



دانشکده مهندسی کامپیوتر

## ناحیه‌بندی تعاملی تومور در تصاویر سه‌بعدی پزشکی با استفاده از یادگیری عمیق

پروژه کارشناسی مهندسی کامپیوتر گروه هوش مصنوعی

علی صداقی ارقون

استاد راهنما:

دکتر محمدرضا محمدی

۱۴۰۱ اسفند

بِسْمِ اللّٰهِ الرَّحْمٰنِ الرَّحِيْمِ

## تأییدیه‌ی هیأت داوران جلسه‌ی دفاع از گزارش پروژه پایانی

نام دانشکده: دانشکده مهندسی کامپیوتر

نام دانشجو: علی صداقی ارقون

عنوان گزارش پروژه پایانی: ناحیه‌بندی تعاملی تومور در تصاویر سه‌بعدی پزشکی با استفاده از یادگیری

عمیق

تاریخ دفاع: اسفند ۱۴۰۱

رشته: مهندسی کامپیوتر

گروه: هوش مصنوعی

ردیف	سمت	نام و نام خانوادگی	مرتبه دانشگاهی	دانشگاه یا مؤسسه دانشگاهی	امضـا
۱	استاد راهنما	دکتر محمد رضا محمدی	دانشیار	دانشگاه علم و صنعت ایران	

ب

## تأییدیه‌ی صحت و اصالت نتایج

با اسمه تعالی

اینجانب علی صداقتی ارقون به شماره دانشجویی ۹۷۵۲۱۳۷۸ دانشجوی رشته مهندسی کامپیوتر مقطع تحصیلی کارشناسی تأیید می‌نمایم که کلیه‌ی نتایج این گزارش پروژه پایانی حاصل کار اینجانب و بدون هرگونه دخل و تصرف است و موارد نسخه‌برداری شده از آثار دیگران را با ذکر کامل مشخصات منبع ذکر کرده‌ام. درصورت اثبات خلاف مندرجات فوق، به تشخیص دانشگاه مطابق با ضوابط و مقررات حاکم (قانون حمایت از حقوق مؤلفان و مصنفان و قانون ترجمه و تکثیر کتب و نشریات و آثار صوتی، ضوابط و مقررات آموزشی، پژوهشی و انضباطی ...) با اینجانب رفتار خواهد شد و حق هرگونه اعتراض درخصوص احراق حقوق مکتب و تشخیص و تعیین تخلف و مجازات را از خویش سلب می‌نمایم. در ضمن، مسئولیت هرگونه پاسخگویی به اشخاص اعم از حقیقی و حقوقی و مراجع ذیصلاح (اعم از اداری و قضایی) به عهده‌ی اینجانب خواهد بود و دانشگاه هیچ‌گونه مسئولیتی در این خصوص نخواهد داشت.

نام و نام خانوادگی: علی صداقتی ارقون

تاریخ و امضا:

پ

## مجوز بهره‌برداری از پایان‌نامه

بهره‌برداری از این پایان‌نامه در چهارچوب مقررات کتابخانه و با توجه به محدودیتی که توسط استاد راهنما

به شرح زیر تعیین می‌شود، بلامانع است:

- بهره‌برداری از این پایان‌نامه برای همگان بلامانع است.
- بهره‌برداری از این پایان‌نامه با اخذ مجوز از استاد راهنما، بلامانع است.
- بهره‌برداری از این پایان‌نامه تا تاریخ ..... ممنوع است.

استاد راهنما: دکتر محمدرضا محمدی

تاریخ:

امضا:

## تشکر و قدردانی

خود را موظف می‌دانم از همه کسانی که در این مدت در کنار اینجانب بوده اند تشکر کنم. ابتدا از خانواده‌ام، بهویژه پدر و مادر عزیزم که در تمام دوران زندگی همواره حامی من بوده اند سپاسگزار می‌باشم.

همچنین از اساتید دانشکده مهندسی کامپیوتر به ویژه جناب آقای دکتر محمدی که نه تنها در تکمیل این پروژه به اینجانب کمک کردند، بلکه در تمام مدت تحصیل راهنمایی‌ها و نکاتشان راهگشای اینجانب بود، تشکر می‌کنم و برای ایشان و خانواده محترمشان آرزوی سلامتی و موفقیت دارم.

همچنین از جناب آقای دکتر مزینی که برای داوری این پروژه قبول زحمت نمودند و وقت ارزشمند خود را در اختیار اینجانب قرار داند تشکر و قدردانی می‌نمایم.

## 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. اگرچه 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 پژوهه 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 درشت‌دانه بر روی بُرشی 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 پژوهه، 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

د	فهرست تصاویر
ز	فهرست جداول
۱	فصل ۱ : مقدمه
۱	۱-۱ مقدمه
۴	فصل ۲ : مروری بر منابع و پیشینه تحقیق
۴	۱-۲ ناحیه‌بندی تصویر
۵	۱-۱-۲ ناحیه‌بندی معنایی
۵	۲-۱-۲ ناحیه‌بندی نمونه‌ای
۵	۳-۱-۲ ناحیه‌بندی سراسرنما
۶	۲-۲ ناحیه‌بندی تعاملی تصویر
۷	۳-۲ دسته‌بندی روش‌های ناحیه‌بندی تعاملی
۷	۱-۳-۲ دسته‌بندی بر اساس نحوه تعامل کاربر
۸	۱-۱-۳-۲ تعامل مبتنی بر Seed
۹	۲-۱-۳-۲ تعامل مبتنی بر ROI
۱۰	۳-۱-۳-۲ تعامل فعال و منفعل
۱۱	۲-۳-۲ دسته‌بندی بر اساس روش‌شناسی
۱۲	۱-۲-۳-۲ روش مبتنی بر Contour
۱۳	۲-۲-۳-۲ روش مبتنی بر GraphCut

## فهرست مطالب

ج

۱۳ . . . . .	روش مبتنی بر Random Walk ۳-۲-۳-۲
۱۴ . . . . .	روش مبتنی بر Region Growing, Merging ۴-۲-۳-۲
۱۵ . . . . .	روش‌های یادگیری عمیق و ترکیبی ۵-۲-۳-۲
۱۶ . . . . .	معیارهای ارزیابی ۴-۲
۱۷ . . . . .	Error Rate (ERR) ۱-۴-۲
۱۷ . . . . .	True Positive Rate (TPR) ۲-۴-۲
۱۷ . . . . .	Precision ۳-۴-۲
۱۸ . . . . .	F1-Score ۴-۴-۲
۱۸ . . . . .	Intersection over Union (IoU) ۵-۴-۲
۲۰ . . . . .	Dice Similarity Coefficient (DSC) ۹-۴-۲
۲۱ . . . . .	Variation of Information (VI) ۷-۴-۲
۲۱ . . . . .	Number of Clicks (NoC) ۸-۴-۲
۲۱ . . . . .	Number of Failures (NoF) ۹-۴-۲
۲۲ . . . . .	Time per Click (TPC) ۱۰-۴-۲
۲۲ . . . . .	mDice و mIoU ۱۱-۴-۲
۲۲ . . . . .	مجموعه دادگان BraTS ۵-۲
۲۲ . . . . .	۱-۵-۲ مقدمه
۲۳ . . . . .	۲-۵-۲ ویژگی‌های BraTS
۲۷ . . . . .	۳-۵-۲ چالش‌های کار با BraTS
۲۸ . . . . .	۴-۵-۲ کاربردها
۲۸ . . . . .	۶-۲ ناحیه‌بندی تعاملی با رویه پالایشی پسانشار (BRS)
۲۸ . . . . .	۱-۶-۲ مقدمه
۲۹ . . . . .	۲-۶-۲ الگوریتم
۳۰ . . . . .	۳-۶-۲ معماری شبکه
۳۰ . . . . .	۴-۶-۲ مرحله آموزش
۳۱ . . . . .	۵-۶-۲ مرحله استنتاج

## فهرست مطالب

ح

۳۱ . . . . .	رویه پالایشی پس انتشار . . . . .	۶-۶-۲
۳۲ . . . . .	نتیجه‌گیری . . . . .	۷-۶-۲
۳۳ . . . . .	فصل ۳: روش پیشنهادی	
۳۳ . . . . .	۱-۳ مقدمه . . . . .	
۳۴ . . . . .	۲-۳ مشکلات روش‌های قبلی . . . . .	
۳۴ . . . . .	۱-۲-۳ کارایی نامناسب بر روی دستگاه‌های با توان محاسباتی کم . . . . .	
۳۴ . . . . .	۲-۲-۳ عدم استفاده از ماسک از پیش موجود . . . . .	
۳۶ . . . . .	۳-۳ ویژگی‌های الگوریتم ارائه شده . . . . .	
۳۶ . . . . .	۴-۳ الگوریتم در یک نگاه . . . . .	
۳۷ . . . . .	۵-۳ جزئیات الگوریتم . . . . .	
۳۸ . . . . .	۱-۵-۳ بخش Target Crop . . . . .	
۳۸ . . . . .	۲-۵-۳ بخش Coarse Segmentation . . . . .	
۴۰ . . . . .	۳-۵-۳ بخش Focus Crop . . . . .	
۴۱ . . . . .	۴-۵-۳ بخش Local Refinement . . . . .	
۴۲ . . . . .	۵-۵-۳ بخش Progressive Merge . . . . .	
۴۳ . . . . .	۶-۳ تابع خطأ . . . . .	
۴۴ . . . . .	۷-۳ ترکیب کردن مدل‌لیتهای مختلف MRI . . . . .	
۴۶ . . . . .	فصل ۴: آزمایش‌ها و نتیجه‌ها	
۴۶ . . . . .	۱-۴ آماده‌سازی مجموعه داده BraTS . . . . .	
۴۷ . . . . .	۲-۴ توزیع دیتاست به مجموعه‌های آموزش، اعتبارسنجی و تست . . . . .	
۴۷ . . . . .	۳-۴ محیط و هایپرپارامترهای عمومی آموزش . . . . .	
۴۸ . . . . .	۴-۴ نحوه تعریف و اجرای آزمایش‌ها . . . . .	
۴۹ . . . . .	۵-۴ نکات کلی درباره نتایج آزمایش‌ها . . . . .	
۵۰ . . . . .	۶-۴ آزمایش ۱: HRNet18s-FLAIR-WT-Pretrained . . . . .	
۵۲ . . . . .	۷-۴ آزمایش ۲: HRNet32-FLAIR-WT-Pretrained . . . . .	

۵۴ . . . . .	آزمایش ۳ : SegFormerB3-FLAIR-WT-Pretrained . . . . .	۸-۴
۵۶ . . . . .	آزمایش ۴ : SegFormerB3-Mix-WT و ترکیب مдалیته‌ها . . . . .	۹-۴
۵۸ . . . . .	۱۰-۴ آزمایش ۵ : ارزیابی با تسک جدید دیده نشده در آموزش . . . . .	
۶۰ . . . . .	۱۱-۴ آزمایش ۶ : ارزیابی با مDALIte و تسک جدید دیده نشده در آموزش . . . . .	
۶۳ . . . . .	۱۲-۴ مقایسه با نتایج کارهای مرتبط . . . . .	
۶۳ . . . . .	۱-۱۲-۴ مقاله Enhanced U-Net . . . . .	
۶۴ . . . . .	۲-۱۲-۴ مقاله DeepIGeoS . . . . .	
۶۵ . . . . .	۳-۱۲-۴ مقاله MIDeepSeg . . . . .	
۶۶ . . . . .	۴-۱۲-۴ روش‌های سنتی و غیر عمیق . . . . .	
۶۸ . . . . .	فصل ۵: جمع‌بندی و پیشنهادها . . . . .	
۶۸ . . . . .	۱-۵ جمع‌بندی . . . . .	
۶۹ . . . . .	۲-۵ پیشنهادها و کارهای آینده . . . . .	
۶۹ . . . . .	۱-۲-۵ مناسب‌سازی FocalClick برای ورودی سه‌بعدی . . . . .	
۶۹ . . . . .	۲-۲-۵ استفاده از FocalClick در BRS . . . . .	
۶۹ . . . . .	۳-۲-۵ استفاده از تبدیل فاصله ژئودزیک نمایی . . . . .	
۷۰ . . . . .	۴-۲-۵ پیشنهادهای دیگر . . . . .	
۷۱ . . . . .	مراجع . . . . .	
۷۶ . . . . .	پیوست‌ها . . . . .	

# فهرست تصاویر

۱-۲	مثالی از ناحیه‌بندی تصویر پزشکی	۵
۲-۲	انواع حالت‌های ناحیه‌بندی	۶
۳-۲	انواع حالت‌های تعامل مبتنی بر ROI و Seed	۸
۴-۲	مثال‌هایی از انواع روش‌های تعامل کاربر	۱۰
۵-۲	دسته‌بندی روش‌های ناحیه‌بندی تعاملی بر اساس روش‌شناسی	۱۲
۶-۲	نمایش ریاضی معیار IoU	۱۹
۷-۲	مثالی از تصویر S و G	۱۹
۸-۲	مثالی از ناحیه اشتراک و اجتماع	۱۹
۹-۲	نمایش ریاضی معیار Dice	۲۰
۱۰-۲	مقایسه سه مدالیته در تصاویر پزشکی	۲۵
۱۱-۲	یکی از نمونه‌های BraTS در هر چهار مدالیته	۲۶
۱۲-۲	نواحی مربوط به کلاس‌های مختلف در BraTS	۲۶
۱۳-۲	نگاه کلی به الگوریتم BRS	۲۹
۱۴-۲	معماری ارائه شده بر پایه DenseNet	۳۰
۱۵-۲	مثالی از نقطه‌های تولید شده توسط خوشبند در مرحله آموزش	۳۱
۱۶-۲	نمایش ریاضی رابطه BRS	۳۲
۱-۳	انواع کلیک	۳۳
۲-۳	تغییرات ناخواسته هنگام تولید خروجی جدید	۳۵
۳-۳	نگاه کلی به مراحل الگوریتم FocalClick	۳۷

۳۹ . . . . .	۴-۳	معماری قسمت Coarse Segmentation
۳۹ . . . . .	۵-۳	معماری شبکه SegFormer
۴۰ . . . . .	۶-۳	معماری شبکه U-Net
۴۰ . . . . .	۷-۳	معماری شبکه HRNet
۴۱ . . . . .	۸-۳	معماری قسمت Focus Crop
۴۱ . . . . .	۹-۳	معماری قسمت Local Refinement
۴۳ . . . . .	۱۰-۳	معماری قسمت Progressive Merge
۴۵ . . . . .	۱۱-۳	نمایش تصویر ایجاد شده در حالت ترکیب مدل‌لیتهای
۴۸ . . . . .	۱-۴	نحوه تعریف متغیرهای فاز آموزش
۴۹ . . . . .	۲-۴	نحوه تعریف متغیرهای فاز ارزیابی
۵۱ . . . . .	۳-۴	خروجی الگوریتم در یکی از مراحل فاز آموزش آزمایش ۱
۵۱ . . . . .	۴-۴	خروجی الگوریتم در یکی از نمونه‌های ارزیابی آزمایش ۱
۵۱ . . . . .	۵-۴	بازیابی جزئیات تصویر یکی از نمونه‌های ارزیابی در آزمایش ۱
۵۳ . . . . .	۶-۴	خروجی الگوریتم در یکی از مراحل فاز آموزش آزمایش ۲
۵۳ . . . . .	۷-۴	خروجی الگوریتم در یکی از نمونه‌های ارزیابی آزمایش ۲
۵۳ . . . . .	۸-۴	بازیابی جزئیات تصویر یکی از نمونه‌های ارزیابی در آزمایش ۲
۵۵ . . . . .	۹-۴	خروجی الگوریتم در یکی از مراحل فاز آموزش آزمایش ۳
۵۵ . . . . .	۱۰-۴	خروجی الگوریتم در یکی از نمونه‌های ارزیابی آزمایش ۳
۵۵ . . . . .	۱۱-۴	بازیابی جزئیات تصویر یکی از نمونه‌های ارزیابی در آزمایش ۳
۵۷ . . . . .	۱۲-۴	خروجی الگوریتم در یکی از مراحل فاز آموزش آزمایش ۴
۵۷ . . . . .	۱۳-۴	خروجی الگوریتم در یکی از نمونه‌های ارزیابی آزمایش ۴
۵۷ . . . . .	۱۴-۴	بازیابی جزئیات تصویر یکی از نمونه‌های ارزیابی در آزمایش ۴
۵۹ . . . . .	۱۵-۴	خروجی الگوریتم در یکی از مراحل فاز آموزش آزمایش ۵ یا همان ۳
۵۹ . . . . .	۱۶-۴	خروجی الگوریتم در یکی از نمونه‌های ارزیابی آزمایش ۵
۶۰ . . . . .	۱۷-۴	بازیابی جزئیات تصویر یکی از نمونه‌های ارزیابی در آزمایش ۵

فهرست تصاویر

ر

- ۱۸-۴ خروجی الگوریتم در یکی از مراحل فاز آموزش آزمایش ۶ یا همان ۳ . . . . . ۶۲
- ۱۹-۴ خروجی الگوریتم در یکی از نمونه‌های ارزیابی آزمایش ۶ . . . . . ۶۲
- ۲۰-۴ بازیابی جزئیات تصویر یکی از نمونه‌های ارزیابی در آزمایش ۶ . . . . . ۶۲
- ۱-۵ مقایسه تبدیل‌های مختلف برای کلیک‌های کاربر . . . . . ۷۰

# فهرست جداول

۱-۲	خلاصهای از توزیع داده‌های BraTS در سال‌های مختلف	۲۴
۲-۲	میزان روشنایی قسمت‌های مختلف مغز در هر مدالیته	۲۵
۱-۴	توزیع مجموعه‌های آموزش، اعتبارسنجی و تست	۴۷
۲-۴	معیار NoC و NoF در آزمایش ۱	۵۰
۳-۴	معیار U mIoU و mDice در آزمایش ۱	۵۰
۴-۴	معیار NoC و NoF در آزمایش ۲	۵۲
۵-۴	معیار U mIoU و mDice در آزمایش ۲	۵۲
۶-۴	معیار C NoF و NoC در آزمایش ۳	۵۴
۷-۴	معیار U mIoU و mDice در آزمایش ۳	۵۴
۸-۴	معیار C NoF و NoC در آزمایش ۴	۵۶
۹-۴	معیار U mIoU و mDice در آزمایش ۴	۵۶
۱۰-۴	معیار C NoF و NoC در آزمایش ۵	۵۸
۱۱-۴	معیار U mIoU و mDice در آزمایش ۵	۵۸
۱۲-۴	معیار C NoF و NoC در آزمایش ۶	۶۰
۱۳-۴	معیار U mIoU و mDice در آزمایش ۶	۶۱
۱۴-۴	نتایج مقاله Enhanced U-Net	۶۴
۱۵-۴	نتایج مقاله DeepIGeoS	۶۵
۱۶-۴	توزیع مجموعه داده در مقاله MIDeepSeg	۶۵
۱۷-۴	نتایج مقاله MIDeepSeg	۶۶

فهرست جداول

ژ

۱۸-۴ نتایج روش‌های قدیمی و غیر عمیق . . . . . ۶۷

# فصل ۱

## مقدمه

### ۱-۱ مقدمه

حوزه تصویربرداری پزشکی<sup>۱</sup> در دهه‌های اخیر با استفاده از یادگیری ماشین<sup>۲</sup> و الگوریتم‌های یادگیری عمیق<sup>۳</sup> برای ناحیه‌بندی تصاویر<sup>۴</sup> پزشکی و شناسایی تومورهای مغزی<sup>۵</sup> یا سایر بیماری‌ها دچار تحول عظیم شده است. این حوزه به پزشکان در تشخیص<sup>۶</sup> زودهنگام بدخیمی‌های مغزی و بهبود پیش‌بینی‌ها درباره بیماری کمک می‌کند. بزرگسالان اغلب به گلیوما<sup>۷</sup> مبتلا می‌شوند که این مورد از سلول‌های گلیال<sup>۸</sup> منشا می‌گیرد و در ادامه حتی باعث حمله به بافت‌های سالم مجاور می‌شود [۲۱]. در حالی تصاویر حاصل از مدل‌الیه‌های<sup>۹</sup> تصویربرداری تشدید رزونانس مغناطیسی<sup>۱۰</sup> 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<sup>۱۱</sup> فرایندی زمان‌بر، با هزینه بالا، مستعد خطأ و نیازمند نیروی انسانی متخصص

<sup>1</sup>Medical Imaging

<sup>2</sup>Machine Learning (ML)

<sup>3</sup>Deep Learning (DL)

<sup>4</sup>Image Segmentation (IS)

<sup>5</sup>Brain Tumor

<sup>6</sup>Diagnosis

<sup>7</sup>Glioma

<sup>8</sup>Glia Cells

<sup>9</sup>Modalities

<sup>10</sup>Magnetic Resonance Imaging (MRI)

<sup>11</sup>3D

می باشد. علاوه بر این موارد ممکن است تشخیص های متفاوتی توسط متخصصان مختلف داده شود. این چالش زمانی جدی تر می شود که تومورها از نظر اندازه، شکل و مکان دارای گوناگونی باشند.

بهبود مواردی مانند تشخیص بیماری، برنامه ریزی درمان، نظارت و مسائل بالینی به شدت به ناحیه بندی تومورهای مغزی در تصاویر پزشکی وابستگی دارد و نیازمند تعیین محل و اندازه دقیق تومور مغزی است. با این حال، ویژگی های تومورهای مغزی، فرایند ناحیه بندی صحیح و دقیق را دشوار می کند [۹]. این تومورها می توانند در هر ناحیه ای از مغز ایجاد شوند و در اندازه ها 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 [۲۷].

روش های ناحیه بندی خود کار<sup>۱۲</sup> و توانایی فناوری های تصویربرداری در کنار یادگیری عمیق به متخصصان اجازه می دهد بسیار موثر تر و دقیق تر کار کنند. در نتیجه این امر، ناحیه بندی دقیق و سریع تومورهای مغزی می تواند 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 ناحیه بندی خود کار معماری های مختلف شبکه های عصبی پیچشی<sup>۱۳</sup> بر روی تصاویر سه بعدی پزشکی و MRI در حال استفاده شدن است. پیش از این، استفاده از شبکه های عمیق به دلیل عدم وجود داده های کافی پزشکی و حجم عظیم قدرت پردازش، کار دشواری بود. انتشار مجموعه دادگانی مانند مجموعه داده ناحیه بندی تومور مغزی<sup>۱۴</sup> که دارای تصاویر MRI سه بعدی شامل برچسب<sup>۱۵</sup> می باشد، این کار را آسان تر کرد. همچنین، با در دسترس قرار گرفتن واحدهای پردازش گرافیکی<sup>۱۶</sup> قوی تر امکان پردازش سریع تر داده ها به وجود آمد. این پیشرفت ها باعث می شود محققان به کار در این حوزه برای کمک به پزشکان توجه بیشتری داشته باشند.

با این حال، از آنجایی که سلول های تومور می توانند در هر جایی از بافت مغز ایجاد شوند و می توانند از نظر اندازه، ظاهر و شکل متفاوت باشند، ناحیه بندی خود کار تومور مغزی و مناطق فرعی و مرزی آن دشوار است [۱۷]. همچنین با توجه به این که در زمینه پزشکی جان و سلامت انسان ها مورد بحث می باشد، نیاز به دخالت و تعامل پزشک در فرایند ناحیه بندی خود کار الزامی می باشد و در نتیجه فرایند ناحیه بندی نیمه خود کار

<sup>12</sup>Automatic Segmentation

<sup>13</sup>Convolutional Neural Network

<sup>14</sup>Brain Tumor Segmentation (BraTS)

<sup>15</sup>Label

<sup>16</sup>Graphics Processing Unit (GPU)

## فصل ۱ . مقدمه

۱۷ مطرح می‌شود.

