

دانشگاه صنعتی خواجه نصیرالدین طوسی

پروژه درس یادگیری ماشین

علی صدیقی ارقون

استاد درس:

دکتر بشری پیشگو

بهمن ۱۴۰۴

چکیده

ناحیه‌بندی دقیق تصاویر سه‌بعدی مغز انسان برای تشخیص تومور، برنامه‌ریزی درمان و بسیاری از کاربردهای پزشکی مرتبط با سلامت انسان‌ها ضروری می‌باشد. طی سال‌های اخیر شبکه‌های عصبی عمیق به پیشرفته‌ترین روش‌های ناحیه‌بندی خودکار تبدیل شده‌اند. با این حال، نتایج مبتنی بر روش‌های خودکار ممکن است نیازمند اصلاح باشد تا به حدی دقیق شود که مورد استفاده بالینی قرار گیرد. ناحیه‌بندی تعاملی به کاربران اجازه می‌دهد تا با تعامل خود، نواحی مورد نظر خود را با دقت بیشتری استخراج کنند. اگر چه این عرصه توسط کارهای دیگری بررسی شده است، اما همچنان شکافی بین دنیای آکادمیک و جهان واقعی و صنعتی وجود دارد. روش‌های موجود، مناسب استفاده در سیستم‌های کم‌توان نمی‌باشند. همچنین این روش‌ها برای تولید خروجی مطلوب، نیازمند تلاش زیادی از سمت کاربر می‌باشند. روش استفاده شده در این پروژه هر دو مشکل را با به‌روزرسانی خروجی در نواحی محلی حل می‌کند. در این روش به جای ناحیه‌بندی تصویر به صورت کامل و یکجا، که نیازمند توان محاسباتی بالا می‌باشد، یک ناحیه‌بندی درشت دانه بر روی بُرش‌های کم وضوح از ناحیه مورد نظر کاربر صورت می‌گیرد. سپس یک اصلاح محلی صورت می‌گیرد تا وضوح از دست رفته ناحیه هدف بازیابی شود. همچنین برای جلوگیری از تولید مجدد خروجی برای تمامی نواحی تصویر، از یک تحلیل مورفولوژی استفاده می‌شود تا تنها نواحی نیازمند تغییر اصلاح شوند و جزئیات سایر نواحی با توجه به خروجی قبلی حفظ شود. در این پروژه، کارایی مناسب این روش تحت آزمایش‌های متعددی اثبات می‌شود و نتایج تجربی به دست آمده نشان دهنده دو نکته مثبت می‌باشد. نکته مثبت اول این است که، این روش در مقایسه با سایر روش‌ها، با وجود تعامل کمتر کاربر به نتایج دقیق‌تری دست می‌یابد. همچنین در مقایسه با روش‌های دیگر، این نتایج در زمان کمتری تولید می‌شود و نیازمند توان محاسباتی کمتری می‌باشد. نکته مثبت دوم تعمیم قابل قبول این روش در وظایفی است که در مرحله آموزش با آنها مواجه نشده است.

واژگان کلیدی: یادگیری عمیق، بینایی کامپیوتر، ناحیه‌بندی تعاملی تصویر، ناحیه‌بندی تومور مغزی، تصاویر سه‌بعدی پزشکی

فهرست مطالب

۱-۰	مقدمه	۱
۲-۰	ناحیه‌بندی تصویر	۴
۱-۲-۰	ناحیه‌بندی معنایی	۴
۲-۲-۰	ناحیه‌بندی نمونه‌ای	۵
۳-۲-۰	ناحیه‌بندی سراسرنما	۵
۳-۰	ناحیه‌بندی تعاملی تصویر	۶
۴-۰	دسته‌بندی روش‌های ناحیه‌بندی تعاملی	۶
۱-۴-۰	دسته‌بندی بر اساس نحوه تعامل کاربر	۶
۱-۱-۴-۰	تعامل مبتنی بر Seed	۷
۲-۱-۴-۰	تعامل مبتنی بر ROI	۸
۳-۱-۴-۰	تعامل فعال و منفعل	۹
۲-۴-۰	دسته‌بندی بر اساس روش‌شناسی	۱۰
۱-۲-۴-۰	روش مبتنی بر Contour	۱۱
۲-۲-۴-۰	روش مبتنی بر GraphCut	۱۲
۳-۲-۴-۰	روش مبتنی بر Random Walk	۱۲
۴-۲-۴-۰	روش مبتنی بر Region Growing, Merging	۱۳
۵-۲-۴-۰	روش‌های یادگیری عمیق و ترکیبی	۱۴
۵-۰	معیارهای ارزیابی	۱۵
۱-۵-۰	Error Rate (ERR)	۱۶
۲-۵-۰	True Positive Rate (TPR)	۱۶

۱۶	Precision ۳-۵-۰
۱۷	F1-Score ۴-۵-۰
۱۷	Intersection over Union (IoU) ۵-۵-۰
۱۹	Dice Similarity Coefficient (DSC) ۶-۵-۰
۲۰	Variation of Information (VI) ۷-۵-۰
۲۰	Number of Clicks (NoC) ۸-۵-۰
۲۰	Number of Failures (NoF) ۹-۵-۰
۲۱	Time per Click (TPC) ۱۰-۵-۰
۲۱	mDice و mIoU ۱۱-۵-۰
۲۱	۶-۰ مجموعه دادگان BraTS
۲۱	۱-۶-۰ مقدمه
۲۲	۲-۶-۰ ویژگی‌های BraTS
۲۶	۳-۶-۰ چالش‌های کار با BraTS
۲۷	۴-۶-۰ کاربردها
۲۷	۷-۰ ناحیه‌بندی تعاملی با رویه پالایشی پس‌انتشار (BRS)
۲۷	۱-۷-۰ مقدمه
۲۸	۲-۷-۰ الگوریتم
۲۹	۳-۷-۰ معماری شبکه
۲۹	۴-۷-۰ مرحله آموزش
۳۰	۵-۷-۰ مرحله استنتاج
۳۰	۶-۷-۰ رویه پالایشی پس‌انتشار
۳۱	۷-۷-۰ نتیجه‌گیری

۰-۱ مقدمه

حوزه تصویربرداری پزشکی^۱ در دهه‌های اخیر با استفاده از یادگیری ماشین^۲ و الگوریتم‌های یادگیری عمیق^۳ برای ناحیه‌بندی تصاویر^۴ پزشکی و شناسایی تومورهای مغزی^۵ یا سایر بیماری‌ها دچار تحول عظیم شده است. این حوزه به پزشکان در تشخیص^۶ زودهنگام بدخیمی‌های مغزی و بهبود پیش‌بینی‌ها درباره بیماری کمک می‌کند. بزرگسالان اغلب به گلیوما^۷ مبتلا می‌شوند که این مورد از سلول‌های گلیال^۸ منشا می‌گیرد و در ادامه حتی باعث حمله به بافت‌های سالم مجاور می‌شود [۱۶]. در حالی تصاویر حاصل از مدالیته‌های^۹ تصویربرداری تشدید رزونانس مغناطیسی^{۱۰} به طور دستی توسط رادیولوژیست‌ها برای به دست آوردن اطلاعات کمی و تعیین لبه‌های تومور و نواحی داخلی آن مورد بررسی قرار می‌گیرند، که این بررسی دستی و در ادامه ناحیه‌بندی سه‌بعدی^{۱۱} فرایندی زمان‌بر، با هزینه بالا، مستعد خطا و نیازمند نیروی انسانی متخصص می‌باشد. علاوه بر این موارد ممکن است تشخیص‌های متفاوتی توسط متخصصان مختلف داده شود. این چالش زمانی جدی‌تر می‌شود که تومورها از نظر اندازه، شکل و مکان دارای گوناگونی باشند.

بهبود مواردی مانند تشخیص بیماری، برنامه‌ریزی درمان، نظارت و مسائل بالینی به شدت به ناحیه‌بندی تومورهای مغزی در تصاویر پزشکی وابستگی دارد و نیازمند تعیین محل و اندازه دقیق تومور مغزی است. با این حال، ویژگی‌های تومورهای مغزی، فرایند ناحیه‌بندی صحیح و دقیق را دشوار می‌کند [۵]. این تومورها می‌توانند در هر ناحیه‌ای از مغز ایجاد شوند و در اندازه‌ها و اشکال مختلف باشند. علاوه بر این، سختی بافت یک تومور ممکن است با بافت سالم و حیاتی مغز منطبق باشد و معمولاً در تضاد و تفاوتی ضعیف قرار گیرد. در نتیجه، تمایز بافت سالم از تومور چالش برانگیز است [۲۱].

روش‌های ناحیه‌بندی خودکار^{۱۲} و توانایی فناوری‌های تصویربرداری در کنار یادگیری عمیق به متخصصان اجازه می‌دهد بسیار موثرتر و دقیق‌تر کار کنند. در نتیجه این امر، ناحیه‌بندی دقیق و سریع تومورهای مغزی

¹Medical Imaging²Machine Learning (ML)³Deep Learning (DL)⁴Image Segmentation (IS)⁵Brain Tumor⁶Diagnosis⁷Glioma⁸Glial Cells⁹Modalities¹⁰Magnetic Resonance Imaging (MRI)¹¹3D¹²Automatic Segmentation

می‌تواند به پزشکان در درمان ایمن تومورها، به ویژه در حین جراحی، بدون به خطر انداختن نواحی سالم مغز کمک کند و در نهایت باعث افزایش نرخ بقای بیماران و بهبود کیفیت زندگی آنان شود. در زمینه ناحیه‌بندی خودکار معماری‌های مختلف شبکه‌های عصبی پیچشی^{۱۳} بر روی تصاویر سه‌بعدی پزشکی و MRI در حال استفاده شدن است. پیش از این، استفاده از شبکه‌های عمیق به دلیل عدم وجود داده‌های کافی پزشکی و حجم عظیم قدرت پردازش، کار دشواری بود. انتشار مجموعه دادگانی مانند مجموعه داده ناحیه‌بندی تومور مغزی^{۱۴} که دارای تصاویر MRI سه‌بعدی شامل برجسب^{۱۵} می‌باشد، این کار را آسان‌تر کرد. همچنین، با در دسترس قرار گرفتن واحدهای پردازش گرافیکی^{۱۶} قوی‌تر امکان پردازش سریع‌تر داده‌ها به وجود آمد. این پیشرفت‌ها باعث می‌شود محققان به کار در این حوزه برای کمک به پزشکان توجه بیشتری داشته باشند.

با این حال، از آنجایی که سلول‌های تومور می‌توانند در هر جایی از بافت مغز ایجاد شوند و می‌توانند از نظر اندازه، ظاهر و شکل متفاوت باشند، ناحیه‌بندی خودکار تومور مغزی و مناطق فرعی و مرزی آن دشوار است [۱۲]. همچنین با توجه به این که در زمینه پزشکی جان و سلامت انسان‌ها مورد بحث می‌باشد، نیاز به دخالت و تعامل پزشک در فرایند ناحیه‌بندی خودکار الزامی می‌باشد و در نتیجه فرایند ناحیه‌بندی نیمه‌خودکار^{۱۷} مطرح می‌شود.

