

الگوریتم Mask R-CNN برای استخراج مرز زمین‌های کشاورزی از تصاویر ماهواره‌ای شهرستان کنگاور استان کرمانشاه

ندا رحیم‌پور انارکی^۱، علیرضا آزادبخت^۱، امیر معینی‌راد^۲، حمید صالحی شهرابی^۲،

امید ابراهیمی^۲، آرش زندکریمی^۲

^۱ دانشکده علوم ریاضی دانشگاه شهید بهشتی.

^۲ پژوهشگاه فضایی ایران

نام نویسنده مسئول:

ندا رحیم‌پور انارکی

تاریخ دریافت: ۱۴۰۲/۰۴/۲۸

تاریخ پذیرش: ۱۴۰۳/۰۷/۰۱

چکیده

سرعت بخشیدن به تهیه مرز کاداستر زمین‌های کشاورزی از اهمیت بالایی برخوردار است. امروزه با ترویج تصاویر ماهواره‌ای و هوایی، بهره‌گیری از روش‌های یادگیری عمیق برای استخراج خودکار مرزها از این دسته تصاویر، از جمله راهکارهایی است که برای حل مسأله به کمک ما می‌آیند. در این مقاله، طی دو عملیات پردازش و پس‌پردازش هندسی، به استخراج خودکار مرزهای کاداستر زمین‌های کشاورزی می‌پردازیم. در عملیات پردازش، با هدف قطعه‌بندی نمونه‌ای تصویر، شبکه Mask R-CNN بعنوان هسته مرکزی مورد استفاده قرار می‌گیرد که از طریق رویکرد یادگیری انتقالی با استفاده از شبکه از پیش‌آموزش‌دیده‌ی ResNet-50 آموزش می‌بیند که از قبل روی مجموعه داده‌ی ImageNet آموزش دیده است. در عملیات پس‌پردازش هندسی، علاوه بر ارائه الگوریتمی جدید برای ساده‌سازی چندضلعی‌ها به فرمت برداری، خروجی خام شبکه را در سه مرحله بهبود می‌بخشیم. کسب مقادیر ۷۴٪، ۹۱٪ و ۸۱٪ به ترتیب برای معیارهای Precision، Recall و F-score بر روی تصویری ماهواره‌ای از مناطق زراعی شهرستان کنگاور در استان کرمانشاه، نشان‌دهنده‌ی نتایج قابل اتکایی در این زمینه هستند.

واژگان کلیدی: نقشه برداری کاداستر، شبکه Mask R-CNN، سنجش از دور، قطعه‌بندی نمونه‌ای

مقدمه

ثبت مالکیت و مکان فیزیکی زمین‌ها که از آن بعنوان کاداستر یاد می‌شود، برای مالکان زمین‌ها امنیت تصدی، معیشت پایدار و افزایش فرصت‌های مالی به همراه دارد [1]. بنابراین یکی از موضوعاتی که در حوزه سنجش از دور از توجه ویژه‌ای برخوردار است، نقشه‌برداری کاداستر است. اهمیت این موضوع زمانی بیشتر می‌شود که در نظر بگیریم در حدود ۷۵ درصد از مالکیت زمین‌ها در هیچ سیستم کاداستر رسمی ثبت نشده است که یک چالش بزرگ برای کشورهای توسعه‌یافته و در حال توسعه محسوب می‌شود [2] و [3]. یک سیستم کاداستر موثر در تامین مالیات بر املاک و مستغلات، برنامه‌ریزی شهری، توسعه و نظارت بر بازار زمین، کنترل ساخت و سازهای غیر مجاز، توسعه زیرساخت‌های شهری و تولید داده‌های آماری نقش به‌سزایی دارد و موجب افزایش امنیت در زمین‌های دارای مالکیت می‌شود [4].

بروز نگه‌داشتن اطلاعات کاداستر زمین‌ها در کشورهایی که با صرف دهه‌ها زمان موفق شده‌اند کاداستر کاملی با روش‌های سنتی نقشه‌برداری تهیه کنند، در اولویت بالایی قرار دارد [5]. حال آن که در کشوری هم‌چون ایران، تهیه نقشه‌های کاداستر خود چالش بزرگی است که بهره‌گیری از روش‌های سنتی زمان‌بر، پرهزینه و مشقت‌بار تلقی می‌شود [6]. امروزه با توجه به دسترس بودن تصاویر و وسایل نقلیه هوایی بدون سرنشین (UAV) در بعضی مناطق و ماهواره‌های رصدکننده زمین در سطح جهانی، سنجش از دور برای نقشه‌برداری کاداستر به جای نقشه‌برداری میدانی مورد استفاده قرار می‌گیرد که مورد حمایت سیستم مدیریت زمین است و باعث سرعت بخشیدن و کاهش هزینه‌ها در عملیات نقشه‌برداری می‌شود [2].

اما تشخیص مرزهای کاداستر از تصاویر سنجش از دور نیز چالش‌های خود را به همراه دارد. زیرا از بین تمام مرزها، تنها مرزهایی قابل شناسایی هستند که منطبق بر مرزهای طبیعی یا دست‌ساز هستند. این‌گونه مرزها را مرز مرئی می‌نامیم. بنابراین نقشه‌برداری کاداستر بر پایه تصاویر، با به رسمیت شناختن این موضوع شکل می‌گیرد که بسیاری از مرزهای کاداستر در تصاویر، مرزهای مرئی هستند. نمونه‌ی چنین مرزهایی جاده‌ها، تجمع درخت‌ها، تغییر در الگوی بافت، دیوارهای سنگی، حصارها، دیوارهای ساختمان، جوی‌ها، رودخانه‌ها، زهکشی‌های آب و نوارهای زمین‌های بدون کشت هستند. چنین مرزهایی پتانسیل این را دارند که به صورت خودکار توسط الگوریتم‌های پردازش تصویر استخراج شوند [7]، [8].

راهکار ما برای این مسأله شامل چارچوبی ۳ مرحله‌ای است که در [9] ارائه شده است. در مرحله اول شبکه Mask R-CNN با رویکرد یادگیری انتقالی را به کار می‌گیریم. برای این منظور، شبکه از پیش‌آموزش‌دیده‌ی ResNet-50 که روی مجموعه داده‌ی ImageNet آموزش دیده است، انتخاب شده است. شبکه بعنوان خروجی نقشه احتمالاتی تولید می‌کند که احتمال تعلق هر پیکسل به زمین‌های کشاورزی را مشخص می‌کند. سپس با انتخاب آستانه‌ی تحمل مشخصی، متد باینری کردن اوتسو [10] بر روی خروجی شبکه اعمال می‌شود تا نقشه‌ای بدست آید که پیکسل‌های متعلق به زمین‌های کشاورزی برچسب ۱ و پیکسل‌های پس‌زمینه برچسب ۰ می‌گیرند. در نهایت الگوریتم تشخیص لبه کنی [11] مرزهای این زمین‌ها را استخراج می‌کند. در مرحله دوم که از آن بعنوان پس‌پردازش هندسی یاد می‌کنیم، زمین‌ها با مساحتی کوچک‌تر از آستانه تحمل تعیین شده به همراه چندضلعی‌های شناسایی شده درون زمین‌های بزرگ‌تر حذف می‌شوند و در آخرین گام، مرزها ساده‌سازی می‌شوند. سپس با در نظر گرفتن دو بافر برای مرزهای مرجع، با سه معیار Precision، Recall و F-score به ارزیابی دقت خروجی با نتایج حاصل شده در [9] می‌پردازیم.

در این مقاله، ابتدا به مرور راهکارهای ارائه شده برای پیدا کردن مرز زمین‌ها در دیگر مقالات در فصل پیشینه پژوهش می‌پردازیم. در بخش روش کار، راهکار پیشنهادی خود برای حل مسأله را در دو مرحله‌ی پردازش و پس‌پردازش هندسی با جزئیات مطرح می‌کنیم و سپس دقت مدل را در بخش ارزیابی دقت، بررسی می‌کنیم. در نهایت در بخش آخر، نتیجه‌گیری و کارهای آتی را خواهیم داشت.

