



Islamic Azad University, Shiraz Branch

نشریه تحلیل مدارها، داده ها و سامانه ها
Journal of Circuits, Data and Systems Analysis

sanad.iau.ir/journal/jcdsa



مکان‌یابی خود کار کاروانسراها در تصاویر ماهواره‌ای با بهره‌گیری از

تکنیک‌های پردازش تصویر مبتنی بر یادگیری عمیق

محمد حسین سالاری^{۱*}، محمدامین شایگان^۲، فرناز فرجی^۳

۱- آزمایشگاه فناوری‌های تعاملی، دپارتمان علوم جنگلبانی و تکنولوژی، دانشگاه فنلاند شرقی

Mohammadhossein.salari@uef.fi

۲- گروه مهندسی کامپیوتر، واحد شیراز، دانشگاه آزاد اسلامی، شیراز، ایران

MA.Shayegan@iau.ac.ir

۳- دانشکده علوم انسانی، دانشکده تاریخ، کلاسیک و باستان‌شناسی، دانشگاه نیوکاسل

Farnaz.faraji@newcastle.ac.uk

چکیده: در دهه اخیر، استفاده از یادگیری عمیق در فرآیند پردازش تصاویر به شدت رواج یافته است. استفاده از این تکنیک در زمینه حفاظت و شناخت میراث فرهنگی نیز در مقیاس ملی و بین‌المللی توجه زیادی را به خود جلب کرده است. لیکن مشکلاتی مانند نبود پایگاه‌های داده مورد نیاز، هزینه‌بر بودن ایجاد پایگاه داده جدید، پیچیدگی کار با شبکه‌های عصبی عمیق و نیاز به منابع سخت‌افزاری گران‌قیمت توجه به این مسئله را محدود ساخته است. پژوهش حاضر با ایجاد پایگاه داده‌ای از تصاویر ماهواره‌ای و استفاده از پردازش تصویر مبتنی بر یادگیری عمیق، پایگاه داده‌ای از کاروانسراهای پابرجای ایران ارائه کرده است. به‌همین منظور، پایگاه داده‌ای متشکل از موقعیت جغرافیایی و تصاویر ماهواره‌ای ۲۰۳ کاروانسرای ایران، ایجاد و سپس با استفاده از تکنیک یادگیری انتقالی، الگوریتم یافتن شیء YOLOv5 بر روی پایگاه داده فوق، آموزش داده شده است. به منظور بررسی کارایی شبکه، ۲۵ تصویر جدید از محیط اطراف برخی از کاروانسراها انتخاب و با استفاده از تکنیک پنجره لغزان و وزن‌های آموزش داده شده از قبل، فرایند تشخیص کاروانسراها انجام گرفت. در نهایت، دقت الگوریتم پیشنهادی برابر با ۹۱/۴۳ mAP_0.5٪ محاسبه گردید. روش پیشنهادی می‌تواند به متخصصان میراث فرهنگی در شناخت آثار برجای مانده کمک شایانی نماید.

واژه‌های کلیدی: یادگیری عمیق، کاروانسرا، تصاویر ماهواره‌ای، شبکه‌های عصبی پیچیده، الگوریتم YOLOv5

Automatic Localization of Carvanserais in satellite Images using Image Processing Techniques based on Deep Learning

Mohammad Hossein Salari^{1*}, Mohammad Amin Shayegan², Farnaz Faraji³

¹ Interactive Technologies Lab, School of Computing, Faculty of Science, Forestry and Technology, University of Eastern Finland

Mohammadhossein.salari@uef.fi

² Department of Computer Engineering, Shiraz Branch, Islamic Azad University, Shiraz, Iran

MA.Shayegan@iau.ac.ir

³ Heritage and landscape, Faculty of Humanities, Vrije Universiteit Amsterdam, Research Assistant, School of History, Classics and Archaeology, Newcastle University

Farnaz.faraji@newcastle.ac.uk

Abstract:

In the past decade, the use of deep learning for image-processing task has gained significant popularity. This method has also attracted national and international attention in the field of cultural heritage protection and recognition. However, some issues such as the lack of necessary databases, the costs of creating a new dataset, the complexity of working with deep neural networks, and the need for expensive hardware resources have limited attention to this issue. The present research has provided a database of Iran's remaining caravanserais using a database of satellite images and a deep



learning-based image processing technique. To this end, a database consisting of geolocation and satellite images of 203 Iranian caravanserais has been created. Then, the YOLOv5 object detection algorithm has been trained on this database using the transfer learning technique. To evaluate the performance of this network, 25 new images from the environment around some caravanserais were selected and tested using the sliding window technique and the trained weights. Finally, the accuracy of the proposed algorithm was calculated mAP_{0.5} 43.91%. Therefore, the current method can be useful for cultural heritage experts in recognizing the remaining locations.

Keywords: Deep Learning, Caravansarai, Satellite Images, Complex neural Network, YOLOv5 Algorithm.

DOI: 00.00000/0000

تاریخ چاپ مقاله: ۱۴۰۲/۱۲/۲۵

تاریخ پذیرش مقاله: ۱۴۰۲/۱۲/۱۹

نوع مقاله: پژوهشی / مروری

تاریخ ارسال مقاله: ۱۴۰۲/۱۱/۰۹

۱- مقدمه

مدیریت میراث است. استفاده از یادگیری ماشینی باعث کاهش زمان و هزینه استفاده از نیروی انسانی در مدیریت پروژه‌های میراث فرهنگی و تسهیل فعالیت‌های متخصصان حفاظت و مرمت میراث می‌شود [۵]. برای رفع مشکلات موجود در روش‌های سنتی مستندنگاری بناهای تاریخی، پژوهش حاضر به بررسی استفاده از تکنیک‌های یادگیری عمیق و پردازش تصاویر ماهواره‌ای در مکان‌یابی مجموعه کاروانسراهای موجود در اقلیم ایران پرداخته است. در این مقاله از تصاویر ماهواره‌ای و یادگیری عمیق برای یافتن مکان کاروانسراها استفاده شده است. هرچند که تاکنون تلاش‌های پژوهشی بسیاری توسط افرادی نظیر زنجانی و مالیان جهت مکان‌یابی خودکار کاروانسراها مبتنی بر پیشینه تحقیق ارائه شده صورت گرفته شده است، لیکن پایگاه داده‌ای از موقعیت جغرافیایی و تصاویر ماهواره‌ای از کاروانسراهای ایران با اتکالی به روش ارائه شده در پژوهش حاضر در دسترس عموم قرار نگرفته است. به منظور نیل به اهداف ذکر شده، در این مقاله ابتدا یک پایگاه داده از موقعیت جغرافیایی کاروانسراهای ایران ایجاد و سپس با استفاده از تصاویر ماهواره‌ای، پایگاه داده تصویری مورد نیاز برای آموزش الگوریتم‌های مکان‌یابی اشیاء ایجاد شده است. در گام بعد، با استفاده از پایگاه داده تصاویر ایجاد شده الگوریتم YOLOv5 برای مکان‌یابی کاروانسراها، آموزش داده شده است. در نهایت برای بررسی میزان کارایی روش پیشنهادی، عملیات مکان‌یابی کاروانسراها با ترکیب تکنیک پنجره لغزان و الگوریتم YOLOv5 انجام گرفته است.

ادامه ساختار این مقاله به شرح زیر است. بخش ۲ به معرفی انواع کاروانسراها و ویژگی‌های هندسی آن‌ها، الگوریتم YOLOv5 مورد استفاده در این تحقیق و معیارهای دقت و پیشینه پژوهش‌های مرتبط در این حوزه اختصاص داده شده است. در بخش ۳ الگوریتم پیشنهادی، مراحل ایجاد پایگاه داده‌های مورد نیاز و آموزش الگوریتم پیشنهادی آورده شده است. در بخش ۴ مقایسه بین روش پیشنهادی با سایر ادبیات تحقیق انجام شده است و در نهایت، بخش ۵ این مقاله به تشریح نتایج حاصل از آزمایش‌های انجام شده اختصاص داده شده است.

میراث فرهنگی به عنوان یکی از ابزارهای قدرتمند حفاظت از هویت جامعه در طول تاریخ بشر شناخته شده است. ارزش بی‌بدیل این آثار، ضرورت اقدامات مراقبتی و حفاظتی از آن‌ها را دوچندان می‌سازد. عملیات مستندنگاری، ثبت و ایجاد نقشه از سازه‌ها، ساختارها و ابنیه تاریخی و فرهنگی، به عنوان یکی از اقدامات اساسی در حفاظت از میراث فرهنگی و دانش تاریخی جامعه شناخته می‌شود. طبقه‌بندی و مکان‌یابی این ابنیه، علاوه بر کمک به مستندنگاری کاروانسراها، به شناسایی مسیرهای تاریخی کاروان‌رو و مهم در سرزمین ایران کمک کرده و این امر می‌تواند نقش به‌سزایی در مطالعات کارشناسان تاریخ معماری ایران و حفاظت‌گران و مرمت‌گران میراث فرهنگی داشته باشد. کاروانسراها مجموعه‌ای از ابنیه تاریخی درون‌شهری و برون‌شهری در فلات ایران بوده و به گواه منابع مختلف، کشور ایران در احداث کاروانسراها مبتکر و متقدم بر سایر کشورها بوده است [۱]. متأسفانه بسیاری از کاروانسراهای موجود در ایران، به علت متروک ماندن و عدم رسیدگی و مرمت مستمر، تخریب شده و یا تغییر یافته‌اند و تنها تعداد اندکی از آنها در وضعیتی مطلوبی حفاظت شده‌اند. روش‌های موجود برای شناسایی و مستندنگاری ابنیه تاریخی در کشور، روش‌های سنتی بوده که این امر باعث افزایش هزینه‌ها و خطای محاسباتی می‌شود [۲].