ناحیه‌بندی تعاملی تصویر^{۱۸} روشی است که در آن با دریافت راهنمایی‌هایی از سمت کاربر درباره شی هدف که با عنوان پیش‌زمینه^{۱۹} شناخته می‌شود و همچنین راهنمایی‌هایی درباره تصویر پس‌زمینه^{۲۰}، می‌توان با دقت و سرعت بالاتری ناحیه مورد علاقه^{۲۱} را ناحیه‌بندی کرد. در این روش‌ها تعامل کاربر^{۲۲} می‌تواند تا زمان رسیدن به نتیجه مطلوب ادامه یابد. یکی از اهداف ناحیه‌بندی تعاملی، کمینه کردن تلاش کاربر در عین افزایش دقت است. به نظر می‌رسد استفاده از یادگیری عمیق و شبکه‌های CNN و همچنین ارائه روش‌های خلاقانه برای تزریق دانش کاربر به شبکه یکی از بهترین راه‌های ممکن برای دستیابی به این هدف است.

هدف از این پژوهش ارائه روشی مبتنی بر یادگیری عمیق برای ناحیه‌بندی تعاملی تومور مغزی در مدالیته‌های

¹³Convolutional Neural Network

¹⁴Brain Tumor Segmentation (BraTS)

¹⁵Label

¹⁶Graphics Processing Unit (GPU)

¹⁷Semi-Automatic

¹⁸Interactive Image Segmentation (IIS)

¹⁹Foreground (FG)

²⁰Background (BG)

²¹Region of Interest (ROI)

²²User Interaction

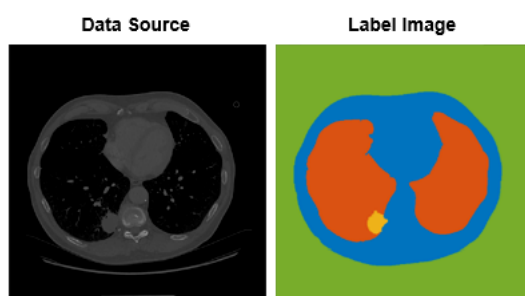
مختلف تصاویر MRI می‌باشد که دارای دقت و سرعت بالا و نیازمند هزینه محاسباتی کمی باشد. افزون‌بر این، یکی از نکات مورد توجه در این پژوهش کمینه کردن تلاش مورد نیاز از سمت کاربر می‌باشد. همچنین بررسی می‌شود که آیا روش ارائه شده می‌تواند با آموزش دیدن روی یک مدالیت و یک وظیفه^{۲۳} خاص بر روی مدالیت و وظایف دیگر نیز نتیجه مطلوب تولید کند یا خیر.

در ادامه این پژوهش و در فصل دوم ابتدا ادبیات موضوع ناحیه‌بندی تعاملی توضیح داده می‌شود و پس از آن ایده‌ها و کارهای مرتبط با موضوع بررسی می‌شود. در فصل سوم جزئیات الگوریتم پیشنهادی شامل معماری شبکه، تابع خطا^{۲۴}، معیارهای ارزیابی^{۲۵} و موارد مشابه ذکر می‌شود. در فصل چهارم نتایج آزمایش‌های مختلف به همراه تحلیل، تفسیر و خروجی تولید شده ارائه می‌شود. در فصل پنجم و آخر نیز به جمع‌بندی و ارائه پیشنهادهایی برای ادامه کار پرداخته می‌شود.

²³Task²⁴Loss Function²⁵Evaluation Metric

۰-۲. ناحیه‌بندی تصویر

ناحیه‌بندی تصویر یکی از وظایف مهم و پایه‌ای در بینایی کامپیوتر^{۲۶} می‌باشد و در زمینه‌های مختلفی کاربرد دارد. هدف اصلی ناحیه‌بندی تصویر، تفکیک یک تصویر به نواحی مشابه است. این نواحی مشابه دارای ویژگی‌های مشترکی نظیر موقعیت مکانی، رنگ، شکل، بافت و حرکت (در حالت ناحیه‌بندی فیلم) هستند [۲۲]. به بیان دقیق‌تر ناحیه‌بندی یک تصویر فرایندی است که در آن به هر پیکسل^{۲۷} تصویر یک برچسب اختصاص داده می‌شود. اختصاص این برچسب‌ها به گونه‌ای است که پیکسل‌های دارای برچسب یکسان دارای ویژگی‌های مشابه و مشترکی هستند که به آن‌ها اشاره شد. شبیه‌سازی سیستم ادراکی انسان برای تفکیک یک تصویر به نواحی منطقی یکی از چالش‌های اساسی در حوزه بینایی کامپیوتر بوده و از روزهای ابتدایی این عرصه مورد مطالعه گسترده قرار گرفته است. دلیل این مطالعات گسترده، کاربردهای گسترده ناحیه‌بندی تصویر می‌باشد. برای مثال در زمینه پزشکی می‌توان از ناحیه‌بندی تصویر برای پردازش تصاویر Computed Tomography (CT) و MRI استفاده کرد و نواحی دارای آسیب و تومور را یافت و از آن برای تشخیص درجه بیماری و عمل جراحی استفاده نمود.



شکل ۱: مثالی از ناحیه‌بندی تصویر پزشکی

۰-۲-۱. ناحیه‌بندی معنایی

در ناحیه‌بندی معنایی^{۲۸} به تمامی پیکسل‌هایی که مربوط به یک کلاس^{۲۹} می‌باشند یک برچسب ثابت اختصاص می‌دهیم [۶]. برای مثال اگر در یک تصویر سه گربه در یک مزرعه وجود داشته باشد، به پیکسل‌های

²⁶Computer Vision (CV)

²⁷Pixel

²⁸Semantic Segmentation

²⁹Class

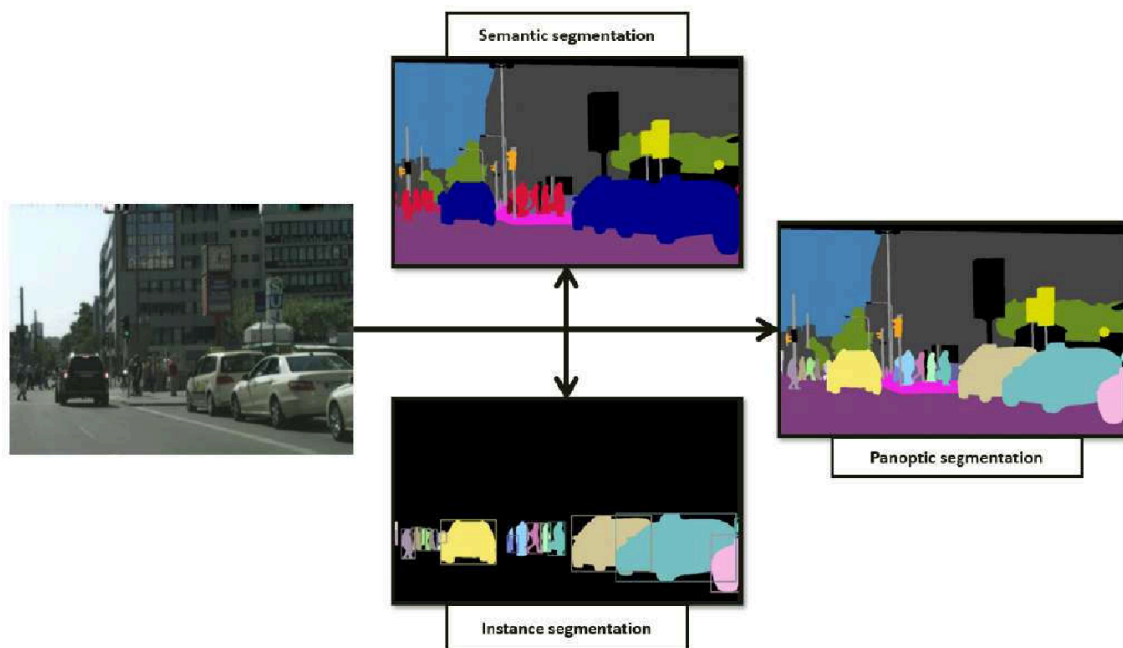
مربوط به گره‌ها برچسب یک و به پیکسل نواحی مزرعه پیکسل صفر اختصاص داده می‌شود.

۲-۲-۰ ناحیه‌بندی نمونه‌ای

در ناحیه‌بندی نمونه‌ای^{۳۰} به تمامی پیکسل‌های هر نمونه از شی مورد نظر یک برچسب متمایز اختصاص داده می‌شود. در واقع هدف شناسایی نمونه‌های موجود از شی مورد نظر می‌باشد [۳۴]. برای مثال در تصویر مثال بالا به پیکسل‌های مربوط به گره اول برچسب صفر، گره دوم برچسب یک و گره سوم برچسب دو اختصاص داده می‌شود.

۳-۲-۰ ناحیه‌بندی سراسرنما

ناحیه‌بندی سراسرنما^{۳۱} ترکیب دو حالت بالا می‌باشد. در این حالت مشابه ناحیه‌بندی معنایی، به هر پیکسل برچسب کلاس مورد نظر اختصاص داده می‌شود. همچنین مشابه ناحیه‌بندی نمونه‌ای، در این حالت نمونه‌های موجود از یک کلاس متمایز می‌شوند [۱۴].



شکل ۲: انواع حالت‌های ناحیه‌بندی [۴]

³⁰Instance Segmentation

³¹Panoptic Segmentation

۳-۰ ناحیه‌بندی تعاملی تصویر

علاوه بر حالت‌های ناحیه‌بندی که در قسمت قبل اشاره شد، کارهای دیگر (و کمتری) روی روش‌های ناحیه‌بندی تعاملی تصویر انجام شده است. ناحیه‌بندی تعاملی تصویر یا ناحیه‌بندی با ناظر^{۳۲} یا ناحیه‌بندی نیمه‌خودکار به معنی استخراج کردن ناحیه‌ای از یک تصویر یا یک شی هدف یا FG از BG به کمک دانش قبلی فراهم شده توسط تعامل کاربر است [۲۵]. تعامل کاربر با مدل می‌تواند به صورت‌های گوناگونی باشد. برای مثال می‌توان از نقاطی^{۳۳} یا از خطوطی^{۳۴} برای نشان دادن شی مورد نظر و یا تصویر BG استفاده کرد. همچنین می‌توان از یک جعبه مرزی^{۳۵} یا یک چند ضلعی^{۳۶} برای مشخص کردن نواحی مورد علاقه یا ROI استفاده کرد تا با مدل تعامل داشت و به فرایند ناحیه‌بندی خودکار کمک کرد. این تعاملات کاربر علاوه بر بهبود نتایج باعث کاهش زمان اجرای ناحیه‌بندی نیز می‌شود و می‌توان گفت ناحیه‌بندی تعاملی دارای دقت و سرعت بیشتری نسبت به ناحیه‌بندی خودکار است [۸]. در واقع در بسیاری از کاربردهای بینایی کامپیوتر مانند پزشکی برای رسیدن به ناحیه‌بندی دقیق نیاز به این تعامل و دخالت کاربر وجود دارد.

۴-۰ دسته‌بندی روش‌های ناحیه‌بندی تعاملی

روش‌های ناحیه‌بندی تعاملی را می‌توان به طریق‌های مختلفی دسته‌بندی کرد که در ادامه به جزئیات دو مورد از این طریق‌ها اشاره می‌شود.