ناحیه‌بندی تعاملی تصویر<sup>۱۸</sup> روشی است که در آن با دریافت راهنمایی‌هایی از سمت کاربر درباره شی هدف که با عنوان پیش‌زمینه<sup>۱۹</sup> شناخته می‌شود و همچنین راهنمایی‌هایی درباره تصویر پس‌زمینه<sup>۲۰</sup>، می‌توان با دقت و سرعت بالاتری ناحیه مورد علاقه<sup>۲۱</sup> را ناحیه‌بندی کرد. در این روش‌ها تعامل کاربر<sup>۲۲</sup> 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 مدل‌های مختلف تصاویر MRI می‌باشد که دارای دقت و سرعت بالا و نیازمند هزینه محاسباتی کمی باشد. افزون بر 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 مدل‌الیه و یک وظیفه<sup>۲۳</sup> خاص بر روی مدل‌الیه 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 خطأ<sup>۲۴</sup>، معیارهای ارزیابی<sup>۲۵</sup> 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.

<sup>17</sup>Semi-Automatic

<sup>18</sup>Interactive Image Segmentation (IIS)

<sup>19</sup>Foreground (FG)

<sup>20</sup>Background (BG)

<sup>21</sup>Region of Interest (ROI)

<sup>22</sup>User Interaction

<sup>23</sup>Task

<sup>24</sup>Loss Function

<sup>25</sup>Evaluation Metric

## فصل ۲

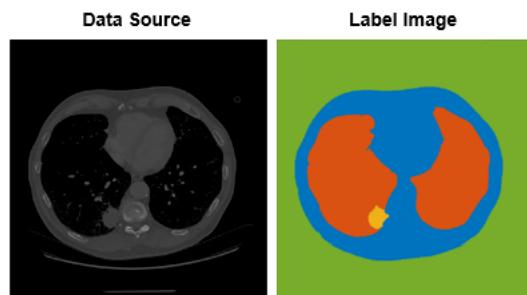
### مروری بر منابع و پیشینه تحقیق

#### ۱-۲ ناحیه‌بندی تصویر

ناحیه‌بندی تصویر یکی از وظایف مهم و پایه‌ای در بینایی کامپیوتر<sup>۱</sup> 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 [۲۸]. به بیان دقیق‌تر ناحیه‌بندی یک تصویر فرایندی است که در آن به هر پیکسل<sup>۲</sup> 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 گستردۀ قرار گرفته است. دلیل این مطالعات گستردۀ، کاربردهای گستردۀ 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.

<sup>1</sup>Computer Vision (CV)

<sup>2</sup>Pixel



شکل ۲-۱: مثالی از ناحیه‌بندی تصویر پزشکی

### ۱-۱-۲ ناحیه‌بندی معنایی

در ناحیه‌بندی معنایی<sup>۳</sup> به تمامی پیکسل‌هایی که مربوط به یک کلاس<sup>۴</sup> می‌باشند یک برچسب ثابت اختصاص می‌دهیم [۱۰]. برای مثال اگر در یک تصویر سه گربه در یک مزرعه وجود داشته باشد، به پیکسل‌های مربوط به گربه‌ها برچسب یک و به پیکسل نواحی مزرعه پیکسل صفر اختصاص داده می‌شود.

### ۲-۱-۲ ناحیه‌بندی نمونه‌ای

در ناحیه‌بندی نمونه‌ای<sup>۵</sup> 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 [۴۶]. برای مثال در تصویر مثال بالا به پیکسل‌های مربوط به گربه اول برچسب صفر، گربه دوم برچسب یک و گربه سوم برچسب دو اختصاص داده می‌شود.

### ۳-۱-۲ ناحیه‌بندی سراسرنما

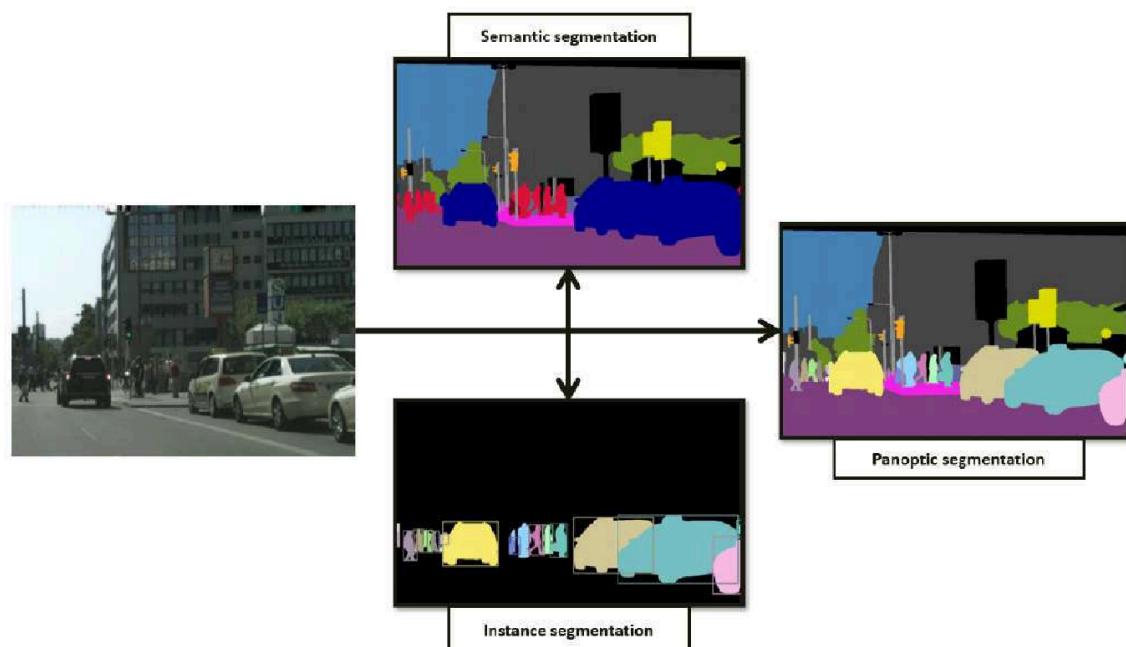
ناحیه‌بندی سراسرنما<sup>۶</sup> ترکیب دو حالت بالا می‌باشد. در این حالت مشابه ناحیه‌بندی معنایی، به هر پیکسل برچسب کلاس مورد نظر اختصاص داده می‌شود. همچنین مشابه ناحیه‌بندی نمونه‌ای، در این حالت نمونه‌های موجود از یک کلاس متمایز می‌شوند [۱۹].

<sup>3</sup>Semantic Segmentation

<sup>4</sup>Class

<sup>5</sup>Instance Segmentation

<sup>6</sup>Panoptic Segmentation



شکل ۲-۲: انواع حالت‌های ناحیه‌بندی [۵]

## ۲-۲ 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<sup>۴</sup> 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 [۳۱]. تعامل کاربر با مدل می‌تواند به صورت‌های گوناگونی باشد. برای مثال می‌توان از نقاطی<sup>۵</sup> یا از خطوطی<sup>۶</sup> برای نشان دادن شی مورد نظر و یا تصویر BG استفاده کرد. همچنین می‌توان از یک جعبه مرزی<sup>۷</sup> یا یک چند ضلعی<sup>۸</sup> 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

<sup>7</sup>Supervised Segmentation<sup>8</sup>Points<sup>9</sup>Scribbles<sup>10</sup>Bounding Box (BB)<sup>11</sup>Polygon

سرعت بیشتری نسبت به ناحیه‌بندی خودکار است [۱۲]. در واقع در بسیاری از کاربردهای بینایی کامپیوتر مانند پژوهشکی برای رسیدن به ناحیه‌بندی دقیق نیاز به این تعامل و دخالت کاربر وجود دارد.

### ۳-۲ دسته‌بندی روش‌های ناحیه‌بندی تعاملی

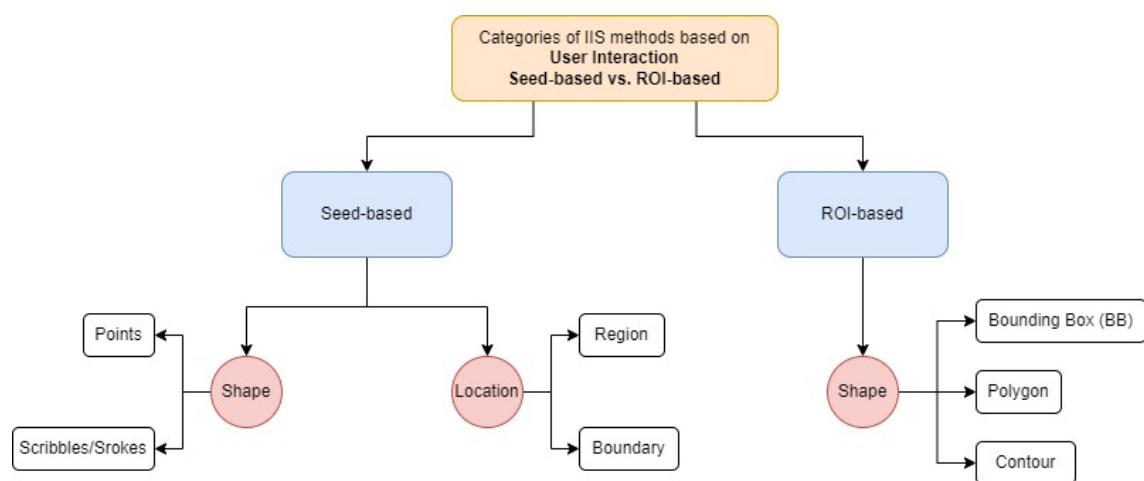
روش‌های ناحیه‌بندی تعاملی را می‌توان به طریق‌های مختلفی دسته‌بندی کرد که در ادامه به جزئیات دو مورد از این طریق‌ها اشاره می‌شود.

#### ۱-۳-۲ 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 [۴۲]<sup>۱۲</sup>. در جنبه دوم تعامل کاربر با مدل می‌تواند به صورت فعال <sup>۱۲</sup> و یا منفعل <sup>۱۳</sup> باشد. از نظر جنبه اول تعامل کاربر با مدل برای مشخص کردن BG و FG می‌تواند به 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 نظری 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.

<sup>12</sup> Active

<sup>13</sup> Passive



شکل ۲-۳: انواع حالت‌های تعامل مبتنی بر Seed و ROI

## ۱-۱-۳-۲ 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<sup>۱۴</sup> 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<sup>۱۵</sup> دست یافت [۳۹]. در این حالت مدل با استفاده از Seed‌های ابتدایی مرزی محیط شکل را تا حدودی مشخص می‌کند و سپس با استفاده از Seed‌های ناحیه‌ای خروجی نهایی بهبود یافته را تولید می‌کند.

<sup>14</sup>Region<sup>15</sup>Hybrid

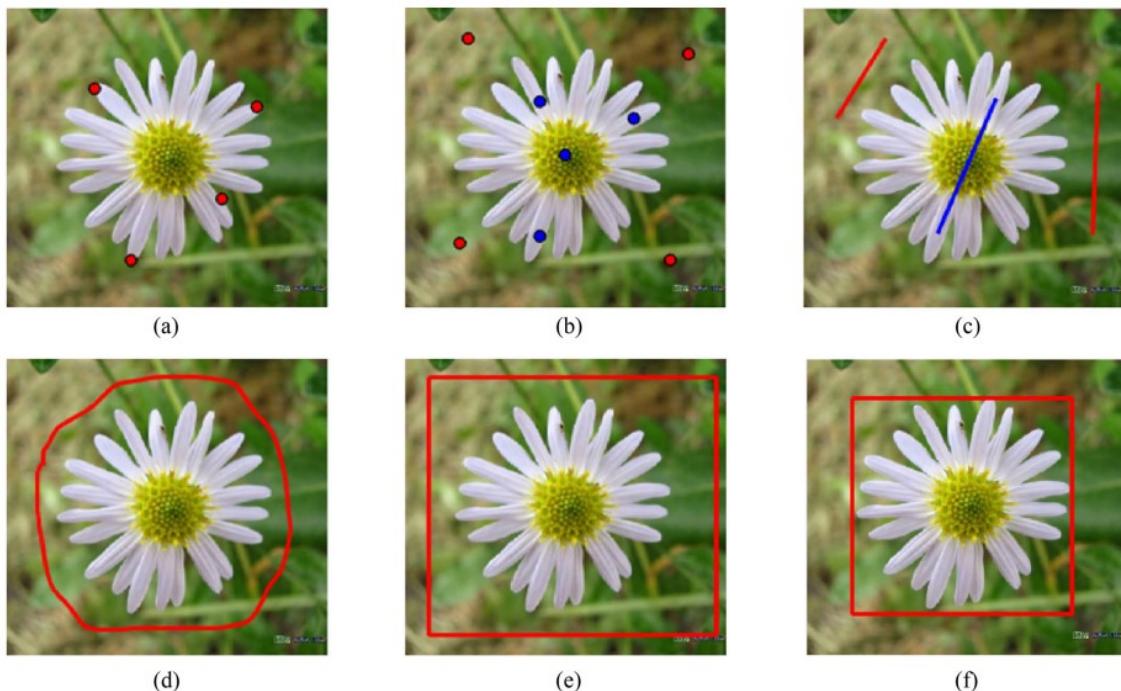
## ۲-۱-۳-۲ تعامل مبتنی بر ROI

در این روش دانش قبلی از محل FG و BG به کمک پیکسل‌های درون و بروان ROI به دست می‌آید. روش Bounding Box [۳۵] معروف‌ترین و پراستفاده‌ترین روش ناحیه‌بندی تعاملی است که به کاربر امکان رسم GrabCut برای جداسازی FG از BG را می‌دهد. عملکرد روش‌های مبتنی بر ROI به کیپ<sup>۱۶</sup> 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 دقیق و عملکرد مدل بسیار وابسته به میزان Tightness خواهد بود. از سوی دیگر الگوریتم‌های مبتنی بر Seed ROI با وجود تعداد کافی می‌توانند FG با شکل‌های پیچیده‌ای را استخراج کنند. به صورت کلی روش Tightness نیازمند تعداد دفعات تعامل بیشتری نسبت به روش Seed می‌باشد [۱۵]. نکته اصلی و چالش کلیدی که بدون توجه به روش ROI یا Seed مطرح است، به دست آوردن دقیق 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 ۴-۲ آورده شده است. تصویر (a) بیانگر حالت Region point seeds، تصویر (b) بیانگر Boundary seeds، تصویر (c) بیانگر ROI contour scribble seeds، تصویر (d) بیانگر ROI BB و تصویر (e) بیانگر ROI tight می‌باشد.

<sup>۱۶</sup>Tight



شکل ۲-۴: مثال‌هایی از انواع روش‌های تعامل کاربر [۳۱]

## ۳-۱-۳-۲ تعامل فعال و منفعل

از جنبه دوم بعضی از مدل‌های ناحیه‌بندی تعاملی امکان تعامل فعالانه و کمک برخط<sup>۱۷</sup> 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<sup>۱۸</sup> (مانند مسابقه بیست سوالی) از کاربر و پردازش پاسخ‌ها نتیجه بهمود یافته نهایی را تولید کند [۳۶]. در روش Active تعامل تا زمانی ادامه می‌یابد که یا کاربر از نتیجه راضی باشد یا مدل به یک آستانه<sup>۱۹</sup> مشخص برسد.

بر خلاف دسته قبلی، تعامل می‌تواند به صورت منفعل یا برون خط<sup>۲۰</sup> صورت گیرد که این دسته به عنوان

<sup>17</sup>Online<sup>18</sup>Binary<sup>19</sup>Threshold<sup>20</sup>Offline

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

## ۲-۳-۲ دسته‌بندی بر اساس روش‌شناسی

الگوریتم‌های ناحیه‌بندی تعاملی را می‌توان بر اساس روش‌شناسی<sup>۲۱</sup> به دو دسته کلی تقسیم کرد. روش‌های سنتی و قدیمی‌تر مانند الگوریتم‌های مبتنی بر Contour، الگوریتم‌های مبتنی بر برش گراف<sup>۲۲</sup>، مبتنی بر الگوریتم ولگشت<sup>۲۳</sup> و مبتنی بر ناحیه مثل 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.

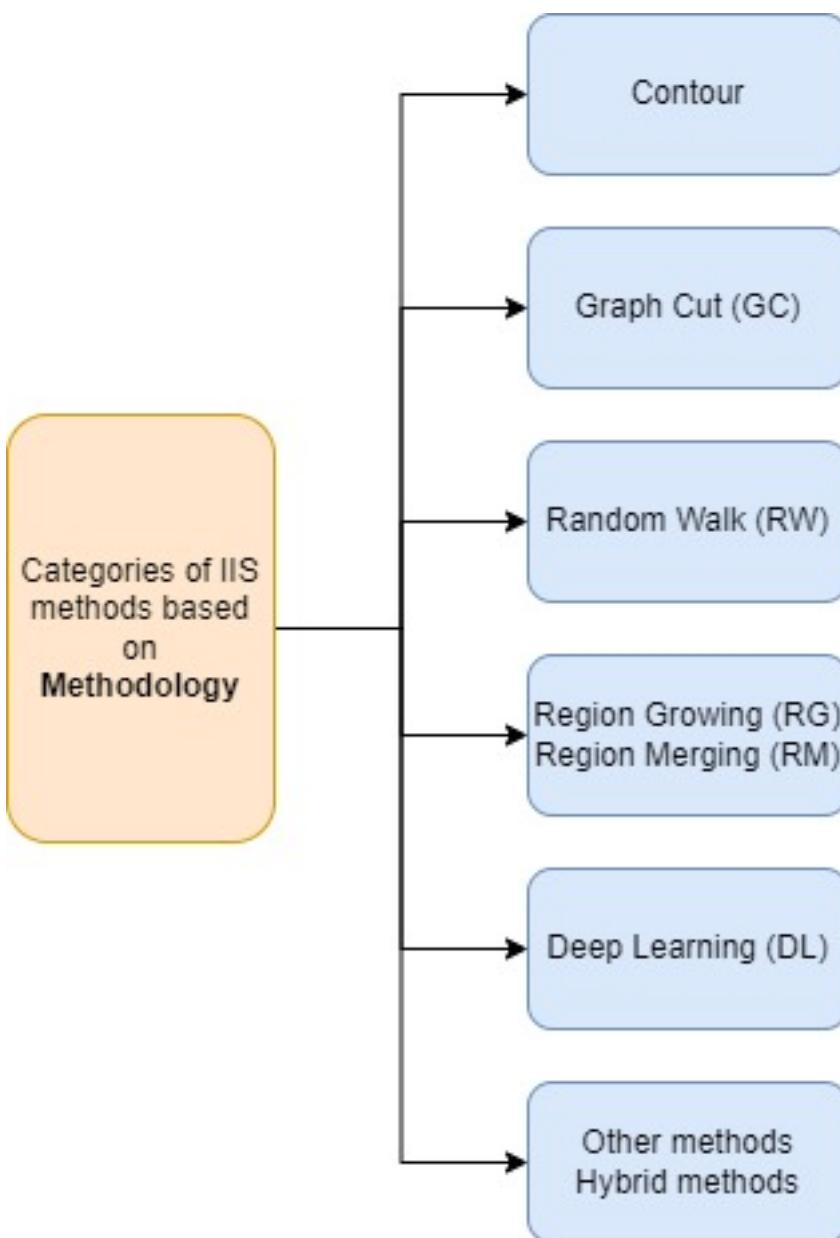
البته در برخی کارها استفاده از هر دو نوع روش مشاهده می‌شود. در این کارها سعی می‌شود از برتری هر روش به خوبی استفاده کرد. برای مثال می‌توان از یادگیری عمیق برای تولید ناحیه‌بندی اولیه (بدون تعامل) کاربر استفاده کرد و سپس تعامل‌های کاربر را با روش GC پردازش کرد و نتیجه را بهبود بخشید [۲۴]. روش مبتنی بر 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 پژوهه 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 Region Merging، Region Growing، Random Walk، GraphCut، Contour و داده خواهد شد.

<sup>21</sup>Methodology

<sup>22</sup>GraphCut (GC)

<sup>23</sup>Random Walk (RW)



شکل ۲-۵: دسته‌بندی روش‌های ناحیه‌بندی تعاملی بر اساس روش‌شناسی

### ۱-۲-۳-۲ روشن مبتنی بر **Contour**

در روش‌های مبتنی بر **Contour** اصل کلیدی مورد توجه، استخراج **Contour**‌های شی هدف با استفاده از ویژگی‌های مربوط به لبه<sup>۲۴</sup> و دانش قبلی فراهم شده توسط تعامل کاربر می‌باشد. **Contour**‌های شی با محاسبه

<sup>24</sup>Edge Features

کوتاهترین مسیر<sup>۲۵</sup> متصل کننده به نقاط Seed استخراج می‌شود. برای محاسبه کوتاهترین مسیر از الگوریتم دایکسترا<sup>۲۶</sup> یا الگوریتم‌های سریع‌تر استفاده می‌شود.

### ۲-۲-۳-۲ روش مبتنی بر GraphCut

روش‌های مبتنی بر GC جزو محبوب‌ترین روش‌های حوزه ناحیه‌بندی تعاملی حساب می‌شوند. در این روش‌ها ابتدا یکتابع انرژی تعریف می‌شود که هدف کمینه کردن این تابع می‌باشد<sup>۲۷</sup>. این کمینه کردن انرژی را می‌توان با حل یک مسئله بیشینه جریان<sup>۲۸</sup> در گراف تخمین زد [۲]. به همین ترتیب طبق نظریه جریان-بیشینه برش-کمینه<sup>۲۹</sup> کافی است برش<sup>۳۰</sup> کمینه در گراف پیدا شود.

روش GrabCut [۳۴] 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 مدل (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<sup>۳۱</sup> مبتنی بر GC برای به دست آوردن این برچسب‌ها اجرا می‌شود.

### ۳-۲-۳-۲ روش مبتنی بر Random Walk

در مدل اصلی و اولیه Random Walk<sup>۳۲</sup> [۳۲] یک گراف بدون جهت<sup>۳۳</sup> ایجاد می‌شود که رئوس<sup>۳۳</sup> این گراف بیانگر پیکسل‌های تصویر ورودی و هر یال<sup>۳۴</sup> آن متصل کننده دو پیکسل (راس) به یکدیگر است. در این گراف وزن هر یال بیانگر احتمال<sup>۳۵</sup> 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

<sup>25</sup>Shortest Path

<sup>26</sup>Dijkstra's Algorithm

<sup>27</sup>Energy Minimization

<sup>28</sup>Max Flow Problem

<sup>29</sup>Max-Flow Min-Cut Theorem

<sup>30</sup>Cut

<sup>31</sup>Optimization

<sup>32</sup>Undirected Graph

<sup>33</sup>Nodes

<sup>34</sup>Edge

<sup>35</sup>Probability

است که در نهایت در راس  $\mathbf{x}$  باقی می‌ماند. راه حل این هدف کمینه کردن تابع زیر است:

$$E(\mathbf{x}) = \mathbf{x}^T \mathbf{L} \mathbf{x} \quad (1-2)$$

در این رابطه  $\mathbf{L}$  بیانگر ماتریس ترکیبی لaplaci<sup>۳۶</sup> 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.  $f$  را به عنوان بردار برچسب پیکسل‌ها در نظر می‌گیریم و ماتریس  $\mathbf{L}$  را به صورت زیر به دو بخش تقسیم می‌کنیم:

$$\mathbf{L} = \begin{bmatrix} \mathbf{L}_m & \mathbf{B} \\ \mathbf{B}^T & \mathbf{L}_u \end{bmatrix} \quad (2-2)$$

در این صورت برای کمینه کردن معادله ۱-۲ کافی است یک سیستم خطی به صورت زیر حل شود:

$$\mathbf{L}_u \mathbf{x}_u = -\mathbf{B}^T f \quad (3-2)$$

#### ۴-۲-۳-۲ روشن 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

<sup>۳۶</sup>Combinatorial Laplacian Matrix

مناسب نیست [۲۹].

