

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشکده مهندسی کامپیوتر و فناوری اطلاعات

پایان نامه کارشناسی ارشد مهندسی هوش مصنوعی

بخش بندی خودکار تصاویر سی تی اسکن ریه به کمک یادگیری عمیق

نگارنده: یگانه جلالی

استاد راهنما

دکتر منصور فاتح

استاد مشاور

دکتر محسن رضوانی

مهر ماه ۹۹

شماره: ۵۷/۷۸
تاریخ: ۹۹/۷/۲۸

باسمه تعالی



فرم شماره (۳) صورتجلسه نهایی دفاع از پایان نامه دوره کارشناسی ارشد

با نام و یاد خداوند متعال، ارزیابی جلسه دفاع از پایان نامه کارشناسی ارشد خاتم یگانه جلالی با شماره دانشجویی ۹۷۰۴۹۷۴ رشته مهندسی کامپیوتر گرایش هوش مصنوعی تحت عنوان بخش بندی خودکار تصاویر سی تی اسکن ریه به کمک یادگیری عمیق که در تاریخ ۱۳۹۹/۷/۱۴ با حضور هیأت محترم داوران در دانشگاه صنعتی شاهرود برگزار گردید به شرح ذیل اعلام می گردد:

الف) درجه عالی: نمره ۱۹-۲۰ ☐ ب) درجه خیلی خوب: نمره ۱۸-۱۹ ☐
 ج) درجه خوب: نمره ۱۷-۱۶ ☐ د) درجه متوسط: نمره ۱۵-۱۴ ☐
 ه) کمتر از ۱۴ غیر قابل قبول و نیاز به دفاع مجدد دارد ☐
 نوع تحقیق: نظری ☐ عملی ☐

عضو هیأت داوران	نام و نام خانوادگی	مرتبه علمی	امضاء
۱- استاد راهنمای اول	دکتر منصور دایخ	استادیار	
۲- استاد راهنمای دوم			
۳- استاد مشاور	دکتر محسن رحمانی	استادیار	
۴- نماینده تحصیلات تکمیلی	مهندس دایخ	مربی	
۵- استاد ممتحن اول	دکتر مرتضی زاهدی	استادیار	
۶- استاد ممتحن دوم	دکتر ناصر دایخ	استادیار	

نام و نام خانوادگی رئیس دانشکده:

تاریخ و امضاء و مهر دانشکده:

تقدیم بہ

خدائی کہ آفرید

جهان را، انسان را، عقل را، علم را، معرفت را، عشق را
و بہ کسانی کہ عشقشان را در وجودم دمید.

و

پدر و مادر کہ از نگاہشان صابست، از رفتارشان محبت
و از صبرشان ایستادگی را آموختم.

و

ہمسرم

کہ سایہ مہربانش سایہ ساز زندگی می باشد او کہ اسوہ صبر و تحمل بودہ
و سختی های مسیر را برایم تسہیل نمود

یکانہ جلالی

مہرماہ ۹۹

از استاد گرامی جناب آقای دکتر منصور فتح بیار پاسکوزارم چرا که بدون راهنمایی ایشان انجام این پایان نامه بسیار مشکل بود و ایشان سپیدی
را بر تخته سیاه زندگی اینجانب نگاشتند.

و در پایان از کلیه اساتید، خانواده و دوستان عزیز که در تمام مراحل انجام این رساله اینجانب را یاری نمودند، بهت همگاری بیدریشان بهت
پیشبر و این پایان نامه پاسکوزارم.

تعمدنامه

اینجانب ..یگانه جلالی..... دانشجوی دوره کارشناسی ارشد (دکتری) رشته ..هوش مصنوعی.... دانشکدهمهندسی کامپیوتر و فناوری اطلاعات.... دانشگاه صنعتی شاهرود نویسنده پایان نامه**"بخش بندی خودکار تصویر سی تی اسکن ریه به کمک یادگیری عمیق"**..... تحت راهنمایی ...دکتر منصور فاتح..... متعهد می شوم.

- تحقیقات در این پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است .
- در استفاده از نتایج پژوهشهای محققان دیگر به مرجع مورد استفاده استناد شده است .
- مطالب مندرج در پایان نامه تاکنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است .
- کلیه حقوق معنوی این اثر متعلق به دانشگاه صنعتی شاهرود می باشد و مقالات مستخرج با نام « دانشگاه صنعتی شاهرود » و یا « Shahrood University of Technology » به چاپ خواهد رسید .
- حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی پایان نامه تأثیرگذار بوده اند در مقالات مستخرج از پایان نامه رعایت می گردد.
- در کلیه مراحل انجام این پایان نامه ، در مواردی که از موجود زنده (یا بافتهای آنها) استفاده شده است ضوابط و اصول اخلاقی رعایت شده است .
- در کلیه مراحل انجام این پایان نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است اصل رازداری ، ضوابط و اصول اخلاق انسانی رعایت شده است .

تاریخ

امضای دانشجو

مالکیت نتایج و حق نشر

کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج ، کتاب ، برنامه های رایانه ای ، نرم افزار ها و تجهیزات ساخته شده است) متعلق به دانشگاه صنعتی شاهرود می باشد . این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود . استفاده از اطلاعات و نتایج موجود در پایان نامه بدون ذکر مرجع مجاز نمی باشد.

چکیده

بخش‌بندی تصاویر سی‌تی‌اسکن ریه، یک فرآیند کلیدی بوده و به‌عنوان اولین و مهم‌ترین قدم، در بسیاری از کاربردها مانند سیستم تشخیص سرطان ریه است. با توجه به ساختار غیریکسان و ناهم‌هنگ ریه و نیز استفاده از وسایل تصویربرداری متفاوت با پروتکل‌های مختلف، بخش‌بندی ناحیه ریه همیشه یک مسئله چالش‌برانگیز بوده‌است. برخی از روش‌های بخش‌بندی موجود، نیمه‌خودکار بوده و به عامل انسانی وابسته هستند. مشکل دیگر این روش‌ها، بالا بودن میزان مثبت کاذب است. در سال‌های اخیر، شبکه‌های مبتنی بر یادگیری عمیق در بخش‌بندی تصاویر پزشکی، به عملکرد پیشرفته‌ای دست یافته‌اند. در بین شبکه‌های موجود، U-Net موفقیت بسیاری در این زمینه کسب کرده‌است. در روش پیشنهادی این پایان‌نامه، ابتدا چندین عمل پیش‌پردازش گسترده بر روی تصاویر خام سی‌تی‌امال شده، سپس ماسک (برچسب) مربوط به این تصاویر، با استفاده از الگوریتم خودکار مبتنی بر عملیات مورفولوژیکی و انجام اندک اصلاحات دستی استخراج می‌شوند. در نهایت، همه تصاویر را به همراه برچسب‌های متناظرشان به یک شبکه عمیق، برای بخش‌بندی داده می‌شوند. معماری شبکه عمیق پیشنهادی، بر پایه یکی از نسخه‌های تغییر یافته U-Net بوده (BCDU-Net) که در بخش رمزگذار آن، با شبکه ResNet34 از پیش آماده شده، جایگزین شده‌است. همچنین در شبکه پیشنهادی، از مکانیزم لایه‌های کانولوشنی متراکم در آخرین لایه کانولوشنی مسیر رمزگذاری استفاده شده و برای ترکیب نقشه‌های ویژگی استخراج شده، در مسیر اتصال پرش از BConvLSTM استفاده شده‌است. در نهایت الگوریتم ارائه شده بر روی مجموعه داده LIDC-IDRI آزمایش شده و نتایج، بیانگر آن است که این روش، علاوه بر رفع چالش‌های مطرح شده، مقدار دایس ۹۷,۳۱٪ را به دست آورده که نسبت به بهترین روش قبل یعنی BCDU-Net، افزایش یک درصدی داشته‌است.

کلمات کلیدی: بخش‌بندی، سی‌تی‌اسکن، ریه، U-Net، ResNet34، BConvLSTM

لیست مقالات مستخرج از پایان نامه

1. Res BCDU-Net: A Deep Learning Framework for Lung CT Image Segmentation

فهرست مطالب

فهرست جداول.....	۵
فهرست اشکال.....	۹
فصل ۱: مقدمه.....	۱
۱-۱ پیش گفتار.....	۲
۲-۱ هدف تحقیق.....	۳
۳-۱ چالش‌ها و راه‌حل‌های آن.....	۴
۴-۱ سوال‌های تحقیق.....	۵
۵-۱ فرضیه‌های تحقیق.....	۶
۶-۱ نوآوری‌های تحقیق.....	۶
۷-۱ ساختار پایان‌نامه.....	۶
فصل ۲: ادبیات موضوع و کارهای پیشین.....	۷
۱-۲ مقدمه.....	۸
۲-۲ ساختار ریه.....	۸
۳-۲ سرطان ریه.....	۱۰
۴-۲ روش‌های مختلف تصویربرداری پزشکی.....	۱۱
۵-۲ قالب ذخیره‌سازی و مقیاس ارزیابی تصاویر سی‌تی.....	۱۴
۶-۲ پیش‌پردازش: نیازمندی‌های بهبود تصاویر.....	۱۷
۱-۶-۲ فیلترهای ریخت‌شناسی ریاضی.....	۱۷
۱-۶-۲-۱ فرسایش و گسترش.....	۱۸

۱۸.....	۲-۱-۶-۲ باز کردن و بستن
۲۰.....	۷-۲ بخش بندی ریه
۲۲.....	۸-۲ مروری بر روش های بخش بندی (کارهای پیشین)
۲۳.....	۱-۸-۲ روش های بخش بندی سنتی
۲۳.....	۱-۱-۸-۲ روش های مبتنی بر آستانه گذاری
۲۶.....	۲-۱-۸-۲ روش های رشد ناحیه های
۳۰.....	۳-۱-۸-۲ روش های مبتنی بر لبه
۳۳.....	۴-۱-۸-۲ روش های مبتنی بر مدل
۳۶.....	۹-۲ جمع بندی روش های سنتی بخش بندی
۳۸.....	۱۰-۲ مقدمه ای بر یادگیری عمیق
۳۹.....	۱۱-۲ جمع بندی
۴۱.....	فصل ۳: ساختارهای یادگیری عمیق
۴۲.....	۱-۳ مقدمه
۴۲.....	۲-۳ شبکه های عصبی کانولوشنی
۴۳.....	۳-۳ انواع لایه های شبکه کانولوشنی
۴۳.....	۱-۳-۳ لایه کانولوشن
۴۴.....	۲-۳-۳ لایه تجمیع (Pooling)
۴۴.....	۳-۳-۳ لایه تماما متصل (Fully Connected)
۴۵.....	۴-۳ مفاهیم و الگوریتم ها در حوزه شبکه های عصبی کانولوشن
۴۵.....	۱-۴-۳ پیش آموزش
۴۶.....	۲-۴-۳ ResNet الگوریتم
۴۷.....	۳-۴-۳ U-Net الگوریتم

۴۸.....	۳-۵ بخش‌بندی تصاویر پزشکی مبتنی بر یادگیری عمیق.....
۴۹.....	۳-۵-۱ ساختار RU-Net.....
۵۰.....	۳-۵-۲ ساختار Attention U-Net.....
۵۱.....	۳-۵-۳ ساختار ResNet34 U-Net.....
۵۳.....	۴-۵-۳ ساختار BCDU-Net.....
۵۸.....	۳-۶ جمع‌بندی.....
۵۹.....	فصل ۴: روش پیشنهادی.....
۶۰.....	۴-۱ مقدمه.....
۶۰.....	۴-۲ مجموعه‌داده‌گان LIDC-IDRI استفاده‌شده در این پایان‌نامه.....
۶۲.....	۴-۳ استخراج برچسب تصاویر سی‌تی.....
۶۶.....	۴-۴ پیش‌پردازش و آماده‌سازی داده‌ها.....
۶۹.....	۴-۵ مدل یادگیری عمیق پیشنهادی.....
۷۳.....	۴-۶ جمع‌بندی.....
۷۵.....	فصل ۵: پیاده‌سازی و نتایج.....
۷۶.....	۵-۱ مقدمه.....
۷۶.....	۵-۲ معیارهای ارزیابی.....
۷۹.....	۵-۳ نتایج آزمایش‌ها.....
۸۰.....	۵-۳-۱ دقت مدل پیشنهادی در مرحله آموزش و آزمایش.....
۸۱.....	۵-۳-۲ ماتریس درهم ریختگی مدل پیشنهادی.....
۸۲.....	۵-۳-۳ سطح زیر نمودار ROC برای مدل پیشنهادی.....
۸۳.....	۵-۳-۴ زمان اجرای الگوریتم تولید برچسب.....
۸۴.....	۵-۴ مقایسه روش پیشنهادی با روش‌های دیگر.....

۵-۵ جمع‌بندی ۸۹

فصل ۶: نتیجه‌گیری و پیشنهادها برای کارهای آتی..... ۹۱

۶-۱ مقدمه ۹۲

۶-۲ جمع‌بندی و نتیجه‌گیری ۹۲

۶-۳ پیشنهاد برای کارهای آتی ۹۳

مراجع ۹۴

فهرست جداول

جدول ۱-۲ مقادیر سی تی متناظر با بافت‌های گوناگون در واحد هانسفیلد ۱۶

جدول ۲-۲ مقایسه روش‌های سنتی مطرح‌شده برای بخش‌بندی ۳۷

جدول ۵-۱ مقایسه روش پیشنهادی با سایر روش‌ها ۸۶

فهرست اشکال

- شکل ۱-۱ نمایشی از ندول‌های متصل به دیواره ریه. فلش‌های موجود در دو تصویر نشان داده شده، نواحی را نشان می‌دهند که بافت ریوی هستند ولی در اکثر روش‌های بخش‌بندی، به‌عنوان نواحی غیرریوی در نظر گرفته می‌شوند. ۵
- شکل ۱-۲ آناتومی ریه ها ۱۰
- شکل ۲-۲ دستگاه اسکنر سی تی اسپیرال. بیمار روی تخت قرار می‌گیرد و وارد محفظه تونلی می‌شود که شامل منبع اشعه ایکس و حسگرهاست. ۱۲
- شکل ۳-۲ نمایشی از ساختار یک فایل DICOM. ۱۵
- شکل ۴-۲ اطلاعات سرآیند فایل DICOM مربوط به یک اسلایس تصویر سی تی ریه. ۱۶
- شکل ۵-۲ مثالی از اعمال عملیات مورفولوژی در بخش‌بندی ریه (آخرین تصویر از ستون سمت چپ، بیانگر آشکار ۲۰
- شکل ۶-۲ تصاویری از برش‌های قسمت‌های مختلف ریه مربوط به یک بیمار. ۲۲
- شکل ۷-۲ شماتیک روش‌های رشد ناحیه ای. (الف): انتخاب بذر اولیه (ب): بررسی نقاط مجاور و انجام فرآیند بخش‌بندی. ۲۷
- نقاط با علامت ضربدر، نشان دهنده قسمت جدا شده هستند. ۲۷
- شکل ۸-۲ بخش‌بندی ریه به روش رشد ناحیه. (الف): تصویر سی‌تی اولیه (ب): تصویر قطعه بندی شده به روش رشد ناحیه (فلش‌های موجود در دو تصویر نشان داده شده، توده چسبیده به دیواره ریه است که در روش رشد ناحیه، در نظر گرفته نشده.) ۲۹
- شکل ۹-۲ ماسک‌های لبه یاب سوبل (Sx، ماسک سوبل عمودی و Sy، ماسک سوبل افقی است) ۳۰
- شکل ۱-۳ یک طرح کلی از معماری یک شبکه عصبی کانولوشن ۴۳
- شکل ۲-۳ نمونه‌ای از یک بلوک باقی‌مانده ۴۶
- شکل ۳-۳ شبکه ResNet. ۴۷
- شکل ۴-۳ معماری شبکه U-Net. ۴۸

- شکل ۳-۵ معماری شبکه RU-Net ۵۰
- شکل ۳-۶ معماری شبکه Attention U-Net ۵۱
- شکل ۳-۷ معماری شبکه ResNet34 U-Net ۵۲
- شکل ۳-۸ (الف) بلوک‌های شبکه ResNet (ب) یک بلوک از شبکه DenseNet ۵۴
- شکل ۳-۹ بلوک Bi-directional ConvLSTM ۵۶
- شکل ۳-۱۰ شبکه BCDU-Net ۵۸
- شکل ۴-۱ تصاویر سی تی مربوط به یک بیمار از مجموعه داده IDRI-LIDC ۶۱
- شکل ۴-۲ گام‌های مرحله تولید ماسک برای تصاویر سی تی ۶۲
- شکل ۴-۳ (الف): تصویر سی تی اصلی (ب): تصویر باینری شده (ج): تصویر حاصل بعد از حذف راه های هوایی متصل به مرز تصویر ۶۳
- شکل ۴-۴ (الف): تصویر برچسب گذاری شده (ب): تصویر برچسب گذاری شده با حفظ دو بزرگ ترین مولفه ۶۴
- شکل ۴-۵ تصویر حاصل بعد از (الف) عملیات فرسایش (ب) عملیات بستن (ج) پرکردن حفره های کوچک ۶۵
- شکل ۴-۶ (الف) : نمایش مشکل غفلت از برخی بافت های ریوی در تصویر ماسک حاصل، به دلیل در نظر گرفتن فقط دو بزرگ ترین ناحیه برچسب زده شده (ب) نمایش مشکل اتصال پیکسل های غیر ریوی در تصویر ماسک حاصل، به دلیل خطا در مرحله بستن ۶۶
- شکل ۴-۷ (الف): تصویر اصلی (ب): تصویر باینری شده ۶۷
- شکل ۴-۸ (الف): تصویر دودویی شده (ب): تصویر پس از اعمال عملگر گسترش ۶۸
- شکل ۴-۹ تصویر پس از عملیات تشخیص لبه با الگوریتم کنی ۶۸
- شکل ۴-۱۰ تصویر جدید تولید شده با ترکیب سه کانال جدید طراحی شده ۶۹
- شکل ۴-۱۱ بلوک دیاگرام الگوریتم ResNet در بخش کدگذار شبکه پیشنهادی ۷۱
- شکل ۴-۱۲ شبکه عمیق پیشنهادی ۷۲
- شکل ۵-۱ شکل کلی ماتریس درهم ریختگی برای یک مسئله دو کلاسه ۷۹

- شکل ۵-۲ نمودار دقت در مرحله آموزش و آزمایش بر اساس تعداد تکرار ۸۱
- شکل ۵-۳ ماتریس درهم ریختگی مدل پیشنهادی ۸۲
- شکل ۵-۴ نمودار ROC مدل پیشنهادی ۸۳
- شکل ۵-۵ زمان اجرای الگوریتم خودکار تولید برچسب ۸۴
- شکل ۵-۶ مقایسه کیفی نتایج قطعه‌بندی با الگوریتم‌های مختلف. ۸۷
- (الف): تصاویر اصلی (ب): تصاویر برچسب (ماسک دودویی) (پ): تصاویر قطعه‌بندی شده توسط U-Net (ت): تصاویر قطعه‌بندی شده توسط BCDU-Net (ث): تصاویر قطعه‌بندی شده توسط روش پیشنهادی ۸۷
- شکل ۵-۷ نمایشی از رفع چالش بالا بودن نرخ مثبت کاذب در روش پیشنهادی ۸۸
- (الف): تصویر اصلی (ب): تصویر برچسب (ماسک دودویی) (پ): تصاویر قطعه‌بندی شده توسط BCDU-Net (ث): تصویر قطعه‌بندی شده توسط روش پیشنهادی (بافت ریز ریوی، با فلش نشان داده شده‌است). ۸۸
- شکل ۵-۸ نمایشی از رفع چالش غفلت از ندول‌های متصل به دیواره ریه در روش پیشنهادی ۸۹
- (الف): تصویر اصلی (ب): تصویر برچسب (ماسک دودویی) (پ): تصاویر قطعه‌بندی شده توسط BCDU-Net (ث): تصویر قطعه‌بندی شده توسط روش پیشنهادی (دور مکان ندول متصل به دیواره ریه، دایره کشیده شده‌است). ۸۹

فصل ۱ : مقدمه

۱-۱ پیش‌گفتار

سرطان یکی از بیماری‌های مهلک در جهان است. در بین انواع سرطان، سرطان ریه، سومین علت اصلی مرگ و میر در سرطان‌ها، پس از سرطان‌های پروستات و سینه محسوب می‌شود [۱]. لذا تشخیص زودهنگام و دقیق سرطان ریه یکی از مهم‌ترین چالش‌ها در دهه‌های اخیر است. بنابراین برای یک بیمار سرطانی در مراحل اولیه بسیار مهم است که بیماری وی تشخیص داده‌شود و لذا تشخیص با دقت و سرعت بالا می‌تواند در رسیدن به این مهم بسیار امیدوارکننده باشد. در این راستا، باید دانست که در صورت قطعه‌بندی مناسب تصاویر ریه، امکان تحلیل دقیق و در نتیجه تشخیص صحیح و به‌هنگام بیماری میسر می‌شود [۲].

قطعه‌بندی^۱ تصویر به فرآیند تفکیک اجزای تصویر اطلاق می‌شود. به‌طور دقیق‌تر، قطعه‌بندی تصویر، فرآیند نسبت دادن یک برچسب به هریک از پیکسل‌های تصویر است، به‌گونه‌ای که پیکسل‌های با برچسب یکسان، مشخصه‌های بصری مشابهی داشته‌باشند [۳]. بخش‌بندی تصاویر پزشکی، گام مهمی در جهت تحلیل آن‌ها محسوب می‌شود. هدف اصلی در قطعه‌بندی تصاویر پزشکی، تقسیم‌بندی آن‌ها به ساختارهای آناتومیکی مختلف و سپس جدا نمودن اجزای مطلوب است. با افزایش استفاده از تصاویر پرتونگاری کامپیوتری (CT)^۲ و تشدید مغناطیسی (MR)^۳، استفاده از سیستم‌های کمک کامپیوتری (CAD)^۴ برای کمک به متخصصان، امری الزامی شده‌است [۴و۵]. معمولاً در همه سیستم‌های کمک تشخیص کامپیوتری تصاویر ریه، قطعه‌بندی ریه انجام می‌شود. نوعاً قسمت اصلی ریه حدود ۲۵٪ از کل تصویر سی‌تی‌اسکن ریه را شامل می‌شود [۶]. در قطعه‌بندی ریه، بافت ریه مشخص شده و بقیه تصویر حذف می‌شود. بنابراین نتیجه گرفته می‌شود که قطعه‌بندی تصاویر ریه تاثیر زیادی بر روی کاهش هزینه محاسبات، قابلیت اطمینان و دقت سیستم کمک تشخیص کامپیوتری دارند [۷]. به‌طور کلی طراحی یک روش قطعه‌بندی موثر، همیشه یک چالش است. تاکنون

¹ Segmentation

² Computed Tomography

³ Magnetic Resonance

⁴ CAD: Computer Aided Design

روش‌های بسیاری برای قطعه‌بندی تصاویر پزشکی به‌کاربرده شده‌است که در فصل‌های آینده به‌طور مفصل شرح داده خواهند شد. اما همواره این مسئله وجود دارد که روش‌های کامپیوتری قطعه‌بندی تصاویر، با چالش‌هایی مانند رزولوشن پایین و کنتراست ضعیف روبه‌رو هستند. وجود نویز و آرتیفکت این مشکل را تشدید می‌کنند. علاوه بر این، با توجه به پیچیدگی و نیز اهمیت قطعه‌بندی تصاویر پزشکی، اغلب سیستم‌های رایانه‌ای به صورت نیمه‌خودکار یا در حالت تعاملی اجرا می‌شوند. طبیعتاً این الگوریتم‌ها علاوه بر محدودیت‌های کاربردی، نیازمند دخالت عامل انسانی نیز هستند. از این‌رو استفاده از الگوریتم‌های خودکار و بدون نیاز به نظارت انسان، علاوه بر کمک در سهولت در تشخیص و قطعه‌بندی، می‌تواند روش‌های پردازشی و تحلیلی را نیز ارتقا بخشد.

۱-۲ هدف تحقیق

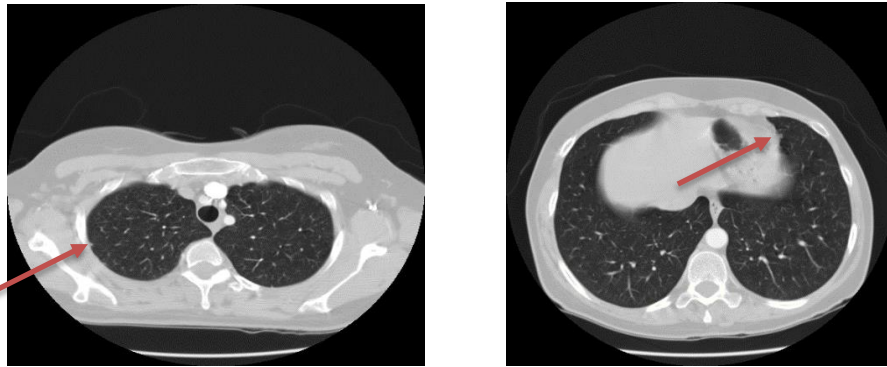
بخش‌بندی تصاویر توموگرافی ریه و شناسایی پیکسل‌ها یا ضایعات ناشی از این تصاویر، یکی از چالش‌ها در حوزه تحلیل این تصاویر است. روش‌های بسیاری برای قطعه‌بندی تصاویر پزشکی به‌کار برده شده‌است که یکی از آن‌ها استفاده از شبکه‌های عصبی عمیق است. اخیراً روش‌های یادگیری به‌خوبی قادر بوده‌اند که نقش عمده و به‌سزایی در مباحث بینایی ماشین و از آن جمله بخش‌بندی نواحی مختلف یک تصویر داشته باشند.

به‌طور خاص، امروزه کاربرد تکنیک‌های یادگیری عمیق برای قطعه‌بندی تصاویر پزشکی، به دلیل توانایی آن‌ها برای یادگیری و پردازش مقادیر حجیم داده‌ها به‌طور سریع و دقیق، مورد توجه زیادی قرار گرفته‌است [۸]. هدف کلی این مقاله، ارائه یک روش کاملاً خودکار است که در آن با استفاده از یک شبکه عصبی عمیق مبتنی بر U-Net (شبکه‌ای کانولوشنی مختص بخش‌بندی تصاویر پزشکی)، تصاویر سی‌تی‌اسکن ریه با دقت مناسب و به‌طور خودکار، بخش‌بندی شوند.

۳-۱ چالش‌ها و راه‌حل‌های آن

تفسیر یک عکس رادیوگرافی یا توموگرافی ریه مشکل است. وجود ساختارهای اضافی و روی هم قرار گرفتن برخی ساختارها در تصویر باعث پیچیده شدن تصویر می‌شوند، به طوری که حتی رادیولوژیست‌های باتجربه هم در تشخیص بعضی موارد از جمله تشخیص ندول‌های^۱ کوچک ریوی، دچار مشکل می‌شوند است [۹]. علاوه بر این، تصاویر سی‌تی‌اسکن بسیار زیاد هستند و همین امر باعث اتلاف وقت متخصصان و خستگی آن‌ها شده و دقت تشخیص را پایین می‌آورد. شکل نامنظم، مکان پیچیده از لحاظ آناتومیک و گاهی اوقات روشنایی کم توده‌های ریوی باعث بروز مشکلاتی در آشکارسازی، جداسازی، تشخیص و محاسبه‌ی حجم توده‌ها به صورت دستی می‌شود. انجام دادن دستی این عملیات بسیار مشکل و وقت‌گیر بوده و با عدم دقت کافی همراه است. بنابراین با در نظر گرفتن حجم بالای تصاویر و چندلایه بودن آن‌ها، پردازش و تفسیر دستی آن‌ها بسیار دشوار است و در این شرایط تشخیص بیماری، همراه با خطای انسانی است [۱۰]. علاوه بر این، مشکلات متنوعی در سر راه الگوریتم‌های پیشنهادی در حوزه بخش‌بندی قرار دارد. به عنوان مثال کیفیت متنوع تصاویر که ناشی از استفاده از انواع دستگاه‌های تصویربرداری با پروتکل‌های غیریکسان است. اما آنچه که کار را سخت‌تر می‌کند، وجود بافت‌هایی درون لوب‌های ریه است. بافت‌هایی که بسیار شبیه به بافت‌های بیرونی ریه هستند اما درون ریه قرار دارند و البته این بافت‌ها زمانی که به دیواره ریه چسبیده‌اند، کار را برای الگوریتم‌های بخش‌بندی بسیار سخت‌تر می‌کنند؛ به طوری که بسیاری از الگوریتم‌ها، آن‌ها را جز بیرونی می‌بینند درحالی که آن‌ها ممکن است ندول‌ها یا غیره باشند. شکل ۱-۱ این موضوع را آشکارتر نشان می‌دهد.

¹ Nodule



شکل ۱-۱ نمایشی از ندول‌های متصل به دیواره ریه. فلش‌های موجود در دو تصویر نشان داده شده، نواحی را نشان می‌دهند که بافت ریوی هستند ولی در اکثر روش‌های بخش‌بندی، به عنوان نواحی غیرریوی در نظر گرفته می‌شوند. با توجه به این مسائل و همچنین گسترش حجم مطالعات سی تی همراه با افزایش حجم داده‌ها، نیاز به طرح سیستم‌های CAD است. این سیستم‌ها، تصاویر پزشکی را به کمک الگوریتم‌هایی مورد بررسی و پردازش قرار می‌دهند تا موارد مشکوک را شناسایی و به پزشک ارائه دهند. درواقع، پزشک از سیستم CAD برای ارائه یک نظر ثانویه جهت دستیابی به تشخیص دقیق‌تر بهره می‌برد. این روش، به طورگسترده برای بهبود اثربخشی درمان، مفید است. یکی از مراحل مهم و ابتدایی در اکثر این سیستم‌ها، بحث بخش‌بندی تصاویر است [۱۰ و ۱۱]. این مرحله، یک پیش‌نیاز برای انجام عملیات تحلیل محتوای تصویر و تشخیص الگوهای مورد جستجو در تصاویر است [۱۱].

با توجه به این‌که در طراحی سیستم‌های CAD، هدف اصلی تولید خروجی قابل فهم برای پزشک است و با توجه به این‌که مرحله قطعه‌بندی ریه، قسمت پایه و ضروری این سیستم‌ها است، لذا هدف این تحقیق، ارائه روشی مناسب جهت رسیدن به قطعه‌بندی نزدیک به تشخیص پزشک است.

۴-۱ سوال‌های تحقیق

- آیا نتایج حاصل از بخش‌بندی تصاویر سی تی اسکن ریه با استفاده از یادگیری عمیق، با تصاویر قطعه‌بندی شده توسط رادیولوژیست هم‌پوشانی دارد؟
- آیا الگوریتم مورد استفاده در این تحقیق برای بخش‌بندی بافت ریه، قادر به تفکیک ندول‌های متصل به دیواره ریه که ممکن است با بافت اصلی ریه اشتباه گرفته شوند، است؟

۵-۱ فرضیه‌های تحقیق

- طبق تحقیقی که در سال ۲۰۱۶ انجام شد، استفاده ترکیبی از بهترین الگوریتم‌های یادگیری ماشین در حوزه پزشکی می‌تواند بهترین نتیجه‌ها را در بر داشته باشد.
- در صورت ترکیب معماری‌های موجود در یادگیری عمیق جهت بخش‌بندی تصاویر و تولید نقشه‌قطعه‌های (GT)^۱ مناسب و صحیح در مرحله اولیه، الگوریتم پیشنهادی ناژول‌های متصل به دیواره ریه را با بافت اصلی متمایز می‌شمارد.