۱-۴-۰ دسته‌بندی بر اساس نحوه تعامل کاربر

نحوه تعامل کاربر با مدل را می‌توان از دو جنبه مورد بررسی قرار داد. در جنبه اول تعامل کاربر با مدل می‌تواند مبتنی بر دو روش Seed و ROI باشد [۳۱]. در جنبه دوم تعامل کاربر با مدل می‌تواند به صورت فعال^{۳۷} و یا منفعل^{۳۸} باشد [۳]. از نظر جنبه اول تعامل کاربر با مدل برای مشخص کردن BG و FG می‌تواند به

^{۳۲}Supervised Segmentation

^{۳۳}Points

^{۳۴}Scribbles

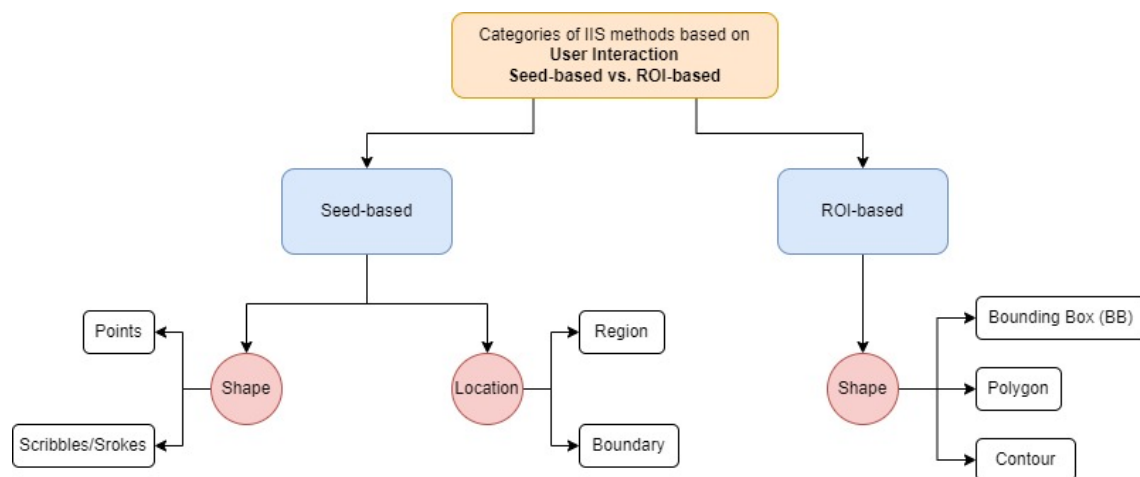
^{۳۵}Bounding Box (BB)

^{۳۶}Polygon

^{۳۷}Active

^{۳۸}Passive

وسیله مجموعه‌ای از نقاط، پاره خط‌ها و خطوط منفصل باشد. در این حالت تعامل کاربر با مدل مبتنی بر Seed نامیده می‌شود. نحوه دیگر تعامل کاربر با مدل با ترسیم محیط‌های تقریبی نظیر Bounding Box یا چند ضلعی و یا هر گونه منحنی بسته صورت می‌گیرد تا شی هدف از BG جدا شود. این محیط‌ها برای مشخص کردن ROI توسط کاربر ترسیم می‌شوند، به همین دلیل در این حالت تعامل کاربر با مدل مبتنی بر ROI نامیده می‌شود.



شکل ۳: انواع حالت‌های تعامل مبتنی بر Seed و ROI

۴-۱-۱ تعامل مبتنی بر Seed

مدل‌های اولیه مبتنی بر Seed نیازمند نقاطی در مرزهای Boundary شی مورد نظر بود. این نقاط تحت نام‌های گوناگونی شناخته می‌شوند که دو مورد از این نام‌ها عبارتند از Seed Points و Anchor Points. همچنین این روش با نام ناحیه‌بندی تعاملی مبتنی بر Boundary Seed شناخته می‌شود. زیردسته دیگری از روش‌های مبتنی بر Seed متکی بر ناحیه^{۳۹} هستند. در این گروه برای کاربر محدودیت خاصی در انتخاب محل نقاط اعمال نمی‌شود. در این حالت خروجی ناحیه‌بندی به تعداد نقاط کاربر بسیار حساس می‌شود و مجموعه نقاط مختلف باعث تولید خروجی‌های بسیار متفاوتی می‌شوند. در این حالت Seedها می‌توانند تنها تعدادی نقطه باشند و یا قطعه‌های پیوسته‌ای از نقاط (Scribble, Stroke) باشند. همچنین می‌توان با ترکیب Seedهای Boundary و Region به یک حالت هیبریدی^{۴۰} دست یافت [۳۰]. در این حالت مدل با استفاده از Seedهای ابتدایی مرزی

^{۳۹}Region

^{۴۰}Hybrid

محیط شکل را تا حدودی مشخص می‌کند و سپس با استفاده از Seedهای ناحیه‌ای خروجی نهایی بهبود یافته را تولید می‌کند.

۴-۱-۲. تعامل مبتنی بر ROI

در این روش دانش قبلی از محل FG و BG به کمک پیکسل‌های درون و برون ROI به دست می‌آید. روش GrabCut [۲۸] معروف‌ترین و پراستفاده‌ترین روش ناحیه‌بندی تعاملی است که به کاربر امکان رسم Bounding Box برای جداسازی FG از BG را می‌دهد. عملکرد روش‌های مبتنی بر ROI به کیپ^{۴۱} بودن منحنی ترسیم شده توسط کاربر وابسته است. اگر منحنی ترسیم شده توسط کاربر ناحیه FG را به صورت کیپ جدا نکند عملکرد مدل بسیار پایین خواهد آمد. در نتیجه این امر مفهوم جدیدی به نام Tightness در این دسته از روش‌ها به وجود آمده است.

به صورت کلی مشخص کردن ROI با ترسیم منحنی دور آن برای کاربر راحت‌تر است اما در این حالت دقت و عملکرد مدل بسیار وابسته به میزان Tightness خواهد بود. از سوی دیگر الگوریتم‌های مبتنی بر Seed با وجود تعداد Seed کافی می‌توانند FG با شکل‌های پیچیده‌ای را استخراج کنند. به صورت کلی روش ROI نیازمند تعداد دفعات تعامل بیشتری نسبت به روش Seed می‌باشد [۱۰]. نکته اصلی و چالش کلیدی که بدون توجه به روش Seed یا ROI مطرح است، به دست آوردن دقت بالا در ازای کمترین تلاش کاربر است. این امر در کارهای اخیر که بر اساس یادگیری عمیق می‌باشد پیشرفت خوبی داشته است.

خلاصه‌ای از دو حالت تعامل مبتنی بر Seed و ROI و انواع زیرحالت‌ها در تصویر ۴ آورده شده است. تصویر (a) بیانگر حالت Boundary seeds، تصویر (b) بیانگر Region point seeds، تصویر (c) بیانگر Region scribble seeds، تصویر (d) بیانگر ROI contour، تصویر (e) بیانگر ROI BB و تصویر (f) بیانگر حالت ROI BB tight می‌باشد.

⁴¹Tight



شکل ۴: مثال‌هایی از انواع روش‌های تعامل کاربر [۲۵]

۳-۱-۴-۰ تعامل فعال و منفعل

از جنبه دوم بعضی از مدل‌های ناحیه‌بندی تعاملی امکان تعامل فعالانه و کمک برخط ^{۴۲} را در طول فرایند ناحیه‌بندی برای کاربر فراهم می‌کنند. این روش‌ها در دسته Active قرار می‌گیرند. برای مثال ابتدا یک ناحیه‌بندی تخمینی توسط مدل خروجی داده می‌شود سپس مدل قسمت‌هایی را که با اطمینان کمی ناحیه‌بندی کرده، به کاربر اعلام می‌کند و کاربر با تعامل با مدل به قصد و نتیجه مطلوب می‌رسد. این تعامل ممکن است در قالب جواب بله یا خیر کاربر به قسمت‌های نامطمئن ناحیه‌بندی باشد. یا ممکن است در ابتدا مدل چندین خروجی تولید کند و با پرسیدن سوالات دودویی ^{۴۳} (مانند مسابقه بیست سوالی) از کاربر و پردازش پاسخ‌ها نتیجه بهبود یافته‌تری را تولید کند [۲۹]. در روش Active تعامل تا زمانی ادامه می‌یابد که یا کاربر از نتیجه راضی باشد یا مدل به یک آستانه ^{۴۴} مشخص برسد.

بر خلاف دسته قبلی، تعامل می‌تواند به صورت منفعل یا برون‌خط ^{۴۵} صورت گیرد که این دسته به عنوان

⁴²Online

⁴³Binary

⁴⁴Threshold

⁴⁵Offline

تعامل Passive شناخته می‌شود [۳]. در حقیقت اکثریت روش‌های ناحیه‌بندی تعاملی به صورت Passive می‌باشند. در این حالت هرگاه کاربر Seed یا ROI جدیدی ایجاد کند یا موارد قبلی را ویرایش کند، خروجی مدل نیز به‌روز می‌شود.

۲-۴-۰ دسته‌بندی بر اساس روش‌شناسی

الگوریتم‌های ناحیه‌بندی تعاملی را می‌توان بر اساس روش‌شناسی^{۴۶} به دو دسته کلی تقسیم کرد. روش‌های سنتی و قدیمی‌تر مانند الگوریتم‌های مبتنی بر Contour، الگوریتم‌های مبتنی بر برش گراف^{۴۷}، مبتنی بر الگوریتم ولگشت^{۴۸} و مبتنی بر ناحیه مثل (RG) Region Growing و Region Merging در دسته اول قرار می‌گیرند. روش‌های جدیدتر که مبتنی بر الگوریتم‌های یادگیری عمیق هستند در دسته دوم قرار می‌گیرند. در این دسته، از شبکه‌های عصبی پیچشی استفاده گسترده‌ای می‌شود.

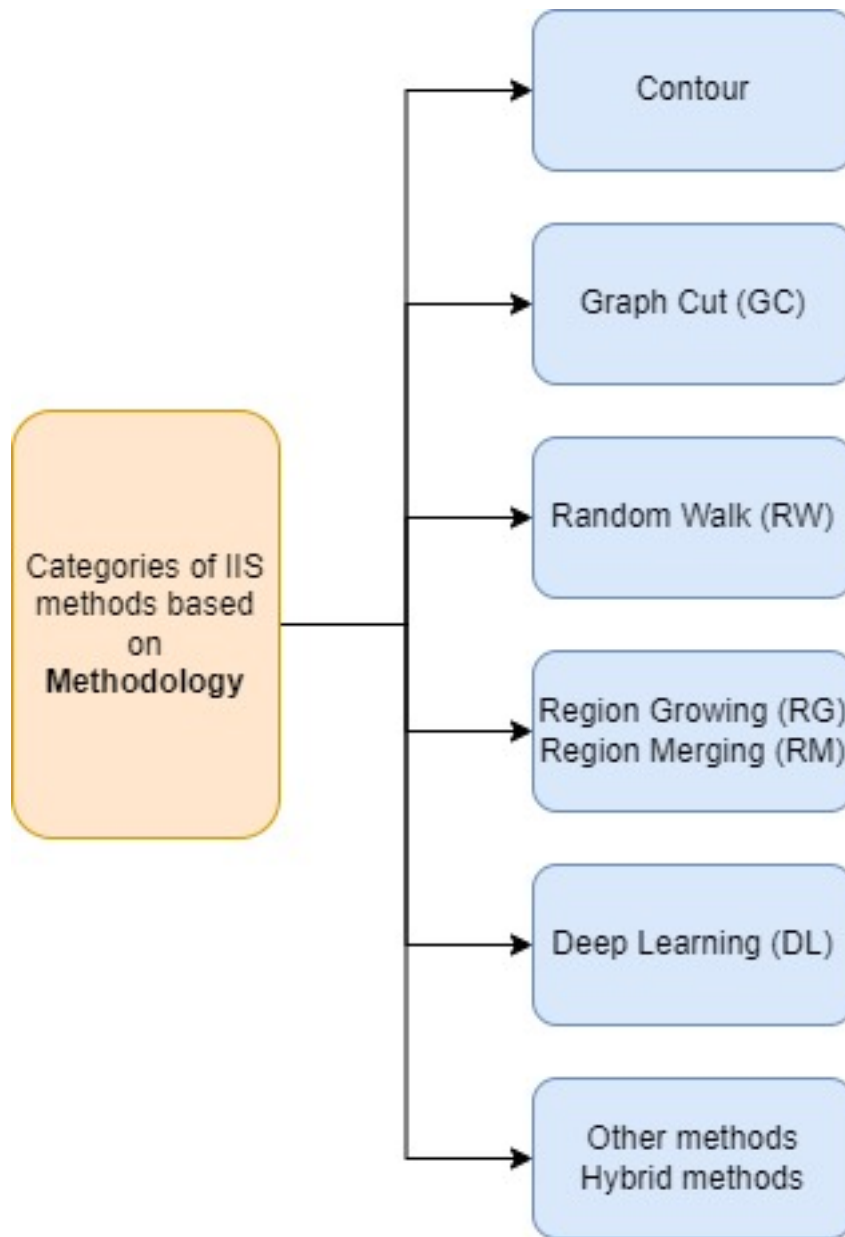
البته در برخی کارها استفاده از هر دو نوع روش مشاهده می‌شود. در این کارها سعی می‌شود از برتری هر روش به خوبی استفاده کرد. برای مثال می‌توان از یادگیری عمیق برای تولید ناحیه‌بندی اولیه (بدون تعامل) کاربر استفاده کرد و سپس تعامل‌های کاربر را با روش GC پردازش کرد و نتیجه را بهبود بخشید [۱۷]. روش مبتنی بر یادگیری عمیق ابتدا خروجی نسبتاً خوبی تولید می‌کند اما به دلیل هزینه محاسباتی بالا نمی‌توان آن را به ازای هر تعامل کاربر دوباره اجرا کرد. به همین دلیل در مرحله تعامل کاربر از الگوریتم‌های کم هزینه‌تر سنتی استفاده می‌شود.