در سال‌های اخیر، استفاده از ابزارها و تکنولوژی‌های نوین، به انتخابی کارآمد، کم‌هزینه و با ریسک کم در مراقبت، حفاظت، مستندنگاری و مدیریت میراث فرهنگی بدل شده است [۳]. از میان تکنولوژی‌ها و فناوری‌های معاصر، ابزارهایی نظیر فتوگرامتری و تصویربرداری از راه دور، واقعیت مجازی، هوش مصنوعی و ... به عنوان بخشی از راهکارهای دیجیتال‌سازی میراث فرهنگی شناخته شده‌اند. در بین ابزارهای ذکر شده، طبقه‌بندی و شناسایی آثار به کمک پردازش تصویر و استفاده از تکنیک‌های شبکه‌های عصبی و یادگیری عمیق، بیشترین کاربرد را در این حوزه داشته‌اند [۴]. از جمله مزایای استفاده از یادگیری ماشینی و یادگیری عمیق در حوزه حفاظت از میراث فرهنگی، امکان گسترش دسترسی کاربران به مجموعه داده‌های بزرگ به منظور طبقه‌بندی و سازمان‌دهی اطلاعات موجود، و همچنین به‌روزرسانی و بازگردانی آرشیو داده‌های تاریخی برای استفاده در عملیات حفاظت و



۲- ادبیات تحقیق

۱-۲- پیشینه پژوهش

زنجانی و مالیان در تحقیق خود به بررسی ویژگی‌های هندسی و پرتوسنجی مناسب برای استخراج کاروانسراها در تصاویر ماهواره‌ای پرداخته‌اند [۶]. آن‌ها با بررسی شکل کاروانسراهای مختلف دریافته‌اند که اکثر کاروانسراها، مربعی و یا مستطیلی شکل بوده و دیواره‌های این بناها دارای اشکال منظمی است. علاوه بر این، برج‌هایی در چهار گوشه اکثر کاروانسراها واقع شده که می‌تواند نشانه‌ای مناسب برای تشخیص این بناهای تاریخی باشد. همچنین، اغلب کاروانسراها دارای محوری متقارن و متعادل بوده و به اتاق‌های متوالی و متعدد تفکیک شده‌اند. در این تحقیق ۱۱ ویژگی هندسی و ۳ ویژگی پرتوسنجی مناسب برای استخراج کاروانسرا از طریق سنجش از دور معرفی شده است. نتایج حاصل نشان داده است که الگوریتم‌های استخراج لبه و شکل می‌تواند به عنوان یک راهکار در تشخیص کاروانسراها در تصاویر ماهواره‌ای به کار گرفته شوند. در تحقیقی دیگر، زنجانی و مالیان با ایجاد پایگاه داده‌ای شامل ۵۰ تصویر ماهواره‌ای از کاروانسراهای ایران، به مسئله شناسایی خودکار کاروانسراهای ایرانی، با بهره‌گیری از الگوریتم تناظریابی الگویی پرداخته‌اند [۲]. آن‌ها در ابتدا با بررسی ساختار کاروانسراهای ایرانی، ویژگی‌های هندسی آن‌ها را از تصاویر موجود استخراج کرده و سپس برای یافتن کاروانسرا در یک تصویر، با استفاده از الگوریتم‌های لبه‌یابی و آستانه‌گذاری و در نهایت با روش‌های قطعه‌بندی تصویر، عملیات مکان‌یابی را انجام داده‌اند. زنجانی و مالیان در تحقیق خود متوجه شده‌اند که به علت شباهت بسیار زیاد دیواره‌های کاروانسرا با محیط اطراف و از سویی وجود سایر اطلاعات نامرتب در تصاویر ماهواره‌ای، الگوریتم طراحی شده بر اساس لبه‌یابی و آستانه‌گذاری قابلیت استفاده به عنوان ابزاری مؤثر برای استخراج کاروانسرا را ندارد. لذا در ادامه، آن‌ها با بهره‌گیری از الگوریتم تناظریابی الگویی و ساخت الگوی کاروانسراها و جستجوی آنها در تصویر، الگوریتم جدیدی معرفی کرده‌اند که قابلیت استخراج دقیق مکان کاروانسراها در ۸۴٪ تصاویر نمونه را دارد. متأسفانه محققین این مقاله، پایگاه داده خود را منتشر نساخته‌اند. همچنین آن‌ها از تصاویر بدون کاروانسرا در پایگاه داده خود استفاده نکرده‌اند و بنابراین نمی‌توان در مورد تعمیم‌پذیری و میزان خطای مثبت کاذب روش پیشنهادی آن‌ها نظری داد. بله‌ی و همکاران نیز به موضوع کاربرد هوش مصنوعی در حفاظت از میراث فرهنگی پرداخته‌اند [۳]. آن‌ها در پژوهش خود بر آن بوده‌اند تا به کمک پروژه موردی CEPROQHA، کاربرد روش‌ها و حوزه‌های مختلف یادگیری عمیق را در حفاظت از میراث فرهنگی مورد واکاوی قرار دهند. به نظر آن‌ها مجموعه رویکردهای یادگیری عمیق، قابلیت طبقه‌بندی و جهت‌دهی به داده‌های میراث فرهنگی را دارند.

کوندرلی و همکاران به تشخیص میراث گمشده در ویدیوهای تاریخی با کمک شبکه‌های عصبی پرداخته‌اند [۵]. این پژوهشگران معتقدند که یکی از روش‌های بازیابی اطلاعات میراث مفقود، رجوع به تصاویر تاریخی آرشیوها است. سلسله مراتب پژوهشی در این تحقیق بدین صورت است که پس از شناسایی میراث مفقود به کمک تکنیک یادگیری عمیق، بازسازی سه‌بعدی مجازی آن‌ها با روش فتوگرامتری انجام می‌شود. این روش بر روی دو نمونه موردی کلیسای تور سنت جکویس و ساختمان لس هالس و بر مبنای وضعیت معماری این دو بنا، پیش از تخریب آن‌ها، انجام شده است. مهم‌ترین مشکل پژوهش فوق، کیفیت پایین تصاویر تاریخی موجود از بناهای ذکر شده و تعداد اندک این تصاویر است که عملکرد شبکه عصبی را محدود ساخته است. یافته‌های این تحقیق نشان از اهمیت بالای نوآوری‌های موجود در زمینه مستندنگاری باده‌مان‌های تخریب‌شده داشته و شیوه‌های جدیدی را برای بازگردانی اطلاعات گذشته در اختیار قرار می‌دهد. دارما و همکاران به مقایسه بین الگوریتم‌های YOLOv5 و mask R-CNN در تشخیص نقوش حکاکی^۱ پرداخته‌اند [۷]. در این تحقیق مشخص گردید که الگوریتم YOLOv5 با رسیدن به دقت [0.95-0.5] mAP برابر با ۰/۹۸۷ و زمان آموزش کمتر، از الگوریتم mask R-CNN بهتر عمل می‌کند. لاماس و همکاران به طبقه‌بندی و دسته‌بندی سبک‌های مختلف ابنیه معماری شناخته شده به عنوان میراث فرهنگی با بهره‌گیری از تکنیک‌های یادگیری عمیق پرداخته‌اند [۸]. در این پژوهش، نویسندگان دو هدف «انتشار پایگاه داده برای دسترسی عموم» و «بررسی امکان استفاده از شبکه‌های عصبی در مستندسازی میراث معماری دارای ارزش» را دنبال کرده‌اند. پایگاه داده ایجاد شده در این پژوهش شامل اجزای میراث معماری^۲ است که برای آموزش و اعتبارسنجی مورد استفاده قرار می‌گیرد.