۶-۱ نوآوری‌های تحقیق

به‌طور کلی، لیست نوآوری‌ها در این پایان‌نامه به شرح ذیل است:

- (۱) ارائه یک روش پیش‌پردازشی جهت تقویت تصاویر خام ورودی.
- (۲) ارائه روشی جهت تولید ماسک‌های دودویی متناظر با تصاویر.
- (۳) ارائه الگوریتم جدید مبتنی بر یادگیری عمیق جهت بخش‌بندی مناسب تصاویر سی‌تی ریه.
- (۴) افزایش دقت نسبت به روش‌های قبلی

۷-۱ ساختار پایان‌نامه

این پایان‌نامه در شش فصل ارائه خواهد شد. در فصل اول، مقدمه و کلیاتی از پژوهش را مطرح کردیم. در فصل دوم، مروری بر مباحث نظری داشته و کارهای پیشین در حوزه‌ی مورد تحقیق را مورد بررسی قرار خواهیم داد. در فصل سوم، به معرفی یادگیری عمیق و چند الگوریتم بخش‌بندی مبتنی بر آن، خواهیم پرداخت. در فصل چهارم روش پیشنهادی تحقیق را به‌طور کامل مطرح کرده و در فصل پنجم به بیان پیاده‌سازی روش پیشنهادی و آزمایش‌ها و تفسیر نتایج آن‌ها خواهیم پرداخت. در نهایت، در فصل ششم، نتیجه‌گیری و پیشنهادات آتی را بیان خواهیم کرد.

¹ Ground Truth

فصل ۲ : ادبیات موضوع و کاربردی پیشین

۲-۱ مقدمه

در این فصل، در ابتدا به توضیح در رابطه با ساختار آناتومی ریه و سرطان ریه پرداخته می‌شود. در قسمت بعد، بعضی از روش‌های مختلف تصویربرداری تشخیصی به‌عنوان روش‌هایی برای تشخیص این بیماری به اختصار مرور می‌شوند. در مرحله بعد، مباحثی در رابطه با قالب ذخیره‌سازی و مقیاس ارزیابی تصاویر سی‌تی مطرح خواهد شد. سپس تعریفی از مفهوم بخش‌بندی و انواع روش‌های آن در قالب کارهای پیشین، خواهیم داشت. در قسمت آخر، به جمع‌بندی خلاصه‌ای از مطالب گفته‌شده در این فصل خواهیم پرداخت.

۲-۲ ساختار ریه

ریه‌ها در قفسه سینه قرار دارند و آنقدر بزرگ هستند که بیشتر فضای سینه را گرفته‌اند. هر فرد سالم دو ریه دارد که یکی سمت راست و دیگری در سمت چپ قفسه سینه واقع شده‌است. ریه سمت راست^۱ به سه قسمت تقسیم می‌شود در حالی که ریه سمت چپ^۲ تنها دو قسمت دارد [۱۲]. هر ریه به وسیله یک پوشش غشایی دولایه (پلورا) محاصره شده که آن را به قفسه سینه متصل می‌کند. فضای میان دو لایه پلورا به وسیله مایعی پر شده‌است. فضای اضافی سمت چپ، جایگاه قلب است. قفسه سینه از ریه‌ها محافظت می‌کند. قفسه سینه از ۱۲ دنده تشکیل شده است این دنده‌ها به ستون فقرات پشت کمر و دور ریه‌ها به هم متصل شده‌اند تا از ریه‌ها محافظت کنند؛ دیا‌فراگم در زیر ریه‌ها قرار دارد. دیا‌فراگم ماهیچه گنبدی شکل است که با ریه کار می‌کند تا عمل دم و بازدم انجام شود.

ریه‌ها از بیرون شبیه اسفنج صورتی رنگ هستند، اما اسرار واقعی درون ریه‌ها است. در انتهای نای^۳ دو لوله بزرگ وجود دارد که به این لوله‌ها، لوله نای اصلی می‌گویند. یکی از این لوله‌ها به سمت چپ

¹ Right Lung

² Left Lung

³ Trachea

و دیگری به سمت راست متصل است. هر شاخه اصلی نایژه، شاخه‌هایی به درون لوله‌ها دارد که این شاخه‌ها مثل شاخه‌های درخت، کوچک و کوچک‌تر می‌شوند، به کوچک‌ترین لوله، نایچه^۱ گفته می‌شود؛ حدود سی‌هزار نایچه در هر ریه وجود دارد و هر نایچه تقریباً به اندازه مو می‌باشد. شکل ۱-۲ قسمت‌های مختلف ریه را نشان می‌دهد.

از آنجایی که ریه‌ها در قفسه سینه محبوس شده‌اند، برای ارتباط داشتن با محیط بیرون نیازمند مسیرهای جریان هوا هستند. در زیر به سیستم‌هایی که در جابه‌جایی هوا نقش دارند، اشاره شده است:

- بینی و دهان: ورودی‌هایی که اجازه ورود هوا به ریه‌ها را می‌دهند.
- حلق^۲: هوا را از طریق بینی و دهان به سمت لارنکس هدایت می‌کند.
- حنجره^۳: هوا را به سمت نای منتقل می‌کند.
- نای: به دو بخش لوله‌های برونشیت راست و چپ تقسیم می‌شود که هوا را به ریه‌های راست و چپ می‌رساند.
- برنشیل^۴: لوله‌های برونشیت کوچک‌تری هستند که هوا را به کسیه‌های کوچک‌تری که آلئول نام دارند می‌رسانند.
- آلئول^۵: کیسه‌های برونشویک که به وسیله مویرگ‌ها احاطه شده‌اند و سطح تنفسی ریه‌ها هستند.

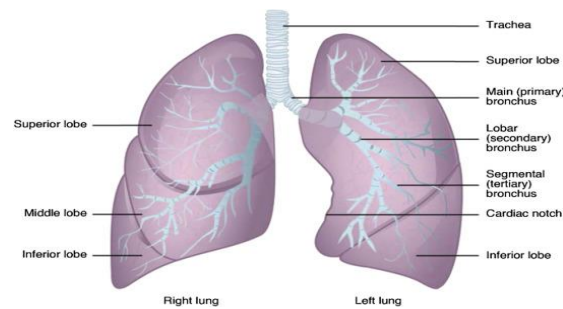
¹ Bronchi

² Throat

³ Larynx

⁴ Bronchial

⁵ Alveoli



شکل ۲-۱ آناتومی ریه‌ها [۱۳].

۳-۲ سرطان ریه

سرطان، به معنای رشد کنترل نشده‌ی سلول‌های غیر طبیعی است که کارکرد سلول‌های طبیعی را ندارند [۱۴]. این سلول‌ها می‌توانند تومورها را شکل دهند، سیستم ایمنی را آسیب بزنند و انحرافات دیگری را ایجاد کنند که از کارکرد صحیح بدن جلوگیری می‌کند. سرطان ریه، یک تومور سرطانی بدخیم است که با رشد کنترل نشده سلولی در بافت‌های ریه شکل می‌گیرد [۱۵]. سرطان ریه، علت اصلی مرگ در بین بیماری‌های ریوی است [۱۶].

گزارشی توسط سازمان بهداشت جهانی اعلام شده است که نشان می‌دهد سرطان ریه باعث مرگ بیش از ۳,۱ میلیون نفر در هر سال می‌شود [۱۷]. در سال ۲۰۱۹ در حدود ۱۱۶,۴۴۰ مرد و ۱۱۰,۷۱۰ زن در ایالات متحده (US)^۱ مبتلا به سرطان ریه تشخیص داده شدند و ۸۶ درصد این جمعیت، در مدت ۵ سال جان خود را از دست دادند [۱۸]. آماری که در سال ۲۰۲۰ از انجمن نشریه سرطان در آمریکا منتشر شد، نشان می‌دهد سرطان ریه منجر به مرگ ۱۳۵,۷۲۰ نفر، در ایالات متحده شده است [۱۹]. به طور کلی سرطان ریه، علت اصلی مرگ ۲۶ درصد از همه مرگ‌های سرطانی در مردان و ۲۵ درصد زنان را تشکیل می‌دهد.

سرطان ریه، در مراحل اولیه تقریباً هیچ نشانه خاصی از خود در تصاویر سی‌تی‌اسکن قفسه سینه بیمار نشان نمی‌دهد و بیماران اغلب زمانی به پزشک مراجعه می‌کنند که بیماری در مراحل پیشرفته خود قرار دارد. باید دانست که سرطان ریه در مراحل اولیه، احتمال درمان بیشتری نسبت به

¹ United States

تومورهای مرحله آخر دارد، در نتیجه تشخیص زودهنگام و درمان برای بهبود وضعیت بالینی بیماران بسیار مهم هستند [۲۰]. بنابراین برای یک بیمار سرطانی فراهم آمدن امکان تشخیص در مراحل اولیه بیماری، بسیار حائز اهمیت است. با توجه به این مسائل، استفاده از روشی که بتواند تشخیص با دقت و سرعت بالا را مهیا سازد، ضروری به نظر می‌رسد.

۲-۴ روش‌های مختلف تصویربرداری پزشکی

رادیوگرافی معمولی قفسه سینه، توموگرافی کامپیوتری، تشدید مغناطیسی و توموگرافی انتشار پوزیترون (PET)^۱ همه تکنیک‌های تصویربرداری هستند که برای تشخیص، شناسایی، طبقه‌بندی و پیگیری سرطان ریه مورد استفاده قرار می‌گیرند. در این تحقیق، عمل بخش‌بندی بر روی تصاویر سی‌تی‌اسکن انجام شده‌اند؛ لذا در ادامه، به بیان مزایا و معایب برخی از این روش‌ها می‌پردازیم.

رادیوگرافی قفسه سینه، رایج‌ترین آزمایش تصویربرداری تشخیصی برای تشخیص بسیاری از بیماری‌های ریوی در سراسر جهان است. سادگی، هزینه پایین، تشعشع کم، مقادیر زیاد اطلاعات و در دسترس بودن گسترده، مزایای این روش تصویربرداری هستند [۲۱]. رادیوگرافی قفسه سینه می‌تواند اندازه ضایعه را ارزیابی کند و ویژگی‌هایی مانند فروپاشی بعد از انسداد، افیوژن پلور و در برخی موارد گسترش خارج سلولی را نشان دهد. با این وجود، رادیوگرافی قفسه سینه قادر به تشخیص مرز دیواره ریه، دیافراگم و مدیاستینوم و محل درگیر شده توسط ندول ریوی نیست [۲۲].

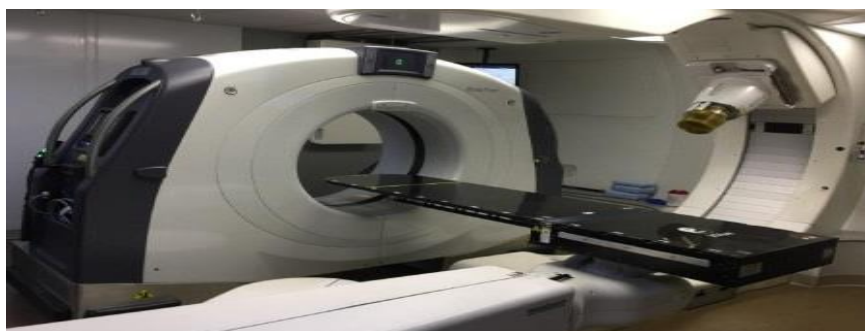
تصویربرداری توموگرافی کامپیوتری، نوعی از تصویربرداری است که با گرفتن تعدادی عکس توسط اشعه‌ی ایکس در ارتباط با الگوریتم‌ها و محاسبات رایانه‌ای، تصویری سه‌بعدی از بدن را نشان می‌دهد. در این روش، بیمار روی تخت متحرکی قرار گرفته و وارد محفظه تونلی می‌شود که این محفظه شامل منبع اشعه ایکس و تعدادی حس‌گر است. سپس اطلاعات متعددی متناسب با شدت اشعه‌ی دریافت شده، توسط آشکارسازها در یافت می‌شود و به رایانه مرکزی سی‌تی‌اسکن ارسال می‌شود. در نهایت

¹ Positron Emission Tomography

این رایانه، اطلاعات دریافتی را پردازش کرده و نتیجه‌ی آن را به صورت نمایش‌های متعددی بر روی مانیتور دستگاه نشان می‌دهد [۲۳]. شکل ۲-۲ یک دستگاه سی‌تی‌اسکن مدرن را نشان می‌دهد که شامل محفظه تونلی، منبع اشعه ایکس و حسگرها است.

قبل از سال ۱۹۹۱، یک سی‌تی‌اسکن معمولی یا افزایشی از قفسه سینه انجام می‌شد و شامل یک سری از برش‌های مقطع به‌دست‌آمده در هنگام تنفس معلق^۱ بود. در سال ۱۹۹۱، سی‌تی مارپیچ در تصویربرداری از قفسه سینه معرفی شد و کیفیت تصاویر CT ساختارهای قفسه سینه را به طرز چشم‌گیری بهبود بخشید [۲۴].

از تصویربرداری سی‌تی برای ارزیابی میزان پیشرفت بیماری‌های ریوی مانند اندازه‌گیری تغییرات حجم تومور در طی زمان، تشخیص توابع اولیه ریوی، ارزیابی تهویه ریه (چرخه تنفسی) و غیره استفاده می‌کنند. این روش تصویربرداری، به‌دلیل ثبت تصاویر از مقاطع مختلف ریه با عمق‌های دلخواه، به‌عنوان استاندارد طلایی برای تشخیص سرطان ریه مورد شناخته شده‌است [۲۵-۲۷]. این نوع تصویربرداری، جزئیاتی را که در تصاویر رادیوگرافی مشخص نیست، آشکار می‌کند. به‌عنوان نمونه، ندول‌های کوچک ریوی که غالباً متاستاتیک‌اند، به این علت که توسط تشکیلات عروقی و سایه‌دهنده‌ها محو می‌شوند، معمولاً در رادیوگرافی معمولی قابل شناسایی نیستند. از میان همه سیستم‌های تصویربرداری پزشکی، سی‌تی‌اسکن به بهترین نحو این ضایعات را آشکار می‌کند.



شکل ۲-۲ دستگاه اسکنر سی‌تی اسپیرال. بیمار روی تخت قرار می‌گیرد و وارد محفظه تونلی می‌شود که شامل منبع اشعه ایکس و حسگرهاست [۲۸].

¹ “stop-and-go” Method

در تصویربرداری تشخیص مغناطیسی، تصاویر به واسطه یک میدان مغناطیسی بسیار قوی و امواج رادیویی از اندام‌ها و بافت‌های بدن ایجاد می‌شوند. این تصاویر از کیفیت بسیار بالایی برخوردارند و تمایز بین بافت نرمال و غیر نرمال را به خوبی نمایان می‌کنند. این روش تصویربرداری نسبت به سی‌تی‌اسکن، اطلاعات دقیق‌تری در مورد اندام‌های داخلی (بافت‌های نرم) ارائه می‌دهد. به عنوان مثال تصاویر ام‌آر‌آی نسبت به سی‌تی‌اسکن، تصاویری با جزئیات بیشتر از ستون فقرات کمری برای تشخیص بیرون‌زدگی دیسک ارائه می‌دهند.

هر دو اسکن ام‌آر‌آی و سی‌تی، می‌توانند ساختار داخلی بدن را مشاهده کنند. با این حال، سی‌تی‌اسکن سریع‌تر بوده و کمتر از پنج دقیقه طول می‌کشد. هم‌چنین می‌تواند تصاویری از بافت‌ها، اندام‌ها و ساختار اسکلتی ارائه دهد؛ در صورتی‌که ام‌آر‌آی معمولاً بیش از ۱۵ دقیقه طول می‌کشد.

اگر در مقام قیاس برآییم، می‌توان گفت به طور کلی سی‌تی‌اسکن بیشتر از ام‌آر‌آی مورد استفاده قرار می‌گیرد. یکی از دلایل آن تصویربرداری از بدن در مدت زمان کمتر است که در مواقع اورژانسی اهمیت زیادی دارد و دلیل دیگر آن، هزینه پایین سی‌تی‌اسکن نسبت به ام‌آر‌آی است. تفاوت اساسی سی‌تی‌اسکن و ام‌آر‌آی در منشا تصویربرداری آن‌ها است به گونه‌ای که در اسکن‌های سی‌تی از اشعه ایکس استفاده می‌شود و در تصویربرداری با ام‌آر‌آی از امواج رادیویی و مگنت استفاده می‌شود. هر دو روش سی‌تی‌اسکن و ام‌آر‌آی غیرتهاجمی و بدون درد هستند با این تفاوت که تصویربرداری با ام‌آر‌آی صدای بیشتری دارد و زمان بیشتری نیز طول می‌کشد.

با توجه به توضیحات مطرح‌شده در رابطه با هر یک از این روش‌ها، ما در این پایان‌نامه از تصاویر سی‌تی‌اسکن استفاده کردیم.

۲-۵ قالب ذخیره‌سازی و مقیاس ارزیابی تصاویر سی تی

معمولا قالب ذخیره‌سازی تصاویر پزشکی، DICOM^۱ است که به صورت "dcm" نوشته می‌شوند. یک فایل تصویری DICOM، نتیجه‌ای از استاندارد تصویربرداری دیجیتال و ارتباطات در پزشکی است. حالت استاندارد در سال ۱۹۸۳ توسط کمیته مشترکی از دانشکده رادیولوژی آمری (ACR)^۲ و انجمن تولیدکنندگان بین المللی الکتریکی (NEMA)^۳ شکل گرفت. هدف اولیه در این پروژه دست یابی به تصاویر دیجیتال پزشکی مستقل از دستگاه های تصویر برداری بود.

انجمن فیزیک پزشکی آمریکا (AAPM)^۴ استاندارد برای ثبت تصاویر بر روی نوارهای مغناطیسی تولید کردند. اولین نسخه از این استاندارد که ACR_NEMA نامیده می‌شد در سال ۱۹۸۵ منتشر شد، همانند هر نسخه ابتدایی ACR_NEMA 1.0 خطاهایی داشت و خیلی زود آشکار شد که این استاندارد به کار و تلاش مداوم بیشتری نیازمند است. به همین دلیل گروهی به منظور بهبود قسمتهایی خاص از این استاندارد در حال پیشرفت تشکیل شد که نتیجه آن نسخه دوم ACR_NEMA 2.0 بود که در سال ۱۹۸۸ منتشر شد. توانایی این وسیله برای ارتباطات پزشکی بسیار محدود بود. برای پاسخ به این نیازمندی‌ها، نسخه سوم استاندارد ACR_NEMA تهیه شد و در سال ۱۹۹۲ به فرم ساده ای ارائه گردید. این نسخه سوم ACR_NEMA DICOM یا DICOM 3.0 شد و این نسخه با نام DICOM معروف شد.

یک فایل DICOM متشکل از یک سرصفحه^۵ و مجموعه داده تصویر در یک فایل واحد است. شکل ۲-۳، ساختار یک فایل DICOM را نشان می‌دهد. اطلاعات درون سرصفحه به صورت یک سری ثابت و استاندارد از برچسب‌ها^۶ تشکیل شده است. با استخراج داده‌ها از این برچسب‌ها می‌توان به اطلاعات مهمی درمورد بیمار، نوع تصویربرداری، ضخامت برش اسلایس‌ها، فاصله‌ی بین پیکسل‌ها در تصویر و

^۱ Digital Imaging and Communication in Medicine

^۲ American College of Radiology

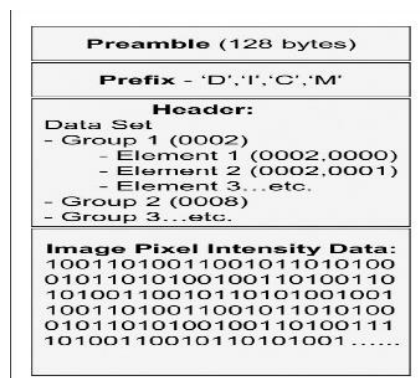
^۳ National Electrical Manufacturers Association

^۴ American Association of Physicists in Medicine

^۵ Header

^۶ Tags

ابعاد تصویر دست‌یافت [۲۹]. در مقایسه با فرمت های فایل تصویر مانند JPEG یا پرونده‌های TIFF، فایل‌های DICOM منفرد توسط ویندوز^۱ به‌عنوان فایل‌های تصویر شناخته نمی‌شوند. در نتیجه، فرد قادر نیست محتویات تصویر را با کلیک ساده بر روی آن‌ها مشاهده کند [۳۰].



شکل ۲-۳ نمایشی از ساختار یک فایل DICOM [۲۹].

چندین بسته نرم‌افزاری رایگان وجود دارد که می‌توانند برای استخراج اطلاعات از سرصفحه یک فایل DICOM استفاده شوند. برای مثال، DicomWorks یکی از نمایشگرهای محبوب DICOM است که می‌تواند اطلاعات سرصفحه را نمایش دهد [۳۱]. MicroDicom [۳۲]، ImageJ [۳۳] و XnView [۳۴]، از دیگر نرم‌افزارهای موجود برای خواندن اطلاعات سرصفحه فایل‌های DICOM محسوب می‌شوند. **Error! Not a valid bookmark self-reference.**، برخی از پارامتر-های سرصفحه‌ی یک فایل سی‌تی‌اسکن ریه را که توسط نرم‌افزار MicroDicom استخراج شده‌است، نشان می‌دهد.

¹ Windows

(Group, El...)	TAG Description	VR	V...	Length	Value
(0002,0000)	FileMetaInformationGroupLength	UL	1	4	196
(0002,0001)	FileMetaInformationVersion	OB	1	2	
(0002,0002)	MediaStorageSOPClassUID	UI	1	26	1.2.840.10008.5.1.4.1.1.2
(0002,0003)	MediaStorageSOPInstanceUID	UI	1	64	1.3.6.1.4.1.14519.5.2.1.6279.6001.143451261327128179989900
(0002,0010)	TransferSyntaxUID	UI	1	20	1.2.840.10008.1.2.1
(0002,0012)	ImplementationClassUID	UI	1	18	1.2.40.0.13.1.1.1
(0002,0013)	ImplementationVersionName	SH	1	14	dcm4che-1.4.35
(0008,0005)	SpecificCharacterSet	CS	1	10	ISO_IR 100
(0008,0008)	ImageType	CS	3	22	ORIGINAL\PRIMARY\AXIAL
(0008,0016)	SOPClassUID	UI	1	26	1.2.840.10008.5.1.4.1.1.2
(0008,0018)	SOPInstanceUID	UI	1	64	1.3.6.1.4.1.14519.5.2.1.6279.6001.143451261327128179989900
(0008,0020)	StudyDate	DA	1	8	20000101
(0008,0021)	SeriesDate	DA	1	8	20000101
(0008,0022)	AcquisitionDate	DA	1	8	20000101
(0008,0023)	ContentDate	DA	1	8	20000101
(0008,0024)	RETIRED_OverlayDate	DA	1	8	20000101
(0008,0025)	RETIRED_CurveDate	DA	1	8	20000101
(0008,002A)	AcquisitionDateTime	DT	1	8	20000101
(0008,0030)	StudyTime	TM	0	0	
(0008,0032)	AcquisitionTime	TM	0	0	
(0008,0033)	ContentTime	TM	0	0	
(0008,0050)	AccessionNumber	SH	1	16	2819497684894126
(0008,0060)	Modality	CS	1	2	CT
(0008,0070)	Manufacturer	LO	1	18	GE MEDICAL SYSTEMS
(0008,0090)	ReferringPhysicianName	PN	0	0	
(0008,1090)	ManufacturerModelName	LO	1	16	LightSpeed Plus
(0008,1155)	ReferencedSOPInstanceUID	UI	1	64	1.3.6.1.4.1.14519.5.2.1.6279.6001.675906998158803995297223

شکل ۴-۲ اطلاعات سرآیند فایل DICOM مربوط به یک اسلایس تصویر سی تی ریه

به منظور ارزیابی تصاویر سی تی اسکن ریه، یک مقیاس کمی به نام هانسفیلد^۱ وجود دارد. برای درک بهتر این تصاویر، می توان آن محدوده ی مشخص از اعداد را بر روی تصویر نشان داد. این مقیاس توسط Godfrey Newbold Hounsfield یکی از مهندسين اصلی و مخترعين سی تی، ایجاد شد. شدت اشعه آب مقطر در دما و فشار استاندارد (STP)^۲ به عنوان صفر واحد هانسفیلد تعريف می شود. مقادير سی تی در بازه ی ۱۰۰۰ تا ۳۰۰۰ واحد هانسفیلد هستند. در جدول ، مقادير سی تی متناظر با بافت های مختلف در واحد هانسفیلد آورده شده است. در تصاویر، پیکسل های با اندازه ۱۰۰۰- هانسفیلد به صورت سیاه و صفر هانسفیلد به صورت سفید ظاهر می شوند.

جدول ۱-۲ مقادير سی تی متناظر با بافت های گوناگون در واحد هانسفیلد [۳۵]

نوع بافت	مقدار سی تی بر حسب هانسفیلد
هوا	-۱۰۰۰
ریه	-۶۰۰ تا -۴۰۰
ندول	-۱۵۰
چربی	-۱۰۰ تا -۶۰
آب	۰
بافت نرم	+۴۰ تا +۸۰
استخوان	+۴۰ تا +۱۰۰۰

¹ Hounsfield

² Standard Temperature and Ppressure

۲-۶ پیش پردازش: نیازمندی‌های بهبود تصاویر

از آنجایی که کار این پایان‌نامه بر مبنای دو عمل مهم پیش‌پردازش و استفاده از یادگیری عمیق است، در این قسمت به توضیحاتی در رابطه با فیلترهای ریخت‌شناسی ریاضی^۱ که از مهم‌ترین روش‌های پیش‌پردازش تصاویر جهت حذف نویز و بهبود لبه‌هاست، می‌پردازیم.

۲-۶-۱ فیلترهای ریخت‌شناسی ریاضی

هدف از فیلتر کردن تصویر، بهبود کیفیت آن با کاهش تاثیر نویز بر روی کیفیت تصویر تا حد ممکن است. معمولاً تصور می‌شود که اشیاء تصویر با سطح خاکستری یکسانی از مناطق هستند، بنابراین مرزها یا لبه‌های اشیاء در جایی که تغییرات شدید در سطح خاکستری و عملیات گرادیان رخ می‌دهد، قرار دارند تا این تغییرات را برجسته کنند. فیلتر مبتنی بر مورفولوژی ریاضی می‌تواند به اطلاعات ویژگی هندسی متکی شود و از عملگرهای مورفولوژیکی برای فیلتر کردن موثر موانع و نگهداری اطلاعات اصلی تصویر استفاده کند. عملیات مورفولوژیک، اطلاعات تصویر را ساده و به این صورت مشخصات شکل‌های تصویر را حفظ و اطلاعات اضافی آن را حذف می‌کنند. مورفولوژی ریاضی را به خصوص در علوم محیطی و مطالعه اکوسیستم‌ها، می‌توان تکنیک نسبتاً جدیدی دانست که بر اساس تئوری مجموعه‌ها و عملیات جبری مینکوفسکی استوار است. از عملیات مورفولوژیک معمولاً برای تشخیص لبه‌ها و رفع نویز تصویر استفاده می‌گردد [۳۶]. برای تشخیص لبه‌های تصاویر پزشکی، می‌بایست عناصر سازنده مناسب از لحاظ اندازه، شکل و جهت با توجه به ویژگی‌های بافتی تصویر انتخاب گردد. عملیات مورفولوژی را می‌توان به چهار دسته فرسایش^۲، گسترش^۳، باز کردن^۴ و بستن^۵ طبقه‌بندی کرد [۳۷].

^۱ Mathematical Morphology Filters

^۲ Erosion

^۳ Dilation

^۴ Opening

^۵ Closing

۲-۱-۶-۱ فرسایش و گسترش

اگر دو مجموعه A و B را در نظر بگیریم، عملگر فرسایش با نماد $A \ominus B$ نمایش داده می‌شود که در آن A ، نمایش تصویر ورودی و B نمایش عنصر ساختمانی است. به طور کلی فرسایش دو اثر مختلف بر تصاویر خاکستری می‌گذارد؛ اولاً اگر تمام درایه‌های عنصر ساختمانی مثبت باشد، تصویر خروجی تیره‌تر از تصویر ورودی است. ثانیاً اثر جزئیات روشن تصویر ورودی که سطحی کوچک‌تر از عنصر ساختمانی دارند، کاهش می‌یابد. به‌طور کلی در عملیات فرسایش، معمولاً نقاط ناخواسته تصویر بانیری حذف می‌شوند و سایر اجزای تصویر نیز به اندازه یک یا چند پیکسل نازک‌تر خواهند شد. در عمل، تمامی نقاط سفید تصویر بررسی شده در صورتی که حداقل یکی از همسایگان انتخابی آن سیاه باشد، آن نقطه نیز خواهد شد.

از طرف دیگر، گسترش برای مجموعه‌های A و B با نماد $A \oplus B$ نمایش داده می‌شود. عملیات گسترش یا افزایش عملیاتی است که باعث افزایش ابعاد اجزاء داخل تصویر به اندازه یک یا چند پیکسل می‌شود. در اثر این عمل ممکن است نقاطی که از یک تصویر بانیری در اثر عواملی چون تاثیر نویز یا اعمال حد آستانه نامطلوب جا افتاده است، تصحیح گردند. الگوریتم اعمال فیلتر افزایش بدین صورت است که تمامی نقاط سیاه تصویر بررسی شده، در صورتی که حداقل یکی از همسایگان انتخابی نقطه مورد بررسی سفید باشد، نقطه مزبور نیز سفید خواهد شد. در غیر این صورت سیاه باقی خواهد ماند.

۲-۱-۶-۲ باز کردن و بستن

از ترکیب‌های مختلف دو عمل فرسایش و گسترش، می‌توان عملیات دیگری ایجاد کرد. مهم‌ترین این عملیات، عملیات باز کردن و بستن است. عملگر باز کردن، تصویر را هموار می‌کند، اتصالات ضعیف را می‌شکند و بیرون‌زدگی‌های باریک را حذف می‌کند. عملگر باز کردن مجموعه A با عنصر ساختمانی B که به شکل $A \circ B$ نشان داده می‌شود، به صورت رابطه (۱-۲) قابل تعریف است:

$$A \circ B = (A \oplus B) \oplus B \quad (1-2)$$

با توجه به رابطه (۱-۲)، عملگر باز کردن ابتدا عمل فرسایش را روی تصویر اعمال می‌کند که این منجر به کاهش اندازه عناصر موجود در تصویر می‌شود و عناصری که کوچک‌تر از عنصر سازنده B باشند، حذف می‌شوند. سپس عملگر گسترش روی نتیجه حاصل‌شده اعمال می‌شود که این عمل در نهایت منجر به بازگشت اندازه عناصر باقی‌مانده در تصویر به اندازه اولیه عناصر می‌شود.

عملگر بستن نیز مشابه عملگر باز کردن، تصویر را هموار می‌کند. این عملگر، معمولاً انفصالات باریک و تورفتگی‌های باریک و بلند را پیوند می‌زند، حفره‌های کوچک را حذف می‌کند و انفصال‌های موجود در منحنی‌ها را پر می‌کند. عملگر بستن مجموعه A با عنصر ساختمانی B که به شکل $A \bullet B$ نشان داده می‌شود، به صورت رابطه (۲-۲) تعریف می‌شود:

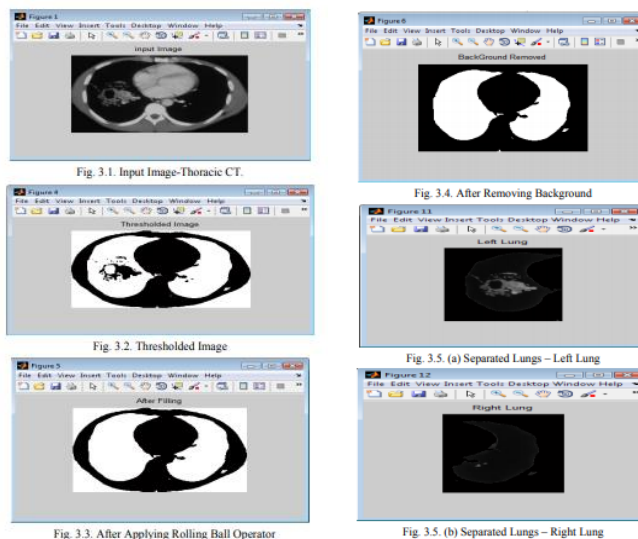
$$A \bullet B = (A \oplus B) \ominus B \quad (2-2)$$

با توجه به رابطه (۲-۲)، عملگر باز کردن ابتدا عمل گسترش را روی تصویر اعمال می‌کند. اعمال این عملگر، موجب حفره‌ها و انفصالات کوچک از تصویر حذف شوند و نیز اندازه عناصر موجود در تصویر بزرگ‌تر می‌شود. در نهایت با اعمال عملگر فرسایش روی تصویر حاصل، اندازه عناصر بزرگ‌شده در تصویر به اندازه اولیه‌شان برمی‌گردد.

