



University of Tehran Press

ECO HYDROLOGY

Online ISSN: 2423-6101

Home Page: <https://ije.ut.ac.ir>

Water Bodies Extraction from Remote Sensing Data by Comparison of Deep Learning Models

Sina Khoshnevisan¹ | Saeid Gharechelou^{2*}  | Reza Mortazavi³ | Seyedeh Fatemeh Khakzad¹

1. MSc Student, Faculty of Civil Engineering, Shahrood University of Technology, Shahrood, Iran
2. Assistant Professor, Faculty of Civil Engineering, Shahrood University of Technology, Shahrood, Iran.
E-mail: sgharachelo@shahroodut.ac.ir
3. Assistant Professor, School of Engineering, Damghan University, Damghan, Iran.

ARTICLE INFO	ABSTRACT
<p>Article type: Research Article</p> <p>Article History: Received July 08, 2024 Revised August 20, 2024 Accepted September 21, 2024</p> <p>Keywords: Water bodies, Semantic image segmentation, Convolutional neural networks, Remote sensing, Deep learning.</p>	<p>In the current century, increasing of greenhouse gases has led to significant changes in the climate. It is causing irreversible impacts on agricultural lands, food production, and drinking water supply. By using remote sensing technology and the processing of satellite data, aerial and drone imagery to collect information from the Earth surface, environmental changes monitoring and analyze water bodies has become an effective tool for planning and optimal management of water resources. Modern and interdisciplinary technologies have enabled water resource specialists to accurately identify, mapping and assess surface water resources. In this study, with the aim of identifying small water bodies using remote sensing data, four deep learning models -ENet, SegNet, SE U-Net, and DeepLabV3+EfficientNet- were trained over 50 epochs using the Binary Cross-Entropy loss function. The results showed that the DeepLabV3+EfficientNet model with a Precision of 96.09% and an IoU of 89.13%, achieved the best performance in detecting agricultural ponds. Additionally, the SegNet model with a Precision of 93.81%, and the DeepLabV3+EfficientNet model with an IoU of 85.58%, demonstrated the best performance in detecting of swimming pools. Based on these results, the DeepLabV3+EfficientNet model is recommended by this research for pools and reservoirs detection.</p>

Cite this article: Khoshnevisan, S., Gharechelou, S., Mortazavi, R., Khakzad, F. (2024). Water Bodies Extraction from Remote Sensing Data by Comparison of Deep Learning Models. *ECOHYDROLOGY*. 11 (3), 321-336. Doi: <https://doi.org/10.22059/ije.2024.378101.1829>



© Sina Khoshnevisan, Saeid Gharechelou, Reza Mortazavi, Seyedeh Fatemeh Khakzad.

Publisher: University of Tehran Press.

DOI: <https://doi.org/10.22059/ije.2024.378101.1829>

استخراج پهنه‌های آبی از داده‌های سنجش از دور با مقایسه مدل‌های یادگیری عمیق

سینا خوشنویسان^۱، سعید قره‌چلو^۲ ، رضا مرتضوی^۳، سیده فاطمه خاکزاد^۱

۱. دانشجوی کارشناسی ارشد، دانشکده مهندسی عمران، دانشگاه صنعتی شاهرود، شاهرود، ایران

۲. استادیار، دانشکده مهندسی عمران، دانشگاه صنعتی شاهرود، شاهرود، ایران. رایانامه: sgharachelo@shahroodut.ac.ir

۳. استادیار، دانشکده فنی و مهندسی، دانشگاه دامغان، دامغان، ایران

اطلاعات مقاله	چکیده
نوع مقاله: پژوهشی	در قرن حاضر، افزایش غلظت گازهای گلخانه‌ای سبب تغییراتی در آب‌وهوا و اقلیم زمین شده است. این تغییرات تأثیرات جبران‌ناپذیری بر اراضی کشاورزی، تولید غذا و تأمین آب آشامیدنی داشته است. استفاده از فناوری سنجش از دور و پردازش تصاویر ماهواره‌ای، هوایی و پهپاد به منظور جمع‌آوری اطلاعات از سطح زمین، تغییرات زیست‌محیطی و تحلیل پهنه‌های آبی، آن را به ابزاری مؤثر برای برنامه‌ریزی، پایش و مدیریت بهینه منابع آب تبدیل کرده است. استفاده از فناوری های نوین و میان‌رشته‌ای امکان شناسایی، پهنه‌بندی و ارزیابی دقیق منابع آب سطحی را برای متخصصان منابع آب فراهم کرده است. در این پژوهش با هدف شناسایی پهنه‌های آبی سطحی با استفاده از داده‌های سنجش از دور چهار مدل یادگیری عمیق شامل SegNet، ENet، SE U-Net و DeepLabV3+EfficientNet تحت ۵۰ دوره آموزشی با تابع خطا باینری کراس-آنترופی (Binary Cross-Entropy) آموزش داده شده‌اند. نتایج نشان می‌دهد که مدل DeepLabV3+EfficientNet با مقدار ۰.۹۶/۰۹ در معیار Precision و ۰.۸۹/۱۳ در معیار IoU، بهترین عملکرد را برای تشخیص استخرهای کشاورزی داشته است. همچنین مدل SegNet با مقدار ۰.۹۳/۸۱ در معیار Precision و DeepLabV3+EfficientNet با مقدار ۰.۸۵/۵۸ در معیار IoU، بهترین عملکرد را به ترتیب برای تشخیص استخرهای شنا داشته‌اند. باتوجه به نتایج، مدل DeepLabV3+EfficientNet پیشنهاد می‌شود.
تاریخ‌های مقاله: تاریخ دریافت: ۱۴۰۳/۰۴/۱۸ تاریخ بازنگری: ۱۴۰۳/۰۵/۳۰ تاریخ پذیرش: ۱۴۰۳/۰۶/۳۱	
کلیدواژه‌ها: پهنه‌های آبی، بخش‌بندی معنایی تصاویر، شبکه‌های عصبی کانولوشنی، سنجش از دور، یادگیری عمیق	

استناد: خوشنویسان، سینا، قره‌چلو، سعید، مرتضوی، رضا، خاکزاد، سیده فاطمه. (۱۴۰۳). استخراج پهنه‌های آبی از داده‌های سنجش از دور با مقایسه مدل‌های یادگیری عمیق. *اِقَد هیدرولوژی*، ۱۱(۳)، ۳۲۱-۳۳۶.

DOI: <https://doi.org/10.22059/ije.2024.378101.1829>

ناشر: مؤسسه انتشارات دانشگاه تهران.

© سینا خوشنویسان، سعید قره‌چلو، رضا مرتضوی، سیده فاطمه خاکزاد.
DOI: <https://doi.org/10.22059/ije.2024.378101.1829>



۱. مقدمه

در دهه‌های اخیر با تصاعدی شدن غلظت گازهای گلخانه‌ای که حاصل از افزایش فعالیت‌های اقتصادی و اجتماعی است، آب‌وهوا و اقلیم کره زمین دستخوش تغییرات چشمگیری شده است [۱]. از جمله دلایل افزایش غلظت گازهایی مانند دی‌اکسید کربن و بخار آب در جو، سوزاندن سوخت‌های فسیلی برای تأمین تقاضای روزافزون انرژی توسط جوامع صنعتی است [۲]. این تغییرات منجر به پدیده‌هایی نظیر گرم شدن کره زمین، افزایش تبخیر و تعرق، و تغییرات در میزان بارش دریافتی هر منطقه می‌شود که با نام تغییر اقلیم شناخته می‌شود [۳ و ۴].

تغییرات اقلیمی و کاهش منابع آب سطحی در دسترس، امروزه به موضوعی مهم برای جوامع بشری تبدیل شده که تأثیر انکارناپذیری بر تولید غذا، کشاورزی و تأمین آب گذاشته است [۵-۷]. هر ساله با کاهش میزان بارندگی و افزایش دوره‌های خشکسالی، راهکارهای نوین برای بهره‌برداری حداکثری از آب و منابع آن در زمینه کشاورزی و افزایش راندمان تولید و کاهش مصرف آب نمود بیشتری پیدا می‌کند [۸]. از روش‌های نوین برای مدیریت مصرف آب در بخش کشاورزی، استفاده از استخرهای کشاورزی پلیمری و بتنی برای ذخیره آب در راستای برنامه‌ریزی مصرف آبیاری کنترل‌شده زمین‌های کشاورزی است [۹]. از طرف دیگر در دهه‌های اخیر، با افزایش سطح رفاه و درآمد جوامع، گرایش به تفریحات آبی نظیر ساخت استخرهای تفریحی و شنای خصوصی در منازل و باغ‌ویلاها افزایش چشمگیری داشته که به دنبال آن اولین تأثیر محسوس این اتفاق مصرف بی‌رویه آب برای بهره‌برداری از این تأسیسات و امکانات رفاهی و ازسوی دیگر، افزایش تبخیر بوده است [۱۰].

پایش و برنامه‌ریزی دقیق برای استفاده بهینه از منابع آب در سطح جهانی، اقدامی مهم و جامع در جهت دستیابی به اهداف توسعه پایدار، حفاظت از محیط‌زیست و حفظ این منبع ارزشمند برای نسل‌های آینده است. سنجش از دور در این زمینه به‌عنوان ابزاری کلیدی و قدرتمند برای گردآوری اطلاعات از راه دور و بدون نیاز به حضور در منطقه مطالعاتی، نقش مهمی برای رصد و نظارت بر منابع طبیعی و به‌طور خاص منابع آبی ایفا می‌کند [۱۱]. به‌وسیله سنجش از دور می‌توان با استفاده از تصاویر ماهواره‌ای یا پهپادی اطلاعات مورد نیاز و مربوط به پهنه‌های آبی را که شامل دریاچه‌ها، سدها، استخرهای کشاورزی و تفریحی است، در مناطق مختلف جهان با دقت بالا تجزیه و تحلیل نمود [۱۲]. اطلاعات کسب‌شده از تصاویر ماهواره‌ای و پهپادی، به شناسایی پهنه‌های آبی و تخمین میزان استفاده از روش‌های نوین آبیاری و کنترل منابع آبی و همچنین تشخیص استخرهای تفریحی برای برنامه‌ریزی سازمان‌های مرتبط جهت کنترل مصرف منابع آب شهری و شرب کمک‌کننده خواهد بود. به‌علاوه، این ابزار بستری را برای امکان تهیه نقشه‌های به‌روز برای شناسایی و تغییرات پهنه‌های آبی موجود به‌منظور تحلیل و ارائه راهکار مناسب در راستای مقابله با تغییرات اقلیمی و بحران آب فراهم می‌کند. در گذشته، کاوش میدانی و بازدید محلی تنها راه برای شناسایی و پایش پهنه‌های آبی بود؛ اما در سالین اخیر با وجود پیشرفت‌های قابل توجه در هوش مصنوعی به‌ویژه در زمینه بینایی ماشین رقم خورده است و می‌توان از این ابزار برای شناسایی دقیق با صرف کمترین زمان و منابع بهره برد [۱۳].

