

تشخیص لکه‌های نفتی در تصاویر ماهواره‌ای سنتینل-1 با استفاده از رویکرد یادگیری عمیق

سعید دهقانی ده‌چشمه¹، مهدی آخوندزاده هنزایی^{2*}

¹ دانشجوی کارشناسی ارشد، دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی، پردیس دانشکده‌های فنی، دانشگاه تهران
saeid.dehghani@ut.ac.ir

² دانشیار دانشکده مهندسی نقشه‌برداری و اطلاعات مکانی، پردیس دانشکده‌های فنی، دانشگاه تهران
makhonz@ut.ac.ir

(تاریخ دریافت: مرداد 1400، تاریخ تصویب: مرداد 1401)

چکیده

آگاهی از مسائل حوزه دریایی برای مدیریت بحران در صورت بروز حوادث بسیار مهم است. نشت نفت یکی از تهدیدهای اصلی برای محیط‌های دریایی و ساحلی محسوب می‌شود و به طور جدی بر اکوسیستم دریایی تأثیر می‌گذارد و نگرانی‌های سیاسی و زیست‌محیطی ایجاد می‌کند، زیرا اکوسیستم شکننده دریایی و ساحلی را به طور جدی تحت تأثیر قرار می‌دهد. میزان تخلیه آلاینده‌ها و اثرات مرتبط با آن بر محیط دریایی، پارامترهای مهمی در ارزیابی کیفیت آب دریاها هستند. نظارت موثر، شناسایی زودهنگام و برآورد حجم این لکه‌های نفتی، اولین و مهمترین مرحله برای یک عملیات پاکسازی موفق است. سنسورهای رادار دریچه مصنوعی (SAR) به دلیل قابلیت عملکرد موثر بدون توجه به وضعیت آب و هوا و شرایط روشنایی محیط و برداشت منطقه وسیعی از زمین، انتخاب بسیار مناسبی برای این منظور هستند. لکه‌های سیاه مربوط به نشت نفت را می‌توان به وضوح توسط سنسورهای SAR ثبت کرد، با این حال، تمایز آن‌ها از نظر ظاهری یک هدف چالش برانگیز است. در این پژوهش، از تصاویر راداری ماهواره سنتینل-1 برای شناسایی نشت نفت استفاده شده است. این مطالعه یک چارچوب یادگیری عمیق برای شناسایی لکه‌های نفتی بر اساس یک مجموعه داده بسیار وسیع از نقاط مختلف دنیا ارائه داده و با استفاده از ساختار شبکه‌های پیچشی *U-Net*، *DeepLabV3+* و *Fc-DenseNet* طبقه‌بندی تصاویر را به دو کلاس انجام می‌دهد. در این پژوهش، با تغییر تابع ضرر و حذف تصاویر تک‌کلاسه نتایج بسیار بهتری نسبت به کارهای مشابه قبلی حاصل شد. به‌طوری که نتایج IoU برای مدل‌های *U-Net*، *DeepLabV3+* و *FC-DenseNet* به ترتیب برابر 0.547، 0.613 و 0.545 بدست آمد.

واژگان کلیدی: لکه‌های نفتی، شبکه عصبی کانولوشن، ماهواره سنتینل-1، *U-Net*، *DeepLabV3+*، *Fc-DenseNet*

1- مقدمه

نشت نفت یک تهدید زیست‌محیطی است که اثرات نامطلوبی بر آب، زمین و هوا دارد [1] و می‌تواند باعث آلودگی دریا شده و به گونه‌های پرندگان، ماهی‌ها و دیگر موجودات آبرزی آسیب برساند. لکه‌های نفتی عمدتاً در اثر تصادفات مربوط به نفتکش‌ها، کشتی‌ها و خطوط لوله‌ای که نفت خام، بنزین، سوخت و فرآورده‌های جانبی نفت را جابه‌جا می‌کنند، ایجاد می‌شوند. با این حال، تخلیه غیرقانونی نفت و نفت باقی‌مانده از تمیز کردن مخازن نفتکش‌ها و کشتی‌ها منابع اصلی وقوع آلودگی نفتی محسوب می‌شوند [2و3]. حذف لکه‌های نفتی برای حفظ محیط امن و تمیز و حفاظت از آبزیان بسیار مهم است. در دهه گذشته حمل و نقل دریایی به طور مداوم در حال رشد بوده و با افزایش تعداد کشتی‌ها تعداد تخلیه غیرقانونی نفت هم افزایش پیدا کرده‌است. از طرفی هم تانکرهای نفت و هم انواع دیگر کشتی‌های نفتی از جمله مظنونین متخلف تخلیه غیرقانونی نفت در دریا هستند. ابزارهای مختلفی از جمله شناورها، هواپیماها و ماهواره‌ها برای شناسایی و نظارت بر نشت نفت استفاده می‌شوند. شناورها، به ویژه اگر به رادارهای تخصصی مجهز باشند، می‌توانند نفت را در دریا تشخیص دهند اما منطقه بسیار محدودی را پوشش می‌دهند. سیستم‌های اصلی برای کنترل آلودگی نفتی در دریا، استفاده از هواپیما و ماهواره‌های مجهز به رادار دیافراگم مصنوعی (SAR)¹ است [4و6]. SAR یک سنسور مایکروویو فعال است که تصاویر دو بعدی را ضبط می‌کند. سنسورهای SAR فضایی به طور گسترده برای تشخیص نشت نفت در محیط دریایی مورد استفاده قرار می‌گیرند؛ زیرا از نور خورشید مستقل هستند، تحت تأثیر هوای ابری قرار نمی‌گیرند، مناطق زیادی را تحت پوشش قرار می‌دهند و مقرون به صرفه‌تر از گشت هوایی هستند. میزان روشنایی تصویر گرفته‌شده، بازتابی از خصوصیات سطح هدف است. تشخیص نشت نفت در یک تصویر SAR متکی به این واقعیت است که لکه نفتی باعث کاهش موج بازگشتی از سطح دریا و در نتیجه تشکیل لکه تاریک در سطح تصویر می‌شود که با دریای بدون نشت نفت در اطراف لکه در تضاد است [3و4]. مقادیر Backscatter رادار از نشت نفت بسیار شبیه به مقادیر Backscatter از

مناطق بسیار آرام دریا و سایر پدیده‌های اقیانوس به نام "شبیه به هم"² (به عنوان مثال جریان‌ها یا گرداب‌ها) می‌باشد. پدیده‌های طبیعی اقیانوسی مانند مناطق با سرعت باد کم³، بسترهای علف‌های هرز⁴ و شکوفه جلبک‌ها⁵، سایه موج در پشت خشکی⁶ و غیره [6و7] نیز می‌توانند به عنوان نقاط تاریک به تصویر کشیده شوند. این مناطق تاریک اغلب به عنوان شبیه به هم دسته‌بندی می‌شوند و مشکل تشخیص نشت نفت را حتی چالش برانگیزتر می‌کند. روش‌های طبقه‌بندی متعددی برای تشخیص نشت نفت از تصاویر SAR ارائه شده‌است. در بیشتر موارد، یک فرایند خاص دنبال می‌شود که شامل سه مرحله اصلی است:

(الف) تشخیص خودکار نقاط تیره در تصویر SAR پردازش‌شده.

(ب) استخراج ویژگی از مناطق شناسایی‌شده در مرحله قبل.

(ج) طبقه‌بندی نقاط تاریک به عنوان لکه نفتی یا شبیه‌به‌هم.

در اولین مرحله، معمولاً تقسیم‌بندی باینری برای نمایش تصویر ورودی به منظور بازیابی لکه‌های سیاه اعمال می‌شود. مرحله دوم شامل استخراج ویژگی‌های آماری از مناطق تقسیم‌شده فوق است که ممکن است نشت احتمالی نفت را شامل شود. برای آخرین مرحله پردازش، کل تصویر یا بخش‌هایی که حاوی مناطق سیاه هستند به عنوان نشت نفت یا شبیه به هم طبقه‌بندی می‌شوند. افراد مختلف روش‌های متفاوتی برای شناسایی و طبقه‌بندی لکه‌های نفتی از اشکال شبیه‌به‌هم پیشنهاد دادند که در ادامه به چند مورد از آن‌ها اشاره می‌شود.

لکه‌های تاریک در تصاویر SAR را می‌توان بر اساس نزدیکی به سکو یا کشتی به‌عنوان لکه نفتی طبقه‌بندی کرد [9]. سولبرگ و همکاران [10] یک چارچوب خودکار را پیشنهاد دادند که لکه‌های تاریک را به سه دسته لکه‌هایی با احتمال زیاد، احتمال متوسط و کم نفت تقسیم کردند. با این وجود، دانش قبلی در مورد وجود لکه‌های نفت در اطراف کشتی‌ها و سکوها نفتی باید در مدل گنجانده شود. علاوه بر این، دانش در مورد شرایط خارجی مانند سرعت باد هم تأثیر زیادی بر روی دقت روش ارائه شده دارد.

² Look-alike

³ Low wind speed regions

⁴ Weed beds

⁵ Algae blooms

⁶ Wave shadows behind land

¹ Synthetic Aperture Radar

روش دیگر برای طبقه‌بندی، بررسی ویژگی‌های رادیومتریکی و هندسی مناطق مورد نظر می‌باشد و مانند روش قبل داشتن اطلاعات درمورد سرعت باد به افزایش دقت طبقه‌بندی کمک می‌کند [11]. در [12]، برای تبعیض بین نفت و اشکال شبیه‌به‌هم از تجزیه و تحلیل شکل، اندازه، شیب پرتوی ارسالی از ماهواره، بافت، سرعت باد لحظه‌ای، جریان آب، موقعیت سکو و خط کشتی استفاده شده‌است، با این حال، داده‌هایی که تغییرات زمانی سرعت باد را در طول روزهای قبل از تصویربرداری SAR نشان می‌دهند نیز ممکن است به طور قابل توجهی در فرایند تفسیر کمک کنند. همچنین، شکل هندسی این لکه‌ها با تاریخ باد مطابقت دارد و با استفاده از آن می‌توان سن لکه را برآورد کرد. می‌توان یک روش احتمالی برای تشخیص نشت نفت از سایر ویژگی‌های اقیانوسی (شبیه‌به‌هم) در تصاویر SAR توسعه داد. در این روش از اطلاعات آماری بدست‌آمده از اندازه‌گیری‌های قبلی مثل مشخصات فیزیکی و هندسی برای تشخیص نشت نفت و ویژگی‌های طبیعی استفاده می‌شود، برای هر تصویر نمونه با استفاده از دو روش مختلف احتمال نشت نفت و شبیه‌به‌هم مورد ارزیابی قرار می‌گیرد، سپس، نتایج دو روش مقایسه می‌شود [11]. از سیستم‌های هوشمند هم می‌توان برای تشخیص لکه‌های نفتی استفاده کرد [13]. این سیستم‌ها تنوع زیادی دارند، از جمله سیستم‌های هوشمند می‌توان به منطق فازی اشاره کرد که از روش‌های هوشمند هوش مصنوعی می‌باشد که برای طبقه‌بندی لکه‌های نفتی هم استفاده می‌شود [14]. در این روش نشت نفت توسط کارشناسان به عنوان الگوهای تیره با اشکال مشخص، در یک زمینه خاص مورد بررسی قرار می‌گیرند، این سیستم تصاویر ماهواره‌ای را تجزیه و تحلیل کرده و احتمال این که تصویر تیره نشت نفت باشد را تعیین می‌کند؛ خروجی شامل چندین تصویر و جدول است که اطلاعات لازم را برای تصمیم‌گیری در اختیار کاربر قرار می‌دهد. یک رویکرد شی‌گرا در ابتدا در [15] معرفی شد و از طبقه‌بندی‌کننده فازی برای برچسب‌زدن نقاط تیره استفاده شد و لکه‌های نفت و شبیه‌به‌هم را به دو کلاس مجزا تقسیم کردند. از مازول درخت تصمیم‌گیری هم می‌توان برای طبقه‌بندی لکه‌های تاریک به چهار دسته: شبیه به هم، نقاط مشکوک، نقاط تار و نشت احتمالی نفت استفاده کرد [16]. یکی از تکنیک‌های جدید برای طبقه‌بندی نشت نفت استفاده از جنگلی از درختان

