نشریه علمی مهندسی فناوری اطلاعات مکانی

سال یازدهم، شماره سوم، پاییز ۱۴۰۲

Vol.11, No.3, Autumn 2023

۱۰۲ -۸۵

مقاله پژوهشی

کاربرد شبکه‌های عصبی کانولوشن آشکارساز چندجعبه تک شات (*SSD*) در تشخیص و استخراج خودکار میله چاه‌های قنات از تصاویر ماهواره‌ای گوگل ارث

مصطفی کابلی‌زاده [۱*]، محمد عباسی [۲]

۱- استادیار گروه سنجش از دور و سیستم اطلاعات جغرافیایی، دانشکده علوم زمین، دانشگاه شهید چمران اهواز

۲- دانشجوی دکتری گروه سنجش از دور و سیستم اطلاعات جغرافیایی، دانشکده علوم زمین، دانشگاه شهید چمران اهواز

تاریخ دریافت مقاله: ۱۴۰۲/۰۴/۲۱ تاریخ پذیرش مقاله: ۱۴۰۲/۰۹/۰۱

چکیده

شناسایی قنات‌ها و تهیه نقشه آن‌ها یکی از وظایف حیاتی در علم مدیریت منابع آب‌های زیرزمینی است. با این حال، تهیه نقشه قنوات دشوار است، زیرا روش‌های مرسوم مانند نقشه‌برداری میدانی پر هزینه و زمانبر هستند و در برخی مواقع بدلیل ساختار دریچه چاه‌های قنات، برداشت زمینی با چالش‌های متعددی مواجه است. شناسایی عوارض از داده‌های تصویری با استفاده از تکنیک‌های پردازش تصاویر کامپیوتری و شبکه-های عصبی یکی از امیدوارکننده‌ترین تکنیک‌ها برای شناسایی چاه‌های قنات است. با توجه به اینکه تصاویر ماهواره‌ای سامانه گوگل ارث، تنها تصاویر ماهواره‌ای با قدرت تفکیک بالای مکانی، بدون هزینه و در دسترس هستند، در این تحقیق از تصاویر ماهواره‌ای گوگل ارث استفاده شده است. در این پژوهش بیش از ۶۰۰ نمونه آموزشی از دهانه میله چاه‌های قنات تهیه شده است و از شبکه عصبی کانولوشون بر مبنای آشکارساز چندجعبه تک شات با پایه شبکه *ResNet*، برای کشف و استخراج خودکار مکان هندسی میله چاه‌های قنوات با توجه به سرعت بالاتر انجام پردازش، استفاده شده است. مدل پیشنهادی توسط نمونه‌های آموزشی بر مبنای ۸۵ درصد داده آموزشی و ۱۵ درصد داده اعتبارسنجی، با ۵۰ دوره تکرار و دقت ۸۹ درصد آموزش داده شده است. مدل آموزش دیده بر روی تصویر منطقه مورد مطالعه برای کشف میله چاه‌های قنات اجرا شده است. نتایج نشان می‌دهد که این مدل می‌تواند با معیار دقت برابر ۰٫۹۱ و معیار بازیابی برابر ۰٫۸۲ و معیار *F1Score* برابر ۰٫۸۶ در کشف موقعیت مکانی دهانه میله چاه‌های قنات که دارای شکل مناسب هستند، عمل نماید. برای مناطقی که شکل دهانه میله چاه‌ها تخریب شده است، دقت تشخیص و استخراج موقعیت میله چاه‌ها برابر ۰٫۶۵ است. نتایج این پژوهش نشان می‌دهد که امکان تشخیص و استخراج اتوماتیک میله چاه‌های قنات با دقت مناسب از تصاویر ماهواره‌ای دانلود شده از سامانه گوگل ارث امکان‌پذیر است.

کلید واژه‌ها : شبکه عصبی کانولوشن، آشکارساز چندجعبه تک شات، قنات، *ResNet*، گوگل ارث.

* نویسنده مکاتبه کننده: اهواز، دانشگاه شهید چمران، دانشکده علوم زمین، گروه سنجش از دور و *GIS* .

تلفن: ۰۹۱۶۳۴۷۰۱۷۷

Email: m.kabolizade@scu.ac.ir

[Downloaded from jgit.kntu.ac.ir on 2026-02-20]

[DOI: 10.61186/jgit.11.3.85]

۱- مقدمه

قنات اصلی‌ترین نماد رابطه انسان و آب است. قنات‌ها علاوه بر آبادانی و توسعه کشاورزی پایدار می‌توانند در رونق گردشگری قنات و در نتیجه اشتغال‌زایی نقش مهمی داشته باشند. قنات‌ها هزاران سال است که نشان‌دهنده یک پیشرفت تکنولوژیکی و راه‌حل عمده برای تامین آب در مناطق خشک و نیمه خشک بوده‌اند و تاریخچه آن‌ها با پیدایش چندین تمدن در طول تاریخ اجین شده است[۱ و ۲]. حفظ قنات به عنوان یک منبع پایدار آب در مناطق خشک بسیار ضروری است. برای مدیریت قنوات نیاز اصلی، تهیه داده‌های مکانی قنوات است. در این راستا اولین گام تهیه نقشه میله چاه‌های قنات برای برنامه‌ریزی بهینه است. تهیه نقشه‌های قنوات به روش‌های مستقیم زمینی مانند برداشت با دوربین‌های نقشه‌برداری یا سیستم‌های تعیین موقعیت ماهواره‌ای معمولا زمانبر و نیاز به هزینه بالایی دارد. همچنین با توجه به ساختار دهانه میله چاه‌های قنات برداشت زمینی با چالش‌های خاصی مواجه است. لذا استفاده از تصاویر هوایی یا ماهواره‌ای با قدرت تفکیک مکانی بالا می‌تواند یک راه‌حل بهینه برای تهیه نقشه قنوات باشد. امضای سیستم‌های قنات زیرزمینی بر روی تصاویر هوایی یا ماهواره‌ای، دهانه‌های دایره‌ای شکل میله‌های عمودی چاه‌های آن‌ها است که می‌تواند در تشخیص خودکار الگو بسیار موثر باشد. نیاز اولیه این روش‌ها دسترسی به تصاویر با قدرت تفکیک بالای مکانی مانند تصاویر پهپاد (پرنده هدایت‌پذیر از دور)، هوایی و ماهواره‌ای با امکان تشخیص میله‌های چاه‌های قنات می‌باشد. تهیه تصاویر پهپاد و بطور کلی عملیات پرواز و پردازش‌های آن‌ها هزینه‌بر می‌باشد، هرچند دارای دقت بالاتری می‌باشند. در این پژوهش از تصاویر ماهواره‌ای برای کشف موقعیت مکانی قنوات استفاده شده است. با توجه به اینکه تنها تصاویر ماهواره‌ای با قدرت تفکیک مکانی بالا که در کشور بدون هزینه در دسترس هستند، تصاویر ماهواره‌ای گوگل ارث می‌باشند. در این تحقیق از تصاویر ماهواره‌ای گوگل ارث برای استخراج موقعیت مکانی میله چاه‌های قنات استفاده شده است. هدف این پژوهش استخراج خودکار میله چاه‌های قنوات از تصاویر ماهواره‌ای گوگل ارث می‌باشد. در این راستا نیاز به استفاده از روش‌های نوین کشف اشیا از تصاویر مبتنی بر بینایی کامپیوتر است.

تشخیص ویژگی‌های بصری متمایز یکی از مهم‌ترین مشکلات در تمام مراحل بینایی کامپیوتر از جمله در زمینه سنجش از دور است. از آنجایی که روش‌های مرسوم مانند ماشین‌های بردار پشتیبان[۳] نیاز به استخراج ویژگی‌های زمان‌بر و طراحی دستی دارند، تلاش‌های قابل توجهی برای توسعه روش‌هایی برای استخراج خودکار ویژگی‌ها انجام شده است. اخیراً یادگیری عمیق به یک راه‌حل جدید برای تشخیص اهداف و طبقه‌بندی تصاویر تبدیل شده است، زیرا این روش جدید نیازی به استخراج دستی ویژگی‌ها ندارد[۴]. یادگیری عمیق نوعی تکنیک یادگیری ماشینی است که از الگوریتم‌هایی الهام گرفته از ساختار و عملکرد مغز انسان استفاده می‌کند که شبکه‌های عصبی مصنوعی نامیده می‌شوند[۵ و ۶]. یادگیری عمیق در دهه گذشته به دلیل توانایی زیاد در یادگیری، به شدت مورد توجه قرار گرفته است. یادگیری عمیق به مدل‌هایی که بر اساس لایه‌های پردازشی متعدد ساخته شده‌اند، اجازه می‌دهد تا نمایش نمونه‌های داده را با چندین سطح انتزاع بیاموزند. یادگیری عمیق مدل‌هایی را که بر اساس لایه‌های چندگانه تشکیل شده‌اند، قادر می‌سازد تا بازنمایی نمونه‌های داده را با چندین محدوده سطوح انتزاعی بیاموزند[۷]. یادگیری عمیق بر اساس عملکرد پیشرفته خود، در حوزه‌های مختلفی مانند بینایی کامپیوتر[۸]، تشخیص گفتار[۹] و بازیابی اطلاعات[۱۰] استفاده شده است. یادگیری عمیق شامل یادگیری همزمان ویژگی‌ها و طبقه‌بندی‌کننده‌ها می‌شود و از داده‌های آموزشی برای دسته‌بندی محتوای تصویر بدون مشخصات قبلی ویژگی‌های تصویر استفاده می‌کند. در میان تمام

[Downloaded from jgit.kntu.ac.ir on 2026-02-20]
[DOI: 10.61186/jgit.11.3.85]