در زیر به یک نمونه از موارد استفاده از عملگرهای مورفولوژی در پردازش تصاویر پزشکی در کارهای پیشین می‌پردازیم:

در [۳۸]، یک الگوریتم کاملاً خودکار برای جداسازی دو ریه چپ و راست از تصاویر سی‌تی اسکن ارائه شده‌است. در این مقاله، ابتدا جهت بخش‌بندی بافت ریه، یک الگوریتم خوشه‌بندی فازی به کار می‌رود. پس از بخش‌بندی، ممکن است به ندرت حفره‌هایی در نزدیکی مرز پارانشیم ریه ظاهر شوند، برای پر کردن این حفره‌ها و چسباندن قسمت‌های مختلف ریه، از عملگر مورفولوژی بستن استفاده

می‌شود. پس از آن، پس‌زمینه تصویر حذف شده و درنهایت با اعمال عملگر مورفولوژی بستن و سپس باز کردن بر تصویر حاصل از مرحله قبل، مرزهای نامنظم موجود صاف^۱ شده و دو ریه چپ و راست از تصویر جدا می‌شوند. در شکل، چگونگی استخراج ریه نمایش داده شده است



شکل ۲-۵ مثال از اعمال عملیات مورفولوژی در بخش‌بندی ریه (آخرین تصویر از ستون سمت چپ، بیانگر آشکار عملگر بستن و پرشدن حفره‌ها از تصویر قبلی آن است). [۳۸].

۲-۷ بخش‌بندی ریه

بخش‌بندی، به فرآیند تفکیک اجزای تصویر اطلاق می‌شود. به‌طور دقیق‌تر، قطعه‌بندی تصویر، فرآیندی است که تصویر را به بخش‌های معنی‌دار افراز می‌کند، به‌نحوی که هر بخش دارای ویژگی‌های بصری مشابه باشد [۳۹]. تحلیل و بررسی تصاویر پزشکی، مستلزم استخراج اطلاعات تصویر و نواحی مجزای هر تصویر است. از این‌رو قطعه‌بندی تصویر، یکی از مهم‌ترین مراحل در پردازش و تحلیل آن‌ها به‌شمار می‌رود. در تصاویر پزشکی، عمل بخش‌بندی برای استخراج ویژگی، اندازه‌گیری و نمایش تصاویر اهمیت دارد و در بسیاری از کاربردهای پردازش تصاویر پزشکی برای

¹ Smooth

جداسازی ارگان‌های بدن مانند مغز، قلب، ریه یا کبد در تصاویر سی‌تی یا ام‌آر‌آی جهت تحلیل و سپس تشخیص نواحی پاتولوژیکی مثل تومورها مفید است.

الگوریتم‌های بخش‌بندی در سیستم‌های CAD، به‌عنوان مرحله‌ی پیش‌نیاز برای انجام عملیات تحلیل محتوای تصویر و تشخیص الگوهای مورد جستجو در تصاویر پزشکی و نیز اولین گام در پردازش خودکار تصاویر جهت تحلیل و ارزیابی آن‌ها محسوب می‌شود؛ بنابراین عملکرد نهایی سیستم‌های CAD به دقت و حساسیت در مرحله بخش‌بندی بستگی دارد.

هدف اصلی در قطعه‌بندی تصاویر پزشکی، تقسیم‌بندی آن‌ها به ساختارهای آناتومیکی مختلف و سپس جدا نمودن اجزای مطلوب است. همان‌طور که می‌دانیم، ناحیه ریه از همه بافت‌های درون سطح ریه یعنی بافت اصلی ریه، رگ‌ها^۱، راه‌های هوایی و ساختارهای غیرطبیعی مانند ندول‌ها تشکیل شده‌است؛ لذا در قطعه‌بندی ریه، هدف جداکردن لوب‌های ریه از سایر نواحی تصویر سی‌تی بوده، چراکه به‌دست‌آوردن نواحی ریه از تصاویر سی‌تی که ندول‌های ریوی ممکن است در آن وجود داشته‌باشند، اولین و مهم‌ترین مرحله برای آشکارسازی^۲ و تشخیص^۳ این ندول‌ها به‌حساب می‌آید.

تعیین و شناسایی یک الگوریتم مناسب که بتواند با دقت بالا، ریه را قطعه‌بندی و مدل ریه را ایجاد کند، بسیار اهمیت دارد. هدف از تحقیقات و بررسی بر قطعه‌بندی تصاویر، افزایش اعتبار، دقت، صحت و کاهش هزینه‌های محاسباتی است. باتوجه به ساختار غیریکسان و پیچیده‌ی ریه در طول و عرض (شکل)، بخش‌بندی ناحیه ریه همیشه یک مسئله چالش‌برانگیز بوده‌است. به‌عبارت دیگر، بخش‌بندی ناحیه ریه، به‌دلیل عدم وجود هماهنگی در این ناحیه، همچون وجود ساختارهای با شدت مشابه مثل عروق خونی^۴، رگ‌ها، نایچه‌ها و ریزنایچه‌ها و نیز استفاده از وسایل تصویربرداری متفاوت با پروتکل‌های مختلف، به‌صورت مداوم به چالش کشیده می‌شود. بنابراین، ایجاد یک سیستم تشخیص خودکار

¹ Veins

² Detection

³ Recognition

⁴ Blood vessels

برای مکان‌یابی و بخش‌بندی ناحیه ریه، به‌منظور داشتن درک بهتر از انواع الگوهای بیماری‌های ریوی و تشخیص و درمان سریع‌تر آن، مسئله‌ای ضروری و حائز اهمیت است.

در مجموع به این نتیجه می‌رسیم که لازم است تا بافت اصلی ریه از پس‌زمینه و سایر نواحی غیر ریه جدا شود. مزیت این کار این است که علاوه بر محدود شدن ناحیه مورد بررسی و افزایش سرعت الگوریتم، میزان دقت نیز افزایش می‌یابد. هم‌چنین با حذف نواحی اضافه، امکان بروز خطا و تغییرات نویز در مراحل بعدی تحلیل و بررسی این تصاویر نیز، کاهش خواهد یافت.



شکل ۲-۶ تصاویری از برش‌های قسمت‌های مختلف ریه مربوط به یک بیمار

۸-۲ مروری بر روش‌های بخش‌بندی (کارهای پیشین)

تاکنون، روش‌های زیادی برای بخش‌بندی ناحیه ریه پیشنهاد شده‌است [۳۷]. روش‌های بخش‌بندی بسته به نوع کاربرد، خودکار بودن یا نیمه خودکار بودن و پارامترهای خاص دیگر، متنوع هستند. روش واحدی برای بخش‌بندی ریه از تصاویر سی‌تی وجود ندارد. بسیاری از محققان، سیستم‌های بخش‌بندی خودکار مختلفی را برای بخش‌بندی ناحیه پارانشیم^۱ ریه پیشنهاد کرده‌اند. برخی از روش‌های موجود، یا نیمه‌خودکار هستند و به‌دخالت عامل انسانی نیاز دارند یا دارای دقت مناسبی نیستند. به‌طور کلی، ما در این پایان‌نامه روش‌های قطعه‌بندی به‌کمک کامپیوتر را ابتدا به دو زیر دسته کلی روش‌های بخش‌بندی سنتی و روش‌های مبتنی بر یادگیری عمیق تقسیم کردیم. روش‌های بخش‌بندی سنتی در این فصل و روش‌های بخش‌بندی مبتنی بر یادگیری عمیق در فصل

^۱ Parenchyma

سوم، مفصل مورد بررسی قرار خواهند گرفت.

۲-۸-۱ روش‌های بخش‌بندی سنتی

در ابتدا به بررسی روش‌های سنتی بخش‌بندی می‌پردازیم. این روش‌ها خود به چهار زیردسته به تقسیم‌بندی می‌شوند که در ادامه نام برده خواهند شد.

- روش‌های مبتنی بر آستانه‌گذاری^۱
- روش‌های رشد ناحیه‌ای
- روش‌های مبتنی بر لبه
- روش‌های مبتنی بر مدل

در ادامه هر یک از این انواع روش‌ها را توضیح داده و به بررسی مزایا و معایبشان می‌پردازیم.

۲-۸-۱-۱ روش‌های مبتنی بر آستانه‌گذاری

در این روش، برچسب یک ناحیه به یک پیکسل، توسط مقایسه مقدار سطح خاکستری آن پیکسل با یک یا چند آستانه اختصاص داده می‌شود. از آنجا که حین تصویربرداری سی‌تی‌اسکن، ریه‌ها پر از هوا هستند، به‌صورت نواحی تیره در تصویر سی‌تی مشخص می‌شوند. بنابراین تباین بین ریه و بافت اطراف آن، اساس بسیاری از روش‌های بخش‌بندی ریه است. به‌همین دلیل در بسیاری از روش‌های بخش‌بندی، از آستانه‌گذاری استفاده می‌شود [۴۰-۴۷].

روش‌های مبتنی بر آستانه‌گذاری برای جداسازی بافت ریه، بر این اصل تکیه دارند که همواره روشنایی بافت ریه از سایر نواحی اطراف، کمتر است [۴۸]. با تکیه بر همین موضوع، با تعیین یک حد آستانه مناسب، اقدام به جداسازی ناحیه ریه می‌کند. در ادامه در مورد سه روش از مجموعه روش‌های آستانه‌گذاری که در مراجع قرار دادیم و بالاتر ذکر شد، توضیح مختصری می‌دهیم.

¹ Threshold Based

در مرجع [۴۶]، Gao و همکاران از روش مبتنی بر آستانه‌گذاری در چهار مرحله استفاده کردند. ابتدا راه‌های هوایی با روش رشد ناحیه (که بعداً توضیح داده می‌شود) حذف شده و سپس از یک آستانه بهینه برای حذف رگ‌ها استفاده می‌شود. در مرحله سوم، با استفاده از آستانه‌گذاری، ریه راست و چپ، جدا شده و سپس از عملگرهای مورفولوژی برای یک‌دست‌سازی مرز ریه استفاده شده است. هو و همکاران [۴۲]، یک روش تقسیم‌بندی مرحله به مرحله^۱ را پیشنهاد کردند که در آن، ابتدا یک آستانه تکراری برای به‌دست آوردن ناحیه تقسیم‌بندی اولیه استفاده می‌شود. سپس ناحیه به‌دست آمده توسط یک عملگر مورفولوژی بازکردن^۲ بسته می‌شود. یک مثال دیگر از این روش در مرجع [۴۹] آورده شده که با استفاده از تبدیل موجک^۳ و یک آستانه بهینه، اقدام به بخش‌بندی اولیه کرده، سپس ناحیه بخش‌بندی‌شده به دست‌آمده را با استفاده از عملیات مورفولوژی تصحیح می‌کند.

آستانه‌گذاری می‌تواند سراسری باشد هنگامی که یک آستانه ثابت روی کل تصویر اعمال گردد و یا می‌تواند متغیر (محلی یا پویا) باشد هنگامی که آستانه‌های متفاوتی برای نواحی متفاوت استفاده شود.

❖ آستانه‌گذاری سراسری^۴

روش‌های آستانه‌گذاری سراسری، سعی در یافتن تنها یک مقدار آستانه برای تمام پیکسل‌های تصویر دارند. یعنی هر پیکسل در پیش‌زمینه یا پس‌زمینه صفحه بر پایه‌ی مقدار خاکستری‌اش، علامت‌گذاری می‌شود [۵۱ و ۵۰]. روش‌های زیادی برای انتخاب آستانه سراسری وجود دارد. برای مثال یکی از این روش‌ها بر مبنای مدل دسته‌بندی است که احتمال خطا را حداقل می‌کند [۵۲]. یعنی اگر ما تصویری فقط شامل پیش‌زمینه و پس‌زمینه داشته باشیم، می‌توان با این روش خطا را به‌عنوان تعداد مجموع پیکسل‌هایی از پیش‌زمینه در نظر گرفت که به اشتباه به‌عنوان پس‌زمینه و پیکسل‌هایی از پس‌زمینه که اشتباهی به‌عنوان پیش‌زمینه، دسته‌بندی شدند محاسبه کنیم. یک نسخه

^۱ Stepwise

^۲ Opening Morphological Operation

^۳ Wavelet Transform

^۴ Global Thresholding

نیمه خودکار از این روش، توسط جانسون و همکارانش می‌کند [۵۳]، برای اندازه‌گیری توده‌های بطنی در تصاویر ام‌آرآی به کار برده شد. روش کاری آن‌ها به این شکل بود که یک عملگر، دو عدد پیکسل را از ناحیه پیش‌زمینه و پس‌زمینه انتخاب می‌کرد و مقدار آستانه به شکل خودکار وقتی که شدت‌های پیکسل در نواحی مدور اطراف پیکسل‌های انتخاب شده مقایسه شد، انتخاب می‌شد و این عمل، متناسب با حداقل تعداد پیکسل‌های اشتباه دسته‌بندی شده، بین دو توزیع است.

آستانه‌گذاری سراسری از لحاظ محاسباتی ساده و سریع است. این نوع آستانه‌گذاری روی تصاویری که شی‌های با مقادیر شدت یکنواخت روی کنتراست پیش‌زمینه دارد، خوب عمل می‌کند. با این وجود اگر یک کنتراست پایین بین پیش‌زمینه و شی وجود داشته باشد، اگر تصویر نویزی و یا اگر شدت پیش‌زمینه بطور عمده در سراسر تصویر تغییر کند، با شکست مواجه می‌شود.

❖ آستانه‌گذاری متغیر یا محلی^۱

آستانه‌گذاری سراسری را نمی‌توان برای موارد خاص، بسط داد. برای مثال، زمانی که پیش‌زمینه ثابت نیست و کنتراست اشیا در سراسر تصویر تغییر کنند، آستانه‌گیری ممکن است در یک بخش از تصویر خوب عمل کند اما ممکن است نتایج نامطلوبی را در نواحی دیگر تولید کند. راه حل این مسئله، استفاده از روش آستانه‌گذاری محلی است [۵۱ و ۵۴ و ۵۵].

آستانه‌گذاری محلی را می‌توان با دو روش مشخص کرد. روش اول این که تصویر را به چندین زیرتصویر تقسیم کنیم و برای هر یک از این زیرتصویرها، مقدار آستانه‌ای به صورت جدا از سایر نواحی انتخاب کنیم. نکته‌ی مورد توجه در این روش این است که حتماً زیرتصویرها باید هر کدام، به اندازه‌ای باشند که قابل تفکیک به دو بخش پیش‌زمینه و پس‌زمینه باشند. در روش دوم، مقدار آستانه با آزمایش کردن شدت‌های تصویر در همسایه هر پیکسل، قابل تعیین است. در [۵۶]، ابتدا یک تصویر را به زیرتصاویر مستطیل شکل که بایکدیگر هم‌پوشانی دارند تقسیم شده، سپس بافت‌نگار^۲هایی برای هر

^۱ Adaptive or Local Thresholding

^۲ Histogram

زیرتصویر محاسبه می‌شود. آنگاه اگر یک زیرتصویر، دارای بافت‌نگار دونماینه باشد (فقط شامل پیش‌زمینه و پس‌زمینه باشد)، مقدار حداقل بین قله‌های بافت‌نگار را به‌عنوان آستانه محلی در نظر می‌گیرد. به‌طور خلاصه، آستانه‌گذاری محلی از لحاظ محاسباتی هزینه‌برتر از روش سراسری است. این روش برای کاربردهایی مثل ناحیه‌بندی تصاویر از پیش‌زمینه‌های متفاوت، مفید است. به‌طور کلی روش‌های آستانه‌گذاری، اگرچه بسیار ساده و سریع هستند و زمانی که تصویر دارای نویز و ناهنجاری‌های ریه نباشد، به خوبی عمل می‌کنند. اما چند ایراد اساسی به آن‌ها وارد است:

- (۱) قادر به حذف موثر نای و نایژه‌ها نیست.

- (۲) قادر به حفظ ندول‌های ریوی متصل به دیواره ریه نیست.

- (۳) دشواری و خطا در یافتن یک حد آستانه مناسب به‌علت وجود سطوح خاکستری نزدیک به هم و مرزهای بسیار [۵۷].

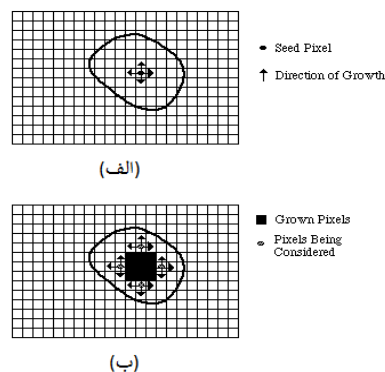
۲-۱-۸-۲ روش‌های رشد ناحیه‌ای

در حالی که روش آستانه‌گذاری روی تفاوت شدت‌های پیکسل تمرکز دارد، روش‌های رشد ناحیه به دنبال گروهی از پیکسل‌ها با شدت‌های مشابه هستند. بخش‌بندی با استفاده از روش‌های رشد ناحیه، روالی است که پیکسل‌ها یا زیرناحیه‌ها را در نواحی بزرگتر گروه‌بندی می‌کند. ساده‌ترین این راه‌ها پیوستن پیکسل‌ها بر اساس یک سری خواص مشابه است. درواقع اساس کار این روش‌ها، دنبال کردن نقطه‌کار^۱ و افزایش تدریجی ناحیه بخش‌بندی‌شده مبتنی بر مقایسه پیکسل‌های مجاور بر اساس معیارهای از پیش تعیین‌شده است. این نقطه‌کار معمولاً به صورت دستی تعیین می‌شود.

نخستین گام برای شروع بخش‌بندی تصاویر به کمک الگوریتم‌های رشد ناحیه‌ای، تعیین یک یا چند نقطه‌کار روی تصویر به‌عنوان بذر یا دانه اولیه است. در مرحله بعدی، پیکسل‌های همسایه برای اضافه شدن به ناحیه‌ی رشد، آزمایش شباهت می‌شوند و آن پیکسل‌هایی از تصویر که بیشترین شباهت به

¹ Seed Point

دانه را دارند تعیین می‌شوند و اگر شرایط را برآورده سازد شامل ناحیه‌ی رشد مورد نظر می‌شود. این الگوریتم تا زمانی ادامه می‌یابد که دیگر هیچ پیکسلی برای اضافه کردن موجود نباشد. در نهایت ناحیه بخش‌بندی‌شده‌ی مطلوب، توسط تمام پیکسل‌هایی که در این رویه رشد پذیرفته شده‌اند، نمایش داده می‌شود [۵۹و۵۸]. همان‌طور که در شکل، مشاهده می‌شود، اول از یک پیکسل دانه، شروع شده و تمام نقاطی که در همسایگی این پیکسل بوده و بیشترین شباهت را به آن دارند، به تدریج به ناحیه‌ی مورد نظر افزوده شده و این ناحیه کم‌کم بزرگ‌تر می‌شود. در نهایت، پیکسل‌هایی که شباهت ظاهری خیلی زیادی به هم دارند، جزء یک مجموعه قرار خواهند گرفت.



شکل ۲-۷ شماتیک روش‌های رشد ناحیه‌ای [۶۰]. (الف): انتخاب بذر اولیه (ب): بررسی نقاط مجاور و انجام فرآیند بخش‌بندی.

نقاط با علامت ضربدر، نشان‌دهنده قسمت جدا شده هستند.

روش‌های رشد ناحیه‌ای، در بسیاری از کاربردهای پزشکی مثل بخش‌بندی حفره‌ها در تصاویر کاردیاک [۶۱]، استخراج رگ‌های خونی در داده‌های آنژیوگرافی [۶۲]، بخش‌بندی کلیه [۶۳]، استخراج سطح مغز [۶۴] و بخش‌بندی تصاویر سی‌تی‌اسکن ریه [۶۵-۶۸] استفاده می‌شود. از آنجا که کار این تحقیق، در رابطه با بخش‌بندی تصاویر ریه است، دو نمونه از کارهای انجام‌شده در این حیطه را مورد بررسی قرار خواهیم داد.

نهاد^۱ و هم‌کاران در [۶۵] یک روش کاملاً خودکار برای استخراج ریه در تصاویر سی‌تی ریه به کمک روش رشد ناحیه استفاده کردند. روش پیشنهادی آن‌ها بدین صورت است که ابتدا تصاویر را خاکستری کرده، سپس الگوریتم آستانه‌گذاری بهینه روی تصاویر سی‌تی ریه اعمال می‌شود تا ماسک‌های اولیه تولید شوند. در مرحله بعد، الگوریتم رشد ناحیه با پیکسل دانه و پیکسل‌های دیگر که دانه را احاطه کرده‌اند شروع می‌شود. سپس پیکسل‌هایی که بیشترین شباهت به پیکسل دانه را دارند تعیین شده و چنانچه شرایط مورد نیاز مسئله را برآورده سازند، شامل ناحیه‌ی موردنظر خواهند شد. این الگوریتم تا زمانی ادامه می‌یابد که پیکسلی برای اضافه کردن موجود نباشد. ناحیه‌ی ایجاد شده، مکرراً توسط بررسی تمام غیرهمسایگی ناحیه رشد می‌کند. این فرآیند تا زمانی پایان می‌یابد که تفاوت شدت بین میانگین ناحیه و پیکسل جدید بیشتر از آستانه معین باشد. پس از اتمام روند الگوریتم رشد ناحیه، تشخیص‌گر لبه بر روی تصویر اعمال می‌شود تا یک ماسک برای استخراج ریه ساخته شود. در نهایت تصویر نهایی را با ترکیب ماسک باینری لبه و تصویر اولیه به دست می‌آورند. مشکلی که در انتهای عملیات استخراج ریه شکل می‌گیرد، این است که در بخش‌بندی ندول‌های متصل به دیواره ریه شکست می‌خورد. شکل ، این مسئله را نشان می‌دهد.

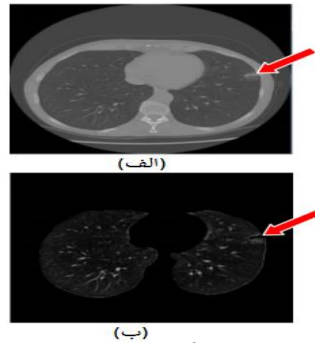
در [۶۶]، الگوریتمی پنج مرحله‌ای برپایه رشد ناحیه استفاده شده است. در مرحله اول این الگوریتم، یک پیکسل اولیه در نظر گرفته می‌شود. سپس طی دو مرحله‌ی بخش‌بندی انطباقی محلی^۲ و پرکردن حفره، پیش‌زمینه و پس‌زمینه از هم جدا می‌شوند. در مرحله بعدی، ابتدا اشیای پیش‌زمینه غیرمتصل حذف شده و سپس ناحیه پیش‌زمینه بسط داده شده و دوبرابر می‌شود. ماسک برای به دست آوردن پارامترهای لازم برای مراحل بعدی که شامل دو الگوریتم رشد ناحیه است استفاده می‌شود و همچنین به عنوان محدودکننده حجم برای ناحیه رو به رشد استفاده می‌شود. ماسک نیز برای به دست آوردن نقطه بذر مطلوب استفاده می‌شود. در مرحله چهارم نقشه اتصال فازی به وسیله روش رشد ناحیه اتصال فازی ایجاد می‌شود. فرآیند رشد تا زمانی ادامه می‌یابد که ماسک کامل که از مرحله قبل

¹ Nihad

² Local Adaptive Segmentation

به دست آمده ایجاد شود. در مرحله آخر، ناحیه رشد یافته کروی شکل بر روی نقشه فازی به منظور

ایجاد نتایج بخش بندی چندتایی تطبیق داده می شود.



شکل ۲-۸ بخش بندی ریه به روش رشد ناحیه [۳۵]. (الف) : تصویر سی تی اولیه (ب) : تصویر قطعه بندی شده به روش رشد ناحیه (فلش های موجود در دو تصویر نشان داده شده، توده چسبیده به دیواره ریه است که در روش رشد ناحیه، در نظر گرفته نشده.)

روش های رشد ناحیه ای، شامل مزایای زیر هستند :

- پیاده سازی آن ها ساده است.
- زمانی که لبه های تصویر واضح باشند، دقت نسبتا بالایی دارند.
- می توانند از چند معیار به طور هم زمان برای ناحیه بندی استفاده کنند.
- تا حد زیادی در برابر نویز مقاوم هستند.

هم چنین از معایب این روش ها، می توان به موارد زیر اشاره کرد :

- ناتوانی در بخش بندی ندول های متصل به دیواره ریه (شکل)
- پیچیدگی محاسباتی زیاد و زمان بر بودن
- وابستگی عملکرد الگوریتم به بذرهای اولیه

۲-۸-۱-۳ روش‌های مبتنی بر لبه

در این دسته از روش‌های قطعه‌بندی، تشخیص لبه^۱ برای بخش‌بندی ناحیه مورد نظر در تصویر استفاده می‌شود. لبه تصویر، نقاطی از تصویر است که در آن شدت روشنایی به طور ناگهانی تغییر کرده است. از آنجایی که قسمت‌های مختلف یک تصویر، توسط لبه از هم جدا می‌شوند، لذا با استخراج لبه‌های تصویر، می‌توان آن را قطعه‌بندی کرد. روش‌های تشخیص لبه، خود به دو زیردسته تقسیم می‌شوند: روش‌های تشخیص لبه در حوزه مکان و روش‌های تشخیص لبه در حوزه تبدیل. برخی از روش‌های تشخیص لبه در حوزه مکان عبارتند از: سوبل^۲، کنی^۳، رابرتز^۴، پرویت^۵ و لاپلاس^۶. ما تنها دو الگوریتم کاربردی و مهم سوبل و کنی را در ادامه توضیح می‌دهیم.

- لبه‌یاب سوبل: این روش، لبه‌ها را با استفاده از تخمین مشتق، آشکار می‌کند. درواقع در این روش، دو ماسک 3×3 برای آشکارسازی لبه‌های افقی و عمودی بر روی تصویر اعمال می‌شود [۶۹]. این ماسک‌ها در شکل نشان داده شده‌اند.

۱-	۰	۱-
-۲	۰	۲
-۱	۰	۱

Sx

۱-	۲-	۱-
۰	۰	۰
۱	۲	۱

Sy

شکل ۲-۹ ماسک‌های لبه‌یاب سوبل (Sx، ماسک سوبل عمودی و Sy، ماسک سوبل افقی است)

- لبه‌یاب کنی: لبه‌یاب کنی، توسط جان اف کنی^۷ در سال ۱۹۸۶ معرفی شد [۷۰] و هنوز یک لبه‌یاب استاندارد و با دقت و کیفیت بالا می‌باشد. لبه‌یاب کنی، از مهم‌ترین الگوریتم‌های لبه‌یاب است که از مشتق مرتبه اول استفاده می‌کند. این الگوریتم، از چهار مرحله اصلی به صورت زیر تشکیل شده است [۷۱]:

¹ Edge Detection

² Sobel

³ Canny

⁴ Roberts

⁵ Prewitt

⁶ Laplace

⁷ John F. Canny

(۱) حذف نویز از تصویر: در این مرحله از یک فیلتر گاوسین که شامل یک ماسک ساده برای حذف نویز است، استفاده می‌شود.

(۲) یافتن لبه‌های قوی: در این مرحله، لبه‌های تصویر با گرفتن گرادیان و با استفاده از عملگر سوبل به دست می‌آیند. سپس لبه‌های مات یافت شده به لبه‌های تیزتر تبدیل می‌شوند.

(۳) سرکوب کردن به حداکثر اندازه^۱: این مرحله، جهت همه لبه‌های یافت‌شده را بررسی می‌کند و آن‌هایی که به اندازه کافی شبیه به لبه نیستند، حذف می‌کند.

(۴) آستانه‌گیری دوتایی در تشخیص لبه: برای تشخیص لبه، از روشی به نام hysteresis استفاده می‌شود. به این منظور، دو آستانه بالا و پایین تعریف می‌شود. سپس هر پیکسل که دارای میزان گرادیان بیشتر از حد بالا باشد به عنوان لبه پذیرفته می‌شود و در صورتی که دارای مقدار کمتر از حد پایین باشد رد خواهد شد. چنانچه دارای مقداری بین این دو حد باشد، در صورتی پذیرفته می‌شود که یکی از همسایه‌های آن پذیرفته شده باشد.

این الگوریتم، بیشتر به کشف لبه‌های ضعیف به درستی می‌پردازد و کمتر فریب نویز را می‌خورد و از بقیه روش‌ها بهتر است. شین و هم‌کاران [۷۲]، بعد از انجام یک ارزیابی از عملکرد الگوریتم‌های مختلف تشخیص لبه، به این نتیجه رسیدند که لبه‌یاب کنی دارای بهترین عملکرد، بهترین هم‌گرایی و سریع‌ترین زمان اجرا است. هم‌چنین معیارهایی که در لبه‌یاب کنی مطرح است، در زیر اشاره شده‌اند [۷۳]:

(۱) پایین آوردن نرخ خطا: یعنی تا حد امکان هیچ لبه‌ای در تصویر نباید از دست برود و هم‌چنین هیچ چیزی که لبه نیست نباید به جای لبه فرض شود.

(۲) لبه در مکان واقعی خود باشد: یعنی تا حد ممکن لبه‌ها کمترین فاصله را با مکان واقعی خود داشته باشند.

(۳) برای هر لبه فقط یک پاسخ داشته باشیم.

¹ Non-maximum suppression

(۴) لبه ها کمترین ضخامت را داشته باشند (در صورت امکان یک پیکسل).

مندونکا و همکاران [۷۴]، از یک الگوریتم لبه یاب فضایی برای بخش بندی ناحیه ریه از تصاویر رادیوگرافی قفسه سینه استفاده کردند. آن ها در روش خود، ابتدا نواحی مورد نظر (ROI)^۱ را به صورت صورت مستطیل شکل برای هر دو ریه چپ و راست جدا کردند. سپس با استفاده از الگوریتم های لبه یابی، لبه های صاف، فوقانی و پایینی آشکار شده و در نهایت با ترکیبی از لبه های باقی مانده، عمل بخش بندی، به طور موثر انجام شد. آن ها در مقاله خود به میزان حساسیت^۲ برابر با ۰,۹۲۲۵ و نیز به ارزش اخباری مثبت^۳ برابر ۰,۹۶۸ دست یافتند.

در مرجع [۷۵]، عمل لبه یابی را بر روی ۴۸۷ تصویر رادیوگرافی قفسه سینه اعمال کردند. آن ها با استفاده از مشتق مرتبه اول فیلتر گاوسی در چهار جهت مختلف و سپس لبه یاب لاپلاسیان گاوسی (LOG)^۴ در سه مقیاس مختلف برای ردیابی لبه های مرز ریه استفاده کردند و در نهایت به دقت ۹۴,۳۷ رسیدند.

مزایای روش های مبتنی بر لبه عبارتند از [۷۶]:

- (۱) به خوبی بر روی پس زمینه و اشیا تصویر کار می کنند.
 - (۲) از آنجایی که عمل بخش بندی در این روش ها شباهت زیادی به طبقه بندی انسانی دارد، جزء روش های سطح بالای بخش بندی محسوب می شوند.
 - (۳) ساختار اشیای طبقه بندی شده را تا حدود زیادی حفظ می کنند.
- با وجود مزایای مطرح شده، باید گفت دو ایراد اصلی این روش ها این است که اولاً، به کارایی الگوریتم لبه یاب بسیار حساس هستند. ثانیاً، وقتی مرزهای ضعیفی در تصویر وجود داشته باشد، یا نواحی تصویر دارای بافت متراکم باشند، یا تصویر نویزی باشد، این روش ها کارایی مطلوبی ندارند [۷۷].