تصمیم‌گیری می‌باشد که نشت نفت و شبیه‌به‌هم در دو دسته تقسیم‌بندی می‌شوند [17]. در [18]، یک روش تشخیص نیمه‌نظارت‌شده با استفاده از تجزیه موجک بر روی تصاویر SAR ارائه شده‌است. استفاده از شبکه‌های عصبی مصنوعی هم برای تشخیص آلودگی‌های نفتی در دریا بسیار مناسب است و دقت قابل قبولی دارد [19]. در اولین شبکه‌های عصبی، ویژگی‌های از پیش تعیین‌شده از تصاویر SAR به منظور برآورد برچسب‌های مربوطه، به شبکه داده می‌شد [1]. در روشی جدید، برای شناسایی نشت نفت از ویژگی‌های دست‌ساز برای آموزش شبکه عصبی موجک استفاده شد و لکه‌های سیاه گرفته‌شده در تصاویر SAR را به نشت نفت و مناطق آب دست نخورده طبقه‌بندی کردند [20]. به طور مشابه در تحقیقی دیگر از یک شبکه‌ی عصبی و الگوریتم ژنتیک، به منظور تعیین یک شبکه عصبی تقریباً مطلوب برای طبقه‌بندی تشکیلات تاریک به نشت نفت یا شبیه‌به‌هم به استفاده شد [21] که هدف آن بهینه‌سازی برای به حداقل رساندن ویژگی‌های ورودی مورد استفاده و در عین حال حداکثر کردن دقت طبقه‌بندی داده‌های مورد نظر است.

نوع پیشرفته‌تری از شبکه‌های عصبی، شبکه‌های عصبی کانولوشن¹ می‌باشد که این شبکه‌ها برای مطالعه و بررسی داده‌های حجیم نظیر تصاویر با بعد زیاد و سری‌های زمانی استفاده می‌شوند [22 و 23]. همچنین، این شبکه‌ها توانایی استخراج ویژگی‌های سطح بالاتری را نسبت به روش‌های قبل دارند [24 و 25]. در شبکه ارائه شده در [26]، داده‌ها به پنج کلاس زمین، آب، کشتی، شبیه‌به‌هم و نفت تقسیم‌بندی شده‌است. در این مقاله از مجموعه داده‌های حاصل از ماهواره‌ی سنتینل-1 که توسط آژانس ایمنی دریانوردی اروپا EMSA² تهیه شده‌است، استفاده می‌شود [27].

هدف این مقاله شناسایی لکه‌های نفتی از تصاویر از پیش‌پردازش‌شده سنجنده سنتینل-1 به منظور افزایش دقت طبقه‌بندی نسبت به کارهای مشابه گذشته می‌باشد. ابتدا با استفاده از داده‌های ورودی که متشکل از 2172 تصویر می‌باشد، شبکه آموزش داده می‌شود و در انتها دقت شبکه طراحی‌شده با استفاده از 238 تصویر مجزا مورد ارزیابی قرار می‌گیرد و با کارهای مشابه مقایسه می‌شود.

1 Convolution neural network

2 European Maritime Safety Agency

در بخش دو این مقاله ابتدا داده‌ها معرفی می‌شوند، سپس چند شبکه یادگیری عمیق طراحی و قسمت‌های مختلف این شبکه‌ها بررسی می‌شود و در ادامه پارامترهای مختلف ارزیابی شبکه معرفی و شرح داده می‌شوند. قسمت‌های بعد این مقاله به شرح زیر تنظیم شده‌است. در بخش 3 نتایج بدست‌آمده مورد ارزیابی قرار گرفته و مورد بحث قرار می‌گیرند و در نهایت، نتیجه‌گیری در بخش 4 ارائه می‌شود.

2- مبانی نظری

در این قسمت، ابتدا درمورد مجموعه داده‌های مورد استفاده در این تحقیق بحث می‌کنیم. سپس، شبکه‌های عصبی کانولوشن بصورت مختصر معرفی می‌شوند و در انتها روش‌های مورد استفاده در این تحقیق را بطور کامل شرح می‌دهیم.

2-1- مجموعه داده‌ها

مجموعه داده مورد استفاده در این مطالعه توسط ماهواره Sentinel-1 به دست آمده و توسط آژانس فضایی اروپا¹ (ESA) از طریق مرکز دسترسی باز Copernicus² توزیع شده‌است [26 و 27]، این داده‌ها حوادث آلودگی نفتی از 28 سپتامبر 2015 تا 31 اکتبر 2017 را پوشش می‌دهند. ماهواره به کار گرفته شده برای مأموریت Sentinel-1 مجهز به سیستم SAR است که در محدوده C-band تصویر برداری می‌کند. محدوده زمینی سنسور SAR تقریباً 250 کیلومتر با دقت مکانی 10×10 متر است. این مشخصات نشان می‌دهد که سنسور SAR می‌تواند مناطق وسیعی از زمین را پوشش دهد و همچنین، عوارض کوچک (نسبتاً کوچک) مانند کشتی‌ها را شناسایی کند، همچنین، پلاریزاسیون این تصاویر دوگانه می‌باشد (پلاریزاسیون VV و VH)؛ برای مثال، در پلاریزاسیون (VV) موج عمودی ارسال می‌شود - عمودی دریافت می‌شود و در پلاریزاسیون (VH)، موج عمودی ارسال می‌شود - افقی دریافت می‌شود. پلاریزاسیون تصاویر مورد استفاده (VV) می‌باشد و همچنین، پیش پردازش‌های لازم روی آن‌ها پیاده‌سازی شده که شامل مراحل زیر می‌باشد:

1- شناسایی هرگونه نشت نفت توسط سازمان EMSA

2- مناطقی که شامل نشت نفت یا پدیده شبیه‌به‌هم بوده از تصاویر خام بریده شده و به ابعاد 650×1250 درآورده شدند.
3- کالیبراسیون رادیومتریکی تصاویر بدست‌آمده.
4- استفاده از یک اسپکل³ فیلتر 7×7 برای از بین بردن نویز تصاویر.
5- اعمال یک تبدیل خطی برای تبدیل مقادیر db به مقادیر واقعی درخشندگی⁴.

بعد از مراحل بالا، 1112 تصویر از مجموعه داده‌های خام SAR بدست آمد، پس از بررسی‌های سازمان EMSA برای تمامی تصاویر ماسک واقعیت‌زمینی تهیه شده که هم بصورت RGB می‌باشد که هر کلاس را با یک رنگ متمایز نشان می‌دهد و هم بصورت تک‌باند برای استفاده در مدل‌های شبکه عصبی ارائه شده‌است. همچنین، ماسک حقیقت زمینی این تصاویر شامل 5 کلاس آب، نشت نفت، شبیه‌به‌هم، کشتی و سطح زمین است.

در شکل (1) یک نمونه از تصاویر SAR با ابعاد 650×1250 همراه با واقعیت زمینی این تصویر نمایش داده شده‌است، نقاط روشن در تصویر SAR کشتی می‌باشند و نقاط تاریک لکه نفتی یا شبیه‌به‌هم هستند و همچنین، قسمت‌های روشن‌تر تصویر سطح زمین می‌باشند. در این شکل بطور واضح چالش بر انگیز بودن مجموعه داده‌ها مشخص می‌باشد و شباهت زیاد لکه‌های نفتی با اجزای شبیه به هم مشخص می‌باشد.

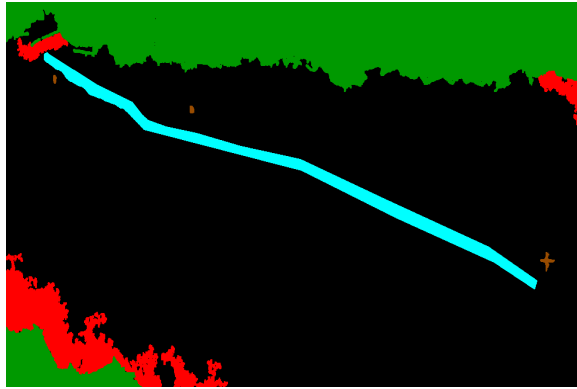
البته این تنها چالش موجود نبوده و چالش بزرگ‌تر پایین‌بودن نسبت پیکسل‌های نفت نسبت به کل پیکسل‌ها می‌باشد که این نسبت برابر با یک درصد بوده و برای حل این مشکل، بعد از تبدیل هر تصویر به هشت تصویر کوچک‌تر، تصاویری که فاقد لکه نفتی بودند حذف شده و این نسبت به سه درصد افزایش پیدا کرد و این موضوع در ادامه بصورت کامل‌تری شرح داده می‌شود. همانطور که قبلاً اشاره شد، ابعاد تصاویر 650×1250 می‌باشد و تصاویر واقعیت زمینی متشکل از پنج کلاس هستند، ولی در این تحقیق هدف تنها تشخیص لکه‌های نفتی می‌باشد، به همین دلیل تصاویر واقعیت زمینی به دو کلاس نفت و غیر نفت تغییر داده شد تا از محاسبات اضافی جلوگیری شود و تمرکز اصلی روی تشخیص نفت باشد.