شبکه‌های مبتنی بر یادگیری عمیق، شبکه عصبی کانولوشنال برای یادگیری ویژگی‌های بصری در زمینه بینایی کامپیوتر، از جمله سنجش از دور، محبوب‌ترین است. تحقیقات اخیر نشان داده است که شبکه عصبی کانولوشنال برای کاربردهای مختلف موثر است[۱۱]. شبکه‌های عصبی کانولوشنال اکنون در حوزه مکانی – طیفی استفاده می‌شوند[۱۲]. شبکه عصبی کانولوشنال یک فناوری امیدوارکننده در زمینه سنجش از دور است[۱۳ و ۱۴]. همچنین اخیراً، یادگیری عمیق به دلیل عملکرد برتر از نظر دقت و تطبیق مدل‌ها، به عنوان یک ابزار قدرتمند برای جامعه سنجش از دور ظاهر شده است [۱۵]. یادگیری عمیق به لطف پیشرفت‌های اخیر مانند افزایش قدرت پردازنده‌ها، بهبود در پردازنده‌های گرافیکی و افزایش حجم داده‌ها (به عنوان مثال، فیلم‌ها و تصاویر) جهش‌های بزرگی به ویژه، در بخش بینایی کامپیوتر داشته است. علم سنجش از دور شاهد افزایش عظیمی در تولید و بهبود قدرت تفکیک تصاویر رقومی اخذ شده از سنجنده‌های نصب شده بر روی سکوهای هوایی یا ماهواره‌هایی است که تقریباً هر زاویه از سطح زمین را پوشش می‌دهند. این رشد در داده‌ها، جامعه علوم زمین و سنجش از دور را به استفاده از الگوریتم‌های یادگیری عمیق برای حل وظایف مختلف سنجش از دور سوق داده است. در این راستا در این تحقیق از روش‌های یادگیری عمیق جهت استخراج خودکار میله چاه‌های قنات از تصاویر گوگل ارث استفاده شده است. پژوهش‌های کمی در خصوص کشف قنات از تصاویر ماهواره‌ای انجام شده است که به مهمترین آن‌ها مانند استفاده از تصاویر ماهواره‌ای کرونا[۱]برای تشخیص خودکار میله چاه‌های قنات با استفاده از شبکه‌های عصبی کانولوشنال عمیق در منطقه کردستان عراق اشاره کرد. مطالعه تطبیقی فناوری قنات در مقیاس منطقه‌ای در حال حاضر غیرممکن است، زیرا هیچ‌کس زمان لازم برای ایجاد نقشه‌های دقیق از میله چاه‌های قنات را به صورت دستی ندارد. می‌توان از یادگیری انتقال و تقویت داده‌ها برای غلبه بر مشکل مجموعه داده‌های کوچک استفاده کرد. محققان در این پژوهش نشان دادند که یادگیری عمیق، حتی با مجموعه داده‌های کوچک هم، می‌تواند با موفقیت برای تشخیص خودکار میله چاه‌های قنات اعمال شود. از آنجایی که زیرساخت‌های قنات در بسیاری از کشورها و زمینه‌ها یافت می‌شوند و چاه‌های آن‌ها دارای امضاهای مشابهی هستند، به‌تدریج می‌توان مجموعه داده‌های آموزشی را از مکان‌های گسترده‌ای در جهان جمع‌آوری و شبکه عصبی کانولوشنال را برای این کار بهینه کرد [۱۶]. همچنین محققان با استفاده از تصاویر گوگل ارث و مدل شما فقط یک‌بار به تصویر نگاه می‌کنید (*YOLOv5*[۲])، بصورت خودکار نقشه قنات در حوضه تورپان در منطقه خودمختار اویغور سین کیانگ را با دقت کلی ۸۶ درصد تهیه نمودند [۱۷]. در برخی موارد ممکن است بصورت دستی عملیات تشخیص و استخراج موقعیت مکانی میله چاه‌های قنات از تصاویر ماهواره‌ای انجام شود که به توجه به محدوده قنوات و زمانبر بودن فرآیند دستی از نظر هزینه نیز مقرون به صرفه نمی‌باشد. همچنین تشخیص و استخراج توسط عامل‌های انسانی می‌تواند یکسان نباشد. لذا با حذف روند رقومی‌سازی دستی با تشخیص و استخراج اتوماتیک، می‌توان زمان ارزشمند و محدود محققین قنات را برروی اعتبارسنجی و تجزیه و تحلیل نتایج متمرکز کرد. علاوه بر این، خودکارسازی تشخیص میله چاه‌های قنات، فرصت‌های تحقیقاتی جدیدی را باز می‌کند. هدف اصلی این پژوهش کشف محل چاه‌های قنات از تصاویر ماهواره‌ای با قدرت تفکیک بالای مکانی گوگل ارث بر مبنای شبکه‌های عصبی کانوولوشن آشکارساز چندجعبه تک شات (*SSD*)[۳]می‌باشد. نوآوری این پژوهش استفاده از مدل آشکارساز چندجعبه تک شات [۱۸] برای یک مدلسازی مجموعه یادگیری عمیق به منظور کاهش پیچیدگی اجرا و افزایش قابلیت انتقال طرح برای الگوهای مکانی با استفاده از تصاویر سامانه گوگل ارث در دسترس در ایران است.

[۱] *CORONA*

[۲] *You Only Look Once*

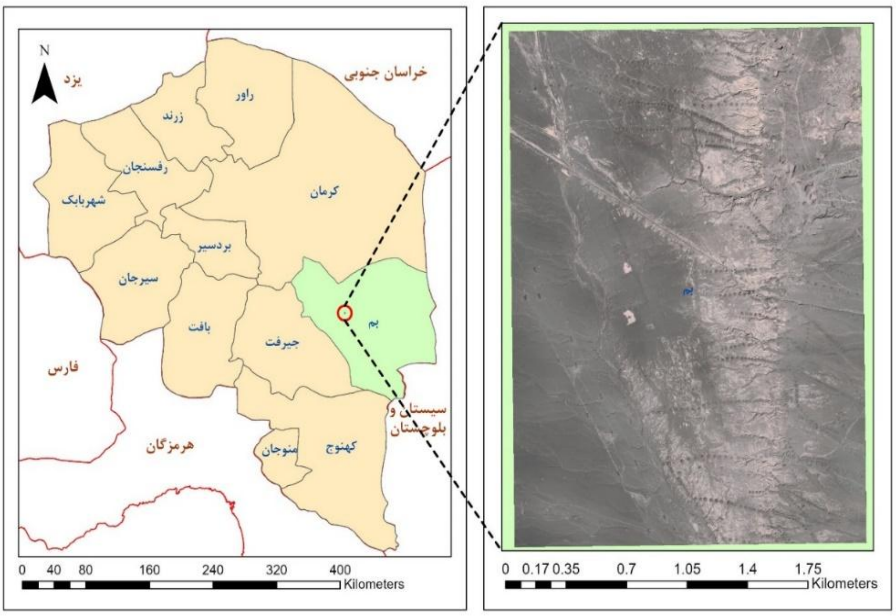
[۳] *Single Shot Multibox Detector*

[Downloaded from jgit.kntu.ac.ir on 2026-02-20]

[DOI: 10.61186/jgit.11.3.85]

۲- مواد و روش‌ها

منطقه مورد مطالعه در این پژوهش قنوات شهرستان بم در استان کرمان است. این شهرستان از نظر جغرافیایی در ۵۸ درجه و ۲۱ دقیقه‌ی طول شرقی و ۲۹ درجه و ۶ دقیقه‌ی عرض شمالی و ارتفاع از سطح دریا ۱۰۵۰ متر می‌باشد و دارای آب و هوای خشک است. بم یکی از شهرهای تاریخی ایران است. قنات یکی از مهمترین تامین‌کننده‌های آب کشاورزی در شهرستان بم می‌باشد و دارای قنوات متعدد می‌باشد. دبی قنوات بم، مهمترین شاخص این قنوات است. دشت بم در حدود ۷۰۲ کیلومتری جنوب شرقی مرکز استان کرمان و در حاشیه کویر لوت واقع شده که جزء مناطق خشک کشور محسوب می‌شود. شکل (۱) موقعیت منطقه مورد مطالعه در این پژوهش را نشان می‌دهد.



شکل ۱: منطقه مورد مطالعه

در این مطالعه، از تصاویر ماهواره‌ای گوگل ارث به عنوان منبع تصاویر ماهواره‌ای استفاده شده است. تصاویر ماهواره‌ای گوگل ارث منابع داده مختلفی دارند، که می‌توان به تصاویر ماهواره لندست با قدرت تفکیک مکانی متوسط تا تصاویر ماهواره‌ای کوئیک برد[1] با قدرت تفکیک مکانی بالا اشاره نمود. در این پژوهش، از تصاویر ماهواره‌ای با قدرت تفکیک مکانی ۰٫۶۵ متر دانلود شده از سامانه گوگل ارث استفاده شده است. تصاویر مورد استفاده مربوط به سال ۱۳۹۹ (۲۰۲۱) در محدوده شهرستان بم شکل (۱) می‌باشد. این تصاویر توسط نرم‌افزارهای دانلود تصاویر از نرم‌افزار گوگل ارث بصورت ژئورفرنس شده در سیستم تصویر *UTM*[2] زون ۴۰ استخراج شده است. شکل (۲) وضوح مکانی تصویر تعدادی از میله چاه‌های قنات در منطقه مورد مطالعه را نشان می‌دهد.

[1] *QuickBird*

[2] *Universal Transverse Mercator*

[Downloaded from jgit.kntu.ac.ir on 2026-02-20]

[DOI: 10.61186/jgit.11.3.85]

شکل ۲: قدرت تفکیک مکانی میله چاه‌های قنات در تصاویر گوگل ارث منطقه مورد مطالعه

۲-۱- شبکه‌های عصبی کانولوشن (*CNN*)[۱]

در این پژوهش، از شبکه‌های عصبی کانولوشن استفاده شده است که نوع خاصی از شبکه‌های عصبی پیشخور هستند که از چندین لایه کانولوشنال[۲] و لایه‌های ادغام[۳] تشکیل شده‌اند (شکل (۳)).