با توسعه هوش مصنوعی و یادگیری عمیق، محققان و پژوهشگران دریافتند که این شبکه‌ها می‌توانند با دقت بالا ویژگی‌های تصویر را برای بخش‌بندی معنایی استخراج کنند [۱۴]. در حال حاضر، دو روش متداول برای بخش‌بندی معنایی تصاویر وجود دارد: دسته اول شامل روش‌های سنتی مبتنی بر پردازش تصویر است که عمدتاً براساس استخراج ویژگی‌های تصویر مانند تشخیص لبه، نظریه گراف و بخش‌بندی براساس منطقه یا ناحیه عمل می‌کنند [۱۵]. از نقاط ضعف این روش‌ها می‌توان به کمبود اطلاعات معنایی در ویژگی‌های استخراج‌شده و نیز عملکرد ضعیف در سناریوهای پیچیده به‌دلیل پیچیدگی کم الگوریتم‌ها اشاره کرد؛ دسته دوم روش‌های بخش‌بندی تصاویر مبتنی بر شبکه‌های عصبی کانولوشنی عمیق^۱ است [۱۶]. شبکه‌های عصبی کانولوشنال نوعی از مدل‌های یادگیری عمیق‌اند که از لایه‌های متعددی از کانولوشن تشکیل شده‌اند. در این مدل‌ها، ویژگی‌های مورد نظر در تصویر با استفاده از لایه‌های کانولوشن که طی آموزش الگوها و ویژگی‌های معنی‌دار را استخراج می‌کنند، شناسایی می‌شوند. برای آموزش یک شبکه کانولوشنی، به مجموعه تصاویر هدف نیاز است. در سال‌های اخیر، با پیشرفت فناوری، استفاده از روش‌های سنتی کاهش یافته و استفاده از روش‌های بخش‌بندی معنایی تصاویر مبتنی بر شبکه‌های عصبی کانولوشنی گسترش یافته است.

دیوگانتی^۱ و همکاران در پژوهش خود با استفاده از مجموعه داده گردآوری شده از ماهواره سنتینل ۲، اقدام به بازسازی پهنه‌های آبی کردند. شبکه هوش مصنوعی انتخابی آن‌ها DeepLabV3+ بود و نتایج حاصل نشان داد که این شبکه دارای دقت بالایی در امر شناسایی است [۱۷]. در تحقیق دیگری که توسط هاریکا^۲ و همکاران به منظور شناسایی پهنه‌های آبی با استفاده از تصاویر ماهواره‌ای سنتینل ۲ صورت گرفت، نتایج به دست آمده نشان داد که DeepLabV3 می‌تواند برای شناسایی آب در تصاویر RGB یا با رنگ واقعی استفاده شود [۱۸]. در پژوهشی دیگر که با استفاده از مجموعه داده‌های ماهواره‌ای گردآوری شده از ماهواره GF1 و شبکه DeepLabV3+ توسط لی^۳ و همکاران صورت پذیرفت، مشخص شد که DeepLabV3+ از دقت بالایی برای شناسایی برخورداری است [۱۹]. ونگ^۴ و همکاران در مطالعه خود به مقایسه عملکرد شبکه‌های هوش مصنوعی مختلف برای شناسایی مناطق آبی در سنجش از دور پرداختند و نتایج نشان داد که شبکه SegNet دارای عملکرد مناسبی بوده است [۲۰]. ما^۵ و همکاران در تحقیقی به بررسی و مقایسه عملکرد شبکه‌های مختلف کانولوشنی برای بخش‌بندی معنایی تصاویر ماهواره‌ای با استفاده از شبکه‌های متنوع کانولوشنی کرد، با توجه به نتایج به دست آمده، استفاده از شبکه ENet را برای انجام بخش‌بندی معنایی تصاویر توصیه نمود [۲۲]. در مطالعه‌ای دیگر که توسط چاوراسیا^۷ و همکاران با استفاده از شبکه ENet برای انجام بخش‌بندی معنایی بر روی تصاویر انجام شد، مشخص شد که این مدل باعث کاهش تعداد پارامترها شده که در نتیجه علاوه بر صرفه‌جویی در زمان، قابلیت اجرا بر روی دستگاه‌های با منابع محدود را داراست [۲۳]. در پژوهش‌های دیگری که دو گروه از محققان با هدف بازسازی پهنه‌های آبی انجام دادند، از شبکه SE-UNet استفاده شد و گزارش گردید که این شبکه برای انجام بخش‌بندی معنایی دارای عملکرد مناسبی است [۲۴ و ۲۵]. السد^۸ و همکاران در پژوهش خود از مدل UNet اصلاح شده با نام SE-UNet که از مکانیزم Squeeze-and-Excitation (SE) بهره می‌برد، برای بخش‌بندی معنایی تصاویر هوایی به منظور بازسازی مناطق پوشیده شده از آب استفاده کردند. ارزیابی نتایج این پژوهش نشان داد که استفاده از مکانیزم SE در UNet به طور قابل توجهی عملکرد را بهبود می‌بخشد [۲۶].

این پژوهش به تحلیل و بررسی روش‌های نوین بخش‌بندی تصویر در حوزه سنجش از دور با تمرکز بر شناسایی دقیق‌تر استخرهای کشاورزی در تصاویر ماهواره‌ای و تهیه نقشه پهنه‌های آبی آن‌ها می‌پردازد. یکی از نوآوری‌های کلیدی این پژوهش، استفاده از برچسب‌زنی دستی دقیق داده‌های ماهواره‌ای است که امکان تهیه داده‌های با کیفیت بالا و بدون خطا را فراهم می‌کند و به‌ویژه در شرایطی که خودکارسازی برچسب‌زنی به دلیل پیچیدگی ویژگی‌ها یا تغییرات محلی دشوار است، اهمیت دارد. با مقایسه و ارزیابی مدل‌های مرسوم بخش‌بندی تصویر بر اساس شاخص‌های مختلف، مدلی بهینه شناسایی شد که توانسته است بر چالش‌های موجود نظیر سایه‌ها، درختان و شباهت برخی نواحی تحت کشت با پهنه‌های آبی به دلیل رشد جلبک‌ها و گیاهان غلبه کند. این چالش‌ها معمولاً منجر به تشخیص‌های نادرست می‌شوند. هدف این پژوهش ارائه راه‌حلی کارآمد برای غلبه بر مشکلات رایج در سنجش از دور و بهبود دقت تشخیص در مواجهه با چالش‌های موجود در تصاویر ماهواره‌ای برای تشخیص استخرهای کشاورزی است. همچنین با بهینه‌سازی مدل‌ها، این پژوهش افزایش قابلیت اجرایی مدل‌های یادگیری عمیق در کاربردهایی مانند مدیریت منابع آب، کنترل و برنامه‌ریزی مصرف منابع آبی و تبخیر و... می‌تواند کاربردهای وسیع‌تری را به دنبال داشته باشد.

۲. مواد و روش‌ها

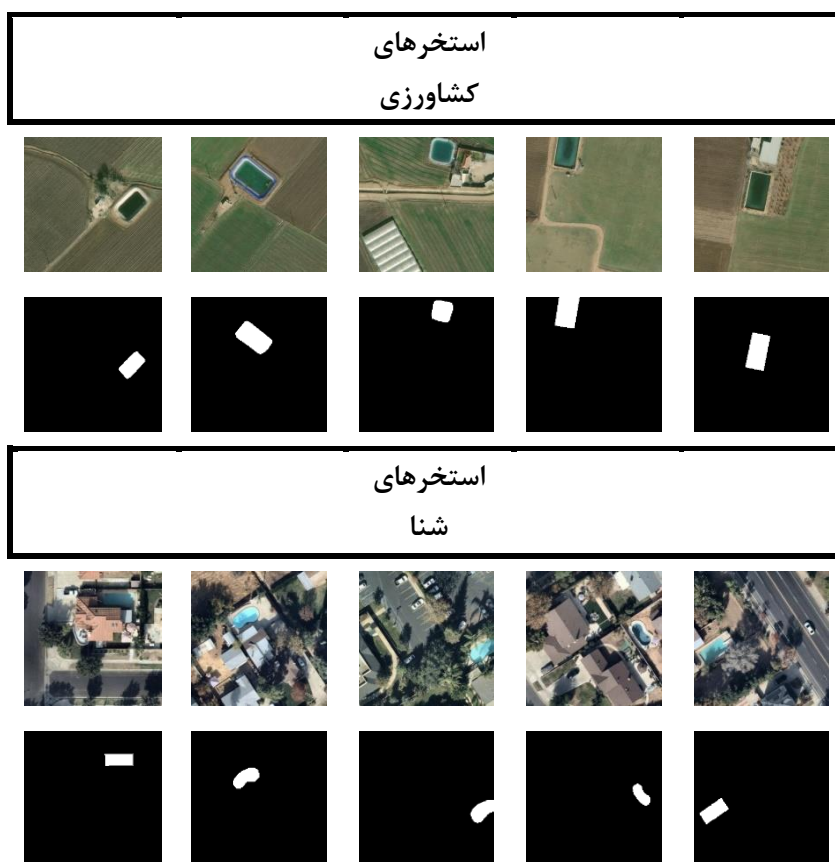
۲.۱. مجموعه داده

در این پژوهش، از دو مجموعه داده شامل عکس‌های ماهواره‌ای استخرهای شنای کالیفرنیا و استخرهای کشاورزی شهرستان شیراز و کرج استفاده شده است. انتخاب کالیفرنیا به عنوان منطقه مطالعاتی برای تشخیص استخرهای شنا به دلیل ابعاد کوچک استخرها و عدم