3 Speckle filter

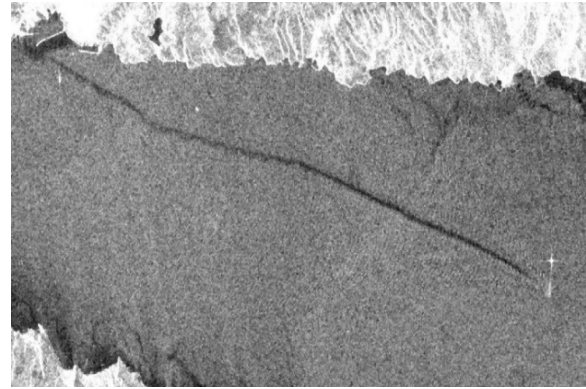
4 Real luminosity

1 European Space Agency

2 Copernicus Open Access Hub



(ب)



(الف)

شکل 1- (الف) نمونه تصویر SAR و (ب) تصویر ماسک واقعیت زمینی مربوطه. رنگ فیروزه‌ای مربوط به نشت نفت، قرمز شبیه به هم، قهوه‌ای کشتی، سبز خشکی و سیاه برای سطح دریا است.

در جدول (1)، تعداد پیکسل‌های کلاس هدف و پس‌زمینه در تصاویر آموزشی و تست آورده شده است که با توجه به تعداد کم پیکسل‌های کلاس هدف، چالش برانگیز بودن کار پیش رو کاملاً مشخص است.

برای توضیح بیشتر پیچیدگی مشکل تشخیص نشت نفت، تعدادی از نمونه‌های تصویر SAR از مجموعه داده، همراه با ماسک حقیقت زمینی مربوطه در شکل (2) ارائه شده است. این مثال‌ها را می‌توان به عنوان بدترین سناریوها در روش تشخیص در نظر گرفت و می‌توان دید که تمایز نشت نفت و شباهت ظاهری کار دشواری است. لکه‌های سیاهی که مناطق وسیعی از تصاویر SAR را پوشش می‌دهند معمولاً شبیه به هم هستند؛ از طرفی، لکه‌های تیره طولانی معمولاً نشت نفت هستند، با این حال این موضوع همیشه صادق نیست. علاوه بر این، تشخیص نشت نفتی که یک منطقه وسیع از تصویر SAR را شامل می‌شود هم یک وضعیت بسیار چالش برانگیز است؛ زیرا شکل لکه نفت مبهم است. با توجه به این که اولین لایه‌های شبکه‌ی عصبی پیچشی عمیق $DCNN^1$ ویژگی‌های سطح پایین مانند لبه‌ها، گوشه‌ها، بافت و غیره را استخراج می‌کند، عوارضی با مرزهای تار را نمی‌تواند به طور دقیق شناسایی کند، البته در لایه‌های بعدی شبکه‌های $DCNN$ چون ویژگی‌های سطح بالاتری مثل اشکال هندسی پیچیده قابل استخراج است، این ویژگی‌ها تا حدودی قابل شناسایی می‌باشند.

همچنین، برای آموزش شبکه‌های مختلف نیاز به تغییر اندازه تصاویر با توجه به شبکه مورد نظر می‌باشد و در این تحقیق سعی شد ابعاد تصاویر برای شبکه‌های مختلف ثابت در نظر گرفته شود، به همین خاطر ابعاد نهایی تصاویر برابر با 320×320 در نظر گرفته شد. هدف تیم، استفاده حداکثری از کل تصاویر در آموزش و تست شبکه بوده است، به همین دلیل هر تصویر 650×1250 را به هشت تصویر با ابعاد 320×320 تبدیل شد که هر تصویر با تصویر کناری 10 پیکسل مشترک دارد و همین امر سبب استحکام بیشتر شبکه شده است، همچنین، ده پیکسل بالا و پایین تصاویر اصلی برای رسیدن به ابعاد مورد نظر حذف شد و در نهایت، 1112 تصویر اصلی تبدیل به 8896 تصویر با ابعاد کوچکتر شد. در مرحله بعد، تصاویری که شامل حداقل یک پیکسل نفت بودند، برای آموزش و تست شبکه انتخاب شدند و بقیه تصاویر از روند آموزش و تست خارج شدند. در نهایت، تعداد تصاویر آموزشی 2172 تصویر و تعداد تصاویر تست برابر 238 تصویر شد.

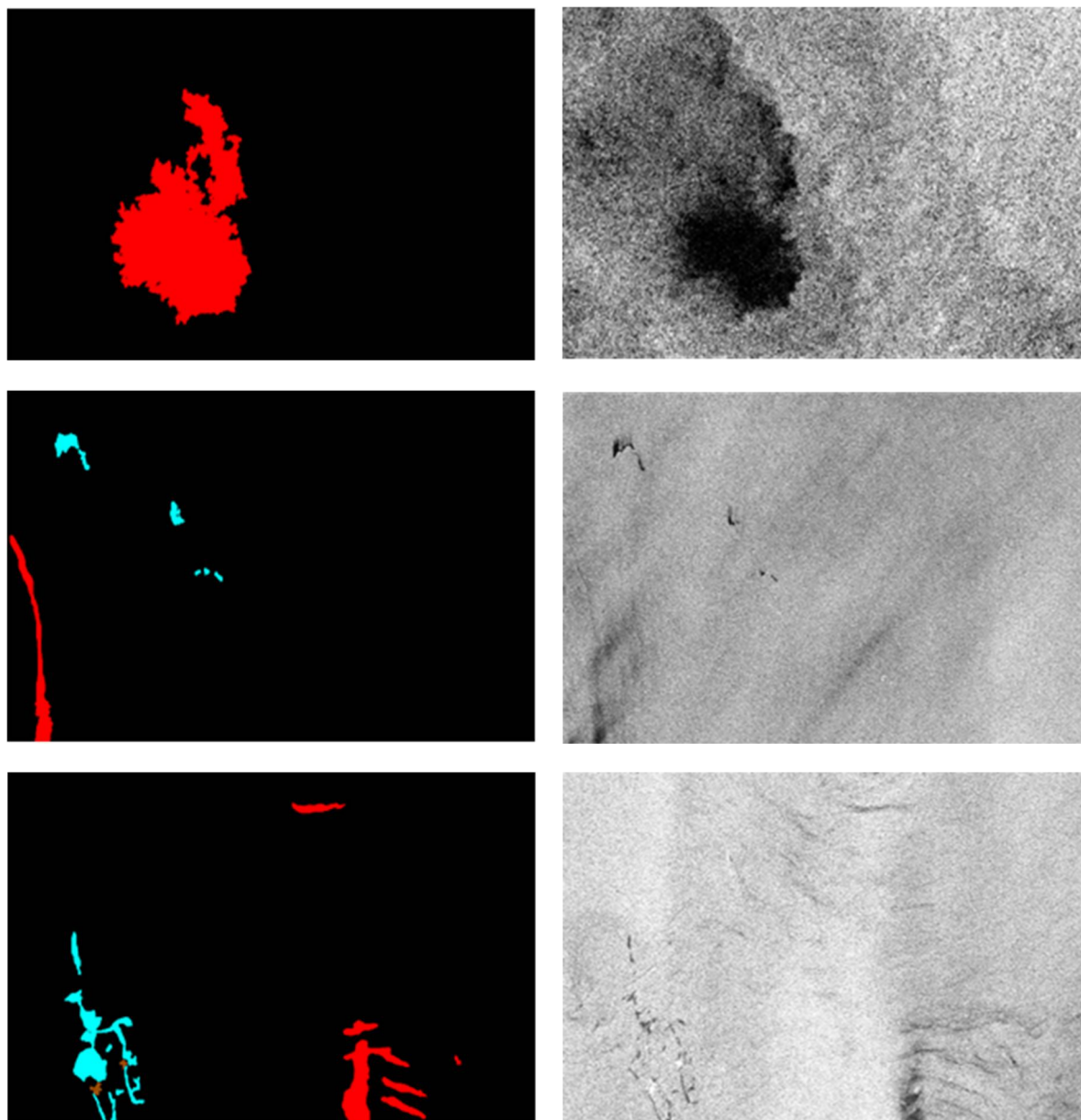
جدول 1- تعداد کل پیکسل‌ها و تعداد پیکسل‌های کلاس هدف (نشت نفت)

تعداد کل پیکسل‌ها		تعداد کل پیکسل‌های کلاس هدف	
آموزش	تست	آموزش	تست
222412,800	24,371,200	8,416,436	727,662

1 Deep Convolutional Neural Network

همانطور که در شکل (2) مشخص می‌باشد، نقاط تاریک در تصاویر SAR بسیار زیاد می‌باشند و باید لکه‌های نفت از این

لکه‌های تیره تفکیک شود تا بتوان یک سیستم هشدار دقیق برای شناسایی زود هنگام لکه‌های نفت طراحی شود.



(ب) واقعیت زمینی

(الف) تصویر SAR

شکل 2- نمونه‌هایی از تصویر SAR و ماسک واقعیت زمینی آن‌ها

2-2- شبکه‌های عصبی کانولوشن

پایه شبکه‌های عصبی مصنوعی، شبکه‌های پرسپترون می‌باشند که جزء شبکه‌های روبه‌جلو به‌شمار می‌آیند. شبکه‌های عصبی کانولوشن نوع خاصی از شبکه‌های عصبی

عمیق پیشرو¹، هستند یادگیری عمیق زیرمجموعه روش‌های یادگیری ماشین محسوب می‌شود و یکی از ویژگی‌های برجسته و متمایزکننده روش‌های یادگیری عمیق، ساختار سلسله‌مراتبی و چند لایه آن است. تفاوت ساختار شبکه‌های عصبی کانولوشن نسبت به شبکه‌های عصبی مصنوعی،

¹ Deep Feedforward Networks

$$\text{ReLU: max}(X, 0) \quad (2)$$

در رابطه بالا X ورودی تابع ReLU می باشد.

لایه Pooling یکی دیگر از لایه های شبکه های عصبی کانولوشن می باشد و هدف اصلی این لایه کوچک کردن تصویر ورودی به منظور کاهش بار محاسباتی، حافظه و تعداد پارامترها (کم کردن ریسک overfitting) است، از طرفی کاهش اندازه تصویر ورودی باعث کاهش حساسیت شبکه عصبی نسبت به جابجایی تصویر (مستقل از موقعیت) می شود. نورون های لایه پولینگ وزن ندارند، در واقع آن ها عملیات ادغام را با استفاده از یک تابع ادغام همانند ماکزیمم یا میانگین انجام می دهند. برای مثال، در تابع ادغام ماکزیمم، نقشه های ویژگی تولید شده در لایه ی قبل به عنوان ورودی وارد لایه پولینگ می شوند و در یک همسایگی مشخص بیشترین مقدار ورودی ها به عنوان ویژگی برجسته انتخاب شده و به عنوان خروجی آن همسایگی معرفی می گردد و این روند برای کل نقشه ویژگی ادامه می یابد تا یک نقشه ویژگی جدید تولید شود.