شبکه عصبی پیشخور یک شبکه عصبی مصنوعی است که در آن اتصالات بین گره‌ها یک چرخه را تشکیل نمی‌دهند. این شبکه‌ها که شبیه‌سازی فعالیت‌های عصبی در مغز را انجام می‌دهند، عموماً به‌عنوان سیستم‌هایی از واحدهای پردازشی به هم پیوسته (نرون‌های مصنوعی) ارائه می‌شوند که می‌توانند مقادیر را از ورودی‌ها محاسبه کنند، و در نتیجه یک خروجی ایجاد می‌کنند که ممکن است در واحدهای بعدی استفاده شود. نورون‌های مصنوعی اساساً واحدهای پردازشی هستند که برخی از عملیات را بر روی چندین متغیر ورودی محاسبه می‌کنند و معمولاً یک خروجی دارند که از طریق تابع فعال‌سازی محاسبه می‌شود. به طور معمول، یک نورون مصنوعی دارای وزن w_i است که نشان‌دهنده درجه ارتباط بین نورون‌های مصنوعی است که دارای متغیرهای ورودی x_i و یک بردار آستانه b است.

از نظر ریاضی، کل ورودی و خروجی نورون‌های مصنوعی را می‌توان به صورت رابطه‌های (۱) و (۲) توصیف کرد[۴].

رابطه(۱)
$$u = \sum_i w_i x_i$$

رابطه(۲)
$$z = f(u+b) = f\left(\sum_i w_i x_i + b\right)$$

در رابطه (۱) و (۲)، پارامترهای u، z، x، w و b به ترتیب کل ورودی، خروجی، متغیرهای ورودی، وزن‌ها و بایاس را نشان می‌دهند. تابع f، یک تابع فعال‌سازی را نشان می‌دهد که می‌تواند یک تابع خطی یا غیرخطی، سیگموئید، هذلولی یا یکسو شده باشد.

در این پژوهش از یک تابع خطی اصلاح شده به عنوان تابع فعال‌سازی استفاده شده است و این تابع به عنوان واحد خطی اصلاح شده (*ReLU*)[۴] نامیده می‌شود. تابع *ReLU* را می‌توان به صورت رابطه (۳) تعریف کرد[۱۹]:

رابطه(۳)
$$f(u) = \max(0,u) = f\begin{Bmatrix} u(u>0) \\ 0(u \le 0) \end{Bmatrix}$$

تابع فعال‌سازی *ReLU* باعث می‌شود خروجی برخی از نورون‌ها صفر شود که باعث پراکندگی شبکه و کاهش وابستگی متقابل پارامترها می‌شود و به طور موثر از مشکل بیش برازش نیز جلوگیری می‌کند[۲۰].

---

[۱] *Convolutional Neural Network*

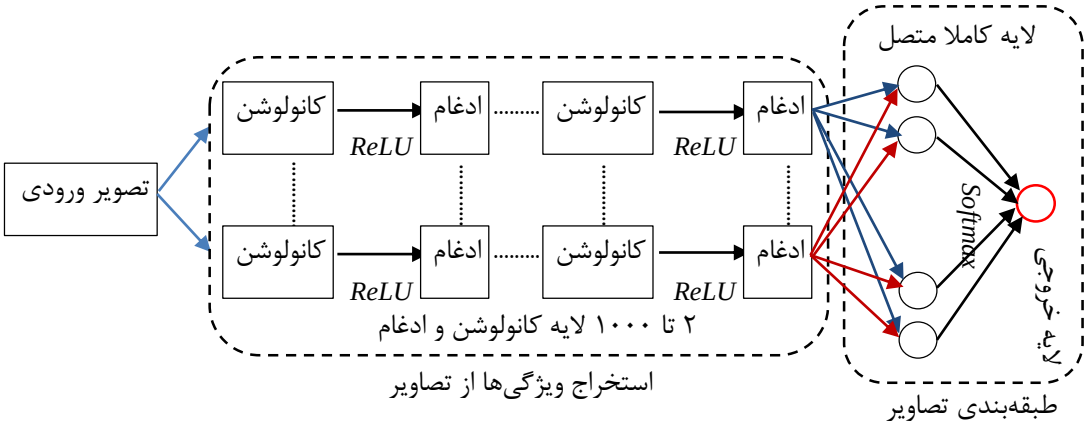
[۲] *Convolutional*

[۳] *Pooling*

[۴] *Rectified Linear Unit*

[Downloaded from jgit.kntu.ac.ir on 2026-02-20]

[DOI: 10.61186/jgit.11.3.85]



شکل ۳: نمایش شماتیک شبکه‌های کانولوشن [۴]

یک شبکه عصبی کانولوشن از یک لایه کانولوشن و یک لایه ادغام تشکیل شده است. لایه کانولوشن در استخراج ویژگی‌ها از تصاویر نقش دارد. در این فرآیند، یک فیلتر با اندازه ثابت برروی تصاویر اعمال می‌شود و الگوهای درجات خاکستری باندها را از تصاویر استخراج می‌کند. بعد از هر لایه کانولوشن، لایه‌های ادغام وجود دارد که برای کاهش واریانس ویژگی‌ها ایجاد می‌شود که با محاسبه عملیات یک ویژگی خاص بر روی یک قسمت از تصویر انجام می‌شود.

لایه ادغام دو عملکرد دارد. اولین عملکرد کاهش حساسیت موقعیت ویژگی‌ای است که از لایه کانولوشن استخراج می‌شود. به طوری که مقدار خروجی لایه ادغام حتی زمانی که موقعیت مقدار ویژگی استخراج شده توسط لایه کانولوشن در تصویر جابجا شود، تغییر نمی‌کند. عملکرد دوم بزرگ کردن میدان پذیرنده برای لایه‌های متوالی کانولوشن است. دو عملیات ممکن است در لایه‌های ادغام انجام شود، یعنی عملیات حداکثر و دیگری عملیات میانگین، که در آن مقادیر حداکثر و میانگین به ترتیب بر روی منطقه ویژگی اعمال می‌شوند. این فرآیند تضمین می‌کند که می‌توان نتایج یکسانی را حتی زمانی که ویژگی‌های تصویر دارای تفسیرها یا چرخش‌های کوچکی هستند، به دست آورد، و این برای طبقه‌بندی و تشخیص اشیا بسیار مهم است. بنابراین، لایه ادغام وظیفه نمونه‌برداری از خروجی لایه کانولوشن و حفظ موقعیت مکانی تصویر و همچنین انتخاب مفیدترین ویژگی‌ها برای لایه‌های بعدی را دارد. پس از چندین لایه کانولوشن و ادغام، لایه‌های کاملاً به هم متصل[۱] هستند که تمام نورون‌های لایه قبلی را گرفته و به تک تک نورون‌های لایه خود متصل می‌کنند.

در نهایت، به دنبال تمام لایه‌های کانولوشن، ادغام و کاملاً متصل، ممکن است از یک لایه طبقه‌بندی‌کننده برای محاسبه احتمال تعلق کلاس هر تصویر استفاده شود. در این پژوهش از تابع *Softmax* استفاده شده است که این تابع احتمالات هر کلاس هدف را بر روی تمام کلاس‌های هدف ممکن محاسبه می‌کند و به صورت رابطه (۴) نوشته می‌شود:

رابطه(۴)

$$y_k = soft\max_k(u_1,u_2,...,u_k) = \frac{e^{uk}}{\sum_{j=1}^{k} e^{u_j}}$$

در رابطه(۴)، k تعداد واحدهای خروجی و u نشان دهنده متغیرهای ورودی است.

برای ارزیابی عملکرد شبکه، باید یک تابع زیان[۲] نیز

[۱] *Fully Connected*

[۲] *Loss Function*

[Downloaded from jgit.kntu.ac.ir on 2026-02-20]
[DOI: 10.61186/jgit.11.3.85]

تعریف شود. تابع زیان، اثربخشی شبکه را در مدل‌سازی مجموعه داده آموزشی ارزیابی می‌کند. هدف از آموزش به حداقل رساندن خطای تابع زیان است.

آنتروپی متقاطع تابع *Softmax* به صورت رابطه (۵) تعریف می‌شود:

$$E = -\sum_{n=1}^{N}\sum_{k=1}^{k} t_{n.k} \log y_k \qquad \text{رابطه(۵)}$$

در رابطه (۵)، t نشان‌دهنده بردار داده‌های آموزشی است، K نشان‌دهنده کلاس محتمل و N نشان‌دهنده تعداد کل نمونه‌ها است.

۲-۲- آشکارساز چندجعبه تک شات

آشکارساز چندجعبه تک شات (*SSD*[۱]) دارای دو جزء است: یک مدل ستون اصلی[۲] و سر[۳]. مدل ستون اصلی معمولاً یک شبکه طبقه‌بندی تصویر از پیش آموزش دیده است که به عنوان استخراج کننده ویژگی عمل می‌کند. آشکارساز چندجعبه تک شات به جای استفاده از پنجره متحرک، تصویر را با استفاده از یک شبکه تقسیم می‌کند و هر سلول شبکه‌ای مسئول تشخیص اشیاء در آن ناحیه از تصویر است. تشخیص اشیاء به معنای پیش‌بینی کلاس و مکان یک شیئ در آن منطقه است. اگر شیئی وجود نداشته باشد، آن را به عنوان کلاس پس زمینه در نظر گرفته و مکان نادیده گرفته می‌شود. شبکه عصبی عمیق آشکارساز چندجعبه تک شات، فضای خروجی جعبه‌های مرزی را به مجموعه‌ای از جعبه‌های پیش‌فرض بر روی نسبت‌ها و مقیاس‌های مختلف در هر مکان نقشه ویژگی تفکیک می‌کند. در زمان پیش‌بینی، شبکه امتیازهایی را برای حضور هر دسته شیئ در هر جعبه پیش‌فرض ایجاد می‌کند و تنظیماتی را برای مطابقت بهتر با شکل شیء ایجاد می‌کند. علاوه بر این، این شبکه پیش‌بینی‌ها را از نقشه‌های چندگانه با وضوح‌های مختلف ترکیب می‌کند تا به طور طبیعی اشیاء با اندازه‌های مختلف را کنترل نماید. این روش نسبت به روش‌هایی دیگر ساده‌تر است، زیرا مراحل نمونه‌برداری مجدد پیکسل را حذف می‌کند و تمام محاسبات را در یک شبکه واحد انجام می‌دهد. در نهایت آموزش آشکارساز چندجعبه تک شات را آسان می‌کند و ادغام آن در سیستم‌هایی که به یک جزء تشخیص نیاز دارند را آسان می‌کند[۲۱].