### ۵-۲-۳-۲ 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 Cut یا به اختصار DEXTR [۲۴] یک معماری مبتنی بر CNN را برای ناحیه‌بندی تعاملی با استفاده از نقاط مفرط<sup>۳۷</sup> در بالا، پایین، چپ و راست شی مورد نظر ارائه می‌کند. این نقاط در کنار کanal‌های RGB به عنوان ورودی به مدل داده می‌شوند. همچنین از شبکه ResNet-101 [۱۴] به عنوان Backbone در این مدل استفاده شده است.

روش Deep Interactive Object Selection [۴۴] کلیک‌های<sup>۳۹</sup> مثبت (روی شی هدف) و منفی (بیرون شی هدف) ارائه شده توسط کاربر را به نقشه‌های<sup>۴۰</sup> مبتنی بر فاصله اقلیدسی<sup>۴۱</sup> مجزا تبدیل می‌کند. سپس این نقشه‌ها با کanal‌های RGB تصویر ورودی ترکیب می‌شوند تا یک جفت<sup>۴۲</sup> تصویر-تعامل ایجاد شود. سپس این جفت وارد یک شبکه کاملاً پیچشی<sup>۴۳</sup> می‌شود تا ناحیه‌بندی اولیه صورت گیرد. در نهایت یک مرحله پالایش<sup>۴۴</sup> با استفاده از بهینه‌سازی GraphCut اعمال می‌شود و خروجی نهایی تولید می‌شود.

روش Deep GrabCut for Object Selection [۴۵] به جای کلیک، یک Bounding Box را از کاربر ورودی می‌گیرد. همچنین این مدل به صورت یکپارچه<sup>۴۵</sup> آموزش داده می‌شود و نیازی به مرحله پس‌پردازش<sup>۴۶</sup> برای پالایش خروجی‌های FCN ندارد. روش SeedNet [۲۰] 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

<sup>37</sup>Extreme Points

<sup>38</sup>Channels

<sup>39</sup>Click

<sup>40</sup>Map

<sup>41</sup>Euclidean Distance

<sup>42</sup>Pair

<sup>43</sup>Fully Convolutional Network (FCN)

<sup>44</sup>Refinement

<sup>45</sup>End-to-End

<sup>46</sup>Post-Processing

سیستم تولید خودکار Seed با استفاده از یادگیری تقویتی عمیق<sup>۴۷</sup> 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<sup>۴۸</sup> 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<sup>۴۹</sup> و همچنین استفاده از روش‌های داده افزایی<sup>۵۰</sup>، روش‌های مبتنی بر یادگیری عمیق با موفقیت در صنعت و کارخانه‌های هوشمند مورد استفاده قرار گرفته اند.

## ۴-۲ معیارهای ارزیابی

برای ارزیابی کمی و قابل مشاهده<sup>۵۱</sup> 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<sup>۵۲</sup> 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.

<sup>47</sup> Deep Reinforcement Learning (Deep RL)

<sup>48</sup> Training Set

<sup>49</sup> Dataset

<sup>50</sup> Data Augmentation

<sup>51</sup> Objective

<sup>52</sup> Ground Truth (GT)

## Error Rate (ERR) ۱-۴-۲

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

$$ERR = \frac{|S \cup G - G|}{\text{Number of pixels in ROI}} \quad (4-2)$$

## True Positive Rate (TPR) ۲-۴-۲

این معیار با نام Recall به معنی یادآوری یا فراخوانی نیز شناخته می‌شود و بیانگر نسبت پیش‌بینی‌های صحیح مدل در یک کلاس مثبت<sup>۵۳</sup> می‌باشد. رابطه Recall به صورت زیر تعریف می‌شود:

$$R = \frac{|S \cap G|}{|G|} \quad (5-2)$$

دلیل استفاده از این معیار به جای معیار دقت ساده<sup>۵۵</sup> 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 ۳-۴-۲

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

<sup>53</sup>Prediction<sup>54</sup>Positive<sup>55</sup>Accuracy

$$P = \frac{|S \cap G|}{|S|} \quad (6-2)$$

**F1-Score ۴-۴-۲**

بیانگر میانگین همساز<sup>۵۶</sup> دو معیار Recall و Precision می‌باشد و این دو معیار را به صورت زیر با هم ترکیب می‌کند:

$$F = \frac{2 \times R \times P}{R + P} \quad (7-2)$$

**Intersection over Union (IoU) ۵-۴-۲**

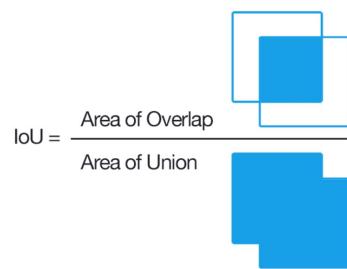
بیانگر نسبت تعداد پیکسل‌هایی که به درستی به عنوان FG برچسب‌گذاری شده‌اند (اشتراک)، به تعداد پیکسل‌هایی که در S یا G (اجتماع) دارای برچسب FG هستند، می‌باشد. مقدار آن از تقسیم این اشتراک و اجتماع به صورت زیر به دست می‌آید:

$$IoU = \frac{|S \cap G|}{|S \cup G|} \quad (8-2)$$

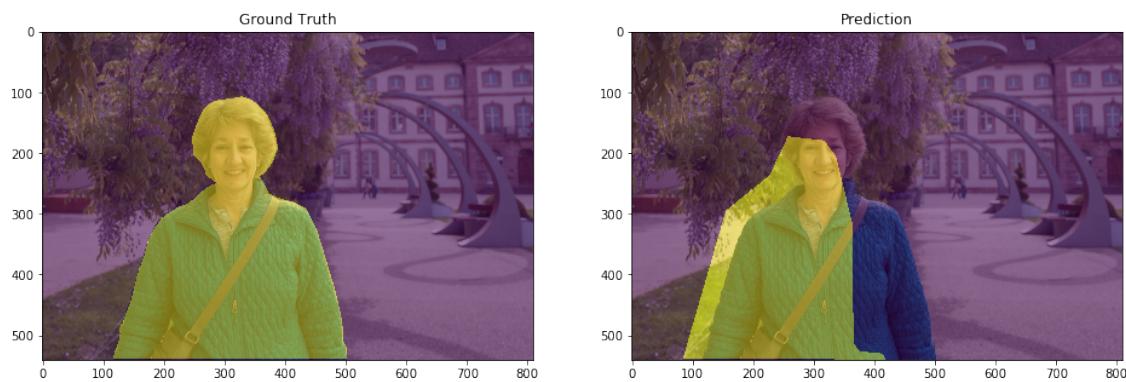
این معیار با نام‌های دیگری نظیر Jaccard Index و Overlap Coefficient نیز شناخته می‌شود و جزو رایج‌ترین معیارهای ارزیابی مدل‌های ناحیه‌بندی می‌باشد.

---

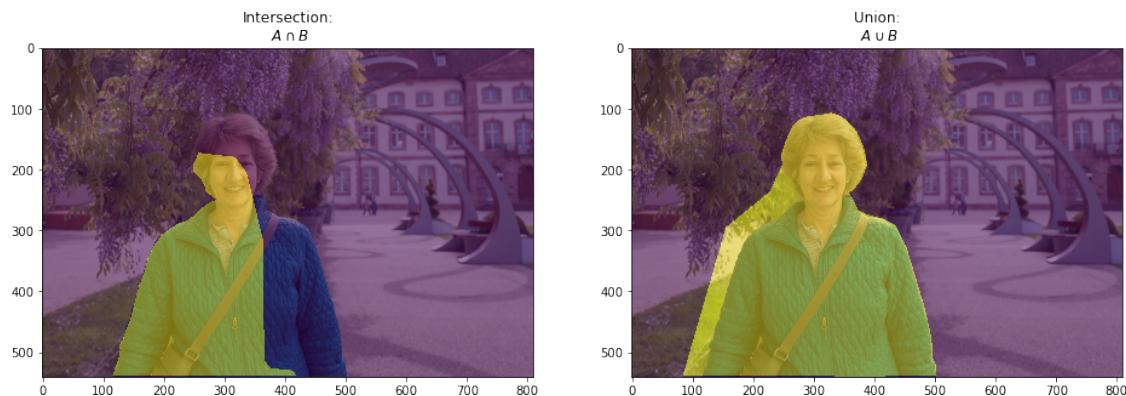
<sup>۵۶</sup>Harmonic Mean



شکل ۲-۶: نمایش ریاضی معیار IoU



شکل ۲-۷: مثالی از تصویر S و G

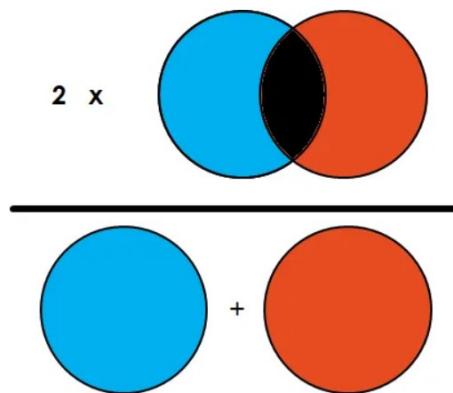


شکل ۲-۸: مثالی از ناحیه اشتراک و اجتماع

**Dice Similarity Coefficient (DSC) ۶-۴-۲**

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.

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



شکل ۹-۲: 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 خطای آموزش شبکه استفاده کرد. دلیل این امر این است که معیار Dice بر خلاف معیار IoU مشتق پذیر می‌باشد. رابطه تابع خطای Dice را می‌توان به صورت زیر تعریف کرد:

$$DiceLoss = 1 - DSC \quad (10-2)$$

**Variation of Information (VI) ۷-۴-۲**

این معیار فاصله بین دو ناحیه‌بندی  $S$  و  $G$  را با استفاده از میانگین آنتروپی شرطی<sup>۵۷</sup> محاسبه می‌کند. همچنین این معیار بر پایه علم نظریه اطلاعات<sup>۵۸</sup> طراحی شده است و بیشتر در مسائل خوشبندی<sup>۵۹</sup> کاربرد دارد و در مسائل ناحیه‌بندی به ندرت مورد استفاده قرار می‌گیرد [۲۵].

**Number of Clicks (NoC) ۸-۴-۲**

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

$$NoC(Dice = 80\%) = 7.87 \quad (11-2)$$

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.

**Number of Failures (NoF) ۹-۴-۲**

اگر فرض کنیم مدل دارای محدودیت max\_clicks روی تعداد کلیک‌های کاربر باشد، در این صورت معیار NoF بیانگر تعداد نمونه‌های<sup>۶۰</sup> است که با حداقل max\_clicks کلیک همچنان به Dice یا IoU مورد نظر نرسیده‌اند. این نمونه‌ها اصطلاحاً نمونه‌های شکست خورده<sup>۶۱</sup> نامیده می‌شوند. برای مثال عبارت زیر:

$$NoF_{max\_clicks=20}(Dice = 85\%) = 851 \quad (12-2)$$

<sup>57</sup>Conditional Entropy

<sup>58</sup>Information Theory

<sup>59</sup>Clustering

<sup>60</sup>Samples

<sup>61</sup>Failed Samples

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

### Time per Click (TPC) ۱۰-۴-۲

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<sup>۶۲</sup> مدل به ازای هر کلیک کاربر می‌باشد. همچنین واحد این معیار ثانیه می‌باشد.

### mDice و mIoU ۱۱-۴-۲

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

$$mDice_{C=20} = 62\% \quad (13-2)$$

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.

### [۲۶] BraTS ۵-۲ مجموعه دادگان

#### ۱-۵-۲ 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.

<sup>62</sup>Response Time

امکان تصویربرداری غیرتهاجمی<sup>۶۳</sup> و دقیق از مغز را فراهم می‌کند و تصاویر سه بعدی 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 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<sup>۶۴</sup> 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.

## ۲-۵-۲ ویژگی‌های BraTS

مجموعه داده BraTS یک مجموعه داده در دسترس عموم است که اسکن‌های<sup>۶۵</sup> MRI تومور مغزی را با حاشیه‌نویسی<sup>۶۶</sup> دستی ارائه می‌دهد. این مجموعه داده در قالب<sup>۶۷</sup> NIfTI که مخفف عبارت Informatics Technology Initiative<sup>۶۸</sup> می‌باشد ارائه می‌شود. نسخه سال ۲۰۲۰ این مجموعه داده شامل اسکن MRI از تومورهای مغزی است که به مجموعه‌های آموزشی (۳۶۹)، اعتبارسنجی<sup>۶۹</sup> (۱۲۵) و آزمایش (۱۶۶)<sup>۷۰</sup> تقسیم می‌شود. در جدول زیر خلاصه‌ای از توزیع مجموعه‌های آموزشی، اعتبارسنجی و آزمایش این مجموعه داده در سال‌های مختلف آورده شده است.

<sup>63</sup>Non Invasive

<sup>64</sup>Benchmark

<sup>65</sup>Scan

<sup>66</sup>Annotation

<sup>67</sup>Format

<sup>68</sup>Validation Set

<sup>69</sup>Testing Set

جدول ۲-۱: 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

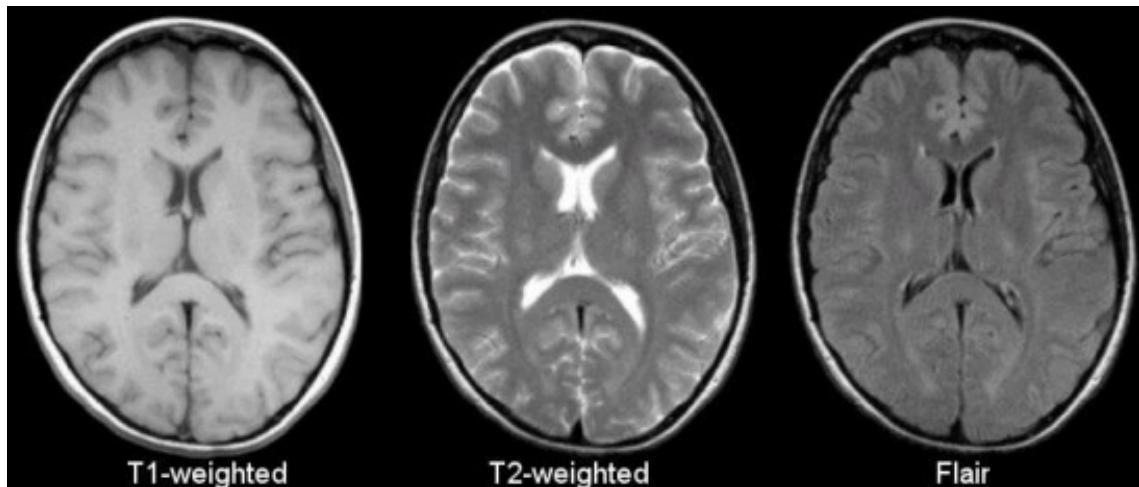
این مجموعه داده شامل تصاویر دوبعدی و سهبعدی MRI از چندین مؤسسه است که هر کدام دارای انواع اسکنر و پروتکل‌های تصویربرداری متفاوتی هستند و مجموعه متنوعی از داده‌های تصویربرداری را ارائه می‌دهند. این مجموعه داده همچنین شامل Annotation برای هر زیرگروه تومور، از جمله درجه، اندازه، مکان و شکل است که می‌تواند برای ناحیه‌بندی و طبقه‌بندی <sup>۷۰</sup> 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)

<sup>70</sup>Classification

در تصویر ۱۰-۲ یک مثال و مقایسه شهودی از سه نوع مدلایته آورده شده است. همچنین جدول ۲-۲ مقایسه دقیق‌تری از این مدلایتهای را ارائه می‌دهد.

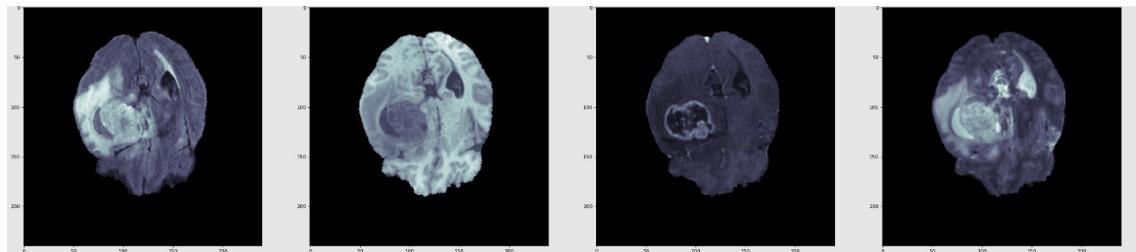


شکل ۲-۲: مقایسه سه مدلایته در تصاویر پزشکی [۳۰]

جدول ۲-۲: میزان روشنایی قسمت‌های مختلف مغز در هر مدلایته [۳۰]

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

در تصویر ۱۱-۲ یکی از نمونه‌های این مجموعه داده در حالت دوبعدی و در هر چهار مدلایته MRI نمایش داده شده است. از سمت چپ اولین تصویر مربوط به مدلایته FLAIR، دومین تصویر مربوط به مدلایته T1، سومین تصویر مربوط به مدلایته T1ce و آخرین تصویر مربوط به مدلایته T2 می‌باشد.



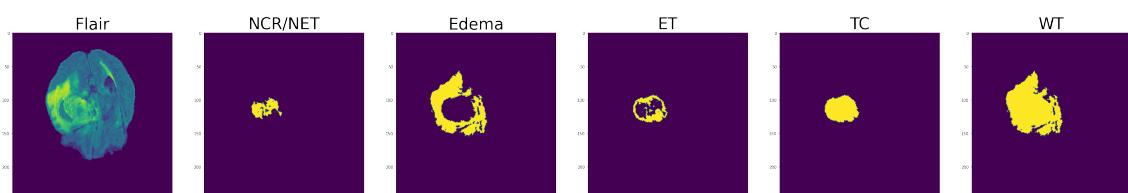
شکل ۲-۱۱: یکی از نمونه‌های BraTS در هر چهار مدالیته

هر اسکن MRI در این مجموعه داده با یک نقشه ناحیه‌بندی دودویی<sup>۷۱</sup> مربوطه همراه است که وکسل‌ها (معادل پیکسل در تصویر سه‌بعدی) را به عنوان یکی از موارد زیر برحسب‌گذاری 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 فزاينده است. مورد دوم به اطرافی اشاره دارد که محل نفوذ تومور به بافت سالم است. مورد سوم نمایانگر بخش فزاينده و در حال تقویت تومور است که معمولاً تهاجمی‌ترین قسمت تومور است.

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



شکل ۲-۱۲: نواحی مربوط به کلاس‌های مختلف در BraTS

<sup>71</sup>Binary Segmentation Mask

<sup>72</sup>Voxel

در هر مдалیته نواحی مربوط به کلاس خاصی از مغز برجسته‌تر می‌باشد. برای مثال در مDALITe T1 قسمت‌های مربوط به تومور فزاينده (ET) برجسته‌تر می‌باشد. یا در MDALITe FLAIR قسمت‌های مربوط به Edema برجسته‌تر می‌باشد. همچنین MDALITهای FLAIR و T1ce دارای مفیدترین و بیشترین اطلاعات در این مجموعه داده می‌باشند [۲۷].

### ۳-۵-۲ چالش‌های کار با BraTS

این مجموعه داده دارای چندین چالش برای توسعه الگوریتم‌های ناحیه‌بندی خودکار تومور می‌باشد. یکی از چالش‌های اصلی عدم تعادل کلاس‌ها<sup>۷۳</sup> در نقشه‌های ناحیه‌بندی است. هسته تومور و نواحی فزاينده تومور در مقایسه با نواحی Edema و پس زمینه، نسبتاً کوچک هستند که این امر ناحیه‌بندی این مناطق را چالش برانگیزتر می‌کند.

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

چالش دیگر اندازه بزرگ این مجموعه داده است. هر نمونه این مجموعه داده دارای عرض<sup>۷۴</sup> ۲۴۰ پیکسل و طول<sup>۷۵</sup> ۲۴۰ پیکسل و عمق<sup>۷۶</sup> ۱۵۵ می‌باشد. با در نظر گرفتن ۴ MDALITه MRI هر نمونه دارای  $240 \times 240 \times 240 \times 155 = 35712000$  عدد است در حالی که برای یک تصویر RGB دو بعدی با ابعاد مشابه، این عدد برابر  $172800 = 240 \times 240 \times 3$  می‌باشد. به عبارت دیگر هر نمونه این مجموعه داده معادل ۲۰۶ تصویر دو بعدی می‌باشد. این امر آموختش مدل‌های عمیق بر روی این مجموعه داده را بدون تجهیزات سخت‌افزاری مخصوص و قدرتمند بسیار مشکل می‌کند. البته می‌توان این مجموعه داده سه بعدی را به صورت دو بعدی 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 سه بعدی باشد، می‌توان آن را در قالب  $57195 = 369 \times 155 \times 240$  نمونه دو بعدی در نظر گرفت. در این صورت ابعاد هر نمونه دو بعدی به صورت  $4 \times 240 \times 240$  خواهد بود.

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

<sup>73</sup>Imbalance of Classes

<sup>74</sup>Width

<sup>75</sup>Height

<sup>76</sup>Depth

## فصل ۲. مروری بر منابع و پیشینه تحقیق ۲-۶. ناحیه‌بندی تعاملی با رویه پالایشی پس انتشار (BRS)

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

### ۴-۵-۲ 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<sup>۷۷</sup> استفاده می‌شود. این مجموعه داده همچنین برای توسعه مدل‌های یادگیری عمیق برای تشخیص خودکار<sup>۷۸</sup>، برنامه‌ریزی درمان<sup>۷۹</sup> و راهنمایی جراحی<sup>۸۰</sup> استفاده شده است.

### ۲-۶ ناحیه‌بندی تعاملی با رویه پالایشی پس انتشار [۱۸]<sup>۸۱</sup>

#### ۱-۶-۲ مقدمه

در این مقاله یک الگوریتم ناحیه‌بندی تعاملی تصویر، که نقطه‌هایی از کاربر در مورد شی هدف و پس زمینه دریافت می‌کند، ارائه شده است. با به دست آوردن فاصله مکانی هر پیکسل تا مکان نقطه‌های داده شده توسط کاربر، نقشه‌های تعامل<sup>۸۲</sup> ایجاد می‌شود. سپس در یک شبکه عصبی کانولوشنی عملیات پاس رو به جلو<sup>۸۳</sup> صورت می‌گیرد و یک نقشه اولیه ناحیه‌بندی به دست می‌آید. ممکن است در نتیجه اولیه نقطه‌های ورودی توسط کاربر برچسب اشتباهی بگیرند. بنابراین در این پژوهش رویه پالایشی پس انتشار توسعه داده شده است. این رویه پیکسل‌هایی که به اشتباه برچسب‌گذاری شده اند را تصحیح می‌کند. نتایج تجربی نشان می‌دهد که

<sup>77</sup>Survival Prediction

<sup>78</sup>Automated Diagnosis

<sup>79</sup>Treatment Planning

<sup>80</sup>Surgical Guidance

<sup>81</sup>Backpropagating Refinement Scheme (BRS)

<sup>82</sup>Interaction Map

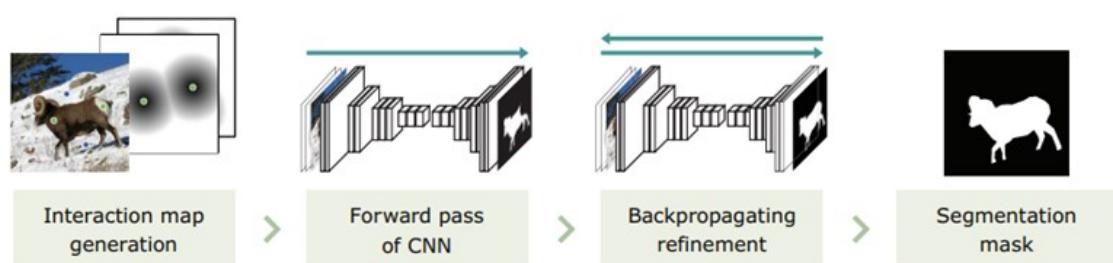
<sup>83</sup>Forward Pass

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

## ۲-۶-۲ الگوریتم

الگوریتم ناحیه‌بندی تعاملی ارائه شده یک ماسک دودویی<sup>۸۴</sup> 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 برچسب مورد انتظار برای پیکسل مربوطه هستند.