بعد از استفاده پی در پی از چند لایه کانولوشن، تابع فعال سازی و تابع پولینگ از یک لایه تماماً متصل (FC) استفاده می شود. همانطور که از نام این لایه پیدا است، نورون های موجود در این لایه به تمامی نورون های موجود در لایه ماقبل خود متصل است؛ به همین دلیل، تعداد پارامترهای قابل تنظیم به طور قابل توجهی بیشتر از سایر لایه ها خواهد بود. پس از لایه های تماماً متصل نیز به صورت معمول از یک تابع فعال ساز استفاده می گردد که اغلب از تابع فعال سازی ReLU استفاده می شود. در لایه ی تماماً متصل انتهایی، ابعاد خروجی به تعداد کلاس های موجود در تصویر خواهد بود و به هر کلاس یک امتیاز تعلق می گیرد که با بررسی این امتیازها کلاس برنده برای نورون مورد نظر مشخص می گردد. در لایه تماماً متصلی که در انتهای شبکه قرار دارد استثنا در نظر گرفته می شود و معمولاً از تابع فعال سازی Softmax استفاده می شود. خروجی این تابع فعال سازی درجه عضویت هر کلاس را برای نورون مورد نظر مشخص می نماید و به صورت زیر تعریف می شود.

$$P(y = j | X) = \frac{e^{X^T W_j}}{\sum_{k=1}^K e^{X^T W_k}} \quad (3)$$

محدودیت های در نظر گرفته شده در ارتباط بین نورون ها در شبکه های کانولوشنی می باشد. این محدودیت ها سبب کاهش عمومیت شبکه های کانولوشن در مواجهه با داده های متنوع شده است، اما از طرفی دیگر ساختار این شبکه ها با توجه به قیود خاص آن ها، کاملاً منطبق و مناسب با ساختار تصویر است. با توجه به محدودیت های اعمال شده در ساختار شبکه های کانولوشن، میزان محاسبات برای حل مسئله طبقه بندی نسبت به شبکه های عصبی به میزان قابل توجهی کاهش یافته و این مسئله سبب شده تا بتوان لایه های بیشتری نسبت به شبکه های عصبی طراحی و پیاده سازی کرد. شبکه های کانولوشن از سه لایه اصلی تشکیل شده اند که وظیفه اصلی این لایه ها استخراج سلسله ای از ویژگی های غیر خطی از تصاویر می باشد. سه نوع لایه مورد استفاده در ساختار CNN عبارتند از: لایه کانولوشن (Conv)، لایه ادغام¹ و لایه کاملاً متصل (FC²)، با قرار دادن این لایه ها به صورت متوالی، ساختار شبکه کانولوشنی ایجاد می شود. در انتهای شبکه، معمولاً لایه ای تحت عنوان Softmax قرار دارد که تصاویر را دسته بندی کرده و برچسب تصاویر را مشخص می کند.

در شبکه های کانولوشنی مشخصه ی اتصال محلی در نظر گرفته شده است و هر نورون به یک محدوده مکانی خاص از ورودی خود متصل است که این امر باعث کاهش تعداد متغیرهای قابل تنظیم در شبکه می گردد و همچنین، میزان محاسبات هم کاهش می یابد. خروجی لایه کانولوشن اعمال شده بر نقشه های ویژگی، از یک تابع فعال سازی غیر خطی عبور می کند. خروجی α_{ij} برای هر موقعیت (i,j) در یک لایه کانولوشن در تصویر از رابطه زیر به دست می آید.

$$\alpha_{ij} = \sigma((W * X)_{ij} + b) \quad (1)$$

در رابطه بالا بردار W بردار وزن، X بردار ورودی به هر نورون، b بایاس و σ تابع فعال سازی برای نورون می باشد. هدف فرایند آموزش، تعیین وزن و بایاس برای هر نورون می باشد. در شبکه های کانولوشن معمولاً از تابع غیر خطی ReLU^3 برای تابع فعال سازی استفاده می شود. مزیت استفاده از ReLU این است که برای همه ی ورودی های بزرگتر از صفر، یک مشتق ثابت دارد که این مشتق ثابت به یادگیری شبکه سرعت می بخشد. تابع ReLU به صورت زیر تعریف می شود.

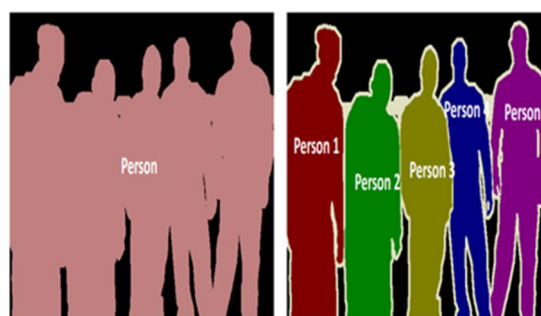
- 1 Pooling
- 2 Fully connected
- 3 Rectified linear unit

در رابطه بالا احتمال عضویت به کلاس λ ام برای بردار نمونه X با وزن W در میان K کلاس موجود بررسی می‌شود. تابع Softmax احتمال مربوط به کلاس‌های مختلف را در بازه 0 تا 1 نرمال می‌کند و کلاسی که دارای بیشترین احتمال باشد به‌عنوان کلاس برنده انتخاب می‌شود.

2-3- تقسیم‌بندی معنایی¹

ایده تقسیم‌بندی معنایی این است که هر پیکسل از تصویر را به یک کلاس خاص طبقه‌بندی کند، روش کار همانند شبکه‌های عصبی است، ولی در اینجا شبکه طراحی شده کمی عمیق‌تر است، چون در اینجا علاوه بر مشخص کردن کلاس هر پیکسل، می‌توان اشکال درون تصاویر را هم تشخیص داد که این قابلیت در شبکه‌های عصبی وجود نداشت، به این روش معمولاً پیش‌بینی متراکم² می‌گویند. در اینجا اشیاء مشابه با برجسب‌های یکسان نشان داده می‌شوند، اما نوع دیگری از تقسیم‌بندی وجود دارد به نام تقسیم‌بندی نمونه³ که می‌تواند اشیاء جداگانه از همان کلاس را پیدا کند و جداگانه تقسیم‌بندی کند.

در شکل (3) تفاوت تقسیم‌بندی معنایی و تقسیم‌بندی نمونه مشخص می‌باشد. در تقسیم‌بندی معنایی نمونه‌هایی که از یک جنس هستند در یک کلاس واحد طبقه‌بندی می‌شوند و برجسب یکسان می‌خورند، در صورتی که در تقسیم‌بندی نمونه، هر نمونه یک برجسب جداگانه می‌خورد و با استفاده از این روش می‌توان حتی نمونه‌هایی که از یک جنس یا یک ماهیت می‌باشند را به کلاس‌های متفاوت طبقه‌بندی کرد.



شکل 3- تقسیم‌بندی معنایی (شکل چپ)، تقسیم‌بندی نمونه (شکل راست)

هدف این پژوهش طبقه‌بندی لکه‌های نفتی که از یک جنس و ماهیت هستند در یک کلاس واحد می‌باشد و نیاز به طبقه‌بندی جداگانه هر لکه نفتی در یک کلاس جداگانه نمی‌باشد، پس برای طبقه‌بندی باید از الگوریتم‌هایی استفاده شود که سازوکارشان بر اساس روش‌های تقسیم‌بندی معنایی می‌باشد. در ادامه، توضیحات مختصری در مورد ساختار شبکه‌های طراحی شده داده می‌شود.

2-4- معماری شبکه U-Net

معماری U-Net در ابتدا برای تقسیم‌بندی تصاویر زیست‌پزشکی پیشنهاد شد [28]، اما کاربرد آن محدود به تصاویر پزشکی نیست [29و30]. U-Net را می‌توان به عنوان یک مدل توسعه یافته‌ای از یک شبکه کاملاً پیچشی (FCN⁴) در نظر گرفت [31] که عمدتاً از دو بخش تشکیل شده‌است: قسمت کدگذاری⁵ و قسمت کدگشایی⁶. U-Net از معماری متقارن بین کدگذاری و کدگشایی شبکه گرفته شده‌است. مسیر کدگذاری از معماری مبتنی بر FCN پیروی می‌کند و هدف آن ثبت محتوای یک تصویر است و برعکس، مسیر کدگشایی، محلی‌سازی⁷ دقیق را امکان‌پذیر می‌کند و در حالی که فیلترهای شبکه کاهش می‌یابد، با نمونه‌برداری از نقشه ویژگی استخراج شده بازنمایی وسیع‌تر اما کم عمق‌تری ایجاد می‌شود.

قسمت سمت چپ شکل (4) مسیر کدگذاری شبکه می‌باشد که در هر لایه آن از دو کانولوشن 3×3 و یک 2×2 max_pooling با گام 2 و یک dropout به اندازه 0.2 استفاده شده‌است که سبب می‌شود ابعاد نقشه ویژگی به تدریج کاسته شود و در مقابل، تعداد کانال‌های ویژگی افزایش پیدا کند. و قسمت سمت راست شبکه مسیر کدگشایی می‌باشد که به تدریج رزولوشن مکانی را افزایش می‌دهد و در هر مرحله از نمونه‌برداری، از کانولوشن 2×2 ("up-convolution") استفاده می‌کند که کانال‌های ویژگی را کاهش می‌دهد. علاوه بر این، در هر مرحله از کدگشایی، نقشه ویژگی نمونه‌برداری شده با نقشه ویژگی مربوط به مرحله کدگذاری برای جلوگیری از زدست رفتن