با توجه به اینکه تمرکز اصلی این پروژه بر روی ناحیه‌بندی تعاملی با استفاده از یادگیری عمیق می‌باشد، در این قسمت روش‌های سنتی و غیر عمیق به صورت اجمالی بررسی می‌شوند و توضیح مختصری درباره الگوریتم‌های مبتنی بر Contour، GraphCut، Random Walk، Region Growing و Region Merging داده خواهد شد.

⁴⁶Methodology

⁴⁷GraphCut (GC)

⁴⁸Random Walk (RW)



شکل ۵: دسته‌بندی روش‌های ناحیه‌بندی تعاملی بر اساس روش‌شناسی

۴-۲-۱ روش مبتنی بر Contour

در روش‌های مبتنی بر Contour اصل کلیدی مورد توجه، استخراج Contourهای شی هدف با استفاده از ویژگی‌های مربوط به لبه^{۴۹} و دانش قبلی فراهم شده توسط تعامل کاربر می‌باشد. Contourهای شی با محاسبه

^{۴۹}Edge Features

کوتاه‌ترین مسیر^{۵۰} متصل کننده به نقاط Seed استخراج می‌شود. برای محاسبه کوتاه‌ترین مسیر از الگوریتم دایکسترا^{۵۱} یا الگوریتم‌های سریع‌تر استفاده می‌شود.

۲-۲-۴-۰ روش مبتنی بر GraphCut

روش‌های مبتنی بر GC جزو محبوب‌ترین روش‌های حوزه ناحیه‌بندی تعاملی حساب می‌شوند. در این روش‌ها ابتدا یک تابع انرژی تعریف می‌شود که هدف کمینه کردن این تابع می‌باشد^{۵۲}. این کمینه کردن انرژی را می‌توان با حل یک مسئله بیشینه جریان^{۵۳} در گراف تخمین زد [۲]. به همین ترتیب طبق نظریه جریان-بیشینه برش-کمینه^{۵۴} کافی است برش^{۵۵} کمینه در گراف پیدا شود.

روش GrabCut [۲۷] معروف‌ترین الگوریتم ناحیه‌بندی تعاملی مبتنی بر GC به شمار می‌رود. در این روش ابتدا کاربر Bounding Box را دور شی هدف ایجاد می‌کند. سپس الگوریتم با استفاده از Gaussian Mixture Model (GMM) تخمینی از توزیع رنگ‌های FG و BG می‌زند. با استفاده از این تخمین یک Markov Random Field (MRF) روی برچسب پیکسل‌ها ایجاد می‌شود که تابع انرژی آن به گونه‌ای است که ترجیح می‌دهد نواحی متصل دارای برچسب یکسان باشند. در نهایت یک بهینه‌سازی^{۵۶} مبتنی بر GC برای به دست آوردن این برچسب‌ها اجرا می‌شود.

۳-۲-۴-۰ روش مبتنی بر Random Walk

در مدل اصلی و اولیه Random Walk [۲۶] یک گراف بدون جهت^{۵۷} ایجاد می‌شود که رئوس^{۵۸} این گراف بیانگر پیکسل‌های تصویر ورودی و هر یال^{۵۹} آن متصل کننده دو پیکسل (راس) به یکدیگر است. در این گراف وزن هر یال بیانگر احتمال^{۶۰} یک قدم تصادفی روی این دو راس می‌باشد. در این الگوریتم هدف یافتن مجموعه‌ای از احتمالات (x) می‌باشد که هر عضو این مجموعه (x_i) احتمال رخداد یک Random Walk

⁵⁰Shortest Path

⁵¹Dijkstra's Algorithm

⁵²Energy Minimization

⁵³Max Flow Problem

⁵⁴Max-Flow Min-Cut Theorem

⁵⁵Cut

⁵⁶Optimization

⁵⁷Undirected Graph

⁵⁸Nodes

⁵⁹Edge

⁶⁰Probability

است که در نهایت در راس i باقی می‌ماند. راه حل این هدف کمینه کردن تابع زیر است:

$$E(x) = x^T Lx \quad (۱)$$

در این رابطه L بیانگر ماتریس ترکیبی لاپلاسی^{۶۱} است. اگر V_m را مجموعه رئوسی (پیکسل) که توسط کاربر به عنوان Seed انتخاب شده اند در نظر گرفت و V_u را رئوسی که انتخاب نشده اند در نظر گرفت، اشتراک آن دو تهی و اجتماع آن دو مجموعه کل پیکسل‌ها می‌شود. f_j را به عنوان بردار برچسب پیکسل‌ها در نظر می‌گیریم و ماتریس L را به صورت زیر به دو بخش تقسیم می‌کنیم:

$$L = \begin{bmatrix} L_m & B \\ B^T & L_u \end{bmatrix} \quad (۲)$$

در این صورت برای کمینه کردن معادله ۱ کافی است یک سیستم خطی به صورت زیر حل شود:

$$L_u x_u = -B^T f \quad (۳)$$

۴-۲-۴-۰ روش مبتنی بر Region Growing, Merging

ایده اصلی در الگوریتم‌های مبتنی بر Region Growing و Region Merging شروع از ناحیه‌هایی که کاربر در آن Seed قرار داده و ادغام این ناحیه‌های مشابه و مجاور است. در تمامی مراحل ادغام باید به اصل همگن بودن این نواحی توجه شود [۱]. فرایند ادغام نواحی به صورت مکرر انجام می‌شود تا زمانی که تمامی نواحی مطابق مطلوب کاربر برچسب زده شود.

روش‌های مبتنی بر ادغام نواحی از نظر فهم، روش‌های ساده‌ای در نظر گرفته می‌شوند اما دارای محدودیت‌ها و چالش‌های بسیاری می‌باشند. برای مثال تغییر ترتیب ادغام‌ها می‌تواند منجر به تولید نتیجه متفاوتی شود [۱]. همچنین در این روش‌ها خروجی دقیق نیازمند ورودی کافی از سمت کاربر می‌باشد تا نواحی اصلی پوشش داده شود. این امر هنگامی که بخش‌هایی از FG بسیار شبیه به BG است باعث چالش‌های زیادی می‌شود. همچنین عملکرد این روش‌ها در تصاویری که دارای سایه یا کیفیت پایین می‌باشند به اندازه کافی

⁶¹Combinatorial Laplacian Matrix

مناسب نیست [۲۳].

۵-۲-۴-۰ روش‌های یادگیری عمیق و ترکیبی

در سال‌های اخیر، شبکه‌های CNN در بسیاری از کاربردهای بینایی کامپیوتر از جمله ناحیه‌بندی تصویر به موفقیت زیادی دست یافته‌اند. همچنین بسیاری از روش‌های ناحیه‌بندی تعاملی مبتنی بر معماری‌های عمیق پیشنهاد شده‌اند که عملکرد بهتری نسبت به مدل‌های سنتی دارند. در ادامه به چند روش عمیق در این حوزه اشاره مختصر می‌شود. همچنین در بخش کارهای مرتبط چند روش عمیق بررسی دقیق‌تر می‌شود.

شبکه Deep Extreme Cut یا به اختصار DEXTR [۱۸] یک معماری مبتنی بر CNN را برای ناحیه‌بندی تعاملی با استفاده از نقاط مفراط^{۶۲} در بالا، پایین، چپ و راست شی مورد نظر ارائه می‌کند. این نقاط در کنار کانال‌های RGB^{۶۳} به عنوان ورودی به مدل داده می‌شوند. همچنین از شبکه ResNet-101 [۹] به عنوان Backbone در این مدل استفاده شده است.

روش Deep Interactive Object Selection [۳۲] کلیک‌های^{۶۴} مثبت (روی شی هدف) و منفی (بیرون شی هدف) ارائه شده توسط کاربر را به نقشه‌های^{۶۵} مبتنی بر فاصله اقلیدسی^{۶۶} مجزا تبدیل می‌کند. سپس این نقشه‌ها با کانال‌های RGB تصویر ورودی ترکیب می‌شوند تا یک جفت^{۶۷} تصویر-تعامل ایجاد شود. سپس این جفت وارد یک شبکه کاملاً پیچشی^{۶۸} می‌شود تا ناحیه‌بندی اولیه صورت گیرد. در نهایت یک مرحله پالایش^{۶۹} با استفاده از بهینه‌سازی GraphCut اعمال می‌شود و خروجی نهایی تولید می‌شود.

روش Deep GrabCut for Object Selection [۳۳] به جای کلیک، یک Bounding Box را از کاربر ورودی می‌گیرد. همچنین این مدل به صورت یکپارچه^{۷۰} آموزش داده می‌شود و نیازی به مرحله پس‌پردازش^{۷۱} برای پالایش خروجی‌های FCN ندارد. روش SeedNet [۱۵] یک سیستم ناحیه‌بندی تعاملی را فراهم می‌کند که به کاربر کمک می‌کند تا یک شی را با استفاده از یک چارچوب تولید خودکار Seed ناحیه‌بندی کند. در این

⁶² Extreme Points

⁶³ Channels

⁶⁴ Click

⁶⁵ Map

⁶⁶ Euclidean Distance

⁶⁷ Pair

⁶⁸ Fully Convolutional Network (FCN)

⁶⁹ Refinement

⁷⁰ End-to-End

⁷¹ Post-Processing

سیستم تولید خودکار Seed با استفاده از یادگیری تقویتی عمیق^{۷۲} صورت می‌گیرد.