شکل(۴) معماری کلی شبکه عصبی کانولوشنون با آشکارساز چندجعبه تک شات را نشان می‌دهد. در این پژوهش شبکه طبقه‌بندی تصویر از پیش آموزش دیده، شبکه *ResNet* می‌باشد.

۲-۳- مدل *ResNet*

معماری *ResNet* اولین بار در سال ۲۰۱۵ توسط هی و همکاران ارائه شد. در آن زمان، این معماری برنده چندین مسابقه شد و رکورد جدیدی را برای طبقه‌بندی مجموعه داده‌های تصاویر ثبت کرد. ایده اصلی مدل *ResNet* این است که بلوک‌های ساختمان آن برای «یادگیری توابع باقیمانده با ارجاع به ورودی‌های لایه، به جای یادگیری توابع غیر مرجع» طراحی شده‌اند[۲۲]. در این مدل یک ورودی دلخواه x به لایه ای از شبکه عصبی و خروجی بهینه y در نظر گرفته شده است. در مورد یک تابع بدون مرجع، شبکه مستقیماً تابع f را که y را تولید می‌کند، تقریب می‌زند: $y = f(x)$. منظور از تابع باقیمانده در این زمینه، این است که شبکه جدید اکنون $y = x + f(x)$ را تقریب می‌زند. بنابراین شبکه به جای اینکه یاد بگیرد که چگونه ورودی باید تبدیل شود تا خروجی مورد نظر را ایجاد کند، تفاوت بین ورودی و خروجی بهینه را یاد می‌گیرد. شکل (۵) نشان می‌دهد که چگونه این مفهوم بر روی لایه‌های یک شبکه کانولوشن اعمال می‌شود.

---

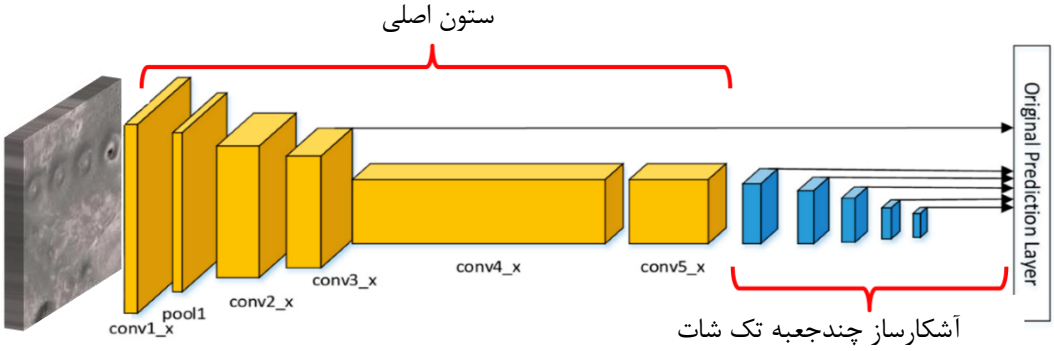
[۱] *Single Shot MultiBox Detector*

[۲] *Backbone*

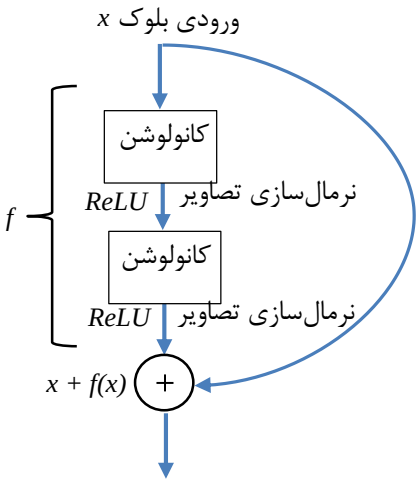
[۳] *Head*

[Downloaded from jgit.kntu.ac.ir on 2026-02-20]

[DOI: 10.61186/jgit.11.3.85]



شکل ۴: معماری کلی شبکه عصبی کانولوشنون با آشکارساز چندجعبه تک شات



شکل ۵: بلوک ساختمان مدل *ResNet* [۲۲]

بلوک شبکه *ResNet* یک ورودی دریافت می‌کند که در آن یک کانولوشن اعمال می‌شود، به دنبال آن نرمال-سازی دسته‌ای و فعال‌سازی *ReLU* انجام می‌شود. این کار برای بار دوم تکرار می‌شود و ورودی به نقشه ویژگی حاصل اضافه می‌شود. با این معماری یادگیری تبدیل هویت آسان‌تر است، زیرا بهینه‌ساز فقط باید وزن‌ها را در لایه‌های کانولوشن به صفر برساند. مشاهدات مدل‌های قبلی این بود که افزودن لایه‌های بیشتر به شبکه موجود لزوماً عملکرد را افزایش نمی‌دهد. نقش کلیدی مدل *ResNet* ایجاد یک لایه یا بلوک است که بتواند به راحتی تغییر هویت را یاد بگیرد. چیدمان این بلوک‌ها در بالای شبکه موجود در حالت ایده‌آل نباید عملکرد آن را بدتر کند، زیرا بلوک‌های جدید تغییر هویت را یاد می‌گیرند و در نتیجه عملکرد مدل پایه را حفظ می‌کنند. اما بلوک‌های جدید اضافه شده تعداد درجات آزادی و عمق شبکه را افزایش می‌دهند، بنابراین ظرفیت یادگیری ویژگی‌های پیچیده را افزایش می‌دهند. در نتیجه اگر نمایش بهتری از داده‌ها وجود داشته باشد، شبکه جدید باید بتواند آن را یاد بگیرد.

۲-۴- برآورد دقت یادگیری

مدل در هر دوره یادگیری با استفاده از توابع دقت و زیان اعتبارسنجی می‌شود. تابع دقت نشان‌دهنده دقت مدل در طبقه‌بندی تصاویر اعتبارسنجی است، در حالی

[Downloaded from jgit.kntu.ac.ir on 2026-02-20]
[DOI: 10.61186/jgit.11.3.85]

که تابع زیان نشان‌دهنده عدم دقت پیش‌بینی توسط مدل است. اگر یادگیری مدل موفقیت‌آمیز باشد، مقدار تابع زیان، کم و مقدار تابع دقت، زیاد است. با این حال، اگر مقدار تابع زیان در طول یادگیری زیاد شود، نشان دهنده بیش برازش است. معیار دقت[1] و معیار بازیابی یا حساسیت[2] رایج‌ترین شاخص‌های ارزیابی برای تشخیص اهداف در یادگیری عمیق هستند. هر چه مقدار این معیارها بالاتر باشد، توانایی پیش‌بینی قوی‌تر است. فرمول‌های محاسبه دقت تشخیص و نرخ بازیابی به ترتیب در رابطه‌های (۶) و (۷) نشان داده شده است:

$$P = \frac{TP}{TP + FP}$$ رابطه(۶)

$$R = \frac{TP}{TP + FN}$$ رابطه(۷)

در رابطه‌های (۶) و (۷)، مثبت واقعی (TP)[3]، منفی کاذب (FN)[4] و مثبت کاذب (FP)[5] بر اساس ماتریس ابهام محاسبه می‌شوند که تعداد تشخیص‌های مختلف را نشان‌می‌دهند. TP تعداد اهدافی را نشان می‌دهد که به درستی شناسایی شده‌اند و FP تعداد اهدافی را نشان می‌دهد که هدف نبوده‌اند، اما به اشتباه به عنوان هدف شناسایی شده‌اند. FN تعداد اهدافی را نشان می‌دهد که هدف هستند، اما شناسایی نشده‌اند [۲۳].

معیار *F1- Score* یک معادل بین این دو پارامتر ارزیابی مدل است و به عنوان میانگین هارمونیک آن‌ها تعریف می‌شود (رابطه (۸)):

$$F1_{Score} = 2 \times \frac{P \times R}{P + R}$$ رابطه(۸)

۳- یافته‌ها

هدف این پژوهش کشف دهانه میله چاه‌های قنات از تصاویر ماهواره‌ای دانلود شده از سامانه گوگل ارث بر مبنای مدل‌های یادگیری عمیق می‌باشد. شکل (۶) ساختار مدل پیشنهادی بر مبنای شبکه عصبی کانولوشن را نشان می‌دهد.

در این تحقیق، ابتدا داده‌های‌آموزشی با استفاده از روش تصویر قطعه قطعه شده آماده شده است[۲۴]. در این روش، ابتدا تصاویر ماهواره‌ای گوگل ارث برای منطقه مورد مطالعه در شهرستان بم، با ابعاد ۲۵۶ در ۲۵۶ پیکسل برای دو کلاس قنات بر اساس شکل، رنگ، بافت و ساختار دهانه میله چاه قنات در نرم‌افزار *ArcGIS Pro* بریده شده است و در مرحله بعد، این تصاویر به مربع‌های با ۵۰٪ همپوشانی هم به صورت عمودی و هم به صورت افقی بریده شده است. در نهایت تصاویر قطعه قطعه شده به عنوان تصاویر آموزشی استفاده شده است. در منطقه مورد مطالعه شکل برخی از میله چاه‌های قنوات تخریب شده است و امکان تشخیص و استخراج توسط عامل انسانی نیز با مشکل مواجه است، لذا در این مطالعه برای آموزش از میله چاه‌های قنواتی استفاده شده است که دارای شکل مناسب هستند. قنوات با شکل مناسب در منطقه مورد مطالعه نمونه‌های آموزشی در دو کلاس (دریچه باز و دریچه بسته) طبقه‌بندی شده‌اند. تعداد میله چاه‌های قنات در منطقه مورد مطالعه به وسعت تقریبی ۳۰ کیلومترمربع برابر ۲۴۹۶ میله چاه بوده است که ۵۵۷ میله چاه آن دارای شکل تخریب شده و ۱۹۳۹ میله چاه دارای شکل مناسب در دو کلاس (دریچه باز و دریچه بسته) بوده است. برای هر کلاس حداقل ۳۰۰ نمونه آموزشی و در مجموع بیش از ۶۰۰ نمونه آموزشی دریچه میله قنات انتخاب شده است.