اسس و همکاران در تحقیق خود به مطالعه تشخیص و طبقه‌بندی آسیب‌های رخ داده بر روی مصالح بنایی ساختمان‌های میراث فرهنگی، به کمک داده‌های تصویری پرداخته‌اند [۹]. هدف از این مطالعه، تعیین نحوه استفاده از فناوری‌های هوش مصنوعی و یادگیری ماشینی در تخمین میزان عملیات حفاظتی لازم بر روی میراث معماری بوده است. آن‌ها در پژوهش خود چهارچوب مورد نیاز برای خودکارسازی فرایند تعیین میزان مرمت مورد نیاز را ارائه نموده و به بررسی عملکرد آن برای ارزیابی چیدمان سنگی دیوارهای بنایی پرداخته‌اند. گرونر و همکاران، به مقایسه مدل‌های مورد استفاده در یادگیری عمیق برای شناسایی اشیاء در تصاویر ماهواره‌ای پرداخته‌اند [۱۰]. آن‌ها در بررسی خود متوجه شدند که الگوریتم‌های تشخیص شیء تک-مرحله‌ای ضمن داشتن دقتی مشابه، با سرعت بسیار بالاتر از الگوریتم‌های دو یا چند مرحله‌ای، عملیات تشخیص شیء در تصاویر ماهواره‌ای را انجام می‌دهند. اما الگوریتم‌های دو یا چند مرحله‌ای برای تشخیص اشیاء کوچک، از دقت بالاتری برخوردارند. همچنین در این تحقیق به استفاده از مدل پنجره لغزان در

^۱ carving motif

^۲ Architectural Heritage Elements Dataset



کاروانسراهای چاه خوشاب و کاروانسرای دوکوهک هستند [۱۶]. شکل (۱) تصویر ماهواره‌ای و پلان نقشه این دو کاروانسرا را نشان می‌دهد. کاروانسراهای چهار ایوانی دارای حیاط مرکزی باز بوده و سابقه‌ای طولانی در تاریخ معماری ایران دارند. اغلب این کاروانسراها با طرح ثابت مربع یا مستطیل و با ایوان‌هایی در مرکز هر جبهه ساخته شده‌اند، لیکن در جزئیات شکل داخل و خارج، دروازه ورودی، برج‌ها و سایر فضاها دارای ویژگی‌های متفاوتی هستند [۱۶]. نمونه این نوع کاروانسراها، کاروانسرای دیرگچین و کاروانسرای مهیار است که نقشه پلان و تصویر ماهواره‌ای آن‌ها در شکل (۲) نشان داده شده است. کاروانسراهای مورد هدف در پژوهش حاضر، اغلب در دسته کاروانسراهای برون‌شهری و ایوان‌دار (دو ایوانی و چهار ایوانی) قرار دارند.



ب

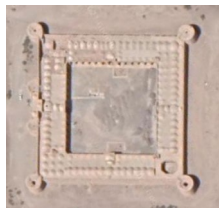
الف



د

ج

شکل (۱): الف) کاروانسرای چاه‌خوشاب، نزدیک فردوس [۱۶]؛ ب) تصویر ماهواره‌ای کاروانسرای چاه‌خوشاب از سرویس نقشه گوگل؛ ج) کاروانسرای دو دهک [۱۶]؛ د) تصویر ماهواره‌ای کاروانسرای دو دهک از سرویس نقشه گوگل



ب

الف



د

ج

شکل (۲): الف) کاروانسرای دیرگچین، نزدیک شهری [۱۶]؛ ب) تصویر ماهواره‌ای کاروانسرای دیرگچین از سرویس نقشه گوگل؛ ج) کاروانسرای مهیار [۱۶]؛ د) تصویر ماهواره‌ای کاروانسرای مهیار از سرویس نقشه گوگل

شناسایی اشیاء در تصاویر ماهواره‌ای اشاره شده است. بالانویک و همکاران عملیات شناسایی موقعیت سدهای ذخیره پسماند معادن در تصاویر ماهواره‌ای را با بهره‌گیری از تکنیک یادگیری عمیق انجام داده‌اند [۱۱]. پایگاه داده این پژوهش مجموعه‌ای متشکل از تصاویر ماهواره‌ای سه کاناله RGB است. یافته‌های این پژوهش نشان می‌دهد که با استفاده از تکنیک یادگیری عمیق، می‌توان هزینه‌های شناسایی و موقعیت سنجی عوارض زمینی را به حداقل رساند و از این تکنیک‌ها، برای شناسایی بناهای ساخته شده توسط انسان بهره گرفت.

در زمینه استفاده از الگوریتم YOLOv5، ونگ و همکاران عملکرد الگوریتم YOLOv5 بهینه شده را در تشخیص رانش زمین در تصاویر ماهواره‌ای مورد بررسی قرار داده‌اند [۱۲]. آن‌ها در مقاله خود ابتدا پایگاه داده‌های مورد نیاز را با استفاده از تصاویر رزولوشن بالای ماهواره‌ای ایجاد و عملیات برچسب‌زنی بر روی تصاویر را با کمک متخصصان انجام داده‌اند. سپس با افزودن تکنیک Adaptively Spatial Feature Fusion (ASFF) به الگوریتم YOLOv5 و ادغام ویژگی‌های موجود در مقیاس‌های مختلف و افزودن Convolutional Block Attention Module (CBAM) برای استخراج بهتر ویژگی‌های کم‌عمق نسبت به بهبود الگوریتم پایه اقدام نموده‌اند. روش آن‌ها ۱/۶۴٪ بهبود را نسبت به روش‌های موجود نشان می‌دهد. جیندال و همکاران به بررسی امکان استفاده از الگوریتم YOLOv5 برای تشخیص هواپیما در تصاویر ماهواره‌ای پرداخته‌اند [۱۳]. آن‌ها برای بررسی دقت الگوریتم YOLOv5 و YOLOv4، آنها را بر روی پایگاه داده CGI plane بررسی نموده و متوجه شده‌اند که دقت YOLOv4 در صورت عدم استفاده از داده‌افزایی^۱، بیشتر از YOLOv5 است. اما با انجام داده‌افزایی، YOLOv5 به دقت بالاتری رسیده، ضمن آنکه سریع‌تر نیز آموزش می‌بیند. ژان‌یینگ و همکاران از الگوریتم YOLOv5 برای تشخیص آتش در جنگل از تصاویر تهیه شده توسط پهپادها استفاده نموده‌اند [۱۴]. همچنین هو و همکاران نیز از الگوریتم YOLOv5 بهبود یافته، برای تشخیص آتش در جنگل استفاده کرده‌اند [۱۵]. هر دو این مقالات توانایی‌های الگوریتم یاد شده در تشخیص آتش را نشان می‌دهند.

۲-۲- مبانی نظری

۲-۲-۱- کاروانسرا و گونه‌های آن

کاروانسراهای ایران از نظر هندسی به پنج دسته کاروانسراهای دو ایوانی، کاروانسراهای چهار ایوانی، کاروانسراها با تالار ستون‌دار، کاروانسراهای با پلان متفرقه و کاروانسراهای چندضلعی تقسیم می‌شوند [۱]. کاروانسراهای دو ایوانی همانند مدارس و مساجد، اغلب به فرم مربع یا مستطیل ساخته شده‌اند. عموماً ایوان‌های این کاروانسراها یکی در مدخل ورودی و دیگری روبه‌روی آن قرار دارد. فضاهای دیگر نیز دور تا دور میان سرای مربع شکل مرکزی جای گرفته‌اند. نمونه این طرح‌مایه،

^۱ Data Augmentation

۲-۲-۲- شبکه عصبی عمیق YOLOv5

است که در پژوهش حاضر، به علت تعداد اندک داده‌های پایگاه داده آموزشی و همچنین نیاز به سرعت بالای محاسبات، از YOLOv5s که کوچک‌ترین و سریع‌ترین معماری خانواده YOLOv5 است استفاده شده است. ساختار YOLOv5s در شکل (۳) ترسیم شده است.

۲-۲-۳- معیارهای خطا

در مقاله حاضر، به منظور ارزیابی عملکرد روش پیشنهادی در مکان‌یابی کاروانسراها، از معیارهای صحت/فراخوانی (۱)، دقت (۲) [۲۴] و mAP^۶ (۶) استفاده شده است.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (۱)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (۲)$$

در (۱، ۲) مقدار TP برابر با تعداد نمونه‌هایی مثبتی که به درستی تشخیص داده شده‌اند، FP تعداد نمونه‌های منفی‌ای که به اشتباه مثبت تشخیص داده شده‌اند و FN تعداد نمونه‌های مثبتی که به اشتباه منفی تشخیص داده شده‌اند، هستند.