پیشینه پژوهش

با توجه به این که در این مسأله به نوعی با تقسیم‌بندی تصویر و یافتن مرز مواجه هستیم، طبیعی است اولین راهکارهایی که برای حل آن مطرح شده باشند بر پایه تقسیم‌بندی و تشخیص لبه باشند. دراگوت و همکاران [12] روشی به نام تقسیم‌بندی

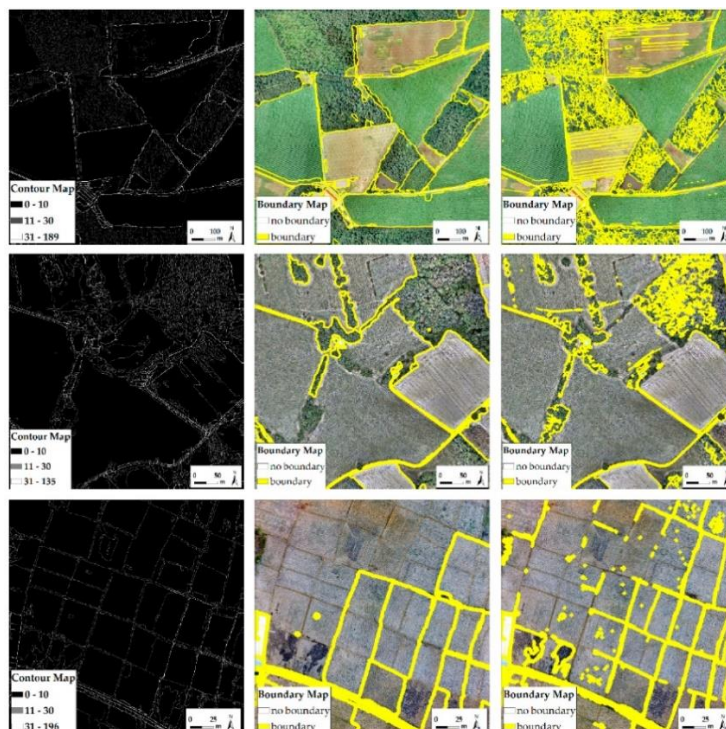
وضوح-چندگانه را به کار بردند که تقسیم‌بندی تصویر مقیاس-چندگانه از لایه‌های متعدد را پارامترسازی می‌کند. این روش متکی بر پتانسیل وارپانس محلی در تشخیص انتقال مقیاس در داده جغرافیایی عمل می‌کند.

در ادامه کروملینک و همکاران [13] تکنیک‌های تشخیص لبه‌ی بینایی ماشین را به نقشه‌برداری کاداستر بر پایه UAV انتقال دادند. آن‌ها نشان دادند مرزهای مرئی می‌توانند با روش‌های بینایی ماشین که کانتورهای اجسام را در تصاویر تشخیص می‌دهند، استخراج شوند. بنابراین انتقال روش تشخیص کانتور gPb که بهترین روش بینایی ماشین در آن زمان بود را به تصاویر UAV سنجش از دور بررسی کردند.

روش تشخیص کانتور gPb توسط آربلاز و همکاران در سال ۲۰۱۱ ارائه شد. این روش ترکیب خطی mPb و sPb است. در یک تعریف کلی می‌توان گفت که mPb تمام لبه‌ها را تشخیص می‌دهد، در حالی که sPb تنها لبه‌های برجسته را از کل تصویر استخراج می‌کند. gPb هر دو این‌ها را ترکیب می‌کند و به طور یکنواخت عملکرد بهتری فراهم می‌کند.

در این مقاله سه تصویر ارتوی UAV از سه منطقه روستایی در کشورهای آلمان، فرانسه و اندونزی انتخاب شده است. علت انتخاب مناطق روستایی این است که نسبت به مناطق شهری با تراکم بالا، معمولاً تعداد مرزهای مرئی بیشتری را شامل می‌شوند. چارچوب [13] برای انتقال روش تشخیص کانتور gPb به نقشه‌برداری کاداستر بر پایه UAV شامل چهار مرحله‌ی زیر می‌شود:

۱. **پیش‌پردازش تصویر:** ابتدا تصویر UAV به دقت پایین‌تر با GSD معادل با ۵ تا ۱۰۰ سانتی‌متر نمونه‌گیری می‌شود.
۲. **ترسیم مرز:** در این مرحله الگوریتم تشخیص کانتور gPb بر روی هر قطعه تصویر با GSD متفاوت پیاده می‌شود که با اعمال فرآیندی به نام gPb-owt-ucm، یک نقشه احتمالاتی کانتور و یک نقشه باینری تولید می‌شود که تعلق هر پیکسل به کلاس "مرز" یا "غیر مرز" را تعیین می‌کند.
۳. **پس‌پردازش تصویر:** تمام قطعات تصویر متعلق به یک مجموعه با هم ادغام می‌شوند تا یک نقشه کانتور و یک نقشه باینری تشکیل دهند و در نتیجه‌ی آن چندضلعی‌هایی برای تمام نواحی متصل بدست می‌آید.
۴. **ارزیابی دقت:** ارزیابی بر پایه پیکسل است و ماتریس در هم‌ریختگی برای هر تصویر در دو مورد تعیین می‌شود: (i) . کیفیت تشخیص: خطاهای Commission و Omission. (ii) کیفیت محلی‌سازی (LQ): دقت مرزها در نگاه هندسی. نمونه‌ای از خروجی‌های تولید شده در این مقاله را می‌توانید در شکل ۱ ببینید.

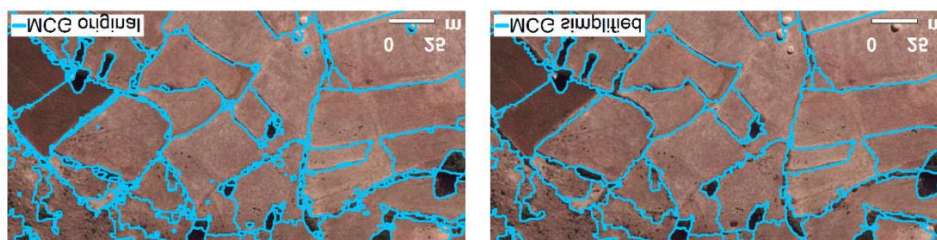


شکل ۱: خروجی‌های تولید شده برای هر یک از سه منطقه تحقیقاتی مقاله [13]. تصاویر سمت چپ نقشه کانتور، تصاویر میانی نقشه مرز با $K=0.1$ بر روی تصویر اصلی و تصاویر سمت راست نقشه مرز با k مشابه اما تقسیم شدن تصویر به تصاویر با وضوح ۱۰۰۰ در ۱۰۰۰ پیکسل است.

در این مسأله از سال ۲۰۱۹ به استفاده از شبکه‌های عصبی روی آوردند. شبکه‌های عصبی پیش‌بینی (CNN) یکی از محبوب‌ترین و موفق‌ترین شبکه‌های عمیق برای مسائل مرتبط با تصاویر هستند. این شبکه‌ها به صورت کارا اشیاء مختلف در تصاویر سنجش از دور را شناسایی می‌کنند.

در سال ۲۰۱۹ کروملینک و همکاران [1] با رویکرد یادگیری انتقالی برای آموزش شبکه و آزمایش شبکه‌های عصبی پیش‌بینی معروفی که تا به آن روز ارائه شده بود، به استخراج مرز زمین‌های کشاورزی از تصاویر UVA پرداختند. تصویری که آن‌ها برای این مقاله انتخاب کردند، یک تصویر هوایی با فاصله نمونه‌برداری زمینی (GSD) برابر با ۰.۲۵ متر از منطقه‌ای روستایی در کشور اتیوپی است. منطقه انتخابی بیشتر شامل کشاورزی خرده مالکی است. علاوه بر این، از دو تصویر دیگر برای ارزیابی مدل استفاده کردند که از دو کشور کنیا و رواندا است و به ترتیب دارای GSD برابر با ۰.۰۲ متر و ۰.۰۶ متر هستند. چارچوب آن‌ها برای استخراج مرز شامل سه مرحله می‌شود:

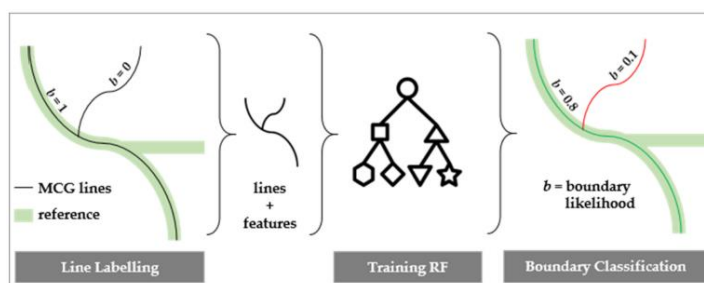
۱) **قطعه‌بندی تصویر برای استخراج خطوط اجسام مرئی:** قطعه‌بندی تصویر بر پایه گروه‌بندی ترکیبی وضوح چندگانه (MCG) انجام می‌شود که کانتورهای بسته را تحویل می‌دهد و خطوط اجسام مرئی را دریافت می‌کند. در این روش یک متغیر k وجود دارد که تنظیم قطعه‌بندی بیش از حد یا کم را کنترل می‌کند. آن‌ها مقدار $k=1.0$ را با هدف قطعه‌بندی بیش از حد انتخاب کردند و هم‌چنین خطوط اطراف نواحی کوچک‌تر از ۳۰ متر مربع با بخش‌های همسایه خود ادغام می‌شوند. در شکل ۲ می‌توان تاثیر این ساده‌سازی را دید.