تصویر ۱۳-۲ 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 برچسب‌های مشخص شده در محل‌های کلیک شده هستند اما ممکن است نقشه احتمال<sup>۸۵</sup>، اطلاعات اشتباہی از محل‌های کلیک شده به دست آورد. بنابراین به وسیله BRS، محل‌های کلیک شده مجبور می‌شوند تا برچسب‌های مشخص شده توسط کاربر را داشته باشند. در نهایت ماسک ناحیه‌بندی شی هدف با اجرای دوباره پاس رو به جلو به دست می‌آید.



شکل ۲-۱۳: نگاه کلی به الگوریتم BRS [۱۸]

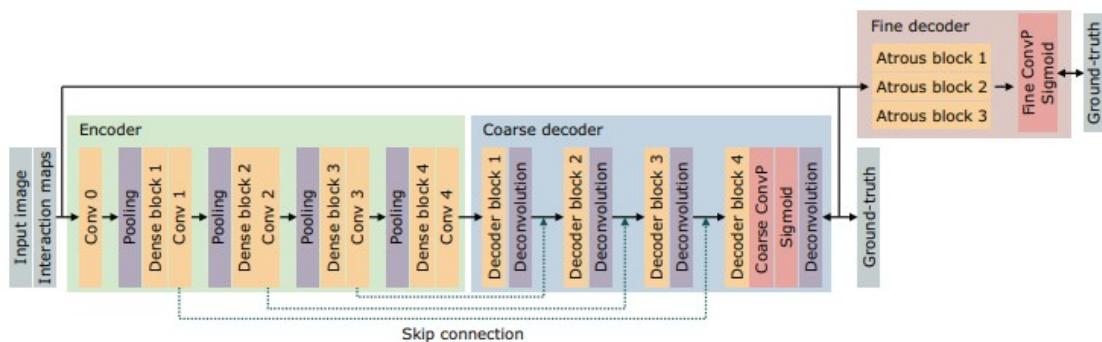
این فرایند هنگامی که کاربر اولین کلیک را روی شی هدف انجام می‌دهد آغاز می‌شود. سپس با به دست آوردن نتیجه ناحیه‌بندی، ممکن است کاربر محل جدیدی را روی شی هدف یا پس‌زمینه کلیک کند. این الگوریتم بارها اجرا می‌شود و ناحیه‌بندی‌های با دقت بالاتر به دست می‌آید. تا زمانی که کاربر دیگر کلیک نکند، این دو مرحله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

<sup>84</sup>Binary Mask

<sup>85</sup>Probability Mask

### ۳-۶-۲ معماری شبکه

شبکه عصبی کانولوشنی ارائه شده از معماری رمزگذار-رمزگشا<sup>۸۶</sup> استفاده می‌کند. معماری این شبکه در تصویر<sup>۱۴</sup> نمایش داده شده است. این شبکه دو نقشه تعامل و تصویر اصلی را ورودی می‌گیرد. از شبکه DenseNet<sup>[۱۶]</sup> به عنوان شبکه رمزگذار استفاده شده است تا ویژگی‌های سطح بالا و سطح پایین استخراج شود. از اتصال پرشی<sup>۸۷</sup> نیز برای نگهداری ویژگی‌های سطح پایین در لایه‌های عمیق شبکه استفاده شده است. در پایان هر بلاک متراکم<sup>۸۸</sup> از مازولهای فشرده‌ساز<sup>۸۹</sup> و برانگیختگی<sup>۹۰</sup> نیز استفاده شده است.



[۱۶] معماری ارائه شده بر پایه DenseNet

### ۴-۶-۲ مرحله آموزش

از مجموعه داده مزهای معنایی<sup>[۱۱]</sup> یا Semantic Boundaries Dataset (SBD) 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 در ۳۶۰ که نمایانگر تصویر و ماسک هدف است بریده<sup>۹۱</sup> شده است. همچنین از داده افزایی به صورت انعکاس افقی<sup>۹۲</sup> استفاده شده است.

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

<sup>86</sup>Encoder-Decoder

<sup>87</sup>Skip Connections

<sup>88</sup>Dense Block

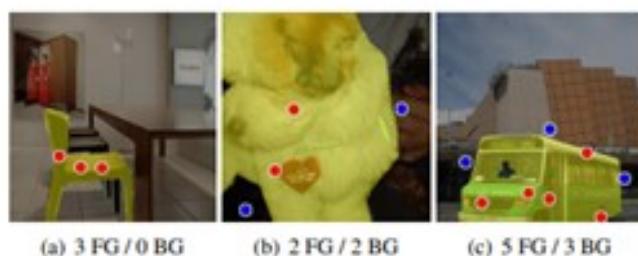
<sup>89</sup>Squeeze Module

<sup>90</sup>Excitation Module

<sup>91</sup>Crop

<sup>92</sup>Horizontal Flip

تصویر ۱۵-۲ آورده شده است.



شکل ۱۵-۲: مثالی از نقطه‌های تولید شده توسط خوشبند در مرحله آموزش [۱۸]

## ۵-۶-۲ 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 به روز می‌شوند. سپس نقشه‌های تعامل به روز 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.

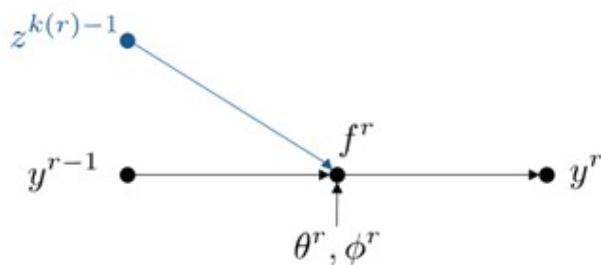
## ۶-۶-۲ رویه پالایشی پس‌انتشار

ممکن است کلیک‌های کاربر به اشتباه برچسب‌گذاری شوند. بنابراین یک معیار برای محاسبه میزان خطأ و یک رویه برای پس‌انتشار خطأ و کاهش آن مورد نیاز است. الگوریتم BRS ارائه شده به صورت تکراری<sup>۹۳</sup> عمل پس‌انتشار خطأ را انجام می‌دهد تا زمانی که تمامی کلیک‌ها برچسب درستی بگیرند. ابتدا نمادهای این الگوریتم توسط چند تنسور<sup>۹۴</sup> تعریف می‌شود که نمای کلی آن در تصویر ۱۶-۲ آورده شده است. تنسورهای  $y^{r-1}$  و  $z^{r-1}$  به هم پیوسته<sup>۹۵</sup> می‌شوند و پارامترهای  $\theta^r$  و  $\phi^r$  برای دستیابی به مورد استفاده قرار می‌گیرد.

<sup>93</sup>Iterative

<sup>94</sup>Tensor

<sup>95</sup>Concatenate



شکل ۲-۱۶: نمایش ریاضی رابطه [۱۸] BRS

این پارامتر بیانگر پاسخ لایه  $r$  شبکه است. مقدار  $y^r$  در رابطه ۱۴-۲ بیان شده است.

$$y^r = f^r(y^{r-1}, z^{r-1}, \theta^r, \phi^r) \quad (14-2)$$

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

## ۷-۶-۲ 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.

## فصل ۳

### روش پیشنهادی

#### ۱-۳ مقدمه

در این قسمت، بر روی روش ناحیه‌بندی تعاملی مبتنی بر کلیک تمرکز می‌شود. تعامل مبتنی بر کلیک در واقع همان تعامل مبتنی بر نقاط Seed می‌باشد که در اینجا با نام کلیک به آن اشاره می‌شود. تحت این روش، کاربر به طور متوالی کلیک‌های مثبت مانند نقاط قرمز در تصویر ۱-۳<sup>۱</sup> را برای تعریف FG و کلیک‌های منفی مانند نقاط سبز در تصویر ۱-۳<sup>۲</sup> را برای تعریف BG قرار می‌دهد و مدل پس از هر کلیک کاربر، پیش‌بینی‌های جدید را خروجی می‌دهد.



شکل ۳-۱: انواع کلیک [۵]

الگو اصلی برای ناحیه‌بندی تعاملی مبتنی بر کلیک، استفاده از نقشه‌ها یا دیسک‌های گاوی<sup>۱</sup> برای نمایش

<sup>۱</sup>Gaussian Maps

## ۲-۳. مشکلات روش‌های قبلی

کلیک‌های کاربر می‌باشد [۴۴]. در این روش یک نقشه گاووسی به مرکزیت کلیک کاربر و شعاع مشخص ایجاد می‌شود. در ادامه نقشه‌های کلیک با تصویر ورودی الحاق می‌شود، و وارد یک مدل ناحیه‌بند می‌شود تا ماسک اولیه پیش‌بینی شود.

بر اساس این روش، کارهای قبلی که مبتنی بر پارادایم‌های دیگر بود از جنبه‌های مختلفی بهبود می‌یابند. با این حال، هنگام اعمال این روش‌ها در سناریوهای عملی و صنعتی، نتیجه در جنبه‌هایی رضایت‌بخش نیست که در ادامه به این جنبه‌ها اشاره می‌شود.

## ۲-۳ مشکلات روش‌های قبلی

## ۱-۲-۳ کارایی نامناسب بر روی دستگاه‌های با توان محاسباتی کم

انتظار می‌رود که یک ابزار ناحیه‌بند خوب، ماسک‌های خوبی را در زمان کوتاهی تولید کند. بیشتر کارهای قبلی فقط بر دقت تمرکز دارند و از مدل‌های بزرگ با ورودی‌های باوضوح بالا<sup>۲</sup> استفاده می‌کنند. با این حال، وقتی روی لپتاپ‌های شخصی، دستگاه‌های لبه<sup>۳</sup> یا برنامه‌های مبتنی بر وب با درخواست‌های حجمی مستقر می‌شوند، با مشکل مواجه می‌شوند. همچنین هنگامی که سعی می‌شود اندازه ورودی برای عملکرد سریع تر کاهش داده شود، دقت این مدل‌ها به طور قابل توجهی کاهش می‌یابد.

## ۲-۲-۳ عدم استفاده از ماسک از پیش موجود

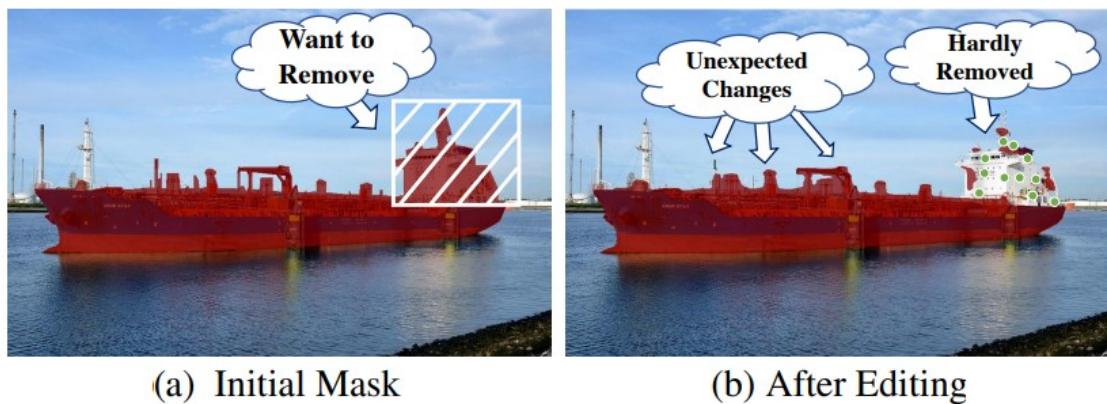
در کاربردهای عملی و در دنیای واقعی، ممکن است تصویری که می‌خواهیم روی آن ناحیه‌بندی صورت گیرد، دارای یک ماسک اولیه ناحیه‌بندی شده باشد. این ماسک‌های موجود می‌توانند توسط یک مدل آفلاین قبلی یا سایر روش‌های پیش‌پردازش ارائه شده باشد. همچنین ماسک قبلی تولید شده توسط خود مدل را نیز می‌توان در نظر گرفت. استفاده از این ماسک‌های موجود و ایجاد تغییرات در آنها می‌تواند عمل ناحیه‌بندی را تسهیل کند. با این وجود، روش‌های قبلی موجود با ماسک‌های خارجی و از پیش موجود سازگار نمی‌باشند. بسیاری از این روش‌ها حتی وجود ماسک قبلی را به عنوان ورودی در نظر نمی‌گیرند. البته در بعضی کارها ماسک قبلی با تصویر ورودی و نقشه‌های به دست آمده از کلیک‌های کاربر به هم ملحق می‌شوند و به عنوان ورودی به

<sup>2</sup>High-Resolution<sup>3</sup>Edge Devices

### فصل ۳. روش پیشنهادی

#### ۲-۳. مشکلات روش‌های قبلی

مدل داده می‌شوند، اما این روش‌ها نیز دارای عملکرد خوبی در استفاده از ماسک از پیش موجود نیستند. عدم استفاده از ماسک از پیش موجود باعث ایجاد مشکلات زیادی می‌شود. برای مثال در تصویر ۲-۳ قصد کاربر حذف کردن کابین کشتی از ماسک قبلی می‌باشد، بنابراین تعدادی کلیک منفی (سبر) بر روی کابین کشتی ایجاد می‌کند تا این قسمت حذف شود. اما پس از اجرای دوباره مدل متوجه می‌شویم مدل علاوه بر حذف کابین، قسمت‌های دیگری که از کلیک‌های کاربر فاصله زیادی دارند و در خروجی قبلی دارای برچسب درستی بوده اند را هم حذف کرده است. یا حتی قسمت‌های جدیدی را به اشتباه به ماسک خروجی اضافه کرده است.



شکل ۲-۳: تغییرات ناخواسته هنگام تولید خروجی جدید [۵]

در واقع هنگامی که کاربر بر روی یک ناحیه کلیک می‌کند، قصد دارد که مدل به آن ناحیه توجه کند. برای مثال در تصویر ۲-۳ کاربر قصد دارد مدل به کابین کشتی توجه کند نه به محیط اطراف کشتی. اما بسیاری از روش‌های قبلی به قصد کاربر توجهی ندارند و در هنگام تولید خروجی جدید با تمامی پیکسل‌های تصویر به صورت برابر رفتار می‌کنند. این امر باعث ایجاد دو پدیده می‌شود. اول اینکه نواحی‌ای که قبلاً به صورت خوبی ناحیه‌بندی شده بودند نیازمند ناحیه‌بندی دوباره هستند که این پدیده افزونگی محاسبات<sup>۴</sup> نامیده می‌شود. دوم اینکه تمامی پیکسل‌های ماسک قبلی دوباره باید ناحیه‌بندی شوند و مدل اطلاعات قبلی موجود از آنها را از دست می‌دهد که این پدیده از دست دادگی اطلاعات<sup>۵</sup> نامیده می‌شود.

<sup>4</sup>Computational Redundancy

<sup>5</sup>Loss of Information

### ۳-۳ ویژگی‌های الگوریتم ارائه شده

ویژگی‌های الگوریتم ارائه شده باقیستی در راستا و متناظر با حل مشکلات اشاره شده باشد. توجه و تمرکز این الگوریتم باید بر روی نواحی محلی<sup>۶</sup> خاصی باشد. همچنین پیش‌بینی جدید تنها برای ناحیه‌هایی که نیازمند محاسبات دوباره هستند صورت گیرد و الگوریتم تنها خروجی ناحیه‌هایی را به روزرسانی کند که کاربر هم قصد به روزرسانی آن را داشته است. علاوه بر این موارد الگوریتم باید در نواحی دیگر خروجی قبلی را حفظ کند. در این پروژه با توجه به نکات، مشکلات و اهداف اشاره شده، از مدل و الگوریتم FocalClick [۵] برای ناحیه‌بندی تعاملی تومور مغزی استفاده خواهیم کرد. در ادامه بخش‌های مختلف این الگوریتم طبق مقاله [۵] توضیح داده می‌شود.

### ۴-۳ الگوریتم در یک نگاه

هنگامی که کاربر کلیک جدیدی انجام می‌دهد، ابتدا یک برش از تصویر اصلی با نام Target Crop ایجاد می‌شود. این برش با توجه به ماسک از پیش موجود ساخته می‌شود. در شکل ۳-۳ این برش با یک زرد رنگ نمایش داده شده است. Box

سپس این Target Crop دچار تغییر اندازه<sup>۷</sup> می‌شود. همچنین از وضوح آن نیز کاسته می‌شود. بر روی این تصویر کوچک و با وضوح کم یک ناحیه‌بندی درشت دانه صورت می‌گیرد که این عمل تحت نام Coarse Segmentaion شناخته می‌شود.

سپس یک ناحیه کوچک دیگر که کاربر قصد تغییر خروجی ناحیه‌بندی در آن جا را داشته، شناسایی می‌شود. با توجه به اینکه تمرکز مدل باید بر روی این ناحیه باشد، آن را Focus Crop می‌نامیم که در شکل ۳-۳ با یک محیط قرمز رنگ مشخص شده است. سپس این ناحیه مورد توجه وارد قسمت پالایش یا Local Refinement می‌شود تا کیفیت و وضوح از دست رفته خود را بازیابی کند. بنابراین عملیات سنگین و زمان‌بر ناحیه‌بندی یک تصویر بزرگ، به دو عملیات محلی و با هزینه کم تبدیل می‌شود.

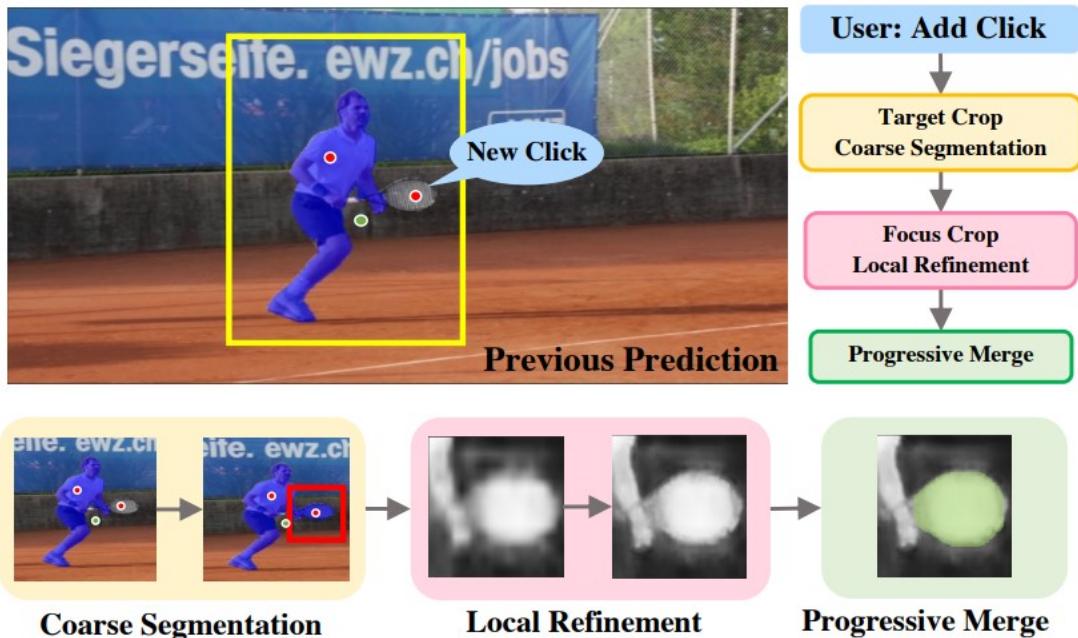
در نهایت ناحیه پالایش شده وارد قسمت ادغام پیش‌رونده یا Progressive Merge می‌شود تا با یک تحلیل مورفولوژیکی<sup>۸</sup> بر روی ماسک از پیش موجود و خروجی جدید، تصمیم بر به روزرسانی یا حفظ برچسب‌های

<sup>6</sup>Local

<sup>7</sup>Resize

<sup>8</sup>Morphological

قبلی گرفته شود. بنابراین نواحی ای که قبلاً درست پیش‌بینی شده بودند دچار تغییر نمی‌شوند.



[۵] شکل ۳-۳: نگاهی کلی به الگوریتم FocalClick

با توجه به توضیحات داده شده می‌توان متوجه شد که روش بالا بسیار متکی به عمل برش یا Crop است که این عمل در بعضی کارهای دیگر ناحیه‌بندی نیز دیده می‌شود. اما تفاوت این الگوریتم با سایر روش‌ها در این است که کارهای دیگر نیازمند حفظ وضوح تصویر برش یافته هستند و ناحیه‌بند آنها یک خروجی با جزئیات دقیق تولید می‌کند که این کار زمانبر است و با سرعت کمی صورت می‌گیرد. اما در این الگوریتم وضوح ناحیه برش یافته کاهش یافته و خروجی ناحیه‌بند سریع و درشت دانه است.

## ۳-۵ جزئیات الگوریتم

در این قسمت، بخش‌ها و مراحل الگوریتم با ذکر جزئیات مورد بررسی قرار می‌گیرد. ابتدا مدل تصویر ورودی، دو نقشه مربوط به کلیک‌های منفی و مثبت و ماسک از پیش موجود را به عنوان ورودی دریافت می‌کند. برای ایجاد نقشه مربوط به کلیک‌های کاربر، از یک دیسک دودویی<sup>۹</sup> با شعاع ۲ استفاده می‌شود. به عبارت دیگر

<sup>9</sup>Binary Disk

### ۳-۵. جزئیات الگوریتم

ناحیه درون دایره‌ای با شعاع ۲ و مرکز کلیک برچسب ۱ و بیرون این دایره برچسب ۰ می‌گیرد. سپس Target Crop پیرامون شی هدف ایجاد می‌شود و اندازه ووضوح آن کاهش می‌یابد. سپس این برش به شبکه ناحیه‌بند داده می‌شود تا خروجی درشت دانه (Coarse) ایجاد شود. ناحیه Focus Crop با محاسبه روی تفاوت نواحی میان ماسک قبلی و ناحیه‌بندی درشت دانه ایجاد می‌شود. از این برش برای پالایش و بازیابی جزئیات تصویر درشت دانه استفاده می‌شود. در نهایت بخش Progressive Merge ماسک قبلی و ناحیه پالایش شده را ادغام می‌کند. در ادامه جزئیات هر مرحله از این الگوریتم توضیح داده می‌شود.