^۱ Region of Interest

^۲ Sensitivity

^۳ positive predictive value

^۴ Laplacian of Gaussian

۲-۸-۱-۴ روش‌های مبتنی بر مدل

روش‌های مبتنی بر مدل، با استفاده از مدل‌های مشخص، ساختار ریه را استخراج می‌کنند. روش‌های مبتنی بر مدل به مدل‌های دگر دیس^۱ یا شکل‌پذیر تقسیم‌بندی می‌شوند. این مدل‌ها، منحنی‌ها یا سطوح تعریف شده در داخل تصویر هستند که می‌توانند تحت تاثیر چند نیرو حرکت کنند. این نیروها طوری تعریف می‌شوند که منحنی را به سمت نواحی مطلوب تصویر، مثل لبه‌ها سوق دهند. ایده اولیه استفاده از مدل‌های دگر دیس، اولین بار توسط کاس و هم‌کارانش در سال ۱۹۸۸ داده شد [۷۸]. این مدل با نام‌های کانتور فعال^۲ و مدل مار^۳ نیز استفاده می‌شود. در واقع بهتر است بگوییم روش‌های دگر دیس، می‌توانند بر اساس این که منحنی‌ها و سطوح چگونه تعریف شوند، به دو زیر دسته تقسیم شوند: مدل پارامتری (صریح)^۴ که کانتور فعال نیز نامیده می‌شود و مدل غیر پارامتری (ضمنی)^۵ که مجموعه سطح^۶ یا کانتور فعال هندسی^۷ نیز نامیده می‌شود.

• مدل‌های دگر دیس پارامتری یا کانتور فعال

کانتور فعال (مار)، خمی است مقید به پیوستگی و همواری و در عین حال توسط یک میدان نیروی خارجی به طرف نقاط مرزی و لبه‌های تصویر کشیده می‌شود. اساس کار کانتورهای فعال، تعریف منحنی‌هایی در داخل تصویر است. در واقع، این روش مرز بافت‌های تصویر را از طریق منحنی‌های پارامتری تشخیص می‌دهد. این منحنی‌های مارپیچی، نقاط کنترلی به هم پیوسته‌ای هستند که با هم کانتور را ایجاد می‌کنند. به ازای هر منحنی، یک تابع انرژی تعریف می‌شود که با توجه به نیروهای داخلی، خارجی و محدودیت (که به مجموع این نیروها، انرژی مار (snake) گفته می‌شود [۷۹])، این تابع انرژی کم و زیاد می‌شود. این نیروها باید به گونه‌ای انتخاب شوند تا منحنی به سمت لبه‌ها

¹ Deformable Models

² Active Contour

³ Snake Model

⁴ Explicit

⁵ Implicit

⁶ Level Set

⁷ Geometric active contour

هدایت گردد. به بیان بهتر، در این مدل، مرز ریه با تکامل نیروهای هدایتی داخلی و خارجی مناسب برای شکل دادن به ریه مشخص می‌شود [۸۰]. درواقع ایده اصلی این روش، تغییر شکل کانتور برای کمینه شدن تابع انرژی کانتور است. باتوجه تابع انرژی تعریف شده، روش‌های مختلفی برای کانتورهای فعال وجود دارد. به کمک دو نیروی داخلی و خارجی عملیات بخش‌بندی انجام می‌شود. در کانتورهای فعال، نیروهای داخلی از داخل منحنی یا سطح بسته محاسبه می‌شوند و در طی جریان تغییر شکل، نرمی سطح یا منحنی را حفظ می‌کنند. نیروهای خارجی ناشی از تصویر بوده و منحنی یا سطح بسته را به سمت ناحیه موردنظر می‌برند. یوشینوری ایتای و هم‌کارانش [۸۱]، از مدل کانتور فعال برای بخش‌بندی نواحی ریه استفاده کردند. در [۸۲]، الگوریتمی که از کانتور فعال الهام گرفته و از دانش‌های قبلی از ساختار غدد استفاده می‌کند، ارائه شده‌است که نتایج خوبی هم در تشخیص ساختار غدد به دست آورد.

مزیت اصلی کانتورهای فعال، سهولت در پیاده‌سازی و بار محاسباتی کم آن‌ها است. از طرفی این روش هم دارای معایبی است که در زیر به آن‌ها اشاره شده‌است:

- ۱) ساخت کانتور اولیه که اغلب نیاز به تعامل انسان دارد [۴۸].
- ۲) ضرورت انتخاب محل اولیه کانتور نزدیک به لبه‌های اصلی
- ۳) هم‌گرایی اندک در لبه‌های مقعر^۱
- ۴) ناتوانی نیروهای خارجی مانند نیروهای مبتنی بر لبه برای در نظرگرفتن نواحی غیریکنواخت ریه
- ۵) انرژی داخلی در مقابل نویز تصویری، با موفقیت عمل نمی‌کند.

• مدل‌های دگر دیس غیر پارامتری یا کانتور فعال هندسی

اگرچه مدل‌های دگر دیس پارامتری در دامنه وسیعی از کاربردها به طور موفقیت‌آمیزی به کار گرفته شده‌است، با این وجود دو محدودیت اصلی در این مدل‌ها وجود دارد: اول، در مواقعی که

¹ Concave

مدل اولیه با مرزهای مطلوب در شکل و اندازه، خیلی متفاوت باشد، این مدل باید مجدداً به شکل پویا پارامتربندی شود تا مرزهای شی را به درستی در برگیرد. محدودیت دوم، بروز مشکل در مواجه شدن با سازگاری‌های توپولوژیکی مانند تقسیم و ادغام بخش‌های مدل، انتخاب یک مشخصه مفید برای دربرگرفتن شی‌های چندگانه^۱ یا توپولوژی ناشناخته می‌باشد [۸۳].

روش دیگر مدل‌های دگردیس، روش‌های غیرپارامتری یا هندسی هستند. مدل‌های دگردیس هندسی، تقریباً هم‌زمان توسط کسل^۲ [۸۴] و ملدی^۳ [۸۵] پیشنهاد شد. مدل هندسی برای کانتورهای فعال، بر اساس تئوری فلو هندسی تکامل منحنی^۴ است. این مدل‌ها معمولاً از الگوریتم‌های عددی مجموعه سطح استفاده می‌کنند. در سال ۱۹۸۸، اوشر^۵ و ستیان^۶ برای اولین بار نمایش روش مجموعه سطح به نام روش سطح‌گذاری را پیشنهاد کردند [۸۶]. در روش مجموعه سطح، مسئله بخش‌بندی ناحیه در یک بعد بالاتر انجام می‌شود. درواقع ایده اصلی روش‌های مجموعه سطح، انتشار دادن منحنی‌هایی به عنوان سطح صفر یک تابع با ابعاد بالاتر است که در سیستم مختصات اولیه بیان شده‌اند.

فرض کنیم یک تصویر دوبعدی داشته باشیم؛ بنابراین منحنی موردنظر در دو بعد به صورت $C(x,y)$ بیان می‌شود. روش مجموعه سطح به جای در نظر گرفتن حرکت منحنی در دو بعد، مسئله را به یک بعد بالاتر می‌برد و منحنی $C(x,y)$ را به عنوان سطح صفر یک تابع سه‌بعدی در نظر می‌گیرد [۸۷]. یعنی یک تابع ضمنی در سه بعد تعریف می‌شود (رابطه (۳-۲)) که دامنه‌اش x و y است و هر نقطه مقدار z را به خود می‌گیرد.

$$Z = \Psi(x,y) : R^2 \rightarrow R \quad (3-2)$$

بنابراین در این روش، یک رویه به منحنی کانتور مورد بحث، نظیر می‌شود. بعد در هر مرحله از تکرار الگوریتم، رویه به نحوی جابه‌جا می‌شود که سطح صفر این رویه هدایت شود تا در نهایت به حد

¹ Multiple objects

² Caselles

³ Malladi

⁴ Curve evolution geometric flow

⁵ Osher

⁶ Sethian

قابل قبولی بر مرزهای هدف مورد نظر، منطبق شود. در این حالت رویه‌ی منطبق‌شده، به‌عنوان منحنی کانتور جدید به‌دست می‌آید. از این‌رو حرکت کردن به جلو به‌طور ضمنی، تحت کنترل تابع مجموعه سطح قرار می‌گیرد.

از جمله مزایای این روش می‌توان به موارد زیر اشاره کرد :

- می‌تواند برای کاربردهای پیچیده همانند گوشه‌های تیز که توپولوژی تغییر می‌کند، به کار گرفته شود.

به‌آسانی قابل ارتقا به ابعاد بالاتر است.

با توجه به این که تابع مجموعه سطح در سیستم مختصات اویلری بیان شده‌است، گریدهای گسسته می‌توانند با روش‌های متفاوت دیگر استفاده شوند تا یک تخمین عددی برای این راه حل به‌دست‌آید.

تا اینجا، به بررسی تعدادی از روش‌های بخش‌بندی سنتی پرداختیم. جدول ، روش‌های شرح داده‌شده در مقاله‌های بیان‌شده را به طور خلاصه مقایسه می‌کند. در این بخش، روش‌های سنتی بخش‌بندی را بسته و در بخش بعدی، به‌صورت خلاصه، مقدمه‌ای بر یادگیری عمیق بیان کرده و در بخش آخر نیز، جمع‌بندی بر کل مطالب بیان‌شده در این فصل خواهیم‌داشت.

۲-۹ جمع‌بندی روش‌های سنتی بخش‌بندی

تا اینجا، به بررسی تعدادی از روش‌های بخش‌بندی سنتی پرداختیم. جدول ، روش‌های شرح داده‌شده در مقاله‌های بیان‌شده را به طور خلاصه مقایسه می‌کند. در این بخش، روش‌های سنتی بخش‌بندی را بسته و در بخش بعدی، به‌صورت خلاصه، مقدمه‌ای بر یادگیری عمیق بیان کرده و در بخش آخر نیز، جمع‌بندی بر کل مطالب بیان‌شده در این فصل خواهیم‌داشت.

جدول ۲-۲ مقایسه روش‌های سنتی مطرح‌شده برای بخش‌بندی

نام روش - شماره مراجع	مزایا	معایب
آستانه‌گذاری - [۴۹و۴۶و۴۲]	- سادگی در محاسبات - سرعت بالا - حافظه کم	- عدم توانایی در حذف نای و نایژه - عدم توانایی در بخش‌بندی ندول‌های چسبیده به دیواره ریه - دشواری در یافتن حد آستانه مطلوب
رشد ناحیه‌ای - [۶۶و۶۵]	- پیاده‌سازی ساده - استفاده هم‌زمان از چند معیار برای ناحیه‌بندی - مقاومت بالا در برابر نویز	- عدم توانایی در بخش‌بندی ندول‌های چسبیده به دیواره ریه - پیچیدگی محاسباتی زیاد و زمان‌بری بالا - وابستگی عملکرد الگوریتم به بذرهای اولیه
مبتنی بر لبه - [۷۵و۷۴]	- شباهت زیاد به روش طبقه‌بندی انسانی - حفظ زیاد اشیای بخش‌بندی شده	- وابستگی عملکرد الگوریتم کلی به الگوریتم لبه‌یاب - مقاومت کم در برابر نویز - شکست هنگام برخورد با بافت‌های متراکم تصویر
مبتنی بر مدل - [۸۲و۸۱]	- پیاده‌سازی ساده - بار محاسباتی کم - قابل ارتقا به ابعاد بالاتر	- نیاز به انسان جهت ساخت کانتور اولیه - مقاومت کم در برابر نویز - هم‌گرایی کم در برابر لبه‌های مقعر

۲-۱۰ مقدمه‌ای بر یادگیری عمیق

یادگیری عمیق نگرشی جدید به ایده‌ی شبکه عصبی است و از کارکرد مغز انسان الهام گرفته‌شده‌است. رویکردهای یادگیری عمیق قادر به پردازش داده‌های طبیعی به شکل خام خود

هستند. در نتیجه، نیاز به ویژگی‌های دستی^۱ را مرتفع کرده‌است [۸۸]. امروزه به دلیل وجود واحدهای پردازش مرکزی سریع‌تر (CPUs)^۲ و واحدهای پردازش گرافیکی (GPUs)^۳ که تا حد زیادی زمان آموزش و اجرا را کاهش داده و موجب تسهیل در روند دسترسی به مجموعه‌های زیادی از داده‌ها شده‌اند، شاهد کاربرد فزاینده رویکردهای یادگیری عمیق در تمامی عرصه‌ها مانند پردازش تصویر، پردازش گفتار، رباتیک، داده‌کاوی و غیره هستیم. به دلیل توانایی الگوریتم‌های یادگیری عمیق در یادگیری و پردازش مقادیر زیادی از داده‌ها با سرعت و دقت بالا، امروزه به‌طور خاص، یادگیری عمیق در حال تبدیل شدن سریع به لبه تکنولوژی^۴ برای پردازش تصاویر پزشکی است [۸۹].

شبکه‌های عمیق را می‌توان به چهار ساختار مهم تقسیم کرد [۹۰]:

- شبکه‌های عصبی کانولوشنی^۵

- شبکه‌های از پیش آموزش دیده بدون ناظر^۶

- شبکه‌های عصبی مکرر^۷

- شبکه‌های عصبی بازگشتی^۸

در فصل سوم شبکه‌های عصبی کانولوشنی و از پیش آموزش دیده را مورد بررسی قرار می‌دهیم.

۲-۱۱ جمع بندی

در این فصل، مفاهیم مورد نیاز در این پایان نامه را تعریف کرده، سپس به بررسی کارهای گذشته انجام شده در حیطه بخش بندی تصاویر پزشکی پرداختیم. دیدیم که به طور کلی، روش‌های بخش بندی تصاویر پزشکی به دو دسته سنتی و مبتنی بر یادگیری عمیق تقسیم می‌شوند. ابتدا به تشریح روش‌های سنتی پرداخته، دریافتیم که اشکال مهم این روش‌ها، اغلب نیاز به عامل انسانی یا عدم

¹ Handcrafted Features

² CPUs: Central Processing Units

³ GPUs: Graphics Processing Units

⁴ State-of-the-art

⁵ Convolutional neural networks (CNNs)

⁶ Unsupervised Pretrained Networks

⁷ Recurrent Neural Networks (RNNs)

⁸ Recursive Neural Network

توانایی در بخش‌بندی ندول‌های چسبیده به دیواره ریه است. بعد از آن، مقدمه‌ای بر یادگیری عمیق ارائه کردیم و دیدیم که یادگیری عمیق، نیاز به ویژگی‌های دستی را مرتفع کرده‌است. در فصل سوم، به توضیح چند تا از معماری‌های یادگیری عمیق پرداخته و روش‌های مبتنی بر آن را برای بخش‌بندی تصاویر، تشریح خواهیم کرد.

فصل ۳ : ساختارهای یادگیری عمیق

۳-۱ مقدمه

در این فصل، ابتدا مهم‌ترین و کاربردی‌ترین ساختار یادگیری عمیق یعنی شبکه‌های عصبی کانولوشنی را مطرح کرده و سپس تعدادی از معماری‌های مبتنی بر این ساختار را مورد بررسی قرار خواهیم داد. در انتها نیز مروری بر روش‌های بخش‌بندی تصاویر پزشکی با معماری‌های مطرح‌شده در یادگیری عمیق خواهیم داشت.

۳-۲ شبکه‌های عصبی کانولوشنی

از بین تمام ساختارهای یادگیری عمیق، شبکه‌های عصبی کانولوشن یکی از مهم‌ترین و پرکاربردترین ساختارها هستند که در آن‌ها چندین لایه به‌شکل سلسله‌مراتبی با روشی قدرتمند آموزش می‌بینند. از سال ۲۰۱۲، شبکه‌های کانولوشنی عمیق در بسیاری از مسائل پردازش تصویر از سایر الگوریتم‌های هوش مصنوعی، سبقت گرفته‌اند. اولین موفقیت قطعی این شبکه‌ها در مسابقه ImageNet توسط هینتون و کریژوسکی^۱ [۹۱]، از طریق آموزش یک شبکه هشت لایه نظارت‌شده بر روی مجموعه‌داده عظیم ImageNet به همراه یک میلیون داده تصویری و تنظیم میلیون‌ها پارامتر ناشی از آن به‌دست‌آمد. کاربرد اصلی این شبکه‌ها در مسائل طبقه‌بندی است. به‌طور کلی، یک شبکه CNN از سه لایه اصلی تشکیل می‌شود که عبارتند از: لایه کانولوشن^۲، لایه تجمیع^۳ و لایه تماماً متصل^۴ [۹۲].

لایه‌های مختلف وظایف مختلفی را انجام می‌دهند. در شکل، یک معماری کلی از شبکه عصبی کانولوشن برای دسته‌بندی تصاویر به‌صورت لایه‌به‌لایه نمایش داده شده‌است. در هر شبکه عصبی کانولوشن، دو مرحله برای آموزش وجود دارد. مرحله تغذیه رو به جلو^۵ و مرحله پس انتشار^۶. در مرحله

^۱ Hinton and Krizhevsky

^۲ Convolution Layer

^۳ Pooling Layer

^۴ Fully Connected Layer

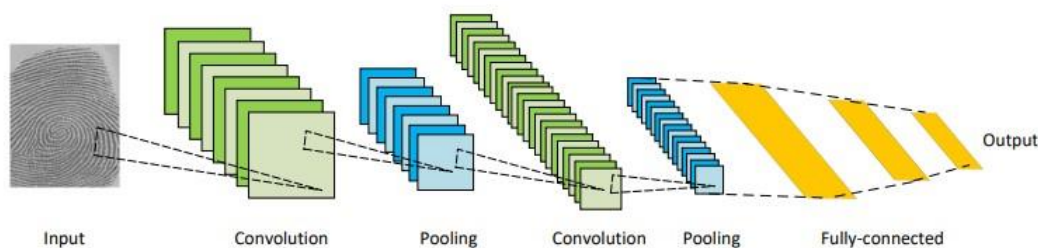
^۵ Feed Forward

^۶ Back Propagation

اول، تصویر ورودی به شبکه تغذیه می‌شود. در این مرحله، ضرب نقطه‌ای بین ورودی و پارامترهای هر نورون انجام شده و پس از اعمال عملیات کانولوشن در هر لایه، خروجی شبکه محاسبه می‌شود. در همین مرحله برای آموزش شبکه، خروجی شبکه را با استفاده از یک تابع خطا^۱ با پاسخ صحیح مقایسه کرده و به این ترتیب، میزان خطا محاسبه می‌شود. سپس بر اساس میزان خطای محاسبه‌شده، مرحله پس انتشار آغاز می‌شود. در این مرحله، گرادیان^۲ هر پارامتر با توجه به قاعده زنجیره‌ای^۳ محاسبه شده و همه پارامترها با توجه به تاثیری که بر خطای ایجادشده در شبکه دارند تغییر پیدا می‌کنند. بعد از به‌روز شدن پارامترها، مرحله رو به جلو شروع می‌شود. بعد از تکرار تعداد مناسبی از این مراحل آموزش شبکه پایان می‌یابد.

۳-۳ انواع لایه‌های شبکه کانولوشنی

در حالت کلی، یک شبکه کانولوشنی، یک شبکه عصبی عمیق سلسله‌مراتبی است که لایه‌های کانولوشنی آن به‌صورت یک‌درمیان با لایه‌های تجمیع، قرار گرفته و بعد از آن‌ها تعدادی لایه تمام‌متصل وجود دارد. در ادامه به توضیح وظایف هر یک از سه لایه مطرح‌شده خواهیم پرداخت.



شکل ۳-۱ یک طرح کلی از معماری یک شبکه عصبی کانولوشن [۹۳].

۳-۳-۱ لایه کانولوشن

لایه‌های کانولوشنی، الگوهای محلی را فرا می‌گیرند. اولین لایه کانولوشنی، الگوهای محلی کوچکی مانند لبه‌ها را فرا می‌گیرد، دومین لایه کانولوشنی، الگوهای بزرگتری که از ویژگی‌های اولین لایه

^۱ Loss Function

^۲ Gradient

^۳ Chain Rule

ساخته شده‌اند مانند چشم، گوش، اشکال هندسی و غیره را فرا می‌گیرد. لایه های بعدی از ترکیب آن‌ها، الگوهای پیچیده‌تری مانند سر انسان، درخت، خودرو و غیره را یاد می‌گیرد. این مسئله به کانولوشن‌ها اجازه می‌دهد که مفاهیم پیچیده و انتزاعی را به شکل کارآمدی یاد بگیرند. یک کانولوشن با لغزاندن پنجره‌هایی با اندازه 3×3 یا 5×5 ، روی نقشه ویژگی^۱ ورودی، کار می‌کند. عملیات کانولوشنی، تکه‌هایی را از نقشه ویژگی ورودی‌اش بیرون می‌کشد و یک نقشه ویژگی جدید شامل مفاهیمی که هدف شبکه است، تولید می‌کند.

۳-۳-۲ لایه تجمیع (Pooling)

پس از هر لایه کانولوشن، عموماً یک لایه تجمیع قرار دارد که وظیفه آن حفظ ویژگی‌های مطلوب و حذف موارد کم‌اهمیت‌تر از میان ویژگی‌هایی است که کرنل (پنجره) کانولوشن از تصویر استخراج نموده‌است. برای پیاده‌سازی این لایه، معمولاً از روش Max Pooling استفاده می‌شود. ابتدا نگاشت ویژگی‌های استخراج‌شده از لایه‌های قبل، به پنجره‌های مجزا از هم تقسیم شده، سپس بیشترین مقدار موجود در این پنجره انتخاب و به‌عنوان مقدار در نگاشت ویژگی جدید انتخاب می‌شود.

۳-۳-۳ لایه تماماً متصل (Fully Connected)

بعد از آخرین لایه Pooling، لایه‌های تماماً متصل وجود دارند. همان‌طور که در Error! Reference source not found، مشاهده می‌شود و از نام این لایه نیز پیداست، تمامی نورون‌های این لایه به نورون‌های لایه قبل متصل هستند. وظیفه اصلی لایه تماماً متصل، ترکیب ویژگی محلی در لایه‌های پایین با ویژگی محلی در لایه‌های بالایی است. در مسئله طبقه‌بندی، آخرین لایه در شبکه، تمام ویژگی‌ها را ترکیب می‌کند تا تصویر را طبقه‌بندی کند. به‌همین دلیل، اندازه خروجی این لایه برابر با تعداد کلاس‌های قابل شناسایی توسط شبکه است. لایه های تماماً متصل همانند همتایان خود در

¹ Feature Map

شبکه های عصبی مصنوعی سنتی عمل می کنند و تقریباً ۹۰٪ پارامترهای یک شبکه کانولوشنی را شامل می شوند. لایه تماماً متصل اجازه می دهد تا نتیجه شبکه در قالب یک بردار با اندازه مشخص ارائه شود. از این بردار می توان در ادامه پردازش های بعدی بهره برد.

۳-۴ مفاهیم و الگوریتم ها در حوزه شبکه های عصبی کانولوشن

در سال های اخیر از شبکه های کانولوشن در حوزه پردازش تصاویر، استفاده های زیادی شده و الگوریتم های معروفی از شبکه های عصبی به وجود آمده است. در این بخش، بنا بر کار موردنظر این پایان نامه در راستای بخش بندی تصاویر، دوتا از مهم ترین الگوریتم ها بر پایه شبکه های کانولوشنی را معرفی می کنیم. قبل از توضیح این الگوریتم ها، به معرفی مفهومی به نام پیش آموزش می پردازیم.

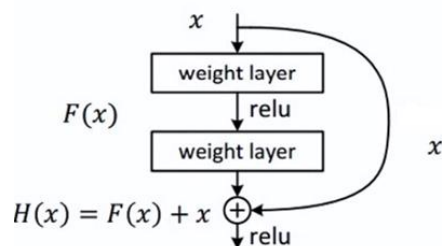
۳-۴-۱ پیش آموزش

آموزش شبکه های عمیق به علت تعداد بالای پارامترها، نیاز به داده های آموزش زیاد، صرف زمان و سخت افزار مناسب دارد. اما وقتی یک شبکه روی یک مجموعه داده خاص مثل ImageNet آموزش می بیند، فقط اطلاعات مربوط به آن مجموعه را یاد نمی گیرد، بلکه اطلاعات کلی راجع به جنس داده را هم یاد می گیرد [۹۴]. این ویژگی شبکه های عمیق، باعث می شود بتوانیم از شبکه ای که برای یک وظیفه خاص مثل تشخیص برچسب در داده های ImageNet آموزش داده شده، برای کاربردهای دیگر مثل بخش بندی تصاویر پزشکی استفاده کنیم. این روش در مدل های مبتنی بر CNN بسیار محبوب است زیرا فرآیند یادگیری را تسریع کرده و قابلیت تعمیم مدل را افزایش می دهد. یکی از انواع این شبکه ها که برنده رقابت سال ۲۰۱۵ نیز بود، شبکه ResNet است.

۳-۴-۲ الگوریتم ResNet

این شبکه‌ی بسیار عمیق که توسط گروه مایکروسافت^۱ طراحی شده، تا ۱۵۲ لایه عمق دارد. این شبکه در سال ۲۰۱۵ در مسابقه ImageNet میزان خطای ۵ کلاس برتر را به ۳٪،۶ کاهش داد [۹۵]. این مسئله همیشه مطرح بوده‌است که با افزایش عمق شبکه، دقت شبکه نیز افزایش می‌یابد و بسیاری پژوهش‌ها نیز این مسئله را نشان داده‌اند. اما در عمل، وقتی عمق شبکه از یک حد خاصی بیشتر می‌شود، اوضاع تغییر کرده و خطا شروع به بالا رفتن می‌کند. به این مشکل، کاهش^۲ گفته می‌شود. برای رفع این مشکل، در شبکه ResNet، بلوک‌های ویژه‌ای طراحی شده‌اند که به آن‌ها بلوک‌های باقی‌مانده^۳ گفته می‌شود. به‌طور کلی، در این شبکه ارتباطاتی خارج از ساختار کانولوشنالی تا ورودی‌های لایه قبلی را بدون واسطه به لایه بعدی منتقل کرده و در مرحله اصلاح شبکه (پس انتشار)، خطای هر لایه را به لایه قبلی انتقال دهد. بدین وسیله، می‌توان شبکه را عمیق‌تر کرده و آن را سریع‌تر آموزش داد. به این ارتباطات، Skip Connection و به ساختار حاصل از آن Residual block می‌گویند. بنا به نیاز و پیچیدگی مسئله مورد استفاده، می‌توان از نسخه‌های مختلفی از این الگوریتم مثلاً با ۲۲ لایه، ۳۴ لایه، ۸۲ لایه و ۱۸۲ لایه استفاده کرد.

همان‌طور که در شکل ۳ نشان داده شده‌است، بلوک رسم شده، ورودی را دریافت کرده و پس از عبور دادن آن از لایه اول و تابع Relu آن را وارد لایه دوم می‌کند. در این مرحله قبل از وارد کردن جریان داده به تابع Relu آن را با مقدار داده‌ی ورودی لایه قبل از خودش جمع می‌کند.



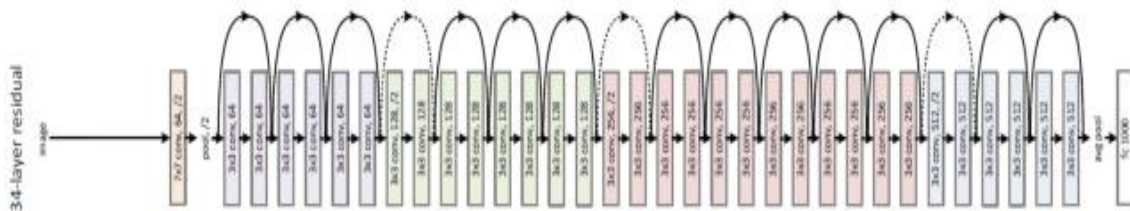
شکل ۳-۲ نمونه‌ای از یک بلوک باقی‌مانده [۹۵].

¹ Microsoft

² Degradation

³ Residual

در نهایت با پشت سر هم قرار دادن چندین بلوک باقی‌مانده، شبکه‌ای بسیار عمیق ساخته می‌شود که با افزایش عمق، با مشکل کاهش مواجه نخواهد شد. شکل نمایشی از ساختار عمیق شبکه‌ی ResNet با ۳۴ لایه ارائه کرده‌است.



شکل ۳-۳ شبکه ResNet [۹۵].

۳-۴-۳ الگوریتم U-Net

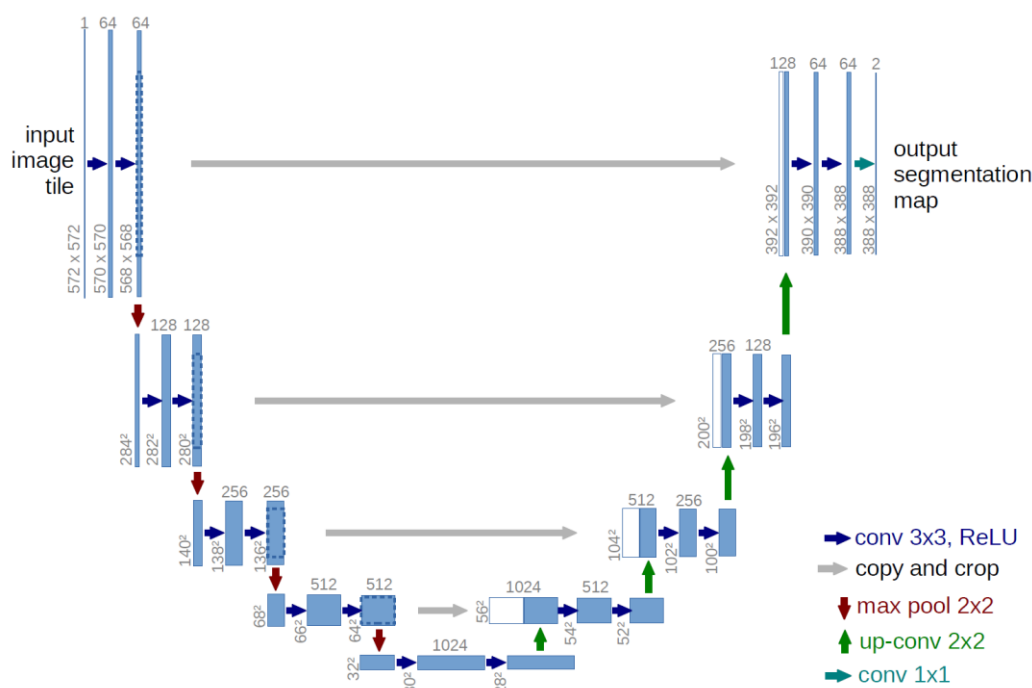
یکی از شبکه‌های عصبی شناخته‌شده در خانواده شبکه‌های کانولوشنی، شبکه‌ای U شکل، با عنوان U-Net است که برای بخش‌بندی تصاویر پزشکی طراحی شده‌است. الگوریتم U-Net [۹۶]، در سال ۲۰۱۵ در دانشگاه فرایبورگ آلمان توسط اولاف رنه‌برگر^۱ و همکاران برای قطعه‌بندی تصاویر پزشکی با افزایش سرعت و دقت پردازش، با استفاده از شبکه‌های تمام کانولوشنی^۲ طراحی شد. این شبکه، هیچ لایه متراکم^۳ یا تمام‌متصلی ندارد. این امر، باعث کاهش نیاز به داده آموزشی شده و در نهایت، منجر به قطعه‌بندی بسیار دقیق اجزای تصویر می‌شود. ایده اصلی نهفته در پشت این الگوریتم، ایجاد یک مسیر انقباضی متوالی است که در آن لایه‌های pooling، با لایه‌های sampling جایگزین شده‌است. پس از تلخیص و گردآوری ویژگی‌های مورد نظر، کار به یک مسیر انبساطی متوالی، سپرده می‌شود که شروع به بازسازی تصویر اولیه و انتشار اطلاعات پس‌زمینه به درون آن می‌نماید. چنان‌که در شکل مشاهده می‌شود، شبکه U-Net متشکل از دو مسیر یا زیرشبکه انقباضی و انبساطی کاملاً متقارن و نظیر یکدیگر است که تعداد طبقه‌های برابری دارند. هرچه در مسیر انقباض حذف شده‌است، در مسیر انبساط دوباره بازیابی می‌شود. مسیر انقباضی، شبیه به شکل عمومی شبکه‌های کانولوشنی است و

^۱ Olaf Ronneberger

^۲ Fully Convolutional

^۳ Dense

شامل یک لایه کانولوشنی است که مسئول اجرای مکرر فیلتر کانولوشنی 3×3 بر روی تصاویر است. پس از هر لایه کانولوشنی، یک لایه فعال‌ساز خطی اصلاح‌شده (Relu) و سپس یک لایه Maxpooling 2×2 با گام ۲ قرار دارد. این سه لایه با یکدیگر تشکیل یک طبقه نمونه بردار کاهنده را می‌دهند. در هر طبقه کاهنده، کانال‌های ویژگی دوبرابر می‌شود. برعکس، در مسیر انبساطی در هر طبقه یک نمونه بردار افزاینده (Upsampling) به همراه یک لایه کانولوشنی 2×2 که مسئول کاهش کانال‌های ویژگی است و داده‌هایی که قرار است به تصویر افزوده شوند را از طبقه نظیر خود در مسیر انقباضی دریافت می‌کند. پس از آن یک لایه کانولوشنی 3×3 به همراه یک لایه Relu قرار دارد. در لایه آخر یک لایه کانولوشنی 1×1 قرار دارد که هر ۶۴ بردار ویژگی را به یک طبقه مورد نظر نگاشت می‌کند. در مجموع، این شبکه شامل ۲۳ لایه کانولوشنی است.