شکل ۲: تصویر سمت چپ، مرزهای استخراج شده توسط MCC و تصویر سمت راست، مرزهای ساده‌سازی شده در مقاله [1].

۲) طبقه‌بندی مرز برای پیش‌بینی احتمال مرز بودن قطعات خطوط استخراج: طبقه‌بندی باینری روی ۶۰۰۰۰۰ خطوط MCC تولید شده در مرحله قبل اعمال می‌شود. آن‌ها در این مرحله دو رویکرد را برای پیدا کردن احتمال مرز بودن هر خط MCC بررسی کردند: جنگل تصادفی (RF) و شبکه‌های عصبی پیچشی. آموزش برای RF شامل خطوط و برای CNN قطعات تصاویر است. برای هر دو رویکرد، کاداستر مرجع بافری به ضخامت ۰.۴ متر می‌گیرد.

- طبقه‌بندی RF مطابق با شکل ۳ انجام می‌شود. یک خط به صورت خودکار در صورتی برچسب "مرز" می‌گیرد که هم‌پوشانی بیش از ۵۰ درصد با کاداستر مرجع که ضخامت ۰.۴ متر گرفته است داشته باشد.
- طبقه‌بندی CNN با آموزش CNN هایی تصویر محور که بهترین عملکرد را دارند انجام شده است. ابتدا قطعات تصاویر با سایز ۲۲۴ در ۲۲۴ پیکسل که خط MCC وسط آن قرار گرفته است از تصویر ارتو جدا می‌شود. سپس این قطعه برچسب "مرز" می‌گیرد اگر پیکسل مرکزی که خط MCC را پوشش می‌دهد، هم‌پوشانی با کاداستر مرجع ضخیم‌شده داشته باشد. هر قطعه با فاصله ۵ متر در طول خط MCC تولید می‌شود. در اینجا از CNN ای که از قبل آموزش دیده است، استفاده شده است و رویکرد یادگیری انتقالی با افزودن لایه‌هایی که قابل آموزش هستند انجام می‌شود. در نهایت به دلیل عملکرد بهتر VGG، این شبکه عصبی انتخاب و مورد استفاده قرار گرفته است.



شکل ۳: طبقه‌بندی خطوط مرزی در مقاله [1] بر پایه RF تا احتمال مرز بودن برای خطوط MCC بدست آید.

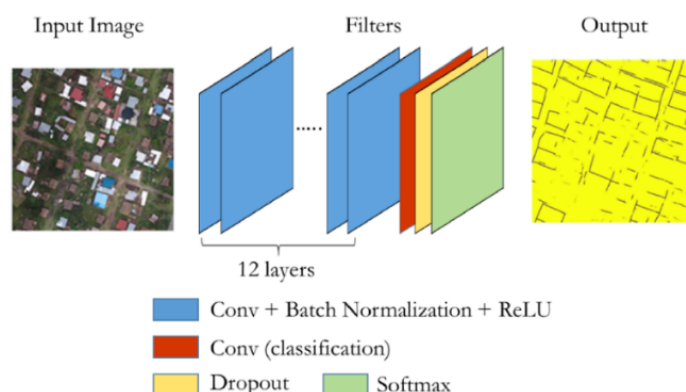
۳) ترسیم تعاملی برای اتصال این خطوط بر پایه پیش‌بینی انجام شده برای احتمال مرز: ترسیم تعاملی، مرزهای نهایی کاداستر را فراهم می‌کند. این افزونه متن-باز است و در QGIS پیاده‌سازی شده است و با عنوان BoundaryDelineation plugin [14] قابل دسترسی است. این افزونه به تکمیل و بهبود مرزهای استخراجی از مرحله قبل می‌پردازد. علت افزوده شدن مرحله سوم این است که یادگیری عمیق و دیگر روش‌هایی که برای حل این مسأله ارائه می‌شوند، هرچقدر هم که موفق عمل کنند، باز هم تنها قادر به یافتن مرزهای مرئی هستند و در پیدا کردن مرزهای نامرئی کاملاً ناتوان هستند. بنابراین نیاز به یک روش تکمیلی که امکان ایجاد تغییرات بر روی مرزهای استخراجی را به ما بدهد همواره در این مسأله حس می‌شود. استخراج مرزها به صورت خودکار و سپس مراجعه به نرم‌افزار برای بهبود و تکمیل آن، سرعت نقشه‌برداری کاداستر را بسیار کاراتر در تعداد کلیک و زمان مصرفی می‌کند تا این که از صفر و همه چیز با نرم‌افزار انجام شود.

مقاله بعدی که به آن می‌پردازیم در سال ۲۰۱۹ و متعلق به زبا و همکاران [8] است که تصمیم به استفاده از یک شبکه عصبی با عنوان شبکه عصبی تماماً پیچشی (FCN) گرفتند. تصاویر استفاده شده در این مقاله نیز از نوع UAV هستند و توسط

پروژه‌ای با نام its4land که در سال ۲۰۱۸ در رواندا انجام شده است، تهیه شده است. تصاویر دو منطقه از این کشور به نام‌های بوسوگو و موهوزا با GSD های به ترتیب ۲.۱۸ سانتی‌متر و ۲.۱۵ سانتی‌متر انتخاب شده‌اند که هر کدام به ۴ قطعه تصویر ۲۰۰۰ در ۲۰۰۰ پیکسل تقسیم شده‌اند. سه تا از این قطعات برای آموزش و دیگری برای آزمایش استفاده شده است. در این مقاله، تشخیص مرز به عنوان مسأله طبقه‌بندی تصویر تحت نظارت بر پایه پیکسل در نظر گرفته شده است که پیکسل‌ها به دو دسته "مرز" و "غیر مرز" دسته‌بندی می‌شوند. چارچوب این مقاله نیز شامل سه مرحله‌ی آماده‌سازی داده، تشخیص مرز و ارزیابی دقت است.

(۱) آماده‌سازی داده: در این مرحله سه قطعه برای آموزش و یکی برای آزمایش از بین چهار تصویری که در اختیار داشتند، انتخاب شده است که برای هر کدام یک تصویر RGB و یک مرز مرجع تهیه شده است. همچنین مرزهای مرجع را به دو دسته مرئی و نامرئی تقسیم کرده‌اند.

(۲) تشخیص مرز: سه روش برای تشخیص مرز کاداستر اعمال شده است که مقایسه ارزیابی عملکرد هر یک بر پایه تک قطعه تصویر آزمایش صورت گرفته است. اولین روش، روش منتخب آن‌ها یعنی FCN است و دو روش دیگر MRS و gPb است. شبکه مورد استفاده، FCN با هسته گسترش یافته FCN with Dilated Kernel (FCN-DK) است [15] که سه تغییر در آن داده شده است. معماری این شبکه تغییر یافته در شکل ۴ ترسیم شده است.



شکل ۴: معماری شبکه FCN استفاده شده برای مقاله [۱۵]

(۳) ارزیابی دقت: با اندازه‌گیری Precision، Recall، و F-score به مقایسه عملکرد سه روش انتخابی برای تشخیص مرز پرداخته‌اند که برای این کار بافر ۰.۴ متر برای مرزهای داده مرجع در نظر گرفته شده است. نتایج حاصل شده از این مقاله را می‌توانید در دو شکل ۵ و شکل ۶ مشاهده کنید. همانطور که در شکل ۱۰ مشخص است، عملکرد FCN در دو تصویر آزمایش بسیار متفاوت بوده است.

Algorithm	Reference	TS1			TS2		
		P	R	F	P	R	F
FCN	visible	0.75	0.65	0.70	0.74	0.45	0.56
	invisible	0.06	0.07	0.06	0.06	0.09	0.07
	all	0.78	0.39	0.52	0.79	0.35	0.48
gPb-owt-ucm	visible	0.21	0.87	0.34	0.23	0.93	0.37
	invisible	0.03	0.19	0.06	0.04	0.39	0.07
	all	0.24	0.57	0.33	0.26	0.78	0.39
MRS	visible	0.19	0.82	0.31	0.18	0.90	0.30
	invisible	0.05	0.27	0.08	0.04	0.56	0.08
	all	0.23	0.57	0.33	0.22	0.80	0.35

شکل ۵: مقایسه عملکرد سه روش gPb، MRS، و FCN با سه معیار Precision، Recall، و F-score در مقاله [8].