شکل (۷) تعدادی از نمونه‌های آموزشی استخراج شده از تصویر اصلی همراه با دو نمونه آموزشی انتخاب شده برای دو کلاس را نشان می‌دهد. به منظور نشان دادن عملکرد مدل مورد استفاده، هیچ پیش‌پردازشی بر روی تصاویر انجام نشده است.

---

[1] *Precision*

[2] *Recall*

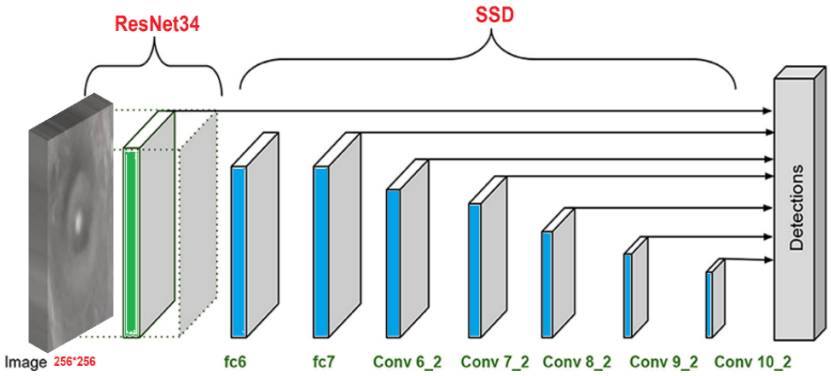
[3] *True Positive*

[4] *False Negative*

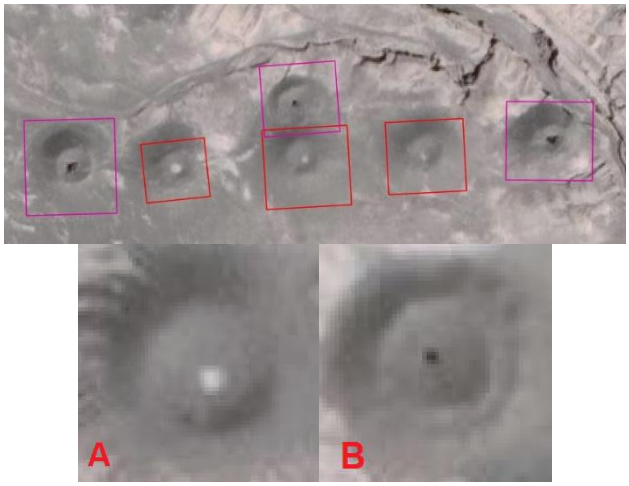
[5] *False Positive*

[Downloaded from jgit.kntu.ac.ir on 2026-02-20]

[DOI: 10.61186/jgit.11.3.85]



شکل ۶: ساختار مدل پژوهش بر مبنای آشکارساز چندجعبه تک شات برای کشف قنات



شکل ۷: نمونه‌های تصاویر آموزشی میله چاه‌های قنوات در منطقه مورد مطالعه در دو کلاس پیشنهادی

قنات‌ها را می‌توان تا حد زیادی به صورت بصری تشخیص و مکان هندسی آن‌ها را استخراج نمود، زیرا دهانه میله چاه‌ها در ترازهای بلند و خطی دیده می‌شوند. در منطقه مورد مطالعه دهانه میله چاه‌ها در تصاویر گوگل ارث قابل تشخیص است. ایجاد مجموعه داده‌های آموزشی بزرگ و با کیفیت بالا یکی از بزرگترین چالش‌ها در مسیر استفاده از یادگیری عمیق است. به منظور افزایش نرخ فراخوان، اشیاء با سطح اطمینان پایین‌تر باید برچسب گذاری شوند. عدم قطعیت، سوگیری و خطا در طول فرآیند برچسب‌گذاری بر عملکرد مدل تأثیر می‌گذارد. بدون داده‌های برچسب‌گذاری شده واقعی، قضاوت در مورد اینکه اشتباهات در داده‌های آموزشی ممکن است در عملکرد ضعیف مدل نقش داشته باشد، دشوار است. در این تحقیق سعی شده است تعداد نمونه آموزشی بهینه برای دستیابی به دقت بهتر در مدل پیشنهادی انتخاب شود.

در این پژوهش از شبکه عصبی کانولوشن بر مبنای آشکارساز چندجعبه تک شات با شبکه *ResNet* برای کشف و استخراج خودکار مکان هندسی میله چاه‌های

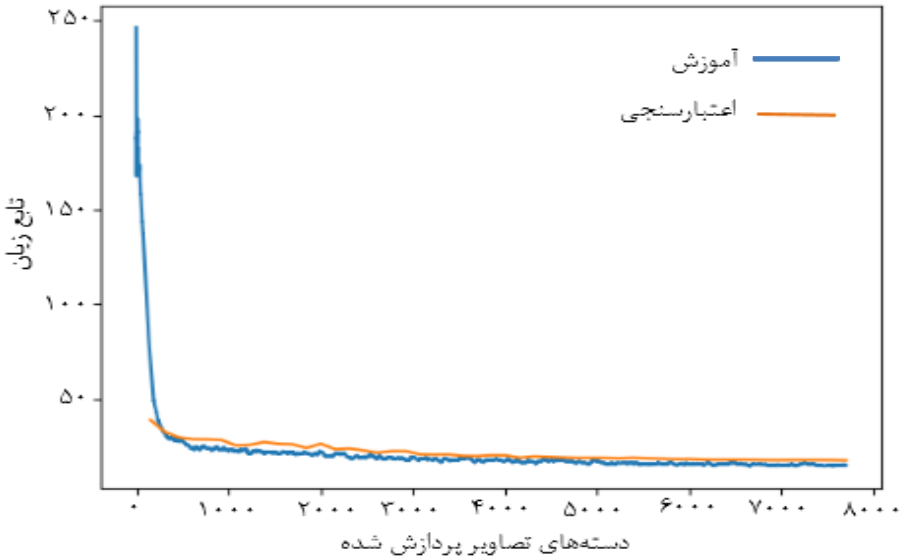
[Downloaded from jgit.kntu.ac.ir on 2026-02-20]
[DOI: 10.61186/jgit.11.3.85]

قنوات استفاده شده است. برخلاف رویکردهای مرسوم برای آموزش طبقه‌بندی‌کننده‌ها با استخراج ویژگی‌های طراحی‌شده دستی، مدل پیشنهادی سلسله‌مراتب ویژگی‌ها را از پیکسل‌ها به طبقه‌بندی‌کننده‌ها می‌آموزد و لایه‌ها را به طور مشترک آموزش می‌دهد. در این پژوهش از لایه نهایی مدل آشکارساز چندجعبه تک شات برای شناسایی قنات از تصاویر گوگل ارث استفاده شده است. به طور تصادفی همه تصاویر با هم مخلوط شده تا از همپوشانی داده‌های آموزشی و داده‌های اعتبارسنجی جلوگیری شود، سپس ۸۵ درصد تصاویر به‌دست‌آمده به‌عنوان داده آموزشی و ۱۵ درصد باقی‌مانده به‌عنوان داده اعتبارسنجی استفاده گردیده است.

برای آموزش از شبکه *ResNet* با ۳۴ لایه (*ResNet 34*) استفاده شده است که یک مدل یادگیری عمیق کلاسیک است، به این علت که در این مطالعه از تصاویر با اندازه کوچک به عنوان داده‌های آموزشی و اعتبارسنجی استفاده می‌کند. این شبکه با دو لایه کانولوشن، دو لایه ادغام و یک لایه کاملا متصل ساخته شده است. *ResNet34*، منجر به تولید ۲۵۶ نقشه ویژگی با ابعاد ۷*۷ برای یک تصویر ورودی می‌شود.

آشکارساز چندجعبه تک شات تنها یک یا چند لایه کانولوشن است که به *ResNet34* اضافه شده است و خروجی‌ها به عنوان جعبه‌های مرزی و کلاس‌های اشیاء در موقعیت مکانی فعال‌سازی لایه‌های نهایی تفسیر می شوند. مدل مورد نظر با برنامه‌نویسی به زبان پایتون و استفاده از کتابخانه‌های یادگیری عمیق و نرم‌افزار *ArcGIS Pro* پیاده‌سازی شده است.

پارامترهای مدل اجرا شده در این مطالعه شامل تعداد دوره‌های آموزشی برابر ۵۰ (*epoch=50*) و نرخ یادگیری به صورت کاهشی استفاده شده است. آموزش مدل پیشنهادی بر روی داده‌های آموزشی با یک کامپیوتر شخصی لپ‌تاپ با پردازشگر مرکزی نسل هشت *Core i7* و *RAM* برابر ۱۲ گیگابایت، حدود ۵ ساعت و ۳۶ دقیقه طول کشیده است. میانگین دقت داده‌های اعتبارسنجی لایه نهایی مدل برابر ۸۹ درصد بوده است. به دنبال افزایش تعداد دوره‌های آموزشی، مقادیر دقت افزایش و مقادیر تابع زیان کاهش یافته است. این نتایج نشان می‌دهد که همه مدل‌ها بیش از حد به مجموعه داده‌ها برازش نشده‌اند و با موفقیت ویژگی‌های تصاویر قطعه شده را یاد گرفته‌اند (شکل (۸)).