1. Edpuganti
2. Harika
3. Li
4. Weng
5. Ma
6. Wang
7. Chaurasia
8. Al-Saad

قابلیت تشخیص دقیق آن‌ها در تصاویر ماهواره‌ای با دقت متوسط انجام شده است، که می‌تواند بر مرحلهٔ برچسب‌زنی دقیق تأثیر منفی بگذارد. به همین دلیل، استفاده از داده‌های ماهواره‌ای با دقت پیکسلی بالا ضروری بوده و کالیفرنیا به دلیل دسترسی به داده‌های مناسب و تعداد زیاد استخرهای کوچک‌مقیاس، منبعی غنی و قابل اعتماد برای این تحقیق به شمار می‌آید. از سوی دیگر، باتوجه به تأثیر شرایط اقلیمی، به‌ویژه زمانی که از تصاویر با باند RGB استفاده می‌شود، تغییرات جوّی، نور، و رنگ‌های طبیعی می‌توانند به تغییرات چشمگیر در ویژگی‌های بصری تصاویر منجر شوند که این موضوع بر عملکرد مدل‌های مورد استفاده تأثیرگذار است. بنابراین، به‌منظور بهبود دقت، پیش‌پردازش‌هایی بر روی مجموعه‌داده‌ها صورت گرفته است. از میان تعداد تصاویر موجود در مجموعه‌داده تعداد ۷۴۶ تصویر با ابعاد ۲۵۶×۲۵۶ دارای کیفیت و وضوح مناسب برای این پژوهش انتخاب شد. در مجموع، ۷۵٪ از این تصاویر برای اهداف آموزشی با استفاده از روش دانه تصادفی^۱ برای تکرارپذیر کردن فرایند تخصیص داده‌ها استفاده شده است و ۲۵٪ باقی‌مانده برای ارزیابی عملکرد مدل‌ها در مرحلهٔ آزمون مورد استفاده قرار گرفته‌اند. مجموعه تصاویر شامل نمونه‌های گسترده‌ای از استخرهای تفریحی و کشاورزی است که شامل استخرهای شنای موجود در مناطق مسکونی و خصوصی، و همچنین استخرهای بتنی و پلیمری کشاورزی می‌شود. در انتخاب این مجموعه‌داده، معیارهایی نظیر جامعیت، پوشش انواع مختلف استخرها، تنوع در شکل و ابعاد آن‌ها و تنوع اقلیمی مورد توجه قرار گرفته است. فرایند برچسب‌گذاری تصاویر به‌صورت دستی انجام شده و به‌دلیل وجود سایهٔ درختان، ساختمان‌ها و ساختارهای نامنظم در تصاویر، این مرحله نیازمند صرف زمان و دقت بسیاری است. در گام بعد، برای اطمینان از صحت برچسب‌گذاری و کاهش خطا، تمامی داده‌ها به‌صورت دستی بازبینی شد و داده‌های دارای نویز و برچسب‌های نامناسب از مجموعه‌داده حذف گردید. به‌منظور افزایش تعداد داده‌ها در مرحلهٔ آموزش، از روش داده‌افزایی شامل برگرداندن و چرخش استفاده شد.

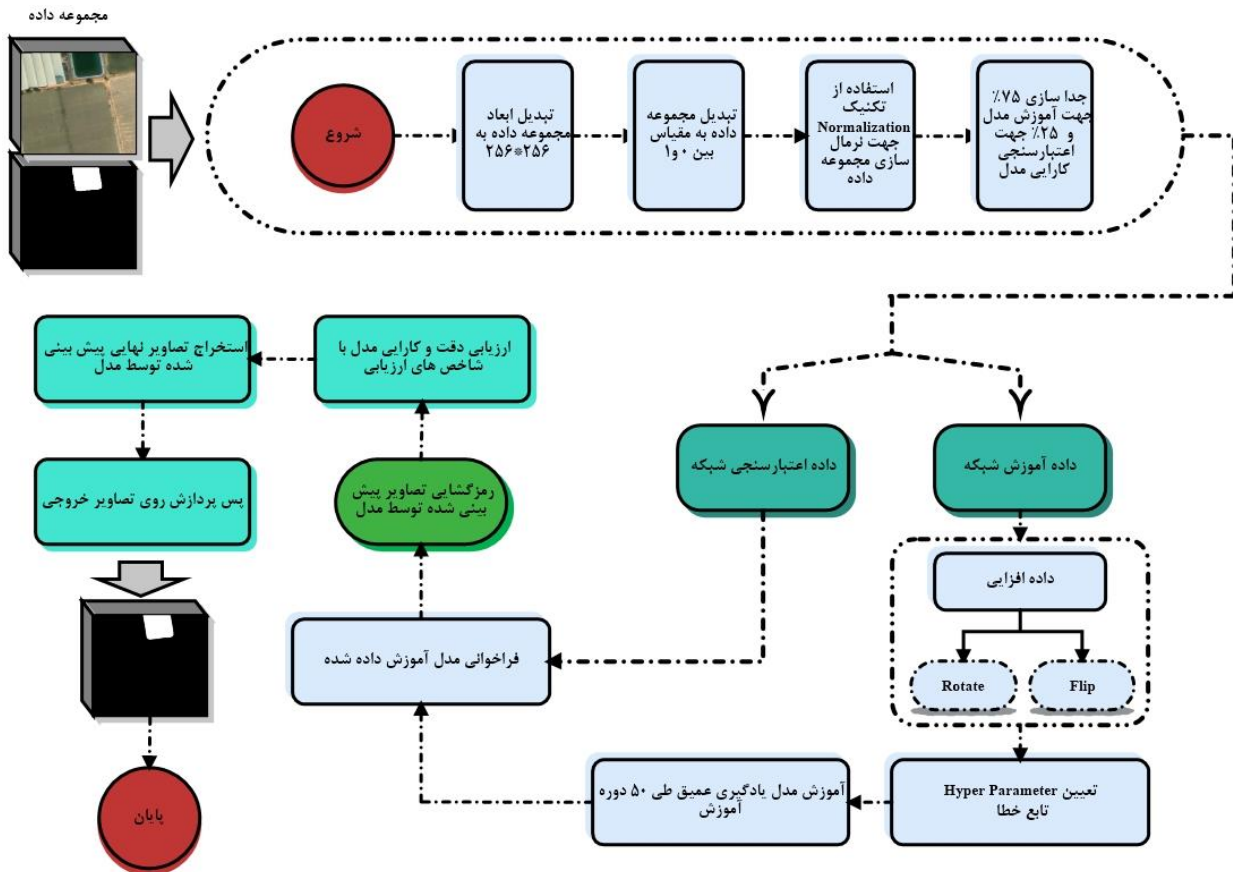
در شکل ۱، نمونه عکس‌های مورد استفاده در این پژوهش که به دو بخش استخرهای کشاورزی شهرستان شیراز و کرج و تفریحی (شنا) مربوط به کالیفرنیا تقسیم گردید، همراه با برچسب‌های مربوط است.



شکل ۱: نمونه تصاویر از مجموعه‌داده

۲.۲. روش کار

شکل ۲ طرح‌واره مراحل آموزش و ارزیابی مدل یادگیری عمیق برای پردازش و پیش‌بینی تصاویر را نمایش می‌دهد. ابتدا تصاویر مجموعه داده نرمال‌سازی و به ابعاد مناسب تغییر اندازه داده می‌شوند. سپس، مدل با استفاده از داده‌افزایی، تعداد و تنوع تصاویر مجموعه داده‌های آموزش را افزایش می‌دهد و در مرحله بعد، مدل، آموزش داده می‌شود و عملکرد آن با شاخص‌های مختلف ارزیابی می‌شود. در نهایت، روی تصاویر پیش‌بینی شده رمزگشایی صورت می‌پذیرد.



شکل ۲: مراحل انجام بخش‌بندی معنایی

۳.۲. معماری شبکه مورد استفاده

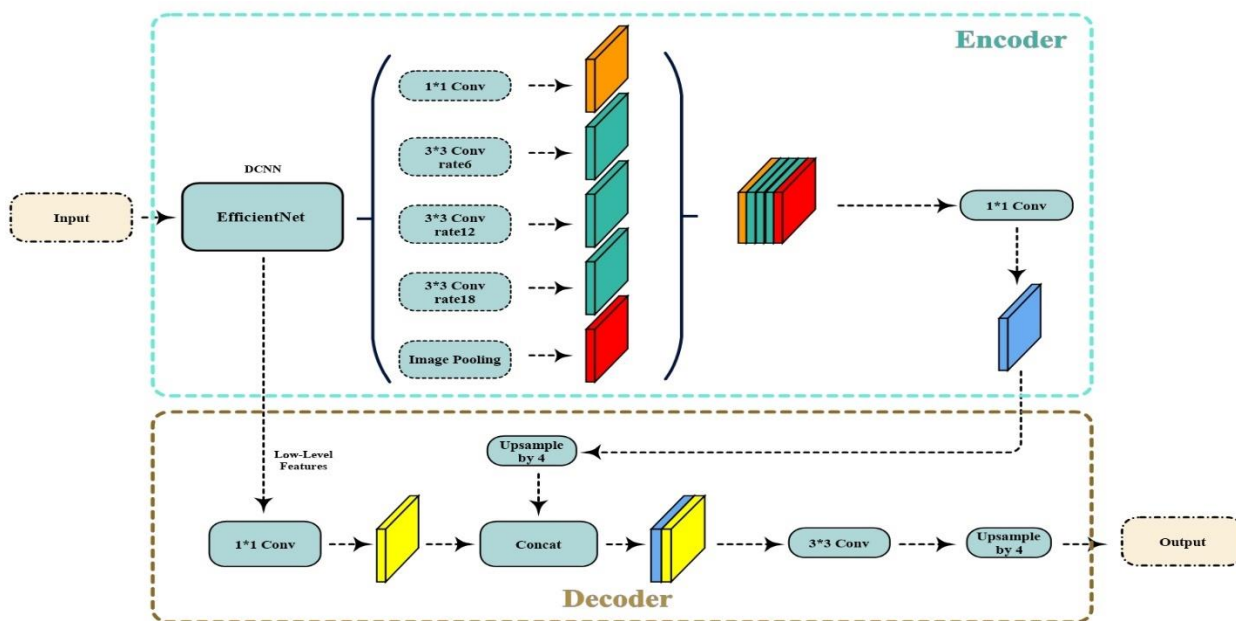
۳.۲.۱. DeepLabV3+EfficientNet

تا به حال، چندین نسخه مختلف از معماری شبکه DeepLab معرفی شده‌اند. از جمله مدل DeepLabV3+ که در سال ۲۰۱۸ توسط گروه تحقیقاتی گوگل ارائه شده است، به‌عنوان یک نسخه بهبودیافته بر پایه شبکه DeepLabV3 عرضه شده است. این مدل با هدف دسته‌بندی و تفکیک تصویر به پیکسل‌ها به چندین کاتلاف یا دسته طراحی شده است. الگوریتم شبکه DeepLabV3+ از یک ساختار رمزگذار-رمزگشا تشکیل شده است و برای جمع‌آوری اطلاعات زمینه‌ای چندمقیاسی، از کانولوشن آتروس^۱ (یا همان کانولوشن دیلیتد^۲) بهره می‌برد.