شکل ۶: مقایسه خروجی سه روش MRS، gPb و FCN در مقاله [8]. مرزهای مرئی با رنگ سبز و مرزهای نامرئی با رنگ قرمز مشخص شده‌اند.

سه مقاله بعدی متعلق به فتای و همکاران است که به ترتیب در سال‌های ۲۰۱۹ [16]، ۲۰۲۱ [17] و ۲۰۲۲ [18] انجام شده‌اند.

در اولین مقاله [16]، آن‌ها از ماژولی به نام ماژول استخراج ویژگی ENVI ENVI Feature Extraction یا به اختصار ENVI FX استفاده کردند که برای پردازش داده و استخراج ویژگی سریع و آسان از تصاویر با قدرت تفکیک‌پذیری بالا استفاده می‌شود. آن‌ها به اهمیت این که مرزها پس از استخراج توسط مدل، باید از فیلتری عبور کنند تا نتایج را بهبود ببخشند، اشاره کرده‌اند. برای دستیابی به هدف این مقاله، تصویر منطقه‌ای روستایی در اسلونی با GSD برابر با ۲ سانتی‌متر انتخاب شده است که شمار مرزهای مرئی در آن در مقایسه با مناطق متمرکز شهری بیشتر است.

ماژول ENVI FX ترکیب فرآیند قطعه‌بندی تصویر و طبقه‌بندی است که تمرکز این مقاله تنها بر روی قطعه‌بندی تصویر و محاسبه ویژگی‌های فضایی است. چارچوب آن‌ها برای حل مسأله نقشه‌برداری کاداستر بر پایه UAV شامل چهار مرحله است: (۱) پیش‌پردازش تصویر: اولین قدم، نمونه‌برداری کردن تصویر ارتوی UAV از ۲ سانتی‌متر به قدرت تفکیک‌پذیری مکانی پایین‌تر (در اینجا GSD برابر با ۲۵، ۵۰ و ۱۰۰ سانتی‌متر) است.

(۲) تشخیص و استخراج مرز: در اینجا ماژول ENVI FX اعمال می‌شود که نقشه‌گردانی توسط الگوریتم تشخیص لبه سوپل تولید می‌کند. در ادامه در سه مرحله خروجی ENVI FX تولید می‌شود.

i. هر پیکسل در نقشه‌گردانی یک سطحی دارد که هرچه مقدار آن بزرگ‌تر انتخاب شود، تعداد قطعه‌بندی‌ها کم‌تر می‌شود و اشیاء با بیشترین مرز مجزا، نگهداری می‌شوند و در نتیجه‌ی آن تصویری قطعه‌بندی شده می‌شود که هر شیء با مقدار میانگین طیفی تمام پیکسل‌های متعلق به آن شیء رنگ‌آمیزی می‌شود.

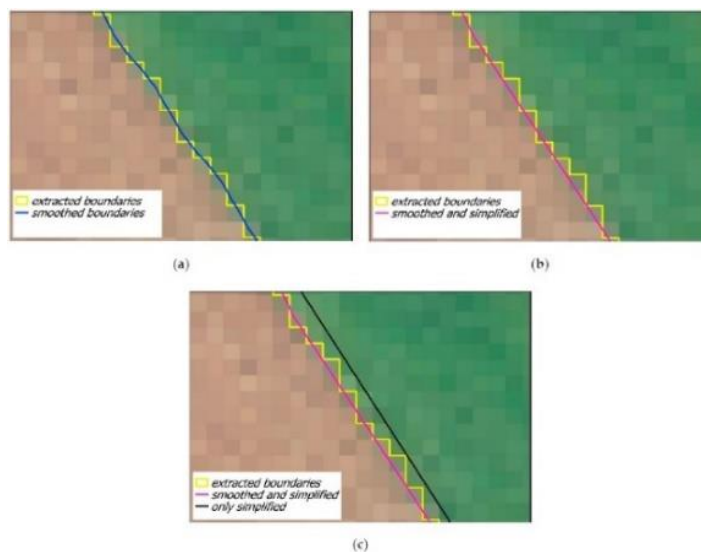
ii. در این مرحله ادغام انجام می‌شود: از بین بردن قطعه‌بندی‌های بیش از حد. در واقع اشیاء با مناطق بافت‌دار بزرگ‌تر مثل درخت‌ها و پهنه‌ها براساس ترکیب اطلاعات طیفی و فضایی تجمیع می‌شود.

iii. در مرحله نهایی صدور مرز اشیاء در فرمت برداری و یک تصویر قطعه‌بندی‌شده در فرمت رستری انجام می‌شود.

(۳) پس‌پردازش داده: این فرآیند شامل دو مرحله است، (i) فیلتر کردن اشیاء استخراجی و (ii) ساده‌سازی اشیاء استخراجی.

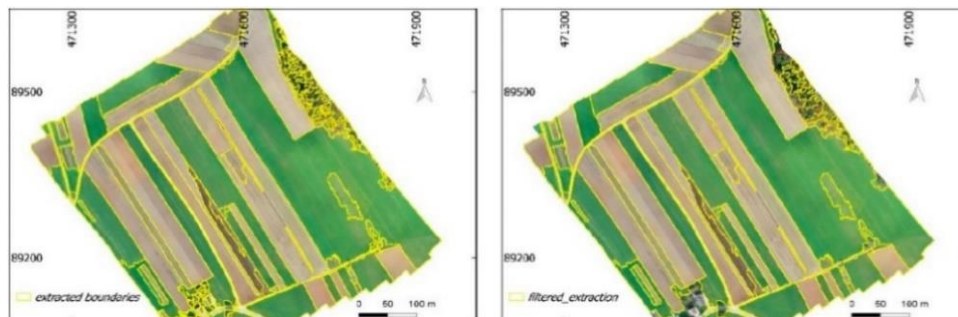
i. کم‌ترین مساحت اشیاء و کل تعداد اشیاء شناسایی شده در داده مرجع برای تعیین مقدار بهینه متغیرها استفاده می‌شود و مقادیر متغیرهایی که موجب دور شدن هرچه بیشتر از این دو مقدار شود، حذف می‌شوند. از طرفی در این مرحله اشیائی که مساحتی کوچک‌تر از مساحت کمینه دارند حذف می‌شوند.

ii. مرزهای اشیاء هموار و ساده می‌شوند. اولی با الگوریتم مار [19] و دومی با اعمال الگوریتم داگلاس-پکر [20] روی خطوط هموار شده انجام می‌شود. نتیجه اعمال دو الگوریتم مار و داگلاس-پکر را می‌توانید در شکل ۷ مشاهده کنید.



شکل ۷: تاثیر اعمال الگوریتم مار و داگلاس-پکر به همراه ترتیب به کارگیری آن‌ها در مقاله [19]

(۴) ارزیابی دقت: بر پایه شیء است چرا که نتایج به فرمت برداری هستند. آن‌ها برای محاسبه خطای Completeness، بافری برای خروجی مدل، و برای محاسبه خطای Correctness، بافری برای داده مرجع به طول ۲۵، ۵۰ و ۱۰۰ سانتی‌متر در نظر گرفتند. نتایج حاصل شده از این مقاله در شکل ۸ و شکل ۹ قابل مشاهده است.



شکل ۸: خروجی تولید شده توسط مدل در مقاله [16]. در شکل سمت چپ خروجی خام و در شکل سمت راست خروجی فیلتر شده را داریم.