4 Fully Convolution Network

5 Encoder

6 Decoder

7 Precise Localization

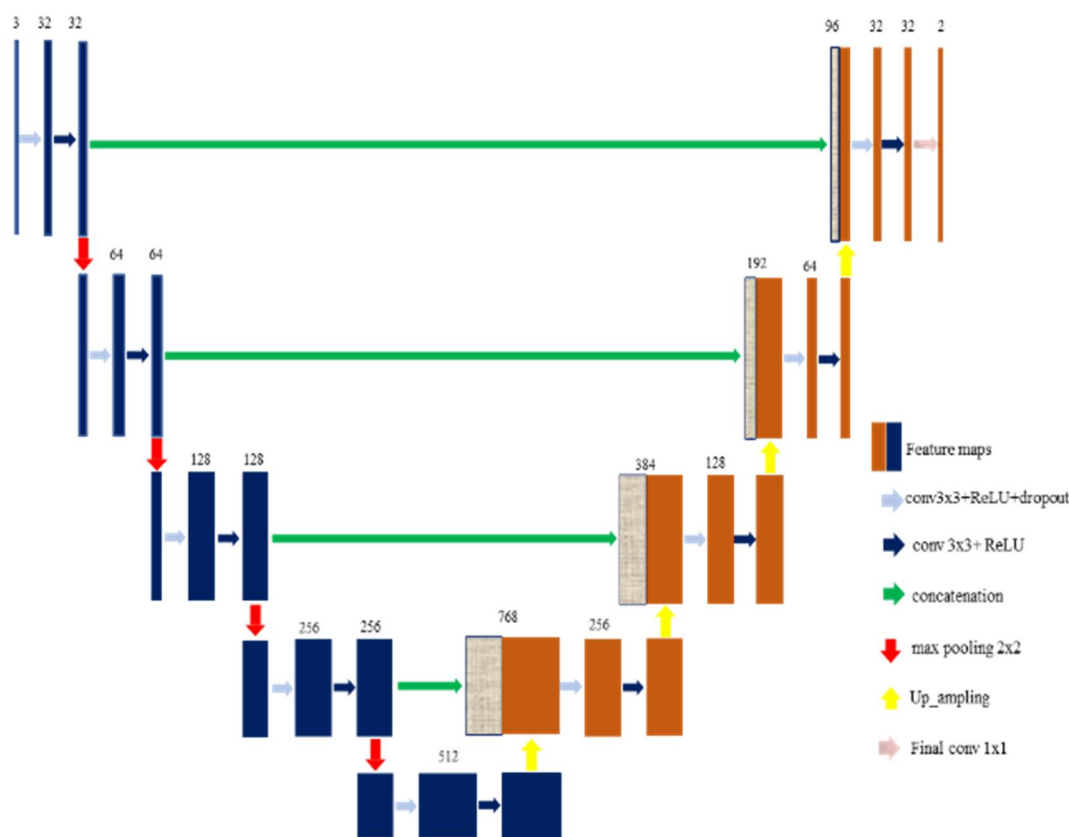
1 Semantic Segmentation Models

2 Dense Prediction

3 Instance Segmentation

شود تا بردار ویژگی هر پیکسل را به تعداد کلاس‌های دلخواه بسط دهد و یک ماسک تقسیم‌بندی برای هر پیکسل تولید کند. با توجه به دو کلاس بودن کار پیش‌رو، خروجی لایه آخر یک تصویر به ابعاد تصویر ورودی و در دو باند می‌باشد.

اطلاعات ترکیب می‌شوند، و به دنبال آن، دو کانولوشن پی‌درپی 3×3 اعمال می‌شود که این کانولوشن‌ها باعث می‌شوند ابعاد کانال‌های نقشه ویژگی نصف شود و ابعاد تصویر به تدریج کاهش یابد. در آخرین مرحله از این شبکه یک کانولوشن 1×1 روی خروجی کدگشایی شده اعمال می‌شود.



شکل 4- معماری شبکه U_Net طراحی شده

اساس این اصل است که اگر هر لایه به طور مستقیم به لایه‌های دیگر به صورت روبه‌جلو متصل شود، شبکه دقیق تر و آموزش آسان‌تر خواهد بود. شکل (5) معماری شبکه FC-DenseNet طراحی شده می‌باشد که در ادامه ساختار شبکه را توضیح می‌دهیم.

شبکه FC-DenseNets یک الگوی اتصال پیچیده طراحی کرده‌است که به طور مکرر همه ویژگی‌های خروجی به صورت پیش‌خور² به هم متصل می‌شوند. برای مثال خروجی لایه ℓ^{th} به صورت زیر تعریف می‌شود.

$$x_{\ell} = H_{\ell}([x_{\ell-1}, x_{\ell-2}, \dots, x_0]) \quad (4)$$

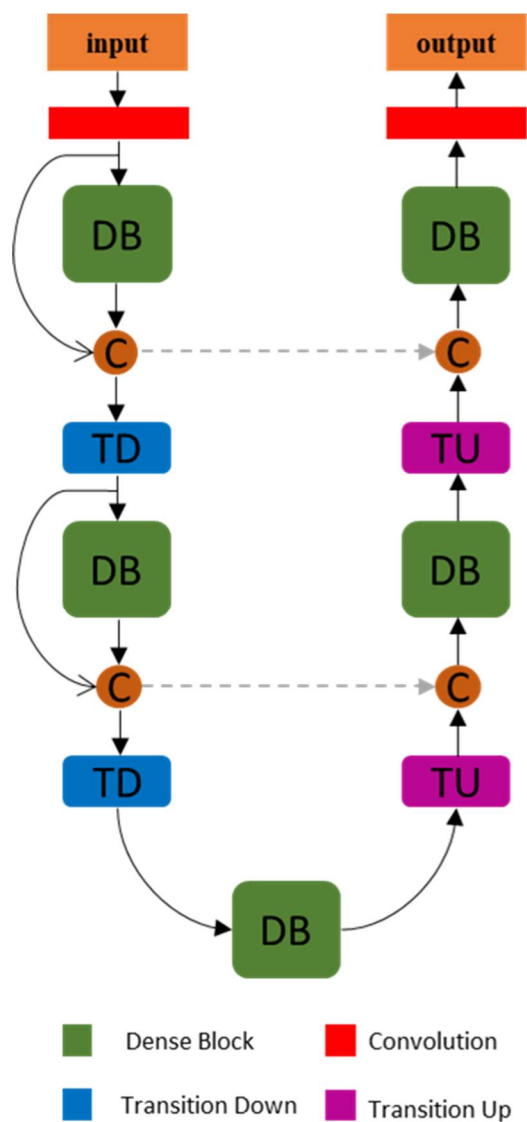
2 Feedforward

2-5- معماری شبکه FC-DenseNet

معماری تقسیم‌بندی معنایی معمولی متشکل از یک مسیر نمونه‌برداری پایین¹ که مسئول استخراج ویژگی‌های معنایی درشت و به دنبال آن یک مسیر نمونه‌برداری دیگر که برای بازیابی وضوح تصویر ورودی در خروجی مدل است استفاده می‌شود. بعد از گذشت چند سال، مدل‌های تقسیم‌بندی معنایی معمولی جای خودشان را به مدل‌های تقسیم‌بندی پیشرفته تری دادند که کارایی بهتری داشتند. یکی از این مدل‌ها که نتایج قابل قبولی هم داشته مدل FC-DenseNet می‌باشد [32]. ایده این شبکه بر

1 Downsampling

که در آن [...] عملیات الحاق¹ را نشان می‌دهد. این الگوی ارتباطی به شدت سبب افزایش استفاده مجدد از ویژگی‌ها می‌شود. خروجی هر لایه دارای k نقشه ویژگی است که به این پارامتر نرخ رشد² گفته می‌شود و معمولاً روی یک مقدار کوچک (به عنوان مثال $k=32$) تنظیم می‌شود.



شکل 5- معماری شبکه FC-DenseNet طراحی شده

در شکل (5) دایره نشان‌دهنده الحاق می‌باشد و فلش‌ها نشان‌دهنده الگوهای اتصال³ در شبکه می‌باشد که قسمت‌های مختلف شبکه را به هم متصل می‌کنند و فلش‌های افقی

خاکستری نشان‌دهنده اتصالات پرش هستند، یعنی نقشه‌های ویژگی از مسیر نمونه‌برداری پایین⁴ با نقشه‌های ویژگی مربوطه در مسیر نمونه‌برداری بالا⁵ به هم متصل می‌شوند. دو لایه کانولوشن با ابعاد 1×1 در ابتدا و انتهای شبکه وجود دارد.

بلوک Dense شامل چهار لایه مختلف می‌باشد. در ابتدا ورودی x_0 (تصویر ورودی یا خروجی transition down) با تعداد بعد m وارد این بلوک می‌شود و با اعمال یک عملیات کانولوشن و سپس یک یکسو کننده غیرخطی مانند ReLU) و در اغلب موارد یک dropout، خروجی x_1 با تعداد بعد k را تولید می‌کند و این خروجی به ورودی x_0 الحاق ($[x_0, x_1]$) می‌شود و وارد لایه بعدی می‌شود. این عملیات چهار بار تکرار می‌شود که در انتها ابعاد ویژگی‌های خروجی بلوک Dense برابر با $4 \times k$ می‌باشد.

لایه transition down برای کاهش رزولوشن مکانی نقشه‌های ویژگی می‌باشد که به ترتیب متشکل از یک تابع نرمال‌سازی، یک یکسو کننده غیرخطی مانند ReLU)، یک کانولوشن 1×1 ، یک dropout و یک ادغام 2×2 می‌باشد. توجه داشته باشید که در مسیر نمونه‌برداری پایین تعداد ویژگی‌ها بصورت خطی افزایش می‌یابد و ابعاد نقشه‌های ویژگی (رزولوشن مکانی) کاهش می‌یابد.

لایه transition Up از یک کانولوشن معکوس⁶ تشکیل شده که نقشه ویژگی دریافتی را نمونه‌برداری می‌کند و این نقشه ویژگی را با نقشه ویژگی دریافتی از اتصال پرش الحاق می‌کند تا ورودی بلوک Dense بعد تشکیل شود. در قسمت نمونه‌برداری بالا، چون رزولوشن مکانی نقشه‌های ویژگی در حال افزایش می‌باشند، نیاز به حافظه بیشتری می‌باشد. به همین دلیل، در مسیر نمونه‌برداری ورودی بلوک‌های Dense به خروجی بلوک الحاق نمی‌شود.

6-2- معماری شبکه DeepLabV3+

آخرین ارتقاء معماری DeepLab، به نام DeepLabV3+ [33] نامگذاری شد و با افزودن یک رمزگشای موثر و در عین حال ساده به نسخه‌های قبلی به‌وجود آمده که این مدل با اصلاح نتایج طبقه‌بندی سبب تولید مرزهای متمایز

4 Downsampling

5 Upsampling

6 Transposed convolution

1 Concatenation

2 Growth rate

3 Connectivity

یک لایه فعال‌سازی ReLU باشد. خروجی قسمت رمزگذاری به یک نسخه تقویت‌شده ASPP³ ارسال می‌شود. هدف ASPP ارائه اطلاعات چند مقیاسی به مدل است. جدای از چهار کانولوشن ASPP، Atrous از یک شاخه اضافی تشکیل شده‌است که با اعمال یک اپراتور ادغام میانگین بر روی نقشه ویژگی قسمت رمزگذاری اطلاعات بستر global را فراهم می‌کند. خروجی های ASPP با هم متصل می‌شوند و از یک لایه کانولوشن 1x1 عبور می‌کنند و خروجی قسمت رمزگذاری تولید می‌شود.

ماژول رمزگشایی دارای دو ورودی است، ورودی اول نتیجه شاخه اصلی رمزگذاری است که به صورت دوخطی با ضرب چهار نمونه‌برداری شده‌است و ورودی دوم نقشه ویژگی سطح پایین استخراج‌شده از قسمت رمزگذاری می‌باشد که توسط یک لایه کانولوشن 1x1 پردازش شده است. این کانولوشن به منظور متعادل کردن اهمیت بین ویژگی‌های سطح پایین تصویر قسمت رمزگذاری و ویژگی‌های فشرده‌شده معنایی قسمت رمزگذاری اعمال می‌شود. هر دو ورودی در یک نقشه ویژگی به هم پیوسته‌اند که در آن دو کانولوشن 3x3 متوالی برای اصلاح ویژگی‌های پیوسته اعمال می‌شود. در نهایت، خروجی به صورت دوخطی با ضرب چهار نمونه‌برداری می‌شود تا ابعاد اولیه تصویر ورودی بازیابی شود.