معیار mAP مقادیر کادر محدودکننده^۷ یک تصویر (با استفاده از اطلاعات ground-truth آن تصویر) را با کادرهای محدودکننده تشخیص داده شده توسط الگوریتم مقایسه کرده و یک امتیاز ایجاد می‌کند. هرچه مقدار این امتیاز بیشتر باشد بیانگر عملکرد بهتر الگوریتم است. برای محاسبه mAP ابتدا دقت میانگین (AP)^۸ برای هر یک از کلاس‌های داده‌ها محاسبه می‌شود. AP راهکاری برای خلاصه‌سازی نمودار precision-recall به یک مقدار عددی بر اساس (۳) است [۲۵].

$$AP = \sum_{k=0}^{k=n-1} [\text{Recalls}(k) - \text{Recalls}(k+1)] * \text{Percisions}(k) \quad (۳)$$

$\text{Recalls}(n) = 0, \text{Percisions}(n) = 1$
 $n : \text{Number of thresholds}$

در کاربردهای تشخیص شیء، یک مقدار آستانه برای محاسبه مثبت یا منفی بودن خروجی الگوریتم انتخاب می‌گردد و اگر معیار IoU^۹ بر اساس (۴) از این مقدار بیشتر باشد، طبق (۵) آن خروجی به عنوان نمونه مثبت و در غیر اینصورت به عنوان نمونه منفی طبقه‌بندی می‌گردد [۲۶].

$$IoU = \frac{\text{Area Of Overlap}}{\text{Area of Union}} \quad (۴)$$

$$\text{class}(IoU) = \begin{cases} \text{Positive if } IoU \geq \text{Threshold} \\ \text{Negative if } IoU < \text{Threshold} \end{cases} \quad (۵)$$

پس از محاسبه AP برای تمام کلاس‌ها، میانگین این مقادیر به همراه مقدار آستانه IoU انتخاب شده، به عنوان mAP بر اساس (۶) گزارش می‌گردد. به شکل معمول، مدل‌های تشخیص شیء با چند مقدار آستانه متفاوت ارزیابی می‌شوند. یکی از متداول‌ترین حد آستانه‌ها،

الگوریتم YOLO توسط ردمون و همکاران برای تشخیص سریع و بلادرنگ اشیاء در تصاویر پیشنهاد شده است [۱۷] و تاکنون نسخه‌های مختلفی از این الگوریتم معرفی شده است. این الگوریتم، تک‌مرحله‌ای بوده و بر خلاف بسیاری از الگوریتم‌های تشخیص شیء، مرحله پیشنهاد نواحی را طی نمی‌کند و فقط تعداد محدودی از کادر محصورکننده را پیشنهاد کرده و به همین علت، دقت تشخیص بلادرنگ^۱ مناسب و سریعی دارد [۱۸]. در مقاله حاضر، از نسخه YOLOv5 [۱۹] به‌عنوان یکی از پر کاربردترین نسخه‌های خانواده YOLO استفاده شده است. اصلی‌ترین تفاوت YOLOv5 با سایر معماری‌های خانواده YOLO، پیاده‌سازی کامل آن با استفاده از زبان برنامه‌نویسی Python و کتابخانه PyTorch است. برخلاف نسخه‌های پیشین YOLO که بر مبنای کتابخانه DarkNet توسعه یافته‌اند، این تغییر، باعث سهولت در نصب، اجرا و آموزش مجدد شبکه YOLOv5 شده است. الگوریتم YOLOv5 ساختاری مشابه با سایر الگوریتم‌های خانواده YOLO داشته و از چهار بخش ورودی، neck، backbone و پیش‌بینی^۲ تشکیل می‌شود [۲۰]. ورودی الگوریتم YOLOv5 از تکنیک بهینه‌سازی داده به صورت موازی^۳ بهره می‌گیرد که باعث بهبود عملکرد آن در تشخیص اشیاء کوچک می‌شود. در هنگام آموزش شبکه YOLOv5، اندازه بهینه برای anchor frame‌ها برای هر مجموعه داده آموزشی، به صورت تطبیقی محاسبه می‌شود [۲۱]. قسمت Backbone در شبکه YOLOv5 ویژگی‌های اصلی تصویر ورودی را استخراج می‌کند. در YOLOv5 از CSPDarknet به عنوان Backbone استفاده شده که با ادغام CSPNet^۴ در معماری Darknet ایجاد شده است [۲۲]. استفاده از CSPNet در YOLOv5 باعث افزایش سرعت، در عین دست‌یابی به دقتی مشابه و حتی بیشتر نسبت به معماری Darknet53 شده است [۲۳]. معماری YOLOv5 برای ترکیب ویژگی‌های استخراج شده توسط Backbone خود، از ساختاری مشابه با PANet^۵ برای بخش Neck استفاده می‌کند. استفاده از این ساختار باعث می‌شود تا مدل آموزش دیده، بتواند اشیاء با اندازه‌های متفاوت را شناسایی کرده و در برابر تغییر اندازه مقاوم باشد [۲۱]. لایه خروجی در YOLOv5 از سه شبکه کانولوشنی تشکیل می‌شود که هر کدام عملیات یافتن شیء را به شکل مجزا انجام می‌دهند. سپس بر روی خروجی حاصل از این سه شبکه، عملیات maximum value suppression انجام گرفته تا خروجی نهایی ایجاد گردد [۲۰]. با تغییر تعداد هسته‌های کانولوشن Focus و CBL و تعداد residual modules در CSP، می‌توان به Backbone‌های با عمق و عرض مختلف برای شبکه YOLOv5 دست یافت که هر کدام داری سرعت و دقتی متفاوت هستند [۲۰]. لازم به ذکر

^۶ Mean Average Precision (mAP)

^۷ Bounding Box

^۸ Average Precision (AP)

^۹ Intersection over Union (IoU)

^۱ Real-Time

^۲ Prediction

^۳ Mosaic Data Enhancement

^۴ Cross Stage Partial Networks (CSPNet)

^۵ Path Aggregation Network (PANet)



IoU mAP@.5 بوده که به معنی متوسط AP کلاس‌های مختلف برای IoU با آستانه ۰/۵ است.

$$mAP = \frac{1}{n} \sum_{k=1}^{k=n} AP_k \quad (6)$$

$AP_k =$ the AP of class k

$n :$ the number of classes

۲-۲-۴- معرفی LabelImg

LabelImg یک ابزار رایگان و منبع باز است که برای برچسب‌زنی و حاشیه‌نگاری گرافیکی تصاویر طراحی شده است. این ابزار با استفاده از Python توسعه یافته و از Qt برای رابط گرافیکی خود استفاده می‌کند. LabelImg امکان ذخیره حاشیه‌نگاری‌ها را در هر دو فرمت XML VOC و فایل متنی YOLO به پژوهشگران می‌دهد که این امر آن را به ابزاری منعطف برای پروژه‌های مختلف تشخیص شیء تبدیل می‌کند. رایگان و منبع باز بودن، سهولت استفاده و فرمت‌های مختلف پشتیبانی شده، باعث کاربرد گسترده این ابزار در ایجاد مجموعه داده‌های با کیفیت بالا برای آموزش مدل‌های یادگیری عمیق شده است.

۳- روش پیشنهادی

روش پیشنهادی پژوهش حاضر برای تشخیص کاروانسراها از روی تصاویر ماهواره‌ای با ابعاد بزرگ طراحی شده است. در این روش، ابتدا تصاویر ورودی با استفاده از تکنیک پنجره لغزان و با همپوشانی متناسب با قدرت تفکیک تصاویر و ابعاد کاروانسراها، برش زده شده و سپس به ابعاد مورد نیاز الگوریتم YOLOv5 کاهش اندازه داده می‌شوند. سپس به وسیله الگوریتم YOLOv5 و با استفاده از وزن‌های از پیش آموزش داده شده، عملیات استخراج مکان‌های کاندیدا انجام می‌شود. در گام بعد، مجدداً تصویر ورودی به مرکز مکان‌های کاندیدا برش خورده منتقل و بار دیگر عملیات مکان‌یابی کاروانسراها توسط الگوریتم YOLOv5 و حد آستانه بالاتر از مرحله قبل، انجام می‌شود. با توجه به این نکته که دو کاروانسرا نمی‌توانند همپوشانی داشته باشند، با ادغام نواحی کاندیدای کاروانسرا با همپوشانی بالا، خروجی نهایی برای بررسی افراد خبره ایجاد می‌گردد. شکل (۳) مراحل کلی روش پیشنهادی در این مقاله جهت مکان‌یابی کاروانسراها را نشان می‌دهد. در این پژوهش تمامی داده‌های مورد نیاز ایجاد شده است. مراحل آماده‌سازی پایگاه داده مورد نیاز و جزئیات گام‌های انجام گرفته برای آموزش و معیارسنجی الگوریتم پیشنهادی در شکل (۴) نمایش داده شده است.

۳-۱- ایجاد پایگاه داده موقعیت جغرافیایی کاروانسرا

به منظور ایجاد پایگاه داده تصاویر ماهواره‌ای از کاروانسراها، لازم بود ابتدا موقعیت جغرافیایی کاروانسراها استخراج شود. با توجه به آنکه سابق

بر این تحقیق، هیچ داده دیجیتال و دقیقی از مختصات جغرافیایی کاروانسراها ایجاد نشده بود، ابتدا فهرستی از کاروانسراها شناخته شده از منابع فهرست کاروانسراها ایران [۶]، [۲۷] و همچنین جست‌وجوی عبارت «کاروانسرا» در سرویس نقشه گوگل، ایجاد شد. سپس مختصات دقیق این کاروانسراها به صورت دستی و با کمک نقشه بینگ^۱ از سرویس‌های نقشه گوگل^۲ بدست آمد. پس از راستی آزمایی و حذف موارد تکراری، فهرست نهایی موقعیت جغرافیایی کاروانسراها، که شامل موقعیت جغرافیایی ۲۰۳ کاروانسرا می‌شود، ایجاد شد. نمونه‌ای از این داده‌ها در جدول (۱) نشان داده شده است.

