Sci. Res. Notes* 2018, 13-18. <http://hdl.handle.net/11025/34647>.