Buffer width [cm]	Completeness [%]		Correctness [%]		Quality [%]	
	Extracted	Filtered	Extracted	Filtered	Extracted	Filtered
25	58	37	18	26	16	20
50	73	48	28	39	26	31
100	78	56	38	50	36	41
200	81 (81) ¹	61 (62) ¹	48 (49) ¹	59 (61) ¹	46 (46) ¹	50 (48) ¹

Buffer width [cm]	Completeness [%]		Correctness [%]		Quality [%]	
	Extracted	Filtered	Extracted	Filtered	Extracted	Filtered
25	45	40	28	35	21	23
50	64	55	46	56	38	41
100	71	61	57	68	48	52
200	75 (74) ¹	65 (67) ¹	65 (66) ¹	76 (77) ¹	56 (53) ¹	59 (56) ¹

Buffer Width [cm]	Completeness [%]		Correctness [%]		Quality [%]	
	Extracted	Filtered	Extracted	Filtered	Extracted	Filtered
25	31	27	21	24	14	15
50	53	47	39	43	29	30
100	67	59	58	64	47	47
200	73 (71) ¹	63 (67) ¹	66 (66) ¹	72 (73) ¹	55 (52) ¹	55 (52) ¹

شکل ۹: نتایج ارزیابی دقت مدل در مقاله [16]. جدول بالا با $GSD = 25$ ، جدول وسط با $GSD=50$ و جدول پایین با $GSD=100$ است

حال به سراغ دومین مقاله فتای و همکاران [17] که در سال ۲۰۲۱ ارائه شده است، می‌رویم. در این مقاله آن‌ها به بررسی شبکه U-Net و ماژول نرم‌افزار-محوری به نام ENVI یادگیری عمیق (ENVI Deep Learning) که رویکردهای یادگیری عمیق را با داده‌های جغرافیایی ساده‌سازی و پیاده کرده است می‌پردازند. داده‌هایی که در این مقاله استفاده شده‌اند، همگی تصاویر UAV هستند و از دو منطقه مختلف در اسلوونی و یک منطقه از مكدونیای شمالی تهیه شده‌اند. تصویر مربوط به اودرانسی اسلوونی دارای GSD برابر با 2.35 سانتی‌متر است و برای هر دوی آموزش و آزمایش استفاده شده است. تصویر تهیه شده از پونووا واس اسلوونی GSD برابر با 2.01 سانتی‌متر دارد و تنها برای آموزش استفاده شده است. در نهایت تصویر تتوو در مكدونیای شمالی GSD ای برابر با 2.85 سانتی‌متر دارد و تنها برای آموزش استفاده شده است.

چارچوب این مقاله شامل سه مرحله اصلی است. آماده‌سازی داده، تشخیص مرزهای مرئی زمین‌ها و ارزیابی دقت. ابتدا ساختار دو مدل U-Net و یادگیری عمیق ENVI را بررسی کرده و سپس به بررسی مرحله سوم می‌پردازیم.

- **U-Net**: در این مقاله مدل U-Net به روش یادگیری از ابتدا آموزش دیده است و تصاویر به قطعات ۲۵۶ در ۲۵۶ پیکسل تقسیم شده‌اند و برای این که محدوده دید برای هر قطعه بیشتر شود، دقت تصاویر ارتوی UAV اصلی به GSD بزرگ‌تر یعنی از ۲ تا ۳ سانتی‌متر به ۲۵ سانتی‌متر افزایش یافته است. هم‌چنین برای مرزهای مرجع، بافری به طول ۵۰ سانتی‌متر در نظر گرفته شده است.

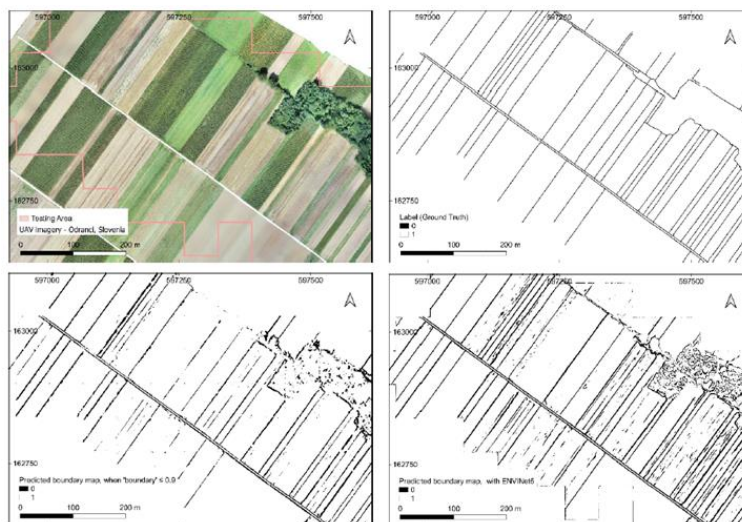
- **ENVI Deep Learning**: در این نرم‌افزار مدلی U-Net شکل پیاده‌سازی شده است که با عنوان ENVINet5 شناخته می‌شود. رویکرد آموزش این شبکه بر پایه پیچ است، به این معنا که داده آموزش UAV می‌تواند بعنوان ورودی داده شود و خود مدل می‌تواند بر اساس پیکسل‌های بخصوصی در پیچ، یادگیری را انجام دهد که سایز آن ۲۵۶ در ۲۵۶ تنظیم شده است.

- **ارزیابی دقت**: سنجش نهایی در این مقاله برای مقایسه دو مدل با سه Precision, Recall و F-Score انجام شده است که نتایج آن را می‌توانید در شکل ۱۰ مشاهده کنید.

Predicted Boundary Map	Overall Accuracy (%)	Recall	Precision	F1 Score
U-Net	94.5	0.654	0.413	0.506
ENViNet5	96.2	0.838	0.351	0.494

شکل ۱۰: مقایسه عملکرد دو مدل U-Net و ENViNet5 در مقاله [17].

در شکل ۱۱ می‌توانید تفاوت عملکرد این دو مدل در تولید خروجی را مشاهده کنید.



شکل ۱۱: شکل بالا سمت چپ، تصویر اصلی. شکل بالا سمت راست مرز مرجع. شکل پایین سمت چپ خروجی مدل U-Net و شکل پایین سمت راست خروجی مدل ENViNet5. همگی برای مقاله [17].

نوبت به بررسی آخرین مقاله فتای و همکاران [18] در سال ۲۰۲۲ رسیده است که چارچوب آن‌ها شامل چهار مرحله‌ی رویکرد آموزش، تشخیص مرز مرئی، ارزیابی دقت و برداری کردن نقشه مرز است.

۱. رویکرد آموزش و مجموعه داده: در این مقاله مجدداً از مدل U-Net استفاده شده است اما این بار تغییری در معماری اصلی شبکه داده شده است و اولین سطح از شبکه حذف شده است. برای آموزش مدل از مجموعه داده قطعه‌بندی Berkeley500 (BSDS500) که در [] قابل دسترسی است، استفاده کرده‌اند. این مجموعه داده برای قطعه‌بندی تصویر و مسائل تشخیص لبه یا مرز است که می‌تواند برای آموزش CNN استفاده شود، که با دامنه مسأله این مقاله هم‌خوانی دارد.

۲. تشخیص مرز مرئی از روی داده UAV: دو منطقه روستایی برای آزمایش مدل U-Net انتخاب شده است که یکی در پونووا و دیگری در اودرانسی است که هر دو در کشور اسلونی واقع شده‌اند. GSD تصویر اول 2.01 سانتی‌متر و دومی 2.35 سانتی‌متر است.

در اینجا علاوه بر مرزهای دستی، مرزهای کاداستر نیز برای هر تصویر تهیه شده است. مرزهای دستی شامل مرزهای مرئی می‌شوند که انتظار داریم مدل، آن‌ها را در طول آموزش بیاموزد و قادر به تشخیص آن‌ها باشد. اما مرزهای کاداستر، مرزهای حقیقی تصویر هستند که ممکن است بر روی مرزهای مرئی واقع نشده باشند.

۳. ارزیابی دقت: در اینجا ارزیابی مجدداً با سه فرمول Precision، Recall و F-score انجام شده است و نتایج بدست آمده از این مقاله با نتایج به دست آمده از مقاله [17] که با U-Net اصلی انجام شده بود، مقایسه شده است. این نتایج برای هر دو تصویر آزمایش به صورت جداگانه در شکل ۱۲ و شکل ۱۳ قابل مشاهده است.

Test Area	Predictions	Customised U-Net (CNN)			Original U-Net (CNN)		
		<i>recall</i>	<i>precision</i>	<i>F1 score</i>	<i>recall</i>	<i>precision</i>	<i>F1 score</i>
Ponova vas	threshold 0.3	0.600	0.597	0.598	0.565	0.520	0.542
	0.5	0.454	0.714	0.555	0.412	0.624	0.496

شکل ۱۲: مقایسه نتایج U-Net اصلی با U-Net تغییر یافته در مقاله [18] برای تصویر پونوا واس.

Test Area	Predictions	Customised U-Net (CNN)			Original U-Net (CNN)		
		<i>recall</i>	<i>precision</i>	<i>F1 score</i>	<i>recall</i>	<i>precision</i>	<i>F1 score</i>
Odranci	Threshold 0.3	0.534	0.549	0.598	0.495	0.494	0.494
	0.5	0.438	0.615	0.511	0.355	0.592	0.444

شکل ۱۳: مقایسه نتایج U-Net اصلی با U-Net تغییر یافته در مقاله [18] برای تصویر اودرانسی.

۴. برداری کردن نقشه مرز پیش‌بینی شده - تبدیل هاف: مرزهای استخراجی در فرمت رستر هستند که توسط تبدیل هاف تبدیل به بردار می‌شوند. خروجی تولید شده توسط این مقاله را برای دو تصویر پونوا واس و اودرانسی به صورت مجزا در شکل ۱۴ و شکل ۱۵ و همچنین نتایج بدست آمده از مقایسه مرزهای پیش‌بینی شده و مرزهای دستی با مرزهای کاداستر را برای این دو تصویر در شکل ۱۶ و شکل ۱۷ می‌توانید مشاهده کنید.