با رواج یادگیری عمیق و کیفیت بالای نتایج ناحیه‌بندی که توسط روش‌های مبتنی بر یادگیری عمیق تنها با چند کلیک به دست می‌آید، تحقیقات در این زمینه بسیار امیدوار کننده است. با این حال، یک مشکل رایج در اکثر روش‌های مبتنی بر یادگیری عمیق که برخی کارها تلاش کرده اند بر آن غلبه کنند، این است که نتایج عالی بر روی مجموعه آموزشی^{۷۳} به دست می‌آید، در حالی که برای کلاس‌های دیده نشده توسط مدل، نتایج ضعیف است.

علاوه بر مورد بالا، نیازمند بودن این روش‌ها به تعداد داده آموزشی زیاد برای آموزش مدلی با تعداد پارامترهای فراوان، عملی بودن استفاده از چنین روش‌هایی را برای برنامه‌های کاربردی و صنعتی مورد تردید قرار می‌دهد. البته با ظهور فزاینده مجموعه دادگان^{۷۴} و همچنین استفاده از روش‌های داده افزایی^{۷۵}، روش‌های مبتنی بر یادگیری عمیق با موفقیت در صنعت و کارخانه‌های هوشمند مورد استفاده قرار گرفته اند.

۵-۰ معیارهای ارزیابی

برای ارزیابی کمی و قابل مشاهده^{۷۶} روش‌های ناحیه‌بندی از جمله روش‌های ناحیه‌بندی تعاملی، یک مقایسه میان نتایج به دست آمده از مدل و برچسب‌های مشخص شده توسط انسان یا به عبارت دیگر حقیقت مبنا^{۷۷} مورد نیاز است. این مقایسه بایستی بر اساس معیاری صورت گیرد که قادر باشد شباهت GT و خروجی مدل را اندازه بگیرد. در این بخش، خلاصه‌ای از معیارهای مختلف رایج مورد استفاده برای چنین ارزیابی‌ای گزارش داده می‌شود. در روابط زیر عبارت S بیانگر خروجی مدل ناحیه‌بند و عبارت G بیانگر خروجی صحیح می‌باشد.

⁷²Deep Reinforcement Learning (Deep RL)

⁷³Training Set

⁷⁴Dataset

⁷⁵Data Augmentation

⁷⁶Objective

⁷⁷Ground Truth (GT)

Error Rate (ERR) ۱-۵-۰

نرخ خطا بیانگر نسبت یا درصد پیکسل‌های درون ROI که به صورت اشتباه برچسب‌گذاری شده اند، می‌باشد و رابطه آن به صورت زیر می‌باشد:

$$ERR = \frac{|S \cup G - G|}{\text{Number of pixels in ROI}} \quad (۴)$$

True Positive Rate (TPR) ۲-۵-۰

این معیار با نام Recall به معنی یادآوری یا فراخوانی نیز شناخته می‌شود و بیانگر نسبت پیش‌بینی‌های ^{۷۸} صحیح مدل در یک کلاس مثبت ^{۷۹} می‌باشد. رابطه Recall به صورت زیر تعریف می‌شود:

$$R = \frac{|S \cap G|}{|G|} \quad (۵)$$

دلیل استفاده از این معیار به جای معیار دقت ساده ^{۸۰} این است که ممکن است برای مثال ۹۸ درصد پیکسل‌های تصویر مغز فردی سالم باشد، در این صورت اگر مدل تمام پیکسل‌ها را سالم پیش‌بینی کند، دارای دقت ۹۸ درصد است که عدد خوبی به نظر می‌رسد. اما در کاربرد پزشکی ما به دنبال آن ۲ درصد پیکسل دارای تومور می‌باشیم و این ۲ درصد برای متخصصان دارای اهمیت بسیار زیادی می‌باشد، چرا که با یافتن آن می‌توان برای درمان بیماری اقدام کرد.

Precision ۳-۵-۰

به معنی صراحت و درستی می‌باشد و بیانگر نسبت پیش‌بینی‌های صحیح در کل پیش‌بینی‌های مربوط به یک کلاس مثبت است. رابطه آن هم به صورت زیر می‌باشد:

⁷⁸ Prediction⁷⁹ Positive⁸⁰ Accuracy

$$P = \frac{|S \cap G|}{|S|} \quad (۶)$$

F1-Score ۴-۵-۰

بیانگر میانگین همساز^{۸۱} دو معیار Recall و Precision می‌باشد و این دو معیار را به صورت زیر با هم ترکیب می‌کند:

$$F = \frac{2 \times R \times P}{R + P} \quad (۷)$$


Intersection over Union (IoU) ۵-۵-۰

بیانگر نسبت تعداد پیکسل‌هایی که به درستی به عنوان FG برچسب‌گذاری شده اند (اشتراک)، به تعداد پیکسل‌هایی که در S یا G (اجتماع) دارای برچسب FG هستند، می‌باشد. مقدار آن از تقسیم این اشتراک و اجتماع به صورت زیر به دست می‌آید:

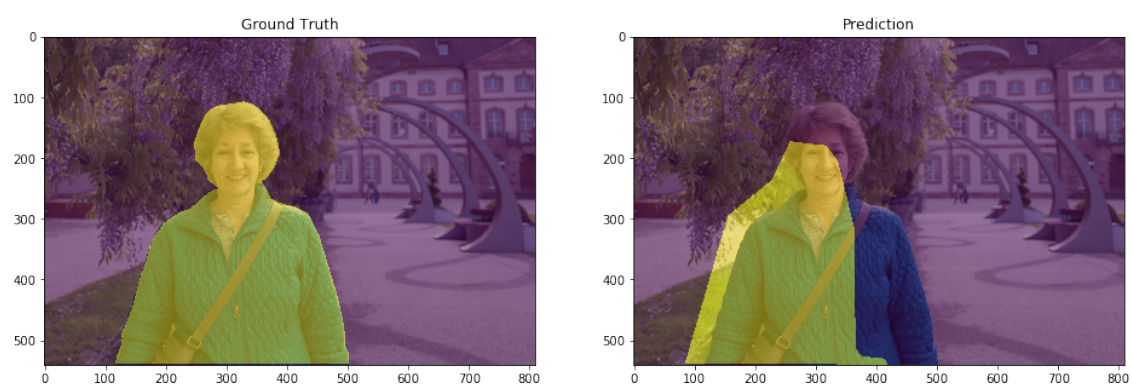
$$IoU = \frac{|S \cap G|}{|S \cup G|} \quad (۸)$$

این معیار با نام‌های دیگری نظیر Jaccard Index و Overlap Coefficient نیز شناخته می‌شود و جزو رایج‌ترین معیارهای ارزیابی مدل‌های ناحیه‌بندی می‌باشد.

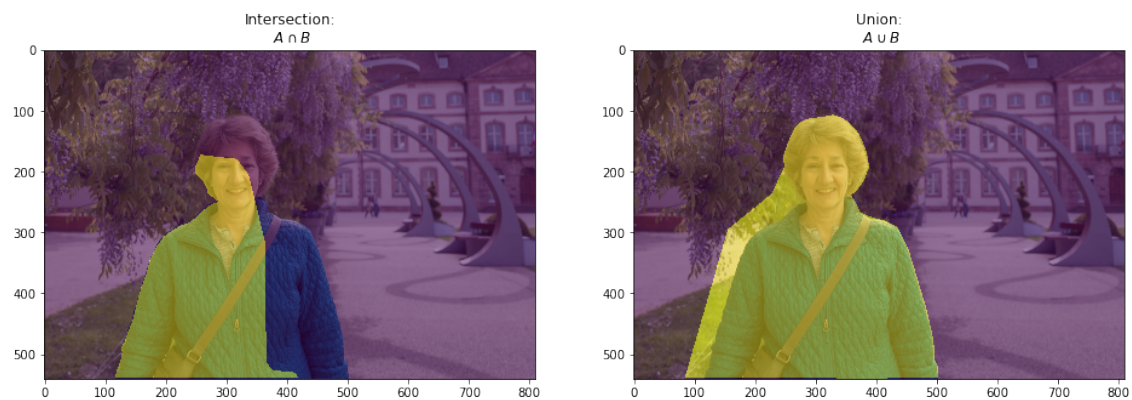
⁸¹Harmonic Mean

$$IoU = \frac{\text{Area of Overlap}}{\text{Area of Union}}$$


شکل ۶: نمایش ریاضی معیار IoU



شکل ۷: مثالی از تصویر S و G

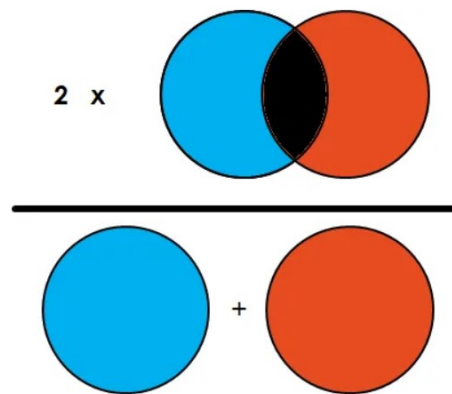


شکل ۸: مثالی از ناحیه اشتراک و اجتماع

۵-۶. Dice Similarity Coefficient (DSC)

این معیار مشابه معیار IoU می‌باشد با این تفاوت که در مخرج رابطه به جای عملیات اجتماع، از مجموع تمامی پیکسل‌های FG درون S و G استفاده می‌شود. علاوه بر آن، رابطه در عدد ۲ نیز ضرب می‌شود.

$$DSC = 2 \times \frac{|S \cap G|}{|S| + |G|} \quad (9)$$



شکل ۹: نمایش ریاضی معیار Dice

از تعریف و رابطه این معیار می‌توان متوجه شد که این معیار معادل معیار F1 می‌باشد. همچنین برخلاف معیار IoU از معیار Dice می‌توان به عنوان تابع خطا برای آموزش شبکه استفاده کرد. دلیل این امر این است که معیار Dice بر خلاف معیار IoU مشتق پذیر می‌باشد. رابطه تابع خطا Dice را می‌توان به صورت زیر تعریف کرد:

$$DiceLoss = 1 - DSC \quad (10)$$

۷-۵-۰ Variation of Information (VI)

این معیار فاصله بین دو ناحیه‌بندی S و G را با استفاده از میانگین آنتروپی شرطی^{۸۲} محاسبه می‌کند. همچنین این معیار بر پایه علم نظریه اطلاعات^{۸۳} طراحی شده است و بیشتر در مسائل خوشه‌بندی^{۸۴} کاربرد دارد و در مسائل ناحیه‌بندی به ندرت مورد استفاده قرار می‌گیرد [۱۹].

۸-۵-۰ Number of Clicks (NoC)

این معیار بیانگر میانگین تعداد کلیک مورد نیاز از سمت کاربر برای رسیدن به IoU یا Dice مورد نظر می‌باشد. برای مثال عبارت زیر:

$$NoC(Dice = 80\%) = 7.87 \quad (11)$$

به این معنی است که برای رسیدن معیار Dice به مقدار ۸۰ درصد، کاربر باید به صورت میانگین ۷/۸۷ کلیک (تعامل کلیک محور) انجام دهد.

۹-۵-۰ Number of Failures (NoF)

اگر فرض کنیم مدل دارای محدودیت max_clicks روی تعداد کلیک‌های کاربر باشد، در این صورت معیار NoF بیانگر تعداد نمونه‌هایی^{۸۵} است که با حداکثر max_clicks کلیک همچنان به IoU یا Dice مورد نظر نرسیده‌اند. این نمونه‌ها اصطلاحاً نمونه‌های شکست خورده^{۸۶} نامیده می‌شوند. برای مثال عبارت زیر:

$$NoF_{max-clicks=20}(Dice = 85\%) = 851 \quad (12)$$

⁸²Conditional Entropy

⁸³Information Theory

⁸⁴Clustering

⁸⁵Samples

⁸⁶Failed Samples

به این معنی است که ۸۵۱ نمونه حتی با وجود ۲۰ کلیک، از رسیدن معیار Dice به مقدار ۸۵ درصد، ناتوان بوده اند.