شکل ۸: تابع زیان بر اساس دوره‌های یادگیری (دسته‌های تصاویر پردازش شده)

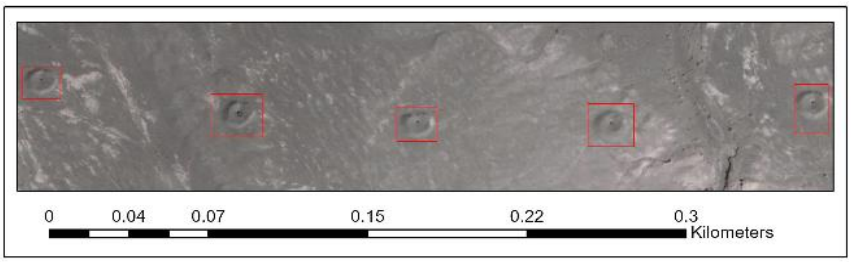
[Downloaded from jgit.kntu.ac.ir on 2026-02-20]
[DOI: 10.61186/jgit.11.3.85]

پس از آموزش مدل، مدل آموزش دیده پیشنهادی بر روی تصویر ماهواره‌ای گوگل ارث دانلود شده به مساحت ۳۰ کیلومتر مربع برای کشف میله چاه‌های قنوات اجرا شده است. برای هر میله چاه قنات کشف شده توسط مدل، جعبه مستطیلی برای نمایش مکان قنوات بدست آمد. هر جعبه مرزی دارای امتیاز اطمینان از ۰ تا ۱ است که درجه اطمینان برای حضور یک میله چاه قنات را نشان می‌دهد. پس از آن، برای حذف کادرهای مرزی اضافی که با همدیگر همپوشانی بالایی داشتند، با نگه‌داشتن یکی با بالاترین امتیاز اطمینان استفاده شد. در نهایت، تمام جعبه‌های مرزی با امتیاز اطمینان کمتر از ۰/۲ (۲۰ درصد) حذف شدند، دامنه اطمینان قنوات کشف شده بین۲۰/۰۳ تا ۹۹/۸۷ درصد می‌باشد. بنابراین بهترین گزینه‌های هدف برای اعتبارسنجی الگوریتم تشخیص حفظ شدند.

با توجه به شکل و تراکم دهانه میله چاه‌های قنات عملکرد مدل پیشنهادی مورد بررسی قرار گرفت که در ادامه به آن‌ها اشاره شده است:

۱) مناطق با تراکم کم و شکل مناسب دهانه میله چاه‌های قنات

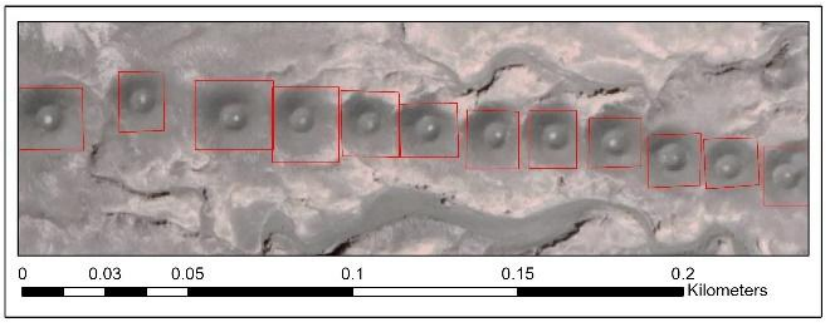
مدل پیشنهادی در مناطقی که شکل دهانه میله چاه‌ها تغییر نکرده و فاصله بین آن‌ها زیاد است، دارای دقت بالایی بوده است و تمام میله چاه‌های قنات را تشخیص می‌دهد. شکل (۹) یک نمونه از نتیجه عمکرد مدل بر روی مناطق با تراکم کم و شکل مناسب میله چاه‌ها را نشان می‌دهد.



شکل ۹: نتیجه مدل اعمال شده در مناطق با شکل مناسب و تراکم کم میله چاه‌ها

۲) مناطق با تراکم بالا و شکل مناسب دهانه میله چاه‌های قنات

مدل پیشنهادی در مناطقی که شکل دهانه میله چاه‌ها تغییر نکرده و فاصله بین آنها کم است، دارای دقت بالایی بوده و بیش از ۹۰ درصد میله چاه‌های قنات را تشخیص می‌دهد، در این مورد میزان همپوشانی بین جعبه‌های شناسایی مجاز است. شکل (۱۰) یک نمونه از نتیجه عمکرد مدل بر روی مناطق با تراکم بالا و شکل مناسب میله چاه‌ها را نشان می‌دهد.

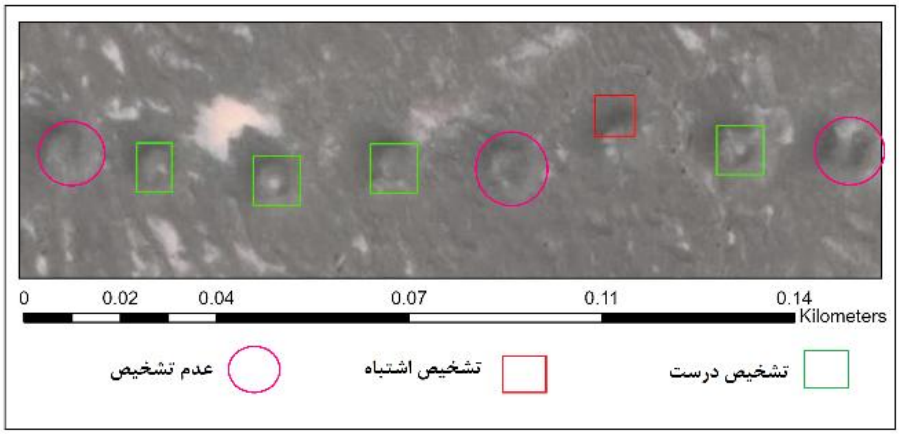


شکل ۱۰: نتیجه مدل اعمال شده در مناطق با شکل مناسب و تراکم بالای میله چاه‌ها

[Downloaded from jgit.kntu.ac.ir on 2026-02-20]
[DOI: 10.61186/jgit.11.3.85]

۳) مناطق با تراکم کم و شکل تخریب شده دهانه میله چاه‌های قنات

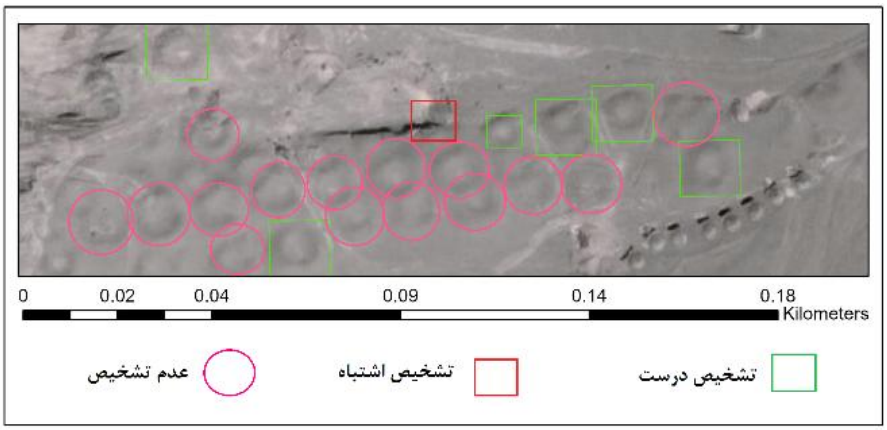
مدل پیشنهادی در مناطقی که شکل دهانه میله چاه‌ها در گذر زمان تخریب شده و بطور کلی شکل دهانه میله چاه‌ها به سختی قابل شناسایی است و همچنین فاصله بین آن‌ها زیاد و در نتیجه تراکم کم است، دارای دقت پایینی می‌باشد. شکل (۱۱) یک نمونه از نتیجه عمکرد مدل بر روی مناطق با تراکم کم و شکل تخریب شده میله چاه‌ها را نشان می‌دهد.



شکل ۱۱: نتیجه مدل اعمال شده در مناطق با شکل تخریب شده و تراکم کم میله چاه‌ها

۴) مناطق با تراکم زیاد و شکل تخریب شده دهانه میله چاه‌های قنات

مدل پیشنهادی در مناطقی که شکل دهانه میله چاه‌ها تخریب شده و فاصله بین آن‌ها کم و در نتیجه تراکم زیاد است، دارای دقت خیلی پایینی می‌باشد. در این مورد میزان همپوشانی بین جعبه‌های شناسایی مجاز بوده است. شکل (۱۲) یک نمونه از نتیجه عمکرد مدل بر روی مناطق با تراکم زیاد و شکل تخریب شده میله چاه‌ها را نشان می‌دهد. در این حالت حتی تشخیص و استخراج دهانه میله چاه‌ها برای عامل انسانی نیز چالشی بوده و دارای خطای انسانی است.



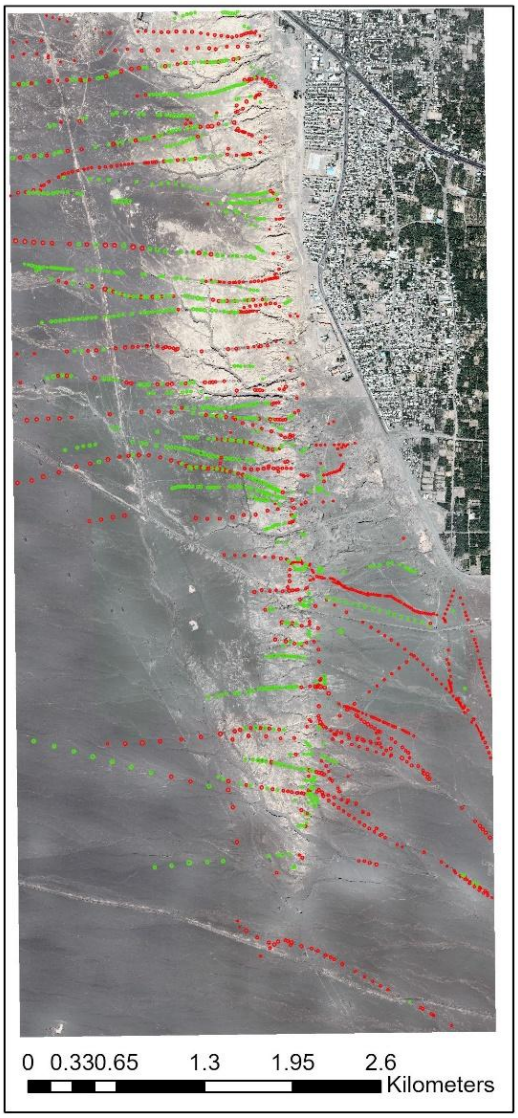
شکل ۱۲: نتیجه مدل اعمال شده در مناطق با شکل تخریب شده و تراکم بالای میله چاه‌ها

[Downloaded from jgit.kntu.ac.ir on 2026-02-20]
[DOI: 10.61186/jgit.11.3.85]

در این پـژوهش هـدف تشـخیص و اسـتخراج موقعیـت مکانی دهانه میله چاه‌هایی بوده است کـه دارای شـکل مناسب بوده‌اند، بـدیهی اسـت در صـورت اعمـال دقـت برای چاه‌هـای قنـات تخریـب شـده، دقـت تشـخیص و استخراج کاهش خواهد یافت.