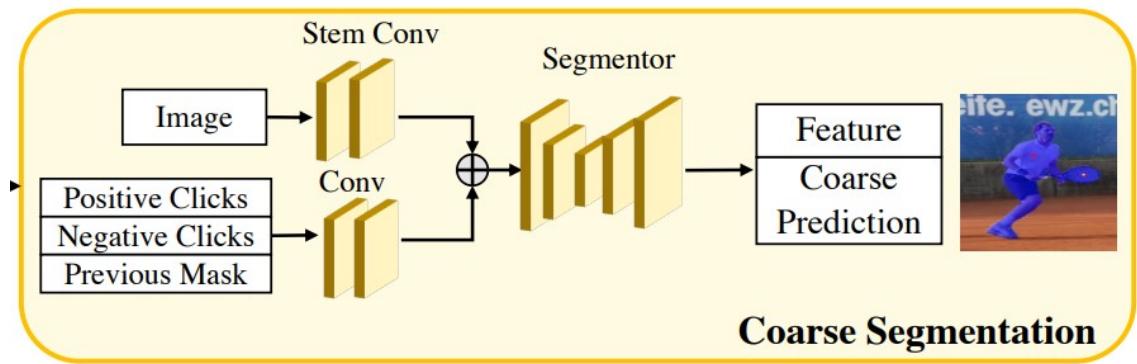
### ۱-۵-۳ بخش Target Crop

هدف این قسمت فیلتر کردن ناحیه BG از ناحیه مربوط به شی هدف است. برای ایجاد این ناحیه یک Bounding Box خارجی کمینه که شامل ماسک (شی هدف) از پیش موجود و کلیک جدید کاربر می‌باشد ایجاد می‌شود. در ادامه این Bounding Box با هایپرپارامتر <sup>۱۰</sup> ratio\_TC مشاهده می‌شود مستطیل زرد رنگ شامل کلیک جدید نظر گرفته شده است. همانطور که در تصویر ۱-۳ مشاهده می‌شود مستطیل زرد رنگ شامل کلیک جدید کاربر و شی هدف می‌باشد و ابعاد آن کمی بسط یافته است. پس از گسترش این Bounding Box و به دست آمدن مختصات آن، سه تنسور مربوط به تصویر، ماسک قبلی و نقشه‌های کلیک کاربر با توجه به این Bounding Box برش می‌خورند. در نهایت وضوح و اندازه این برش‌های به دست آمده کم می‌شود تا آمده قسمت بعدی شود.

### ۲-۵-۳ بخش Coarse Segmentation

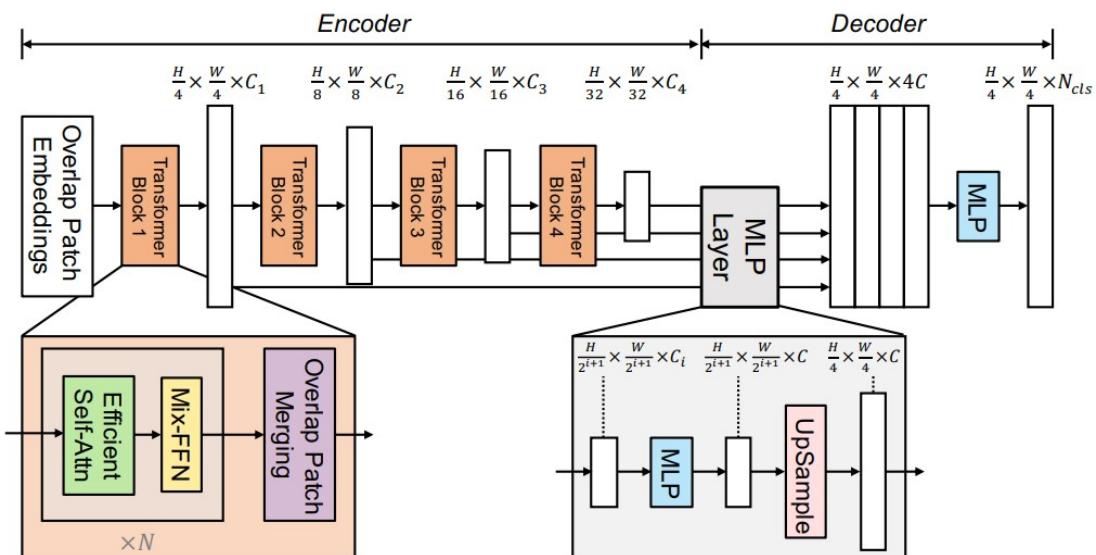
در این قسمت ورودی‌های دودویی مربوط به نقشه کلیک‌های منفی، کلیک‌های مثبت و ماسک از پیش موجود وارد دو لایه کانولوشنی می‌شوند تا در بعد کanal مشابه تصویر ورودی شوند. تصویر ورودی هم از دو لایه کانولوشنی عبور می‌کند تا ویژگی‌های سطح پایین آن به دست آید. سپس ویژگی‌های به دست آمده از تصویر و سه ورودی دودویی با هم ترکیب می‌شوند.

<sup>10</sup>Hyperparameter

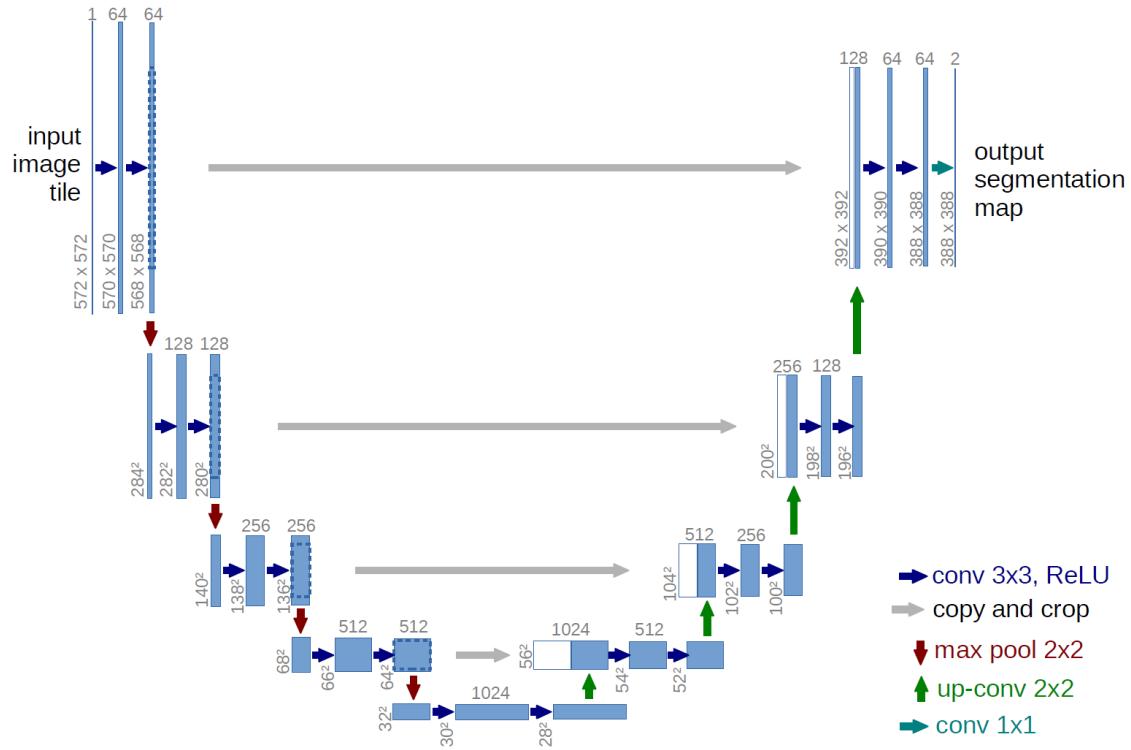


[۴] شکل ۳-۴: معماری قسمت Coarse Segmentation

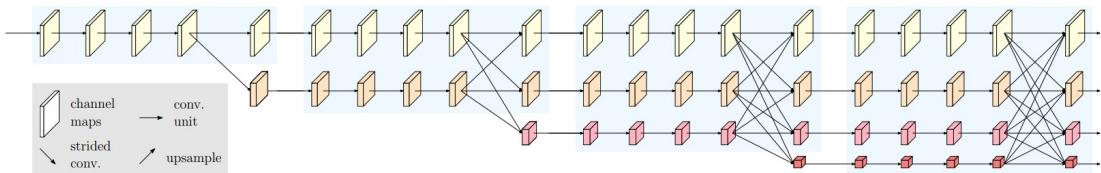
این ویژگی‌های ترکیب شده وارد یک شبکه ناحیه‌بند می‌شوند. این شبکه می‌تواند هر شبکه ناحیه‌بندی مانند Transformer [۴۱] (مناسب تصاویر با وضوح بالا) و یا شبکه‌های ناحیه‌بند مبتنی بر U-Net [۳۲] مانند SegFormer [۴۳] باشد. ویژگی مشترک این شبکه‌ها وجود معماری Encoder-Decoder و همچنین اتصالات پرشی برای حفظ اطلاعات تصویر ورودی می‌باشد. در نهایت خروجی ناحیه‌بند Coarse تولید می‌شود.



[۴۳] شکل ۳-۵: معماری شبکه SegFormer



شکل ۳-۶: معماری شبکه U-Net



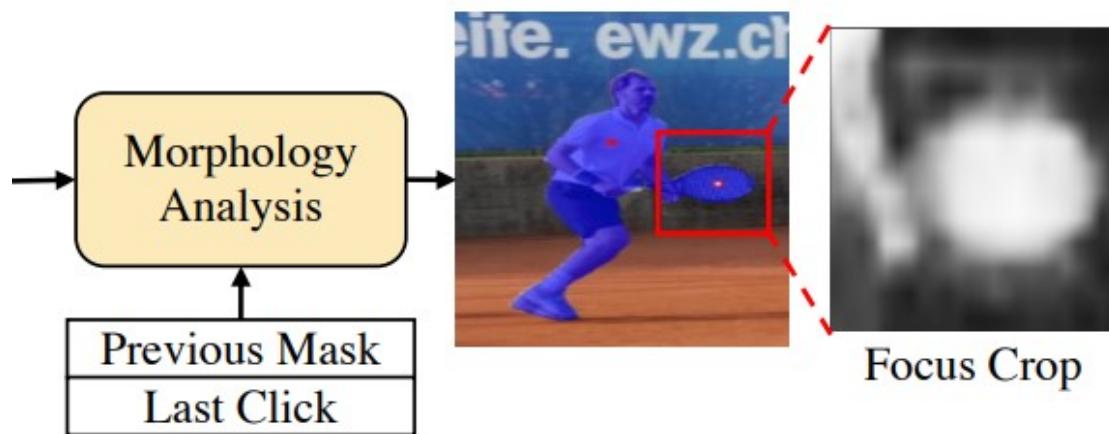
شکل ۳-۷: معماری شبکه HRNet

### ۳-۵-۳ بخش Focus Crop

این قسمت ابتدا تفاوت ماسک از پیش موجود و خروجی قسمت Coarse Segmentation را محاسبه می‌کند. این مقایسه به وسیله تابع منطقی XOR صورت می‌گیرد و نتیجه آن یک ماسک جدید با نام Difference Mask (DM) می‌باشد. سپس بیشینه ناحیه متصل درون این ماسک که شامل کلیک جدید کاربر می‌باشد محاسبه می‌شود. سپس یک Bounding Box دور بیشینه ناحیه متصل ایجاد می‌شود. مشابه قسمت Target Crop، این به دست آمده با هایپرپارامتر ratio\_FC بسط می‌یابد. مقدار این هایپرپارامتر را هم برابر  $1/4$  در Bounding Box

## ۳-۵. جزئیات الگوریتم

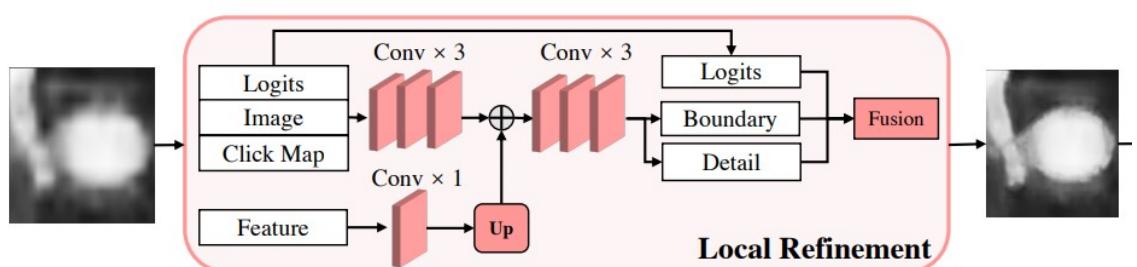
نظر می‌گیریم. در نهایت Focus Crop که در تصویر ۳-۲ با رنگ قرمز نمایش داده شده است به دست می‌آید. مطابق Focus Crop به دست آمده، تصویر ورودی و نقشه‌های کلیک کاربر نیز Crop می‌شوند. همچنین از ماژول [۱۲] برای ایجاد برش روی Feature ها و Logit های خروجی ناحیه‌بند استفاده شده است.



شکل ۳-۸: معماری قسمت [۵] Focus Crop

## ۴-۵-۳ بخش Local Refinement

این قسمت، جزئیات پیش‌بینی درشت دانه صورت گرفته بر روی Focus Crop را بازیابی و پالایش می‌کند. ابتدا ویژگی سطح پایین با استفاده از لایه‌های کانولوشنی Xception [۶] از تنسورهای برش یافته استخراج می‌شود. در همین حال، تعداد کانال‌های مربوط به ویژگی‌های RoiAlign تنظیم می‌شود و سپس با ویژگی‌های سطح پایین به دست آمده ادغام می‌شود.



شکل ۳-۹: معماری قسمت [۵] Local Refinement

### فصل ۳. روش پیشنهادی

#### ۳-۵. جزئیات الگوریتم

برای به دست آوردن پیش‌بینی‌های پالایش شده، از دو Head برای پیش‌بینی نقشه جزئیات  $M_d$  و نقشه مرزی  $M_b$  استفاده می‌کنیم. سپس نقشه پیش‌بینی پالایش شده  $M_r$  با استفاده از رابطه زیر به دست می‌آید. در این رابطه  $M_l$  بیانگر Logit‌های پیش‌بینی شده توسط شبکه Coarse Segmentor می‌باشد.

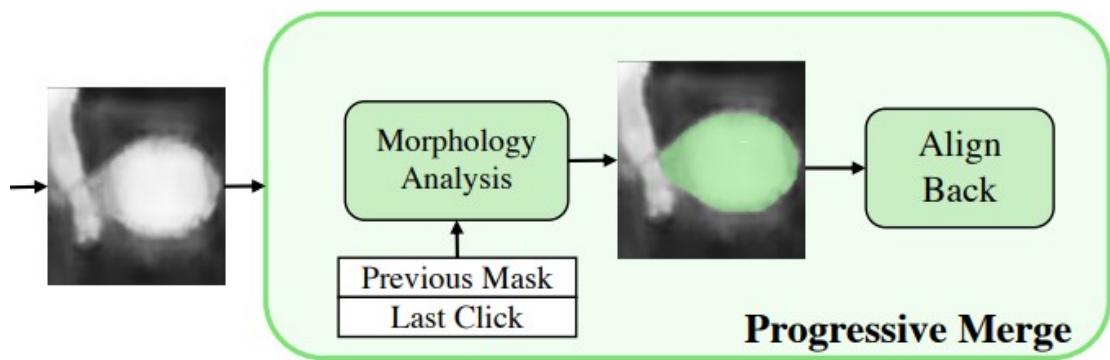
$$M_r = \text{Sigmoid}(M_b) * M_d + (1 - \text{Sigmoid}(M_b)) * M_l \quad (1-3)$$

#### ۵-۵-۳ بخش Progressive Merge

هنگام برچسب‌گذاری یا ویرایش ماسک‌ها، انتظار نداریم که مدل پس از هر کلیک، ماسک را برای همه پیکسل‌ها به روز کند. در غیر این صورت تمامی جزئیاتی که الگوریتم به خوبی برچسب زده دوباره به روزرسانی می‌شود و حتی ممکن است برچسب صحیح قبلی از بین برود. کاربر فقط می‌خواهد در مناطق محدودی که قصد اصلاح آن‌ها را دارد به روزرسانی صورت گیرد. در قسمت Progressive Merge نیز مشابه قسمت‌های Focus Crop و Target Crop از آنالیز مبتنی بر مورفولوژی استفاده می‌شود تا ناحیه‌ای که نیازمند اصلاح است به دست آید.

پس از وارد شدن یک کلیک کاربر، به سادگی ماسک پیش‌بینی شده جدید با آستانه <sup>۱۱</sup> برابر ۰/۵ بازتری می‌شود و ناحیه متفاوتی بین پیش‌بینی جدید و ماسک موجود محاسبه می‌شود. سپس بیشینه ناحیه متصل که حاوی کلیک جدید است به عنوان منطقه به روز رسانی انتخاب می‌شود (ناحیه سیز رنگ در تصویر ۱۰-۳). در این ناحیه، ماسک جدید پیش‌بینی شده روی ماسک قبلی قرار داده می‌شود و به روز می‌شود. همچنین ماسک قبلی در نواحی دیگر دست نخورده باقی می‌ماند.

<sup>11</sup> Threshold

شکل ۳-۱۰: معماری قسمت <sup>۵</sup> Progressive Merge

هنگامی که الگوریتم با یک ماسک از قبل موجود اجرا می‌شود یا یک ماسک اولیه از سایر ابزارهای ناحیه‌بندی وجود دارد، قسمت Progressive Merge برای حفظ جزئیات صحیح اعمال می‌شود. اما هنگام برچسب‌گذاری اولیه و بدون وجود ماسک قبلی از ابتدا، قسمت Progressive Merge پس از ۱۰ کلیک فعال می‌شود.

### ۳-۶ تابع خطای

با توجه به این که این الگوریتم دارای قسمت‌های مختلف می‌باشد، برای هر کدام از این بخش‌ها یک تابع خطای در نظر گرفته می‌شود و با ترکیب این توابع، تابع نهایی خطای به دست می‌آید و الگوریتم طبق نظارت <sup>۱۲</sup> این تابع آموزش داده می‌شود. تابع خطای نقشه مرزی  $M_b$  از ۸ مرتبه Down Sample کردن ناحیه‌بندی مربوط به Ground Truth و سپس Up Sample کردن آن به اندازه قبلی به دست می‌آید. در این حالت پیکسل‌های تغییر یافته می‌توانند نمایانگر ناحیه‌هایی باشند که نیازمند دقت و جزئیات بیشتری هستند. از تابع خطای Binary Cross Entropy (BCE) برای نظارت روی Boundary Head استفاده می‌شود. همچنین برای نظارت بر روی خروجی ناحیه‌بند Coarse NFL از تابع خطای Normalized Focal Loss (NFL) استفاده می‌شود. برای نظارت بر روی قسمت Local Refinement نیز از تابع خطای NFL وزن‌دار استفاده می‌شود. این وزن برابر  $b = 1.5$  در نظر گرفته می‌شود و تابع جدید وزن‌دار به اختصار BNFL نامیده می‌شود. در نهایت تابع خطای نهایی کل شبکه از جمع این ۳ تابع خطای به دست می‌آید.

<sup>12</sup>Supervision

$$L = L_{BCE} + L_{NFL} + L_{BNFL} \quad (2-3)$$

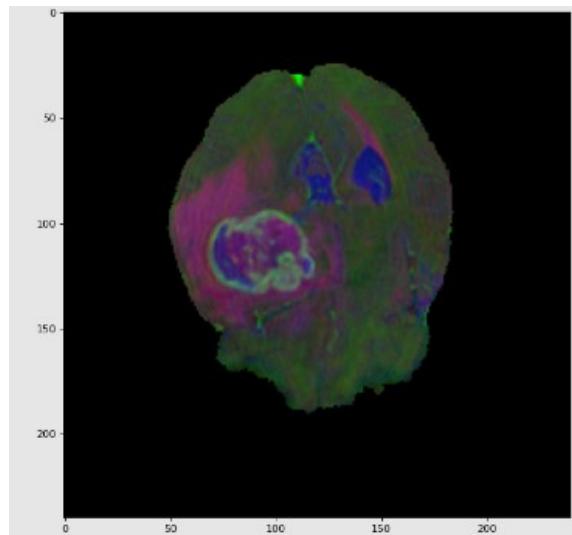
## ۷-۳ ترکیب کردن مدلیتهای مختلف MRI

همانطور که در فصل دوم اشاره شد، هر نمونه تصویر MRI مجموعه داده BraTS در چهار مدلیته FLAIR، T1ce، T1 و T2 موجود است. همچنین هر کدام از این مدلیتهای بخش‌های خاصی از مغز یا تومور مغزی را بهتر برجسته می‌کنند. این مدلیتهای Grayscale وجود دارند و هر پیکسل آنها دارای شدتی بین ۰ تا ۱ می‌باشد. در کارهای دیگر که بر روی ناحیه‌بندی تصاویر MRI مغزی تمرکز دارند معمولاً از یکی از این مدلیتهای استفاده می‌شود و مدلیته استفاده شده اغلب از نوع FLAIR یا T1ce می‌باشد [۲۳، ۴۰].

یکی از نوآوری‌های اثربخش به کار گرفته شده در این پژوهه استفاده همزمان از سه مدلیته T1ce، FLAIR و T2 می‌باشد. برای انجام این کار، مشابه تصاویر معمولی RGB که از ادغام سه کanal قرمز، سبز و آبی شکل گرفته اند عمل شده است. یعنی سه مدلیته FLAIR، T1ce و T2 با هم Concatenate شده و یک تصویر سه کanalه ایجاد شده است. برای اشاره راحت‌تر به این حالت آن را حالت Mix می‌نامیم.

هر چند تصویر تولید شده از ادغام این سه مدلیته که نمونه‌ای از آن در تصویر ۱۱-۳ آورده شده است برای انسان دارای معنی و اطلاعات خاصی نیست و حتی باعث اشتباه می‌شود، اما مدل مبتنی بر یادگیری عمیق می‌تواند از هر کanal، ویژگی‌های مورد نیاز خود را استخراج کند تا خروجی بهتری تولید کند. همچنین مزیت دیگر این کار این است که تصویر ورودی به مدل مشابه تصاویر RGB می‌باشد که در این صورت می‌توان از مدل‌های از پیش آموخته<sup>۱۳</sup> شده که بر روی مجموعه دادگان بزرگی مثل ImageNet [۵] آموزش دیده شده اند راحت‌تر استفاده کرد. پیش از این برای استفاده از مدل‌های از پیش آموخته بایستی مدلیته انتخاب شده سه بار تکرار می‌شد که نه تنها برای مدل اطلاعات جدیدی نداشت، بلکه تعداد پارامترهای مدل را هم بسیار زیادتر می‌کرد.

<sup>13</sup>Pretrained



شکل ۳-۱۱: نمایش تصویر ایجاد شده در حالت ترکیب مدلیتهای

## فصل ۴

### آزمایش‌ها و نتیجه‌ها

#### ۱-۴ آماده‌سازی مجموعه داده BraTS

الگوریتم FocalClick مناسب تصاویر رنگی طبیعی با وضوح بالا می‌باشد و ورودی قسمت‌های Refiner و Segmentor آن حتی پس از انجام عمل Crop به صورت  $256 \times 256$  می‌باشد. همچنین این الگوریتم مناسب تصاویر دو بعدی RGB می‌باشد. در حالی که مجموعه داده BraTS شامل تصاویر سه بعدی در چهار مدالیته با اندازه  $155 \times 240 \times 240$  می‌باشد. بنابراین منطبق کردن این الگوریتم و این مجموعه داده نیازمند تغییراتی در هر دو می‌باشد. ابتدا اندازه ورودی قسمت‌های Refiner و Segmentor را از  $256 \times 256$  به  $96 \times 96$  تبدیل می‌کنیم تا مناسب اندازه تصاویر BraTS شود. همچنین با توجه به این که پردازش تصاویر و آموزش آن‌ها به صورت سه بعدی نیازمند قدرت محاسباتی بسیار قوی و حافظه زیادی می‌باشد، و همچنین عدم وجود مأذول‌های سه بعدی نظیر RoiAlign در ابزارهای یادگیری عمیق مثل TensorFlow و PyTorch، مجموعه داده BraTS را به صورت دو بعدی وارد شبکه می‌کنیم. برای این کار هر اسلایس مربوط به Depth این مجموعه داده را یک نمونه جدید در نظر می‌گیریم. از مدالیتهای FLAIR، T1ce و حالت ابداعی Mix نیز در آزمایش‌ها استفاده شده است.

## ۴-۲ توزیع دیتاست به مجموعه‌های آموزش، اعتبارسنجی و تست

با توجه به در دسترس عموم نبودن مجموعه تست BraTS، از مجموعه داده آموزشی BraTS سال ۲۰۲۰ برای آموزش، اعتبارسنجی و تست الگوریتم استفاده شده است. در این نسخه از مجموعه داده ۳۶۹ نمونه سه‌بعدی در چهار مدالیته وجود دارد. از این ۳۶۹ نمونه، ۷۵ درصد برای آموزش مدل، ۲۰ درصد برای اعتبارسنجی مدل و ۵ درصد برای تست مدل جدا شده است. در جدول ۱-۴ خلاصه‌ای از تعداد نمونه‌ها در حالت دو بعدی و سه‌بعدی مربوط به مجموعه‌های آموزش، اعتبارسنجی و تست آورده شده است.

جدول ۱-۴: توزیع مجموعه‌های آموزش، اعتبارسنجی و تست

	Percentage	3D Cases	2D Cases
<b>Training Set</b>	75 %	276	42780
<b>Validation Set</b>	20 %	73	11315
<b>Testing Set</b>	5 %	20	3100
<b>Total</b>	100 %	369	57195

## ۴-۳ محیط و هایپرپارامترهای عمومی آموزش

تعدادی از هایپرپارامترها و مشخصه‌ها در تمامی آزمایش‌های صورت گرفته ثابت می‌باشد که در این قسمت به این موارد اشاره می‌شود. تمامی مدل‌های برای ۲۰ ایپاک<sup>۱</sup> آموزش داده شده اند. همچنین نمونه‌ها در هایی با اندازه ۳۲ وارد مدل می‌شوند تا مدل قادر باشد عملیات را موازی‌سازی کند. استفاده از Batch های بزرگ‌تر باعث پرشدن حافظه و توقف الگوریتم می‌شود. همچنین Crop و کاهش وضوح به نحوی انجام می‌شود که نواحی Focus Crop و Target Crop دارای اندازه  $96 \times 96$  شوند. هایپرپارامترهای ratio\_FC و ratio\_TC که برای بسط دادن ابعاد Bounding Box زرد و قرمز بود نیز برابر  $1/4$  در نظر گرفته شده است. آزمایش‌ها در دو محیط Google Colab با پردازشگر گرافیکی Nvidia K80 / T4 و Nvidia P100 / Kaggle با پردازشگر گرافیکی Nvidia T4 / ۱ Epoch انجام شده است. لینک این محیط‌ها، نتایج آزمایش، تصاویر خروجی

<sup>1</sup>Epoch

مدل در هر مرحله از آموزش (Visualization)، وزن مدل‌ها و پیاده‌سازی الگوریتم در قسمت پیوست‌ها آورده شده است.

## ۴-۴ نحوه تعریف و اجرای آزمایش‌ها

برای تعریف و اجرای آزمایش‌ها، کافی است یکی از نوتبوک‌های Google Colab یا Kaggle اجرا شود. طراحی این نوتبوک‌ها به گونه‌ای می‌باشد که با تغییر متغیرهای موجود، می‌توان آزمایش‌های مختلفی انجام داد. برای مثال می‌توان از شبکه‌های ناحیه‌بند موجود یکی را به عنوان Backbone انتخاب کرد. همچنین امکان انتخاب مدالیته‌های گوناگون هم در فاز آموزش و هم در فاز ارزیابی وجود دارد. تسلیک مورد نظر کاربر نیز قابل تعریف می‌باشد و کاربر می‌تواند ناحیه و ماسک مطلوب خود را تعریف کند. جزئیات مربوط به نحوه تعریف آزمایش در تصویر ۱-۴ و تصویر ۲-۴ نمایش داده شده است. علاوه بر موارد بالا، امکان آموزش مدل بر روی چندین به صورت همزمان نیز وجود دارد.

```

MODE = "FocalClick" # CDNet, FocalClick
BACKBONE = "segformerB3" # cdnet_res34, hrnet18s, hrnet32, segformerB3

# CDNet-ResNet34, FocalClick-HRNet18s, FocalClick-HRNet32, FocalClick-SegFormerB3
EXP_NAME = "FC-SF-Mix-WT"

CHANNEL = "mix" # flair, t1, t1ce, t2, mix
LABEL = "wt" # net, ed, et, wt, tc
EPOCHS = 20
BATCH_SIZE = 32
NGPUS = 1
WORKERS = 2

MODEL_FILE = f"{REPO_PATH}/models/brats/{BACKBONE}.py"
MODEL_DIR = f"{REPO_PATH}/experiments/brats/{BACKBONE}/000_{EXP_NAME}/checkpoints/"

# Set --pretrained in below cell

```

شکل ۴-۱: نحوه تعریف متغیرهای فاز آموزش

```

MODE = "FocalClick" # CDNet, FocalClick
BACKBONE = "segformerB3" # cdnet_res34, hrnet18s, hrnet32, segformerB3
EXP_NAME = "000_FocalClick-SegFormerB3"
CHECKPOINT = "epoch-19-val-loss-1.28.pth"

CHANNEL = "t1ce" # flair, t1, t1ce, t2, mix
LABEL = "tc" # net, ed, et, wt, tc

MODEL_DIR = f"{REPO_PATH}/experiments/brats/{BACKBONE}/{EXP_NAME}/checkpoints/"

```

شکل ۴-۲: نحوه تعریف متغیرهای فاز ارزیابی

## ۵-۴ نکات کلی درباره نتایج آزمایش‌ها

همواره معیار Dice در مقایسه با معیار IoU دارای مقدار بیشتری می‌باشد که این اتفاق به دلیل وجود ضریب ۲ در رابطه Dice می‌باشد. در تحلیل و تفسیر نتایج نیز تمرکز بیشتر بر روی معیار Dice می‌باشد و معیار IoU جهت مقایسه نتایج این الگوریتم با سایر مقالاتی که نتایج خود را فقط در قالب IoU ارائه کرده‌اند، می‌باشد. همچنین زیاد بودن معیار NoF به این دلیل است که بسیاری از اسلامیس‌های تصویر سه‌بعدی تبدیل به دو‌بعدی شده فاقد پیکسل FG می‌باشند. در این موقع، الگوریتم این نمونه‌ها را از نوع Failure شمرده است. در تصویر اول ارائه شده برای هر آزمایش، در سمت بالا و با شروع از سمت چپ هر قسمت به ترتیب بیانگر Previous Full Prediction سپس Full Ground Truth و در سمت راست Full Image می‌باشد. در پایین تصویر و با شروع از سمت چپ نیز هر قسمت به ترتیب بیانگر Refine Image سپس Refine Prediction و در سمت راست Trimap Prediction سپس Focus Prediction (ترکیب نقشه‌های کلیک منفی و مثبت) می‌باشد.

در تصویر دوم ارائه شده برای هر آزمایش، ناحیه و IoU به دست آمده به ازای ۲۰ کلیک نشان داده شده است. در این تصویر محل کلیک‌های مثبت و منفی نیز نمایش داده شده است. همچنین محل مربوط به Bounding Box با Focus Crop و Target Crop و تصویر Ground Truth و خروجی نهایی مدل و خروجی ماقبل نهایی مدل نیز آورده شده است. در تصویر سوم و آخر ارائه شده برای هر آزمایش نیز، ناحیه Focus Crop به دو صورت Coarse و Refined

آورده شده است.

## HRNet18s-FLAIR-WT-Pretrained : ۱-۶ آزمایش ۱

در این آزمایش، الگوریتم FocalClick با استفاده از Backbone و ناحیه‌بند HRNet18s که از پیش بر روی مجموعه آموزش داده شده است، اجرا می‌شود. همچنین از مدلیته FLAIR برای ایجاد مجموعه داده استفاده شده است. تسك و هدف مورد نظر آزمایش نیز پیدا کردن ناحیه مربوط به کل تومور یا Whole Tumor می‌باشد. پس از اینکه آموزش مدل کامل شد، ارزیابی مجموعه تست، بر روی همان مدلیته و همان تسك مرحله آموزش صورت می‌گیرد.

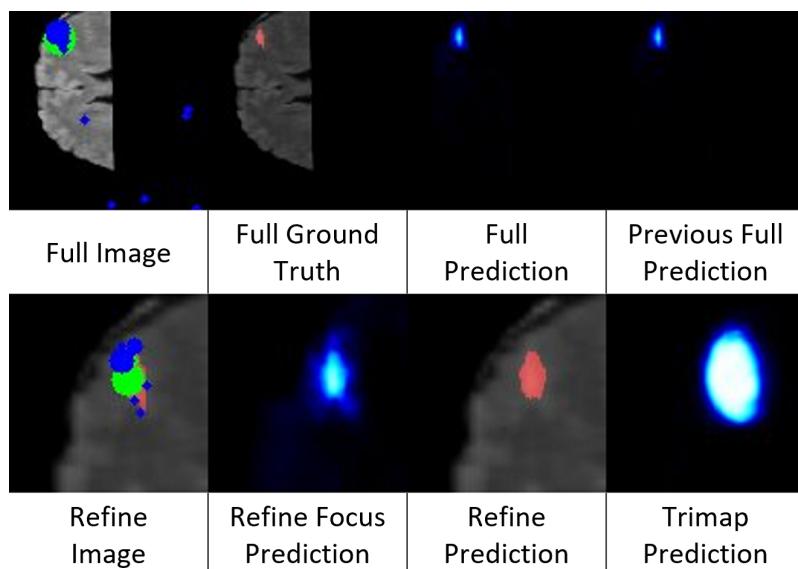
جدول ۲-۱: معیار NoC و NoF در آزمایش ۱

Metric	NoC80%	NoC85%	NoC90%	NoF85%	NoF90%	TPC	Time
<b>IoU</b>	3.47	4.31	6.00	325	568	0.061	0:56:38
<b>Dice</b>	2.50	2.92	3.72	180	259	0.061	0:56:38

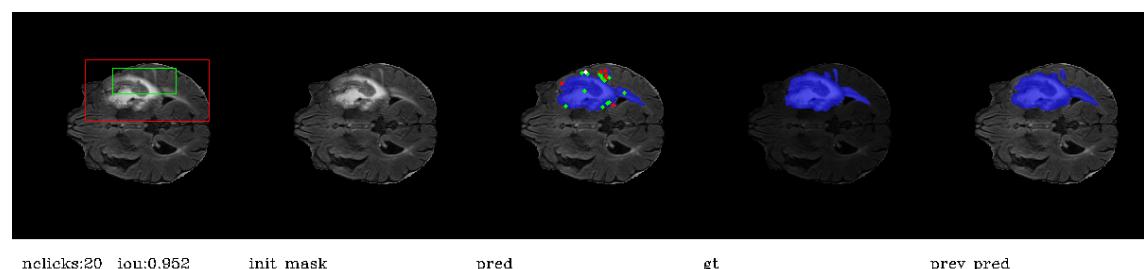
جدول ۳-۱: معیار mIoU و mDice در آزمایش ۱

Metric	Clicks=1	Clicks=2	Clicks=3	Clicks=5	Clicks=10	Clicks=20
<b>mIoU</b>	84.27 %	85.77 %	86.59 %	88.08 %	89.79 %	90.94 %
<b>mDice</b>	88.39 %	89.67 %	90.28 %	91.36 %	92.63 %	93.55 %

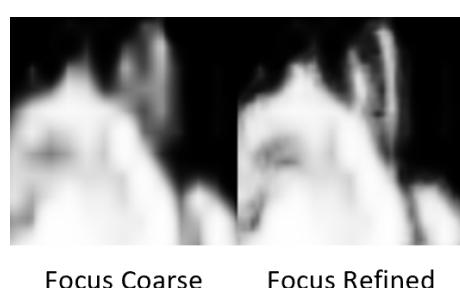
با توجه به اینکه ناحیه‌بند سبک با تعداد پارامتر کم می‌باشد، الگوریتم پس از دریافت کلیک جدید توانسته در زمان خوبی خروجی جدید را تولید کند. همچنین برای رسیدن معیار Dice به مقدار ۹۰ درصد به صورت میانگین ۳/۷۲ کلیک مورد نیاز است که تا حد خوبی قابل قبول می‌باشد. الگوریتم با دریافت ۲۰ کلیک نیز توانسته به میانگین Dice برابر با ۹۳/۵۵ درصد برسد که باز هم نشان دهنده عملکرد خوب الگوریتم می‌باشد. همچنین هرچه تعداد کلیک بیشتر می‌شود عملکرد مدل نیز بهتر می‌شود، که نشان دهنده تاثیر کلیک‌های کاربر و استفاده مناسب مدل از این کلیک‌ها می‌باشد.



شکل ۴-۳: خروجی الگوریتم در یکی از مراحل فاز آموزش آزمایش ۱



شکل ۴-۴: خروجی الگوریتم در یکی از نمونه‌های ارزیابی آزمایش ۱



شکل ۴-۵: بازیابی جزئیات تصویر یکی از نمونه‌های ارزیابی در آزمایش ۱

## ۷-۴ آزمایش ۲ : HRNet32-FLAIR-WT-Pretrained

در این آزمایش، الگوریتم FocalClick با استفاده از Backbone و ناحیه‌بند HRNet32 که از پیش بر روی مجموعه داده ImageNet آموزش داده شده است، اجرا می‌شود. همچنین از مDALIte FLAIR برای ایجاد مجموعه داده استفاده شده است. تسک و هدف مورد نظر آزمایش نیز پیدا کردن ناحیه مربوط به کل تومور یا Whole Tumor (WT) می‌باشد. پس از اینکه آموزش مدل کامل شد، ارزیابی مجموعه تست، بر روی همان م DALIte و همان تسک مرحله آموزش صورت می‌گیرد.

جدول ۴-۴: معیار NoC و NoF در آزمایش ۲

Metric	NoC80%	NoC85%	NoC90%	NoF85%	NoF90%	TPC	Time
<b>IoU</b>	4.23	5.11	6.53	515	697	0.103	1:35:37
<b>Dice</b>	3.02	3.53	4.52	326	443	0.103	1:35:37

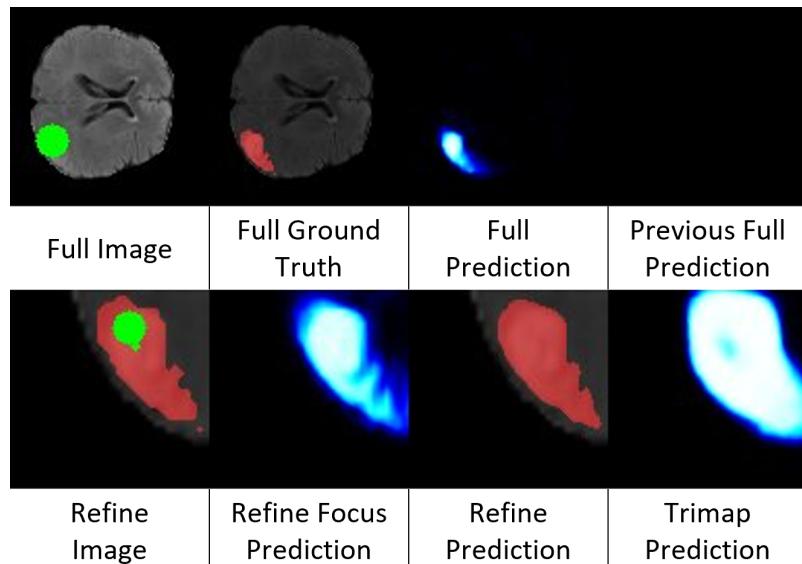
جدول ۴-۵: معیار mIoU و mDice در آزمایش ۲

Metric	Clicks=1	Clicks=2	Clicks=3	Clicks=5	Clicks=10	Clicks=20
<b>mIoU</b>	85.58 %	79.77 %	79.02 %	79.39 %	80.66 %	81.69 %
<b>mDice</b>	89.03 %	82.69 %	81.84 %	81.65 %	82.46 %	83.34 %

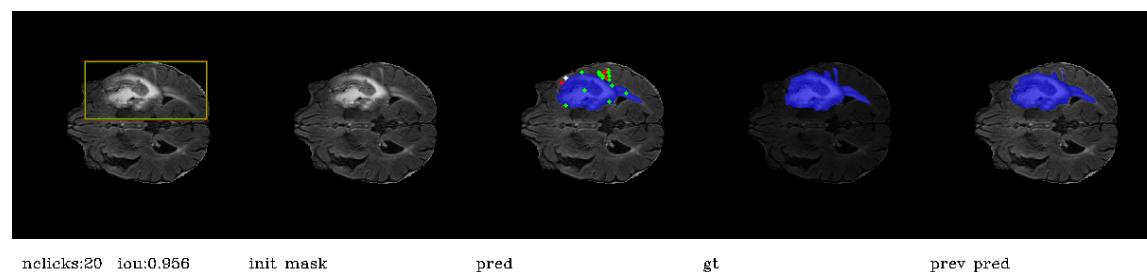
با توجه به این نکته که شبکه بزرگ با تعداد پارامتر بسیار زیاد می‌باشد و همچنین مناسب بودن این شبکه برای تصاویر با وضوح بالا، مشاهده می‌شود که زمان پاسخ الگوریتم برابر  $1/0^3$  ثانیه شده است که نسبت به آزمایش‌های دیگر افزایش دو برابر داشته است. همچنین این شبکه در تمامی معیارهای موجود نسبت به آزمایش قبلی عملکرد بدتری داشته است. با توجه به اینکه شبکه HRNet32 نسخه بزرگتر می‌باشد، می‌توان حدس زد که مدل در این آزمایش Overfit شده است.

نکته جالب دیگر این است که با افزایش تعداد کلیک‌ها، مقادیر IoU و Dice بدتر می‌شوند. این بدان معنی است که شبکه نتوانسته از تعاملات کاربر استفاده مطلوب کند و این راهنمایی‌ها حتی باعث بدتر شدن نتیجه شده است. با بررسی تصاویر خروجی مدل در هر مرحله می‌توان متوجه شد که با توجه به ابعاد مدل، ابتدا یک

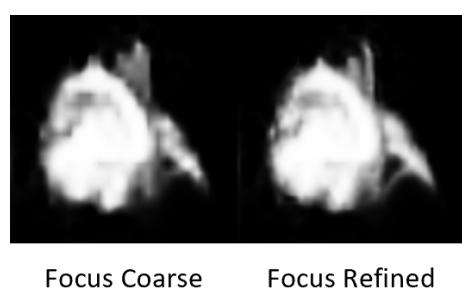
ناحیه‌بندی مناسب ایجاد می‌شود، اما با گذشت زمان پدیده Overfit بر کلیک‌های کاربر غلبه می‌کند و نتیجه با گذشت زمان و افزایش تعداد کلیک بدتر می‌شود.



شکل ۴-۶: خروجی الگوریتم در یکی از مراحل فاز آموزش آزمایش ۲



شکل ۴-۷: خروجی الگوریتم در یکی از نمونه‌های ارزیابی آزمایش ۲



شکل ۴-۸: بازیابی جزئیات تصویر یکی از نمونه‌های ارزیابی در آزمایش ۲

## SegFormerB3-FLAIR-WT-Pretrained : ۴-۸ آزمایش ۳

در این آزمایش، الگوریتم FocalClick با استفاده از Backbone و ناحیه‌بند SegFormerB3 که از پیش بر روی مجموعه داده ImageNet آموزش داده شده است، اجرا می‌شود. همچنین از مدلیته FLAIR برای ایجاد مجموعه داده استفاده شده است. تسک و هدف مورد نظر آزمایش نیز پیدا کردن ناحیه مربوط به کل تومور یا Whole Tumor (WT) می‌باشد. پس از اینکه آموزش مدل کامل شد، ارزیابی مجموعه تست، بر روی همان مدلیته و همان تسک مرحله آموزش صورت می‌گیرد.

جدول ۴-۶: معیار NoC و NoF در آزمایش ۳

Metric	NoC80%	NoC85%	NoC90%	NoF85%	NoF90%	TPC	Time
<b>IoU</b>	3.28	4.03	5.75	320	530	0.059	0:55:07
<b>Dice</b>	2.45	2.81	3.48	176	246	0.059	0:55:07

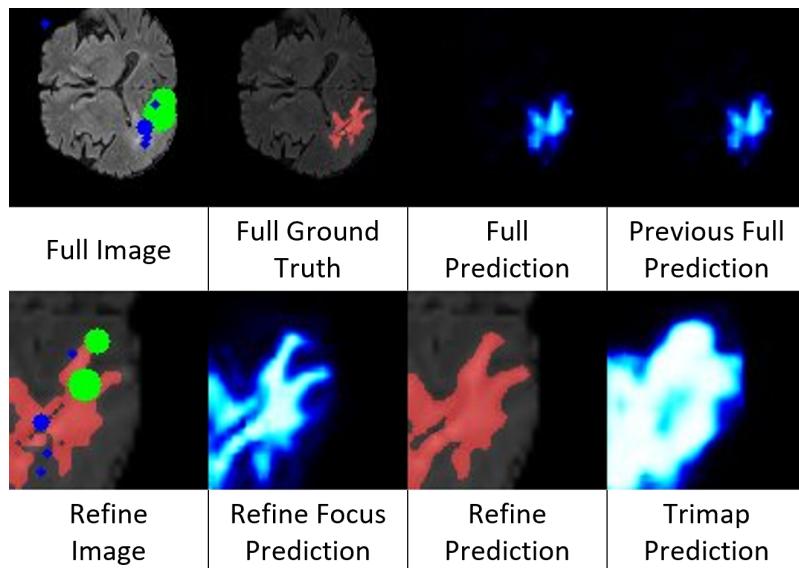
جدول ۴-۷: معیار mIoU و mDice در آزمایش ۳

Metric	Clicks=1	Clicks=2	Clicks=3	Clicks=5	Clicks=10	Clicks=20
<b>mIoU</b>	85.35 %	85.30 %	86.46 %	88.37 %	90.29 %	91.75 %
<b>mDice</b>	88.96 %	88.84 %	89.75 %	91.16 %	92.78 %	94.12 %

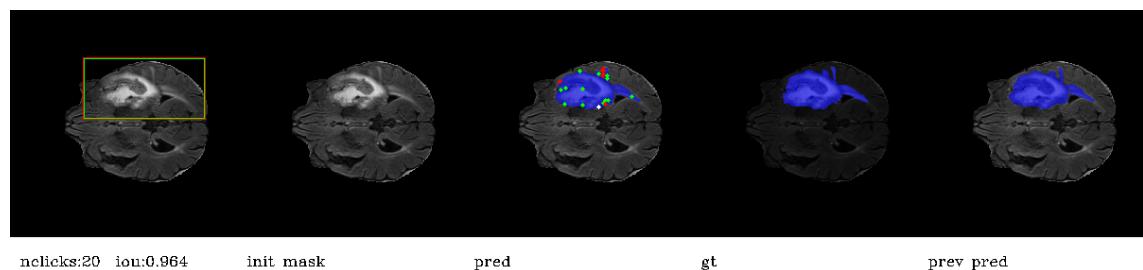
با توجه به اینکه ناحیه‌بند SegFormer یک ناحیه‌بند سبک‌تر با تعداد پارامتر کم‌تر نسبت به HRNet32 می‌باشد، الگوریتم پس از دریافت کلیک جدید توانسته در زمان خوبی (تقريباً مشابه آزمایش ۱ اما کمی بهتر) خروجی جدید را تولید کند همچنین برای رسیدن معیار Dice به مقدار ۹۰ درصد به صورت ميانگين ۳/۴۸ کلیک مورد نياز است که نسبت به دو آزمایش قبلی بهبود خوبی داشته است. الگوریتم با دریافت ۲۰ کلیک نيز توانسته به ميانگين Dice برابر با ۹۴/۱۲ درصد برسد که باز هم نشان دهنده عملکرد خوب الگوریتم نسبت به دو آزمایش قبلی می‌باشد. همچنین هرچه تعداد کلیک بيشتر می‌شود عملکرد مدل نيز بهتر می‌شود، که نشان دهنده تاثير کلیک‌های كاري و استفاده مناسب مدل از اين کلیک‌ها می‌باشد. به صورت خلاصه می‌توان گفت استفاده از ناحیه‌بند SegFormerB3 باعث بهبود عملکرد مدل در تمامی جنبه‌ها شده است.

#### فصل ۴. آزمایش‌ها و نتیجه‌ها

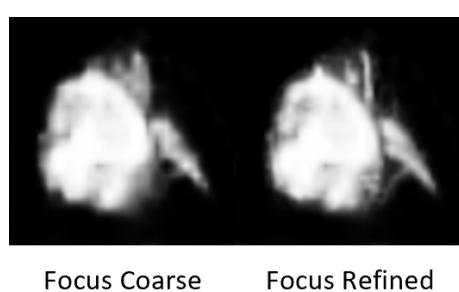
SEGFORMERB3-FLAIR-WT-PRETRAINED : آزمایش ۳-۸



شکل ۹-۴: خروجی الگوریتم در یکی از مراحل فاز آموزش آزمایش ۳



شکل ۱۰-۴: خروجی الگوریتم در یکی از نمونه‌های ارزیابی آزمایش ۳



شکل ۱۱-۴: بازیابی جزئیات تصویر یکی از نمونه‌های ارزیابی در آزمایش ۳

## ۹-۴ آزمایش ۴: SegFormerB3-Mix-WT و ترکیب مدل‌الیته‌ها

در این آزمایش، الگوریتم FocalClick با استفاده از Backbone و ناحیه‌بند SegFormerB3 و بدون اینکه از FLAIR, T1ce, T2 برای ایجاد مجموعه داده استفاده شده است. همچنین از حالت Mix مدل‌الیته‌ها یعنی ترکیب T2 مدل‌الیته‌ها در کنار هم چه تاثیری بر عملکرد مدل دارد. هدف در این آزمایش این است که بینیم استفاده همزمان از کل تومور یا (WT) Whole Tumor می‌باشد. پس از اینکه آموزش مدل کامل شد، ارزیابی مجموعه تست، بر روی همان مدل‌الیته‌ها و همان تسک مرحله آموزش صورت می‌گیرد.

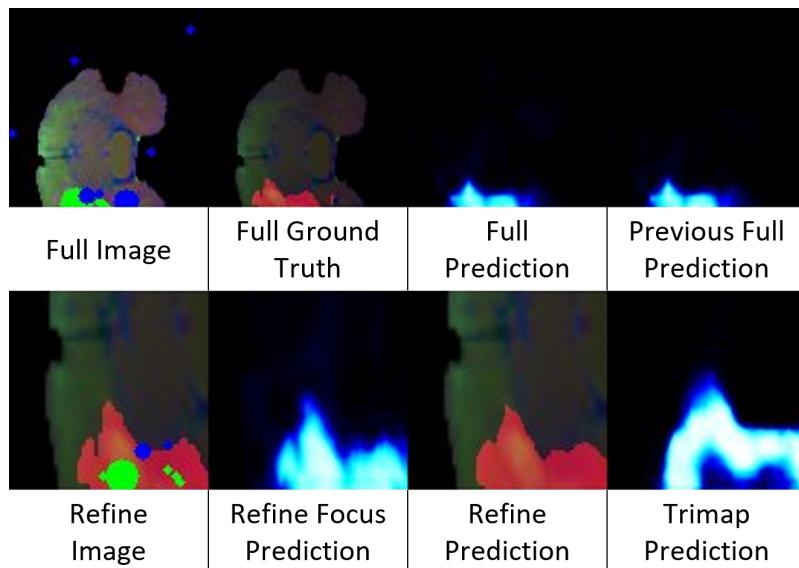
جدول ۸-۴: معیار NoC و NoF در آزمایش ۴

Metric	NoC80%	NoC85%	NoC90%	NoF85%	NoF90%	TPC	Time
<b>IoU</b>	3.24	4.15	5.63	319	533	0.061	0:56:54
<b>Dice</b>	2.21	2.66	3.53	138	233	0.061	0:56:54

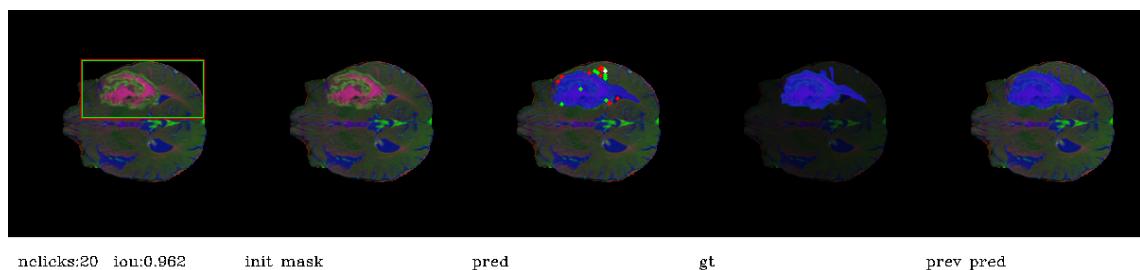
جدول ۹-۴: معیار mIoU و mDice در آزمایش ۴

Metric	Clicks=1	Clicks=2	Clicks=3	Clicks=5	Clicks=10	Clicks=20
<b>mIoU</b>	86.17 %	87.31 %	88.36 %	89.73 %	91.74 %	93.07 %
<b>mDice</b>	89.57 %	90.86 %	91.74 %	92.82 %	94.49 %	95.56 %

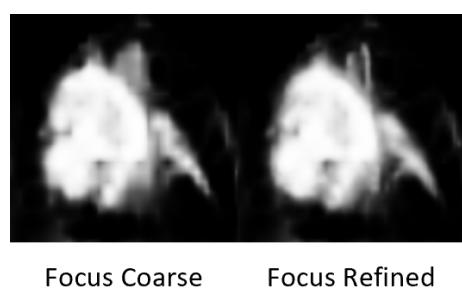
با توجه به این که در این آزمایش مشابه با آزمایش ۳ از ناحیه‌بند SegFormerB3 استفاده شده است و در تعداد پارامترهای مدل تفاوتی ایجاد نشده است، زمان پاسخ و اجرای مدل مشابه آزمایش ۳ بوده است که زمان بسیار مناسبی است. همچنین از مقایسه نتایج مربوط به NoC و mDice این آزمایش با آزمایش قبلی متوجه می‌شویم که مدل در هر دو معیار به صورت کلی بهبود یافته است. این امر ثابت کننده این است که استفاده همزمان و ترکیبی از مدل‌الیته‌ها اثربخش می‌باشد. برای مثال در آزمایش ۳ برای رسیدن معیار IoU به مقدار ۹۰ درصد مدل به صورت میانگین به ۵/۷۵ کلیک نیاز داشته، در حالی که این مقدار در این آزمایش برابر ۵/۶۳ می‌باشد. همچنین در آزمایش ۳ با دریافت ۲۰ کلیک مقدار mIoU ۹۱/۷۵ درصد شده است. در حالی که در این آزمایش این مقدار بهبود یافته است و برابر ۹۳/۰۷ درصد می‌باشد.



شکل ۱۲-۴: خروجی الگوریتم در یکی از مراحل فاز آموزش آزمایش ۴



شکل ۱۳-۴: خروجی الگوریتم در یکی از نمونه‌های ارزیابی آزمایش ۴



شکل ۱۴-۴: بازیابی جزئیات تصویر یکی از نمونه‌های ارزیابی در آزمایش ۴

## ۴-۱۰ آزمایش ۵: ارزیابی با تسک جدید دیده نشده در آموزش

در این آزمایش قصد داریم عملکرد مدل را بر روی تسکی که در مرحله آموزش با آن مواجه نشده<sup>۲</sup> است بررسی کنیم. در این راستا از وزن‌های مدل آموزش داده شده در آزمایش ۳ برای ارزیابی تسکی جدید استفاده می‌کنیم. در آزمایش ۳ تسک تعریف شده پیدا کردن ناحیه مربوط به کل تومور (WT) بود و مدل بر اساس آن تسک آموزش داده شده بود. در حالی که در این آزمایش تسک مورد نظر پیدا کردن ناحیه مربوط به هسته تومور (TC) می‌باشد. با تحلیل و تفسیر نتایج این آزمایش می‌توان متوجه شد که کلیک‌های کاربر با وجود تغییر تسک، چقدر برای مدل مفید بوده است تا مدل را در راستای تسک جدید راهنمایی کند.

جدول ۴-۵: معیار NoF و NoC در آزمایش ۵

Metric	NoC80%	NoC85%	NoC90%	NoF85%	NoF90%	TPC	Time
IoU	5.96	6.39	6.82	667	790	0.059	0:54:53
Dice	4.93	5.49	6.13	456	596	0.059	0:54:53

جدول ۴-۶: معیار mDice و mIoU در آزمایش ۵

Metric	Clicks=1	Clicks=2	Clicks=3	Clicks=5	Clicks=10	Clicks=20
mIoU	78.04 %	78.29 %	78.96 %	80.30 %	82.85 %	86.92 %
mDice	81.76 %	82.12 %	82.71 %	83.83 %	86.09 %	89.75 %

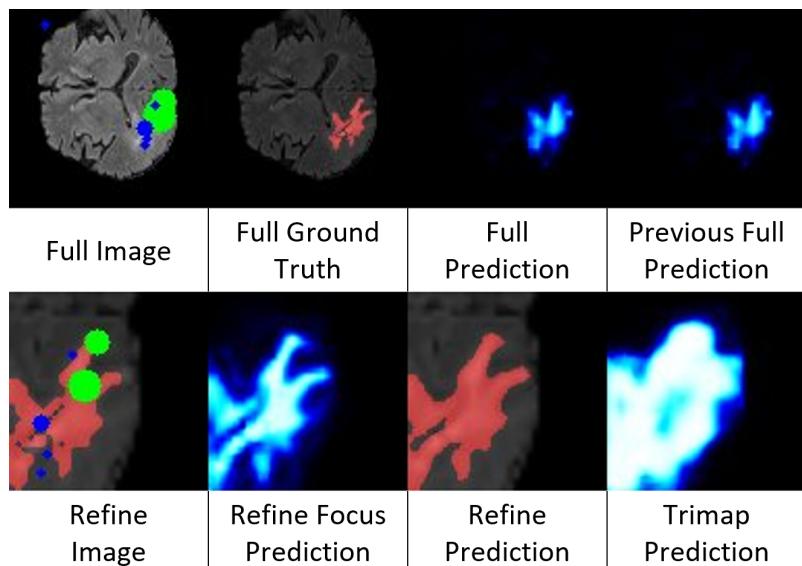
ابتدا باید گفت ناحیه مربوط به TC اندازه کوچکتری نسبت به ناحیه WT دارد و همین موضوع ناحیه‌بندی قسمت TC را دشوارتر می‌کند. بنابراین کمتر بودن مقادیر IoU و Dice نسبت به آزمایش‌های قبلی امری طبیعی است. در واقع حتی اگر مدل روی تسک TC هم آموزش ببیند باز هم مقادیر IoU و Dice کمتری خواهد داشت.

با مشاهده نتایج به دست آمده متوجه می‌شویم در ابتدا که مدل تنها یک کلیک دریافت کرده است، دارای عملکرد بسیار بدی بوده است و مقدار mDice آن برابر ۸۱/۷۶ درصد بوده است. در واقع این اتفاق مربوط به تغییر تسک مدل می‌باشد. اما با افزایش تعداد کلیک متوجه می‌شویم مدل توانسته متوجه تسک جدید

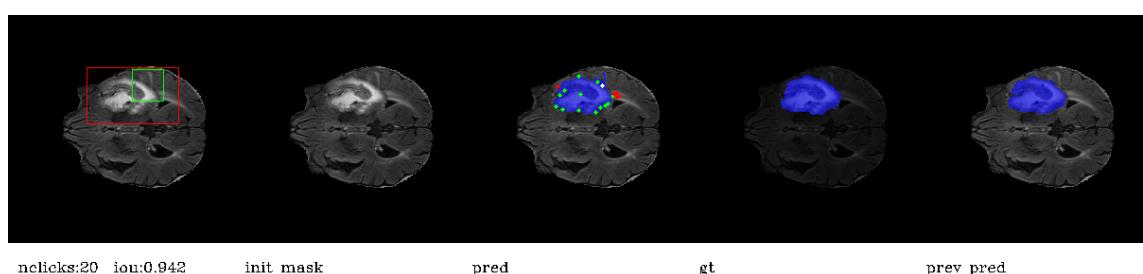
<sup>2</sup>Unseen Task

#### فصل ۴. آزمایش‌ها و نتیجه‌ها

شود و مطابق آن عمل کند. به طوری که با رسیدن تعداد کلیک به عدد ۲۰ مقدار mDice به  $89/75$  درصد رسیده است که برای تسک TC مقدار مناسبی است. همچنین در این آزمایش تاثیر تعداد کلیک‌های کاربر بر مقادیر U و mDice نسبت به آزمایش‌های قبلی مشهودتر می‌باشد. از تصاویر خروجی تولید شده در مراحل مختلف الگوریتم نیز می‌توان متوجه این نکته شد که هنگامی که تعداد کلیک کاربر کم است، مدل ناحیه مربوط به WT را خروجی می‌دهد و با افزایش تعداد کلیک متوجه می‌شود که یک ناحیه دیگر یعنی TC مورد نظر کاربر بوده است.

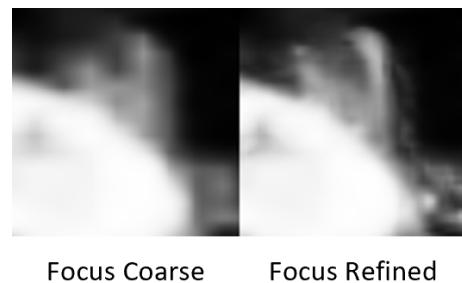


شکل ۴-۱۵: خروجی الگوریتم در یکی از مراحل فاز آموزش آزمایش ۵ یا همان ۳



شکل ۴-۱۶: خروجی الگوریتم در یکی از نمونه‌های ارزیابی آزمایش ۵

#### فصل ۴. آزمایش‌ها و نتیجه‌ها ۱۱-۴. آزمایش ۶: ارزیابی با مدل‌الیه و تسک جدید دیده نشده در آموزش



شکل ۱۷-۴: بازیابی جزئیات تصویر یکی از نمونه‌های ارزیابی در آزمایش ۵

#### ۱۱-۴ آزمایش ۶: ارزیابی با مدل‌الیه و تسک جدید دیده نشده در آموزش

در این آزمایش قصد داریم به صورت همزمان عملکرد مدل را بر روی مدل‌الیه و تسکی که در مرحله آموزش با آن مواجه نشده<sup>۳</sup> است بررسی کنیم. در این راستا از وزن‌های مدل آموزش داده شده در آزمایش ۳ برای ارزیابی مدل‌الیه و تسکی جدید استفاده می‌کنیم. در آزمایش ۳ تسک تعریف شده پیدا کردن ناحیه مربوط به کل تومور (WT) از روی مدل‌الیه FLAIR بود و مدل بر اساس آن مدل‌الیه و تسک آموزش داده شده بود. در حالی که در این آزمایش تسک مورد نظر پیدا کردن ناحیه مربوط به هسته تومور (TC) با استفاده از مدل‌الیه T1ce می‌باشد. با تحلیل و تفسیر نتایج این آزمایش می‌توان متوجه شد که کلیک‌های کاربر با وجود تغییر مدل‌الیه و تسک، چقدر برای مدل مفید بوده است تا مدل را در راستای مدل‌الیه و تسک جدید راهنمایی کند.

جدول ۱۲-۴: معیار NoC و NoF در آزمایش ۶

Metric	NoC80%	NoC85%	NoC90%	NoF85%	NoF90%	TPC	Time
<b>IoU</b>	4.56	4.87	5.42	480	576	0.058	0:54:22
<b>Dice</b>	4.20	4.38	4.68	421	458	0.058	0:54:22

<sup>3</sup>Unseen Modality and Task

#### فصل ۴. آزمایش‌ها و نتیجه‌ها ۱۱-۴. آزمایش ۶: ارزیابی با مдалیته و تسک جدید دیده نشده در آموزش

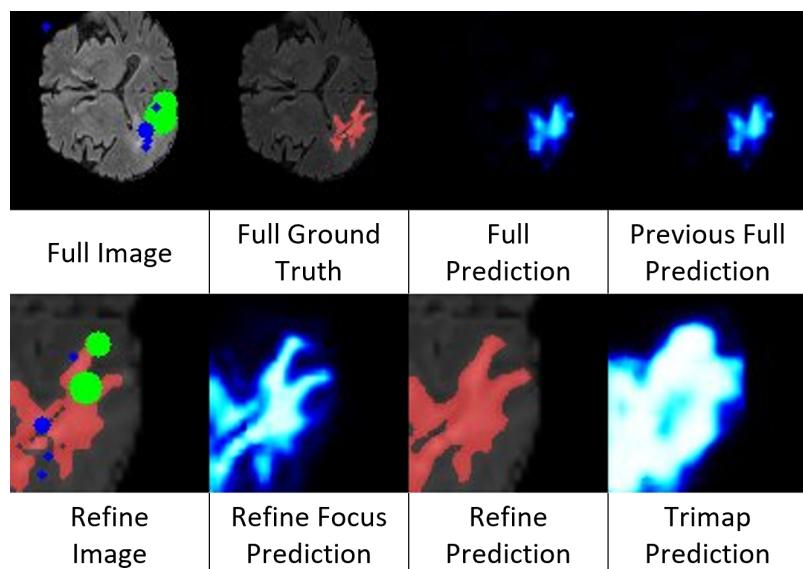
جدول ۱۳-۴: معیار mIoU و mDice در آزمایش ۶

Metric	Clicks=1	Clicks=2	Clicks=3	Clicks=5	Clicks=10	Clicks=20
<b>mIoU</b>	79.80 %	79.27 %	80.06 %	81.31 %	82.96 %	85.49 %
<b>mDice</b>	81.52 %	80.77 %	81.47 %	82.58 %	84.21 %	86.77 %

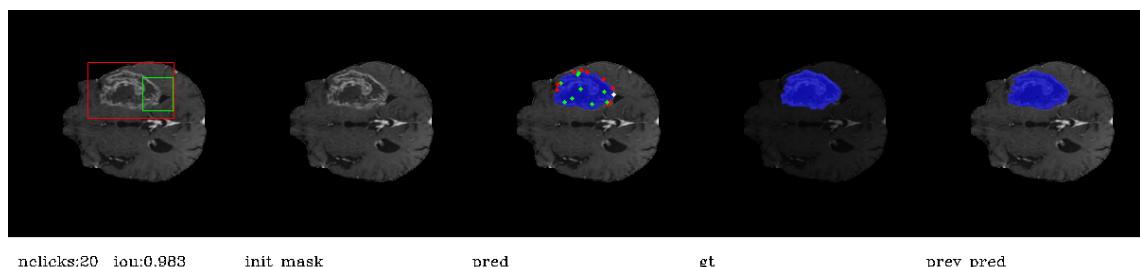
ابتدا باید گفت ناحیه مربوط به TC اندازه کوچکتری نسبت به ناحیه WT دارد و همین موضوع ناحیه‌بندی قسمت TC را دشوارتر می‌کند. بنابراین کمتر بودن مقادیر IoU و Dice نسبت به آزمایش‌های قبلی امری طبیعی است. در واقع حتی اگر مدل روی تسک TC هم آموزش ببیند باز هم مقادیر IoU و Dice کمتری خواهد داشت. علاوه بر مورد بالا، همانطور که در فصل ۲ اشاره شد مдалیته FLAIR ناحیه مربوط به Edema را برجسته‌تر می‌کند. ناحیه مربوط به کل تومور شامل ناحیه Edema می‌باشد اما ناحیه مربوط به هسته تومور شامل آن نمی‌باشد. مдалیته T1ce نیز بخش فزاینده تومور را برجسته‌تر می‌کند که این ناحیه در هر دو ماسک و وجود دارد. همه این موارد باعث می‌شود مدل با یک تسک چالش برانگیز و سخت مواجه شود. با مشاهده نتایج به دست آمده متوجه می‌شویم در ابتدا که مدل تنها یک کلیک دریافت کرده دارای عملکرد بسیار بدی بوده است و مقدار mDice آن برابر ۸۱/۵۲ درصد بوده است. در واقع این اتفاق مربوط به تغییر مдалیته و تسک مدل می‌باشد. اما با افزایش تعداد کلیک متوجه می‌شویم مدل توانسته متوجه مдалیته و تسک جدید شود و مطابق آن عمل کند. به طوری که با رسیدن تعداد کلیک به عدد ۲۰ مقدار mDice به ۸۶/۷۷ درصد رسیده است. این مقدار نسبت به آزمایش ۵ که مдалیته تغییر نکرده بود کمتر می‌باشد و تغییر مдалیته اثر مخرب خود را داشته است. با این حال در نهایت مدل به ازای تعداد کلیک زیاد به نتیجه بدی رسیده است. البته یکی از اهداف این پژوهه کم کردن کاربر نیاز از سمت کاربر می‌باشد که نتیجه این آزمایش در تضاد با آن هدف می‌باشد و با تغییرات مдалیته و تسک، کاربر برای رسیدن به نتیجه مطلوب خود باید تلاش زیادی کند.

همچنین در این آزمایش تاثیر تعداد کلیک‌های کاربر بر مقادیر mIoU و mDice نسبت به آزمایش‌های قبلی مشهودتر می‌باشد. از تصاویر خروجی تولید شده در مراحل مختلف الگوریتم نیز می‌توان متوجه این نکته شد که هنگامی که تعداد کلیک کاربر کم است، مدل ناحیه مربوط به WT را خروجی می‌دهد و با افزایش تعداد کلیک متوجه می‌شود که یک ناحیه دیگر یعنی TC مورد نظر کاربر بوده است.

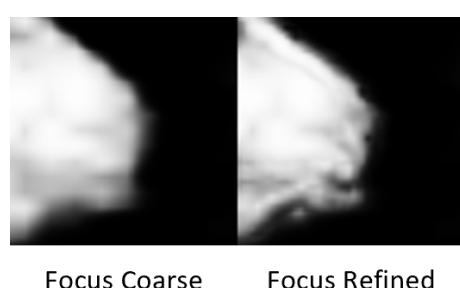
#### فصل ۴. آزمایش‌ها و نتیجه‌ها ۱۱-۴. آزمایش ۶: ارزیابی با مدل‌الیه و تسک جدید دیده نشده در آموزش



شکل ۱۸-۴: خروجی الگوریتم در یکی از مراحل فاز آموزش آزمایش ۶ یا همان ۳



شکل ۱۹-۴: خروجی الگوریتم در یکی از نمونه‌های ارزیابی آزمایش ۶



شکل ۲۰-۴: بازیابی جزئیات تصویر یکی از نمونه‌های ارزیابی در آزمایش ۶

## ۱۲-۴ مقایسه با نتایج کارهای مرتبط

وجود ۴ نکته باعث شده امکان مقایسه قطعی و دقیق میان کارهای مختلف بر روی مجموعه داده BraTS با مشکل همراه شود. این چهار نکته عبارت اند از:

۱. وجود نسخه‌های سالانه از مجموعه داده BraTS (از سال ۲۰۱۲ تا ۲۰۲۲) و استفاده مقالات از نسخه‌های متفاوت

۲. رفتار متفاوت مقالات در توزیع مجموعه داده BraTS به مجموعه‌های آموزشی، اعتبارسنجی و تست

۳. وجود چهار مدلیتی برای هر تصویر و استفاده مقالات از مدلیت‌های مختلف

۴. وجود چندین ماسک مختلف و استفاده مقالات از ماسک‌های مختلف این مجموعه داده

با وجود این مشکلات همچنان می‌توان یک مقایسه تقریبی میان نتیجه روش پیشنهادی ما و سایر روش‌ها را ارائه کرد. در این قسمت نتایج روش پیشنهادی ما با نتایج سه مقاله Enhanced U-Net [۲۷]، DeepIGeoS [۲۸] و MIDeepSeg [۲۹] مقایسه می‌شود. در آخر هم یک مقایسه اجمالی با نتایج مقالات قدیمی‌تر و غیر عمیق صورت می‌گیرد.

### ۱۲-۴-۱ مقاله Enhanced U-Net

مقاله Enhanced U-Net [۲۷] یک روش غیر تعاملی ناحیه‌بندی را ارائه می‌کند. در این روش از شبکه U-Net برای ناحیه‌بندی استفاده شده است. همچنین از نسخه‌های سالانه ۲۰۱۷ تا ۲۰۲۰ این مجموعه داده به صورت مجزا استفاده شده است. این مقاله نسخه ۲۰۱۹ این مجموعه داده را به صورت ۶۸ درصد آموزش، ۲۰ درصد اعتبارسنجی و ۱۲ درصد تست تفکیک کرده است. همچنین در این مقاله از مدلیت‌های FLAIR و T1ce استفاده شده است. همانطور که در جدول ۱۴-۴ مشاهده می‌شود این مقاله به صورت تقریبی به Dice برابر با ۸۴ درصد رسیده است که نسبت به روش پیشنهادی و تعاملی ما دارای مقدار کمتری می‌باشد. اختلاف زیاد نتیجه این مقاله با روش ما به دلیل غیر تعاملی بودن این مقاله می‌باشد و از این نکته می‌توان به اهمیت و تاثیر تعامل کاربر پی برد.

جدول ۱۴-۴: نتایج مقاله [۲۷] Enhanced U-Net

Dataset	Loss	Accuracy	Mean IoU	Precision	Sensitivity	Specificity	Dice Score
BraTS 2017	0.0056	0.9980	0.9637	0.9973	0.9970	0.9972	0.8453
BraTS 2018	0.0057	0.9979	0.8927	0.9972	0.9970	0.9940	0.8160
BraTS 2019	<b>0.0054</b>	<b>0.9981</b>	<b>0.9130</b>	<b>0.9974</b>	<b>0.9971</b>	<b>0.9991</b>	<b>0.8409</b>
BraTS 2020	0.0056	0.9980	0.8935	0.9973	0.9970	0.9983	0.8300

## ۲-۱۲-۴ مقاله DeepIGeoS

مقاله [۴۰] DeepIGeoS یک روش تعاملی ناحیه‌بندی را ارائه می‌کند. در این روش از فاصله ژئودزیک<sup>۴</sup> برای تبدیل کلیک‌های کاربر به نقشه‌های تعامل استفاده شده است. همچنین در این روش از شبکه‌های P-Net و CRF-Net [۲۲] و R-Net [۳۸] به صورت همزمان برای ناحیه‌بندی استفاده شده است.

در این مقاله از نسخه ۲۰۱۵ مجموعه داده BraTS استفاده شده است. این مقاله نسخه ۲۰۱۵ این مجموعه داده را به صورت ۸۵ درصد آموزش و ۱۵ درصد اعتبارسنجی و تست تفکیک کرده است. همچنین در این مقاله فقط از مدلیته FLAIR استفاده شده است. تسک تعریف شده نیز بر روی ماسک WT می‌باشد. همانطور که در جدول ۱۵-۴ مشاهده می‌شود این مقاله در بهترین حالت خود به Dice برابر با ۸۹/۹۳ درصد رسیده است که نسبت به روش پیشنهادی ما دارای مقدار کمتری می‌باشد.

<sup>4</sup>Geodesic

جدول ۴-۱۵: نتایج مقاله [۴۰] DeepIGeoS

Method	Dice (%)	ASSD (pixels)
Before refinement	$87.55 \pm 6.72$	$2.04 \pm 1.70$
Min-cut user-editing	$88.41 \pm 7.05$	$1.74 \pm 1.53$
3D R-Net (Euc)	$88.82 \pm 7.68$	$1.60 \pm 1.56$
3D R-Net	$89.30 \pm 6.82$	$1.52 \pm 1.37$
3D R-Net (Euc) + CRF-Net (fu)	$89.27 \pm 7.32$	$1.48 \pm 1.22$
3D R-Net + CRF-Net (fu)	<b><math>89.93 \pm 6.49</math></b>	<b><math>1.43 \pm 1.16</math></b>

## ۳-۱۲-۴ مقاله MIDeepSeg

مقاله MIDeepSeg [۲۳] یک روش تعاملی ناحیه‌بندی را ارائه می‌کند. در این روش از فاصله ژئودزیک نمایی برای تبدیل کلیک‌های کاربر به نقشه‌های تعامل استفاده شده است. همچنین در این روش از شبکه U-Net [۳۳] برای ناحیه‌بندی اولیه استفاده شده است. برای اعمال کلیک‌های کاربر بر روی ناحیه‌بندی اولیه تولید شده نیز از روش غیر عمیق Graph Cut استفاده شده است.

در این مقاله از نسخه ۲۰ مجموعه داده BraTS استفاده شده است. این مقاله نسخه ۱۸ این مجموعه داده را به صورت ۷۸ درصد آموزش و ۲۲ درصد اعتبارسنجی و تست تفکیک کرده است. همچنین در این مقاله از مدل‌الیته‌های FLAIR و T1ce به صورت مجزا استفاده شده است. تسك تعریف شده نیز بر روی ماسک TC و WT می‌باشد.

جدول ۴-۱۶: توزیع مجموعه داده در مقاله [۲۳] MIDeepSeg

Object	Modality	$N_{\text{train}}$	$N_{\text{test}}$	Dataset
Tumor core	MRI (T1ce)	170	47	BraTS2018
Whole tumor	MRI (FLAIR)	0	60	BraTS2018

این مقاله نیز مشابه روش پیشنهادی ما عملکرد خود را بر روی مدل‌الیته و تسك دیده نشده می‌سنجد.

<sup>۵</sup>Exponentialized Geodesic Distance (EGD)

همانطور که در جدول ۱۷-۴ مشاهده می‌شود این مقاله در بهترین حالت خود به Dice برابر با ۸۸/۷۱ درصد رسیده است که نسبت به روش پیشنهادی ما دارای مقدار کمتری می‌باشد.

[۲۳] جدول ۱۷-۴: نتایج مقاله MIDeepSeg

Method	Dice (%)	ASSD (mm)	Time (s)
3D Graph Cuts	78.91 $\pm$ 14.98	3.46 $\pm$ 5.10	99.4 $\pm$ 36.7
ITK-SNAP	82.34 $\pm$ 11.42	1.99 $\pm$ 1.31	173.0 $\pm$ 75.5
DeepIGeoS	82.47 $\pm$ 10.78	3.34 $\pm$ 3.81	82.2 $\pm$ 44.7
DeepGrabCut	82.58 $\pm$ 11.79	2.89 $\pm$ 2.37	<b>10.5 <math>\pm</math> 8.3</b>
DIOS	83.34 $\pm$ 10.47	2.57 $\pm$ 1.79	67.5 $\pm$ 23.6
DEXTR	86.39 $\pm$ 9.03	1.59 $\pm$ 1.11	34.7 $\pm$ 18.6
MIDeepSeg	<b>88.71 <math>\pm</math> 7.00</b>	<b>1.24 <math>\pm</math> 0.88</b>	28.6 $\pm$ 12.2

#### ۱۲-۴-۴ روشهای سنتی و غیر عمیق

در این قسمت مقدار Dice برای روشهای سنتی و غیر عمیق آورده شده است. با توجه به تعداد زیاد این روشهای جزئیات آنها ذکر نمی‌شود. نکته قابل توجه اختلاف زیاد مقدار Dice در این روشهای نسبت به روشهای عمیق می‌باشد. طوری که برای تسک WT مقدار Dice بین ۴۴ تا ۸۲ درصد می‌باشد. همچنین برای تسک TC نیز این معیار بین ۱۷ تا ۷۰ درصد می‌باشد.

جدول ۴-۱۸: نتایج روش‌های قدیمی و غیر عمیق [۲۶]

	whole	core	active	time (min) (arch)
Dice (in %)	<i>LG/HG</i>	<i>LG/HG</i>		
Bauer	68 49/74	48 30/54	57	8 (CPU)
Buendia	57 19/71	42 8/54	45	0.3 (CPU)
Cordier	68 60/71	51 41/55	39	20 (Cluster)
Doyle	74 63/78	44 41/45	42	15 (CPU)
Festa	62 24/77	50 33/56	61	30 (CPU)
Geremia	62 55/65	32 34/31	42	10 (Cluster)
Guo	74 71/75	65 59/67	49	< 1 (CPU)
Hamamci	72 55/78	57 40/63	59	20 (CPU)
Meier	69 46/77	50 36/55	57	6 (CPU)
Menze (D)	78 81/76	58 58/59	54	20 (CPU)
Menze (G)	69 48/77	33 9/42	53	10 (CPU)
Reza	70 52/77	47 39/50	55	90 (CPU)
Riklin Raviv	74 <i>na</i> /74	50 <i>na</i> /50	58	8 (CPU)
Shin	30 28/31	17 22/15	5	8 (CPU)
Subbanna	75 55/82	70 54/75	59	70 (CPU)
Taylor	44 24/51	28 11/34	41	1 (Cluster)
Tustison	75 68/78	55 42/60	52	100 (Cluster)
Zhao (I)	82 78/84	66 60/68	49	15 (CPU)
Zhao (II)	76 67/79	51 42/55	52	20 (CPU)
Zikic	75 62/80	47 33/52	56	2 (CPU)
Best Combination	88 86/89	78 66/82	71	
Fused	82 68/87	73 62/77	65	

## فصل ۵

### جمع‌بندی و پیشنهادها

#### ۱-۵ جمع‌بندی

در این پژوهه روش و الگوریتم FocalClick [۵] به عنوان یک روش ناحیه‌بندی تعاملی عمیق مبتنی بر کلیک استفاده شد تا تصاویر پزشکی MRI از مغز را ناحیه‌بندی و نواحی دارای تومور را مشخص کند. آزمایش‌های صورت گرفته نشان دادند که الگوریتم ارائه شده عملکرد بسیار بهتری نسبت به الگوریتم‌های غیر تعاملی و تعاملی قبلی دارد. این آزمایش‌ها بر روی مجموعه داده BraTS سال ۲۰۲۰ صورت گرفت و در تمامی آزمایش‌ها نتیجه به دست آمده بهتر از روش‌های پیشین بود. این عملکرد بهتر به دلیل وجود کلیک‌های کمکی کاربر بود که با ارائه روشی برای بهره برداری مناسب از این کلیک‌ها، ثابت شد که وجود تعامل کاربر تا چه میزان کمک کننده است. همچنین عملکرد مناسب الگوریتم FocalClick در معیارهایی نظیر TPC و NoC نشان دهنده این است که این الگوریتم بر خلاف روش‌های پیشین کاملاً مناسب استفاده در صنعت و دنیای واقعی می‌باشد. الگوریتم FocalClick به طور قابل توجهی کارایی Pipeline‌های موجود را بهبود می‌بخشد و امکان استقرار آنها در دستگاه‌های کم مصرف را فراهم می‌کند. علاوه بر موارد بالا، نتایج آزمایش‌ها نشان داد که الگوریتم مورد نظر بر روی تسك‌های متفاوت به خوبی تعمیم<sup>۱</sup> می‌باشد. این تعمیم به گونه‌ای است که مدل با آموخته بر روی یک تسك، می‌تواند نتیجه مطلوبی در تسك دیگر دیده نشده داشته باشد. در ادامه به کارهایی که می‌تواند قدم‌های بعدی پژوهه باشد، اشاره می‌شود.

<sup>۱</sup>Generalization

## ۵-۲ پیشنهادها و کارهای آینده

### ۱-۲-۵ مناسب‌سازی FocalClick برای ورودی سه‌بعدی

در این پژوهه به دلیل وجود نداشتن مازول‌های مورفولوژی سه‌بعدی و همچنین تعریف نشدن قسمت‌های مختلف الگوریتم FocalClick برای حالت سه‌بعدی و چالش برانگیز بودن طراحی سه‌بعدی این قسمت‌ها، مجموعه داده سه‌بعدی BraTS با انجام عمل Slicing بر روی Depth تبدیل به حالت دو‌بعدی شد. این کار می‌تواند باعث از دست رفتن اطلاعات مکانی<sup>۲</sup> و در نهایت باعث کاهش عملکرد شود. البته به دلیل محدودیت‌های حافظه‌ای و بزرگ بودن هر نمونه سه‌بعدی، امکان ورودی دادن این نمونه‌ها به صورت کامل به مدل وجود ندارد. برای حل این مشکل می‌توان به جای عمل Crop که در FocalClick از آن استفاده می‌شود، از ترکیب دو عمل Crop و Slice استفاده کرد. در این صورت به جای اینکه هر نمونه مجموعه داده که دارای ابعاد  $155 \times 240 \times 240$  می‌باشد به ابعاد  $96 \times 96 \times 96$  تبدیل شود و بعد سوم خود را کاملاً از دست دهد، می‌توان آن نمونه را به ابعاد  $48 \times 96 \times 96$  تبدیل کرد تا اطلاعات مکانی از بعد سوم نیز وجود داشته باشد. این عمل در مقالات DeepIGeoS [۴۰] و MIDeepSeg [۲۳] انجام شده است و نتیجه‌بخش بوده است.

### ۲-۲-۵ استفاده از FocalClick در BRS

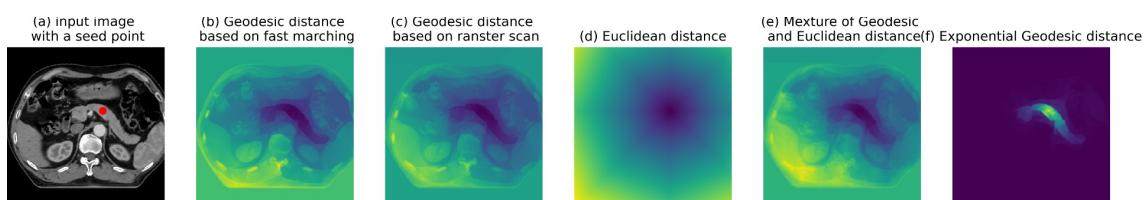
در مقاله BRS [۱۸] تطبیق‌پذیری و تاثیر مثبت روش BRS بر روی چندین شبکه CNN نشان داده می‌شود و ثابت می‌شود که استفاده از BRS باعث کاهش معیار NoC می‌شود. کم کردن تعداد کلیک مورد نیاز و همچنین کاهش تلاش کاربر، یکی از مهم‌ترین اهداف تعریف شده برای طراحی مدل‌های ناحیه‌بندی تعاملی می‌باشد. یکی از کارهای آینده در این پژوهش می‌تواند به کار گیری روش BRS در الگوریتم FocalClick باشد. البته این کار نیازمند تغییرات زیادی در الگوریتم موجود می‌باشد.

### ۲-۳-۵ استفاده از تبدیل فاصله ژئودزیک نمایی

در الگوریتم استفاده شده در این پژوهه، کلیک‌های کاربر به نقشه‌های Binary تبدیل می‌شوند. به عبارت دیگر هر کلیک کاربر به یک دایره با شعاع ۲ تبدیل می‌شود. این تبدیل جزو ساده‌ترین تبدیل‌های موجود می‌باشد و

<sup>2</sup>Spatial

حتی نسبت به سایر روش‌های معمول دیگر مانند فاصله اقلیدسی یا گاوی اطلاعات کمتری را برای مدل فراهم می‌کند. در مقاله DeepIGeoS از فاصله ژئودزیک برای تبدیل کلیک‌های کاربر به نقشه‌های تعامل استفاده شده است. در ادامه مقاله MIDeepSeg از حالت بهبود یافته این فاصله یعنی فاصله ژئودزیک نمایی (EGD) استفاده کرده است. در تصویر ۱-۵ می‌توان نقشه تعامل ساخته شده در روش‌های مختلف را مشاهده کرد. همانطور که مشاهده می‌شود تبدیل EGD بهترین نقشه تعامل را برای مدل فراهم کرده است و تقریباً ناحیه مورد نظر کاربر را به خوبی نمایش داده است. استفاده از تبدیل EGD را می‌توان به عنوان یکی از کارهای آینده در نظر گرفت.



شکل ۱-۵ : مقایسه تبدیل‌های مختلف برای کلیک‌های کاربر

#### ۴-۲-۵ پیشنهادهای دیگر

علاوه بر موارد بالا، استفاده از نسخه ۲۰۲۲ مجموعه داده BraTS که مجموعاً دارای ۱۰ برابر نسخه ۲۰۲۰ (تقریباً ۵۸۸۰) اسکن MRI می‌باشد می‌تواند برای آموزش مدل بسیار مناسب باشد. در این پژوهه از روش‌های داده افزایی بسیار ساده نظیر Flip و Rotaion استفاده شده است. در آینده می‌توان از روش‌های بیشتری در قسمت داده افزایی استفاده کرد. استفاده از ناحیه‌بندها و Backbone‌های دیگر هم می‌تواند جزو کارهای آینده باشد.

## 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] Chen, S., Chen, Z., Yan, L., and Guo, W. R-net: A novel fully convolutional network–based infrared image segmentation method for intelligent human behavior analysis. *Infrared Physics & Technology* 123 (04 2022), 104164.
- [5] Chen, X., Zhao, Z., Zhang, Y., Duan, M., Qi, D., and Zhao, H. Focalclick: Towards practical interactive image segmentation. in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (Los Alamitos, CA, USA, jun 2022), IEEE Computer Society, pp. 1290–1299.
- [6] Chollet, F. Xception: Deep learning with depthwise separable convolutions. in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (Los Alamitos, CA, USA, jul 2017), IEEE Computer Society, pp. 1800–1807.
- [7] Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L. Imagenet: A large-scale hierarchical image database. in *2009 IEEE Conference on Computer Vision and Pattern Recognition* (2009), pp. 248–255.
- [8] Elharrouss, O., Al-Maadeed, S. A., Subramanian, N., Ottakath, N., Almaadeed, N., and Himeur, Y. Panoptic segmentation: A review. *ArXiv abs/2111.10250* (2021).

- [9] 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.
- [10] 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.
- [11] 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.
- [12] He, J., Kim, C.-S., and Kuo, C.-C. J. *Interactive Image Segmentation Techniques*. Springer Singapore, Singapore, 2014, pp. 17–62.
- [13] He, K., Gkioxari, G., Dollár, P., and Girshick, R. Mask r-cnn. in *2017 IEEE International Conference on Computer Vision (ICCV)* (2017), pp. 2980–2988.
- [14] 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.
- [15] 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.
- [16] Huang, G., Liu, Z., and Weinberger, K. Q. Densely connected convolutional networks. *CoRR abs/1608.06993* (2016).
- [17] 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).
- [18] 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).
- [19] 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.

- [20] 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.
- [21] Li, A. Y., and Iv, M. Conventional and advanced imaging techniques in post-treatment glioma imaging. *Frontiers in Radiology* 2 (2022).
- [22] Li, H., and Peers, P. Crf-net: Single image radiometric calibration using cnns. in *Proceedings of the 14th European Conference on Visual Media Production (CVMP 2017)* (12 2017), pp. 1–9.
- [23] 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.
- [24] 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.
- [25] Meila, M. Comparing clusterings: An axiomatic view. in *Proceedings of the 22nd International Conference on Machine learning* (01 2005), pp. 577–584.
- [26] Menze, B. H. e. a. The multimodal brain tumor image segmentation benchmark (brats). *IEEE Transactions on Medical Imaging* 34, 10 (2015), 1993–2024.
- [27] 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).
- [28] 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.
- [29] 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.
- [30] 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.

- [31] Ramadan, H., Lachqar, C., and Tairi, H. A survey of recent interactive image segmentation methods. *Computational Visual Media* (08 2020).
- [32] Rayleigh. The problem of the random walk. *Nature* 72 (1905), 318–318.
- [33] Ronneberger, O., Fischer, P., and Brox, T. U-net: Convolutional networks for biomedical image segmentation. in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* (Cham, 2015), Springer International Publishing, pp. 234–241.
- [34] Rother, C., Kolmogorov, V., and Blake, A. Grabcut -interactive foreground extraction using iterated graph cuts. *ACM Transactions on Graphics (SIGGRAPH)* (August 2004).
- [35] Rother, C., Kolmogorov, V., and Blake, A. Grabcut: Interactive foreground extraction using iterated graph cuts. *ACM Trans. Graph.* 23 (08 2004), 309–314.
- [36] 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.
- [37] Sofiiuk, K., Petrov, I. A., and Konushin, A. Reviving iterative training with mask guidance for interactive segmentation. in *2022 IEEE International Conference on Image Processing (ICIP)* (2022), pp. 3141–3145.
- [38] Song, W., Yu, H., and Wu, J. Patch network for medical image segmentation, 2023.
- [39] 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.
- [40] Wang, G., Zuluaga, M., Li, W., Aughwane, R., Patel, P., Aertsen, M., Doel, T., David, A., Deprest, J., Ourselin, S., and Vercauteren, T. Deepigeos: A deep interactive geodesic framework for medical image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence PP* (07 2017).
- [41] Wang, J., Sun, K., Cheng, T., Jiang, B., Deng, C., Zhao, Y., Liu, D., Mu, Y., Tan, M., Wang, X., Liu, W., and Xiao, B. Deep high-resolution representation learning for visual recognition, 2019.
- [42] Xian, M., Zhang, Y., Cheng, H.-D., Xu, F., and Ding, J. Neutro-connectedness cut. *Trans. Img. Proc.* 25, 10 (oct 2016), 4691–4703.

- [43] Xie, E., Wang, W., Yu, Z., Anandkumar, A., Alvarez, J. M., and Luo, P. Segformer: Simple and efficient design for semantic segmentation with transformers. in *Advances in Neural Information Processing Systems 34 (NeurIPS 2021)* (12 2021).
- [44] 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.
- [45] xu, N., Price, B., Cohen, S., Yang, J., and Huang, T. Deep grabcut for object selection. in *2017 British Machine Vision Conference* (07 2017).
- [46] 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.

## پیوست‌ها

گیتهاب پیاده‌سازی پروژه: [لینک](#) Repository

نتایج آزمایش‌های مختلف به همراه فایل مدل آموزش دیده و تصاویر خروجی: [لینک](#)

گزارش‌ها و ارائه‌های پروژه: [لینک](#)

مجموعه داده BraTS استفاده شده: [لینک](#)

محیط و نوتبوک آموزش و ارزیابی مدل بر روی Google Colab: [لینک](#)

محیط و نوتبوک آموزش و ارزیابی مدل بر روی Kaggle: [لینک](#)

محیط و نوتبوک ارزیابی مدل از پیش آموزش داده شده بر روی Kaggle: [لینک](#)

**Abstract:**

Accurately segmenting 3D images of the human brain is crucial for medical applications such as tumor diagnosis and treatment planning. Deep neural networks have emerged as the most advanced automatic segmentation method, but they may require modifications for clinical use. Interactive segmentation allows for greater accuracy by incorporating user interactions, but current methods are not suitable for low-power systems and may require significant user effort. This study utilizes a novel method that performs a coarse segmentation on a low-resolution section of the target region, followed by a local refinement to restore lost resolution. To optimize efficiency, morphological analysis is used to modify only areas that need updating while preserving previously obtained results for other regions. Experimental results demonstrate that this method achieves more accurate results with less user interaction and requires less computing power and time compared to other methods. Additionally, this method shows acceptable generalization in tasks not encountered during the training phase.

**Keywords:** Deep Learning, Computer Vision, Interactive Image Segmentation, Brain Tumor Segmentation (BraTS), 3D Medical Images



**Iran University of Science and Technology  
Computer Engineering Department**

# **Interactive Brain Tumor Segmentation in 3D Medical Images via Deep Learning**

**Bachelor of Science Thesis in Computer Engineering**

**By:**

**Ali Sedaghi Arghoun**

**Supervisor:**

**Dr. Mohammad Reza Mohammadi**

**March 2023**