۱۰-۵-۰ Time per Click (TPC)

این معیار بیانگر میانگین زمانی است که مدل پس از دریافت کلیک جدید، خروجی جدید را تولید می‌کند. در واقع این معیار معادل زمان پاسخ^{۸۷} مدل به ازای هر کلیک کاربر می‌باشد. همچنین واحد این معیار ثانیه می‌باشد.

۱۱-۵-۰ mDice و mIoU

بیانگر میانگین IoU یا Dice می‌باشد که مدل با تعداد مشخصی کلیک به ازای هر نمونه، به آن دست می‌یابد. برای مثال عبارت زیر:

$$mDice_{C=20} = 62\% \quad (۱۳)$$

به این معنی است که مدل پس از دریافت ۲۰ کلیک به ازای هر نمونه، به صورت میانگین به Dice برابر ۶۲ درصد، دست یافته است.

۶-۰. مجموعه دادگان BraTS [۲۰]

۱-۶-۰ مقدمه

تومورهای مغزی یکی از تهدیدکننده‌ترین بیماری‌ها هستند که عامل مرگ و میر قابل توجهی در سراسر جهان هستند. تشخیص زودهنگام و تشخیص دقیق تومورهای مغزی برای درمان موثر و بهبود نرخ بقا ضروری است. تکنیک‌های تصویربرداری پزشکی مانند MRI نقش حیاتی در تشخیص تومورهای مغزی دارند. MRI

⁸⁷Response Time

امکان تصویربرداری غیرتهاجمی^{۸۸} و دقیق از مغز را فراهم می‌کند و تصاویر سه‌بعدی با وضوح بالا از ساختار مغز و هرگونه ناهنجاری ارائه می‌دهد. با این حال، ناحیه‌بندی دستی تومورهای مغزی از تصاویر MRI یک کار چالش برانگیز و وقت‌گیر است. بنابراین، ناحیه‌بندی خودکار تومورهای مغزی از تصاویر MRI یک حوزه تحقیقاتی حیاتی در تصویربرداری پزشکی است.

مجموعه داده ناحیه‌بندی تومور مغزی یا به اختصار BraTS مجموعه‌ای از داده‌های در دسترس عموم برای توسعه و ارزیابی الگوریتم‌های ناحیه‌بندی خودکار تومور مغزی است. این مجموعه داده به طور گسترده برای تحقیق و محک^{۸۹} در زمینه تصویربرداری پزشکی، به ویژه در توسعه مدل‌های ناحیه‌بندی تومور مغزی مبتنی بر یادگیری عمیق، استفاده شده است. در این قسمت، مروری بر مجموعه داده BraTS از جمله ویژگی‌ها، چالش‌ها و کاربردهای بالقوه آن ارائه می‌شود.

۰-۶-۲ ویژگی‌های BraTS

مجموعه داده BraTS یک مجموعه داده در دسترس عموم است که اسکن‌های^{۹۰} MRI تومور مغزی را با حاشیه‌نویسی^{۹۱} دستی ارائه می‌دهد. این مجموعه داده در قالب^{۹۲} NIfTI که مخفف عبارت Neuroimaging Informatics Technology Initiative می‌باشد ارائه می‌شود. نسخه سال ۲۰۲۰ این مجموعه داده شامل ۶۶۰ اسکن MRI از تومورهای مغزی است که به مجموعه‌های آموزشی^{۹۳} (۳۶۹)، اعتبارسنجی^{۹۴} (۱۲۵) و آزمایش^{۹۴} (۱۶۶) تقسیم می‌شود. در جدول زیر خلاصه‌ای از توزیع مجموعه‌های آموزشی، اعتبارسنجی و آزمایش این مجموعه داده در سال‌های مختلف آورده شده است.

⁸⁸Non Invasive

⁸⁹Benchmark

⁹⁰Scan

⁹¹Annotation

⁹²Format

⁹³Validation Set

⁹⁴Testing Set

جدول ۱: خلاصه‌ای از توزیع داده‌های BraTS در سال‌های مختلف

Year	Total Data	Training Data	Validation Data	Testing Data
2012	50	35	NA	15
2013	60	35	NA	25
2014	238	200	NA	38
2015	253	200	NA	53
2016	391	200	NA	191
2017	477	285	46	146
2018	542	285	66	191
2019	626	335	125	166
2020	660	369	125	166
2021	2040	1251	219	577

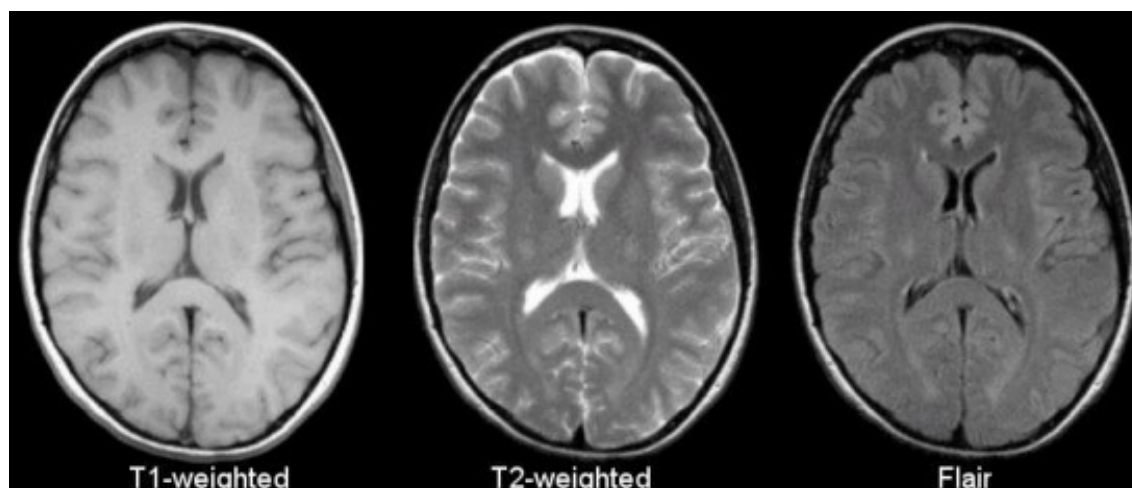
این مجموعه داده شامل تصاویر دوبعدی و سه‌بعدی MRI از چندین مؤسسه است که هر کدام دارای انواع اسکالر و پروتکل‌های تصویربرداری متفاوتی هستند و مجموعه متنوعی از داده‌های تصویربرداری را ارائه می‌دهند. این مجموعه داده همچنین شامل Annotation برای هر زیرگروه تومور، از جمله درجه، اندازه، مکان و شکل است که می‌تواند برای ناحیه‌بندی و طبقه‌بندی^{۹۵} تومور استفاده شود.

این مجموعه داده شامل چهار مدالیته یا روش برای تصاویر MRI است که عبارت اند از:

- T2 Fluid Attenuated Inversion Recovery (FLAIR)
- Native T1-weighted (T1)
- T1-weighted contrast-enhanced (T1ce)
- T2-weighted (T2)

⁹⁵Classification

در تصویر ۱۰ یک مثال و مقایسه شهودی از سه نوع مدالیته آورده شده است. همچنین جدول ۲ مقایسه دقیقی‌تری از این مدالیته‌ها را ارائه می‌دهد.

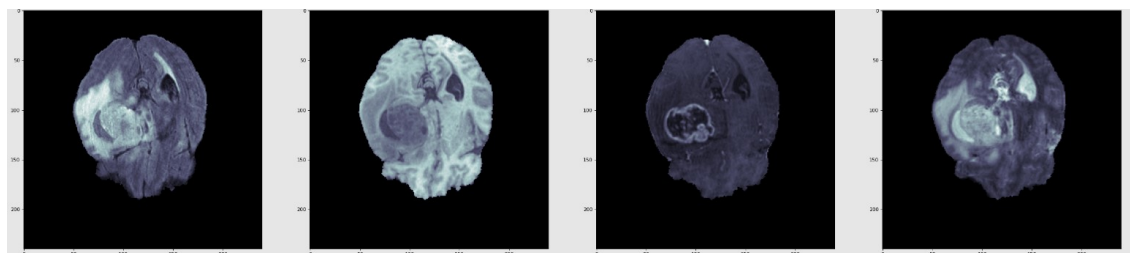


شکل ۱۰: مقایسه سه مدالیته در تصاویر پزشکی [۲۴]

جدول ۲: میزان روشنایی قسمت‌های مختلف مغز در هر مدالیته [۲۴]

	T1	T2	FLAIR
Cerebrospinal fluid (CSF)	Dark	Bright	Dark
White Matter	Light	Dark Gray	Dark Gray
Cortex	Gray	Light Gray	Light Gray
Fat	Bright	Light	Light
Inflammation	Dark	Bright	Bright

در تصویر ۱۱ یکی از نمونه‌های این مجموعه داده در حالت دوبعدی و در هر چهار مدالیته MRI نمایش داده شده است. از سمت چپ اولین تصویر مربوط به مدالیته FLAIR، دومین تصویر مربوط به مدالیته T1، سومین تصویر مربوط به مدالیته T1ce و آخرین تصویر مربوط به مدالیته T2 می‌باشد.



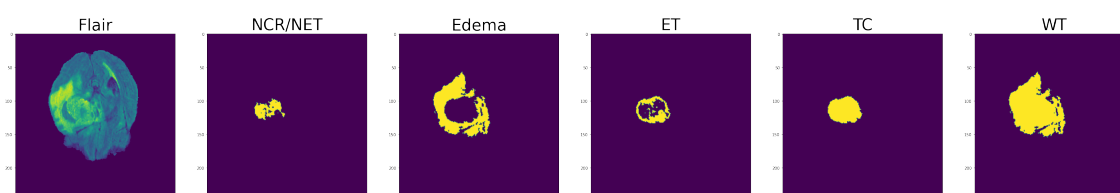
شکل ۱۱: یکی از نمونه‌های BraTS در هر چهار مدالیته

هر اسکن MRI در این مجموعه داده با یک نقشه ناحیه‌بندی دودویی^{۹۶} مربوطه همراه است که وکسل‌ها^{۹۷} (معادل پیکسل در تصویر سه‌بعدی) را به عنوان یکی از موارد زیر برچسب‌گذاری می‌کند:

- Necrotic and non-enhancing tumor core (NCR/NET - Label 1)
- Peritumoral edema (ED - Label 2)
- GD-enhancing tumor (ET - Label 4)

مورد اول شامل قسمت مرکزی تومور است که شامل تومور Necrotic و غیر فزاینده است. مورد دوم به Edema اطراف اشاره دارد که محل نفوذ تومور به بافت سالم است. مورد سوم نمایانگر بخش فزاینده و در حال تقویت تومور است که معمولاً تهاجمی‌ترین قسمت تومور است.

همچنین می‌توان با تجمیع نواحی NCR، ED و ET به ناحیه مربوط به کل تومور یا Whole Tumor (WT) دست یافت. از تجمیع نواحی NCR و ET نیز ناحیه مربوط به هسته تومور یا Tumor Core (TC) به دست می‌آید.