شکل(۱۳) نتیجه نهـایی اعمـال مـدل بـر روی منطقـه مورد مطالعه به وسعت ۳۰ کیلومتر مربـع در تشـخیص قنوات با شکل مناسب را نشان می‌دهد.

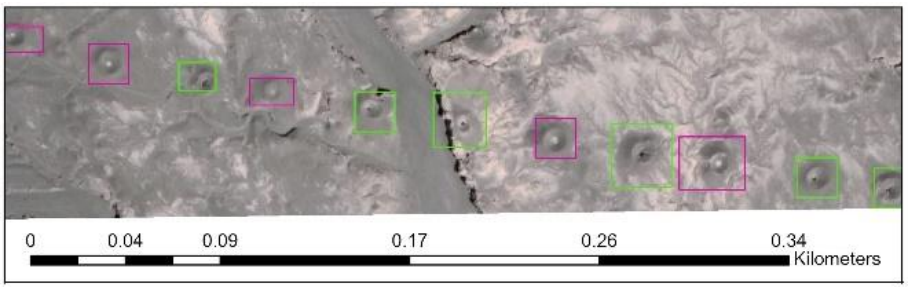
شکل(۱۴) بخشی از محدوده تصـویر مـاهواره‌ای گوگـل ارث با وضوح بالا، حاوی میله چاه‌های قنات کشف شده توسط مدل را برای منـاطق دارای میلـه چـاه بـا شـکل مناسب نشان می‌دهد.



شکل ۱۳: نتیجه نهایی مدل اعمال شده برای کشف میله چاه‌های قنوات در منطقه مورد مطالعه

[Downloaded from jgit.kntu.ac.ir on 2026-02-20]

[DOI: 10.61186/jgit.11.3.85]



شکل ۱۴: نمونه میله چاه‌های قنوات کشف شده توسط مدل در منطقه مورد مطالعه

جدول (۱) نتیجه ارزیابی دقت مدل استفاده شده را برای قنات‌ها با شکل مناسب نشان می‌دهد. بدیهی است در مناطقی که شکل میله چاه‌های قنات تخریب شده و حتی توسط عامل انسانی استخراج و تشخیص آن مشکل است، دقت مدل پیشنهادی خیلی کمتر خواهد بود (جدول(۲)).

جدول (۲) نتیجه ارزیابی دقت مدل استفاده شده را برای کلیه قنات‌های تخریب شده و با شکل مناسب نشان می‌دهد.

جدول ۱: معیارهای ارزیابی دقت مدل برای میله چاه‌های قنوات با شکل مناسب

معیار *F1 Score*	معیار بازیابی (*Recall*)	معیار دقت (*Precision*)
۰٫۸۶	۰٫۸۲	۰٫۹۱

جدول ۲: معیارهای ارزیابی دقت مدل برای تمام میله چاه‌های قنوات

معیار *F1 Score*	معیار بازیابی (*Recall*)	معیار دقت (*Precision*)
۰٫۶۲	۰٫۶۰	۰٫۶۵

۴- نتیجه‌گیری

قنات به عنوان یک سازه تامین آب پایدار در مناطق کم آب یا خشک شناخته شده است. مدیریت این سازه‌های آبی با توجه به شرایط آب و هوایی ضروری است. اولین گام در این راستا تهیه نقشه قنوات موجود است. روش‌های زمینی تهیه نقشه قنوات با توجه به شرایط آب و هوایی و نیاز به صرف هزینه و زمان بالا، هر چند دارای دقت بالایی می‌باشند، مقرون به صرفه نیستند. لذا راه‌کار بعدی استفاده از تصاویر هوایی و تصاویر ماهواره‌ای با قدرت تفکیک بالا است. علاوه بر هزینه تهیه تصاویر هوایی و همچنین نیاز به پردازش و عملیات مثلث‌بندی باید اپراتور تبدیل عوارض مورد نظر را بصورت بصری در فضای سه بعدی یا دوبعدی بر روی تصاویر ارتوفتو، تشخیص و استخراج نماید که همین مرحله تبدیل (تشخیص و استخراج عوارض) زمانبر است و در نتیجه هزینه تهیه نقشه زیاد خواهد شد. در این راستا استفاده از تصاویر ماهواره‌ای با قدرت تفکیک مکانی بالا برای تشخیص و استخراج نقشه قنوات پیشنهاد می‌شود. با توجه به هزینه نسبتا بالای اخذ تصاویر ماهواره‌ای می‌توان از تصاویر ماهواره‌ای در دسترس در سامانه گوگل ارث استفاده نمود. در صورت استخراج بصری عوارض از تصاویر ماهواره‌ای نیاز به اپراتور تبدیل و تفسیر تصاویر وجود دارد و ممکن است اپراتورهای مختلف تشخیص و حتی نحوه استخراج متفاوتی داشته باشند. در این راستا

[Downloaded from jgit.kntu.ac.ir on 2026-02-20]
[DOI: 10.61186/jgit.11.3.85]

در این پژوهش برای تشخیص و استخراج خودکار دهانه چاه‌های قنات از تصاویر ماهواره‌ای با قدرت تفکیک بالای در دسترس گوگل ارث استفاده شده است. برای تشخیص اهداف از تصاویر بهترین روش در عصر حاضر روش‌های یادگیری عمیق می‌باشد. مزیت استفاده از یادگیری عمیق این است که می‌تواند ویژگی‌ها را به تنهایی یاد بگیرد که در مقایسه با الگوریتم‌های کلاسیک یادگیری ماشین باعث صرفه جویی در زمان می‌شود. یکی دیگر از ویژگی‌های یادگیری عمیق دقت بالاتر آن است، به این معنی است که آموزش بیشتر یا ورودی داده بیشتر اغلب منجر به دقت بالاتر نسبت به مدل‌های قبلی می‌شود. اما، مدل‌های یادگیری عمیق به زمان بیشتری برای آموزش نیاز دارند، ولی زمان کوتاه‌تری برای استنباط در مقایسه با الگوریتم‌های یادگیری ماشین معمولی نیاز دارند. برای دقت بهتر، یادگیری عمیق به مقادیر زیادی مجموعه داده به عنوان ورودی در طول آموزش نیاز دارد. یادگیری عمیق و شبکه‌های عصبی کانولوشنال اخیراً به راه‌حل جدیدی برای تشخیص اهداف و طبقه‌بندی تصاویر تبدیل شده‌اند. در عمل، دو نوع الگوریتم اصلی تشخیص اشیا وجود دارد. نوع اول الگوریتم‌هایی مانند *R-CNN* و *R-CNN Fast(er)* هستند که از یک رویکرد دو مرحله‌ای استفاده می‌کنند. یعنی ابتدا برای شناسایی مناطقی که انتظار می‌رود اشیاء در آنجا کشف شوند و سپس اشیاء را فقط در آن مناطق با استفاده از شبکه کانولوشنی شناسایی می‌کنند. نوع دوم، الگوریتم‌هایی مانند *YOLO* (شما فقط یک بار نگاه می‌کنید) [۲۵] و آشکارساز چندجعبه تک شات از یک رویکرد کاملاً کانولوشنال استفاده می‌کنند که در آن شبکه می‌تواند تمام اشیاء درون یک تصویر را در یک نگاه پیدا کند. الگوریتم‌های اول معمولاً دقت بهتری دارند، اما اجرای آن‌ها کندتر است، در حالی که الگوریتم‌های تک‌مرحله‌ای کارآمدتر هستند و دقت قابل قبولی هم دارند و به همین دلیل در این پژوهش از این الگوریتم استفاده شده است. آشکارساز چندجعبه تک شات از نسبت ابعاد پیش‌فرض استفاده می‌کند و می‌تواند برای آشکارسازی شیئ استفاده شود [۲۶].

در این پژوهش ابتدا نمونه‌های آموزشی دهانه میله چاه‌های قنات از تصاویر ماهواره‌ای دانلود شده از سامانه گوگل ارث استخراج شده است. بیش از ۶۰۰ نمونه آموزشی در تصاویر با ابعاد ۲۵۶ در ۲۵۶ پیکسل برای آموزش مدل تهیه شده است. سعی گردیده مدل طوری انتخاب گردد که بتوان آن‌را با نمونه آموزشی کمتر و در زمان کمتر آموزش داد. مدل پیشنهادی در این پژوهش یک مدل شبکه عصبی کانولوشن بر مبنای آشکارساز چندجعبه تک شات بر پایه شبکه *ResNet34* می‌باشد. مدل با برنامه‌نویسی پایتون و کتابخانه‌های آماده یادگیری عمیق پیاده‌سازی شده است. پس از آموزش مدل، مدل آموزش دیده بر روی یک تصویر ماهواره‌ای گوگل ارث ژئورفرنس شده در محدوده منطقه مورد مطالعه اجرا شده است و موقعیت مکانی میله چاه‌های قنات در سیستم تصویر *UTM* استخراج شده است. نتایج ارزیابی مدل بروی میله چاه‌های قنات با شکل مناسب بر اساس معیار دقت برابر ۰/۹۱ و معیار بازیابی برابر ۰/۸۲ و معیار *F1Score* برابر ۰/۸۶ می‌باشد. بطور کلی نتایج نشان می‌دهد که مدل یادگیری عمیق پیشنهادی می‌تواند در تشخیص و استخراج میله چاه‌های قنات از تصاویر ماهواره‌ای موجود در سامانه گوگل ارث استفاده شود.