شکل ۱۴: خروجی تولید شده توسط مقاله [18] برای تصویر پونوا واس. تصویر سمت چپ مرز کاداستر، تصویر میانی مرز دستی و تصویر سمت راست مرز پیش‌بینی شده توسط مدل.



شکل ۱۵: خروجی تولید شده توسط مقاله [18] برای تصویر اودرانسی. تصویر سمت چپ مرز کاداستر، تصویر میانی مرز دستی و تصویر سمت راست مرز پیش‌بینی شده توسط مدل.

Cadastral Map				
	Boundary Map	<i>recall</i>	<i>precision</i>	<i>F1 score</i>
Ponova vas	Manually	0.051	0.090	0.065
	Predicted	0.055	0.095	0.070

شکل ۱۶: مقایسه با مرزهای کاداستر برای تصویر پونوا واس در مقاله [18].

Cadastral Map				
	Boundary Map	recall	precision	F1 score
Odranci	Manually	0.372	0.627	0.467
	Predicted	0.207	0.491	0.291

شکل ۱۷: مقایسه با مرزهای کاداستر برای تصویر اودرانسی در مقاله [18].

روش کار

در این جا راهکاری که در مقاله رحیم‌پور و همکاران [9] بررسی شده است را به کار می‌گیریم. فرآیندی که در آن مرز زمین‌های کشاورزی را از تصاویر ماهواره‌ای و هوایی استخراج می‌کنیم شامل دو مرحله‌ی پردازش و پس پردازش هندسی است.

پردازش

در فرآیند پردازش، هسته اصلی مدل، شبکه Mask R-CNN است. این مدل بعنوان ورودی، تصاویر با سایز 400 x 400 پیکسل را دریافت کرده و مرز هر زمین را در آن تشخیص می‌دهد. این تشخیص در دو مرحله توسط این شبکه صورت می‌گیرد. ابتدا زیرمجموعه‌ای از نواحی در تصویر را که ممکن است شامل یک شیء (در اینجا زمین‌ها) باشند شناسایی می‌کند و در مرحله دوم هر زمین در هر ناحیه را طبقه‌بندی می‌کند و نقشه احتمالاتی از امکان زمین بودن به هر پیکسل اختصاص می‌دهد. از آن جایی که تعداد داده‌هایی که برای آموزش در اختیار داریم محدود است، برای آموزش شبکه از روش یادگیری انتقالی استفاده می‌کنیم. برای این کار مدل Mask RCNN از شبکه از پیش آموزش دیده‌ی ResNet-50 که روی مجموعه داده‌ی Imagenet [21] آموزش دیده است بهره می‌گیرد. علت انتخاب شبکه ResNet-50 بعنوان هسته اصلی شبکه، توانایی آن در تبدیل تصاویر ورودی به ویژگی‌های برداری با سایز 4096 است. این ویژگی‌های برداری نقش ویژه‌ای در پیش‌بینی جعبه مرزی، کلاس و ماسک هر شیء شناسایی شده دارند. واضح است که بدون استفاده از شبکه ResNet-50، مدل Mask RCNN نیاز به حجم زیادی از داده برای آموزش دارد.

تصاویر ورودی برای آموزش شبکه، تصاویر RGB ماهواره‌ای خام تهیه شده از Google Earth یا تصاویر UAV از مناطق زراعی ایران هستند که هیچ‌گونه برچسب یا داده اضافی ندارند. برای پیدا کردن مرز زمین‌های کشاورزی به روش یادگیری با نظارت، به صورت دستی، ماسک‌ها توسط نرم‌افزار LabelMe [22] تهیه می‌شوند. با استفاده از این ماسک‌ها می‌توانیم مدل قطعه‌بندی نمونه‌ای را برای تشخیص هر زمین به کار گیریم.

بنابراین مدل Mask RCNN هر زمین را بعنوان یک شیء مجزا شناسایی می‌کند و بعنوان خروجی نقشه احتمالاتی می‌دهد که مشخص می‌کند هر پیکسل داخل زمین قرار دارد یا خیر. اما در اینجا ما نیاز به مرزهای زمین‌ها داریم و نه نواحی داخلی هر زمین. در ادامه در ۵ مرحله، خروجی خام را تولید می‌کنیم تا بتوانیم وارد مرحله پس پردازش هندسی شده و خروجی نهایی بدست آید.

۱. تصویر ورودی به بلوک‌های ۴۰۰ در ۴۰۰ پیکسل با قدم‌های ۲۰۰ پیکسلی شکسته می‌شود. از این طریق، بلوک‌های تصویر با یکدیگر هم‌پوشانی دارند. انتخاب اندازه بلوک‌ها در اینجا اهمیت زیادی دارد. زیرا مقدار کم‌تر موجب می‌شود تا یک شیء در بلوک‌های مختلف قرار گیرد و شناسایی آن سخت‌تر شود، در حالی که اندازه بزرگ‌تر باعث می‌شود مدل ویژگی‌های کوچک و در نتیجه زمین‌های کوچک را نادیده بگیرد و آن را بعنوان نویز شناسایی کند.

۲. هر کدام از این بلوک‌های تصویر به مدل Mask RCNN داده می‌شود و بعنوان خروجی، نقشه احتمالاتی دریافت می‌کنیم که مشخص می‌کند با چه احتمالی هر پیکسل داخل زمین قرار گرفته است.

۳. ماسک‌های خروجی شبکه با هم ادغام می‌شوند تا خروجی نهایی با سایز اصلی تصویر بدست آید. در نواحی هم‌پوشانی بلوک‌های تصویر، متدهای زیادی می‌توانند اعمال شوند. از جمله حداکثر ادغام، میانگین‌گیری، مجموع‌یابی و میانگین هارمونیک. در نهایت ماسک بزرگی از تصویر ورودی بدست می‌آید که مقادیر نزدیک به ۱ با احتمال بیشتری زمین هستند.

۴. برای استخراج مرز زمین‌ها از این تصویر خروجی، متد بابنری کردن اوتسو اعمال می‌شود تا نقشه احتمالات را بابنری کند.

۵. در نهایت الگوریتم تشخیص لبه کنی بر روی نقشه باینری اجرا می‌شود تا تصویری باینری که تنها شامل مرز زمین‌ها است، بدست آید.

حال با داشتن خروجی خامی که از این فرآیند بدست می‌آید، آماده هستیم تا با ورود به بخش پس پردازش هندسی، خروجی نهایی را بدست آوریم.

پس پردازش هندسی

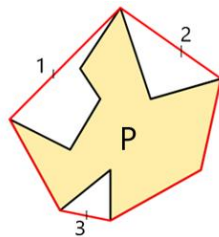
تصویر خروجی بدست آمده از مرحله قبل دارای فرمت رستری است که در واقع یک تصویر باینری است که مرز بودن یا نبودن هر پیکسل را مشخص می‌کند. در ادامه این خروجی را به فرمت برداری تبدیل می‌کنیم. بنابراین مرزهای بدست آمده در مرحله قبل را که برای هر زمین شامل یک چندضلعی است را استخراج کرده و این چندضلعی‌ها را در یک Shapefile به فرمت برداری ذخیره می‌کنیم. سه عملیات پس پردازش می‌تواند برای بهبود کیفیت مرزهای استخراجی روی چندضلعی‌های بدست آمده اجرا شود. این سه مرحله به صورت زیر است:

مرحله اول: چندضلعی‌هایی که مساحتی کوچک‌تر از مساحت یک زمین کشاورزی واقعی دارند حذف می‌شوند.

مرحله دوم: شبکه ممکن است چندضلعی‌های بسیار کوچکی درون چندضلعی‌های دیگر تشخیص دهد که از جمله دلایل آن می‌تواند وجود ساختمان در داخل زمین باشد. این چندضلعی‌ها شناسایی شده و حذف می‌شوند.

مرحله سوم: مرزهای استخراجی حالت پلکانی دارند و باید ساده‌سازی شوند که برای این کار روش جدیدی به نام الگوریتم ساده‌سازی بر پایه پاکت ارائه می‌دهیم.

پاکت مفهومی در هندسه محاسباتی است که در ادامه با تعریف پوسته محدب به آن می‌رسیم. زیر مجموعه S از فضا را محدب نامیم اگر و تنها اگر برای هر زوج نقاط $p, q \in S$ ، پاره خط \overline{pq} تماماً داخل S بیفتد. پوسته محدب مجموعه S ، کوچک‌ترین مجموعه محدبی است که شامل S باشد و با $CH(P)$ نمایش داده می‌شود. پاکت‌های یک چندضلعی ساده، نواحی چندضلعی شکل خارج از چندضلعی هستند که داخل پوسته محدب قرار گرفته‌اند. بعنوان مثال سه چندضلعی سفید رنگ در شکل ۱۸، پاکت‌های چندضلعی p هستند [23].



شکل ۱۸: چندضلعی زرد رنگ P و سه پاکت آن به رنگ سفید.

الگوریتم ساده‌سازی بر پایه پاکت به این شکل عمل می‌کند:

برای چندضلعی P ، تمام پاکت‌های آن را بدست می‌آوریم، $P_1 \dots P_k$ در ادامه برای هر پاکت P_i که $1 \leq i \leq k$ و $k \in \mathbb{N}$ ، دو فاصله را محاسبه می‌کنیم که آن‌ها را به ترتیب $dist$ و d می‌نامیم. تنها ضلعی از پاکت را که متعلق به چندضلعی P نیست، ضلع احتمالی و طول این ضلع را $dist$ می‌نامیم. d نیز مجموع طول اضلاع پاکت به جز ضلع احتمالی است. در ادامه باید از بین ضلع احتمالی و سایر اضلاع پاکت انتخاب کنیم که کدام یک به چندضلعی P' که ساده‌سازی شده‌ی چندضلعی P است افزوده شود. برای این کار آستانه تحمل t را انتخاب می‌کنیم و اگر $d < t \times dist$ ، آن‌گاه تمام اضلاعی که در محاسبه d لحاظ شده بودند را بعنوان ضلعی از چندضلعی P' در نظر می‌گیریم و در غیر این صورت، تک ضلع احتمالی را به P' می‌افزاییم. بنابراین چندضلعی P' به جای چندضلعی P در shapefile قرار می‌گیرد.

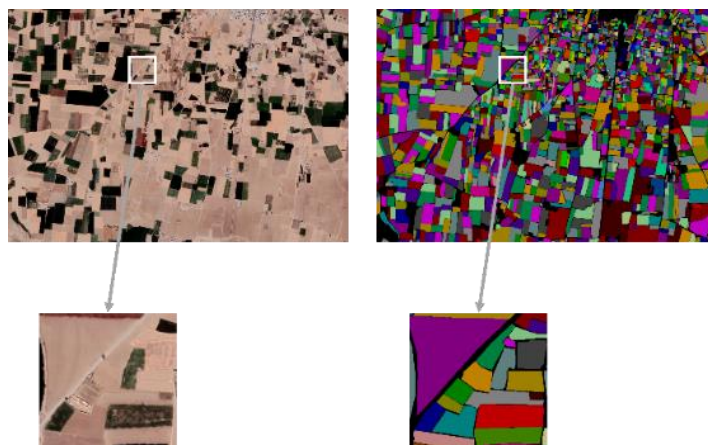
برای تمام چندضلعی‌هایی که از مرحله قبل استخراج شده‌اند این عملیات تکرار می‌شود تا در نهایت مجموعه‌ای از چندضلعی‌های P' که ساده‌سازی شده‌ی چندضلعی P هستند، در فایل نهایی قرار گیرند.

مطالعه تجربی

داده‌ها:

سه تصویری که برای آموزش شبکه مورد استفاده قرار گرفتند aerial image, ortho image و Famenin Irrigated نام دارند که دو تصویر اول از نوع UAV و تصویر سوم یک تصویر ماهواره‌ای بدست آمده Google Earth است. تصویر اول و سوم از مناطق زراعی در ایران هستند و تصویر دوم از کشور اتیوپی است. سایز این تصاویر به ترتیب برابر با 4936 در 2819، 3999 در 3999 و 5520 در 3776 پیکسل است.

از آنجایی که تصاویر توسط خودمان تهیه شده است و مجموعه داده‌ای در اینترنت برای رفع نیازهای ما یافت نشد، نیازمند این بودیم که ماسک تصاویر آموزش را خودمان تولید کنیم. با توجه به این که برای حل مسأله از رویکرد تقسیم‌بندی نمونه‌ای بهره بردیم، بنابراین ماسک‌ها نیز باید بر این اساس ساخته شوند. برای این کار از نرم‌افزاری به نام LabelMe استفاده کردیم که یک ابزار تفسیر گرافیکی تصویر است و به زبان پایتون نوشته شده است. در این‌جا با دقت مرزهای هر زمین را به صورت دستی برای هر تصویر رسم کرده و به هر یک از چندضلعی‌های تولید شده برای هر زمین رنگ مجزایی اختصاص می‌دهیم. لازم به ذکر است که زمین‌هایی که در همسایگی یکدیگر قرار ندارند (ضلع مشترک ندارند)، می‌توانند رنگ مشابهی بگیرند. شکل ۱۹ ماسک تهیه شده برای یکی از تصاویر آموزش را نشان می‌دهد.



شکل ۱۹: نمونه‌ای از تصویر آموزش به همراه ماسک آن.

تصویری که برای تست مورد استفاده قرار گرفته است، از منطقه‌ای زراعی در شهرستان کنگاور استان کرمانشاه با GSD برابر با 0.56 متر است که از آن با عنوان تصویر "کارخانه" یاد می‌کنیم و در آموزش شبکه به کار گرفته نشده است. سایز این تصویر 3930 در 3247 پیکسل است. مرز مرجع این تصویر به صورت دستی و به فرمت رستری تهیه شده است تا خروجی بدست آمده توسط شبکه با آن مقایسه شود و امکان ارزیابی دقت شبکه فراهم شود. دو بافر ۷ و ۸ پیکسلی به ترتیب معادل با خطای ۱۹۶ و ۲۲۴ سانتی‌متر برای مرز مرجع در نظر گرفته شده است در حالی که مرز خروجی شبکه با همان ضخامت ۱ پیکسلی مورد ارزیابی قرار می‌گیرد.

تصویری که برای مقایسه با خروجی در نظر گرفته شده است، تصویر مقاله رحیم‌پور و همکاران [9] است که از آن بعنوان تصویر "نگارخاتون" نام می‌بریم. این تصویر دارای سایز 3432 در 3621 با GSD برابر با 0.72 متر است و از منطقه‌ای زراعی در شهرستان فامنین استان همدان است. مرز مرجع این تصویر بافرهای ۵ و ۶ پیکسلی به ترتیب معادل با خطای ۱۸۰ و ۲۱۶ سانتی‌متر گرفته است.

تنظیمات سیستم

زبان برنامه‌نویسی استفاده شده برای تمام کدها پایتون است. با بهره‌گیری از GPU ی K80، آموزش شبکه تقریباً ۲ ساعت زمان می‌برد.

ارزیابی

سه متریکی که برای ارزیابی در نظر گرفته شده است Precision، Recall و F-score است که معیارهای معمول در این حوزه هستند و فرمول‌های هر یک به شرح زیر است:

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

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

$$\frac{2 \times Precision \times Recall}{Precision + Recall} F - score = (3)$$

Precision بیانگر این است که از بین تمام مرزهای استخراج‌شده، چند درصد آن به درستی تشخیص داده شده است. Recall مشخص می‌کند از بین تمام مرزهای حقیقی، چند درصد آن توسط مدل شناسایی شده است. F-score نیز میانگین هارمونیک دو معیار قبلی است که برای ارزیابی نهایی مورد استفاده قرار می‌گیرد.

با در نظر گرفتن طبقه‌بندی باینری و تشکیل ماتریس درهم‌ریختگی به طوری که کلاس مثبت پیکسل‌ها برچسب مرز داشته باشند و کلاس منفی برچسب غیر مرز، دو تصویر مرز استخراجی توسط مدل و مرز مرجع را روی هم انداخته و مقادیر را برای سه معیار معرفی شده محاسبه می‌کنیم. مقدار BF که در فرمول Recall مشاهده می‌شود، در واقع همان بافری است که برای مرز مرجع در نظر گرفته شده است. علت حضور این مقدار در فرمول این است که TP+FP نشان می‌دهد تعداد پیکسل‌هایی که مرز در نظر گرفته شده‌اند چقدر است. در حالی که مجموع صحیح مثبت (TP) با نادرست منفی (FN) نمایانگر تمام پیکسل‌های مرزی در مرز بافردار مرجع است و نه مرز مرجع اصلی با ضخامت یک پیکسل. بنابراین برای بدست آوردن مرزهای مرجع یک پیکسلی، باید حاصل جمع TP+FN را بر بافر تقسیم کنیم [8].