شکل ۳-۴ معماری شبکه U-Net [۹۶].

۳-۵ بخش‌بندی تصاویر پزشکی مبتنی بر یادگیری عمیق

در فصل دوم، روش‌های بخش‌بندی سنتی برای تصاویر پزشکی را مورد بررسی قرار دادیم. در این بخش، قصد داریم تا مروری بر روش‌های مبتنی بر یادگیری عمیق در حیطه بخش‌بندی تصاویر

پزشکی داشته باشیم. همان طور که در بخش قبل مطرح کردیم، بهترین و پرکاربردترین شبکه عصبی عمیق در زمینه پردازش تصاویر و از آن جمله بخش بندی آن ها، شبکه های عصبی کانولوشنی هستند. در میان معماری های مختلف از خانواده شبکه های کانولوشن، U-Net متداول ترین و مناسب ترین شبکه ی طراحی شده جهت بخش بندی تصاویر پزشکی محسوب می شود. معماری U-Net، یک معماری شبکه عصبی محبوب برای انجام تقسیم بندی تصویر پزشکی است. این شبکه امروزه در حوزه بخش بندی تصاویر پزشکی، به لبه تکنولوژی دست یافته است. به عنوان مثال، سواستوپولسکی^۱ و همکاران [۹۷]، از این شبکه استفاده کردند تا به طور مستقیم دیسک نوری را در تصاویر فوندوس شبکیه برای تشخیص بیماری آب سیاه، بخش بندی کنند. روی^۲ و همکاران [۹۸] برای تقسیم بندی لایه شبکیه در تصاویر توموگرافی انسجام نوری (OCT)^۳ از شبکه U-Net استفاده کردند. در مرجع [۹۹] نیز از U-Net برای تقسیم غضروف و منیسک^۴ از تصاویر ام آر آی زانو استفاده شده است.

در سال های اخیر، نسخه های مختلفی مبتنی بر معماری شبکه U-Net ارائه شده است که در واقع شامل دست کاری و ایجاد تغییراتی در راستای بهبود شبکه اصلی U-Net هستند. به عنوان مثال برخی مقالات، برای افزایش توانایی قابلیت یادگیری U-Net، ماژول های جدیدی را برای جایگزینی بلوک های اصلی این شبکه پیشنهاد کرده اند [۱۰۰ و ۱۰۱]، برخی دیگر از مقالات، روی تغییر در اتصالات موجود در این شبکه متمرکز شده اند [۱۰۲ و ۱۰۳].

در ادامه چند نسخه تغییر یافته از شبکه U-Net را مورد بررسی قرار می دهیم.

۳-۵-۱ ساختار RU-Net

یکی از نسخه های تغییر یافته U-Net، شبکه ای به نام RU-Net^۵ است [۱۰۲]. معماری این شبکه،

¹ Sevastopolsky

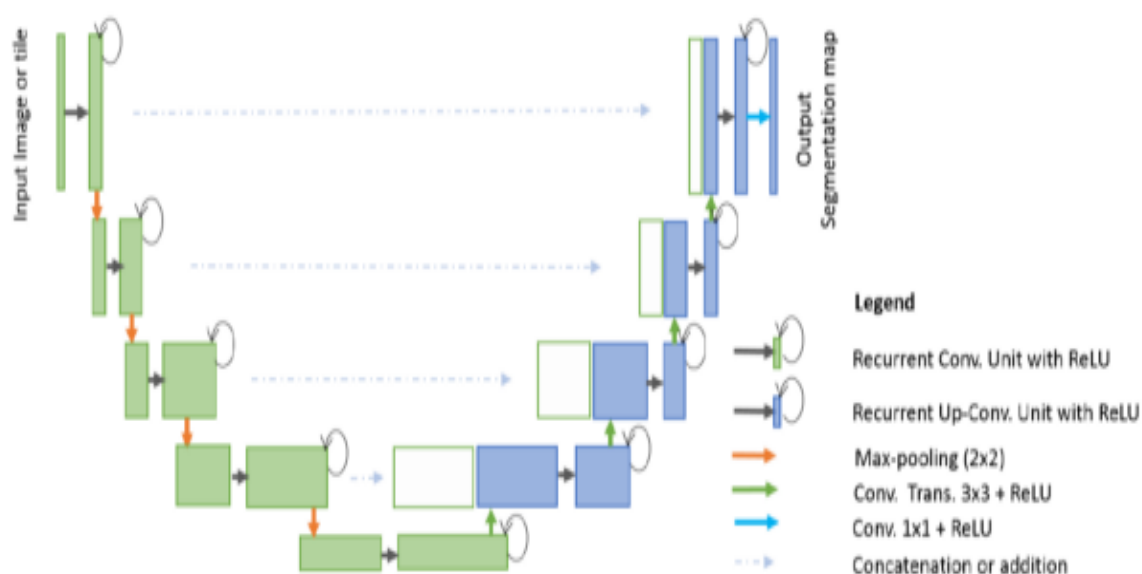
² Roy

³ Optical Coherence Tomography

⁴ Meniscus

⁵ Recurrent Convolutional Neural Network based on U-Net

تلفیقی از دو ساختار کانولوشنی RCNN^۱ و U-Net است. در واقع اسکلت این شبکه، کاملاً مبتنی بر ساختار U-Net بوده و نوآوری آن، استفاده از لایه‌های کانولوشنی مکرر در مسیر انقباض، قبل از لایه‌های DownSampling و در مسیر انبساط، قبل از لایه‌های UpSampling و در نهایت، قبل از تولید تصویر قطعه‌بندی‌شده در آخرین لایه از مسیر انبساطی است. این شبکه دو مزیت مهم نسبت به ساختار اصلی U-Net دارد. اول این که، استفاده از لایه‌های کانولوشنی مکرر، باعث استخراج ویژگی‌های بهتری در مسیر انبساطی شبکه می‌شود. دوم این که، این ساختار جدید به ما اجازه می‌دهد تا شبکه‌ای با عملکرد بهتر، مبتنی بر همان U-Net و تعداد پارامترهای قبلی آن، ارائه کنیم. شکل، معماری کلی این شبکه را نشان می‌دهد.



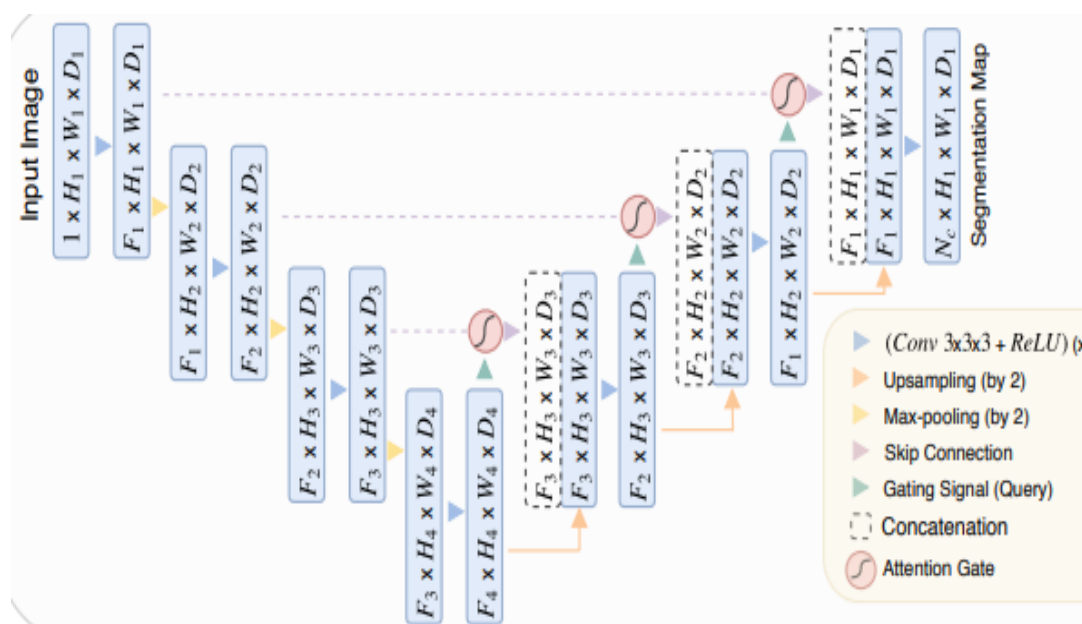
شکل ۳-۵ معماری شبکه RU-Net [۱۰۲].

۳-۵-۲ ساختار Attention U-Net

در ساختار شبکه اصلی U-Net، جهت بازیابی اطلاعات مکانی تصویر در هر لایه از مسیر انبساط، محتوای لایه‌ی قبلی در مسیر انبساط با محتوای لایه‌ی معادل از مسیر انقباض، الحاق می‌شود (Skip

¹ Recurrent Convolutional Neural Network

connection). با این حال، این امر موجب استخراج بسیاری از ویژگی‌های سطح پایین و اضافی می‌شود، به‌طوری که نمایش ویژگی در لایه‌های اولیه ضعیف است. برای حل این مشکل، یک ساختار مبتنی بر U-Net به نام Attention U-Net ارائه شد [۱۰۳]. ایده این شبکه، استفاده از دروازه‌هایی به نام Attention در اتصالات پرش است، به‌نحوی که نقشه‌ویژگی‌های^۱ استخراج‌شده در اتصالات پرش، ابتدا توسط دروازه‌های Attention پردازش‌شده و سپس با محتوای لایه قبلی در مسیر انبساط، الحاق می‌شوند. عملکرد این دروازه‌ها، کاهش منابع محاسباتی هدر رفته بر توابع فعال‌ساز نامربوط و ارائه شبکه، با قدرت تعمیم بهتر است. اساساً، این شبکه می‌تواند توجه^۲ را به قسمت‌های خاصی از تصویر متمرکز کند. شکل، معماری این شبکه را نشان می‌دهد.



شکل ۳-۶ معماری شبکه Attention U-Net [۱۰۳].

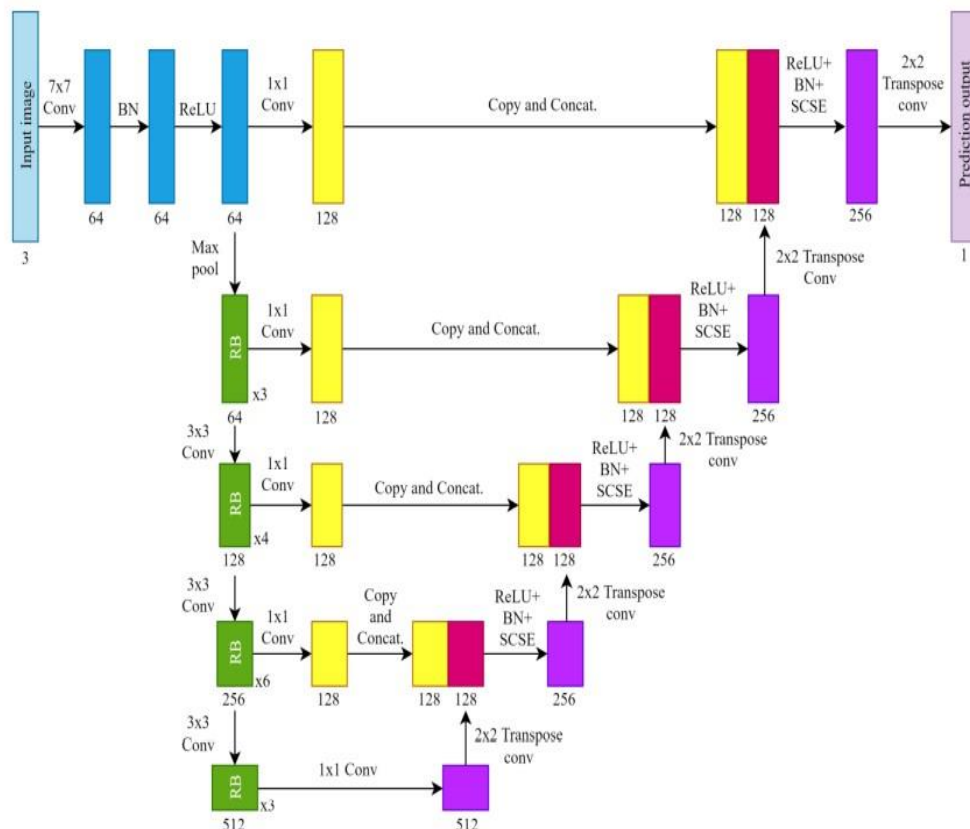
۳-۵-۳ ساختار ResNet34 U-Net

با توجه به آزمایش‌های انجام شده در شبکه‌های مختلف و مقایسه نتایج آن‌ها، استفاده از لایه‌ی کانولوشن به‌جای لایه‌ی Pooling ترجیح داده شده‌است؛ زیرا لایه‌های Pooling منجر به از دست

^۱ Feature Maps

^۲ Attention

دادن اطلاعات معنایی در تصاویر می‌شوند [۱۰۴]. در معماری ResNet، فقط در لایه اول از Pooling استفاده شده‌است و در لایه‌های بعدی برای حذف اطلاعات، از لایه کانولوشن با مقدار گام دو در جهت کاهش حجم اطلاعات استفاده شده‌است. از طرف دیگر، استفاده از شبکه‌هایی مانند ResNet که از قبل روی مجموعه داده‌های بزرگی مانند ImageNet آموزش دیده‌اند، فرآیند یادگیری را تسریع کرده و قابلیت تعمیم مدل را افزایش می‌دهد. به همین منظور در جهت بهبود شبکه اصلی U-Net، می‌توان به جای استفاده از لایه‌های کانولوشن دلخواه، از معماری ResNet در بخش کدگذار این شبکه استفاده کرد. به همین دلیل در مرجع [۱۰۵]، شبکه‌ای مبتنی بر U-Net ارائه شده‌است که در مسیر انقباضی آن (سمت چپ U)، از مدل پیش‌آموزش ResNet با ۳۴ لایه استفاده شده‌است. شکل ۷-۳ معماری این شبکه را نشان می‌دهد.



شکل ۷-۳ معماری شبکه ResNet34 U-Net [۱۰۵].

۳-۵-۴ ساختار BCDU-Net

یکی از شبکه‌های مبتنی بر U-Net که در سال ۲۰۱۹ در مرجع [۱۰۶] ارائه شد، ساختاری تحت‌عنوان BCDU-Net است. طبق نتایج ارائه‌شده در این مقاله، از سایر نسخه‌های تغییر یافته الگوریتم U-Net، بهتر است.

مسیر کدگذاری این شبکه، شامل چهار مرحله است. هر مرحله شامل دو فیلتر کانولوشنی 3×3 و پس از آن، یک تابع Pooling 2×2 و فعال‌ساز Relu است. تعداد نقشه‌های مشخصه در هر مرحله دوبرابر می‌شود. مسیر انقباض، به تدریج بازنمایی تصاویر را استخراج می‌کند و ابعاد لایه‌های بازنمایی را لایه به لایه افزایش می‌دهد. در نهایت لایه نهایی در مسیر کدگذاری، یک نمایش تصویر ابعادی بالا با اطلاعات معنایی بالا تولید می‌کند. آخرین مرحله از مسیر کدگذاری در نسخه اصلی U-Net، شامل دنباله‌ای از لایه‌های کانولوشنی است؛ این امر موجب یادگیری ویژگی‌های مختلف تصویر می‌شود. با این وجود، شبکه ممکن است ویژگی‌های اضافی را در این کانولوشن‌های پیوسته، یاد بگیرد. برای جلوگیری از این امر در این نسخه تغییر یافته، از مکانیزم لایه‌های کانولوشنی متراکم^۱ در آخرین لایه کانولوشنی مسیر کدگذاری استفاده شده است.

مکانیزم شبکه کانولوشنی متراکم، اولین بار در سال ۲۰۱۷ و در مرجع [۱۰۷] مطرح شد. ایده این شبکه، از روی شبکه ResNet گرفته شده است. همان‌طور که در بخش ۳-۴-۲ مطرح شد، این شبکه ورودی را دریافت کرده و پس از عبور دادن آن از لایه اول و تابع Relu آن را وارد لایه دوم می‌کند. در این مرحله، قبل از وارد کردن جریان داده به تابع Relu آن را با مقدار داده‌ی ورودی لایه قبل از خودش جمع می‌زند. برای بهبود بیشتر جریان اطلاعات بین لایه‌ها، ایده اتصالات مستقیم از هر لایه به همه لایه‌های بعدی معرفی شد. در این روش در مقایسه با ResNet، ترکیب ویژگی‌ها از طریق جمع کردن، رخ نمی‌دهد، بلکه با الحاق ویژگی‌ها، انجام می‌شود. به‌طور کلی، ایده استفاده از لایه‌های

¹ Densely connected convolutions

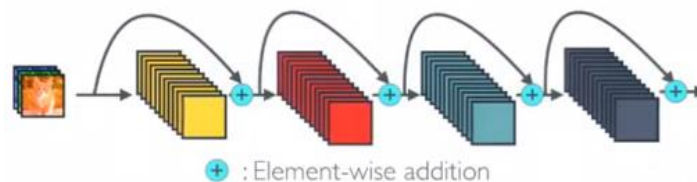
کانولوشنی متراکم به جای لایه‌های کانولوشنی معمولی، دارای مزایایی است که در زیر به آن‌ها اشاره شده است:

(۱) این ایده به شبکه کمک می‌کند تا مجموعه متنوعی از نقشه‌های مشخصه را به جای ویژگی‌های اضافی، یاد بگیرد.

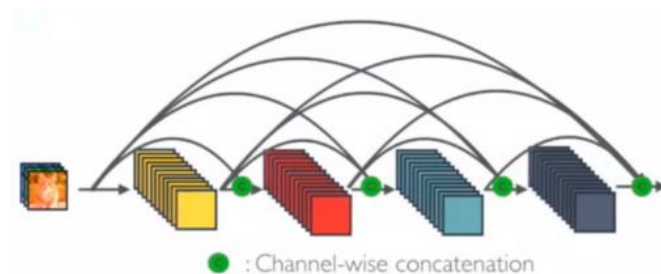
(۲) این ایده با اجازه انتشار جریان اطلاعات از طریق شبکه و استفاده مجدد از ویژگی‌ها، قدرت بازنمایی شبکه را بهبود می‌بخشد.

(۳) از اثر محوشدگی گرادیان جلوگیری می‌کند.

با توجه به موارد مزایای بیان شده برای لایه‌های کانولوشنی متراکم، با استفاده از این مکانیزم در آخرین لایه کانولوشنی مسیر کدگذاری، ویژگی‌هایی که در هر بلوک یاد گرفته می‌شوند، به بلوک بعدی منتقل می‌شوند. لذا این استراتژی به شبکه کمک می‌کند تا از یادگیری ویژگی‌های اضافی اجتناب کند. شکل، تفاوت بین اتصالات در بلوک Res و Dense را نشان می‌دهد.



(الف)



(ب)

شکل ۳-۸ (الف) بلوک‌های شبکه ResNet (ب) یک بلوک از شبکه DenseNet [۱۰۸].

در مسیر رمزگشایی (انبساط) جهت بازیابی اطلاعات مکانی تصویر در هر لایه، محتوای لایه قبلی در

مسیر انبساط با محتوای لایه‌ی معادل از مسیر انقباض، الحاق می‌شود (Skip connection). برای ترکیب این نقشه‌های ویژگی، به‌جای یک الحاق ساده در مسیر اتصال پرش، از ConvLSTM دو جهته (BConvLSTM)^۱ استفاده شده که به روش پیچیده‌تری این دو نوع نقشه‌ویژگی را پردازش می‌کند [۱۰۹]. برای درک چگونگی کار در مسیر کدگشایی، این بخش را با آوردن روابط ریاضی و شکل، بیشتر توضیح می‌دهیم.

فرض کنیم X_e ، نقشه‌ویژگی‌هایی باشد که از مسیر کدگذاری کپی شده و X_d ، نقشه‌ویژگی‌های لایه‌ی کانولوشنی قبلی در مسیر کدگشایی باشد. همچنین F_L ، تعداد نقشه‌های ویژگی و $W_L \times H_L$ ، اندازه هر نقشه‌ویژگی در لایه L است. X_d مطابق شکل، ابتدا به یک لایه Up-convolution داده می‌شود که در آن یک تابع Up-sample و سپس یک کانولوشن 2×2 اعمال شده‌است. خروجی این قسمت X_d^{up} است. سپس X_d^{up} تحت تابع نرمال‌سازی دسته‌ای (BN)^۲ قرار می‌گیرد. BN را می‌توان تکنیکی برای بهبود عملکرد، افزایش سرعت و پایداری شبکه‌های عصبی دانست [۱۱۰]. ایده این کار، نرمال‌سازی ورودی هر لایه با کمک میانگین و واریانس بوده به شکلی که میانگین اعداد بردار نهایی برابر با صفر و واریانس آن‌ها برابر با یک شود. خروجی حاصل از تابع نرمال‌سازی دسته‌ای، X_d^{up} بوده که آماده ورود به لایه BConvLSTM است. این لایه خود از دو ConvLSTM برای پردازش داده‌های ورودی به مسیرهای رو به جلو^۳ و رو به عقب^۴ استفاده می‌کند. بنابراین با مدیریت وابستگی به داده‌ها در هر دو جهت، تصمیمی برای ورودی فعلی گرفته می‌شود. علت استفاده از ConvLSTM به‌جای LSTM استاندارد، عدم لحاظ کردن همبستگی مکانی در شبکه‌های LSTM است. در واقع مشکل شبکه‌های LSTM، این است که همبستگی مکانی را در نظر نگرفته و زیرا از اتصالات کامل در انتقال‌های ورودی به حالت و حالت به حالت استفاده می‌کنند. برای حل این مشکل، ConvLSTM [۱۱۱] پیشنهاد شده که از عملیات کانولوشنی در انتقال‌های ورودی به حالت و حالت به حالت، بهره

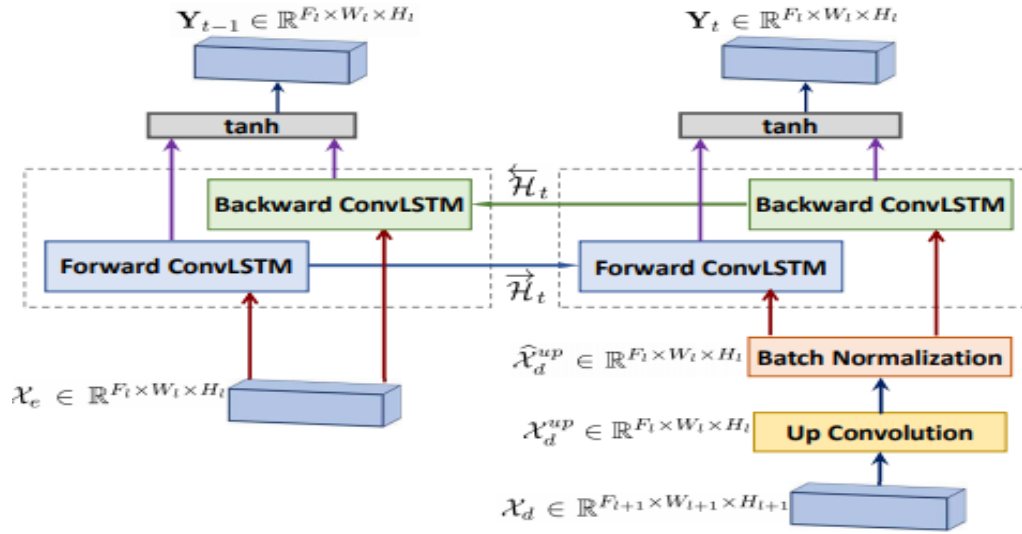
^۱ Bi-Directional ConvLSTM

^۲ Batch Normalization

^۳ Forward

^۴ Backward

می‌برد.



شکل ۳-۹ بلوک Bi-directional ConvLSTM [۱۰۶].

همان‌طور که در شکل مشاهده می‌شود، در بلوک BConvLSTM از دو ConvLSTM برای پردازش داده‌های ورودی استفاده شده‌است که هر ConvLSTM خود شامل اجزای زیر است:

- گیت ورودی (i_t)
- گیت خروجی (o_t)
- گیت فراموشی (f_t): این گیت در واقع یک شبکه عصبی کوچک است که سه ورودی و یک بایاس را دریافت کرده و از یک تابع sigmoid به‌عنوان تابع فعال‌ساز استفاده می‌کند. خروجی این شبکه مقدار فراموشی است.
- سلول حافظه (C_t)

به‌طور کلی، روابط ConvLSTM به صورت رابطه (۳-۱) هستند.

$$\begin{aligned}
 i_t &= \sigma(W_{xi} * X_t + W_{hi} * H_{t-1} + W_{ci} * C_{t-1} + b_i) \\
 f_t &= \sigma(W_{xf} * X_t + W_{hf} * H_{t-1} + W_{cf} * C_{t-1} + b_f) \\
 C_t &= f_t \circ C_{t-1} + i_t \tanh(W_{xc} * X_t + W_{hc} * H_{t-1} + b_c) \\
 o_t &= \sigma(W_{xo} * X_t + W_{ho} * H_{t-1} + W_{co} * C_t + b_o) \\
 H_t &= o_t \circ \tanh(C_t)
 \end{aligned} \tag{۳-۱}$$

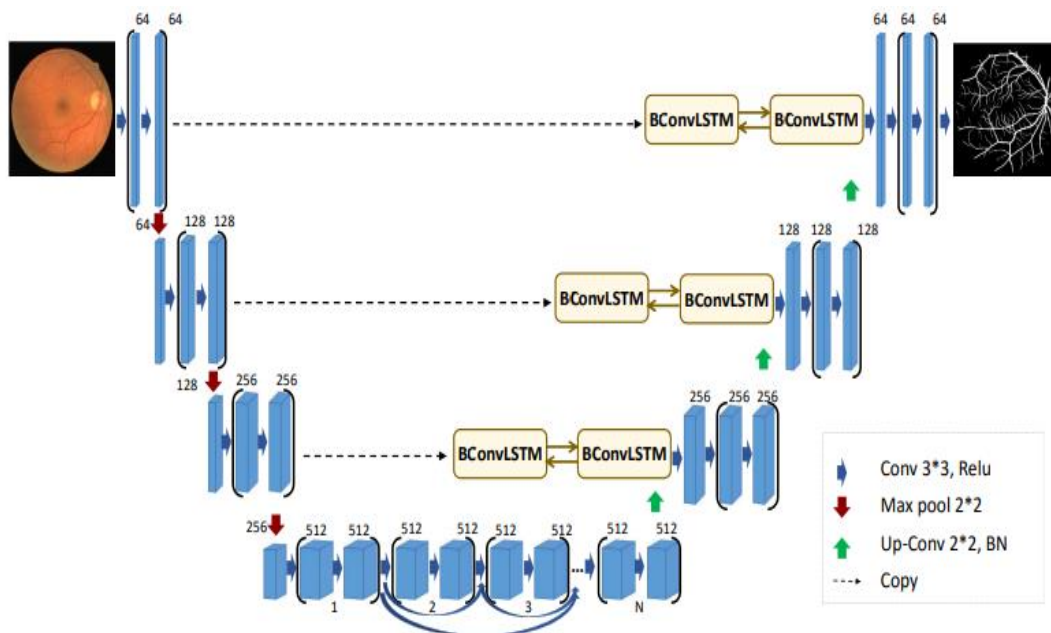
در روابط (۳-۱)، * و O ، به ترتیب نمایانگر توابع کانولوشن و هادامارد^۱ هستند. X_t معادل تانسور ورودی (در اینجا X_t معادل X_e و X_d^{up} است)، H_t ، معادل تانسور مخفی، C_t تانسور سلول حافظه، W_{h*} و W_{x*} ، هسته‌های کانولوشنی دوبعدی به ترتیب متناظر با حالت ورودی و پنهان هستند. b_c ، b_o ، b_f ، b_i نیز شروط بایاس هستند. در این شبکه، X_e و X_d^{up} توسط BConvLSTM رمزگشایی می‌شوند. خود BConvLSTM از دو ConvLSTM استفاده می‌کند تا داده‌های ورودی را به دو جهت مسیرهای رو به جلو و عقب پردازش کند و سپس با برخورد با وابستگی‌های داده‌ها در هر دو جهت، تصمیم می‌گیرد.

در یک ConvLSTM استاندارد، تنها وابستگی‌های جهت جلو پردازش می‌شود. با این حال، تمام اطلاعات در یک توالی باید به طور کامل در نظر گرفته شود. بنابراین، محتمل است که در نظر گرفتن وابستگی‌های متقابل، امری موثر باشد. ثابت شده است که تحلیل هر دو دیدگاه رو به جلو و عقب، عملکرد پیش‌بینی‌کننده را بهبود می‌بخشد [۱۱۲]. هر یک از ConvLSTM‌های رو به جلو و عقب را می‌توان یک استاندارد در نظر گرفت. بنابراین، ما دو مجموعه پارامتر برای حالت‌های رو به عقب و رو به جلو داریم. خروجی BConvLSTM از طریق رابطه (۳-۲) محاسبه می‌شود:

$$Y_t = \tanh(W_Y^{\bar{H}} * \bar{H}_t + W_Y^{\bar{H}} \bar{H}_t + b) \quad (۳-۲)$$

در رابطه (۳-۲)، $\bar{H}_t$ و $\bar{H}_t$ به ترتیب نمایانگر تانسورهای مخفی برای حالت رو به جلو و رو به عقب هستند، b ، نشان بایاس و Y_t ، نشان‌دهنده خروجی نهایی با در نظر گرفتن اطلاعات دوسویه است. علاوه بر این، از $\tanh$ برای ترکیب خروجی هر دو حالت جلو و عقب از طریق یک روش غیر خطی استفاده می‌شود. شکل، شماتیک نهایی شبکه BCDU-Net را نشان می‌دهد. همچنین، استفاده از مکانیزم لایه‌های کانولوشنی متراکم در آخرین لایه از مسیر کدگذاری و BConvLSTM در مسیر اتصال پرش، مشهود است.

¹ Hadamard



شکل ۳-۱۰ شبکه BCDU-Net [۱۰۶].