مدل در مناطقی که شکل دهانه میله چاه بهتر حفظ شده و تخریب نشده باشد، بهترین عملکرد را داشته است. بیشترین میزان مثبت کاذب به مثبت واقعی در مناطقی بود که میله چاه قنات در گذر زمان تخریب شده است یا قطعیت شناسایی پایین است. همچنین در مناطقی که فاصله بین میله چاه‌ها کمتر از ۱۰ متر بوده است، بدلیل قدرت تفکیک مکانی تصویر و نزدیکی نسبی میله چاه‌های قنات، مدل نتایج ضعیف‌تری داشته است. نتایج نشان می‌دهد که امکان تشخیص و استخراج اتوماتیک میله چاه‌های قنات با دقت مناسب از تصاویر ماهواره‌ای دانلود شده از سامانه گوگل ارث

[Downloaded from jgit.kntu.ac.ir on 2026-02-20]
[DOI: 10.61186/jgit.11.3.85]

امکان‌پذیر است که با نتایج پژوهش لی و همکاران مطابقت دارد. لی و همکاران در حوضه تورپان در منطقه خودمختار اویغور سین کیانگ با استفاده از مدل یادگیری عمیق *YOLOv5* نقشه قنوات را با دقت کلی ۸۶ درصد از تصاویر ماهواره‌ای گوگل ارث تهیه نمودند [۱۷]. در صورت استفاده از تصاویر هوایی و همچنین افزایش تعداد نمونه‌های آموزشی دقت مدل افزایش خواهد داشت. همچنین می‌توان با ادغام تصاویر چندطیفی کارکرد مدل را بهبود بخشید. در صورت افزایش تعداد نمونه‌های آموزشی با شرایط متفاوت می‌توان به نتایج بهتری دست یافت. برای تحقیقات آتی پیشنهاد می‌گردد که از سایر مدل‌های یادگیری عمیق و همچنین تصاویر پهپاد فتوگرامتری استفاده گردد.

مراجع

[1] M. Khaneiki, "Territorial Water Cooperation in the Central Plateau of Iran", Springer Nature, Switzerland, 2019.

[2] A. Yazdi and M. Khaneiki, "Qanat Knowledge: construction and maintenance", Springer, Dordrecht, 2017.

[3] H. Drucker, C. J. Burges, L. Kauffman, A. Smola and V. Vapnik, "Support vector regression machines. Neural information processing systems", Eds Mozer MC, Jordan JI & Petsche T. pp. 155–161, 1997.

[4] S. Watanabe, K. Sumi and T. Ise, "Identifying the vegetation type in Google Earth images using a convolutional neural network: a case study for Japanese bamboo forests", BMC Ecol 20, 65, 2020.

[5] Y. Bengio, "Learning deep architectures for AI", Foundat Trends Mach Learn, 2:1–127, 2009.

[6] I. Goodfellow, Y. Bengio and A. Courville, "Deep Learning", MIT press, 2016.

[7] Y. LeCun, Y. Bengio and G. Hinton, "Deep learning", Nature, vol. 521, pp. 436_444, 2015.

[8] A. Voulodimos, N. Doulamis, A. Doulamis and E. Protopapadakis, "Deep learning for computer vision: A brief review, Comput. Intell", Neurosci., vol. 2018, pp. 1_13, 2018.

[9] L. Deng, G. Hinton and B. Kingsbury, "New types of deep neural network learning for speech recognition and related applications: An overview", in Proc. IEEE Int. Conf. Acoust., Speech Signal Process., pp. 8599_8603, 2013.

[10] H. Palangi, L. Deng, Y. Shen, J. Gao, X. He, J. Chen, X. Song and R. Ward, "Deep sentence embedding using long short-term memory networks: Analysis and application to information retrieval", IEEE/ACM Trans, Audio, Speech, Language Process., vol. 24, no. 4, pp. 694_707, Apr, 2016.

[11] A. Karpathy, G. Toderici, S. Shetty, T. Leung, R. Sukthankar and L. Fei-Fei, "Largescale video classification with convolutional neural networks", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (Columbus). 1725–1732, 2014.

[12] S. Alberto, M. Savardi, A. Baronio and S. Benini, "Deep learning meets hyperspectral image analysis: A multidisciplinary review", Journal of Imaging 5, no. 5: 52, 2019.

[13] K. Nogueira, O.A.B. Penatti and J.A. Dos Santos, "Towards better exploiting convolutional neural networks for remote sensing scene classification", Pattern Recogn. 61:539–56, 2017.

[14] X.X. Zhu, D. Tuia, L. Mou, G.S. Xia, L. Zhang and F. Xu, "Deep Learning in remote sensing: a comprehensive review and list of resources", IEEE Geosci Remote Sens Mag. 5:8–36, 2017.

[15] L. Ma, Y. Liu, X. Zhang, Y. Ye, G. Yin and B. A. Johnson, "Deep learning in remote sensing applications: A meta-analysis and review", ISPRS J. Photogram. Remote Sens. 152, 166–177, 2019.

[Downloaded from jgit.kntu.ac.ir on 2026-02-20]

[DOI: 10.61186/jgit.11.3.85]

[16] S. Mehrnoush, A. Mehrtash, E. Khazraee and A. Jason, "Deep learning in archaeological remote sensing: Automated qanat detection in the kurdistan region of Iraq", Remote Sensing 12, no. 3: 500, 2020.

[17] Q. Li, H. Guo, L. Luo and X. Wang, "Automatic Mapping of Karez in Turpan Basin Based on Google Earth Images and the YOLOv5 Model", Remote Sens, 14, 3318, 2022.

[18] D. Erhan, C. Szegedy, S. Reed, C.Y. Fu and A.C. Berg, "SSD: Single shot multibox detector", In Proceedings of the Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer: Cham, Switzerland, Volume 9905, pp. 21–37, 2016.

[19] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition", in CVPR, 2016.

[20] B. Xu, N. Wang, T. Chen and M. Li, "Empirical evaluation of rectified activations in convolutional network", arXiv Prepr. arXiv:00853/1505, 2015.

[21] W. Liu, D. Anguelov, D. Erhan and et al., "SSD: single shot multibox detector", In: European conference on computer vision. Cham: Springer, p. 21–37, 2016.

[22] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition", In: arXiv e-prints, 03385/1512, 2015.

[23] T.Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Doll´ar and C.L. Zitnick, "Microsoft coco: Common objects in context", In European Conference on Computer Vision; Springer: Berlin/Heidelberg, Germany, pp. 740–755, 2014.

[24] T. Ise, M. Minagawa and M. Onishi, "Classifying 3 moss species by deep learning, using the "chopped picture" method", Open J Ecol. 8:166–73, 2018.

[25] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection", arXiv:02640/1506, 2015.

[26] J. Jeong, H. Park and N. Kwak, "Enhancement of SSD by concatenating feature maps for object detection", In: BMVC'17, 2017.

[Downloaded from jgit.kntu.ac.ir on 2026-02-20]

[DOI: 10.61186/jgit.11.3.85]

K. N. TOOSI UNIVERSITY OF TECHNOLOGY
FACULTY OF GEODESY AND GEOMATICS ENGINEERING

Journal of Geospatial Information Technology
Vol.11, No.3, Autumn 2023

Research Paper

Application of convolutional neural networks single-shot multibox detector (SSD) in automatic detection and extraction of Qanat from Google Earth satellite images

Mostafa Kabolizadeh [1], Mohammad Abbasi [2]*

1- Assistant Professor, Department of Remote Sensing and Geographical Information System, Faculty of Earth Sciences, Shahid Chamran University of Ahvaz.
2- Phd. Student of Remote Sensing and Geographical Information System, Faculty of Planning and Environmental Sciences, University of Shahid Chamran Ahvaz.

Abstract

Identifying qanat and mapping them is one of the vital tasks in the science of groundwater resources management. However, it is difficult to map qanats because conventional methods such as field mapping are very costly and time consuming, and in some cases, due to the structure of qanat, land extraction faces many challenges. Identifying objects from image data using computer image processing techniques and neural networks is one of the most promising techniques for identifying qanats. Due to the fact that the satellite images of Google Earth system are the only satellite images with high spatial resolution, free of charge and available, the Google Earth satellite images have been used in this research. In this research, more than 600 educational samples have been prepared from the openings of qanats. In this research, due to its higher processing speed, the convolutional neural network single-shot multi-box detector based on ResNet network has been used to automatically detect and extract the geometric location of qanats. The proposed model has been taught by training samples based on 85% of the training data and 15% of the validation data, with 50 repetition courses and an accuracy of more than 90%. The trained model is implemented on the image of the study area in order to discover the shafts of qanats. The results show that this model can work with the accuracy criterion equal to 0.91, the recovery criterion equal to 0.82, and the F1Score criterion equal to 0.86 in discovering the location of the rod of the aqueduct wells with a suitable shape. For the areas where the shape of the rod of the wells is destroyed, the accuracy of detecting and extracting the position of the wells is equal to 0.65. This research shows that it is possible to automatically identify and extract aqueduct well rods with appropriate accuracy from satellite images downloaded from the Google Earth system.

Key words: *Convolution Neural Network, Single Shot Multi-box Detector, Qanat, ResNet, Google Earth.*

Correspondence Address *: Department of Remote Sensing and Geographical Information System, Faculty of Earth Sciences, Shahid Chamran University of Ahvaz, Iran.*
Tel*: +98 916 3470177.*
Email*: m.kabolizade@scu.ac.ir*

[Downloaded from jgit.kntu.ac.ir on 2026-02-20]

[DOI: 10.61186/jgit.11.3.85]