جدول ۱: ارزیابی دقت دو تصویر نگارخاتون و کارخانه با توجه به مقادیر مختلف بافر

F-Score	Recall	Precision	بافر	تصویر
76	87	67	۵ پیکسل	نگارخاتون
82	95	72	۶ پیکسل	نگارخاتون
77	85	71	۷ پیکسل	کارخانه
81	91	74	۸ پیکسل	کارخانه

همان‌طور که در جدول ۱ مشاهده می‌کنید، با وجود این که تصویر نگارخاتون دارای GSD بالاتر است و زمین‌های خرده‌مالکی زیادی را شامل می‌شود، به این معنا که محدوده‌ی دید در هر قطعه تصویر بزرگ‌تر است و زمین‌های بیشتری را شامل می‌شود، اما هم‌چنان تصویر کارخانه با GSD پایین‌تر که مساحت زمین‌ها به نسبت بزرگ‌تر است، به این معنا که محدوده‌ی دید در هر قطعه تصویر کم‌تر است و تنها بخشی از هر زمین ممکن است در لحظه توسط شبکه مشاهده شود، اما هم‌چنان شبکه در شناسایی مرزها خوب عمل کرده است و دارای Precision و Recall تقریباً برابر با تصویر نگارخاتون است. از

طرفی نزدیک بودن مقادیر Precision و Recall به یکدیگر با حفظ مقادیر بالا و Precision بالاتر به نسبت نگارخاتون، نشان می‌دهد که نیاز به اصلاحات در تصویر کارخانه کم‌تر است.

نتیجه‌گیری

در این مقاله رویکرد طبقه‌بندی نمونه‌ای برای استخراج مرز کاداستر از زمین‌های کشاورزی برگزیده شد. به این منظور، با به کارگیری شبکه عصبی Mask R-CNN این رویکرد را تحقق بخشیدیم. این شبکه که از طریق یادگیری انتقالی آموزش دیده است، از شبکه از پیش‌آموزش‌دیده‌ی ResNet-50 که بر روی مجموعه داده‌ی ImageNet آموزش دیده بود، در ساختار خود بهره برده است. خروجی شبکه نقشه‌ای احتمالاتی است که تعلق هر پیکسل به هر زمین کشاورزی را تعیین می‌کند. سپس روش باینری کردن اوتسو اعمال می‌شود تا نقشه احتمالاتی به نقشه‌ای باینری تبدیل شود. در نهایت الگوریتم تشخیص لبه کنی به استخراج مرز زمین‌های کشاورزی از روی تصویر خروجی شبکه می‌کند. در ادامه سه عملیات پس‌پردازش هندسی بر روی خروجی نهایی مرحله پردازش که از فرمت رستری به برداری تبدیل شده است اجرا می‌شود تا مرزهای تمیزتری بدست آیند. ابتدا زمین‌ها با مساحتی کوچک‌تر از مساحت یک زمین کشاورزی واقعی حذف می‌شوند، سپس به حذف زمین‌هایی که اشتباهات داخلی زمین‌های بزرگ‌تر شناسایی شده بودند می‌پردازیم و در نهایت با اجرای الگوریتم جدیدی به نام الگوریتم ساده‌سازی بر پایه پاکت که خودمان ارائه داده‌ایم، مرزها را ساده‌سازی می‌کنیم تا خروجی نهایی بدست آید.

راهکار ارائه شده در این مقاله، توانایی جداسازی مناطق شهری از مناطق زراعی را ندارد که موجب کاهش دقت در نتایج می‌شود. بنابراین می‌توان با حذف مناطق شهری از تصاویر پیش از این که تصویر بعنوان ورودی به شبکه Mask R-CNN داده شود، نتایج را بهبود ببخشد.

نتایج بدست آمده نشان می‌دهد قطعه‌بندی نمونه‌ای می‌تواند بعنوان راهکاری مفید برای حل مسأله استخراج خودکار مرز زمین‌های کشاورزی از تصاویر ماهواره‌ای یا هوایی در نظر گرفته شود. بنابراین بهره‌گیری از سایر شبکه‌های عصبی که در قطعه‌بندی نمونه‌ای عملکرد خوبی دارند، می‌تواند نتایج بهتری را به دنبال داشته باشد.

منابع و مراجع

- [1] S. Crommelinck, M. Koeva, M. Y. Yang and G. Vosselman, "Application of deep learning for delineation of visible cadastral boundaries from remote sensing imagery," *Remote sensing*, vol. 11, p. 2505, 2019 .
- [2] S. Enemark, K. C. Bell, C. Lemmen and R. McLaren, *Fit-for-purpose land administration*. International Federation of Surveyors (FIG), Copenhagen: International Federation of Surveyors (FIG), 2014 .
- [3] X. Luo, R. Bennett, M. Koeva, C. Lemmen and N. Quadros, "Quantifying the overlap between cadastral and visual boundaries: A case study from Vanuatu," *Urban Science*, vol. 1, p. 32, 2017 .
- [4] S. Crommelinck, R. Bennett, M. Gerke, F. Nex, M. Y. Yang and G. Vosselman, "Review of automatic feature extraction from high-resolution optical sensor data for UAV-based cadastral mapping," *Remote Sensing*, vol. 8, p. 689, 2016 .
- [5] K. Kocur-Bera and H. Frąszczak, "Coherence of cadastral data in land management—a case study of rural areas in Poland," *Land*, vol. 10, p. 399, 2021 .
- [6] M. Koeva, C. Stöcker, S. Crommelinck, M. Chipofya, K. Kundert, A. Schwering, J. Sahib, T. Zein, C. Timm, M. I. Humayun and others, "Innovative geospatial solutions for land tenure mapping," *Rwanda Journal of Engineering, Science, Technology and Environment*, vol. 3, 2020 .
- [7] X. Luo, R. M. Bennett, M. Koeva and C. Lemmen, "Investigating semi-automated cadastral boundaries extraction from airborne laser scanned data," *Land*, vol. 6, p. 60, 2017 .
- [8] X. Xia, C. Persello and M. Koeva, "Deep fully convolutional networks for cadastral boundary detection from UAV images," *Remote sensing*, vol. 11, p. 1725, 2019 .
- [9] A. A. M. T. H. F. S. R. K. A. J. Neda Rahimpour Anaraki, "Automatic CAdastral boundary Detection of Very High Resolution Images Using Mask-RCNN," *Journal of Electrical and Computer Engineering Innovations*, 2024 .
- [10] J.-D. Yang, Y.-S. Chen and W.-H. Hsu, "Adaptive thresholding algorithm and its hardware implementation," *Pattern Recognition Letters*, vol. 15, p. 141–150, 1994 .
- [11] J. Canny, "A computational approach to edge detection," *IEEE Transactions on pattern analysis and machine intelligence*, p. 679–698, 1986 .
- [12] L. Drăguț, O. Csillik, C. Eisank and D. Tiede, "Automated parameterisation for multi-scale image segmentation on multiple layers," *ISPRS Journal of photogrammetry and Remote Sensing*, vol. 88, p. 119–127, 2014 .
- [13] S. Crommelinck, R. Bennett, M. Gerke, M. Y. Yang and G. Vosselman, "Contour detection for UAV-based cadastral mapping," *Remote sensing*, vol. 9, p. 171, 2017 .
- [14] I. I. Sophie Crommelink, "QGIS Plugin Repository: BoundaryDelineation," 2019 .
- [15] C. Persello and A. Stein, "Deep fully convolutional networks for the detection of informal settlements in VHR images," *IEEE geoscience and remote sensing letters*, vol. 14, p. 2325–2329, 2017 .
- [16] B. Fetai, K. Oštir, M. Kosmatin Fras and A. Lisec, "Extraction of visible boundaries for cadastral mapping based on UAV imagery," *Remote sensing*, vol. 11, p. 1510, 2019 .
- [17] B. Fetai, M. Račić and A. Lisec, "Deep learning for detection of visible land boundaries from UAV imagery," *Remote Sensing*, vol. 13, p. 2077, 2021 .
- [18] B. Fetai, D. Grigillo and A. Lisec, "Revising cadastral data on land boundaries using deep learning in image-based mapping," *ISPRS International Journal of Geo-Information*, vol. 11, p. 298, 2022 .
- [19] GRASS Development Team, "Geographic Resources Analysis Support System (GRASS GIS) Software, Version 8.2," 2022.
- [20] D. H. Douglas and T. K. Peucker, "Algorithms for the reduction of the number of points required to represent a digitized line or its caricature," *Cartographica: the international journal for geographic information and geovisualization*, vol. 10, p. 112–122, 1973 .
- [21] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*, 2009 .
- [22] K. Wada, *labelme: Image Polygonal Annotation with Python*, GitHub, 2018 .

- [23]M. De Berg, Computational geometry: algorithms and applications, Springer Science & Business Media, 2000 .