۳-۶ جمع‌بندی

در این فصل، ابتدا به توضیح یادگیری عمیق پرداخته و سپس یکی از مهم‌ترین ساختارهای آن یعنی شبکه‌های عصبی کانولوشنی را با جزئیات مورد بررسی قرار دادیم. در ادامه، کاربردی‌ترین و مهم‌ترین معماری‌های مبتنی بر شبکه‌های کانولوشنی را مطرح کرده و در این راستا، به بخش‌بندی تصاویر پزشکی با تکیه بر معماری‌های یادشده پرداختیم. درنهایت نتیجه گرفتیم که شبکه U-Net، مهم‌ترین شبکه‌ی کانولوشنی جهت بخش‌بندی تصاویر پزشکی محسوب می‌شود و دارای مزایای بسیار است. لذا در بخش آخر، نسخه‌های مبتنی بر ساختار U-Net ارائه‌شده را مورد تجزیه و تحلیل قرار دادیم. هر یک از این نسخه‌ها، سعی در بهبود نسخه اصلی این شبکه داشته و ایراداتی را در آن برطرف می‌کردند. با توجه به بررسی‌های مفصل و بحث‌های مطرح‌شده، در فصل چهارم، مدلی مبتنی بر یادگیری عمیق جهت بخش‌بندی تصاویر سی‌تی ریه ارائه خواهیم‌داد که هم مبتنی بر BCDU-Net بوده و هم از مزایای شبکه پیش‌آموزش‌دیده ResNet در آن، حداکثر استفاده می‌شود.

فصل ۴ : روش پیشنهادی

۴-۱ مقدمه

کار اصلی ما در این پایان‌نامه، از سه بخش استخراج برچسب تصاویر، پیش‌پردازش و آماده‌سازی داده‌ها و بخش‌بندی مبتنی بر یادگیری عمیق تشکیل شده‌است. در این فصل، ابتدا به معرفی مجموعه‌دادگان مورد استفاده خود در این تحقیق پرداخته، سپس روند تولید برچسب تصاویر مجموعه‌داده مذکور با روشی نیمه‌خودکار را توضیح داده و عملیات پیش‌پردازش جهت آماده‌سازی داده‌ها را به صورت گام‌به‌گام شرح می‌دهیم. در نهایت به تشریح روش مبتنی بر یادگیری عمیق جهت بخش‌بندی این تصاویر و برچسب‌های متناظرشان می‌پردازیم.

۴-۲ مجموعه‌دادگان LIDC-IDRI استفاده شده در این پایان‌نامه

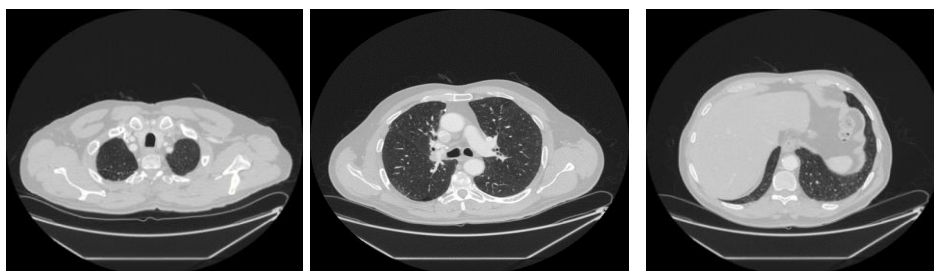
در این بخش قصد داریم تا به معرفی از مجموعه‌دادگان مورد استفاده خود در این پایان‌نامه بپردازیم. مجموعه‌داده LIDC^۱، یک مجموعه‌داده قابل دسترسی عمومی شامل ۳۳۹ تصویر سی‌تی‌اسکن قفسه سینه با گزارشات تشخیص و اندازه است که به عنوان یک منبع تحقیق تصویربرداری پزشکی به کار می‌رود. چهار رادیولوژیست، هر اسکن را با استفاده از پردازش‌های زیر مورد بررسی قرار داده‌اند:

ابتدا هر رادیولوژیست، تصاویر سی‌تی را به طور مستقل مورد بررسی قرار دادند. در مرحله دوم، نتایج از تمام چهار مرور کور گردآوری شدند و به هر رادیولوژیست برای مرور دوم داده شدند؛ تا رادیولوژیست‌ها گزارش‌های خود را با گزارش‌های دیگر رادیولوژیست‌ها مرور کنند. نتایج مرور هر رادیولوژیست جمع‌آوری شده تا مرور نهایی را شکل دهند. طبق تعبیرهای رادیولوژیست‌ها LIDC، شامل غده‌های بزرگ‌تر مساوی ۳ میلی‌متر است. تفسیرها همچنین شامل یک علامت تنها (یک مرکز تقریبی) از غده‌های کوچک‌تر مساوی ۳ میلی‌متر (قطر) همانند غده‌های بزرگ‌تر از ۳ میلی‌متر است

[۱۱۳].

^۱LIDC: Lung Image Database Consortium

با این فرض که می‌دانیم همکاری بخش خصوصی و دولتی برای سرعت بخشیدن به کشف علوم برای سلامتی انسان‌ها و موفقیتشان در این حوزه لازم است،^۱ FNIH که بنیادی برای موسسات بین‌المللی سلامت است، در سال ۲۰۰۴ مجموعه‌داده‌ی IDRI^۲ را به منظور تلاش‌های بیشتر برای پیشرفت LIDC ساخت. IDRI به پنج موسسه LIDC به همراه دو مرکز دانشگاهی و هشت شرکت تصویربرداری پزشکی ملحق شد. از طریق IDRI، این شرکت‌ها منابع کمکی (افزوده) را فراهم ساختند تا مجموعه‌داده LIDC را به مجموعه‌ای از ۱۰۰۰ سی‌تی اسکن گسترش دهند و یک مجموعه‌داده کامل، با تقریباً ۳۰۰ تصویر رادیوگرافی قفسه سینه دیجیتال، ایجاد کنند. در نتیجه، مجموعه‌داده LIDC-IDRI که امروز یک مرجع دردسترس عموم برای جامعه تحقیقاتی تصویربرداری پزشکی است با ۱۰۱۸ نمونه به اتمام رسید. هریک از این نمونه‌ها حاوی تصاویری از یک سی‌تی اسکن کلینیکی و یک فایل XML که حاوی نتایج تفسیر تصویر دو مرحله‌ای توسط چهار رادیولوژیست قفسه سینه باتجربه است، می‌باشد. همچنین میانگین تعداد برش‌ها برای هر نمونه، ۲۵۴ برش است. همه تصاویر این مجموعه‌داده، در قالب DICOM بوده و در واحد هانسفیلد اندازه‌گیری شده‌اند، همچنین دارای سه کانال بوده و در اندازه ۵۱۲×۵۱۲ پیکسل هستند. شکل ۴، چند نمونه از تصاویر این مجموعه‌داده را نشان می‌دهد.



شکل ۴-۱ تصاویر سی‌تی مربوط به یک بیمار از مجموعه‌داده IDRI-LIDC

^۱ Foundation for the National Institutes of Health

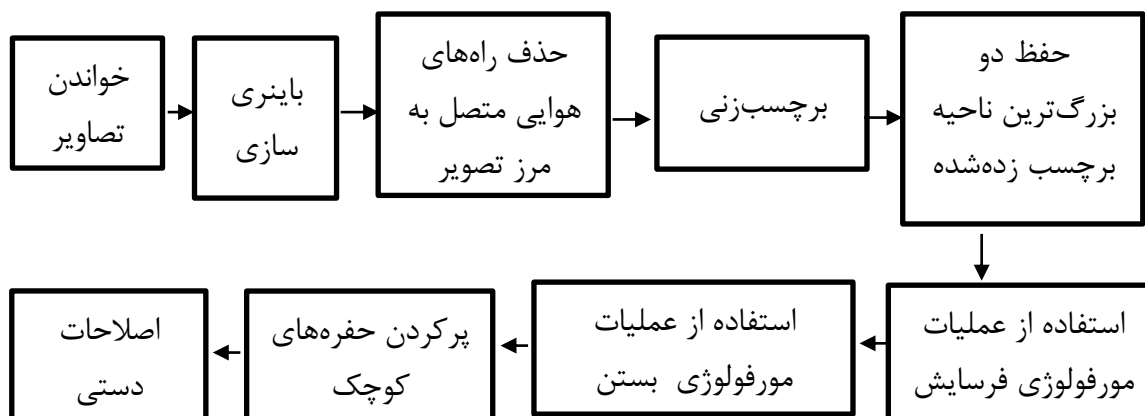
^۲ Image Database Resource Initiative

۳-۴ استخراج برچسب تصاویر سی تی

همان طور که در بخش قبل بیان شد، مجموعه داده LIDC-IDRI به ازای هر ۱۰۱۸ نمونه، تعدادی تصویر سی تی اسکن به همراه یک فایل XML دارد. این فایل، در واقع حاوی برچسبی برای ندول‌های موجود در تصاویر سی تی ریه مربوط به هر نمونه است. بخش‌بندی تصاویر سی تی ریه، نیازمند خود تصاویر به همراه برچسب‌های دودویی متناظر آن‌ها است؛ به نحوی که در این برچسب‌ها، فضای ریه از غیرریه کاملاً تفکیک شده باشد. در نهایت هم این تصاویر به همراه برچسب‌های باینری متناظرشان به یک شبکه عصبی عمیق خورانده شده و عمل بخش‌بندی صورت می‌گیرد.

در این بخش، به علت نیاز مبرم به برچسب هر یک از تصاویر و عدم وجود برچسب برای بافت ریه در این مجموعه داده، اقدام به تولید برچسب‌ها به شکلی نیمه خودکار کرده‌ایم. به طور کلی ما برای تولید ماسک حقیقی^۱ هر یک از تصاویر سی تی، از الگوریتم ارائه شده در مرجع [۳۵] استفاده کردیم.

شکل، گام‌های اصلی این مرحله را به صورت شماتیک نشان می‌دهد.



شکل ۲-۴ گام‌های مرحله تولید ماسک برای تصاویر سی تی

(۱) خواندن تصاویر DICOM: در اولین مرحله، اقدام به خواندن هر یک از برش‌ها از تصاویر

ورودی می‌شود.

(۲) تبدیل تصویر به باینری: در این مرحله، طبق رابطه (۴-۱) از آستانه‌گذاری، برای دودویی

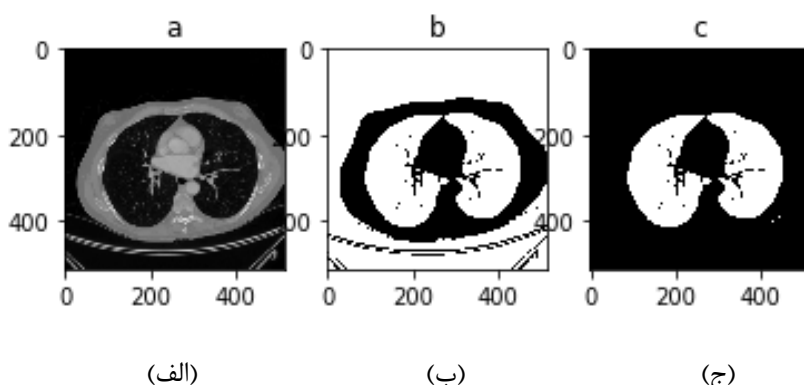
کردن تصویر استفاده شده است. از آنجایی که در آزمایش‌ها نشان داده شده که تمامی

¹ Ground Truth

مناطق مطلوب (ROI)، شدت بزرگتر از ۶۰۴ را دارند (400 HU-) [۱۱۴]، از این آستانه برای فیلتر کردن مناطق تاریکتر استفاده شده است. شکل بخش (ب)، تصویر باینری حاصل را نشان می‌دهد.

$$\text{binary}(i,j) = \begin{cases} 1 & \text{if } f(i,j) < T \\ 0 & \text{otherwise} \end{cases} \quad T=604 \quad (۱-۴)$$

۳) حذف راه‌های هوایی متصل به مرز تصویر: بعد از عمل باینری‌سازی، یک‌سری پیکسل‌های غیرریوی وجود دارند که با مرز تصویر اتصال داشته و باید از بافت اصلی ریه جدا شوند. در این مرحله، باید هوای پیرامون ریه و هم‌چنین ستون فقرات در برگیرنده بافت اصلی ریه جدا شوند. نتیجه این بخش در شکل بخش (ج)، نمایش داده شده است.



شکل ۳-۴ (الف): تصویر سی تی اصلی (ب): تصویر باینری شده (ج): تصویر حاصل بعد از حذف راه‌های هوایی متصل به مرز تصویر

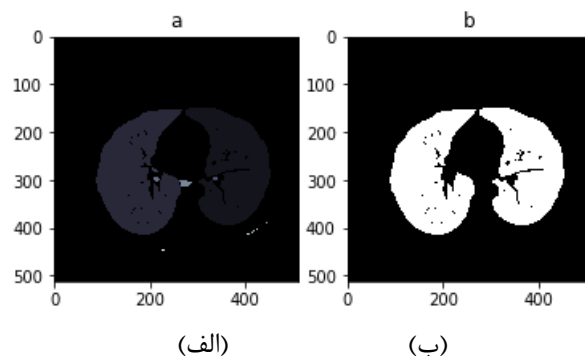
۱) برچسب‌زنی به تصویر: بعد از مراحل پردازشی بالا، فقط بافت ریه، نای، نایژه‌ها و برخی نواحی کوچک در اطراف بافت ریه، در تصویر باقی می‌ماند. به علاوه، تصویر برش سی تی ریه معمولاً شامل معده نیز می‌شود. بنابراین ممکن است که بخشی از معده به عنوان ریه طبقه‌بندی شود. به هر حال، نواحی باقی‌مانده از بافت ریه کوچک‌تر هستند. در این مرحله، به منظور پیدا کردن برچسب مربوط به حجم ریه و حذف باقی قسمت‌ها، از برچسب‌زنی استفاده شده به نحوی که دو پیکسل همسایه با ارزش یکسان متصل می‌شوند. نتیجه این

مرحله در شکل بخش (الف)، نشان داده شده است.

(۲) حفظ دو بزرگ‌ترین ناحیه برچسب زده شده: در این مرحله، دو مولفه از بزرگ‌ترین مولفه‌های برچسب زده شده در مرحله قبل را به عنوان دو بخش راست و چپ ریه طبق رابطه (۲-۴)، انتخاب کرده و سایر مؤلفه‌های کوچک که قطعا ریه نیستند، حذف می‌شوند. خروجی این مرحله در شکل بخش (ب)، نشان داده شده است.

$$\text{LungRegion} = \begin{cases} \max(\text{areaI}) & \text{lung1} \\ (\max-1)(\text{areaI}) & \text{lung2} \end{cases} \quad (2-4)$$

در رابطه (۲-۴)، منظور از R، نواحی با برچسب‌های متفاوت است.



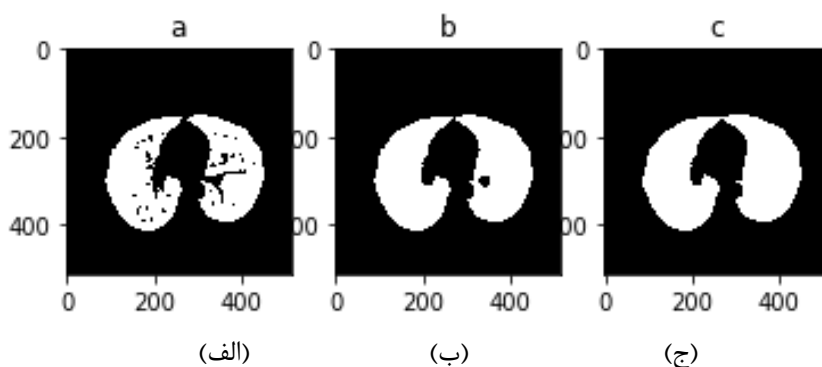
شکل ۴-۴ (الف): تصویر برچسب‌گذاری شده (ب): تصویر برچسب‌گذاری شده با حفظ دو بزرگ‌ترین مولفه

(۱) استفاده از عملیات فرسایش: در این مرحله، به منظور جدا کردن ندول‌های متصل به رگ‌های خونی، از عملیات فرسایش به کمک دیسکی با مقدار شعاعی ۲ پیکسل، استفاده کردیم. فرسایش، نواحی روشن کوچک‌تر از ماسک مربوطه را حذف کرده و مناطق تاریک را مطابق آنچه در شکل بخش (الف)، نشان داده شده، گسترش می‌دهد.

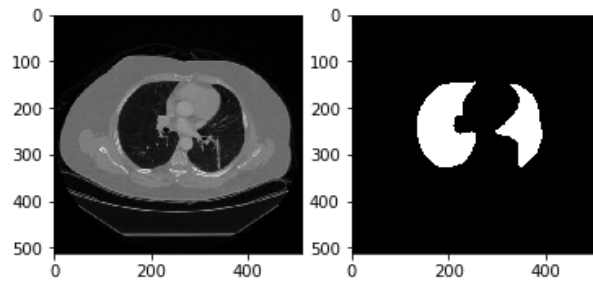
(۲) استفاده از عملیات بستن: در این مرحله، به منظور حفظ ندول‌های ریوی متصل به دیواره ریه، از عملیات بستن به کمک دیسکی با مقدار شعاعی ۱۰ پیکسل، استفاده

شده‌است. با استفاده از این عملگر، نقاط تاریک کوچک در تصویر حذف شده و شکاف‌های کوچک و روشن موجود در تصویر، به هم وصل می‌شوند. نتیجه حاصل از این عمل، در شکل بخش (ب)، نشان داده شده‌است.

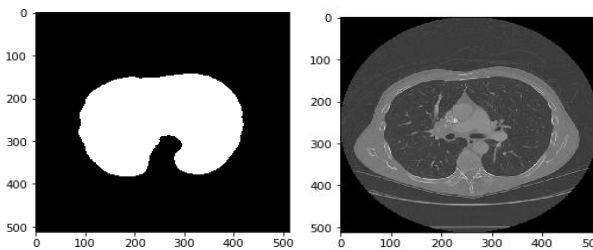
۳) **پرکردن حفره‌های کوچک:** گاهی به دلیل عدم آستانه‌گذاری مناسب هنگام تبدیل دودویی، مجموعه‌ای از نقاط سیاه، در تصویر تشکیل می‌شوند. این مناطق شناخته‌شده به عنوان حفره، ممکن است مفید باشند. بنابراین مطابق شکل بخش (ج)، باید این مناطق را با پر کردن آن‌ها در نظر گرفت.



شکل ۴-۵ تصویر حاصل بعد از (الف) عملیات فرسایش (ب) عملیات بستن (ج) پرکردن حفره‌های کوچک پس از پایان این مراحل، در حقیقت برچسب‌های تصاویر ما تولید شده‌اند. نتیجه حاصل از این هشت مرحله، همیشه به درستی کار نمی‌کند. چون اولاً گاهی در مرحله پنجم، یعنی حفظ برچسب‌ها با دو قسمت بزرگ، بافت‌های ریز شامل اجزای ریه حذف می‌شوند. دوم، گاهی مرحله هفتم، یعنی استفاده از عملیات بستن جهت اتصال شکاف‌های موجود کوچک در تصویر، منجر به اتصال دو پیکسل غیر ریوی می‌شود. شکل، یک نمونه از هر یک از مشکلات مطرح‌شده را نشان می‌دهد.



(الف)



(ب)

شکل ۴-۶ (الف) : نمایش مشکل غفلت از برخی بافت‌های ریوی در تصویر ماسک حاصل، به دلیل در نظر گرفتن فقط دو بزرگ‌ترین ناحیه برچسب زده شده (ب) نمایش مشکل اتصال پیکسل‌های غیر ریوی در تصویر ماسک حاصل، به دلیل خطا در مرحله بستن

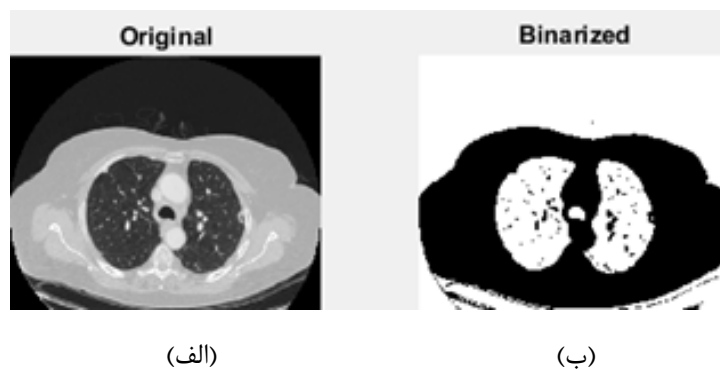
با توجه به مشکلات اشاره شده، در صورت لزوم باید از اصلاحات دستی پس از تولید ماسک های باینری، استفاده کرد. در این پایان نامه، ما با استفاده از این روش نیمه خودکار، ۱۷۱۴ ماسک باینری برای ۱۰ بیمار (به طور متوسط، ۱۷۰ نمونه برای هر بیمار) استخراج کردیم. مزیت اصلی این روش، صرفه جویی زیاد در زمان بوده که درباره این مسئله، در فصل پنجم به طور مفصل توضیحاتی خواهیم داد. در نهایت بعد از تولید تمام ماسک‌ها، آن‌ها را با فرمت 'bmp' ذخیره کردیم.

۴-۴ پیش پردازش و آماده سازی داده ها

هدف این مرحله، آماده سازی تصاویر خام ورودی در راستای بهبود آموزش شبکه یادگیری عمیق، با استفاده از چند مرحله عملیات پیش پردازش است. همان طور که در بخش ۴-۲ مطرح شد، تمام تصاویر سی تی در مجموعه داده LIDC-IDRI، دارای سه کانال بوده و در اندازه 512×512 پیکسل هستند. در این بخش، می خواهیم این سه کانال پر شده به طور پیش فرض را با سه کانال جداگانه که توسط خودمان طراحی شده است، پر کنیم. در این راستا ما از چند عملگر پیش پردازشی مانند توابع تشخیص

لبه و عملیات مورفولوژیکی استفاده می‌کنیم. مهم‌ترین مزیت این ایده، یعنی اعمال عملیات پیش‌پردازش بر روی تصاویر خام، بهبود در آموزش شبکه یادگیری عمیق و افزایش دقت و نیز کاهش نرخ خطای مثبت کاذب، خواهد بود. در ادامه به توضیح روند و مراحل کار برای طراحی کانال‌های جدید تصاویر می‌پردازیم:

(۱) **دودویی کردن تصویر:** این فرآیند، تصویر دودویی را ایجاد می‌کند که دو مقدار در سطح خاکستری (سیاه و سفید) دارد [۱۱۵]. بافت‌های ریوی، شامل رنگ سیاه با مقدار منطقی صفر هستند. در این مرحله، فرآیند دودویی کردن، برای جدا کردن قسمت‌های ریه‌ها انجام شده‌است. شکل، فرآیند دودویی کردن تصویر را نشان می‌دهد.



شکل ۴-۷ (الف): تصویر اصلی (ب): تصویر باینری شده

(۱) **استفاده از عملیات گسترش:** از آن‌جا که پس از فرآیند دودویی‌سازی تصویر، هنوز مناطق سفیدی در اطراف ناحیه ریه و حفره‌های سیاه کوچک داخل ریه‌ها یافت می‌شود که در واقع نوعی نویز محسوب می‌شوند، لذا باید آن‌ها را با عملیات مورفولوژیکی حذف کنیم. چنان‌چه قبلاً در بخش ۲-۶-۱ بیان شد، استفاده از عملگر گسترش باعث می‌شود نقاطی از یک تصویر باینری که در اثر عواملی چون تاثیر نویز یا اعمال حد آستانه نامطلوب جا افتاده‌است، تصحیح گردند. بنابراین ما عملیات گسترش را برای پاک کردن قسمت‌های اضافی سفید اطراف شش و حفره‌های سیاه کوچک درون شش اعمال کردیم. شکل، نتیجه حاصل از این عملیات را نشان می‌دهد.



شکل ۴-۸ (الف): تصویر دودویی شده (ب): تصویر پس از اعمال عملگر گسترش

همان‌طور که در شکل ۴-۸ مشاهده می‌شود، پس از اعمال عملگر گسترش بر روی تصویر باینری، قسمت سفید اطراف ریه در بخش (الف) که با فلش مشخص شده‌است، حذف گردیده و حفره‌های سیاه ریز داخل دو ریه چپ و راست، نیز محو شدند.

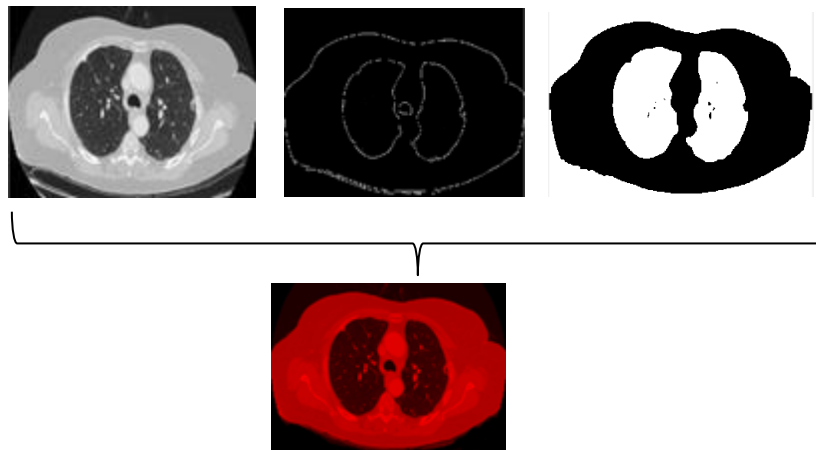
۲) استفاده از عملیات تشخیص لبه: همان‌طور که در بخش ۲-۸-۱-۳ ذکر شد، فرآیند تشخیص لبه، رئوس یک جسم و مرزهای بین اشیاء و پس‌زمینه در تصویر را شناسایی می‌کند. از عملیات تشخیص لبه می‌توان برای بهبود ظاهر یک تصویر مات استفاده کرد. الگوریتم لبه‌یاب canny، یک روش بسیار مهم برای یافتن لبه‌ها با حذف نویز از تصویر است. با توجه به مزایای بیان شده در رابطه با الگوریتم کنی، در این مرحله ما از این الگوریتم برای شناسایی لبه‌ها در تصاویر مرجع استفاده کردیم. شکل، نتیجه فرآیند تشخیص لبه را نشان می‌دهد.



شکل ۴-۹ تصویر پس از عملیات تشخیص لبه با الگوریتم کنی

اکنون قصد داریم تا کانال‌های جدید تصویر را پرکنیم. اولین کانال تصویر جدید، با تصویر اصلی پر شده‌است. دومین کانال از تصویر جدید، تصویری است که حاوی فرآیند تشخیص لبه است. تشخیص لبه، اطلاعات بی‌هوده و نویز را از تصویر حذف کرده و درعین حال، خصوصیات ساختاری لازم را در

یک تصویر حفظ می‌کند. در پایان، سومین کانال جدید، تصویر حاصل از عملیات گسترش است. این امر به کاهش ناحیه اطراف جسم کمک کرده و هم‌چنین نویز تصویر حذف می‌شود. شکل ، نتیجه ترکیب کانال‌های طراحی شده را نشان می‌دهد. ما ۱۷۱۴ تصویر سی‌تی جدید ریه را برای ۱۰ بیمار با استفاده از روش پردازش فوق، تولید کردیم.



شکل ۴-۱۰ تصویر جدید تولیدشده با ترکیب سه کانال جدید طراحی‌شده

۴-۵ مدل یادگیری عمیق پیشنهادی

از آن‌جا که هدف اصلی این پایان‌نامه، استخراج خودکار ریه‌ها از تصاویر سی‌تی است، مدل مورد نظر باید با موفقیت مشکلات بخش‌بندی را حل کند. یکی از بهترین معماری‌های مبتنی بر یادگیری عمیق برای این حوزه، شبکه U-Net است. این شبکه، مزایای بیشتری نسبت به دیگر شبکه‌ها در حوزه قطعه‌بندی تصاویر پزشکی دارد؛ زیرا اولاً قادر است با یک مجموعه داده نسبتاً کوچک، به‌خوبی آموزش ببیند. دوماً، قادر به استفاده هم‌زمان از ویژگی‌های محتوایی و مکانی تصویر است. از طرفی، می‌دانیم اگر یک مدل از پیش آموزش‌دیده مورد استفاده قرار گیرد، سرعت آموزش و قدرت تعمیم شبکه یادگیری عمیق، بسیار بالا می‌رود. بنابراین استفاده از مدل و وزن‌هایی که در مجموعه داده بزرگی مثل ImageNet از قبل آموزش‌دیده‌اند، ایده خوبی محسوب می‌شود. در فصل سوم گفتیم که در سال‌های اخیر، نسخه‌های مختلفی از شبکه U-Net ارائه شده‌است که

هر یک از این نسخه‌ها، با دست‌کاری و ایجاد تغییرات، سعی در بهبود شبکه اصلی U-Net دارند. هم‌چنین چند تا از مهم‌ترین شبکه‌های تغییریافته مبتنی بر U-Net را در فصل سوم بررسی کردیم. با بررسی‌های انجام‌شده، دیدیم که ساختار BCDU-Net از سایر نسخه‌های تغییریافته U-Net که قبلاً پیشنهاد شده‌اند، بهتر است. با بررسی مفصلی که همراه با جزئیات روی این ساختار در بخش ۳-۵-۴ داشتیم، دیدیم در این شبکه چند نوآوری مهم به‌کار رفته که در این بخش، به‌طور خلاصه آن‌ها را مرور می‌کنیم:

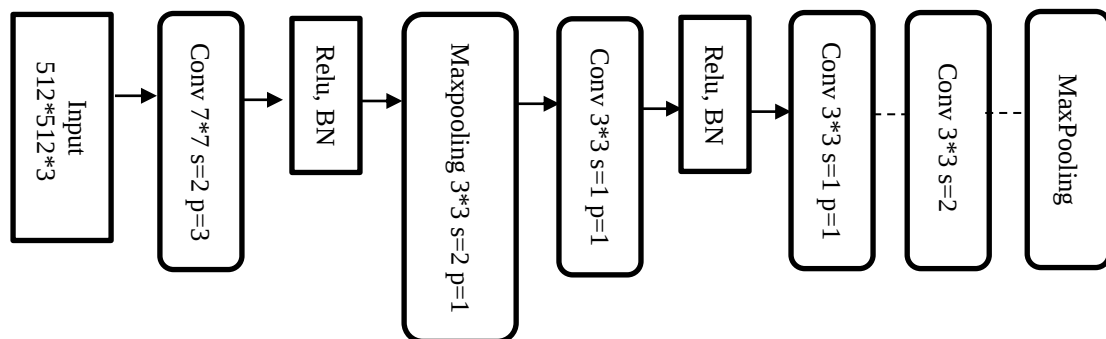
اول این که از مکانیزم لایه‌های کانولوشنی متراکم در آخرین لایه کانولوشنی مسیر کدگذاری استفاده شده‌است. این کار موجب جلوگیری از یادگیری ویژگی‌های اضافی در مسیر کانولوشن‌های پیوسته می‌شود. ایده استفاده از لایه‌های کانولوشنی متراکم به‌جای لایه‌های کانولوشنی معمولی، دارای چندین مزیت است که در بخش ۳-۵-۴ به آن‌ها اشاره شد.

دوم، در مسیر کدگذاری بعد از هر لایه up-sample، از نرمال‌سازی دسته‌ای استفاده شده‌است. نرمال‌سازی دسته‌ای به بهبود عملکرد، افزایش سرعت و پایداری شبکه‌های عصبی کمک می‌کند. خروجی حاصل از تابع نرمال‌سازی دسته‌ای، به یک لایه BConvLSTM داده می‌شود. یعنی در بخش اتصال پرش، به‌کمک لایه BConvLSTM، از پردازشی پیچیده‌تر از حالت عادی U-Net جهت الحاق نقشه‌ویژگی‌ها استفاده می‌کند. این لایه، خود از دو ConvLSTM برای پردازش داده‌های ورودی به مسیرهای رو به جلو و رو به عقب استفاده کرده و با مدیریت وابستگی داده‌ها در هر دو جهت، تصمیمی برای داده ورودی گرفته می‌شود.