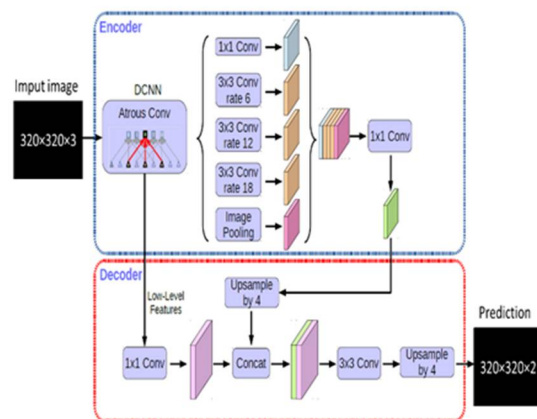
تابع ضرر⁴ برای محاسبه میزان مناسب بودن الگوریتمی که آموزش داده‌ایم می‌باشد و با مقایسه مقادیر پیش‌بینی شده و واقعی بدست می‌آید. اگر مقادیر پیش‌بینی شده با مقادیر واقعی فاصله زیادی داشته باشد، تابع ضرر افزایش می‌یابد. توابع ضرر انواع مختلفی دارد که با توجه نوع داده‌ها انتخاب می‌شود. در این پژوهش از تابع ضرر Dice استفاده شده‌است. ضریب Dice یک معیار پرکاربرد در بینایی کامپیوتر می‌باشد که برای محاسبه شباهت بین دو تصویر است. این ضریب در سال 2016، به عنوان یک تابع ضرر معرفی شد [37] و با عنوان ضرر Dice نامیده شده‌است که بصورت زیر محاسبه می‌شود.

$$Dice\ Loss(y, \hat{y}) = 1 - \frac{2 \times y\hat{y} + 1}{y + \hat{y} + 1} \quad (5)$$

3 Atrous spatial pyramid pooling

4 Loss

بین عوارض مختلف شده‌است. قسمت رمزگذاری متکی بر چارچوب DeepLabv3 [34] است، در حالی که از یک مدل اصلاح‌شده و همسو با Xception [35 و 36] به‌عنوان ستون فقرات مدل استفاده شده که به‌درستی با تقسیم‌بندی معنایی سازگار شده‌است. معماری مدل DeepLabv3 در شکل (6) آورده شده‌است که در ادامه این مدل را شرح داده می‌شود.



شکل 6- معماری شبکه DeepLabV3+ طراحی شده

مدل Xception با هدف ساختن یک شبکه عمیق‌تر، با لایه‌های بیشتری گسترش یافته‌است. علاوه بر این، با گام برداشتن هر لایه ادغام (max pooling) با یک کانولوشن تفکیک‌پذیر عمقی¹ جایگزین شده‌است. کانولوشن تفکیک‌پذیر عمقی به عنوان یک تکنیک کارآمد برای محاسبه کانولوشن استاندارد با هزینه محاسباتی کمتر پیشنهاد شده‌است. برای این منظور، عملیات کانولوشن به دو مرحله تقسیم می‌شود: (الف) کانولوشن عمقی و (ب) کانولوشن نقطه‌ای. کانولوشن عمقی یک فرآیند مکانی بر روی هر کانال ورودی به صورت جداگانه است و برعکس، کانولوشن نقطه‌ای یک کانولوشن ساده با اندازه هسته 1x1 است که در خروجی کانولوشن عمقی اعمال می‌شود تا اطلاعات مکانی را در هر کانال ترکیب کند. کانولوشن عمقی توسط کانولوشن Atrous پشتیبانی می‌شود که منجر به جداسازی کانولوشن می‌شود و می‌تواند پیچیدگی محاسباتی مدل را کاهش دهد و در نهایت از نقشه‌های ویژگی وضوح دلخواه را استخراج کند. آخرین اصلاح مدل Xception این است که بعد از هر کانولوشن عمقی 3x3 یک نرمال‌سازی دسته‌ای² و

1 Depthwise separable convolution

2 Batch normalization

2-7- معیارهای ارزیابی شبکه

خروجی مدل‌ها حاصل از دو نوع تابع فعال‌سازی می‌باشند: (الف) حاصل از تابع فعال‌سازی Sigmoid، که تصاویری تک‌بندیده بوده که احتمال تعلق هر پیکسل به کلاس هدف را نشان می‌دهد. بنابراین، با گذاشتن یک حد آستانه احتمالات با مقدار بیشتر از 0.5 را برای تولید تصاویر تقسیم‌بندی باینری انتخاب شد. (ب) حاصل از تابع فعال‌سازی Softmax، که تصاویری دو باندیده بوده که هر باند احتمال تعلق داشتن پیکسل به کلاس نفت یا پس‌زمینه را مشخص می‌کند.

جدول 2- اطلاعات مدل‌ها

مدل	U_Net	FC_DenseNet	DeepLabV3+
تعداد پارامترها	12,559,170	1,394,544	11,852,610
تابع ضرر	Dice loss	Dice loss	Dice loss

در این حالت، با مقایسه مقادیر این دو باند برای هر پیکسل و انتخاب مقدار بیشتر، کلاس هر پیکسل مشخص شد و تصاویر نهایی تقسیم‌بندی باینری تولید شد. با مقایسه تصاویر تقسیم‌بندی باینری و تصاویر واقعیت زمینی می‌توان ماتریس درهم‌ریختگی¹ را محاسبه کرد. از ماتریس درهم‌ریختگی، چندین معیار استخراج می‌شود، از جمله مقادیر مثبت صحیح (TP)، منفی صحیح (TN)، مثبت کاذب (FP) و منفی کاذب (FN). با استفاده از این معیارها می‌توان معیارهای دقت²، حساسیت³، صحت⁴، امتیاز⁵ F1 و IoU را برای ارزیابی مدل محاسبه کرد. در بین این پنج معیار، معیار IoU سخت‌گیرانه‌ترین معیار است و معمولاً مقدار کمتری نسبت به بقیه معیارها دارد [38]. روابط محاسبه این پنج معیار در معادلات (6)، (7)، (8)، (9) و (10) تعریف شده‌است.

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + TN + FP} \quad (6)$$

- 1 Confusion matrix
- 2 Accuracy
- 3 Sensitivity
- 4 Precision
- 5 F1-score

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (7)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (8)$$

$$\text{F1 - score} = \frac{2 \times \text{Precision} \cdot \text{Sensitivity}}{\text{Precision} + \text{Sensitivity}} \quad (9)$$

$$\text{IoU} = \frac{TP}{TP + FP + FN} \quad (10)$$

3- نتایج

در مجموع از 2172 تصویر برای آموزش شبکه و 238 تصویر برای تست شبکه استفاده شده‌است. ابعاد تصاویر ورودی و خروجی برابر بوده و برابر 320×320 می‌باشد. همچنین، ده درصد داده‌های آموزشی بصورت رندوم برای ارزیابی مدل استفاده شده‌است. تابع ضرر مورد استفاده در شبکه‌ها ضرر Dice می‌باشد. در جدول (2)، تعداد پارامترهای آموزش شبکه‌ها و نوع تابع ضرر مورد استفاده آورده شده‌است. در مدل‌های یادگیری عمیق یکی از عوامل موثر در سرعت آموزش شبکه‌ها، میزان پارامترهای شبکه می‌باشد [26] و با افزایش پارامترها معمولاً زمان آموزش شبکه افزایش می‌یابد.

با توجه به جدول (2)، به‌ترتیب مدل‌های U_Net، FC_DenseNet و DeepLabV3+ بیشترین پارامترهای آموزشی را برای شبکه‌های طراحی‌شده دارند. در روند آموزش مدل‌ها به دلیل محدودیت حافظه، اندازه دسته⁶ ثابت گرفته نشد و برای مدل U_Net برابر با 4 و برای دو مدل‌های دیگر برابر با 12 در نظر گرفته شد. تجربه ثابت کرده هرچه اندازه دسته بیشتر باشد، دقت‌های بدست آمده نیز بیشتر می‌باشند [25].

جدول (3) نتایج بدست‌آمده از مدل‌ها می‌باشد. همانطور که قبلاً اشاره شد، در بین معیارهای ارزیابی، معیار IoU معمولاً مقدار کمتری دارد، چون این معیار سخت‌گیرانه‌تر از سایر معیارها می‌باشد و در اکثر موارد از این معیار برای ارزیابی عملکرد مدل‌ها استفاده می‌شود. بهترین مقدار برای هر معیار ارزیابی را درشت‌تر کرده‌ایم که به راحتی قابل تشخیص باشد. با توجه به معیار دقت، مدل‌های

6 Batch size

خوبی تشخیص نداده‌اند ولی بطور کلی مدل DeepLabV3+ توانسته نتایج خیلی بهتری نسبت به سایر مدل‌ها بدست آورد. یکی از مهمترین چالش‌هایی که در یادگیری ماشین با آن روبه‌رو هستیم مشکل کمبود درصد کلاس هدف نسبت به پس‌زمینه می‌باشد که در مقالات مختلف سعی شده با روش‌های مختلفی بر آن غلبه کنند. برای مثال، شعبان و همکاران [39] تصاویر SAR را به تکه‌هایی با ابعاد 64x64 تبدیل کردند که میزان لکه نفت نسبت به پس‌زمینه بین 40 تا 60 درصد کل تصویر بوده و با استفاده از معماری U-Net در یک طبقه‌بندی دو کلاسه به دقت 92٪ برای تشخیص لکه‌های نفتی دست یافتند. از طرفی برای به دست آوردن تقسیم‌بندی دقیق موارد نشت نفت، تعادل بین نشت نفت و کلاس پس‌زمینه ممکن است امکان‌پذیر نباشد که برای جبران آن می‌توان از استخراج دستی ویژگی‌های خاص مثل انحراف استاندارد شیء¹، انحراف استاندارد پس‌زمینه² و نسبت قدرت شیء به میانگین³ استفاده کرد که در [20 و 40]، به دقت بیشتر از 92٪ دست یافته‌اند. معمولاً در روش‌های مبتنی بر یادگیری عمیق [41 و 42] که از ساختارهای CNN برای استخراج ویژگی‌های خودکار و همچنین طبقه‌بندی تصاویر SAR استفاده کرده‌اند، برای کاهش نسبت پس‌زمینه به کلاس هدف در تصاویر آزمایش‌شده، از تقسیم تصویر به تکه‌های کوچک‌تر استفاده کرده‌اند.