شکل ۱۲: نواحی مربوط به کلاس‌های مختلف در BraTS

^{۹۶} Binary Segmentation Mask

^{۹۷} Voxel

در هر مدالیته نواحی مربوط به کلاس خاصی از مغز برجسته‌تر می‌باشد. برای مثال در مدالیته T1 قسمت‌های مربوط به تومور فزاینده (ET) برجسته‌تر می‌باشد. یا در مدالیته FLAIR قسمت‌های مربوط به Edema برجسته‌تر می‌باشد. همچنین مدالیته‌های FLAIR و T1ce دارای مفیدترین و بیشترین اطلاعات در این مجموعه داده می‌باشند [۲۱].

۳-۶-۰ چالش‌های کار با BraTS

این مجموعه داده دارای چندین چالش برای توسعه الگوریتم‌های ناحیه‌بندی خودکار تومور می‌باشد. یکی از چالش‌های اصلی عدم تعادل کلاس‌ها^{۹۸} در نقشه‌های ناحیه‌بندی است. هسته تومور و نواحی فزاینده تومور در مقایسه با نواحی Edema و پس‌زمینه، نسبتاً کوچک هستند که این امر ناحیه‌بندی این مناطق را چالش برانگیزتر می‌کند.

چالش دیگر تنوع در ظاهر تومورها در مدالیته‌های مختلف تصاویر MRI است. تومورها می‌توانند در مدالیته‌های مختلف به طور متفاوت ظاهر شوند که این امر تشخیص ناحیه دقیق تومورها در هر مدالیته را دشوار می‌کند.

چالش دیگر اندازه بزرگ این مجموعه داده است. هر نمونه این مجموعه داده دارای عرض^{۹۹} ۲۴۰ پیکسل و طول^{۱۰۰} ۲۴۰ پیکسل و عمق^{۱۰۱} ۱۵۵ می‌باشد. با در نظر گرفتن ۴ مدالیته MRI هر نمونه دارای $35712000 = 240 \times 240 \times 155 \times 4$ عدد است در حالی که برای یک تصویر RGB دوبعدی با ابعاد مشابه، این عدد برابر $172800 = 240 \times 240 \times 3$ می‌باشد. به عبارت دیگر هر نمونه این مجموعه داده معادل ۲۰۶ تصویر دوبعدی می‌باشد. این امر آموزش مدل‌های عمیق بر روی این مجموعه داده را بدون تجهیزات سخت‌افزاری مخصوص و قدرتمند بسیار مشکل می‌کند. البته می‌توان این مجموعه داده سه‌بعدی را به صورت دوبعدی در نظر گرفت. برای این کار کافی است بعد مربوط به عمق را جداگانه در نظر گرفت. برای مثال اگر مجموعه داده ما دارای ۳۶۹ نمونه سه‌بعدی باشد، می‌توان آن را در قالب $57195 = 369 \times 155$ نمونه دوبعدی در نظر گرفت. در این صورت ابعاد هر نمونه دوبعدی به صورت $240 \times 240 \times 4$ خواهد بود.

علاوه بر این مشکلات، حاشیه‌نویسی تومورهای مغزی، امری Subjective و وابسته به طرز تفکر شخص

⁹⁸ Imbalance of Classes

⁹⁹ Width

¹⁰⁰ Height

¹⁰¹ Depth

است و حاشیه‌نویسان مختلف ممکن است تومورها را به طور متفاوتی برچسب‌گذاری کنند که این امر می‌تواند منجر به پدیده و چالش Inter-Annotator Variability شود.

در نهایت، مجموعه داده BraTS فقط شامل تعداد محدودی اسکن MRI می‌باشد که ممکن است برای نشان دادن تنوع کامل تومورهای مغزی کافی نباشد. البته این چالش با روش‌های داده افزایی تا حدی قابل حل می‌باشد.

۴-۶-۰ کاربردها

مجموعه داده BraTS کاربردهای متعددی در تصویربرداری پزشکی و بینایی کامپیوتر دارد. این مجموعه داده به طور گسترده‌ای برای توسعه و ارزیابی الگوریتم‌هایی برای ناحیه‌بندی تومور مغزی، طبقه‌بندی تومور و پیش‌بینی بقا^{۱۰۲} استفاده می‌شود. این مجموعه داده همچنین برای توسعه مدل‌های یادگیری عمیق برای تشخیص خودکار^{۱۰۳}، برنامه‌ریزی درمان^{۱۰۴} و راهنمایی جراحی^{۱۰۵} استفاده شده است.

۷-۰. ناحیه‌بندی تعاملی با رویه پالایشی پس‌انتشار^{۱۰۶} [۱۳]

۱-۷-۰ مقدمه

در این مقاله یک الگوریتم ناحیه‌بندی تعاملی تصویر، که نقطه‌هایی از کاربر در مورد شی هدف و پس‌زمینه دریافت می‌کند، ارائه شده است. با به‌دست آوردن فاصله مکانی هر پیکسل تا مکان نقطه‌های داده شده توسط کاربر، نقشه‌های تعامل^{۱۰۷} ایجاد می‌شود. سپس در یک شبکه عصبی کانولوشنی عملیات پاس رو به جلو^{۱۰۸} صورت می‌گیرد و یک نقشه اولیه ناحیه‌بندی به‌دست می‌آید. ممکن است در نتیجه اولیه نقطه‌های ورودی توسط کاربر برچسب اشتباهی بگیرند. بنابراین در این پژوهش رویه پالایشی پس‌انتشار توسعه داده شده است. این رویه پیکسل‌هایی که به اشتباه برچسب‌گذاری شده‌اند را تصحیح می‌کند. نتایج تجربی نشان می‌دهد که

¹⁰²Survival Prediction

¹⁰³Automated Diagnosis

¹⁰⁴Treatment Planning

¹⁰⁵Surgical Guidance

¹⁰⁶Backpropagating Refinement Scheme (BRS)

¹⁰⁷Interaction Map

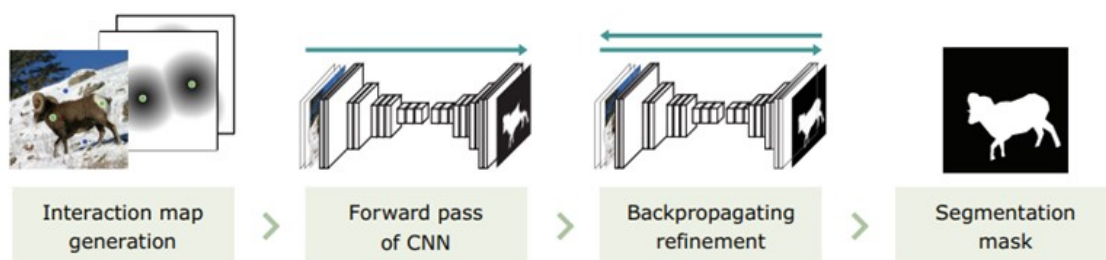
¹⁰⁸Forward Pass

الگوریتم ارائه شده عملکرد بهتری نسبت به روش‌های معمول قبلی بر روی چهار مجموعه داده چالشی دارد.

۷-۲ الگوریتم

الگوریتم ناحیه‌بندی تعاملی ارائه شده یک ماسک دودویی^{۱۰۹} از شی مشخص شده توسط کاربر خروجی می‌دهد. این ماسک از روش Scribble به دست می‌آید که نیازمند کلیک‌های پیش‌زمینه و پس‌زمینه است. این کلیک‌ها نشان دهنده برجسب مورد انتظار برای پیکسل مربوطه هستند.

تصویر ۱۳ خلاصه‌ای از الگوریتم ارائه شده است. با ورودی گرفتن نقطه‌های کاربر، ابتدا نقشه‌های تعامل پیش‌زمینه و پس‌زمینه تولید می‌شوند. سپس تصویر ورودی و این نقشه‌های تعامل وارد یک شبکه CNN می‌شوند. اگر چه این نقشه‌های تعامل به طور شفاف نمایانگر برجسب‌های مشخص شده در محل‌های کلیک شده هستند اما ممکن است نقشه احتمال^{۱۱۰}، اطلاعات اشتباهی از محل‌های کلیک شده به دست آورد. بنابراین به وسیله BRS، محل‌های کلیک شده مجبور می‌شوند تا برجسب‌های مشخص شده توسط کاربر را داشته باشند. در نهایت ماسک ناحیه‌بندی شی هدف با اجرای دوباره پاس رو به جلو به دست می‌آید.



شکل ۱۳: نگاه کلی به الگوریتم BRS [۱۳]

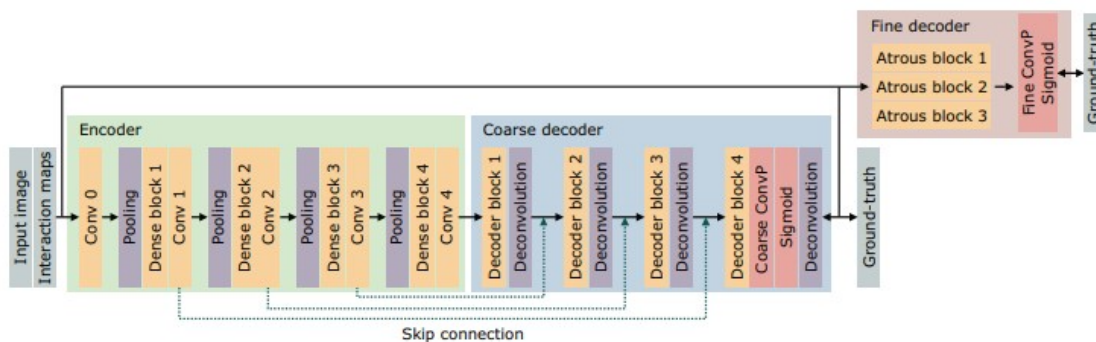
این فرایند هنگامی که کاربر اولین کلیک را روی شی هدف انجام می‌دهد آغاز می‌شود. سپس با به دست آوردن نتیجه ناحیه‌بندی، ممکن است کاربر محل جدیدی را روی شی هدف یا پس‌زمینه کلیک کند. این الگوریتم بارها اجرا می‌شود و ناحیه‌بندی‌های با دقت بالاتر به دست می‌آید. تا زمانی که کاربر دیگر کلیک نکند، این دو مرحله به صورت بازگشتی اجرا می‌شوند.

¹⁰⁹ Binary Mask

¹¹⁰ Probability Mask

۳-۷-۰ معماری شبکه

شبکه عصبی کانولوشنی ارائه شده از معماری رمزگذار-رمزگشا^{۱۱۱} استفاده می‌کند. معماری این شبکه در تصویر ۱۴ نمایش داده شده است. این شبکه دو نقشه تعامل و تصویر اصلی را ورودی می‌گیرد. از شبکه DenseNet [۱۱] به عنوان شبکه رمزگذار استفاده شده است تا ویژگی‌های سطح بالا و سطح پایین استخراج شود. از اتصال پرشی^{۱۱۲} نیز برای نگهداری ویژگی‌های سطح پایین در لایه‌های عمیق شبکه استفاده شده است. در پایان هر بلاک متراکم^{۱۱۳} از ماژول‌های فشرده‌ساز^{۱۱۴} و برانگیختگی^{۱۱۵} نیز استفاده شده است.



شکل ۱۴: معماری ارائه شده بر پایه DenseNet [۱۱]

۴-۷-۰ مرحله آموزش

از مجموعه داده مرزهای معنایی یا Semantic Boundaries Dataset (SBD) [۷] برای آموزش شبکه ارائه داده شده استفاده شده است. این مجموعه داده شامل ۸۴۹۸ تصویر آموزشی است. پیرامون هر مورد شی به صورت تصادفی یک بخش ۳۶۰ در ۳۶۰ که نمایانگر تصویر و ماسک هدف است بریده^{۱۱۶} شده است. همچنین از داده افزایی به صورت انعکاس افقی^{۱۱۷} استفاده شده است.