با توجه به توضیحات بالا و بیان جزئیات و بررسی‌هایی که در فصل دوم در رابطه با این معماری انجام دادیم، در این پایان‌نامه قصد داریم از شبکه BCDU-Net تغییر یافته برای بخش‌بندی خودکار تصاویر سی‌تی‌اسکن ریه استفاده کنیم. درواقع، اسکلت و اساس شبکه یادگیری عمیق مورد استفاده ما در این پایان‌نامه، همان معماری BCDU-Net است. از طرفی از آن‌جا که پیش‌تر نیز مطرح شد، استفاده از شبکه‌های از پیش آموزش‌دیده مانند ResNet که از قبل روی مجموعه‌داده ImageNet

آموزش دیده‌اند، سرعت آموزش و قدرت تعمیم شبکه را بالاتر می‌برد. از طرف دیگر، چنان‌که در فصل دوم بیان شد، با توجه به آزمایش‌های انجام شده در شبکه‌های مختلف و مقایسه نتایج آن‌ها، استفاده از لایه‌ی کانولوشن به جای لایه‌ی Pooling ترجیح داده شده‌است؛ زیرا لایه‌های Pooling منجر به از دست دادن اطلاعات معنایی در تصاویر می‌شوند. در معماری ResNet فقط در لایه اول از Pooling استفاده شده و در لایه‌های بعدی برای حذف اطلاعات، از لایه کانولوشن با مقدار گام دو، در جهت استفاده شده‌است.

با توجه به تمام بحث‌ها و دلایل مطرح‌شده، تغییری که ما در شبکه BCDU-Net اعمال کردیم، استفاده از معماری ResNet34 به جای لایه‌های کانولوشنی معمولی، در بخش کدگذار این شبکه است. شکل ، بلوک دیاگرام الگوریتم ResNet استفاده شده در بخش کدگذار شبکه پیشنهادی را به همراه لایه‌های مورد استفاده، نشان می‌دهد.



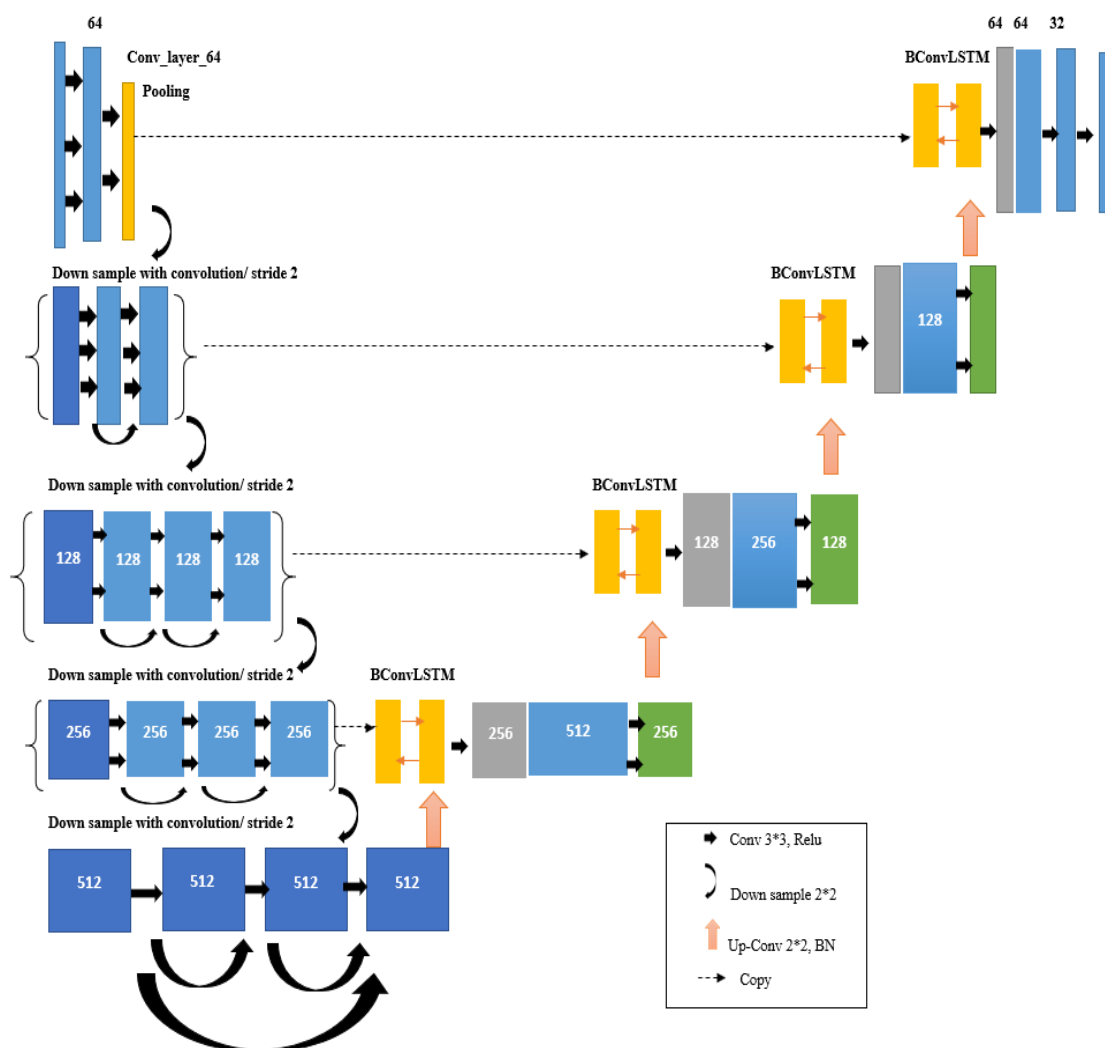
شکل ۴-۱۱ بلوک دیاگرام الگوریتم ResNet در بخش کدگذار شبکه پیشنهادی

شکل ۴-۱۱ بلوک دیاگرام بخش کدگذار شبکه پیشنهادی در این پایان‌نامه را نشان می‌دهد، چنان‌که گفته شد، در آخرین لایه از بخش کدگذار، به جای لایه‌های کانولوشنی ساده، از لایه‌های کانولوشنی متراکم استفاده شده‌است.

جزئیات این لایه و مابقی مسیر کدگشا که در این شبکه استفاده شده نیز به صورت کامل، در بخش ۳-

۴-۵ بیان شده‌است. با توجه به توضیحات ذکر شده، شبکه نهایی پیشنهادی مورد استفاده در این

پایان‌نامه، در شکل نشان داده شده‌است.



شکل ۴-۱۲ شبکه عمیق پیشنهادی

با توجه به شکل مسیر کدگذاری شبکه، از ساختار ResNet بهره برده و در آخرین لایه از این مسیر، از سه بلاک کانولوشنی و دو اتصال متراکم استفاده شده است. نقشه ویژگی‌های برگرفته از هر لایه در مسیر کدگذاری به کمک BConvLSTM با نقشه ویژگی‌های هر لایه از مسیر متناظر (کدگشایی)، الحاق می‌شوند. در ضمن نقشه ویژگی‌های هر لایه از مسیر کدگشایی، قبل از ورود به BConvLSTM، تحت تابع نرمال‌سازی دسته‌ای قرار می‌گیرند. این عمل با فلش‌های صورتی رنگ رو به بالا، مشخص شده است.

۴-۶ جمع‌بندی

در این فصل، به توضیح روش پیشنهادی پرداخته‌شد. گفتیم که روش پیشنهادی از سه بخش تولید برچسب تصاویر، پیش‌پردازش تصاویر و بخش‌بندی تصاویر با یادگیری عمیق تشکیل شده‌است. در بخش تولید برچسب‌ها، الگوریتمی نیمه‌خودکار مبتنی بر عملیات مورفولوژی و اصلاحات دستی ارائه‌شد؛ برای پیش‌پردازش تصاویر، کانال‌های جدیدی برای هر یک از تصاویر با عملیات مورفولوژی و لبه‌یابی طراحی و درنهایت برای مرحله بخش‌بندی، شبکه‌ی عصبی عمیقی با ترکیب دو ساختار BCDU-Net و ResNet34 ارائه گردید.

فصل ۵ : پیاده سازی و نتایج

۵-۱ مقدمه

در این تحقیق، یک روش جدید برای بخش‌بندی خودکار تصاویر سی‌تی‌اسکن ریه ارائه شد. این روش، شامل سه مرحله تولید ماسک (برچسب) تصاویر مرجع با روشی نیمه‌خودکار، پیش‌پردازش و بهبود تصاویر سی‌تی اصلی و نیز بخش‌بندی خودکار تصاویر سی‌تی مبتنی بر یادگیری عمیق است. در این روش، تمام بافت ریوی به‌طور دقیق و کامل از تصویر اصلی استخراج می‌شود. کارایی روش پیشنهادی بر روی تصاویر سی‌تی‌اسکن پایگاه داده LIDC-IDRI بررسی شده است. این مجموعه داده، بزرگ‌ترین مجموعه داده‌ی ایجادشده در این زمینه است که شامل ۱۰۱۸ اسکن ریه و ۲۶۶۹ ندول با اندازه بیش از ۳ میلیمتر بوده و از ۱۰۱۰ بیمار مختلف به‌دست آمده است. در این پایان‌نامه، ۱۷۱۴ تصویر از این مجموعه داده به‌همراه برچسب‌های تولیدشده‌ی متناظرشان مورد استفاده قرار گرفته است. در این فصل، به ارزیابی روش پیشنهادی بر اساس معیارهای رایج می‌پردازیم. با توجه به این‌که سیستم پیشنهادی از لحاظ کیفی و کمی قابل ارزیابی است، لذا هر یک از این دو حالت به‌صورت جداگانه مورد ارزیابی قرار می‌گیرند. در ادامه این فصل، به مقایسه نتایج به‌دست‌آمده با سیستم‌های پیشنهادی و سایر روش‌های موجود، می‌پردازیم.

۵-۲ معیارهای ارزیابی

دانشی که در مرحله یادگیری مدل تولید می‌شود، می‌بایست در مرحله ارزیابی، مورد تحلیل قرار گیرد تا بتوان ارزش آن را تعیین نمود و در پی آن کارایی الگوریتم یادگیرنده مدل را نیز مشخص کرد. این معیارها را می‌توان هم برای مجموعه داده‌های آموزشی در مرحله یادگیری و هم برای مجموعه رکوردهای آزمایشی در مرحله ارزیابی محاسبه کرد. در ادامه معیارهای ارزیابی عملکرد الگوریتم پیشنهادی در این پایان‌نامه را معرفی می‌کنیم:

- امتیاز شباهت دایس^۱: معیار Dice (DSC)، یکی از پرکاربردترین معیارهایی است که برای محاسبه

¹ Dice Similarity Score

شباهت دو تصویر استفاده می‌شود [۱۱۶]. این معیار، یکی از معیارهای اصلی برای ارزیابی نتایج بخش‌بندی تصاویر پزشکی نیز محسوب می‌شود. مقدار Dice، نشان می‌دهد که نتایج مدل بخش‌بندی، چقدر به ماسک حقیقی (برچسب) بافت، شباهت دارند. مقدار Dice مطابق رابطه زیر به صورت دو برابر اشتراک دو تصویر، تقسیم بر اجتماع دو تصویر، محاسبه می‌شود:

$$DSC = 2 \times \frac{|A \cap B|}{|A| + |B|} \quad \text{رابطه (۵-۱)}$$

در رابطه (۵-۱)، A بیانگر مجموعه پیکسل‌های تصویر قطعه‌بندی شده توسط مدل یادگیری عمیق پیشنهادی و B، بیانگر مجموعه پیکسل‌های تصویر ماسک است. هر چه شباهت دو مجموعه A و B به هم بیشتر باشد، مقدار DSC به یک، نزدیک‌تر می‌شود.

- معیار صحت^۱: این معیار، مشهورترین و عمومی‌ترین معیار محاسبه کارایی الگوریتم‌های طبقه‌بندی بوده و نشان می‌دهد که طبقه‌بند طراحی شده، چند درصد از کل مجموعه رکوردهای آزمایشی را به درستی پیش‌بینی کرده است [۱۱۷]. در واقع این معیار، بیانگر تعداد «پیش‌بینی‌های صحیح انجام شده» توسط دسته‌بند، تقسیم بر تعداد «کل پیش‌بینی‌های انجام شده» توسط همان دسته‌بند است. با این حال، این معیار به تنهایی، معیار مناسبی برای ارزیابی عملکرد یک دسته‌بند نیست. زمانی که کلاس‌های موجود در داده‌ها نامتوازن^۲ باشند (یعنی، تعداد داده‌های متعلق به کلاس (برچسب) خاص از کلاس‌های دیگر بسیار بیشتر باشد)، ممکن است سیستم با پدیده خاصی به نام تناقض صحت^۳ مواجه شود. در نتیجه این تناقض، مدل دسته‌بند به احتمال زیاد عملکرد بسیار خوبی در پیش‌بینی کلاس (برچسب) داده‌ها از خود نشان می‌دهد؛ زیرا، اکثریت داده‌ها تنها به یکی از کلاس‌ها تعلق دارند. در صورتی که چنین پدیده‌ای رخ دهد، بهتر است که معیارهای دیگری مثل یادآوری^۴ و دقت^۵ برای ارزیابی عملکرد سیستم در

^۱ Accuracy

^۲ Imbalanced

^۳ Paradox Accuracy

^۴ Recall

^۵ Precision

نظر گرفته شوند.

- معیار یادآوری: این معیار، بیان‌کننده نسبت «تعداد پیکسل‌های درست دسته‌بندی شده» در یک کلاس خاص، به تعداد کل پیکسل‌هایی است که باید در همان کلاس خاص دسته‌بندی شوند [۱۱۸]. مقدار بالا برای معیار یادآوری، بیانگر تعداد کم داده‌هایی است که به اشتباه، در آن کلاس خاص دسته‌بندی نشده‌اند. البته استفاده از این معیار به تنهایی، برای ارزیابی عملکرد سیستم درست نیست و باید در کنار معیار دقت، مورد استفاده قرار بگیرد.
- معیار دقت: این معیار، نسبت تعداد «پیش‌بینی‌های صحیح انجام شده» برای نمونه‌های یک کلاس خاص، به تعداد «کل پیش‌بینی‌ها» برای نمونه‌های همان کلاس خاص را (این تعداد، مجموع تمامی پیش‌بینی‌های صحیح و پیش‌بینی‌های نادرست را شامل می‌شود) ارزیابی می‌کند [۱۱۸]. مقدار بالا برای معیار دقت، بیانگر تعداد کم داده‌هایی است که به اشتباه، در کلاس خاص دسته‌بندی شده‌اند. شایان توجه است که معیار دقت، فقط برای مواردی ارزیابی می‌شود که در آن‌ها، مدل دسته‌بندی تعلق یک نمونه به یک کلاس خاص را پیش‌بینی کرده باشد.
- معیار امتیاز F1 (F1-score): این معیار در آمار برای ارزیابی عملکرد سیستم‌ها کاربرد دارد. همان‌طور که در رابطه (۵-۱) نشان داده شده، یک نوع میانگین بین پارامتر دقت و یادآوری است [۱۱۷]. این معیار، نسبت به معیار صحت، تصویر دقیق‌تری از نحوه عملکرد مدل دسته‌بند روی تمامی کلاس‌های موجود در داده‌ها ترسیم می‌کند.

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad \text{رابطه (۵-۲)}$$

برای درک بهتر این معیارها، ابتدا مفهوم ماتریس درهم‌ریختگی^۱ را شرح می‌دهیم:

ماتریس درهم‌ریختگی: در مباحث هوش مصنوعی، از این ماتریس برای تعیین مقدار شاخص‌های ارزیابی مانند دقت و صحت استفاده می‌شود [۱۱۹]. در شکل، ماتریس درهم‌ریختگی نمایش داده

¹ Confusion Matrix

شده است. این ماتریس، تعداد نمونه‌هایی که به درستی یا به اشتباه تشخیص داده شده‌اند را نشان می‌دهد.

کلاس واقعی	مثبت	مثبت صحیح (True Positive)	منفی غلط (False Negative)
	منفی	مثبت غلط (False Positive)	منفی صحیح (True Negative)

شکل ۵-۱ شکل کلی ماتریس درهم ریختگی برای یک مسئله دو کلاسه

با توجه به تعریف ماتریس درهم ریختگی، می‌توان معیارهای صحت، دقت و یادآوری را به صورت زیر تعریف کرد:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad \text{رابطه (۵-۳)}$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad \text{رابطه (۵-۴)}$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad \text{رابطه (۵-۵)}$$

۵-۳ نتایج آزمایش‌ها

در این بخش، به بررسی برخی اجراها و آزمایش‌ها و نتایج حاصل از مدل شبکه عصبی عمیق پیشنهاد شده می‌پردازیم. هم‌چنین در این بخش، نتایج را در دو حالت کمی و کیفی مورد تحلیل قرار می‌دهیم. ماتریس درهم ریختگی و به تبع آن نمودار ROC برای بررسی نتایج حاصل از آزمون مدل آموزش دیده و هم‌چنین عملکرد آن تحلیل شده و در نهایت، مدل ما تحت هریک از معیارهای ارزیابی مطرح شده در بخش قبل، با روش‌های دیگر مقایسه می‌شود. تاثیر پیش‌پردازش تصاویر خام اولیه و نیز استفاده از شبکه از پیش آموزش دیده در بخش کدگذار شبکه پیشنهادی در بهبود عملکرد مدل

پیشنهادی نیز در این بخش، مورد بررسی قرار می گیرند

در این پایان نامه، به صورت تصادفی، ۹۰ درصد از کل داده ها (۱۵۴۲ تصویر) به عنوان داده آموزشی و ۱۰ درصد از داده (۱۷۲ تصویر) نیز به عنوان داده آزمون در نظر گرفته شده است. همه پیاده سازی های ما در این پایان نامه، با استفاده از چارچوب کراس^۱ در زبان برنامه نویسی پایتون با پس زمینه تانسورفلو^۲ با نسخه ۳,۶ طراحی شده است. اجرای کدهای برنامه نیز بر روی GPU انجام شده است. شرایط محیط آزمایش به صورت زیر است:

گرافیک: NVIDIA GTX 1050 / پردازنده: CUDA Cores 640 / حافظه: ۸ گیگابایت

هم چنین تعداد تکرارهای الگوریتم برابر با ۵۰ و اندازه هر دسته^۳ برابر با ۳۲ در نظر گرفته شده است.

۵-۳-۱ دقت مدل پیشنهادی در مرحله آموزش و آزمایش

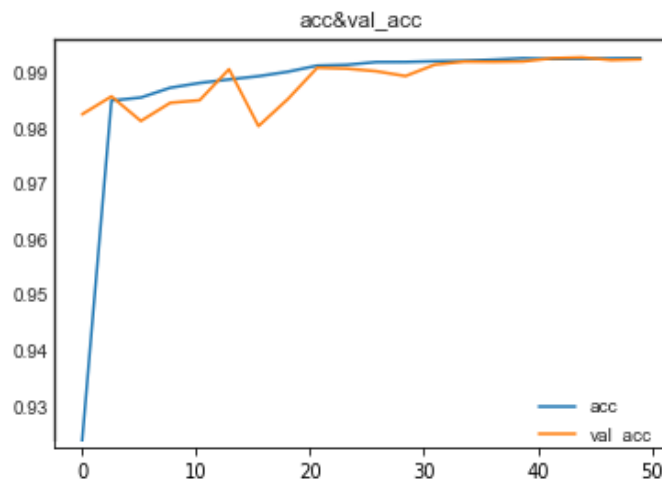
در مرحله آموزش، ورودی شبکه پیشنهادی شامل تصاویر مورد نظر و برچسب های آنها در فایل های جداگانه است. در شبکه های عصبی کانولوشن، هدف اصلی کاهش مقدار تابع loss از طریق اصلاح وزن ها است که این اصلاح وزن از طریق تابع بهینه ساز Adam انجام می شود که یک الگوریتم بهینه سازی بر اساس شیب است. هدف از بهینه سازی، تعیین متغیرهای طراحی است، به گونه ای که تابع هدف کمینه یا بیشینه شود. تابع انرژی شبکه توسط یک تابع softmax در سطح پیکسل که در ترکیب با تابع هزینه cross-entropy به نقشه خروجی اعمال می شود، محاسبه می شود. در مرحله آزمایش، ما یک تصویر به عنوان ورودی به شبکه داده و در خروجی آن، انتظار تولید طرح قطعه گذاری شده صحیح را داریم. پس از مرحله آموزش، مرحله آزمایش انجام می شود و دقت مدل در پیش بینی برچسب ها ارزیابی می شود. در مرحله آزمایش یک تصویر جدید به شبکه تغذیه شده و همه کارهایی که در فاز آموزش انجام می شود، در این مرحله هم انجام می شود. تنها تفاوت این مرحله با مرحله آموزش این است که دیگر فاز پس انتشار خطا وجود ندارد. صرفاً ورودی در مقادیر وزن ها ضرب

¹ Keras

² Tensorflow

³ batch size

می‌شوند. نمودار دقت شبکه را در فاز آموزش و آزمایش در شکل، نشان داده شده‌است.



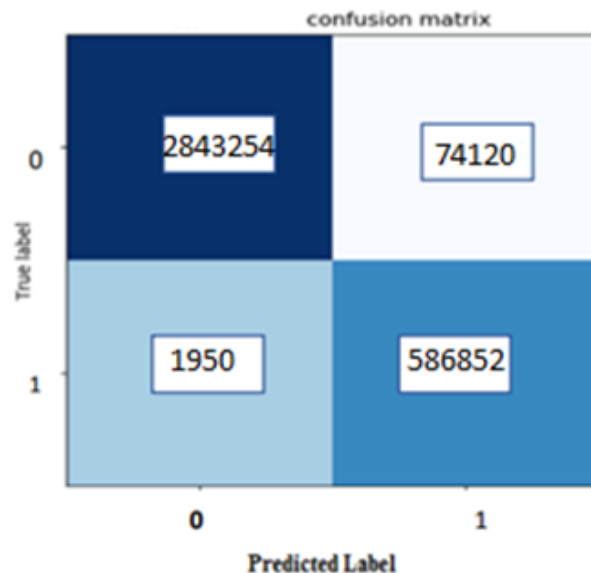
شکل ۵-۲ نمودار دقت در مرحله آموزش و آزمایش بر اساس تعداد تکرار

همان‌طور که در شکل مشخص شده، تعداد تکرار یا دوره، ۵۰ مرتبه بوده و طبق نمودار، مشاهده می‌شود که دقت فاز آموزشی، تقریباً تا دوره ۸، دارای سیر صعودی شدید بوده و پس از آن، روند تدریجی صعودی به‌خود می‌گیرد تا جایی که از تکرار ۴۰ تا ۵۰، دقت شبکه به بالای ۹۹ درصد می‌رسد و این، نشان‌دهنده آموزش مناسب شبکه است. در فاز آزمایش نیز، از تکرار صفر تا ۳۰، دارای سیر صعودی و نزولی است که نشان‌دهنده انتخاب وزن‌های نامناسب است. اما از تکرار سی‌ام به‌بعد، دقت به صورت تدریجی، سیر صعودی پایداری را طی کرده‌است. از تکرار ۳۵ تا ۵۰، نمودار دقت آموزش و آزمایش، با یکدیگر هم‌پوشانی داشته و روند آموزش شبکه به‌خوبی صورت می‌گیرد.

۵-۳-۲ ماتریس درهم ریختگی مدل پیشنهادی

در شکل، ماتریس درهم ریختگی نشان داده شده که بیانگر عملکرد الگوریتم پیشنهادی است. هر ستون از ماتریس، نمونه‌ای از مقدار پیش‌بینی‌شده را نشان می‌دهد. در صورتی که هر سطر، نمونه‌ای واقعی (درست) را دربر دارد. از این ماتریس برای تعیین مقدار شاخص‌های ارزیابی صحت، دقت و یادآوری استفاده می‌شود. مطابق روابط (۵-۳) تا (۵-۵)، به‌ترتیب مقدار محاسبه‌شده برای معیارهای

صحت، دقت و یادآوری، مقادیر ۹۷٪،۸۳، ۹۹٪،۹۳ و ۹۷٪،۴۵ به ازای مدل پیشنهادی به دست می‌آید. هم‌چنین مقدار F1-score، برابر ۹۸٪،۶۷ است.



شکل ۵-۳ ماتریس درهم ریختگی مدل پیشنهادی

۵-۳-۳ سطح زیر نمودار ROC برای مدل پیشنهادی

یکی از مهم‌ترین معیارها برای تعیین میزان کارایی یک دسته‌بند، میزان AUC^۱ است. AUC، نشان‌دهنده سطح زیر نمودار ROC^۲ است که هر چه این مقدار بزرگ‌تر باشد، کارایی نهایی دسته‌بند مطلوب‌تر ارزیابی می‌شود [۱۲۰]. نمودار ROC، روشی برای بررسی کارایی دسته‌بندها است. در واقع منحنی‌های ROC، منحنی‌های دو بعدی هستند که در آن‌ها، نرخ تشخیص صحیح دسته مثبت (TPR^۳) روی محور Y و به‌طور مشابه، نرخ تشخیص غلط دسته منفی (FPR^۴) روی محور X رسم می‌شوند. مقدار AUC برای یک دسته بند تصادفی، برابر ۰/۵ است. هم‌چنین بیشترین مقدار این معیار برابر یک بوده و برای دسته‌بند ایده‌آل رخ می‌دهد. معیار AUC برخلاف دیگر معیارهای تعیین کارایی دسته‌بندها، مستقل از آستانه تصمیم‌گیری دسته‌بند است. بنابراین این معیار، نشان‌دهنده

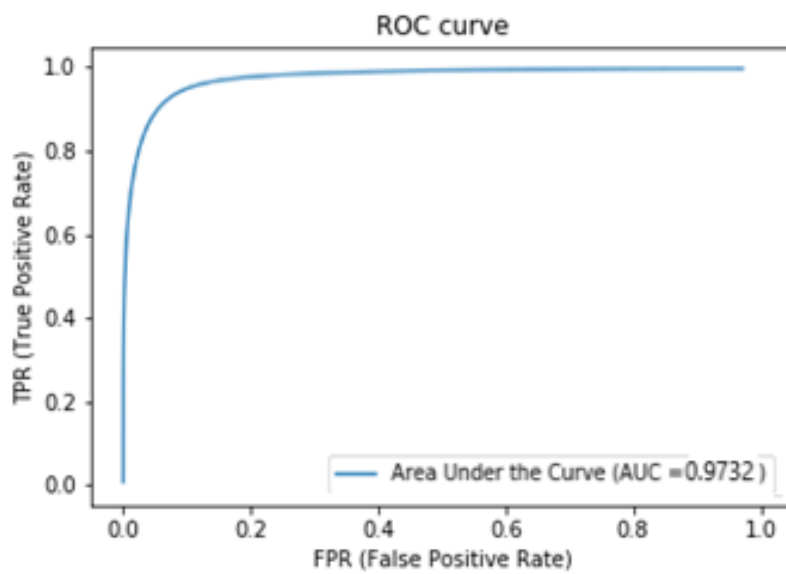
^۱ Area Under Curve

^۲ Receiver Operating Characteristic

^۳ True Positive Rate

^۴ False Positive Rate

میزان قابل اعتماد بودن خروجی یک دسته‌بند مشخص به ازای مجموعه داده‌های متفاوت است که این مفهوم توسط سایر معیارهای ارزیابی کارایی دسته‌بندها قابل محاسبه نیست. شکل ، نمودار ROC را برای مدل پیشنهادی، نشان می‌دهد.



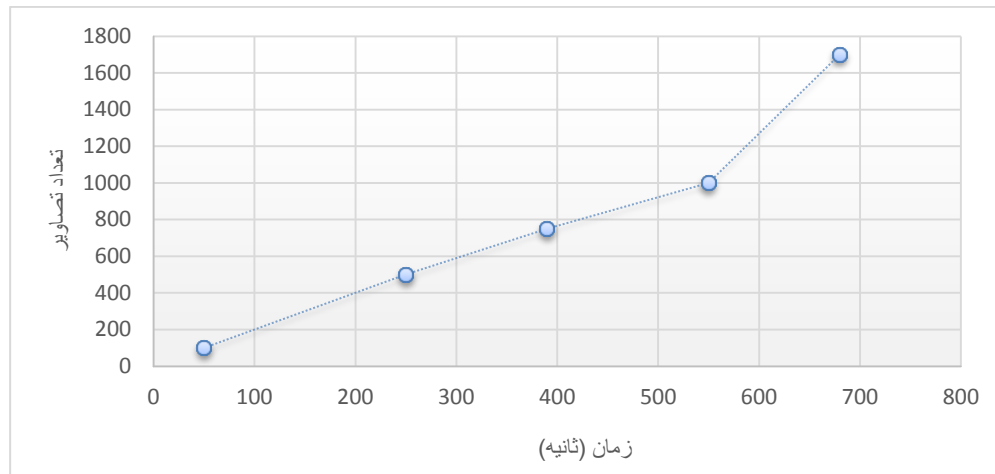
شکل ۵-۴ نمودار ROC مدل پیشنهادی

مطابق شکل ، مقدار AUC که برابر سطح زیر نمودار ROC است، برابر با ۰,۹۷۳۲ به دست آمده که این مقدار، بر کارا بودن عملکرد مدل پیشنهادی دلالت دارد.

۵-۳-۴ زمان اجرای الگوریتم تولید برچسب

همان‌طور که در فصل چهارم بیان شد، ابتدا از یک الگوریتم خودکار شامل عملیات مورفولوژی ریاضی برای تولید برچسب تصاویر سی‌تی استفاده کرده و بعد در صورت نیاز، اصلاحات دستی را اعمال کردیم. کار برچسب زدن هر تصویر سی‌تی توسط متخصصان رادیولوژیست، ساعت‌ها زمان می‌برد. در روش پیشنهادی، تولید همه ماسک‌ها بدون نیاز به اصلاح دستی، ۱۰ دقیقه به طول انجامید. با در نظر گرفتن بدترین شرایط و نیاز به اصلاح دستی و بررسی یکایک تصاویر تولیدشده توسط الگوریتم فوق، هر ماسک نیاز به زمان سه دقیقه‌ای برای تولید دارد. شکل ، زمان اجرای

اندازه‌گیری شده توسط روش پیشنهادی را بر روی بعد مجموعه داده از ۵۰ تا ۱۷۰۰ تصویر، نشان می‌دهد. علاوه بر این، با در نظر گرفتن تنها زمان محاسبات خالص بدون بارگذاری تصویر از دیسک، زمان اجرا تا ۲۰ درصد کاهش می‌یابد.



شکل ۵-۵ زمان اجرای الگوریتم خودکار تولید برچسب

با توجه به شکل با افزایش تعداد تصاویر، زمان اجرا به صورت خطی افزایش می‌یابد، در حالی که زمان مورد نیاز برای تجزیه و تحلیل تصاویر توسط یک متخصص با افزایش تعداد تصاویر به دلیل تاثیر عوامل انسانی مانند خستگی و غیره، با شدت زیادی افزایش خواهد یافت.

۴-۵ مقایسه روش پیشنهادی با روش‌های دیگر

همان‌طور که در فصل چهارم گفتیم، مدل پیشنهادی ترکیبی از ساختار شبکه عصبی عمیق ResNet و BCDU-Net است. استفاده از مدل ResNet دو دلیل عمده دارد. اولاً این شبکه از قبل بر روی ImageNet آموزش دیده‌است که منجر به دقت و تعمیم بالاتر مدل می‌شود، دوماً معماری ResNet فقط در لایه اول از Pooling استفاده شده‌است؛ حال آن که می‌دانیم لایه‌های Pooling، منجر به از دست دادن اطلاعات معنایی در تصاویر می‌شوند.