مدل‌های از پیش آموزش‌دیده، مانند شبکه‌های ResNet-101، VGG-16 و شبکه‌های GAN در [41 و 43] یا شبکه‌های CNN چند سطحی مانند [24]، برای طبقه‌بندی اتوماتیک وصله‌ها، با عملکردی متوسط معرفی شده‌اند و دقت خیلی زیادی ندارند ولی باید به این نکته توجه داشت که در حالت کلی در مباحث یادگیری ماشین می‌توان با انجام تغییرات کوچک در ساختار شبکه و داده‌ها، دقت مدل را تا مقدار قابل قبولی افزایش داد.

در یادگیری ماشین برای مقایسه میزان کارآمدی هر مدل، معیارهای ارزیابی مدل پیاده‌سازی شده را با معیارهای ارزیابی مدل‌های دیگر مقایسه می‌کنند.

DeepLabV3+، FC-DenseNet و U-Net بترتیب بهترین نتایج را داشته‌اند، با توجه به معیار حساسیت، مدل‌های DeepLabV3+ و U-Net بترتیب بهترین نتایج را داشته‌اند، با توجه به معیار صحت، مدل‌های DeepLabV3+ و FC-DenseNet بترتیب بهترین نتایج را داشته‌اند، با توجه به معیار ارزیابی F1-score، مدل‌های DeepLabV3+ و FC-DenseNet بترتیب بهترین نتایج را داشته‌اند و در نهایت مدل‌ها با معیار IoU مورد ارزیابی قرار گرفتند که با توجه به نتایج مدل‌های DeepLabV3+، U-Net و FC-DenseNet بترتیب بهترین نتایج را داشته‌اند. در بین این سه مدل، مدل DeepLabV3+ بیشترین مقادیر را برای همه معیارهای ارزیابی بدست آورده‌است.

جدول 3- نتایج پارامترهای ارزیابی شبکه

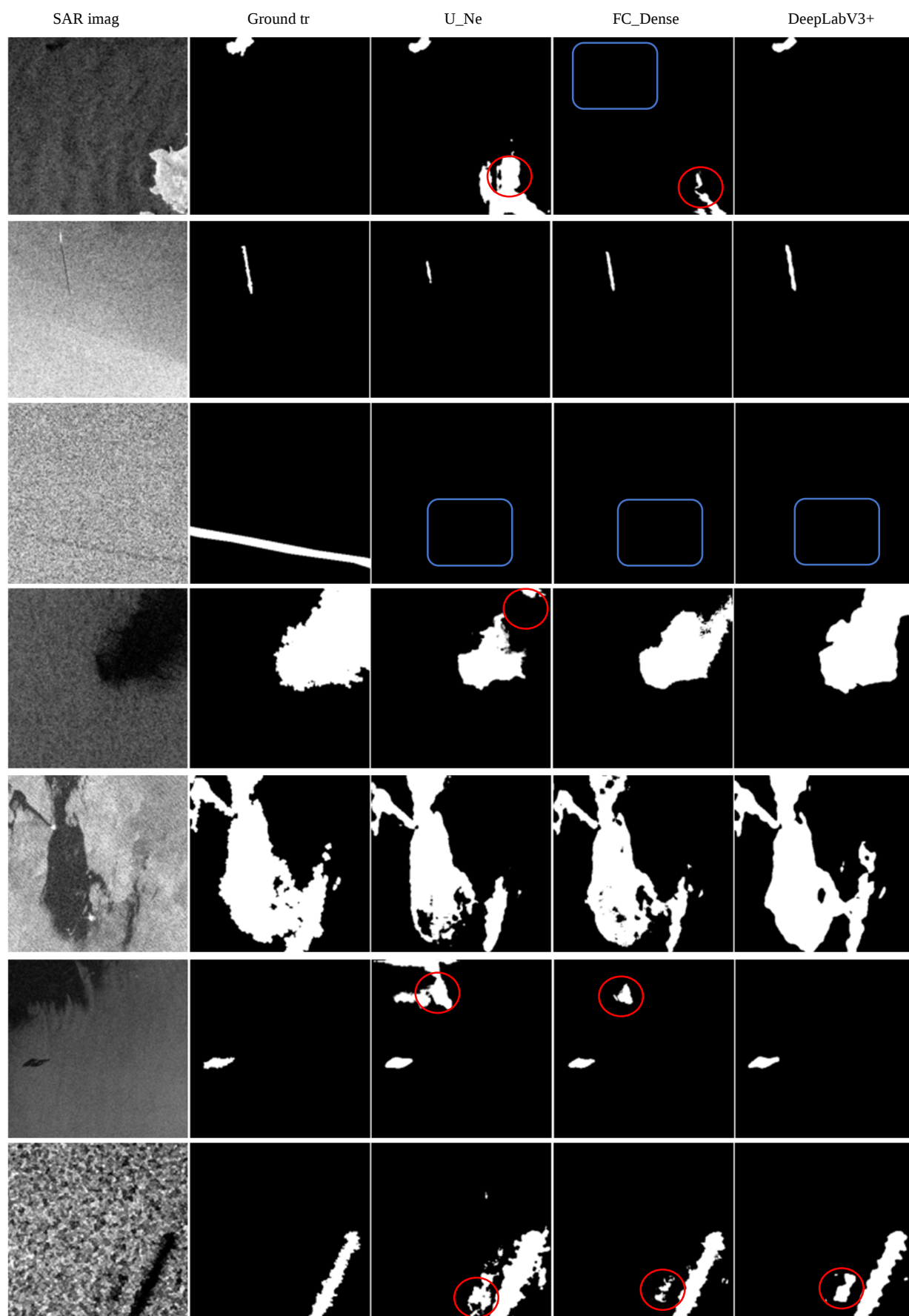
مدل	U_Net	FC_DenseNet	DeepLabV3+
Mean Accuracy (%)	0.976	0.979	0.982
Mean Sensitivity (%)	0.655	0.623	0.695
Mean Precision (%)	0.730	0.814	0.838
Mean F1-score (%)	0.691	0.706	0.760
Mean IoU (%)	0.547	0.545	0.613

شکل (7) نتایج پیش‌بینی‌شده مدل‌ها برای 7 تصویر متفاوت را نشان می‌دهد. با توجه به خروجی‌های بدست‌آمده و با مقایسه نتایج با تصاویر واقعیت زمینی می‌بینیم که شبکه DeepLabV3+ نسبت به سایر مدل‌ها خروجی بهتری داشته‌است. با توجه به پایین بودن مقدار پیکسل‌های هدف نسبت به پس‌زمینه و همچنین عدم وضوح کافی تصویر SAR، در بعضی از تصاویر هیچ‌کدام از مدل‌ها قادر به تشخیص درست نفت نبوده‌اند؛ این درحالی است که در بعضی از تصاویر، هر سه مدل‌ها توانسته‌اند بطور کامل نفت را تشخیص دهند، در شکل (7) نواحی که با دایره قرمز نشان داده شده‌است قسمت‌هایی می‌باشد که هر شبکه به اشتباه نفت تشخیص داده‌است و نواحی که با مستطیل آبی نشان‌داده شده‌است نواحی می‌باشد که شبکه نتوانسته‌است لکه نفت را تشخیص دهد. در بعضی موارد هر سه مدل در پیدا کردن جزئیات تصاویر ضعیف بوده‌اند و جزئیات را به

1 Object standard deviation

2 Background standard deviation

3 Object power to mean ratio



شکل 7- نتایج پیش‌بینی شده مدل‌ها برای 7 تصویر متفاوت بعد از 50 دور آموزش شبکه، نرخ یادگیری برای کل مدل‌ها برابر با 0.0001 بوده است.

درضمن، این ارزیابی زمانی درست و قابل قبول می‌باشد که داده‌های مورد استفاده در همه مدل‌ها یکی باشد و نوع داده‌ها تغییری نکند. به دلیل این که داده‌های مورد استفاده در این مقاله سال 2019 تولید شده‌است، تحقیق زیادی روی آن انجام نشده‌است، ما توانستیم تنها یک مقاله مرتبط پیدا کنیم [26]. نویسندگان این مقاله داده‌های مورد استفاده در مقاله پیش‌رو را جمع‌آوری و شش مدل مختلف را پیاده‌سازی کرده‌اند که از بین این مدل‌ها، مدل‌های U-Net و DeepLabV3+ بین هر دو مقاله مشترک می‌باشد و مدل FC-DenseNet برای اولین باری روی این دیتاست پیاده‌سازی شده‌است. همچنین، برای کل مدل‌ها از تابع ضرر cross-entropy استفاده کرده بودند و باتوجه به نوع مدل، اندازه دسته را برابر با 12 یا 16 در نظر گرفته بودند. با توجه به بزرگ‌بودن ابعاد تصویر (650×1250)، در مرحله آموزش شبکه از هر تصویر آموزشی بصورت رندوم یک تصویر با سایز کوچک‌تر (معمولاً 320×320) از یک قسمت تصویر اصلی انتخاب کردند. در این مقاله نتایج معیار IoU گزارش شده‌است که بالاترین مقدار این معیار مربوط به شبکه U-Net بوده که برابر با 0.538 است و بعد از مدل U-Net بیشترین مقدار IOU مربوط به مدل DeepLabV3+ بوده که برابر با 0.534 است. این درحالی است که مقادیر IoU بدست‌آمده از مدل‌های ما برابر با 0.545، 0.547 و 0.613 می‌باشد و همه مقادیر بدست‌آمده بزرگتر از حداکثر مقدار IoU بدست‌آمده در [26] می‌باشد که این موضوع نشان می‌دهد شبکه‌های طراحی شده توانسته‌اند عملکرد خیلی بهتری نسبت به کارهای مشابه گذشته داشته باشند و این موضوع بسیار امیدوار کننده بود.

4- نتیجه‌گیری

نشت نفت یکی از بزرگترین تهدیدهای محیط‌های دریایی و ساحلی محسوب می‌شود و به طور جدی بر اکوسیستم دریایی تأثیر می‌گذارد و نگرانی‌های سیاسی و زیست‌محیطی ایجاد می‌کند، زیرا اکوسیستم دریایی و ساحلی را به طور جدی تحت تأثیر قرار می‌دهند. در

سال‌های اخیر روش‌های زیادی برای شناسایی لکه‌های نفتی با استفاده از تصاویر ماهواره‌ای ارائه شده‌است. در این مقاله، یک چارچوب یادگیری عمیق برای شناسایی موارد نشت نفت در تصاویر SAR معرفی شد. ابتدا هر یک از تصاویر با ابعاد 650×1250×3 به هشت تصویر با ابعاد 320×320×3 تبدیل شد. در ادامه تصاویر تک‌کلاسه از روند آموزش حذف شدند که تأثیر بسیار مثبتی در روند آموزش داشت، سپس، با طراحی شبکه‌های U-Net، DeepLabV3+ و FC-DenseNet و تنظیم پارامترهای مختلف این شبکه‌ها، لکه‌های نفت تشخیص و شناسایی داده‌شد. همچنین، از تابع ضرر Dice برای آموزش شبکه‌ها استفاده شده که نتایج خیلی بهتری نسبت به کارهای قبلی برای همه مدل‌ها داشته‌است. IoU بهترین معیار برای مقایسه عملکرد مدل‌ها می‌باشد، چرا که نسبت به سایر معیارها سخت‌گیرانه‌ترین است. نتایج IoU برای مدل‌های U-Net، DeepLabV3+ و FC-DenseNet به ترتیب برابر 0.547، 0.613 و 0.545 می‌باشد. همچنین، نتایج F1-score برای مدل‌های U-Net، DeepLabV3+ و FC-DenseNet به ترتیب برابر 0.691، 0.760 و 0.706 می‌باشد. بر اساس این نتایج مدل‌های DeepLabV3+، FC-DenseNet و U-Net به ترتیب بهترین کارایی را برای شناسایی لکه‌های نفتی داشته‌اند.

جهت بهبود عملکرد مدل‌ها در تشخیص لکه‌های نفتی در کارهای دیگر می‌توان تأثیر پارامترهای هواشناسی مانند سرعت باد را نیز مورد بررسی قرار داد. همچنین، می‌توان ویژگی‌های راداری و استخراج لبه را در کنار تصاویر موجود ورودی وارد مدل کرد. همچنین، در صورت آموزش مدل‌های مختلف می‌توان اثر یادگیری تجمعی¹ را در بهبود نتایج مورد ارزیابی قرار داد.

سپاسگزاری

نویسندگان این مقاله لازم می‌دانند از آژانس فضایی اروپا (ESA)، به دلیل در اختیار گذاشتن داده‌های نشت نفت کمال تقدیر و تشکر را داشته باشند.

¹ Ensemble Learning

مراجع

- [1] F. Del Frate, A. Petrocchi, J. Lichtenegger, and G. Calabresi, "Neural networks for oil spill detection using ERS-SAR data," *IEEE Transactions on geoscience and remote sensing*, vol. 38, no. 5, pp. 2282-2287, 2000.
- [2] C. Brekke and A. H. Solberg, "Oil spill detection by satellite remote sensing," *Remote sensing of environment*, vol. 95, no. 1, pp. 1-13, 2005.
- [3] A. H. Solberg, C. Brekke, and P. O. Husoy, "Oil spill detection in Radarsat and Envisat SAR images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 3, pp. 746-755, 2007.
- [4] K. N. Topouzelis, "Oil spill detection by SAR images: dark formation detection, feature extraction and classification algorithms," *Sensors*, vol. 8, no. 10, pp. 6642-6659, 2008.
- [5] A. H. S. Solberg, "Remote sensing of ocean oil-spill pollution," *Proceedings of the IEEE*, vol. 100, no. 10, pp. 2931-2945, 2012.
- [6] M. Fingas and C. Brown, "Review of oil spill remote sensing," *Marine pollution bulletin*, vol. 83, no. 1, pp. 9-23, 2014.
- [7] M. F. Fingas and C. E. Brown, "Review of oil spill remote sensing," *Spill Science & Technology Bulletin*, vol. 4, no. 4, pp. 199-208, 1997.
- [8] I. A. Kapustin et al., "On capabilities of tracking marine surface currents using artificial film slicks," *Remote Sensing*, vol. 11, no. 7, p. 840, 2019.
- [9] H. Espedal and O. Johannessen, "Cover: Detection of oil spills near offshore installations using synthetic aperture radar (SAR)," 2000.
- [10] A. S. Solberg, G. Storvik, R. Solberg, and E. Volden, "Automatic detection of oil spills in ERS SAR images," *IEEE Transactions on geoscience and remote sensing*, vol. 37, no. 4, pp. 1916-1924, 1999.
- [11] B. Fiscella, A. Giancaspro, F. Nirchio, P. Pavese, and P. Trivero, "Oil spill detection using marine SAR images," *International Journal of Remote Sensing*, vol. 21, no. 18, pp. 3561-3566, 2000.
- [12] H. A. Espedal, "Satellite SAR oil spill detection using wind history information," *International Journal of Remote Sensing*, vol. 20, no. 1, pp. 49-65, 1999.
- [13] D. L. d. Souza, A. D. Neto, and W. d. Mata, "Intelligent system for feature extraction of oil slick in sar images: Speckle filter analysis," in *International Conference on Neural Information Processing*, 2006: Springer, pp. 729-736.
- [14] I. Keramitsoglou, C. Cartalis, and C. T. Kiranoudis, "Automatic identification of oil spills on satellite images," *Environmental modelling & software*, vol. 21, no. 5, pp. 640-652, 2006.
- [15] V. Karathanassi, K. Topouzelis, P. Pavlakis, and D. Rokos, "An object-oriented methodology to detect oil spills," *International Journal of Remote Sensing*, vol. 27, no. 23, pp. 5235-5251, 2006.
- [16] M. Konik and K. Bradtke, "Object-oriented approach to oil spill detection using ENVISAT ASAR images," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 118, pp. 37-52, 2016.
- [17] Topouzelis, K.; Psyllos, A. Oil spill feature selection and classification using decision tree forest on SAR image data. *ISPRS J. Photogramm. Remote Sens.* 2012, 68, 135–143.
- [18] Mercier, G.; Girard-Ardhuin, F. Partially supervised oil-slick detection by SAR imagery using kernel expansion. *IEEE Trans. Geosci. Remote Sens.* 2006, 44, 2839–2846.
- [19] K. Topouzelis and A. Psyllos, "Oil spill feature selection and classification using decision tree forest on SAR image data," *ISPRS journal of photogrammetry and remote sensing*, vol. 68, pp. 135-143, 2012.
- [20] D. Song, Y. Ding, X. Li, B. Zhang, and M. Xu, "Ocean oil spill classification with RADARSAT-2 SAR based on an optimized wavelet neural network," *Remote Sensing*, vol. 9, no. 8, p. 799, 2017.
- [21] D. Stathakis, K. Topouzelis, and V. Karathanassi, "Large-scale feature selection using evolved neural networks," in *Image and Signal Processing for Remote Sensing XII*, 2006, vol. 6365: SPIE, pp. 342-350.
- [22] S. M. Seyed Mousavi and M. Akhoondzadeh Hanzaei, "Monitoring and Prediction of the changes in water zone of wetlands using an intelligent neural-fuzzy system based on data from Google Eearth Engine system (Case study of Anzali Wetland, 2000-2019)," *Engineering Journal of Geospatial Information Technology*, pp. 31-40.
- [23] A. Rostami, R. Shah-Hosseini, S. Asgari, A. Zarei, M. Aghdami-Nia, and S. Homayouni, "Active Fire Detection from Landsat-8 Imagery Using Deep Multiple Kernel Learning," *Remote Sensing*, vol. 14, no. 4, p. 992, 2022.
- [24] G. Orfanidis, K. Ioannidis, K. Avgerinakis, S. Vrochidis, and I. Kompatsiaris, "A deep neural network for oil spill

- semantic segmentation in Sar images," in 2018 25th IEEE International Conference on Image Processing (ICIP), 2018: IEEE, pp. 3773-3777.
- [25] M. Krestenitis, G. Orfanidis, K. Ioannidis, K. Avgerinakis, S. Vrochidis, and I. Kompatsiaris, "Early identification of oil spills in satellite images using deep CNNs," in International Conference on Multimedia Modeling, 2019: Springer, pp. 424-435.
- [26] M. Krestenitis, G. Orfanidis, K. Ioannidis, K. Avgerinakis, S. Vrochidis, and I. Kompatsiaris, "Oil spill identification from satellite images using deep neural networks," Remote Sensing, vol. 11, no. 15, p. 1762, 2019.
- [27] Copernicus Open Access Hub. Available online: <https://scihub.copernicus.eu/> (accessed on 20 July 2020).
- [28] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in International Conference on Medical image computing and computer-assisted intervention, 2015: Springer, pp. 234-241.
- [29] V. Iglovikov and A. Shvets, "Ternausnet: U-net with vgg11 encoder pre-trained on imagenet for image segmentation," arXiv preprint arXiv:1801.05746, 2018.
- [30] V. Iglovikov, S. Mushinskiy, and V. Osin, "Satellite imagery feature detection using deep convolutional neural network: A kaggle competition," arXiv preprint arXiv:1706.06169, 2017.
- [31] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 3431-3440.
- [32] S. Jégou, M. Drozdal, D. Vazquez, A. Romero, and Y. Bengio, "The one hundred layers tiramisu: Fully convolutional densenets for semantic segmentation," in Proceedings of the IEEE conference on computer vision and pattern recognition workshops, 2017, pp. 11-19.
- [33] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," in Proceedings of the European conference on computer vision (ECCV), 2018, pp. 801-818.
- [34] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking atrous convolution for semantic image segmentation," arXiv preprint arXiv:1706.05587, 2017.
- [35] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 1251-1258.
- [36] J. Dai et al., "Deformable convolutional networks," in Proceedings of the IEEE international conference on computer vision, 2017, pp. 764-773.
- [37] C. H. Sudre, W. Li, T. Vercauteren, S. Ourselin, and M. Jorge Cardoso, "Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations," in Deep learning in medical image analysis and multimodal learning for clinical decision support: Springer, 2017, pp. 240-248.
- [38] M. Aghdami-Nia, R. Shah-Hosseini, A. Rostami, and S. Homayouni, "Automatic coastline extraction through enhanced sea-land segmentation by modifying Standard U-Net," International Journal of Applied Earth Observation and Geoinformation, vol. 109, p. 102785, 2022.
- [39] M. Shaban et al., "A deep-learning framework for the detection of oil spills from SAR data," Sensors, vol. 21, no. 7, p. 2351, 2021.
- [40] G. Calabresi, F. Del Frate, J. Lichtenegger, A. Petrocchi, and P. Trivero, "Neural networks for the oil spill detection using ERS-SAR data," in IEEE 1999 International Geoscience and Remote Sensing Symposium. IGARSS'99 (Cat. No. 99CH36293), 1999, vol. 1: IEEE, pp. 215-217.
- [41] W.-K. Baek, H.-S. Jung, and D. Kim, "Oil spill detection of Kerch strait in November 2007 from dual-polarized TerraSAR-X image using artificial and convolutional neural network regression models," Journal of Coastal Research, vol. 102, no. SI, pp. 137-144, 2020.
- [42] M. Ansari, S. Homayouni, A. Safari, and S. Niazmardi, "A New Convolutional Kernel Classifier for Hyperspectral Image Classification," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 14, pp. 11240-11256, 2021.
- [43] M. Nieto-Hidalgo, A.-J. Gallego, P. Gil, and A. Pertusa, "Two-stage convolutional neural network for ship and spill detection using SLAR images," IEEE Transactions on Geoscience and Remote Sensing, vol. 56, no. 9, pp. 5217-5230, 2018.