۳-۲- دانلود تصاویر ماهواره‌ای متناظر با موقعیت‌های

جغرافیایی کاروانسراها

در این گام برای ایجاد پایگاه داده تصاویر ماهواره‌ای از کاروانسراها، موقعیت‌های جغرافیایی استخراج شده در مرحله ۳-۱ به عنوان مرکز تصویر انتخاب شده و با استفاده از سرویس‌های نقشه گوگل و بینگ و به صورت نیمه خودکار، یک تصویر از هر کاروانسرا در سطح نزدیکی به زمین^۳ ۱۰۰ متر و ابعاد ۴۰۰ × ۴۰۰ متر تهیه شد. برای ایجاد تصاویر نمونه منفی، از محیط اطراف هر کاروانسرا دو تصویر و از ۲۰ مکان تاریخی ایران (مانند حافظیه شیراز، میدان آزادی تهران و ...) در همان سطح نزدیکی به زمین مرحله قبل، ۷۴۵ تصویر دانلود و در پایگاه داده ذخیره شد. تصاویر نمونه منفی به گونه‌ای انتخاب شده‌اند که بافتی مشابه با کاروانسراها داشته باشند تا شبکه عصبی عمیق آموزش داده شده در مراحل بعد، دارای تعمیم‌پذیری بیشتری باشد. شکل (۵) نمونه‌ای از داده‌های موجود در پایگاه داده اولیه را نشان می‌دهد.

۳-۳- ترسیم کادرهای محدود کننده

در این مرحله با استفاده از ابزار LabelImg [۲۸] اطراف هر کاروانسرا یک کادر^۴ ترسیم شد. تصاویر فاقد کاروانسرا بدون ترسیم کادر محدودکننده، به عنوان نمونه منفی در پایگاه داده تشخیص شیء قرار داده شده‌اند. شکل (۶) نمونه‌ای از تصاویر موجود در پایگاه داده به همراه کادرهای محدود کننده رسم شده برای آنها را نشان می‌دهد.

جدول (۱): نمونه داده موقعیت جغرافیایی کاروانسراها ایران

شماره	نام کاروانسرا	موقعیت جغرافیایی بر حسب Decimal degrees (DD)
۰۰۱	کاروانسرای یام	۳۸.۳۴۴۳۲۸ و ۴۵.۸۳۴۹۶۶
۰۰۲	کاروانسرای خواجه نظر	۳۸.۹۷۷۴۵۲ و ۴۵.۵۷۷۰۳۸
۰۰۳	کاروانسرای جمال آباد	۲۷.۲۷۱۹۲۳ و ۸۴۳۰۷۵.۴۷
۰۰۴	کاروانسرای رباط شرق	۲۶۶۴۹۰.۳۶ و ۶۵۵۲۵۳.۶۰
...
۲۰۳	کاروانسرای رباط ده چاهک	۳۳.۲۸۱۲۵۴ و ۹۱۰۱۱۱.۵۸

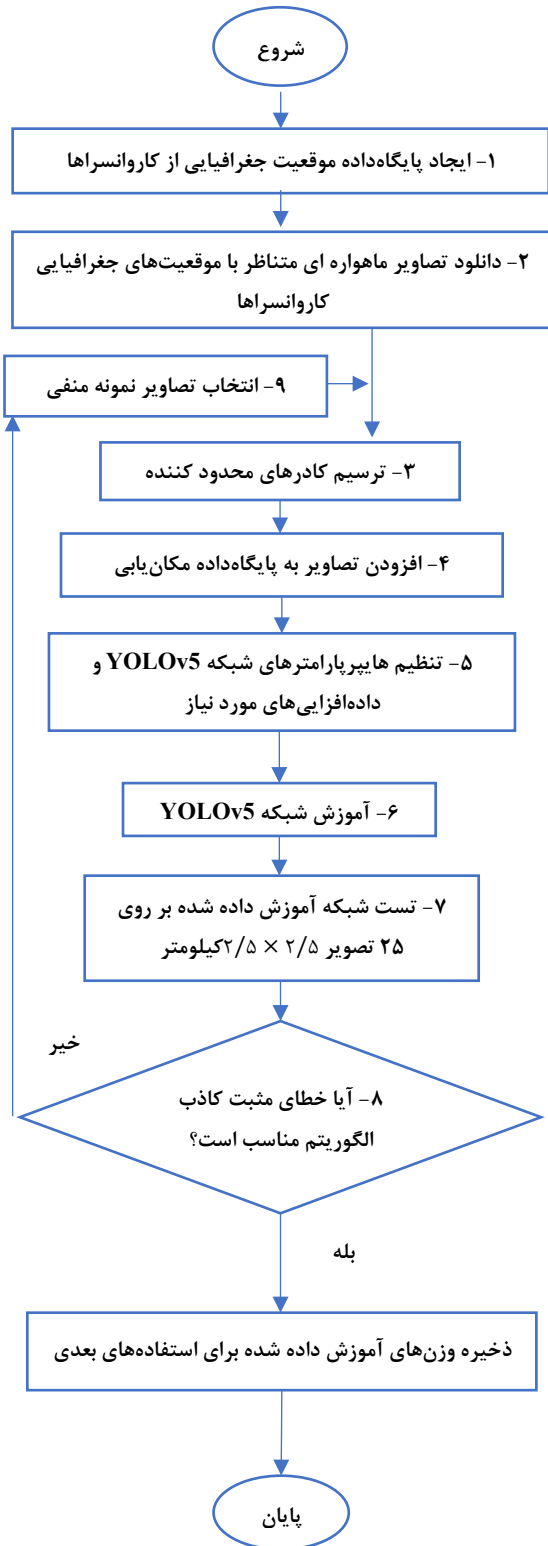
³ Zoom

⁴ Bonding box

¹ <https://www.bing.com/maps>

² <https://www.google.com/maps>





شکل (۴): فلوجارت مراحل ایجاد پایگاه داده و آموزش شبکه عصبی YOLOv5

هایپرپارامترهای تنظیم شده برای آموزش الگوریتم YOLOv5 در جدول (۲) و داده‌افزایی‌های انجام شده هنگام آموزش شبکه در جدول (۳) آمده است. نمونه‌ای از داده‌افزایی‌های انجام شده بر روی تصاویر پایگاه‌داده در شکل (۷) نشان داده شده است.



شکل (۳): مراحل کلی روش پیشنهادی

۳-۴- افزودن تصاویر به پایگاه‌داده مکان‌یابی

با توجه به اینکه برای برخی از کاروانسراها دو تصویر (یکی از سرویس نقشه گوگل و دیگری از سرویس نقشه بینگ) در پایگاه داده وجود دارد، جهت جلوگیری از بروز خطای Group Leakage، زوج تصویر کاروانسرا تنها در یکی از دو گروه تصاویر اعتبارسنجی یا آموزش قرار داده شده‌اند.

۳-۵- تنظیم هایپرپارامترهای شبکه YOLOv5 و

داده‌افزایی‌های مورد نیاز

به منظور افزایش تعداد تصاویر جهت آموزش بهتر شبکه عمیق مورد استفاده، از فرایند داده‌افزایی در این پژوهش استفاده شده است. مقدار انتخاب شده برای پارامترهایی مانند چرخش تصادفی، جابه‌جایی تصویر، تغییر ابعاد و ... به گونه‌ای انتخاب شده‌اند تا بیشترین نزدیکی را با شرایط دنیای واقعی داشته باشند.



جدول (۲): هایپر پارامترهای آموزش

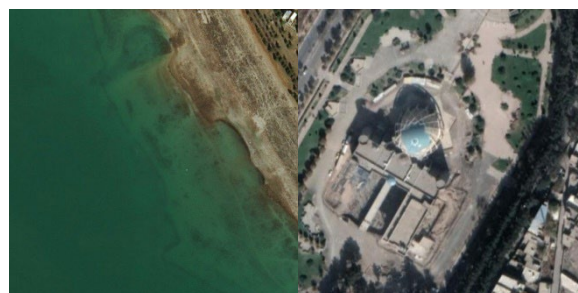
مقدار	پارامتر
lr0: 0.001	مقدار اولیه نرخ یادگیری
lrf: 0.2	OneCycleLR مقدار نهایی پارامتر (lr0 * lrf)
momentum: 0.937	SGD momentum/Adam بتا ۱
weight_decay: 0.05	کاهش وزن بهینه ساز
warmup_epochs: 3.0	دوره‌های گرم کردن
warmup_momentum: 0.8	جنبش اولیه گرم کردن
warmup_bias_lr: 0.1	lr بایاس اولیه گرم کردن
box: 0.05	box loss gain
cls: 0.5	cls loss gain
cls_pw: 1.0	وزن‌های مثبت BCELoss
obj: 1.0	obj loss gain (متناسب با پیکسل‌ها)
obj_pw: 1.0	وزن‌های مثبت BCELoss
iou_t: 0.20	IoU حد آستانه آموزش
anchor_t: 4.0	حد آستانه لنگرهای چندگانه
anchors: 3	لنگرها در هر لایه خروجی
fl_gamma: 0.0	focal loss گاما

جدول (۳): هایپر پارامترهای داده‌افزایی

مقدار	پارامتر
hsv_h: 0.015	داده افزایی HSV-Hue تصویر
hsv_s: 0.7	اشباع داده افزایی HSV-تصویر
hsv_v: 0.4	داده افزایی HSV-Value تصویر
degrees: 15.0	چرخش تصویر (+- درجه)
translate: 0.35	انتقال تصویر
scale: 0.3	ابعاد تصویر
shear: 0.001	برش تصویر
perspective: 0.0	چشم انداز تصویر
flipud: 0.5	وارونگی تصویر بالا-پایین
fliplr: 0.5	وارونگی تصویر چپ-راست
mosaic: 0.0	موزاییک تصویر
mixup: 0.0	ترکیب تصویر
copy_paste: 0.0	رونوشت و الصاق بخش‌ها

۳-۶- آموزش شبکه YOLOv5

برای آموزش شبکه YOLOv5 از کدهای Python منتشر شده توسط توسعه‌دهندگان اصلی^۱ این شبکه استفاده شده است. کارت گرافیک مورد استفاده برای آموزش شبکه از نوع Tesla K80 با ۱۲ گیگابایت رم بوده و از محیط Google Colab^۲ برای آموزش الگوریتم YOLOv5 استفاده شده است.