از طرفی، استفاده از ساختار BCDU-Net نیز به جای نسخه اصلی U-Net، دو دلیل دارد. دلیل اول، جلوگیری از یادگیری ویژگی‌های اضافی در هنگام آموزش شبکه و همچنین جلوگیری از اثر

محوشدگی گرادیان به کمک استفاده از لایه‌های کانولوشنی متراکم در آخرین لایه از مسیر کدگذاری بوده؛ دلیل دوم، استفاده از لایه BConvLSTM در مسیر اتصال پرش، برای الحاق نقشه‌ویژگی‌ها به روشی پیچیده‌تر از حالت الحاق ساده، است. خود BConvLSTM از دو ConvLSTM استفاده می‌کند تا داده‌های ورودی را در هر دو جهت مسیرهای رو به جلو و عقب پردازش کرده و سپس در برخورد با وابستگی‌های داده‌ها در هر دو جهت، تصمیم بگیرد. از آنجایی که در یک ConvLSTM استاندارد، تنها وابستگی‌های در جهت جلو پردازش می‌شوند؛ درحالی که تمام اطلاعات در یک توالی باید به‌طور کامل در نظر گرفته‌شود، لذا محتمل است که در نظرگرفتن وابستگی‌های متقابل با استفاده از لایه BConvLSTM در مسیر اتصال پرش شبکه، امری موثر در بهبود عملکرد شبکه پیش‌بینی‌کننده باشد. هم‌چنین قبل از ورود تصاویر و ماسک‌های متناظرشان به شبکه عمیق بیان‌شده، تغییراتی در تصاویر ورودی ایجاد کردیم. هر یک از سه کانال تصاویر سی‌تی اصلی را با سه کانال طراحی‌شده با استفاده از عملیات پیش‌پردازشی پر کردیم. کانال اول هر یک از تصاویر، خود تصویر اصلی سی‌تی بوده، کانال دوم با تصویر لبه‌یابی شده از تصویر اصلی، پر شده و در نهایت کانال سوم، با اعمال عملگر dilation، بعد از دودویی کردن بر روی تصویر اصلی، پر شده‌است. این فرآیند، موجب می‌شود شبکه عصبی عمیق، با نگاه به ویژگی‌هایی که در هر یک از کانال‌های جدید، پررنگ‌تر شده و مقایسه آن‌ها با برچسب متناظر این تصاویر، بهتر آموزش ببیند و نرخ خطای مثبت کاذب آن نیز کاهش یابد. در جدول ، مقایسه‌ای کمی به کمک معیارهای ارزیابی ارائه شده‌است. برای هر یک از معیارهای ارزیابی موجود در این جدول، بهترین مقدار، پررنگ¹ شده‌است.

¹ Bold

جدول ۵-۱ مقایسه روش پیشنهادی با سایر روش‌ها

روش‌ها و مراجع	دقت	یادآوری	امتیاز F1	صحت	امتیاز شباهت دایس
U-Net [۹۶]	۹۶,۱۱	۹۶,۳۴	۹۶,۲۲	۹۵,۱۸	۹۵,۰۲
RU-Net [۱۰۲]	۹۵,۵۲	۹۷,۲۱	۹۶,۳۵	۹۷,۱۵	۹۴,۹۳
ResNet34 U-Net [۱۰۵]	۹۷,۳۲	۹۸,۳۵	۹۷,۸۳	۹۶,۷۳	۹۵,۲۸
BCDU-Net [۱۰۶]	۹۹,۰۲	۹۸,۰۳	۹۸,۵۲	۹۷,۲۱	۹۶,۳۲
روش پیشنهادی (ورودی شبکه = تصاویر با کانال‌های پیش‌فرض)	۹۹,۱۲	۹۷,۰۱	۹۸,۰۵	۹۷,۵۸	۹۷,۱۵
روش پیشنهادی (ورودی شبکه = تصاویر با کانال‌های جدید)	۹۹,۹۳	۹۷,۴۵	۹۸,۶۷	۹۷,۸۳	۹۷,۳۱

با توجه به جدول ، متوجه می‌شویم که روش پیشنهادی در حالتی که تصاویر (ورودی شبکه عمیق)، با کانال‌های طراحی شده جدید پر شوند، بهترین عملکرد را داشته‌است. هم‌چنین می‌توان چند نتیجه مهم گرفت:

(۱) استفاده از ساختار ResNet34 در بخش کدگذار از شبکه U-Net، باعث بهبود شبکه می‌شود.

این امر، غالباً به دلیل از پیش آموزش دیدن این شبکه است.

(۲) مدل BCDU-Net در حالت کلی، بهتر از حالت به کارگیری شبکه ResNet در بخش کدگذار

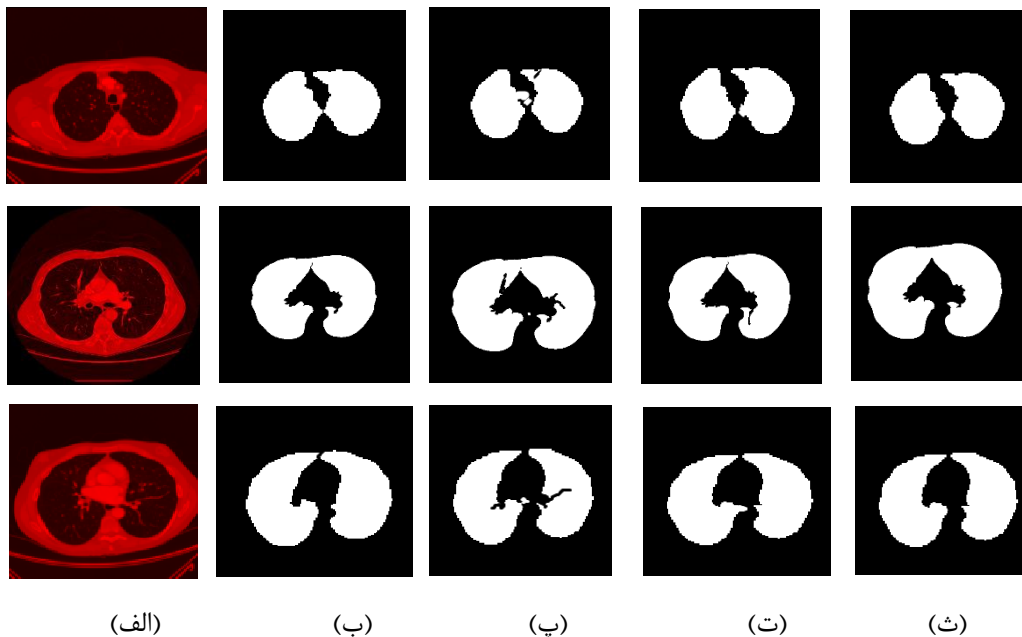
شبکه U-Net عمل کرده‌است. چنان‌که پیش‌تر مطرح شد، این به دلیل استفاده از

BConvLSTM و نیز مکانیزم لایه‌های کانولوشنی متراکم است.

(۳) ترکیب این دو شبکه، یعنی استفاده از شبکه ResNet در بخش کدگذار شبکه BCDU-Net،

امتیاز شباهت دایس بیشتری از حالت استفاده تک‌تک این شبکه‌ها کسب کرده‌است.

(۴) استفاده از تصاویر تحت کانال‌های طراحی شده در روش پیشنهادی، موجب بهبود نتایج کمی در همه معیارهای ارزیابی، نسبت به حالت استفاده کانال‌های پیش‌فرض ایجاد کرده‌است. برای نتیجه‌گیری کیفی نیز در شکل ، چند نمونه از تصاویر را به همراه برچسب اصلی آن‌ها نشان داده و با تصویر قطعه‌بندی تولید شده توسط برخی از الگوریتم‌های نام‌برده در جدول ، مقایسه کرده‌ایم.

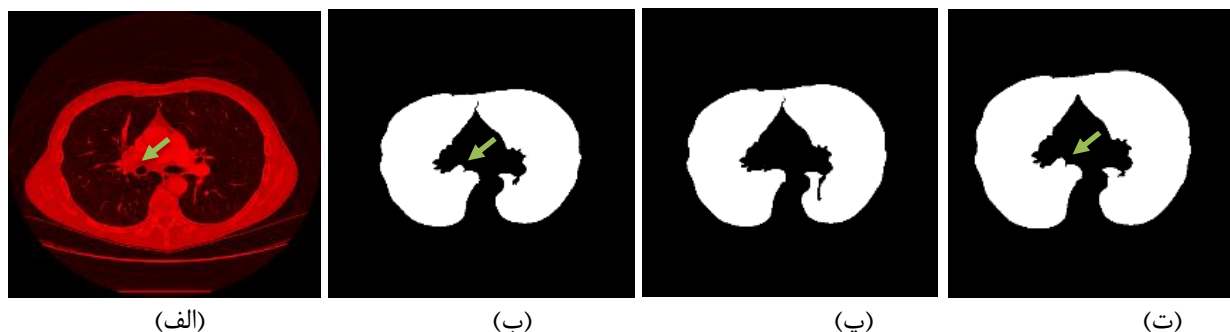


شکل ۵-۶ مقایسه کیفی نتایج قطعه‌بندی با الگوریتم‌های مختلف.

(الف): تصاویر اصلی (ب): تصاویر برچسب (ماسک دودویی) (پ): تصاویر قطعه‌بندی شده توسط U-Net (ت): تصاویر قطعه‌بندی شده توسط BCDU-Net (ث): تصاویر قطعه‌بندی شده توسط روش پیشنهادی

همان‌طور که در شکل مشاهده می‌شود، الگوریتم U-Net خیلی موفق عمل نکرده‌است که به علت اشکالات وارد شده به آن است. مدل BCDU-Net، بسیاری از کاستی‌های موجود در تصویر قطعه‌بندی شده توسط U-Net را از بین برده اما گاهی حالت تشخیص مثبت کاذب در آن نمایان شده‌است. مطابق شکل در روش پیشنهادی، این مشکل هم تا حد زیادی رفع شده و تصویر قطعه‌بندی شده نهایی، شباهت بسیار زیادی به تصویر ماسک متناظرش دارد. با این بررسی می‌توان نتیجه گرفت که ترکیب کانال‌های جدید برای تولید تصاویر اولیه و تاکید بر اجزایی مانند لبه‌ها و حذف موارد اضافی نامربوط در کانال‌های جدید پرشده، به بهبود پیش‌بینی شبکه عصبی عمیق بسیار کمک کرده‌است.

علاوه بر این، روش پیشنهادی، چالش مطرح شده در بخش ۲-۸ یعنی عدم حفظ ندول‌های متصل به دیواره ریه، را نیز حل کرده است. شکل، این موضوع را بهتر نشان می‌دهد.



شکل ۵-۷ نمایشی از رفع چالش بالا بودن نرخ مثبت کاذب در روش پیشنهادی
(الف): تصویر اصلی (ب): تصویر برچسب (ماسک دودویی) (پ): تصاویر قطعه‌بندی شده توسط BCDU-Net (ت):
تصویر قطعه‌بندی شده توسط روش پیشنهادی (بافت ریز ریوی، با فلش نشان داده شده است).

همان‌طور که در شکل مشاهده می‌شود، در تصویر اصلی بافت‌های ریز ریوی وجود دارند که جزئی از ریه بوده ولی در تصویر قطعه‌بندی شده در روش‌های دیگر، به‌عنوان بافت غیر ریوی در نظر گرفته می‌شوند. این امر، یکی از مواردی است که موجب افزایش نرخ کاذب مثبت در این روش‌ها شده است. در شکل نیز ندول متصلی به دیواره ریه در تصویر اصلی مشاهده می‌شود که باید جزء بافت ریوی در نظر گرفته شود؛ در حالی که اکثر روش‌ها مثل BCDU-Net، دچار مشکل ناتوانی در بخش‌بندی ندول‌های متصل به دیواره ریه هستند. این دو چالش در روش پیشنهادی، حل شده‌اند.



(الف)

(ب)

(پ)

(ت)

شکل ۵-۸ نمایشی از رفع چالش غفلت از ندول‌های متصل به دیواره ریه در روش پیشنهادی
(الف): تصویر اصلی (ب): تصویر برچسب (ماسک دودویی) (پ): تصاویر قطعه‌بندی شده توسط BCDU-Net (ث):
تصویر قطعه‌بندی شده توسط روش پیشنهادی (دور مکان ندول متصل به دیواره ریه، دایره کشیده شده‌است).

۵-۵ جمع‌بندی

در این فصل، به بررسی نتایج حاصل از پیش‌پردازش تصاویر ورودی و شبکه یادگیری عمیق پرداخته شد. با انجام پیاده‌سازی و آزمایش‌ها، دیدیم که استفاده از کانال‌های جدید طراحی‌شده برای هر یک از تصاویر CT ورودی، منجر به بهبود در روند آموزش و نتایج حاصل در شبکه یادگیری عمیق خواهد شد. شبکه عمیق مورد استفاده در این پایان‌نامه، با ترکیبی از الگوریتم‌های BCDU-Net و ResNet34 پیاده‌سازی شده‌است. استفاده از الگوریتم ResNet34 در بخش کدگذار BCDU-Net، به دلیل استفاده از وزن‌های بهینه‌ای که در پایگاه ImageNet ذخیره شده، نسبت به الگوریتم‌های دیگر دارای سرعت بیشتر و خطای کمتری است. در مجموع، با ترکیب این دو شبکه عمیق و استفاده از تصاویر پیش‌پردازش‌شده برای ورودی شبکه، دیدیم که روش ارائه‌شده علاوه بر کسب دقت و امتیاز شباهت دایس بالاتر، تعداد خطای مثبت کمی نیز دارد. این افزایش دقت، در مقایسه با روش‌های دیگر بهترین بود. هم‌چنین دیدیم با استفاده از الگوریتم خودکار استفاده‌شده در بخش تولید برچسب برای دادگان، زمان اجرا نسبت به تولید دادگان توسط متخصص، بسیار کمتر است و این امر، یکی از مزیت‌های مهم این روش محسوب می‌شود.

فصل ۴ : نتیجه گیری و پیشنهاد برای کاربرهای آتی

۱-۶ مقدمه

در این فصل، ابتدا الگوریتم ارائه شده در این پایان نامه را به طور مختصر توضیح داده و جمع بندی می کنیم. سپس پیشنهادهایی برای کارهای آتی در این حوزه ارائه خواهیم داد.

۲-۶ جمع بندی و نتیجه گیری

الگوریتم های بخش بندی تصاویر پزشکی در سیستم های CAD، به عنوان اولین گام در پردازش خودکار تصاویر جهت تحلیل و ارزیابی آنها به شمار آمده و عملکرد نهایی سیستم های CAD، به دقت و حساسیت در مرحله بخش بندی بستگی دارد. در فصل دوم بیان کردیم که مشکلات مختلفی مثل کیفیت متنوع تصاویر که ناشی از استفاده از انواع دستگاه های تصویربرداری با پروتکل های غیریکسان بوده، بر سر راه الگوریتم های بخش بندی موجود قرار دارد. آنچه که کار بخش بندی را چالش برانگیز می کند، وجود بافت هایی در اطراف ناحیه ریه است که در واقع جزو بافت ریوی بوده اما اکثر روش های بخش بندی، آن را نادیده گرفته و هنگام قطعه بندی ریه، آن را جزء نواحی غیرریوی می شمارند. از طرفی، چالش بعدی مربوط به سختی در تفسیر یک تصویر سی تی ریه به صورت دستی و توسط رادیولوژیست است. وجود ساختارهای اضافی و روی هم قرار گرفتن برخی ساختارها در یک تصویر رادیوگرافی یا توموگرافی ریه، باعث پیچیده شدن تفسیر و تحلیل این تصاویر می شود؛ چنان که حتی رادیولوژیست های باتجربه هم در تشخیص بعضی موارد از جمله تشخیص ندول های کوچک ریوی، دچار مشکل می شوند. علاوه بر این، تصاویر سی تی اسکن بسیار زیاد هستند و همین امر باعث اتلاف وقت متخصصان و خستگی آنها شده و دقت تشخیص را پایین می آورد.

در فصل چهارم از این پایان نامه، یک روش جدید برای بخش بندی خودکار تصاویر سی تی اسکن ریه ارائه شد. این روش، شامل سه مرحله تولید ماسک (برچسب) تصاویر مرجع با روشی نیمه خودکار، پیش پردازش و بهبود تصاویر سی تی اصلی و نیز بخش بندی خودکار تصاویر سی تی مبتنی بر یادگیری عمیق بود. در فصل پنجم، دیدیم که قسمت تولید ماسک به صورت نیمه خودکار برای تصاویر اولیه،

چالش دوم مطرح شده در پاراگراف قبل را مرتفع کرده است. هم‌چنین دیدیم که استفاده ترکیبی از ساختار ResNet34 و BCDU-Net در بخش شبکه یادگیری عمیق معرفی شده، با دقت و سرعت بالایی قادر به بخش‌بندی دقیق از بافت ریه بود. در ادامه تاثیر وجود کانال‌های جدید طراحی شده در تصاویر ورودی شبکه را نیز بر کاهش چشم‌گیر میزان مثبت کاذب و رسیدن به ضریب دایس بالاتر مشاهده کردیم. در نهایت دیدیم این شبکه عصبی عمیق با ورودی‌های بهبودیافته، علاوه بر دارا بودن مزیت‌های بیان شده، قادر به رفع چالش مطرح شده در بخش ۲-۸ نیز بود.

۳-۶ پیشنهاد برای کارهای آتی

یکی از کارهایی که می‌توان در آینده به آن پرداخت، استفاده از تصاویر سه بعدی سی‌تی ریه برای بخش‌بندی است. یکی دیگر از کارهایی ممکن به بهبود نتایج این پایان‌نامه کمک کند، استفاده از ترکیبی از سایر شبکه‌های یادگیری عمیق با یک‌دیگر است. هم‌چنین می‌توان در صورت نیاز، از روش‌های افزایش داده نیز جهت کسب دقت‌های بالاتر بهره برد.

[1] Hossain, Mir Rayat Imtiaz, Imran Ahmed, and Md Hasanul Kabir. "Automatic lung tumor detection based on GLCM features." In Asian Conference on Computer Vision, pp. 109-121. Springer, Cham, 2014.

[۲] قیصری پرنیا؛ منصور فاتح و محسن رضوانی، ۱۳۹۶، "بخش بندی خودکار تصاویر سی تی اسکن ریه به کمک یادگیری تقویتی." سومین کنفرانس پردازش سیگنال و سیستم های هوشمند ایران، شاهرود، دانشگاه صنعتی شاهرود

[3] Saini, Sujata, and Komal Arora. "A study analysis on the different image segmentation techniques." International Journal of Information & Computation Technology 4, no. 14 (2014): 1445-1452.

[4] Sharma, Neeraj, and Lalit M. Aggarwal. "Automated medical image segmentation techniques." Journal of medical physics/Association of Medical Physicists of India 35, no. 1 (2010): 3.

[5] Hu, Shiyang, Eric A. Hoffman, and Joseph M. Reinhardt. "Automatic lung segmentation for accurate quantitation of volumetric X-ray CT images." IEEE transactions on medical imaging 20, no. 6 (2001): 490-498.

[6] Shikata, Hidenori, Geoffrey McLennan, Eric A. Hoffman, and Milan Sonka. "Segmentation of pulmonary vascular trees from thoracic 3D CT images." International journal of biomedical imaging 2009 (2009).

[7] Badrinarayanan, Vijay, Ankur Handa, and Roberto Cipolla. "Segnet: A deep convolutional encoder-decoder architecture for robust semantic pixel-wise labelling." arXiv preprint arXiv:1505.07293 (2015).

[8] Skourt, Brahim Ait, Abdelhamid El Hassani, and Aicha Majda. "Lung CT image segmentation using deep neural networks." Procedia Computer Science 127 (2018): 109-113.

[9] Skourt, Brahim Ait, Abdelhamid El Hassani, and Aicha Majda. "Lung CT image segmentation using deep neural networks." Procedia Computer Science 127 (2018): 109-113.

[10] Kulkarni, Anjali, and Anagha Panditrao. "Classification of lung cancer stages on CT scan images using image processing." In 2014 IEEE International Conference on Advanced Communications, Control and Computing Technologies, pp. 1384-1388. IEEE, 2014.

[11] Jithy, P. K., and Philomina Simon. "Review on Automatic Segmentation Techniques in Medical Images." (2017).

[12] Wang, Wenjia, Junxuan Chen, Jie Zhao, Ying Chi, Xuansong Xie, Li Zhang, and Xiansheng Hua. "Automated segmentation of pulmonary lobes using coordination-guided deep neural networks." In 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), pp. 1353-1357. IEEE, 2019.

[13] Aung, H. H., A. Sivakumar, S. K. Gholami, S. P. Venkateswaran, and Bapi Gorain. "An Overview of the Anatomy and Physiology of the Lung." In Nanotechnology-Based Targeted Drug Delivery Systems for Lung Cancer, pp. 1-20. Academic Press, 2019.

[14] Nasser, Ibrahim M., and Samy S. Abu-Naser. "Lung Cancer Detection Using Artificial Neural Network." International Journal of Engineering and Information Systems (IJEAIS) 3, no. 3 (2019): 17-23.

[15] Di Natale, Corrado, Antonella Macagnano, Eugenio Martinelli, Roberto Paolesse, Giuseppe D'Arcangelo, Claudio Roscioni, Alessandro Finazzi-Agro, and Arnaldo D'Amico. "Lung cancer identification by the analysis of breath by means of an array of non-selective gas sensors." Biosensors and Bioelectronics 18, no. 10 (2003): 1209-1218.

[16] R. Nall, "Medical News Today," 12 11 2018. [Online].

Available: <https://www.medicalnewstoday.com/articles/323648.php>. [Accessed 14 3 2019]

[17] Sun, Shanhui, Christian Bauer, and Reinhard Beichel. "Automated 3-D segmentation of lungs with lung cancer in CT data using a novel robust active shape model approach." IEEE transactions on medical imaging 31, no. 2 (2011): 449-460.

[18] Jameson, J. Larry. Harrison's principles of internal medicine. McGraw-Hill Education,, 2018.

[19] American Cancer Society's publication, Cancer Facts & Figures 2020, Available: http://www.cancer.org/research/cancerfactsstatistics/cancerfacts_figures2020. /accessed on 2nd November, 2020

- [20] Armato, Samuel G., Maryellen L. Giger, Catherine J. Moran, James T. Blackburn, Kunio Doi, and Heber MacMahon. "Computerized detection of pulmonary nodules on CT scans." *Radiographics* 19, no. 5 (1999): 1303-1311.
- [21] Schaefer-Prokop, C., and M. Prokop. "New imaging techniques in the treatment guidelines for lung cancer." *European Respiratory Journal* 19, no. 35 suppl (2002): 71s-83s.
- [22] MacDonald, Sharyn LS, and David M. Hansell. "Staging of non-small cell lung cancer: imaging of intrathoracic disease." *European journal of radiology* 45, no. 1 (2003): 18-30.
- [23] Yamamoto, Shinji, Ippei Tanaka, Masahiro Senda, Yukio Tateno, Takeshi Iinuma, Toru Matsumoto, and Mitsuomi Matsumoto. "Image processing for computer- aided diagnosis of lung cancer by CT (LSCT)." *Systems and Computers in Japan* 25, no. 2 (1994): 67-80.
- [24] Kalra, Mannudeep K., Michael M. Maher, Roy D'Souza, and Sanjay Saini. "Multidetector computed tomography technology: current status and emerging developments." *Journal of computer assisted tomography* 28 (2004): S2-S6.
- [25] Swierczynski, Piotr, Bartłomiej W. Papież, Julia A. Schnabel, and Colin Macdonald. "A level-set approach to joint image segmentation and registration with application to CT lung imaging." *Computerized Medical Imaging and Graphics* 65 (2018): 58-68.
- [26] Chen, Chii-Jen. "Image segmentation for lung lesions using ant colony optimization classifier in chest CT." In *International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pp. 283-289. Springer, Cham, 2017.
- [27] Alilou, Mehdi, Niha Beig, Mahdi Orooji, Prabhakar Rajiah, Vamsidhar Velcheti, Sagar Rakshit, Niyoti Reddy et al. "An integrated segmentation and shape- based classification scheme for distinguishing adenocarcinomas from granulomas on lung CT." *Medical physics* 44, no. 7 (2017): 3556-3569.
- [28] Sun, Baozhou, Deshan Yang, Dao Lam, Tiezhi Zhang, Thomas Dvergsten, Jeffrey Bradley, Sasa Mutic, and Tianyu Zhao. "Toward adaptive proton therapy guided with a mobile helical CT scanner." *Radiotherapy and Oncology* 129, no. 3 (2018): 479-485.
- [29] Varma, Dandu Ravi. "Managing DICOM images: Tips and tricks for the radiologist." *The Indian journal of radiology & imaging* 22, no. 1 (2012): 4.
- [30] Graham, Richard NJ, Richard W. Perriss, and Andrew F. Scarsbrook. "DICOM demystified: a review of digital file formats and their use in radiological practice." *Clinical radiology* 60, no. 11 (2005): 1133-1140.
- [31] Puech, Philippe A., Loic Bousset, Samir Belfkih, Laurent Lemaitre, Philippe Douek, and Régis Beuscart. "DicomWorks: software for reviewing DICOM studies and promoting low-cost teleradiology." *Journal of Digital Imaging* 20, no. 2 (2007): 122.
- [32] Huff, Alyssa, Mitchell Reed, and Teresa Pitts. "Role of distention in swallow phase transition." *The FASEB Journal* 32, no. 1_supplement (2018): 913-5.
- [33] Abràmoff, Michael D., Paulo J. Magalhães, and Sunanda J. Ram. "Image processing with ImageJ." *Biophotonics international* 11, no. 7 (2004): 36-42.
- [34] Huang, H. K., and Osman M. Ratib. "Medical Imaging 2003: PACS and Integrated Medical Information Systems: Design and Evaluation." In *SPIE*, vol. 5033. 2003.
- [35] Sasidhar, B., D. R. Ramesh Babu, M. Ravishankar, and N. Bhaskar Rao. "Automated segmentation of lung regions using morphological operators in CT scan." *Int. J. Sci. Eng. Res* 4, no. 9 (2013).
- [36] Vanitha, U., P. Prabhu Deepak, N. PonNageswaran, and R. Sathappan. "Tumor detection in brain using morphological image processing." *Journal of Applied Science and Engineering Methodologies* 1, no. 1 (2015): 131-136.
- [37] Huidrom, Ratishchandra, Yambem Jina Chanu, and Kh Manglem Singh. "A Study on Lung Segmentation for the Detection and Diagnosis of Lung Cancer." *International Journal* 6, no. 11 (2016).
- [38] Karthikeyan, C., B. Ramadoss, and S. Baskar. "Segmentation algorithm for CT images using morphological operation and artificial neural network." *international journal of signal processing, image processing and pattern recognition* 5, no. 2 (2012): 115-122.
- [39] Parveen, S. Shaik, and C. Kavitha. "A review on computer aided detection and diagnosis of lung cancer nodules." *International Journal of Computers & Technology* 3, no. 3a (2012): 393-400.
- [40] Armato III, Samuel G., and William F. Sensakovic. "Automated lung segmentation for thoracic CT: Impact on computer-aided diagnosis1." *Academic Radiology* 11, no. 9 (2004): 1011-1021.
- [41] Leader, Joseph K., Bin Zheng, Robert M. Rogers, Frank C. Sciurba, Andrew Perez, Brian E. Chapman, Sanjay Patel, Carl R. Fuhrman, and David Gur. "Automated lung segmentation in X-ray

- computed tomography: development and evaluation of a heuristic threshold-based scheme1." *Academic radiology* 10, no. 11 (2003): 1224-1236.
- [42] Hu, Shiyong, Eric A. Hoffman, and Joseph M. Reinhardt. "Automatic lung segmentation for accurate quantitation of volumetric X-ray CT images." *IEEE transactions on medical imaging* 20, no. 6 (2001): 490-498.
- [43] Sun, Xuejun, Haibo Zhang, and Huichuan Duan. "3D computerized segmentation of lung volume with computed tomography." *Academic radiology* 13, no. 6 (2006): 670-677.
- [44] Brown, Matthew S., Michael F. McNitt-Gray, Nicholas J. Mankovich, Jonathan G. Goldin, John Hiller, Laurence S. Wilson, and D. R. Aberie. "Method for segmenting chest CT image data using an anatomical model: preliminary results." *IEEE transactions on medical imaging* 16, no. 6 (1997): 828-839.
- [45] Kemerink, Gerrit J., Rob JS Lamers, Bas J. Pellis, Han H. Kruize, and J. M. A. Van Engelshoven. "On segmentation of lung parenchyma in quantitative computed tomography of the lung." *Medical Physics* 25, no. 12 (1998): 2432-2439.
- [46] Gao, Qixin, ShengJun Wang, Dazhe Zhao, and Jiren Liu. "Accurate lung segmentation for X-ray CT images." In *Third International Conference on Natural Computation (ICNC 2007)*, vol. 2, pp. 275-279. IEEE, 2007.
- [47] Ukil, Soumik, and Joseph M. Reinhardt. "Smoothing lung segmentation surfaces in three-dimensional x-ray ct images using anatomic guidance1." *Academic radiology* 12, no. 12 (2005): 1502-1511.
- [48] Memon, Nisar Ahmed, Anwar Majid Mirza, and S. A. M. Gilani. "Segmentation of lungs from CT scan images for early diagnosis of lung cancer." In *Proceedings of world academy of science, engineering and technology*, vol. 14, pp. 228-233. 2006.
- [49] Korfiatis, Panagiotis, S. Skiadopoulos, P. Sakellaropoulos, C. Kalogeropoulou, and Lena Costaridou. "Combining 2D wavelet edge highlighting and 3D thresholding for lung segmentation in thin-slice CT." *The British journal of radiology* 80, no. 960 (2007): 996-1004.
- [50] Sonka, Milan, Vaclav Hlavac, and Roger Boyle. *Image processing, analysis, and machine vision*. Cengage Learning, 2014.
- [51] Davies, E. Roy. *Computer and machine vision: theory, algorithms, practicalities*. Academic Press, 2012.
- [52] Mitiche, Amar, and Jake K. Aggarwal. "Image segmentation by conventional and information-integrating techniques: a synopsis." *Image and Vision Computing* 3, no. 2 (1985): 50-62.
- [53] Johnson, L. A., J. D. Pearlman, C. A. Miller, T. I. Young, and K. R. Thulborn. "MR quantification of cerebral ventricular volume using a semiautomated algorithm." *American journal of neuroradiology* 14, no. 6 (1993): 1373-1378.
- [54] Haralick, Robert M., and Linda G. Shapiro. "Image segmentation techniques." *Computer vision, graphics, and image processing* 29, no. 1 (1985): 100-132.
- [55] Kundu, Amlan. "Local segmentation of biomedical images." *Computerized Medical Imaging and Graphics* 14, no. 3 (1990): 173-183.
- [56] Chow, C. K., and T. Kaneko. "Automatic boundary detection of the left ventricle from cineangiograms." *Computers and biomedical research* 5, no. 4 (1972): 388-410.
- [57] Talakoub, Omid, Javad Alirezaie, and Paul Babyn. "Lung segmentation in pulmonary ct images using wavelet transform." In *2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07*, vol. 1, pp. I-453. IEEE, 2007.
- [58] Brown, Matthew S., Michael F. McNitt-Gray, Jonathan G. Goldin, Robert D. Suh, James W. Sayre, and Denise R. Aberle. "Patient-specific models for lung nodule detection and surveillance in CT images." *IEEE Transactions on Medical Imaging* 20, no. 12 (2001): 1242-1250.
- [59] Croisille, Pierre, Miguel Souto, Maria Cova, Susan Wood, Yohannes Afework, Janet E. Kuhlman, and Elias A. Zerhouni. "Pulmonary nodules: improved detection with vascular segmentation and extraction with spiral CT. Work in progress." *Radiology* 197, no. 2 (1995): 397-401.
- [60] Abd-Elaziz, O. Fekry, M. Sharaf Sayed, and M. Ibrahim Abdullah. "Liver tumors segmentation from abdominal CT images using region growing and morphological processing." In *2014 international conference on engineering and technology (ICET)*, pp. 1-6. IEEE, 2014.
- [61] Singleton, H. Ross, and Gerald M. Pohost. "Automatic cardiac MR image segmentation using edge detection by tissue classification in pixel neighborhoods." *Magnetic resonance in medicine* 37, no. 3 (1997): 418-424.
- [62] