Shahid Sattari Air University

Journal of Emerging Defense Knowledge and Research Policy

Volume 1,Issue2
June 2024
p.p 55-70

Object recognition in aerial images using deep learning

Amirhosein zanganeh[1], Ehsan Sharifi[2], Mohammad Lotfi[3]

Abstract

Background & Purpose: Lately, identification and recognition of objects in an image and video has been attended by many researchers. Aerial images are one of those images supplied by multiple ground-based and space-based tools. Todays, many aerial and space imaging systems are used to provide large amounts of visual data for monitoring and controlling the terrestrial effects.
Methodology: This research is applicable in terms of purpose and the results of this study are presented using simulation - based methods.
Findings: The accuracy of the basic neural network is increased by 7.68% to recognize the objects in aerial images by making some changes in the basic architecture and using two error functions to recognize the objects.
Conclusion: Given that image processing is one of the most reliable automated methods available in defense and surveillance systems to identify and detect objects, increasing the accuracy of these systems will increase the efficiency of these systems.

Keywords: *Image Processing, Aerial Images, Deep Learning,Object Recognition,NeuralNetwork,DenseNet*

1. Ph.D. Student, Department of Computer Engineering, North Tehran Branch, Islamic Azad University, Tehran, Iran. **E-mail: amirhosein@aut.ac.ir**
2. Ph.D. Student, Faculty of Computer, Amirkabir University of Technology, Tehran, Iran. **E-mail: esharif@aut.ac.ir**
3. Instructor**,** Faculty of Computer, Ayatollah Boroujerdi University, Borujerd, Iran. **E-mail: m.lotfi@abru.ac.ir**

Received:2022/06/05
Accepted: 2022/06/26
Corresponding Author: Amirhossein zangeneh

فصلنامهٔ سیاستگذاری دانش و پژوهش‌های نوین دفاعی
دوره ۱، شماره ۲
تیر ۱۴۰۳
صص۵۵-۷۰

معاونت پژوهش و فناوری

تشخیص اشیا در تصاویر هوایی با استفاده از یادگیری عمیق

امیرحسین زنگنه[۱]، احسان شریفی[۲]، محمد لطفی[۳]

چکیده

زمینه و هدف: اخیرا شناسایی و تشخیص اشیا و تحلیل تصاویر، توجه بسیاری از محققان را به سوی خود جلب کرده‌است. از جمله تصاویر با اهمیت این حوزه، تصاویر هوایی هستند که توسط ابزار های مختلف هواپایه و فضا پایه تهیه می‌شوند. امروزه با استفاده از سیستم‌های تصویر برداری هوایی و فضایی حجم بسیار زیادی از داده‌های بصری برای نظارت، کنترل و پایش عوارض زمینی فراهم شده است.

روش شناسی: این تحقیق از نظر هدف کاربردی است و نتایج این پژوهش با استفاده از روش‌های مبتنی بر شبیه‌سازی ارائه شده است.

یافته‌ها: با ایجاد تغییر در معماری پایه و استفاده از دو تابع خطا، دقت شبکه‌های عصبی پایه را در شناسایی و تشخیص اشیاء در تصاویر هوایی را به میزان ۷٫۶۸ درصد افزایش داده‌ایم.

نتیجه‌گیری: با توجه به اینکه پردازش تصویر از قابل اعتمادترین روش‌های خودکار موجود در سیستم‌های دفاعی و نظارتی برای شناسایی و تشخیص اشیاء می‌باشد، افزایش دقت این سامانه‌ها منجر به افزایش کارایی این سیستم‌ها خواهد شد.

کلیدواژه‌ها: *پردازش تصویر، تصاویر هوایی، یادگیری عمیق، تشخیص اشیاء، شبکه عصبی، شبکه دنس‌نت*

---

۱. دانشجوی دکتری، دانشکده مهندسی کامپیوتر، واحد تهران شمال، دانشگاه آزاد اسلامی، تهران، ایران. **رایانامه**: anirhosein@aut.ac.ir

۲. دانشجوی دکتری، دانشکده مهندسی کامپیوتر، دانشگاه صنعتی امیرکبیر، تهران، ایران. **رایانامه**: esharif@aut.ac.ir

۳. مربی، دانشکده مهندسی برق و کامپیوتر، دانشگاه علامه بروجردی، بروجرد، ایران. **رایانامه**: m.lotfi@abru.ac.ir

تاریخ دریافت مقاله: ۱۴۰۱/۰۳/۱۵
تاریخ پذیرش نهایی مقاله: ۱۴۰۱/۰۴/۱۵
نویسنده مسئول مقاله: امیرحسین زنگنه

مقدمه

شناسایی، تشخیص و ردیابی اشیا در سال‌های اخیر به دلیل کاربردهای فراوان، تحقیقات فراوان انجام‌شده در این زمینه و پیشرفت‌های فراوانی که در سیستم‌های سخت‌افزاری و نرم‌افزاری روی داده است، توجه بسیاری از محققان را به خود جلب کرده‌است. با توجه به پیشرفت روزافزون تجهیزات و فن‌آوری‌های حوزه تصویربرداری، امروزه تجهیزات تصویربرداری هوایی و فضایی گوناگونی وجود دارد.

با استفاده از سیستم‌های سنجش از دور هوایی و فضایی حجم عظیمی از داده‌های ارزشمند بصری برای نظارت، کنترل و پایش اجسام و پدیده‌های سطح زمین فراهم شده است. عکس‌هوایی، یک تصویر واقعی از سطح زمین است، که توسط هواپیما یا کوادکوپتر و با استفاده از یک دوربین عکس‌برداری هوایی گرفته شده‌است. تصاویر هوایی در مقایسه با تصاویر ماهواره‌ای به دلیل وضوح مکانی بیشتر در کاربردهای شناسایی و تشخیص اشیاء در سطح زمین بسیار به کارگرفته می‌شوند(Grabner et al., ۲۰۰۸). با در دسترس قرار گرفتن تعداد فزاینده‌ای از تصاویر هوایی، تشخیص اشیا، که هدف آن شناسایی و مکان یابی اشیاء در تصاویر هوایی است، به یک وظیفه کلیدی در تجزیه و تحلیل تصاویر هوایی تبدیل شده است(Azimi et al., ۲۰۱۸; Razakarivony and Jurie, ۲۰۱۶).

اما با توجه به ثبت تصاویر هوایی از فواصل بسیار دور، شناسایی و تشخیص اشیاء در تصاویر هوایی با دشواری و چالش‌های مختلفی روبرو می‌باشد. تشخیص اشیاء در تصاویر هوایی به دلیل اندازه کوچک آنها و جهت‌گیری‌های مختلف شی در تصویر، چالش برانگیز است. وضعیت روشنایی تصویر نیز یکی دیگر از عوامل مهم برای تشخیص اشیاء در تصاویر هوایی است. وجود چالش‌های بیان شده موجب شده که تشخیص اشیاء در تصاویر هوایی و فضایی مورد توجه بسیاری از محققین قرار گیرد.

(Deng et al., ۲۰۱۷; Long et al., ۲۰۱۷).

شناسایی و تشخیص اشیاء در زمینه های مختلف از جمله امور امنیتی و نظارتی، نظامی، حمل و نقل، پزشکی و حوزه‌های سرگرمی کاربرد گسترده ای دارد.

(Jiao et al., ۲۰۱۹).

شناسایی و تشخیص اشیا در فریم را می‌توان با استفاده از روش‌های سنتی و مدرن انجام داد. با توسعه سریع شبکه‌های یادگیری عمیق و همچنین توسعه توان محاسباتی

واحدهای پردازش گرافیکی[۳]، عملکرد و کارایی سیستم‌های شناسایی رویداد، ردیابی اشیاء که مبتنی بر سخت‌افزار هستند، به شدت بهبود یافته‌است.

شبکه‌های عصبی عمیق[۴] برخلاف روش‌های سنتی برای استخراج ویژگی، از چندین لایه‌ی مخفی استفاده می‌کنند، تا بتوانند از ویژگی‌های سطح پایین و سطح بالای موثر در شناسایی و تشخیص اشیا، استفاده کنند. در این شبکه‌ها، از شبکه‌های عصبی کانولوشنال[۵] به عنوان ستون فقرات، برای استخراج ویژگی‌های کلیدی یک تصویر ورودی استفاده می‌شود [۷].

علیرغم عملکرد قابل قبول شبکه‌های عصبی کانولوشن در شناسایی و تشخیص اشیاء در تصاویر، عملکرد آنها در تصاویر هوایی به دلیل وجود چالش‌هایی مانند اندازه‌های کوچک و جهت‌گیری‌های مختلف اشیاء، پس زمینه پیچیده در تصاویر هوایی و مشکلات در تشخیص سریع با توجه به منطقه بزرگ تحت پوشش این تصاویر دچار محدودیت می‌شود. و این شبکه‌ها در شناسایی و تشخیص اشیاء در تصاویر هوایی با خطای زیادی روبرو هستند. بطور کلی یکی از چالش‌های عمده استفاده از شبکه‌های عصبی عمیق برای پردازش و تحلیل تصاویر افزایش دقت این نوع شبکه‌ها می‌باشد، که ما در این مقاله سعی کرده‌ایم با استفاده از دو تابع خطا دقت شبکه‌های عصبی عمیق را در شناسایی و تشخیص اشیاء در تصاویر هوایی افزایش دهیم.

مقاله در ادامه به شرح زیر سازماندهی شده‌است: در بخش ۲، کارهای انجام شده در زمینه شناسایی و تشخیص اشیاء در تصاویر هوایی مورد بررسی قرار می‌گیرد. در بخش ۳، معماری پایه و معماری پیشنهادی برای شبکه عصبی عمیق بهینه شرح داده می‌شود. در بخش ۴ نتایج تجربی ارائه شده و در نهایت، نتیجه‌گیری در بخش ۵ ذکر شده است.

پیشینهٔ پژوهش

شناسایی و تشخیص اشیاء در تصاویر هوایی، یک کار چالش برانگیز است. این کار با استخراج ویژگی‌ها در سطوح مختلف انجام می‌شود. از ویژگی‌های سطح پایین تصویر مانند رنگ، شکل، بافت و همچنین از ویژگی‌های سطح بالای موجود در تصویر از قبیل شناسایی ساختارهای هندسی موجود در تصویر، می‌توان برای شناسایی و تشخیص اشیاء در تصاویر هوایی استفاده کرد.

---

[3] Graphic process Unit
[4] Deep neural networks
[5] Convolutional Neural Network

تشخیص اشیاء در تصاویر سنجش از دور دارای کاربردهای بسیار زیادی می‌باشد. بسیاری از محققان برای شناسایی و تشخیص اشیاء در حوزه‌ی بینایی ماشین شبکه‌های عصبی کانولوشنی عمیق را مورد استفاده قرار داده اند. یکی از اولین کارهای انجام شده در این خصوص، استفاده از شبکه عصبی عمیق کانولوشنی بازگشتی[۶] (Girshick et al., ۲۰۱۴) است که در آن ابتدا نواحی کاندید بالقوه برای حضور شی، را در تصویر تولید کرده و سپس با ارزیابی نواحی کاندید مشخص شده در مرحله قبل با استفاده از طبقه بندی کننده اقدام به شناسایی و تشخیص اشیاء می‌کنند. زمان اجرا در این روش مخصوصا برای تولید نواحی بالقوه در تصویر بسیار زیاد است و به همین دلیل استفاده از این روش‌ها مخصوصا برای کارهای برخط غیرممکن به نظر می‌رسد.

از جمله کارهای انجام شده دیگر در این زمینه روش‌های مبتنی بر شبکه عصبی کانولوشنی بازگشتی سریع[۷] (Ren et al., ۲۰۱۵) هستند که از شبکه عصبی عمیق کانولوشنی برای تولید نواحی بالقوه در تصویر استفاده کرده‌اند.

در حوزه تشخیص اشیاء در تصاویر هوایی تهیه شده با پهپاد نیز برخی از محققان از شبکه‌های عصبی عمیق کانولوشنی استفاده کرده‌اند.

در(De Oliveira and Wehrmeister, ۲۰۱۶) از شبکه‌های عصبی عمیق کانولوشنی برای شناسایی و تشخیص مردم در تصاویر هوایی تهیه شده با پهپاد، استفاده شده‌است. در کار انجام شده برای کاهش مناطق کاندید و در نتیجه کاهش نواحی تحت جستجو از پردازش تصویر حرارتی استفاده شده است.

لی و همکاران(Jangwon et al., ۲۰۱۷) تشخیص اشیاء از شبکه‌های عصبی کانولوشنی و شبکه عصبی کانولوشنی بازگشتی استفاده کردند. آن‌ها برای افزایش سرعت فرآیند تشخیص اشیاء از پردازش ابری استفاده کردند.

سان هانگ و همکاران(Han et al., ۲۰۱۶) نیز برای شناسایی، تشخیص و ردیابی اشیاء از شبکه‌های کانولوشنی استفاده کردند. زنگنه و همکاران(Zanganeh et al., ۲۰۲۲) نیز برای شناسایی و تشخیص رویداد در ویدیو از شبکه‌های عصبی عمیق کانولوشنی استفاده کرده‌اند.

ردمان و همکاران(Redmon et al., ۲۰۱۶) معماری یولو[۸] را برای شناسایی و تشخیص اشیاء معرفی کردند. معماری معرفی شده توسط ردمان و همکاران از لحاظ دقت و

---

[6] R-CNN
[7] Fast R-CNN
[8] YOLO

سرعت در شناسایی و تشخیص اشیاء از بسیاری ازسیستم‌های شخیص شیء پیشی گرفته-است.

ژو[9] و همکاران از روش‌های سنتی مانند حذف پس زمینه، تغییر فریم و جریان نوری برای شناسایی وسایل نقلیه در تصاویر هوایی استفاده کرده‌اند (Yongzheng et al., ۲۰۱۶.). اما این روش‌ها به پیچیدگی صحنه حساس بوده و تنها می‌توانند وسایل نقلیه در حال حرکت را شناسایی کنند و برای تشخیص وسایل نقلیه ساکن در ترافیک محدودیت دارند.

خوش برش و همکاران (Khoshboresh Masouleh, M., Shah-Hosseini, R., ۲۰۲۰) براساس ترکیب شبکه‌های عصبی کانولوشنی عمیق و ماشین بولتزمن، یک روش ترکیبی برای استفاده از مزایای قطعه بندی معنایی تصاویر هوایی برای شناسایی و تشخیص وسایل نقلیه پیشنهاد کرده‌اند.

لاوانیا و همکاران(Vijayaraghavan, V. et al ., ۲۰۱۹) نیز از یک شبکه عصبی کانولوشن برای تشخیص و طبقه بندی وسایل نقلیه به سه کلاس اتوبوس، خودروی سواری و موتورسیکلت استفاده کردند. در این روش کل تصویر به عنوان ورودی الگوریتم درنظر گرفته می‌شود و سپس احتمال تعلق تصاویر ورودی به هر یک از کلاس‌های ورودی، به عنوان خروجی الگوریتم تعیین می‌گردد.

روش پیشنهادی

در این بخش با توجه به ضرورت توسعه روش‌های خودکار دقیق و موثر برای شناسایی اشیاء در تصاویر هوایی، به معرفی روش پیشنهادی می‌پردازیم. ما از یک مدل پایه یادگیری عمیق برای استخراج ویژگی‌ها استفاده می‌کنیم. استفاده از شبکه‌های عصبی عمیق برای شناسایی و تشخیص اشیاء معمولا با دوچالش عمده، بار پردازشی بالا شبکه عصبی عمیق و افزایش دقت شبکه عصبی عمیق در شناسایی و تشخیص اشیاء روبرو هستند. در نتیجه استفاده از روش‌های ترکیبی، می‌توانند به افزایش دقت شبکه‌های عصبی در شناسایی اشیاء منجر شود. در ادامه ما ابتدا معماری پایه شبکه عصبی عمیق دنس‌نت[10]-۱۲۱ که در این مقاله به عنوان شبکه عصبی عمیق پایه، مورد استفاده قرار گرفته شده است را شرح می‌دهیم، سپس در زیربخش بعدی سیستم ترکیب توابع خطا را به منظور افزایش دقت شبکه عصبی عمیق معرفی و تشریح می‌کنیم.

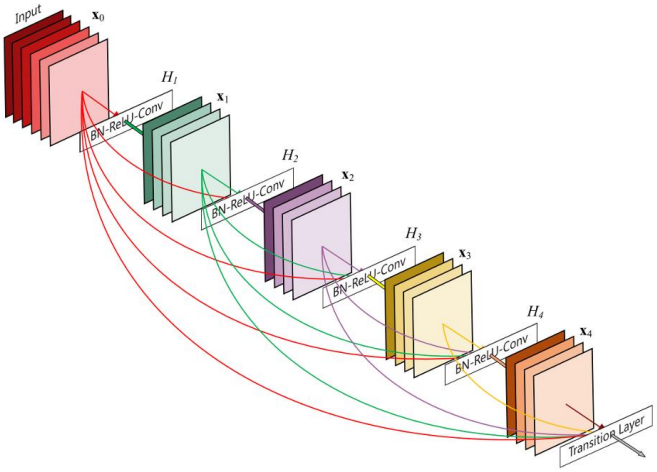
---

[9] Xu
[10] Dense net

۳-۱- معماری مدل پایه

در این مقاله از مدل پایه شبکه عصبی عمیق دنس‌نت برای شناسایی و تشخیص اشیاء در تصاویر هوایی استفاده کرده‌ایم. شبکه عصبی عمیق دنس‌نت توسط هاونگ[11] و همکاران (Huang et al., ۲۰۱۷) در سال ۲۰۱۷ معرفی گردید. شبکه عصبی یادشده یکی از آخرین شبکه‌های ارائه شده برای اهداف شناسایی و تشخیص اشیاء می‌باشد.

این شبکه از نظر معماری مانند شبکه عصبی رزنت[12] می‌باشد، اما دارای چند تفاوت اساسی است. همانطور که در شکل۱ نشان داده شده است، فرق عمده معماری شبکه دنس-نت با معماری شبکه‌های عصبی عمیق قبلی استفاده از skip connection برای تمامی لایه‌ها و لایه‌های قبل از آن است. براساس نتایج بدست آمده برای شناسایی اشیاء، این معماری برای دیتابیس ImageNet نسبت به معماری رزنت به تعداد پارامتر کمتری نیاز دارد در حالی که دقت دو روش مشابه است.



شکل ۱. معماری شبکه عصبی عمیق دنس‌نت (Huang et al., ۲۰۱۷)

در شبکه های عصبی کانولوشن لایه‌های ابتدایی ویژگی‌های سطح پایین مانند لبه‌ها را استخراج و لایه‌های انتهایی این زنجیره، ویژگی‌های سطح بالا مثل بافت‌ها، و اشکال پیچیده را استخراج می‌کنند. در برخی موارد ممکن است ویژگی‌های سطح پایین استخراج شده در عملیات طبقه‌بندی یک کلاس از ویژگی‌های سطح بالای استخراج شده کاراتر و

---

[11] Huang
[12] Res Net

موثر تر باشند، در نتیجه با توجه به ارتباط لایه‌های ابتدایی و لایه‌های انتهایی در معماری شبکه عصبی عمیق دنس‌نت، این شبکه می‌تواند یاد بگیرد که برای کلاس مورد نظر فقط از ویژگی های سطح پایین، ویژگی‌های سطح بالا یا از ترکیب ویژگی‌های سطح پایین و بالا استفاده کند.

۳-۲- افزایش دقت شبکه با ترکیب توابع خطا

هنگام آموزش شبکه‌های عصبی، سعی می‌کنیم وزن‌های شبکه عصبی به گونه‌ای تنظیم شوند که که خروجی را هرچه دقیقتر پیش‌بینی کنند. مقدار دقت در شبکه‌های عصبی توسط تابع هزینه محاسبه می‌شود. این تابع سعی می‌کند هنگامی که شبکه در تخمین خروجی خطا می‌کند، آن را جریمه کند. و در حالت ایده‌آل بهینه‌ترین خروجی، خروجی‌ای است که کمترین میزان هزینه و خطا را داشته باشد.

الگوریتم‌های زیادی با استفاده از یادگیری عمیق برای شناسایی و تشخیص اشیاء در تصاویر معرفی شده است. در میان ساختارهای موجود در این زمینه، ساختار شبکه‌های کانولوشنی[۱۳] توسط محققان زیادی برای تشخیص اشیاء استفاده شده اند. علیرغم عملکرد قابل قبول شبکه‌های عصبی کانولوشن در تشخیص عوارض از تصاویر، عملکرد آنها در تصاویر هوایی به دلیل اندازه‌های کوچک و جهت‌گیری‌های مختلف وسایل نقلیه، پس زمینه پیچیده در تصاویر هوایی و مشکلات در تشخیص سریع با توجه به منطقه بزرگ تحت پوشش این تصاویر دچار محدودیت می‌شود و این شبکه در شناسایی و تشخیص اشیاء در تصاویر هوایی با خطای زیادی روبرو هستند. در این مقاله ما با استفاده از شبکه عصبی عمیق بهبود یافته اقدام به شناسایی و تشخیص اشیاء در تصاویر هوایی نموده‌ایم.

تابع هزینه، معیاری برای سنجش و ارزیابی مدل تهیه شده از نظر قابلیت و توانایی در پیشگویی و تخمین مقدارهای جدید و برچسب داده‌های جدید می‌باشد. شناسایی اشیاء در تصاویر هوایی جزء مسائل طبقه‌بندی چند کلاسه هستند. در این نو مسائل نمونه‌های ورودی هنگام دسته‌بندی در بیش از دو کلاس طبقه‌بندی می‌شوند.

برای دسته‌بندی چند کلاسه معیارهای مختلفی را می‌توان به عنوان تابع هزینه استفاده نمود. ما در این مقاله با استفاده همزمان از دو تابع هزینه آنتروپی متقاطع و معیار کولبک لیبلر سعی در افزایش دقت شبکه داریم. در ادامه هر کدام از این معیار ها را توضیح داده شده‌اند.

---

[13] Convolutional neural networks

آنتروپی متقاطع

تابع آنتروپی متقاطع، تابع هزینه پیش‌فرض است که برای مسائل طبقه‌بندی چند کلاسه از آن استفاده می‌شود. از نظر ریاضی، این تابع هزینه تحت چارچوب استنتاج حداکثر احتمال است. تابع آنتروپی متقاطع میانگین اختلاف بین توزیع‌های احتمالی و احتمال پیش‌بینی‌شده برای همه کلاس‌ها را در مساله محاسبه می‌کند.

میزان ضرر برای بردار ورودی Xi و بردار هدف Yi برابر است با:

$$l(X_i, Y_i) = -\sum_{j=1}^{e} y_{ij} * \log(p_{ij})$$

$$y_{ij} = \begin{cases} 1, & if\ i_{th}\ element\ is\ in\ class\ j \\ 0, & otherwise \end{cases}$$

$$p_{ij} = f(X_i) = Probability\ that\ i_{th}\ element\ is\ in\ class\ j$$

در رابطه بالا Y_i بردار هدف و برابر با $(y_{i1}, y_{i2}, y_{i3}, \ldots, y_{ie})$ می‌باشد.

معیار کولبک لیبلر[۱۴]

این معیار برای اندازه‌گیری میزان تفاوت یک توزیع احتمال با یک توزیع احتمال ثانویه، استفاده می‌شود در صورتیکه مقدار این معیار برای دو توزیع صفر شود، ما می‌توانیم انتظار رفتار مشابهی (نه دقیقاً یکسان) از دو توزیع داشته باشیم در حالی که مقدار ۱ برای این معیار نشان می‌دهد که دو توزیع رفتارهای متضادی دارند.

در حوزه یادگیری ماشین از $D_{KL}(P||Q)$ عموماً به نام بهره اطلاعاتی حاصله به شرطی که به جای توزیع P از توزیع Q استفاده شود، یاد می‌شود.

برای دو توزیع احتمالاتی گسسته P و Q این معیار به صورت زیر تعریف می‌شود:

$$D_{KL}(P||Q) = \begin{cases} -\sum_{j=1}^{e} y_{ij} * \log(p_{ij}), & for\ discrete\ distribution \\ -\int P(x).\log\frac{Q(x)}{P(x)}.dx = \int P(x).\log\frac{P(x)}{Q(x)}.dx, & for\ continious\ distribution \end{cases}$$

[14] Kullback–Leibler

جدول ۱. پارامترهای مربوط به شبیه سازی

پارامتر	مقدار
تابع بهینه سازی شبکه عصبی	Adam
تابع هزینه شبکه عصبی	categorical_crossentropy, kullback_leibler_divergence
معیار عملگرد شبکه	Accuracy
تعداد تکرار	۵۰
سایز دسته	۶۴

یافته‌های پژوهش

در این بخش ابتدا مشخصات پایگاه داده تصاویر مورد استفاده را معرفی کرده و سپس روش پایه و ساختار معماری پیشنهادی را مورد ارزیابی قرار می‌دهیم. در انتهای این بخش مقایسه‌ای بین نتایج معماری پایه و معماری پیشنهادی انجام شده‌است. پارامترهای مربوط به شبیه‌سازی نیز در جدول شماره ۱ ارائه شده است.

در این مقاله برای آموزش و تست روش پیشنهادی از تصاویر موجود در پایگاه داده داتا[۱۵] (Xia et al., ۲۰۱۸) استفاده شده است. اندازه تصاویر در این پایگاه داده از حدود ۸۰۰ × ۸۰۰ تا حدود ۴۰۰۰*۴۰۰۰ متغیر است. این پایگاه داده شامل ۲۸۰۶ تصویر می‌باشد که در ۱۴ کلاس طبقه بندی شده‌اند.

۴-۱- معیارهای ارزیابی روش پیشنهادی

برای ارزیابی عملکرد روش پیشنهادی از ۴ معیار ارزیابی ۱- بازیابی[۱۶] (رابطه ۳)، ۲- وضوح[۱۷] (رابطه ۴)، ۳- معیار-اف[۱۸] (رابطه ۵) و ۴- دقت[۱۹] (رابطه ۶) استفاده کرده‌ایم. هدف ما در این ارزیابی‌ها پیدا کردن اشیاء در تصاویر هوایی است.

$$Recall = \frac{TP}{TP+FP} \quad (۳)$$

$$Precision = \frac{TP}{TP+FN} \quad (۴)$$

[15] DOTA
[16] Recall
[17] Precision
[18] F-measure
[19] Accuracy

که در آن TP[20] تعداد نمونه های مثبتی است که به درستی مثبت شناسایی شده‌اند، TN[21] تعداد نمونه‌های منفی که به درستی منفی شناسایی شده‌اند، FP[22] تعداد شناسایی‌های مثبت کاذب و FN[23] تعداد شناسایی‌های منفی کاذب است. در ادامه مقدار معیار-اف f- و دقت نیز به شرح زیر تعریف می‌شوند:

$$\mathrm{f}-measure = \frac{2*Precision*Recall}{Precision+Recall} \quad (۵)$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (۶)$$

۴-۲- ارزیابی و مقایسه روش ارائه شده

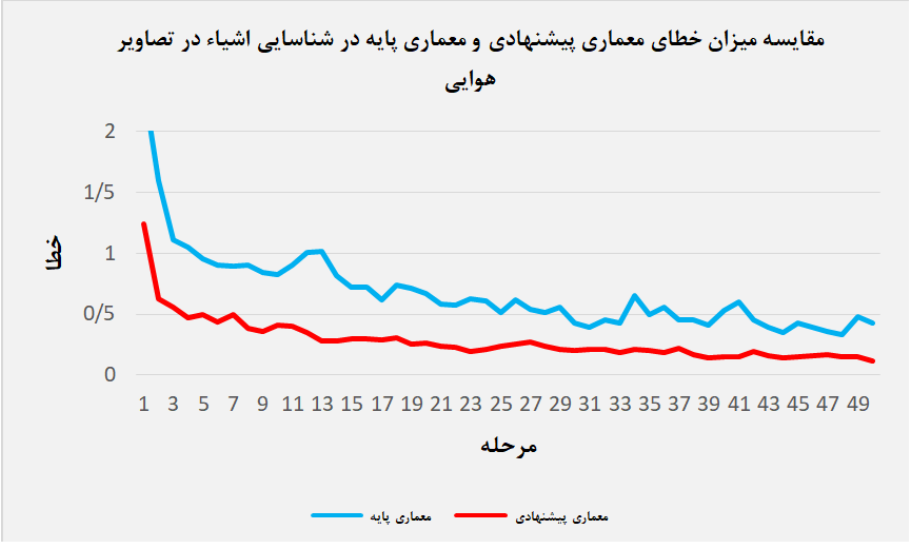
در این بخش، نتایج ارزیابی روش پیشنهادی در افزایش دقت و کاهش میزان خطای شبکه ارائه شده است. طی این ارزیابی با شرایط یکسان، معماری پایه شبکه عصبی دنس‌نت که مدل پایه معماری ارائه شده می‌باشد و معماری پیشنهادی روی پایگاه‌داده معرفی شده، مورد ارزیابی قرار داده‌ایم. در این ارزیابی که هدف آن شناسایی اشیاء در تصاویر هوایی می‌باشد، مجموعه داده‌های تست که شامل تصاویر دربردارنده‌ی شی مورد نظر هستند، به شبکه عصبی عمیق پایه و شبکه عصبی عمیق با معماری پیشنهادی ارائه شده در بخش ۳، به عنوان ورودی داده شده و سپس نتایج روش پیشنهادی روی مجموعه تصاویر تست محاسبه و مقایسه شده است. جدول شماره ۲ این نتایج را نشان می‌دهد چنانکه مشاهده می‌شود دقت روش پیشنهادی برای شناسایی تصاویر حاوی اشیاء مورد نظر بهبود یافته و از ۸۴,۹ درصد به ۹۲,۵۸ درصد افزایش یافته است.

جدول ۲. مقایسه دقت معماری پایه شبکه دنس نت و معماری پیشنهادی

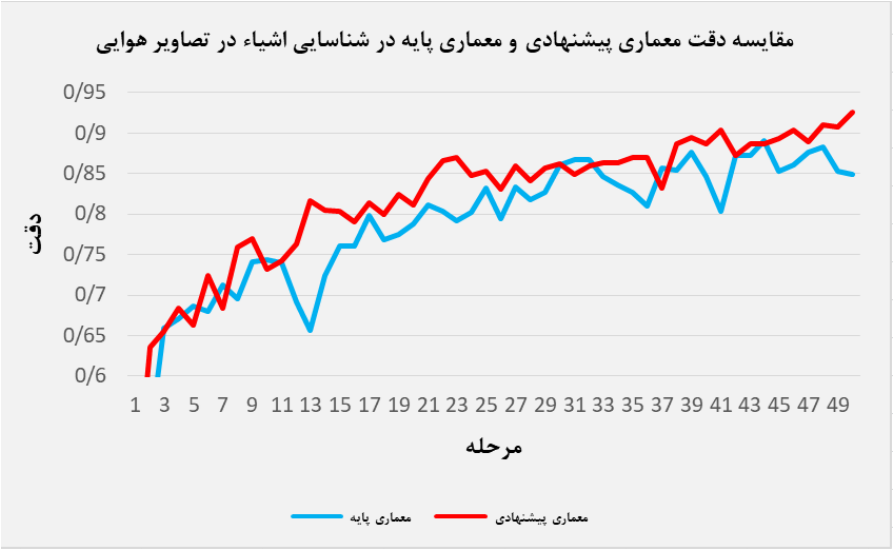
روش	خطا	دقت (برحسب درصد)
معماری پایه	۰,۴۲۴۳	۸۴,۹۰
روش پیشنهادی	۰,۱۱۱۹	۹۲,۵۸

[20] True Positive
[21] True Negative
[22] False Positive
[23] False Negative

نتایج حاصل از شبکه عصبی عمیق پیشنهادی در شناسایی اشیاء مورد نظر براساس معیارهای معرفی شده در بخش ۴-۲ در جدول شماره ۳ ارائه شده‌است.
نمودارهای مربوط به مقایسه میزان خطا و دقت حاصل از معماری پیشنهادی برای شبکه عصبی عمیق شبکه عصبی عمیق پایه در شناسایی و تشخیص اشیاء به ترتیب در شکل‌های شماره ۲ و شکل شماره ۳ ارائه شده‌است.
همانطور که در این شکل‌ها قابل مشاهده است، معماری پیشنهادی برای شبکه عصبی عمیق نسبت به معماری پایه دارای خطای کمتر و دقت بهتری به میزان ۷٫۶۸ درصد می-



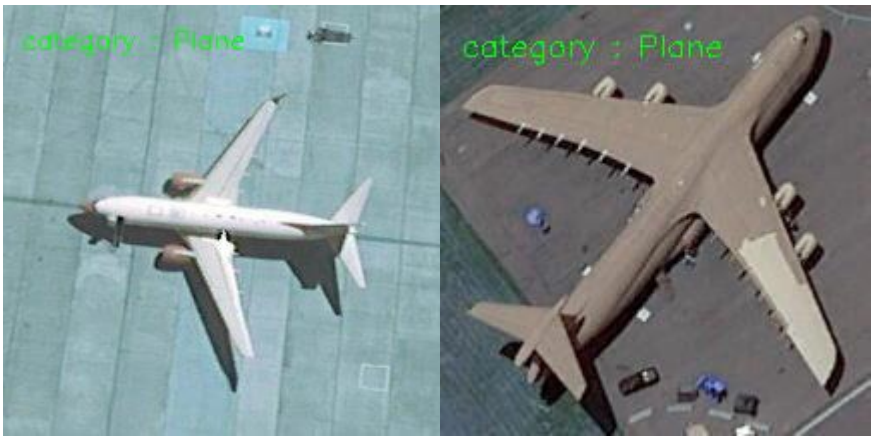
شکل ۲. مقایسه میزان خطای روش پیشنهادی با روش پایه



شکل ۳. مقایسه دقت روش پیشنهادی با روش پایه

باشد.

نتایج ارزیابی روش پیشنهادی در دسته بندی مجموعه تصاویر آزمایشی برای ۲ تصـویر در شکل ۴ ارائه شده‌است. شکل ۴، الف و ب به توانایی این روش را در دسته‌بنـدی تصـاویر دارای اشیاء مورد نظر نشان می‌دهد و روش پیشنهادی با دقت بالا، اشیاء موجود در تصویر را شناسایی و تفکیک کرده است، که گویای کارایی و دقت روش پیشنهادی می‌باشد.



الف **ب**

شکل ۴. ارزیابی روش پیشنهادی در شناسایی اشیاء در تصاویر هوایی

بحث و نتیجه‌گیری

یکی از روش‌های نظارت و کنترل محیط تحلیل و بررسی تصاویر تهیه شده توسط سیستم-های تصویر برداری نظارتی هوایی و فضایی می‌باشد. با استفاده از سیستم‌های تصویربرداری از راه دور هوایی و فضایی حجم بسیار زیادی از داده‌های بصری برای نظارت، کنترل و پایش سطح زمین که دارای کاربردهای مختلف تجاری، امنیتی و دفاعی می‌باشد، تهیه می‌شود. امروزه با توجه به لزوم استفاده از سیستم‌های خودکار در صحنه نبرد، پردازش تصویر از قابل اعتمادترین روش‌های موجود در سیستم‌های دفاعی برای شناسایی و تشخیص هدف می-باشد. اما به دلیل اینکه تصاویر هوایی از فواصل بسیار دور تهیه و ثبت می‌شوند، شناسایی و تشخیص اشیاء در این تصاویر، با چالش‌های مختلفی رویرو می‌باشد.

رشد چشمگیر روش‌های مبتنی بر یادگیری عمیق توانسته نتایج قابل قبولی را در حوزه تحلیل تصویر فراهم کنند. علیرغم عملکرد قابل قبول شبکه‌های عصبی عمیق در شناسـایی و تشخیص اشیاء، عملکرد آنها در تشخیص اشیاء در تصـاویر هـوایی بـه دلیـل چـالش‌هـای

مختلف مانند اندازه‌های کوچک و جهت‌گیری‌های مختلف اشیاء، پس زمینه پیچیده در تصاویر هوایی، وسیع بودن منطقه تحت پوشش این تصاویر در شناسایی و تشخیص اشیاء در تصاویر هوایی با خطای زیادی روبرو هستند.

در این مقاله با استفاده از دو تابع خطا دقت شبکه‌های عصبی عمیق در شناسایی و تشخیص اشیاء در تصاویر هوایی افزایش داده شده‌است و براساس نتایج ارائه شده معماری پیشنهادی برای شبکه عصبی عمیق نسبت به معماری پایه دارای خطای کمتر و دقت بهتری به میزان ۷٫۶۸ درصد می‌باشد.

منابع

Azimi, S.M., Vig, E., Bahmanyar, R., Körner, M., Reinartz, P., 2018. Towards multi-class object detection in unconstrained remote sensing imagery, in: Asian Conference on Computer Vision. Springer, pp. 150–165.

De Oliveira, D.C., Wehrmeister, M.A., 2016. Towards real-time people recognition on aerial imagery using convolutional neural networks, in: 2016 Ieee 19th International Symposium on Real-Time Distributed Computing (Isorc). IEEE, pp. 27–34.

Deng, Z., Sun, H., Zhou, S., Zhao, J., Zou, H., 2017. Toward fast and accurate vehicle detection in aerial images using coupled region-based convolutional neural networks. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 10, 3652–3664.

Girshick, R., Donahue, J., Darrell, T., Malik, J., 2014. Rich feature hierarchies for accurate object detection and semantic segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 580–587.

Grabner, H., Nguyen, T.T., Gruber, B., Bischof, H., 2008. On-line boosting-based car detection from aerial images. ISPRS J. Photogramm. Remote Sens. 63, 382–396.

Han, S., Shen, W., Liu, Z., 2016. Deep drone: Object detection and tracking for smart drones on embedded system. URL Httpsweb Stanf. Educlasscs231aprevprojects2016deepdrone-Object2 Pdf.

Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q., 2017. Densely connected convolutional networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 4700–4708.

Jangwon, L., Wang, J., Crandall, D., Šabanovic, S., Fox, G., 2017. Real-time object detection for unmanned aerial vehicles based on cloud-based convolutional neural networks. J. Concurr. Comput. Pract. Exp. 29.

Jiao, L., Zhang, F., Liu, F., Yang, S., Li, L., Feng, Z., Qu, R., 2019. A survey of deep learning-based object detection. IEEE Access 7, 128837–128868.

Khoshboresh Masouleh, M., Shah-Hosseini, R. (2020) "A hybrid deep learning–based model for automatic car extraction from high-resolution airborne imagery." Appl Geomatics, Vol. 12, PP. 107–119.

Long, Y., Gong, Y., Xiao, Z., Liu, Q., 2017. Accurate object localization in remote sensing images based on convolutional neural networks. IEEE Trans. Geosci. Remote Sens. 55, 2486–2498.

Razakarivony, S., Jurie, F., 2016. Vehicle detection in aerial imagery: A small target detection benchmark. J. Vis. Commun. Image Represent. 34, 187–203.

Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2016. You only look once: Unified, real-time object detection, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 779–788.

Redmon, J., Farhadi, A., 2017. YOLO9000: better, faster, stronger, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 7263–7271.

Ren, S., He, K., Girshick, R., Sun, J., 2015. Faster r-cnn: Towards real-time object detection with region proposal networks. Adv. Neural Inf. Process. Syst. 28.

Vijayaraghavan, V. and Laavanya, M. (2019) "'Vehicle Classification and Detection using Deep Learning.'" International Journal of Engineering and Advanced Technology (IJEAT) Vol. 9(1S5), PP. 24-28.

Xia, G.-S., Bai, X., Ding, J., Zhu, Z., Belongie, S., Luo, J., Datcu, M., Pelillo, M., Zhang, L., 2018. DOTA: A large-scale dataset for object detection in aerial images, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 3974–3983.

Xu, Yongzheng, Guizhen Yu, Yunpeng Wang, Xinkai Wu, and Yalong Ma. 2016. "A Hybrid Vehicle Detection Method Based on Viola-Jones and HOG + SVM from UAV Images" Sensors 16, no. 8: 1325. https://doi.org/10.3390/s16081325, n.d.

Zanganeh, A., Jampour, M., Layeghi, K., n.d. IAUFD: A 100k images dataset for automatic football image/video analysis. IET Image Process. 1–10 (2022). https://doi.org/10.1049/ipr2.12543.