شکل (۵): نمونه تصاویر موجود در پایگاه داده اولیه



شکل (۶): نمونه تصاویر پایگاه داده تشخیص شیء



شکل (۷): نمونه‌ای از تغییرات ایجاد شده بر روی تصاویر کاروانسرا در هنگام انجام تکنیک داده‌افزایی

² <https://colab.research.google.com/>

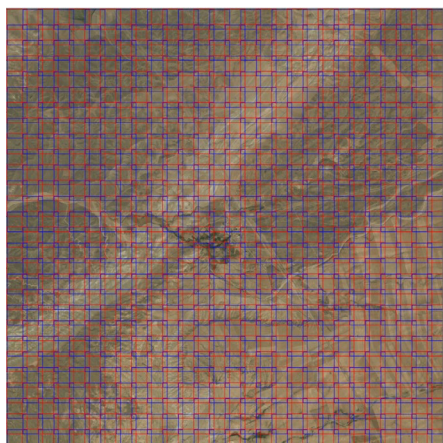
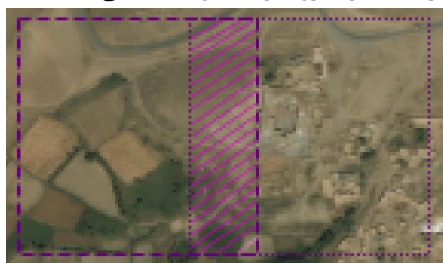
¹ <https://github.com/ultralytics/yolov5>



سطح اطمینان مرحله انتخاب شده است. نتایج اعمال سطوح اطمینان مختلف در گام دوم تشخیص کاروانسراها، در جدول (۵) درج شده است.

۳-۸- بررسی خطای مثبت کاذب الگوریتم

با اجرای گام‌های شرح داده شده در مرحله قبل، کاروانسراهای کاندیدا از تصاویر نمونه استخراج می‌شوند. به علت تعداد کم داده‌های آموزشی و شباهت سایر بناها با ساختار کاروانسرا، خطای مثبت کاذب الگوریتم آموزش داده شده، زیاد است. به منظور کاهش این خطا، مکان‌هایی که در این گام به اشتباه کاروانسرا تشخیص داده شده‌اند، به عنوان تصاویر نمونه منفی، به پایگاه‌داده افزوده شده و آموزش شبکه عصبی مجدداً تکرار می‌گردد. مرحله افزودن تصاویر نمونه منفی به پایگاه‌داده و آموزش مجدد الگوریتم YOLOV5 تا زمانی که افزودن تصاویر جدید باعث کاهش معیار تشخیص کاروانسراها نشود، ادامه می‌یابد.



شکل (۸): نمونه‌ای تصاویر دانلود شده در مرحله ۷ و نحوه برش خوردن آن (کادرهای سبز و قرمز نواحی برش خوردن را نمایش می‌دهد)



شکل (۹): حساسیت شبکه آموزش داده شده به محل قرارگیری کاروانسرا در تصویر

۳-۷- تست شبکه آموزش داده شده بر روی ۲۵

تصویر $2/5 \times 2/5$ کیلومتر

پس از ایجاد پایگاه داده و آموزش شبکه عصبی YOLOV5 بر روی آن، جهت کاهش خطای مثبت کاذب الگوریتم پیشنهادی، وزن‌های آموزش داده شده بر روی تصاویر دنیای واقعی بررسی شده و خطاهای شبکه به عنوان تصاویر جدید به پایگاه‌داده افزوده می‌شود و آموزش YOLOV5 مجدداً تکرار می‌گردد و در نهایت بهترین وزن‌های شبکه YOLOV5 برای استفاده‌های بعدی ذخیره می‌گردد. بدین منظور، موقعیت جغرافیایی ۲۵ کاروانسرا از پایگاه داده موقعیت‌های جغرافیایی، به صورت تصادفی انتخاب و تصاویر با ابعاد $2/5 \times 2/5$ کیلومتری با مرکزیت آن‌ها دانلود شدند. سپس این تصاویر با استفاده از تکنیک پنجره لغزان در ابعاد 800×800 پیکسل با همپوشانی ۲۰۰ پیکسل برش داده شدند (شکل (۸)). در این بخش ابعاد هر تصویر به 526×526 پیکسل کاهش داده شدند. پس از کاهش ابعاد تصاویر برش خورده و با استفاده از وزن‌های آموزش داده شده در مرحله ۶، الگوریتم YOLOV5 بر روی این تصاویر جدید اجرا شد.

با وجود استفاده از داده‌افزایی جابه‌جایی تصاویر، الگوریتم YOLOV5 به محل کاروانسرا در تصویر باز هم حساس بود (شکل (۹)). لذا برای افزایش دقت نهایی، الگوریتم پیشنهادی در دو مرحله، عملیات تشخیص کاروانسراها را انجام می‌دهد. ابتدا با آستانه اطمینان^۱ برابر با ۰/۲۵، کاروانسراهای کاندیدا استخراج می‌شوند. سپس محل کاندیدهای نمونه به عنوان مرکز یک تصویر جدید انتخاب شده و پس از انجام عملیات حذف تصاویر با همپوشانی بالا، پیش‌بینی نهایی الگوریتم YOLOV5 بر روی این تصاویر با آستانه اطمینان برابر با ۰/۵ اجرا می‌شود. مقادیر آستانه اطمینان ۰/۲۵ و ۰/۵ استفاده شده با آزمون و خطا به دست آمده‌اند. برای رسیدن به این مقادیر، به صورت تصادفی ۱۰ کاروانسرا از پایگاه‌داده انتخاب شده و مرحله استخراج کاروانسراهای کاندیدا با confidence thresholdهای مختلف بر روی این ۱۰ تصویر انجام شده است. با اجرای تکنیک پنجره لغزان در ابعاد 800×800 پیکسل با همپوشانی ۲۰۰ پیکسل بر روی ۱۰ تصویر $5/2 \times 5/2$ کیلومتری انتخاب شده، به ازای هر تصویر ۱۱۵۶ تصویر جدید ایجاد می‌شود. با توجه به همپوشانی ۲۰۰ پیکسلی و ابعاد کاروانسراها، ممکن است هر کاروانسرا در حداکثر ۴ تصویر قرار گرفته باشد. همچنین در بسیاری از تصاویر هیچ کاروانسرای وجود نخواهد داشت. نتیجه حاصل از اجرای الگوریتم یافتن کاروانسرا با حد آستانه‌های مختلف مرحله اول در جدول ۴ آورده شده است. لازم به ذکر است که منظور از تشخیص صحیح در جدول ۴، تشخیص کاروانسرای منحصر به فرد است که حداکثر می‌تواند برابر با تعداد ۱۰ کاروانسرا باشد. بر اساس نتایج جدول (۴)، حد آستانه ۰/۲۵ که در آن تمامی کاروانسراها تشخیص داده می‌شوند به عنوان

^۱ Confidence Threshold



جدول (۴): کاروانسراهای کاندیدای استخراج شده به ازای مقادیر مختلف سطح اطمینان

آستانه اطمینان	تشخیص صحیح	تشخیص اشتباه
۰.۷۵	۲	۳۹۸
۰.۶۵	۳	۱۲۵۳
۰.۵۵	۳	۱۸۹۵
۰.۴۵	۷	۳۴۵۹
۰.۳۵	۸	۳۹۸۱
۰.۲۵	۱۰	۴۹۸۰
۰.۱۵	۱۰	۵۳۴۹
۰.۰۵	۱۰	۷۵۹۸

جدول (۵): کاروانسراها به ازای مقادیر مختلف سطح اطمینان

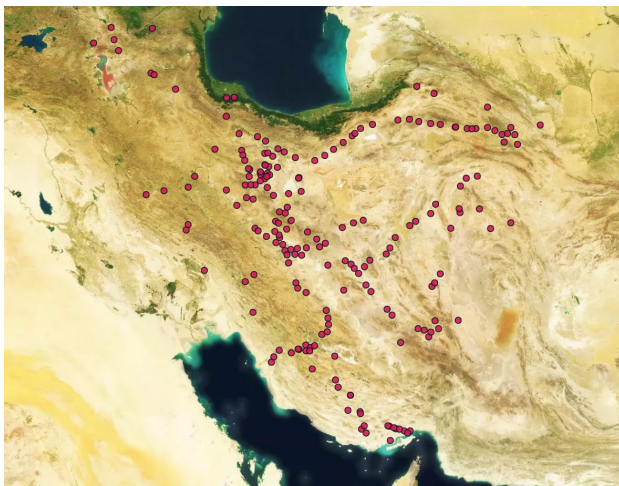
سطح اطمینان	تشخیص صحیح	تشخیص اشتباه
۰.۷۵	۱۰	۲۳۱۰
۰.۵۰	۱۰	۱۱۰۶
۰.۲۵	۹	۸۷۹

۳-۹- انتخاب تصاویر نمونه منفی

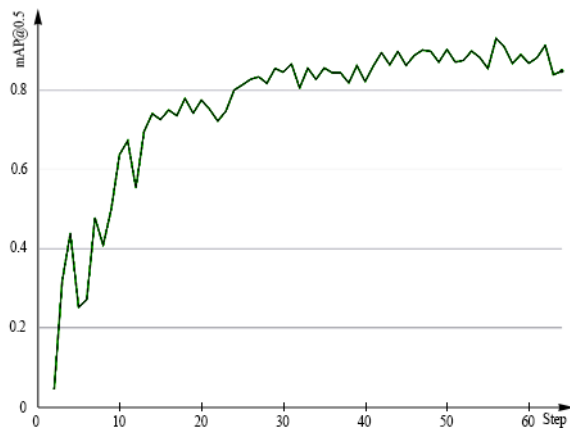
برخی از تصاویر نمونه منفی حاصل از گام ۸ در شکل (۱۱) نمایش داده شده است. لازم به ذکر است که تصاویر افزوده شده به پایگاه داده، فاقد کادرهای محدودکننده در تصاویر شکل (۱۰) هستند.

۴- بررسی نتایج

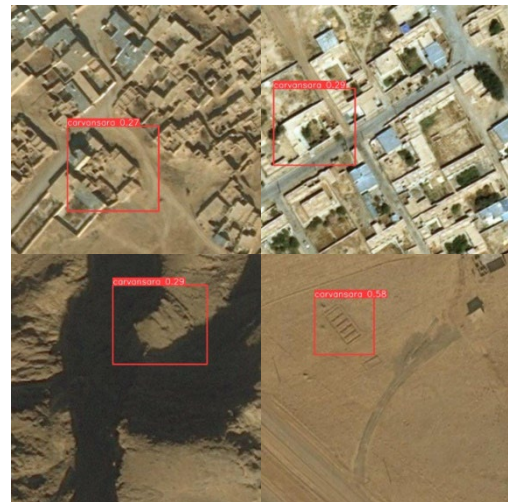
به منظور استفاده بهینه از پایگاه داده موقعیت‌های جغرافیایی ایجاد شده در این مقاله، یک نقشه تعاملی^۱ با استفاده از کتابخانه keplergl ایجاد شده است (شکل (۱۱)). این نقشه در آدرس carvansarai.ai در دسترس است. نمودار تغییرات $mAP@0.5$ در گام‌های مختلف آموزش معماری YOLOv5 در شکل (۱۲) نشان داده شده است. بر اساس شکل (۱۲)، بهترین دقت این معماری در یافتن کاروانسراها، $mAP_{0.5} 91/43\%$ است که در گام ۵۸ آموزش به دست آمده است. مقادیر recall و precision (شکل‌های (۱۳)، (۱۴)) الگوریتم YOLOv5 به ترتیب برابر 0.9629 و 0.8387 بوده که بیانگر توانایی بالای مدل آموزش داده شده در تشخیص کاروانسراها است.



شکل (۱۱): نقشه پراکندگی موقعیت جغرافیایی کاروانسراهای ایران



شکل (۱۲): نمودار تغییرات $mAP@0.5$



شکل (۱۰): برخی از تصاویر نمونه منفی حاصل از گام ۸

^۱ interactive



شکل (۱۵): خروجی الگوریتم YOLOv5 بر روی تصاویر راستی آزمایی



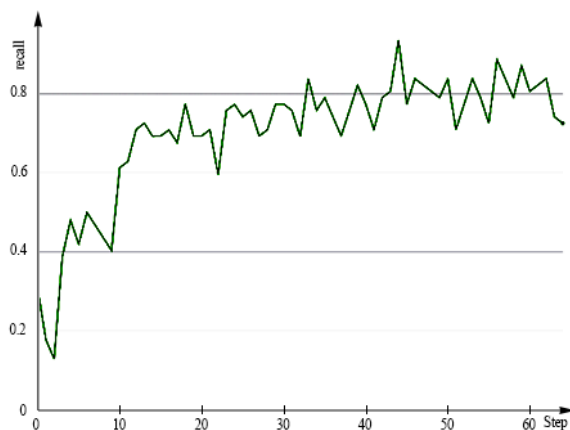
شکل (۱۶): کاروانسرای جدید یافته شده توسط الگوریتم پیشنهادی



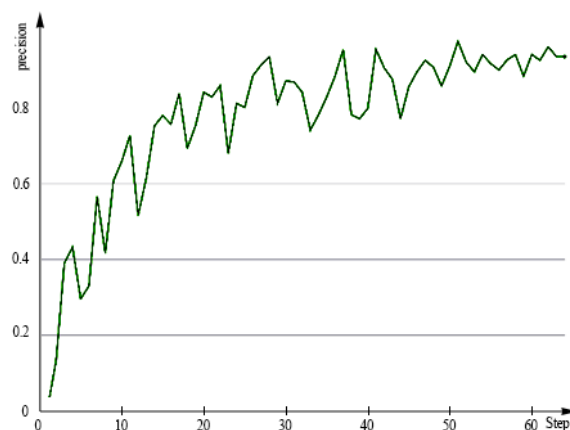
شکل (۱۷): سه مکان تاریخی یافته شده توسط الگوریتم پیشنهادی که کاروانسرا بودن یا نبودن آن‌ها را نمی‌توان از روی تصاویر ماهواره‌ای به شکل قطعی مشخص نمود.

۵ - مقایسه نتایج مدل پیشنهادی با ادبیات تحقیق

الگوریتم تناظریابی الگویی با عملیات لبه‌یابی سوبل با هسته دو مربعی معرفی شده توسط زنجانی و مالیان [۲] تنها پژوهشی است که سابق بر این بر روی استخراج خودکار کاروانسراها بر روی تصاویر ماهواره‌ای انجام شده است. دقت روش زنجانی و مالیان بر روی تنها ۵۰ کاروانسرا بررسی شده است در حالی که در پایگاه داده مقاله حاضر، ۲۰۹ کاروانسرا وجود



شکل (۱۳): نمودار تغییرات recall



شکل (۱۴): نمودار تغییرات precision

نمونه‌ای از خروجی الگوریتم YOLOv5 بر روی تصاویر راستی آزمایی پایگاه‌داده، در شکل (۱۵) نشان داده شده است. با اجرای شبکه بر روی تصویر $2/5 \times 2/5$ کیلومتری دانلود شده، مشخص گردید که الگوریتم هر ۲۵ کاروانسرای مورد بحث را به درستی تشخیص می‌دهد، اما تعداد تشخیص‌های اشتباه بیشتر از حداکثر میزان مورد نظر است (برای نمونه بر روی یک تصویر $2/5$ در $2/5$ کیلومتری، ۸۵۰ تصویر به عنوان کاروانسرای کاندید استخراج می‌گردید). جهت کاهش خطای مثبت کاذب، تصاویر تشخیص اشتباه، به عنوان نمونه منفی به پایگاه‌داده افزوده شد و آموزش شبکه YOLOv5 مجدداً تکرار گردید. با چندین مرتبه انجام مراحل آموزش شبکه YOLOv5 و افزودن تصاویر نمونه منفی به پایگاه‌داده، میزان خطای مثبت کاذب الگوریتم از ۸۵۰ تصویر به ۱۰۱ تصویر کاندید کاهش یافت. در بررسی‌های انجام شده مشخص شد که الگوریتم پیشنهادی این مقاله، موفق به تشخیص یک کاروانسرای جدید در این تصاویر شده است که این کاروانسرا در منابع موجود [۶، ۲۷]، معرفی نشده است (شکل ۱۶). همچنین الگوریتم پیشنهادی، ۳ مکان جدید را یافته که شباهت زیادی به ساختار کاروانسرا دارند (شکل ۱۷)، اما این مورد را از روی تصاویر ماهواره‌ای نمی‌توان به طور قطعی بیان کرد و این مسئله نیاز به بررسی میدانی جهت نتیجه‌گیری نهایی دارد.



مراجع:

- [۱] کیانی، م. ی. معماری ایران: دوران اسلامی. سازمان مطالعه و تدوین کتب علوم انسانی دانشگاهها (سمت)، ۱۳۷۹.
- [۲] زنجانی، ف. مالیان، ع. "استخراج خودکار کاروانسراهای ایرانی از تصاویر ماهواره ای با بهره گیری از الگوریتم تناظریابی الگویی". سیزدهمین سمپوزیوم بین المللی پیشرفت های علوم و تکنولوژی. سرزمین پایدار، معماری و شهرسازی، ۱۳۹۷.
- [3] Belhi, Abdelhak, et al. "Deep learning and cultural heritage: the CEPROQHA project case study." *2019 13th International Conference on Software, Knowledge, Information Management and Applications (SKIMA)*. IEEE, 2019.
- [4] Kersten, T. P., and M. Lindstaedt. "Potential of automatic 3D object reconstruction from multiple images for applications in architecture, cultural heritage and archaeology." *International Journal of Heritage in the Digital Era* 1.3 (2012): 399-420.
- [5] Condorelli, Francesca, et al. "A neural networks approach to detecting lost heritage in historical video." *ISPRS International Journal of Geo-Information* 9.5 (2020): 297.
- [۶] زنجانی، ف. مالیان، ع. "ملاحظات هندسی و پرتوسنجی برای آشکارسازی و مستندنگاری کاروانسراها"، نخستین همایش ملی مستندنگاری میراث طبیعی و فرهنگی، ۱۳۹۶.
- [7] Darma, I. Wayan Agus Surya, Nanik Suciati, and Daniel Siahaan. "A performance comparison of balinese carving motif detection and recognition using YOLOv5 and mask R-CNN." *2021 5th International Conference on Informatics and Computational Sciences (ICICoS)*. IEEE, 2021.
- [8] Llamas, Jose, et al. "Classification of architectural heritage images using deep learning techniques." *Applied Sciences* 7.10 (2017): 992.
- [9] Oses, Noelia, Fadi Dornaika, and Abdelmalik Moujahid. "Image-based delineation and classification of built heritage masonry." *Remote Sensing* 6.3 (2014): 1863-1889.
- [10] Groener, Austen, Gary Chern, and Mark Pritt. "A comparison of deep learning object detection models for satellite imagery." *2019 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*. IEEE, 2019.
- [11] Balaniuk, R., O. Isupova, and S. Reece. "Mining and tailings dam detection in satellite imagery using deep learning." *arXiv 2020.* *arXiv preprint arXiv:2007.01076*.
- [12] Wang, Tao, et al. "Landslide detection based on improved YOLOv5 and satellite images." *2021 4th International Conference on Pattern Recognition and Artificial Intelligence (PRAI)*. IEEE, 2021.
- [13] Jindal, Manik, et al. "Aircraft Detection from Remote Sensing Images using YOLOV5 Architecture." *2022 6th International Conference on Devices, Circuits and Systems (ICDCS)*. IEEE, 2022.
- [14] Zhanying, Zhang, and Chen Xinyuan. "Research on Forest Fire Detection Algorithm Based on Yolov5." *2021 International Conference on Intelligent Computing, Automation and Systems (ICICAS)*. IEEE, 2021.
- [15] Hu, Mingdi, Yaqian Ren, and Haoxin Chai. "Forest Fire Detection Based on Improved YOLOv5." *Proceedings of the 2021 4th International Conference on Artificial Intelligence and Pattern Recognition*. 2021.
- [۱۶] کیانی، م. کلایس، و. فهرست کاروانسراهای ایران، جلد اول، سازمان میراث فرهنگی کشور، ۱۳۶۳.

دارد. همچنین مراحل استخراج ویژگی‌ها در روش زنجانی و مالیان به صورت دستی برای کاروانسراهای مربعی شکل مورد مطالعه انجام شده است که باعث می‌شود روش ارائه شده توسط این محققین برای تشخیص کاروانسراهای چندضلعی و یا با پلان متفرقه، قابل استفاده نباشد. در صورتی که در روش پیشنهادی فرایند استخراج ویژگی‌ها به صورت خودکار و توسط الگوریتم یادگیری عمیق انجام شده و نیازی به تغییر در فرایند پیشنهادی برای تشخیص سایر کاروانسرا نیست. اما شاید بتوان مهم‌ترین ضعف روش زنجانی و مالیان را در عدم بررسی دقت الگوریتم ارائه شده بر روی تصاویر سایر بناها و مکان‌ها دانست، زیرا با توجه به ساختار بسیار مشابه کاروانسرا با برخی از مدارس، مساجد و ابنیه دیگر، و نوع ویژگی‌های انتخاب شده در آن مقاله، می‌توان توقع داشت که حساسیت^۱ روش ارائه شده توسط زنجانی و مالیان کم بوده و در عمل، امکان استفاده از آن روش در دنیای واقعی وجود نداشته باشد. در صورتی که در روش پیشنهادی این مقاله، با افزودن تصاویر نمونه منفی حاصل از اجرای روش بر روی ناحیه‌ای حدود ۱۵۶ کیلومتر مربعی و آموزش چندین باره الگوریتم YOLOv5 می‌توان از عملکرد روش پیشنهادی در دنیای واقعی مطمئن بود. لازم به ذکر است که به دلیل در دسترس نبودن داده‌های مقاله زنجانی و مالیان، مقایسه مستقیم دقت روش پیشنهادی زنجانی و مالیان و الگوریتم پیشنهادی این مقاله امکان پذیر نمی‌باشد.

۶ - نتیجه گیری

دقت $0.5_{\text{map}}/91/43\%$ به دست آمده توسط الگوریتم YOLOv5 در مکان‌یابی کاروانسراها با توجه به محدود بودن تعداد داده‌های در دسترس و یافتن یک کاروانسرای جدید توسط الگوریتم پیشنهادی در مساحت ۱۵۶ کیلومتر مربعی بررسی شده در این تحقیق، بسیار امیدوار کننده است. همچنین پس‌پردازش‌های انجام شده بر روی نتایج هنگام استفاده از تکنیک پنجره لغزان، که باعث کاهش خطای مثبت کاذب الگوریتم پیشنهادی شده است، باعث گردیده تا بتوان از روش پیشنهادی این مقاله به صورتی عملی و در دنیای واقعی استفاده نمود. نتایج حاصل از آزمایش‌های انجام شده بیانگر آن است که تکنیک‌های پردازش تصویر مبتنی بر یادگیری عمیق، می‌توانند برای حل مسائل مرتبط به میراث فرهنگی استفاده شوند. لازم به ذکر است که پایگاه‌داده موقعیت‌های جغرافیایی و پایگاه‌داده تصاویر ماهواره‌ای کاروانسراهای که در این تحقیق ایجاد شده‌اند به شکل منبع باز در آدرس

<https://caravanserai.ai-hue.ir>

<https://caravanserai.ai-hue.ir/map.htm>

برای دانلود در دسترس محققین قرار گرفته‌اند.

¹ Sensitivity

- [17] Redmon, Joseph, et al. "You only look once: Unified, real-time object detection." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
- [18] Mohamed, Hussam El-Din, et al. "Msr-yolo: Method to enhance fish detection and tracking in fish farms." *Procedia Computer Science* 170 (2020): 539-546.
- [19] Jocher, Glenn, et al. *YOLOv5*. Zenodo. <https://doi.org/10.5281/zenodo.5563715> (2021)
- [20] Zhou, Junchi, et al. "Ship target detection algorithm based on improved YOLOv5." *Journal of Marine Science and Engineering* 9.8 (2021): 908.
- [21] Yao, Jia, et al. "A real-time detection algorithm for Kiwifruit defects based on YOLOv5." *Electronics* 10.14 (2021): 1711.
- [22] Fang, Yiming, et al. "Accurate and automated detection of surface knots on sawn timbers using YOLO-V5 model." *BioResources* 16.3 (2021): 5390.
- [23] Wang, Chien-Yao, et al. "CSPNet: A new backbone that can enhance learning capability of CNN." *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*. 2020.
- [24] Powers, David MW. "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation." *arXiv preprint arXiv:2010.16061* (2020).
- [25] Chen, Sheng, Bernard Mulgrew, and Peter M. Grant. "A clustering technique for digital communications channel equalization using radial basis function networks." *IEEE Transactions on neural networks* 4.4 (1993): 570-590.
- [26] Rezatofighi, Hamid, et al. "Generalized intersection over union: A metric and a loss for bounding box regression." *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2019.
- [۲۷] کیانی، م، کلایس، و. و. فهرست کاروانسراهای ایران، جلد دوم، سازمان میراث فرهنگی کشور، ۱۳۸۶.
- [28] LabelIm. Open Annotation Tool. <http://labelme.csail.mit.edu/Release3.0>. Last accessed 24 July 2021

COPYRIGHTS

©2024 by the authors. Published by the Islamic Azad University Shiraz Branch. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 International (CC BY 4.0)

<https://creativecommons.org/licenses/by/4.0>