در مجموعه مدل‌های DeepLab معمولاً از معماری‌هایی مانند VGG، ResNet و EfficientNet به‌عنوان رمزگذار استفاده می‌شود. این معماری‌ها به‌همراه کانولوشن دیلیتد و سپس یک میدان تصادفی شرطی کاملاً متصل (CRF) به‌عنوان مرحله پس از پردازش، مورد استفاده قرار می‌گیرند. برای استخراج ویژگی‌های سطح بالا از تصویر ورودی، از EfficientNetB2 به‌عنوان رمزگذار استفاده می‌شود. این مدل، که یکی از مدل‌های شبکه‌های عصبی کانولوشنی است، در سال ۲۰۱۹ معرفی شد و با استفاده از روش

1. Atrous Convolution
2. Dilated Convolution

مقیاس‌بندی جدید، سعی در به دست آوردن دقت بیشتر با کمترین تعداد پارامترها دارد. این روش مقیاس‌بندی، با استفاده از یک ضریب ترکیبی ساده اما بسیار مؤثر، به‌طور یکنواخت تمام ابعاد عمق، عرض و تفکیک‌پذیری شبکه‌های عصبی را بهینه‌سازی می‌کند. شکل ۳ معماری شبکه DeepLabV3+EfficientNet را نشان می‌دهد.



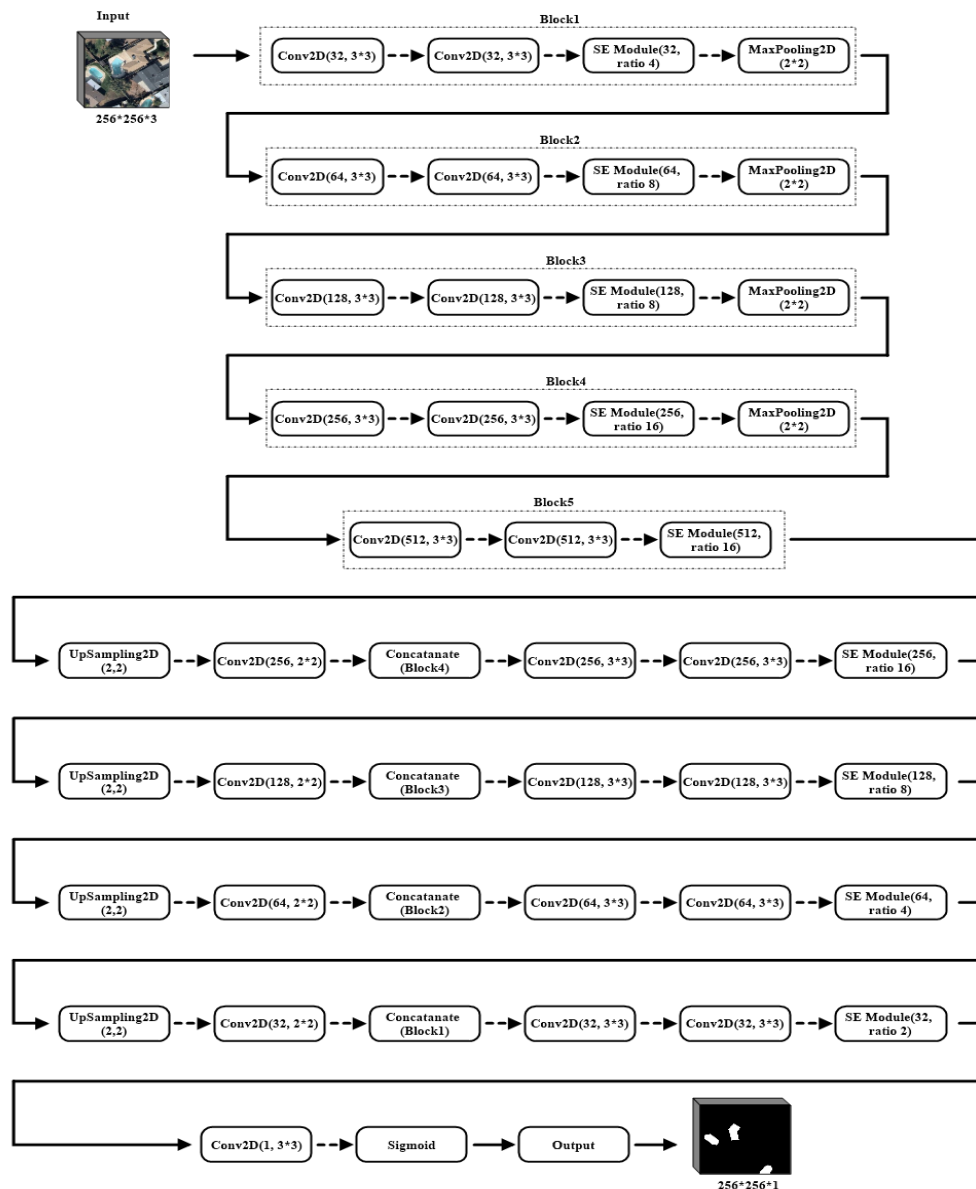
شکل ۳: معماری شبکه DeepLabV3+EfficientNet

۲.۳.۲. ENet

ENet یک معماری شبکه عصبی کانولوشنی است که برای بخش‌بندی تصاویر طراحی شده است. این مدل با توجه به سرعت و دقت بالا، برای کاربردهایی که نیاز به پردازش در زمان کمی دارند مناسب است. مدل ENet از ساختار رمزگذار-رمزگشا استفاده می‌کند که در قسمت رمزگذار، ویژگی‌های مختلفی از تصویر ورودی استخراج می‌شود و سپس در قسمت رمزگشا، این ویژگی‌ها برای پیش‌بینی برچسب هر پیکسل استفاده می‌شود. ENet با استفاده از روش‌های مختلفی مانند کانولوشن‌های نقطه‌ای به‌منظور کاهش تعداد پارامترهای شبکه و ماژول‌های باتلنک برای افزایش کارایی شبکه و منتقل کردن اطلاعات به‌شیوه مؤثرتری از لایه‌های قبلی به لایه‌های بعدی استفاده می‌کند. این کاهش در تعداد پارامترها باعث می‌شود که این مدل را بتوان با صرف منابع و زمان کمتر اجرا کرد.

۲.۳.۲. SE U-Net

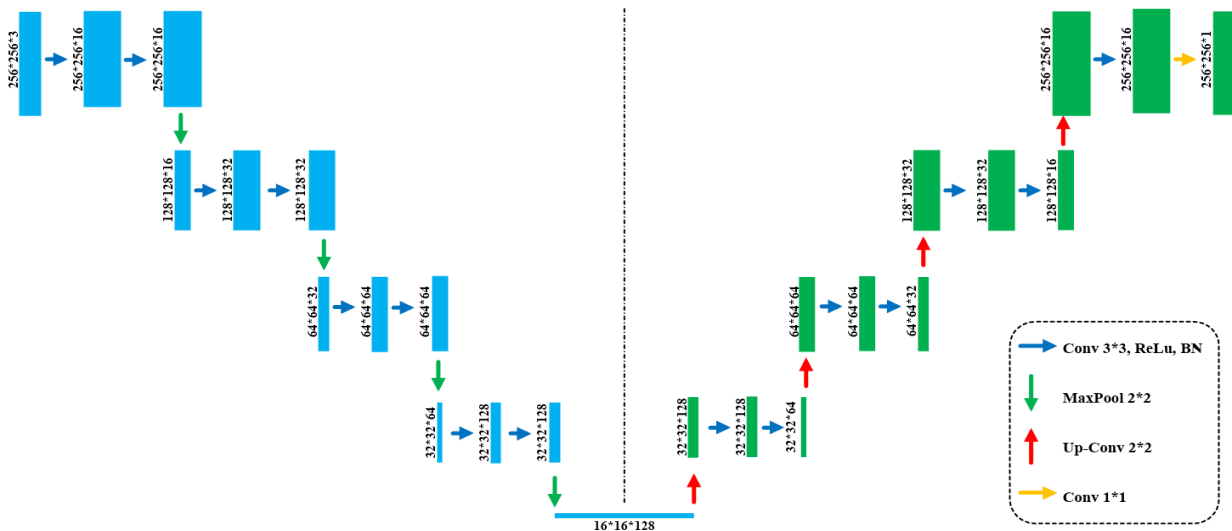
SE-U-Net براساس معماری موفق U-Net با ترکیب بلوک‌های Squeeze-and-Excitation (SE) ساخته شده است. بلوک‌های SE برای بهبود عملکرد مدل در زمینه‌های مختلف از جمله طبقه‌بندی تصویر شناخته می‌شوند. بلوک‌های SE این توانایی را دارند که به‌صورت پویا نتایج استخراج ویژگی را بازکالیبراسیون کنند. این کار که با برجسته‌سازی اطلاعات مهم و کاهش اطلاعات غیرضروری انجام می‌شود، می‌تواند در وظایف تقسیم‌بندی تصاویر مفید واقع شود. این کالیبراسیون مجدد هوشمندانه می‌تواند بهبود قابل توجهی در تمایز بین اشیا در تصویر و دقت بخش‌بندی به‌همراه داشته باشد. در شکل ۴، معماری شبکه SE U-Net نشان داده شده است.



شکل ۴: معماری شبکه SE U-Net

۲.۳.۴. SegNet

SegNet که گروه تحقیقاتی دانشگاه کمبریج در سال ۲۰۱۵ آن را معرفی کرد، یک معماری شبکه عصبی کانولوشنی است که برای انجام وظایف بخش بندی معنایی در زمینه های مختلف مانند تصویربرداری پزشکی و رانندگی خودکار طراحی شده است. این معماری از دو بخش رمزگذار و رمزگشا برای طبقه بندی کردن هر پیکسل در تصویر استفاده می کند. رمزگذار SegNet با استفاده از چندین لایه، ویژگی های سلسله مراتبی از تصاویر ورودی را استخراج می کند که می توانند شامل اطلاعاتی مانند رنگ، بافت، شکل و موقعیت مکانی باشند. سپس، رمزگشا SegNet با استفاده از لایه های برداری و کانولوشنی، خروجی بخش بندی شده را بازسازی می کند تا خروجی با وضوح اصلی حاصل شود که به دقت بالایی در تشخیص شکل و موقعیت اشیا منجر می شود. در نهایت، طبقه بندی دودویی پیکسل به پیکسل انجام می شود تا بین مناطق پیش زمینه و پس زمینه در خروجی بخش بندی شده تمایز قائل شود. این تمایز کمک می کند تا اشیای مختلف در تصویر به صورت دقیق تشخیص داده شود. در شکل ۵ معماری شبکه SegNet رسم شده است.



شکل ۵: معماری شبکه SegNet

۴.۲. شاخص‌های ارزیابی

در این بخش، معیارهای ارزیابی مختلف مورد استفاده در این مقاله معرفی می‌شود.