با توجه به اینکه مجموعه داده SBD دارای نقطه‌های وارد شده توسط کاربر نیست، از یک روش خوشه‌بندی ساده به نام k-Medoids برای فراهم کردن این نقطه‌ها استفاده شده است. نتیجه این الگوریتم خوشه‌بندی در

^{۱۱۱} Encoder-Decoder

^{۱۱۲} Skip Connections

^{۱۱۳} Dense Block

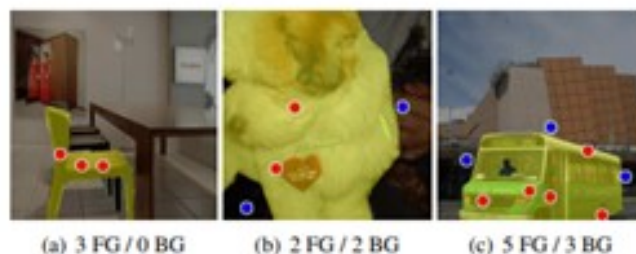
^{۱۱۴} Squeeze Module

^{۱۱۵} Excitation Module

^{۱۱۶} Crop

^{۱۱۷} Horizontal Flip

تصویر ۱۵ آورده شده است.



شکل ۱۵: مثالی از نقطه‌های تولید شده توسط خوشه‌بند در مرحله آموزش [۱۳]

۷-۰-۵ مرحله استنتاج

الگوریتم ارائه شده تصویر اصلی و نقشه‌های تعامل پیش‌زمینه و پس‌زمینه را ورودی می‌گیرد. ابتدا نقشه‌های تعامل با محاسبه فاصله هر پیکسل با نزدیک‌ترین نقطه به‌روز می‌شوند. سپس نقشه‌های تعامل به‌روز شده وارد شبکه عصبی می‌شوند تا نقشه احتمال تولید شود. پیکسل‌هایی که احتمال آن‌ها بالاتر از نصف می‌باشد به عنوان شی هدف تشخیص داده می‌شوند.

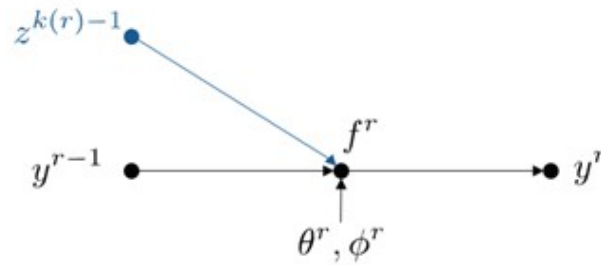
۷-۰-۶ رویه پالایشی پس‌انتشار

ممکن است کلیک‌های کاربر به اشتباه برچسب‌گذاری شوند. بنابراین یک معیار برای محاسبه میزان خطا و یک رویه برای پس‌انتشار خطا و کاهش آن مورد نیاز است. الگوریتم BRS ارائه شده به صورت تکراری^{۱۱۸} عمل پس‌انتشار خطا را انجام می‌دهد تا زمانی که تمامی کلیک‌ها برچسب درستی بگیرند. ابتدا نمادهای این الگوریتم توسط چند تنسور^{۱۱۹} تعریف می‌شود که نمای کلی آن در تصویر ۱۶ آورده شده است. تنسورهای y^{r-1} و z^{r-1} به هم پیوسته^{۱۲۰} می‌شوند و پارامترهای θ^r و ϕ^r برای دستیابی به y^r مورد استفاده قرار می‌گیرد.

¹¹⁸Iterative

¹¹⁹Tensor

¹²⁰Concatenate



شکل ۱۶: نمایش ریاضی رابطه BRS [۱۳]

این پارامتر بیانگر پاسخ لایه r شبکه است. مقدار y^r در رابطه ۱۴ بیان شده است.

$$y^r = f^r(y^{r-1}, z^{r-1}, \theta^r, \phi^r) \quad (۱۴)$$

با استفاده از یک تابع خطا مناسب و انجام عملیات Backpropagation بر روی تابع بالا می‌توان این شبکه را آموزش داد.

۷-۷-۰ نتیجه‌گیری

هرچند که این روش بر روی مجموعه داده BraTS آزمایش نشده است، اما آزمایش‌ها نشان داد که این الگوریتم پیشنهادی از الگوریتم‌های مرسوم و سنتی در چندین مجموعه داده بهتر عمل می‌کند. در واقع استفاده از BRS باعث شد که معیار NoC بهبود چشم‌گیری داشته باشد و نقاط ورودی کاربر به صورت بهینه‌تری مورد استفاده قرار گیرند.

مراجع

- [1] ADAMS, R., AND BISCHOF, L. Seeded region growing. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 16, 6 (1994), 641–647.
- [2] BOYKOV, Y., VEKSLER, O., AND ZABIH, R. Fast approximate energy minimization via graph cuts. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23, 11 (2001), 1222–1239.
- [3] CHEN, D.-J., CHEN, H.-T., AND CHANG, L.-W. Interactive segmentation from 1-bit feedback. in *Computer Vision–ACCV 2016* (Cham, 2017), S.-H. Lai, V. Lepetit, K. Nishino, and Y. Sato, eds. , Springer International Publishing, pp. 261–274.
- [4] ELHARROUSS, O., AL-MAADEED, S. A., SUBRAMANIAN, N., OTTAKATH, N., ALMAADEED, N., AND HIMEUR, Y. Panoptic segmentation: A review. *ArXiv abs/2111.10250* (2021).
- [5] GHAFFARI, M., SOWMYA, A., AND OLIVER, R. Automated brain tumor segmentation using multimodal brain scans: a survey based on models submitted to the brats 2012-2018 challenges. *IEEE Reviews in Biomedical Engineering* 13 (2020), 156–168.
- [6] GUO, D., PEI, Y., ZHENG, K., YU, H., LU, Y., AND WANG, S. Degraded image semantic segmentation with dense-gram networks. *IEEE Transactions on Image Processing* 29 (2020), 782–795.
- [7] HARIHARAN, B., ARBELÁEZ, P., BOURDEV, L., MAJI, S., AND MALIK, J. Semantic contours from inverse detectors. in *2011 International Conference on Computer Vision* (2011), pp. 991–998.
- [8] HE, J., KIM, C.-S., AND KUO, C.-C. J. *Interactive Image Segmentation Techniques*. Springer Singapore, Singapore, 2014, pp. 17–62.

- [9] HE, K., ZHANG, X., REN, S., AND SUN, J. Deep residual learning for image recognition. in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2016), pp. 770–778.
- [10] HERNANDEZ-VELA, A., HERNÁNDEZ-VELA, A., PRIMO, C., AND ESCALERA, S. Automatic user interaction correction via multi-label graph cuts. in *2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops)* (2011), pp. 1276–1281.
- [11] HUANG, G., LIU, Z., AND WEINBERGER, K. Q. Densely connected convolutional networks. *CoRR abs/1608.06993* (2016).
- [12] ISLAM, K. T., WIJEWICKREMA, S., AND O’LEARY, S. A deep learning framework for segmenting brain tumors using mri and synthetically generated ct images. *Sensors* 22, 2 (2022).
- [13] JANG, W.-D., AND KIM, C.-S. Interactive image segmentation via backpropagating refinement scheme. in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (June 2019).
- [14] KIRILLOV, A., HE, K., GIRSHICK, R. B., ROTHER, C., AND DOLLÁR, P. Panoptic segmentation. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2018), 9396–9405.
- [15] LEE, K. M., MYEONG, H., AND SONG, G. Seednet: Automatic seed generation with deep reinforcement learning for robust interactive segmentation. in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2018), pp. 1760–1768.
- [16] LI, A. Y., AND IV, M. Conventional and advanced imaging techniques in post-treatment glioma imaging. *Frontiers in Radiology* 2 (2022).
- [17] LUO, X., WANG, G., SONG, T., ZHANG, J., AERTSEN, M., DEPREST, J., OURSELIN, S., VERCAUTEREN, T., AND ZHANG, S. Mideepseg: Minimally interactive segmentation of unseen objects from medical images using deep learning. *Medical Image Analysis* 72 (2021), 102102.
- [18] MANINIS, K.-K., CAELLES, S., PONT-TUSET, J., AND VAN GOOL, L. Deep extreme cut: From extreme points to object segmentation. in *Computer Vision and Pattern Recognition (CVPR)* (06 2018), pp. 616–625.
- [19] MEILA, M. Comparing clusterings: An axiomatic view. in *Proceedings of the 22nd International Conference on Machine learning* (01 2005), pp. 577–584.

- [20] MENZE, B. H. E. A. The multimodal brain tumor image segmentation benchmark (brats). *IEEE Transactions on Medical Imaging* 34, 10 (2015), 1993–2024.
- [21] NASIM, M. A. A., MUNEM, A., ISLAM, M., PALASH, M. A. H., HAQUE, M. M. A., AND SHAH, F. M. Brain tumor segmentation using enhanced u-net model with empirical analysis. *ArXiv abs/2210.13336* (2022).
- [22] NIELSEN, F., AND NOCK, R. On region merging: the statistical soundness of fast sorting, with applications. in *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings.* (2003), volume 2, pp. II–19.
- [23] NING, J., ZHANG, L., ZHANG, D., AND WU, C. Interactive image segmentation by maximal similarity based region merging. *Pattern Recognition* 43, 2 (2010), 445–456. Interactive Imaging and Vision.
- [24] PRESTON, D. C. Magnetic resonance imaging (mri) of the brain and spine: Basics. <https://case.edu/med/neurology/NR/MRI%20Basics.htm>, 2006. Accessed: 2023-03-07.
- [25] RAMADAN, H., LACHQAR, C., AND TAIRI, H. A survey of recent interactive image segmentation methods. *Computational Visual Media* (08 2020).
- [26] RAYLEIGH. The problem of the random walk. *Nature* 72 (1905), 318–318.
- [27] ROTHER, C., KOLMOGOROV, V., AND BLAKE, A. Grabcut -interactive foreground extraction using iterated graph cuts. *ACM Transactions on Graphics (SIGGRAPH)* (August 2004).
- [28] ROTHER, C., KOLMOGOROV, V., AND BLAKE, A. Grabcut: Interactive foreground extraction using iterated graph cuts. *ACM Trans. Graph.* 23 (08 2004), 309–314.
- [29] RUPPRECHT, C., PETER, L., AND NAVAB, N. Image segmentation in twenty questions. in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2015), pp. 3314–3322.
- [30] SPINA, T. V., DE MIRANDA, P. A. V., AND XAVIER FALCÃO, A. Hybrid approaches for interactive image segmentation using the live markers paradigm. *IEEE Transactions on Image Processing* 23, 12 (2014), 5756–5769.
- [31] XIAN, M., ZHANG, Y., CHENG, H.-D., XU, F., AND DING, J. Neutro-connectedness cut. *Trans. Img. Proc.* 25, 10 (oct 2016), 4691–4703.

- [32] XU, N., PRICE, B., COHEN, S., YANG, J., AND HUANG, T. Deep interactive object selection. in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (Los Alamitos, CA, USA, jun 2016), IEEE Computer Society, pp. 373–381.
- [33] XU, N., PRICE, B., COHEN, S., YANG, J., AND HUANG, T. Deep grabcut for object selection. in *2017 British Machine Vision Conference* (07 2017).
- [34] YI, J., WU, P., JIANG, M., HUANG, Q., HOEPPNER, D. J., AND METAXAS, D. N. Attentive neural cell instance segmentation. *Medical image analysis* 55 (2019), 228–240.