۴.۲.۱. IoU

Intersection over Union (IoU) یک شاخص ارزیابی است که نسبت مساحت تلاقی ماسک بخش‌بندی پیش‌بینی‌شده و ماسک واقعی را به مساحت اجتماع آن‌ها محاسبه می‌کند. تلاقی شامل تعداد پیکسل‌های به‌درستی شناسایی‌شده به‌عنوان پس‌زمینه در هر دو ماسک پیش‌بینی‌شده و واقعی است. اجتماع شامل تعداد کل پیکسل‌ها در هر دو ماسک است و شامل نتایج مثبت واقعی، نتایج مثبت کاذب و نتایج منفی کاذب است. IoU یک معیار ارزشمند است زیرا هم برای مثبت کاذب (پیکسل‌های پس‌زمینه که به‌صورت نادرست به‌عنوان پیش‌زمینه شناسایی شده‌اند) و هم برای منفی‌های کاذب (پیکسل‌های پیش‌زمینه به‌اشتباه به‌عنوان پس‌زمینه برچسب‌گذاری شده‌اند) جریمه اعمال می‌کند. در رابطه این شاخص ارزیابی، مثبت‌های واقعی (TP) به پیکسل‌های پیش‌زمینه اشاره دارد که به‌درستی توسط مدل شناسایی شده است. مثبت‌های کاذب (FP) به پیکسل‌های پس‌زمینه اشاره دارد که به‌اشتباه به‌عنوان پیش‌زمینه طبقه‌بندی شده است. درمقابل، منفی‌های کاذب (FN) به پیکسل‌های پیش‌زمینه اشاره دارد که به‌اشتباه به‌عنوان پس‌زمینه برچسب‌گذاری شده‌اند. این معیار از طریق رابطه (۱) محاسبه می‌شود:

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

۴.۲.۲. Dice Coefficient

Dice Coefficient یک شاخص ارزیابی برای سنجش میزان شباهت بین ماسک خروجی از مدل و ماسک واقعیت زمینی ساخته‌شده توسط ناظر است. این شاخص ارزیابی بر پایه تعداد پیکسل‌هایی که در هر دو ماسک خروجی توسط مدل و ماسک واقعیت زمینی توسط ناظر به‌درستی به‌عنوان پیش‌زمینه شناسایی شده‌اند، محاسبه می‌شود. این معیار اطلاعات دقیقی از عملکرد مدل در تشخیص و جداسازی اجسام مختلف در تصویر ارائه می‌دهد. با استفاده از این معیار، می‌توان با دقت بیشتری عملکرد مدل در تشخیص پهنه‌های آبی و مرزهای استخرها را ارزیابی کرد. این امر کمک می‌کند فهم بهتری از توانایی‌ها و محدودیت‌های مدل‌های مورد استفاده به وجود آید. به عبارت دیگر، این شاخص ارزیابی مهم کمک می‌کند تا درک دقیق‌تری را از توانایی مدل در تشخیص و تفکیک پهنه‌های آبی و مرزهای استخرها داشته باشیم و این امکان ایجاد شود که با اطمینان بیشتری به صحت و دقت نتایج حاصل از خروجی مدل اعتماد کنیم. شاخص ارزیابی Dice Coefficient با استفاده از رابطه (۲) محاسبه می‌شود:

$$Dice\ Coefficient = \frac{2TP}{2TP + FP + FN} \quad (2)$$

۲.۴.۳ Precision

Precision در بخش‌بندی تصویر دودویی، معیاری است که نشان می‌دهد چه تعداد از پیکسل‌هایی که مدل به‌عنوان پیش‌زمینه شناسایی کرده، به‌درستی در دسته پیش‌زمینه قرار دارند. این شاخص، عملکرد مدل در تشخیص صحیح پیکسل‌های پیش‌زمینه را ارزیابی می‌کند. Precision شاخص مهمی در بخش‌بندی تصویر دودویی است که کمک می‌کند تا به درک بهتری بابت اینکه چقدر می‌توان به نتایج بخش‌بندی اعتماد کرد دست یافت؛ به‌خصوص در تمایز بین پیکسل‌های پیش‌زمینه واقعی و کاذب. مقدار بالای دقت، نشان‌دهنده این است که مدل با موفقیت توانسته تعداد پیکسل‌های پیش‌زمینه غلط را کاهش دهد، که منجر به نتایج بخش‌بندی دقیق‌تر می‌شود. در حوزه‌های گوناگونی نظیر تصویربرداری پزشکی، رانندگی خودروهای اتوماتیک و تجزیه و تحلیل تصاویر ماهواره‌ای، دقت بالا در بخش‌بندی تصاویر برای انجام وظایفی مثل شناسایی اشیا، مشخص کردن مرزها و تفسیر تصاویر، ضروری است. بنابراین، دقت به‌عنوان یک شاخص کلیدی از کارایی مدل و مناسب بودن آن برای کاربردهای عملی مورد توجه است. این شاخص ارزیابی با استفاده از رابطه (۳) محاسبه می‌شود:

$$\text{Precision} = \frac{TP}{TP+FN} \quad (3)$$

۲.۵. تابع اتلاف (Loss Function)

در این پژوهش، از تابع خطا باینری کراس-آنترپی (Binary Cross-Entropy) استفاده شد. باینری کراس-آنترپی یک تابع خطاست که فاصله بین مقدار واقعی و پیش‌بینی شده برای هر دسته یا کاتلاف را اندازه‌گیری می‌کند؛ سپس خطاها را براساس کاتلاف، میانگین می‌گیرد تا مقدار نهایی را محاسبه کند. بنابراین این تابع خطا سعی می‌کند تا مقدار پیش‌بینی شده را به مقدار واقعی نزدیک کند. در این پژوهش، باتوجه به اجرای تقسیم‌بندی باینری، می‌توان فرض کرد که فقط دو کاتلاف برای هر پیکسل وجود دارد. این تابع از طریق رابطه (۴) محاسبه می‌شود:

$$(y, \hat{y}) = -\frac{1}{N} \sum_{i=0}^N (y * \log(\hat{y}_i) + (1 - y) * (1 - \hat{y}_i)) \quad (4)$$

در این رابطه، نماد N برابر است با تعداد نمونه‌ها و \hat{y}_i به ترتیب نشان‌دهنده بردار واقعی برجسبها و بردار پیش‌بینی هستند.

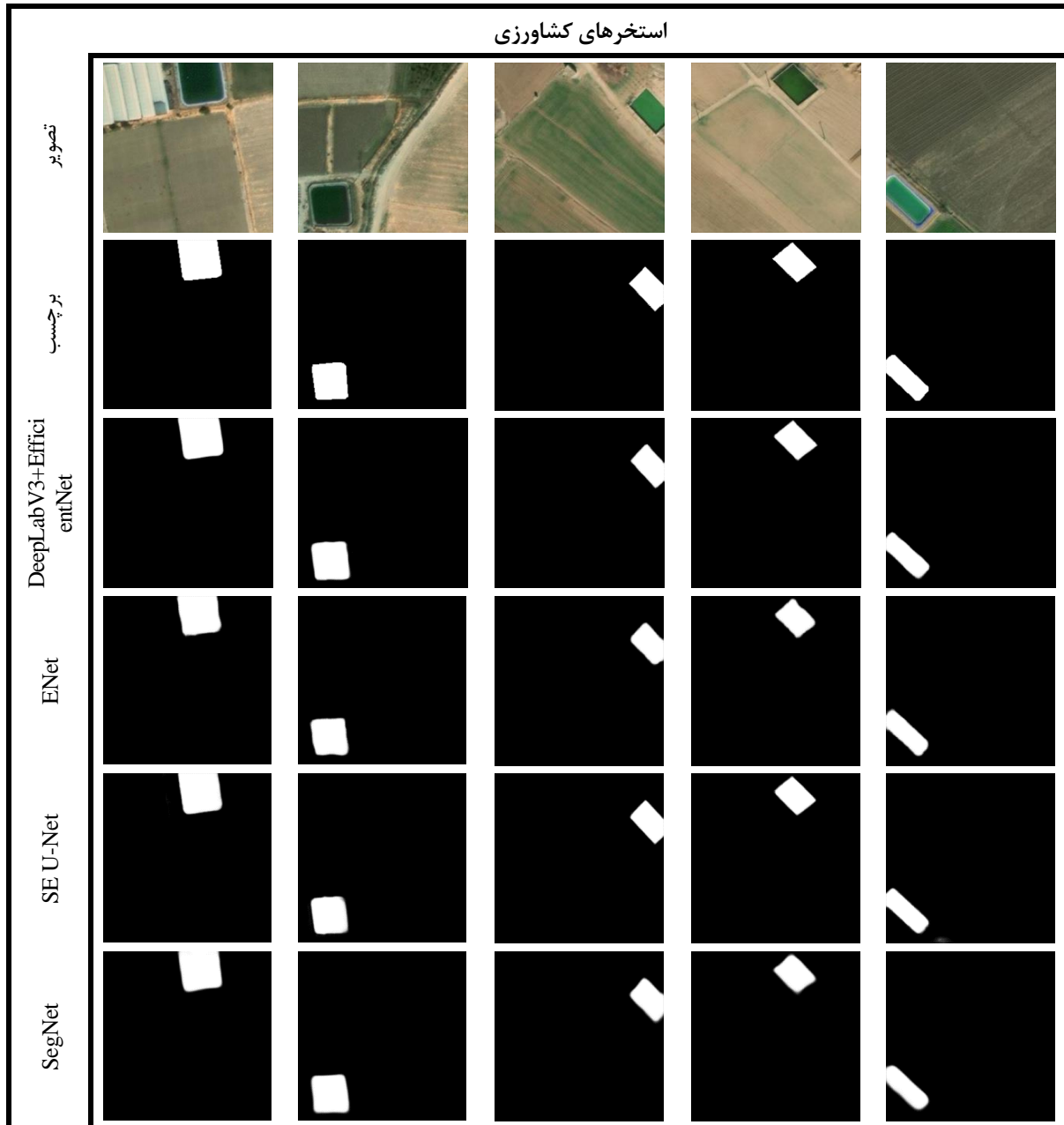
۳. نتایج و بحث

در این بخش، نتایج به‌دست‌آمده از بخش‌بندی معنایی پهناهای آبی مورد بحث و بررسی قرار می‌گیرند. نتایج حاصل از چهار مدل مختلف یادگیری عمیق برای تشخیص استخرهای کشاورزی در داده‌های ماهواره‌ای شهرستان کرج و شیراز در جدول ۱ ذکر شده است. شبکه‌های مورد استفاده در این تحقیق بر روی سیستمی با کارت گرافیک Nvidia RTX 3060 با ۱۲ گیگابایت حافظه گرافیکی و ۱۶ گیگابایت رم، با استفاده از کتابخانه TensorFlow آموزش داده شده‌اند. نرخ یادگیری (learning rate) در طول آموزش مدل‌ها ۰/۰۰۳ تنظیم شده است. به‌منظور ارزیابی مدل‌های حاضر در این پژوهش از معیارهای ارزیابی شامل دقت (Precision)، ضریب تاس (Dice Coefficient) و IoU استفاده شده است. در معیار دقت Precision، مدل DeepLabV3+EfficientNet با دقت ۹۶/۰۹٪ بهترین عملکرد را داشته است. این نتیجه نشان می‌دهد که ۹۶/۰۹٪ از پیکسل‌های شناسایی شده به‌عنوان استخر کشاورزی توسط این مدل، واقعاً استخرهای کشاورزی هستند. مدل‌های SE U-Net، SegNet و ENet نیز در جایگاه‌های بعدی قرار گرفته‌اند. باتوجه به ضریب تاس، که یک معیار برای اندازه‌گیری شباهت بین دو مجموعه است، مدل DeepLabV3+EfficientNet با ۹۳/۶۲٪ بهترین عملکرد را داشته است. این نتیجه نشان می‌دهد که ۹۳/۶۲٪ از پیکسل‌های پیش‌بینی شده توسط این مدل، با پیکسل‌های استخرهای کشاورزی واقعی هم‌پوشانی دارند. درنهایت، از نظر شاخص ارزیابی IoU، بهترین عملکرد مربوط به مدل DeepLabV3+EfficientNet با ۸۹/۱۳٪ است. دو مدل SegNet و SE U-Net نیز با مقادیر ۸۵/۱۹٪ و ۸۴/۷۹٪ به ترتیب در جایگاه‌های بعدی قرار دارند و ضعیف‌ترین مدل با مقدار IoU برابر ۸۳/۶۸٪ مربوط به مدل ENet است.

نتایج ارزیابی چهار مدل یادگیری عمیق برای تشخیص استخرهای کشاورزی در تصاویر ماهواره‌ای شهرستان کرج و شیراز نشان می‌دهد که مدل DeepLabV3+EfficientNet با مقدار ۹۶/۰۹٪ در معیار Precision، عملکرد بهتری نسبت به سایر مدل‌ها داشته است. همچنین در معیارهای ضریب تاس و IoU نیز، این مدل به ترتیب با مقادیر ۹۳/۶۲٪ و ۸۹/۱۳٪ بهترین عملکرد را از خود نشان داده است. در شکل ۶، تصاویر خروجی بخش‌بندی معنایی تصاویر هر مدل به‌منظور مقایسه قرار داده شده است.

جدول ۱: نتایج ارزیابی خروجی شبکه‌های یادگیری عمیق بر روی تصاویر پهنه آبی استخرهای کشاورزی

نتایج بر روی تصاویر استخرهای کشاورزی			نام شبکه
Precision	Dice Coefficient	IoU	
۹۶/۰۹	۹۳/۶۲	۸۹/۱۳	DeeplabV3+ EfficientNet
۹۳/۳۳	۹۰/۵۰	۸۳/۶۸	ENet
۹۵/۱	۹۳/۰۹	۸۴/۷۹	SE U-Net
۹۳/۷۴	۹۳/۱۶	۸۵/۱۹	SegNet



شکل ۶: نمونه تصاویر استخرهای کشاورزی و خروجی از شبکه‌های یادگیری عمیق

نتایج ارزیابی چهار مدل مختلف یادگیری عمیق برای تشخیص استخرهای شنا در تصاویر ماهواره‌ای مربوط به مجموعه داده تصاویر کالیفرنیا در جدول ۲ ذکر شده است. باتوجه به شاخص ارزیابی Precision، مدل SegNet با دقت ۹۳/۸۱٪ بهترین عملکرد را داشت و

پس از آن، از مدل‌های DeepLabV3+EfficientNet، SE U-Net و ENet به ترتیب مقادیر $۹۳/۵۳$ ، $۹۳/۲۲$ و $۹۱/۶۵$ به دست آمد. نتایج بررسی مدل‌ها تحت شاخص ارزیابی ضریب تاس نشان داد که مدل DeepLabV3+EfficientNet با ضریب تاس $۹۹/۷۷$ بهترین عملکرد را داشت و ضعیف‌ترین مدل براساس این شاخص، مدل SegNet با ضریب تاس $۸۹/۸۹$ بود. مدل DeepLabV3+EfficientNet با دقت $۸۵/۵۸$ دارای عملکرد بهتر نسبت به سه مدل دیگر تحت شاخص IoU بود. سه مدل SE U-Net، SegNet و ENet به ترتیب در جایگاه‌های بعدی قرار داشتند.

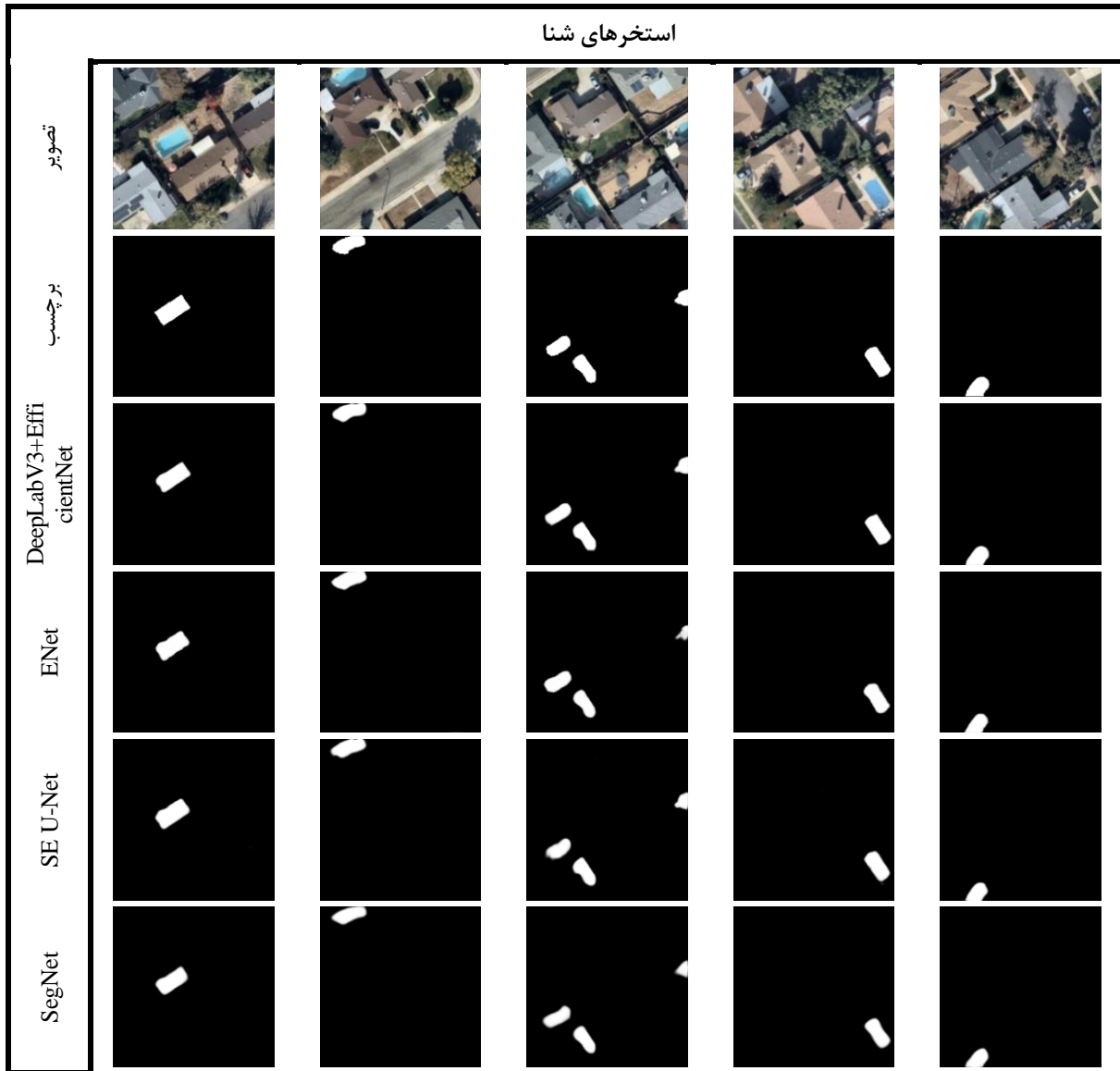
باتوجه به نتایج ارزیابی مدل‌ها برای تشخیص استخرهای شنا در تصاویر ماهواره‌ای مربوط به مجموعه عکس‌های کالیفرنیا مدل SegNet با دقت $۹۳/۸۱$ در معیار Precision، بهترین عملکرد را داشته و مدل DeepLabV3+EfficientNet در شاخص‌های IoU و Dice Coefficient به ترتیب با مقادیر $۸۵/۵۸$ و $۹۲/۷۷$ عملکرد بهتری نسبت به سایر مدل‌ها از خود نشان داده است. در شکل ۷، تصاویر خروجی بخش‌بندی معنایی تصاویر هر مدل به منظور مقایسه قرار داده شده است.

در پژوهش دوان و هو^۱ که به بخش‌بندی پهنه‌های آبی در تصاویر ماهواره‌ای با کیفیت بالا پرداختند، نتایج نشان داد که مدل DeepLabV3+ در شاخص‌های ارزیابی Precision، Dice Coefficient و IoU عملکرد بهتری نسبت به مدل‌های SegNet و U-Net داشت. ما نیز در تحقیق خود به این نتیجه رسیدیم که مدل DeepLabV3+EfficientNet، باتوجه به نتایج و ارزیابی‌های انجام‌شده، بهترین عملکرد را نسبت به دیگر مدل‌های استفاده‌شده در این پژوهش دارد. یکی از دلایل موفقیت این مدل، استفاده از ترکیب معماری DeepLabV3+ و EfficientNet است که موجب افزایش دقت و کارایی در تشخیص ویژگی‌های پیچیده و جزئیات دقیق در تصاویر ماهواره‌ای می‌شود [۲۷]. در پژوهشی دیگر توسط چن^۲ و همکاران به منظور مقایسه عملکرد مدل‌های مختلف در زمینه بخش‌بندی معنایی تصاویر ماهواره‌ای بر روی مجموعه داده‌های مختلف صورت گرفت، مشخص گردید مدل DeepLabV3+ نتایج بهتری نسبت به مدل SegNet کسب کرده است [۲۸]. در شکل ۸، نمونه‌هایی از تصاویر استخرهای شنا و خروجی شبکه‌های یادگیری عمیق در مواجهه با چالش‌های موجود در مجموعه داده نمایش داده شده است. با بررسی و مقایسه این تصاویر مشخص می‌شود که مدل DeepLabV3+EfficientNet در تشخیص جزئیات و غلبه بر چالش‌های مجموعه داده، عملکرد بهتری نسبت به سایر مدل‌ها دارد.

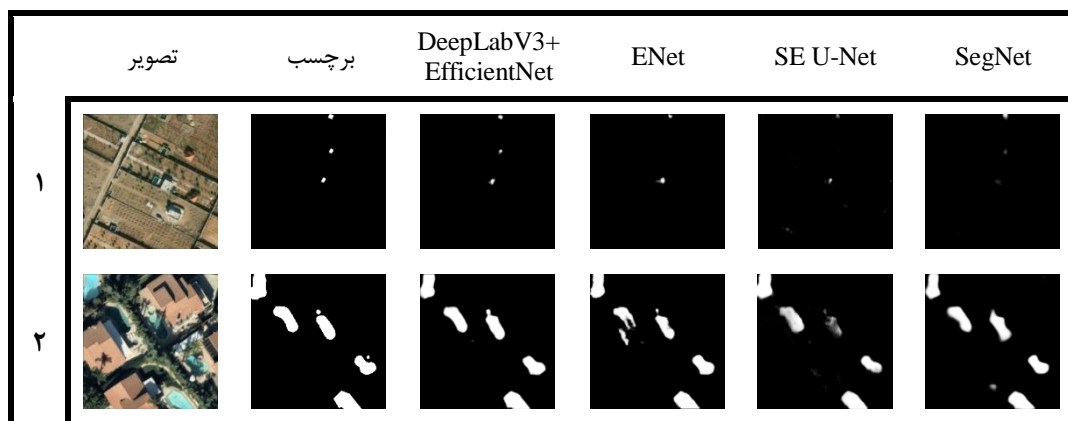
باین حال، محدودیت‌هایی نیز وجود دارد که در کاربردهای واقعی باید در نظر گرفته شوند. یکی از این محدودیت‌ها، عملکرد ضعیف مدل‌ها در شرایطی است که تصاویر شامل سایه‌ها، انعکاسات نوری یا موانع مشابه است. این عوامل می‌توانند باعث ایجاد نویز در تشخیص مرزهای استخرها شوند. به‌ویژه در تصاویر با جزئیات کمتر، مانند تصویر ۲، مدل‌ها در تشخیص دقیق استخرها دچار مشکل می‌شوند. همچنین مدل‌هایی مانند ENet و SegNet، با وجود ساده‌تر بودن ساختارشان نسبت به DeepLabV3+EfficientNet، در شناسایی برخی از جزئیات ظریف عملکرد ضعیف‌تری دارند. این امر می‌تواند به محدودیت‌های معماری و پیچیدگی کمتر این مدل‌ها نسبت داده شود [۲۸ و ۲۹].

جدول ۲: نتایج ارزیابی خروجی شبکه‌های یادگیری عمیق بر روی تصاویر پهنه آبی استخرهای شنا

نتایج بر روی تصاویر استخرهای شنا			نام شبکه
Precision	Dice Coefficient	IoU	
۹۳/۵۳	۹۲/۷۷	۸۵/۵۸	DeeplabV3+ EfficientNet
۹۱/۶۵	۹۰/۴۳	۸۰/۵۱	ENet
۹۳/۲۲	۹۰/۸۴	۸۳/۰۴	SE U-Net
۹۳/۸۱	۸۹/۸۹	۸۱/۳۳	SegNet



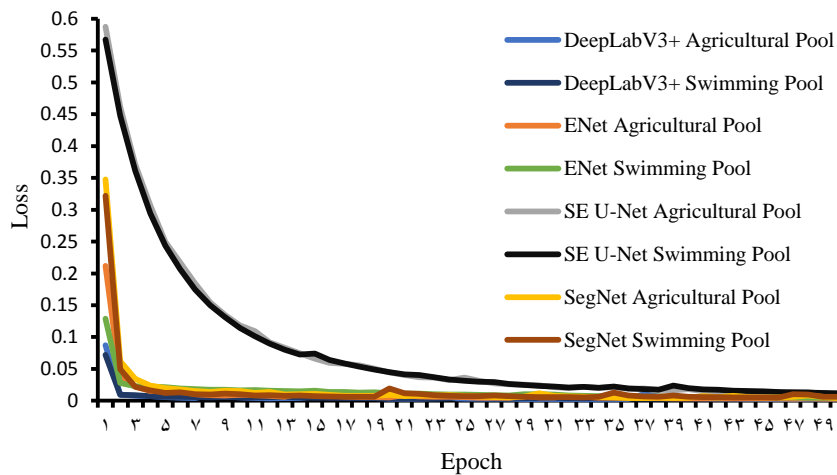
شکل ۷: نمونه تصاویر استخرهای شنا و خروجی از شبکه‌های یادگیری عمیق



شکل ۸: مقایسه نمونه تصاویر مدل‌ها در تشخیص جزئیات

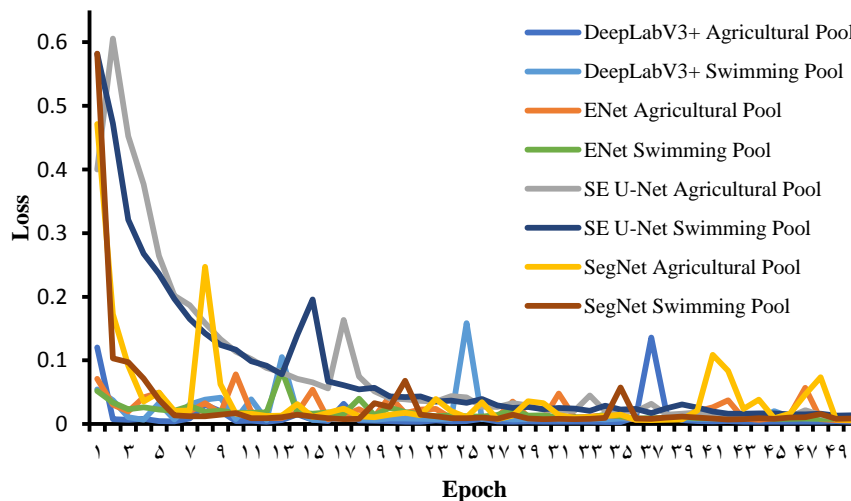
شکل ۹ تغییرات تابع اتلاف (Loss) را بر روی داده‌های آموزشی در طول ۵۰ دوره (Epoch) آموزش برای چند مدل مختلف یادگیری عمیق نشان می‌دهد که بر روی دو نوع مجموعه داده متفاوت (Swimming Pool, Agricultural Pool) آزمایش شده‌اند. محور افقی، تعداد دوره‌های آموزش و محور عمودی، مقدار تابع اتلاف را نمایش می‌دهد. در مراحل اولیه آموزش، تمامی مدل‌ها کاهش سریعی در مقدار اتلاف را نشان می‌دهند که نشان‌دهنده یادگیری سریع در این مراحل است. مدل‌های DeepLabV3+EfficientNet

در هر دو نوع داده (آبی روشن و آبی تیره) و SE U-Net (خاکستری روشن و خاکستری تیره) عملکرد بهتری نسبت به سایر مدل‌ها دارند. این امر نشان‌دهنده کارایی بالای این مدل‌ها در یادگیری و تطابق با داده‌هاست. پس از چند دوره، مقدار اتلاف در بیشتر مدل‌ها به حالت پایدار می‌رسد و تغییرات آن بسیار کم می‌شود که نشان‌دهنده رسیدن مدل به تعادل و اتمام فرایند یادگیری اصلی است. برخی مدل‌ها مانند ENet (نارنجی و سبز) و SegNet (زرد و قهوه‌ای) نوسانات کوچکی در مقدار اتلاف طی دوره‌های آموزش نشان می‌دهند که ممکن است به دلیل حساسیت بیشتر این مدل‌ها به تغییرات داده‌ها باشد. به‌طور کلی، مدل DeepLabV3+EfficientNet از نظر عملکرد، در هر دو نوع مجموعه داده بهبود قابل توجهی نسبت به مدل‌های دیگر از خود نشان می‌دهد. بنابراین این مدل می‌تواند گزینه مناسبی برای فعالیت‌های مرتبط با شناسایی پهنه‌های آبی در تصاویر در نظر گرفته شود (۲۷).



شکل ۹: منحنی رسم‌شده مربوط به تابع خطا مربوط به داده‌های آموزشی

شکل ۱۰ تغییرات تابع اتلاف (Loss) بر روی مجموعه داده اعتبارسنجی (Validation) را در طول دوره آموزش نشان می‌دهد که عملکرد مدل‌های مختلف در دو مجموعه داده کشاورزی و استخر شنا را مقایسه می‌کند. مدل DeepLabV3+EfficientNet در هر دو مجموعه داده بهترین عملکرد را با کمترین مقدار تابع اتلاف دارد. مدل‌های SE U-Net و SegNet نیز عملکرد خوب و مناسبی دارند؛ اگرچه نوسانات بیشتری را در طول آموزش تجربه می‌کنند. مدل ENet با وجود کاهش قابل توجه مقدار تابع اتلاف، نوسانات بیشتری را نشان می‌دهد. به‌طور کلی، تمامی مدل‌ها با کاهش مقدار تابع اتلاف به پایداری نسبی رسیده‌اند. با بررسی نمودار همچنین می‌توان دریافت به دلیل نبود افزایش مداوم و ناگهانی در تابع اتلاف اعتبارسنجی پس از کاهش اولیه نشان می‌دهد که مدل‌ها دچار بیش‌برازش نشده‌اند؛ به‌خصوص، مدل DeepLabV3+EfficientNet با کمترین نوسانات و رسیدن به تابع اتلاف نهایی پایین، تعمیم‌پذیری خوبی را بر روی داده‌های جدید نشان داده است. سایر مدل‌ها نیز با وجود برخی نوسانات، روند کاهشی و پایداری در تابع اتلاف را حفظ کرده‌اند که این امر حاکی از عدم بیش‌برازش مدل‌هاست.



شکل ۱۰: منحنی رسم‌شده مربوط به تابع خطا مربوط به داده‌های اعتبارسنجی

۴. نتیجه‌گیری

در تحقیق حاضر از دو مجموعه داده متفاوت استفاده شده که شامل تصاویر ماهواره‌ای از استخرهای شنای کالیفرنیا و استخرهای کشاورزی شهرستان کرج و شیراز است. انتخاب ۷۴۶ تصویر با ابعاد 256×256 با کیفیت و وضوح مناسب از این دو مجموعه داده، به منظور انجام تحلیل‌های مورد نیاز و تشخیص پهنه‌های آبی توسط مدل‌های هوش مصنوعی انجام شد. در فرایند آموزش مدل‌ها، ۷۵٪ از تصاویر برای آموزش و ۲۵٪ باقی‌مانده برای ارزیابی عملکرد مدل‌ها (آزمون) استفاده شده است. مجموعه تصاویر شامل نمونه‌های متنوعی از استخرهای تفریحی و کشاورزی، از جمله استخرهای شنای مناطق مسکونی و خصوصی و استخرهای بتنی و پلیمری کشاورزی است. برای انتخاب این مجموعه داده، از معیارهایی مانند جامع بودن نمونه‌های آموزشی، پوشش انواع مختلف استخرها، تنوع در شکل و ابعاد استخرها، کیفیت مناسب تصاویر ماهواره‌ای و تنوع اقلیمی به منظور تعمیم‌پذیری یافته‌ها استفاده شد. فرایند برچسب‌گذاری دستی تصاویر نیز باتوجه به وجود سایه درختان، ساختمان‌ها و ساختارهای نامنظم در تصاویر، نیازمند صرف زمان و دقت بسیاری بوده است.

ارزیابی‌های تجربی صورت گرفته در این پژوهش نشان می‌دهند که مدل DeepLabV3+EfficientNet برای تشخیص پهنه‌های آبی، با عبارت دیگر برای بازسازی استخرهای کشاورزی و شنا در تصاویر ماهواره‌ای، دارای بهترین دقت نسبت به سایر مدل‌ها بوده است. این مدل تا ۴ درصد بهبود عملکردی در معیار IoU، تا ۲ درصد بهبود در معیار Dice Coefficient و حدود ۱ درصد بهبود در معیار Precision داشته است. این نتایج نشان‌دهنده قابلیت بالای مدل DeepLabV3+EfficientNet در تشخیص دقیق‌تر و بازسازی بهتر پهنه‌های آبی در مقایسه با سایر مدل‌های مورد بررسی است. در زمان آموزش مدل‌های هوش مصنوعی با چالش‌هایی مانند دشواری تفکیک مرز آب و سازه استخر کشاورزی (پلیمری)، شباهت برخی از نواحی تحت کشت از نظر رنگ با این آب درون استخرها، رشد گیاهان و جلبک‌ها در سطح و همچنین سایه درختان، تشخیص و بازسازی را با خطا مواجه ساخت. این محدودیت‌ها می‌توانند به‌طور بالقوه بر نتایج نهایی تأثیر بگذارند. علاوه بر این، مدل‌های ساده‌تر مانند ENet و SegNet به دلیل ساختار ساده‌تر خود، در شناسایی جزئیات دقیق و تفکیک مرزهای ظریف، عملکرد ضعیف‌تری نسبت به مدل‌های پیچیده‌تر دارند. این مسئله می‌تواند به محدودیت‌های معماری این مدل‌ها نسبت داده شود که در شرایط پیچیده‌تر دچار مشکل می‌شوند. هرچند شایان ذکر است که این مدل‌ها به‌طور کلی از عملکردی قابل قبول برای تشخیص پهنه‌های آبی برخوردار بودند، برای مقابله با این چالش و محدودیت‌ها در هنگام آموزش شبکه از روش داده‌افزایی برای افزایش تنوع داده‌ها استفاده شد تا مدل‌ها نسبت به چالش‌های گفته‌شده مقاوم شوند. نتایج نشان داد که استفاده از روش داده‌افزایی باعث افزایش دقت مدل‌ها در شناسایی و بخش‌بندی استخرهای کشاورزی شد. به‌طور خاص، مدل‌ها توانستند با دقت بیشتری نواحی تحت کشت را که از نظر رنگ مشابه آب درون استخرها بودند، تشخیص دهند و همچنین تأثیر سایه درختان و رشد گیاهان و جلبک‌ها را به حداقل برسانند. این یافته‌ها نشان می‌دهد که داده‌افزایی به‌عنوان یک دستاورد مهم پژوهش، نقش بسزایی در بهبود عملکرد مدل‌ها در مواجهه با چالش‌های مذکور داشته است. یافته‌های این پژوهش می‌تواند به‌عنوان یک راهنما برای محققان و کارشناسان در زمینه سنجش از دور نقش ایفا کند، تا به آن‌ها کمک کند انتخابی مناسب برای معماری شبکه برای انجام بخش‌بندی معنایی داشته باشند. با این حال، برای ادامه و بهبود کار، پیشنهاد می‌شود که پژوهش‌های آینده می‌تواند بر گسترش و بهبود مجموعه داده‌ها تمرکز داشته باشد. افزایش حجم مجموعه داده‌ها می‌تواند با افزودن داده‌های متنوع، مانند تصاویر باند مادون قرمز برای تحلیل بهتر در شرایط نوری شبانه انجام شود. همچنین، استفاده از تصاویر هوایی و پهپادی به منظور افزایش وضوح و جزئیات تصاویر و تنوع بخشیدن به مجموعه داده‌ها پیشنهاد می‌شود. علاوه بر این، توجه به استفاده از روش‌های پیشرفته‌تر داده‌افزایی، ارزیابی مدل‌های جدیدتر و پیشرفته‌تر، و ترکیب چندین مدل یادگیری عمیق نیز ضروری است.

منابع

1. Dietz T, Shwom RL, Whitley CT. Climate change and society. *Annual Review of Sociology*. 2020; 46: 135-158.
2. Höök M, Tang X. Depletion of fossil fuels and anthropogenic climate change—A review. *Energy Policy*. 2013; 52: 797-809.
3. Loarie SR, Duffy PB, Hamilton H, Asner GP, Field CB, Ackerly DD. The velocity of climate change. *Nature*. 2009; 462(7276): 1052-1055.
4. Trenberth KE. Changes in precipitation with climate change. *Climate research*. 2011; 47(1-2): 123-138.
5. Calzadilla A, Rehdanz K, Betts R, Falloon P, Wiltshire A, Tol RS. Climate change impacts on global agriculture. *Climatic change*. 2013; 120: 357-374.
6. Taylor RG, Scanlon B, Döll P, Rodell M, Van Beek R, Wada Y, et al. Ground water and climate change. *Nature climate change*. 2013; 3(4): 322-329.

7. Wheeler T, Von Braun J. Climate change impacts on global food security. *Science*. 2013; 341(6145): 508-513.
8. Chartzoulakis K, Bertaki M. Sustainable water management in agriculture under climate change. *Agriculture and Agricultural Science Procedia*. 2015; 4: 88-98.
9. Rao CS, Rejani R, Rao CR, Rao K, Osman M, Reddy KS, et al. Farm ponds for climate-resilient rainfed agriculture. *Current Science*. 2017; 471-477.
10. Siebrits R. Swimming pools and intra-city climates: influences on residential water consumption in Cape Town. *Water SA*. 2012; 38(1): 133-144.
11. Chi M, Plaza A, Benediktsson JA, Sun Z, Shen J, Zhu Y. Big data for remote sensing: Challenges and opportunities. *Proceedings of the IEEE*. 2016; 104(11): 2207-2219.
12. Huang C, Chen Y, Zhang S, Wu J. Detecting, extracting, and monitoring surface water from space using optical sensors: A review. *Reviews of Geophysics*. 2018; 56(2): 333-360.
13. Isikdogan F, Bovik AC, Passalacqua P. Surface water mapping by deep learning. *IEEE journal of selected topics in applied earth observations and remote sensing*. 2017; 10(11): 4909-4918.
14. Cheng G, Xie X, Han J, Guo L, Xia G-S. Remote sensing image scene classification meets deep learning: Challenges, methods, benchmarks, and opportunities. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 2020; 13: 3735-3756.
15. Cai W, Wei Z. Remote sensing image classification based on a cross-attention mechanism and graph convolution. *IEEE Geoscience and Remote Sensing Letters*. 2020; 19:1-5.
16. Liu X, Song L, Liu S, Zhang Y. A review of deep-learning-based medical image segmentation methods. *Sustainability*. 2021; 13(3): 1224.
17. Edpuganti A, Akshaya P, Gouthami J, Sajith Variyar V, Sowmya V, Sivanpillai R. Effect of data quality on water body segmentation with deeplabv3+ algorithm. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. 2023; 48: 81-85.
18. Harika A, Sivanpillai R, Sajith Variyar V, Sowmya V. Extracting water bodies in rgb images using deeplabv3+ algorithm. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. 2022; 46: 97-101.
19. Li Z, Wang R, Zhang W, Hu F, Meng L. Multiscale features supported DeepLabV3+ optimization scheme for accurate water semantic segmentation. *Ieee Access*. 2019; 7: 155787-155804.
20. Weng L, Xu Y, Xia M, Zhang Y, Liu J, Xu Y. Water areas segmentation from remote sensing images using a separable residual segnet network. *ISPRS international journal of geo-information*. 2020; 9(4): 256.
21. Ma Z, Xia M, Weng L, Lin H. Local feature search network for building and water segmentation of remote sensing image. *Sustainability*. 2023; 15(4): 3034.
22. Wang Y. Remote sensing image semantic segmentation algorithm based on improved ENet network. *Scientific Programming*. 2021; 1-10.
23. Paszke A, Chaurasia A, Kim S, Culurciello E. *Enet: A deep neural network architecture for real-time semantic segmentation*. arXiv preprint arXiv:160602147. 2016.
24. Aburaed N, Al-Saad M, Alkhatib M, Zitouni M, Almansoori S, Al-Ahmad H. Semantic Segmentation of Remote Sensing Imagery Using AN Enhanced Encoder-Decoder Architecture. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. 2023; 10: 1015-1020.
25. Li W, Dong Y, Wang Y, Xu T, Liu Z, Yu K, Xiao C, editors. SE-UNet: Channel Attention Based UNet for Water Body Segmentation from SAR Image. *International Conference on Image, Vision and Intelligent Systems 2023 Aug 16 (pp 100-107)*; 2024; Singapore: Springer Nature Singapore.
26. Al-Saad M, Aburaed N, Alkhatib MQ, Zitouni MS, Al-Ahmad H, editors. An Enhanced UNet Model to Detect Water Bodies From Remote Sensing Data. *International Conference on Information Technology (ICIT) 2023 Aug 9 (pp 411-415) IEEE*; 2023: IEEE.
27. Duan L, Hu X. Multiscale refinement network for water-body segmentation in high-resolution satellite imagery. *IEEE Geoscience and Remote Sensing Letters*. 2019; 17(4): 686-690.
28. Chen J, Xia M, Wang D, Lin H. Double branch parallel network for segmentation of buildings and waters in remote sensing images. *Remote Sensing*. 2023; 15(6): 1536.
29. Pellis E, Masiero A, Cortesi I, Tucci G, Betti M, Grussenmeyer P, editors. Performance Comparison Between Segnet and DEEPLABV3+ on the Semantic Segmentation of Heritage Buildings. *12th International Symposium on Mobile Mapping Technology (MMT 2023)*, May 24-26, 2023, Padua, Italy 2023 May 25 (Vol 48, pp 379-386); 2023: TCopernicus GmbH.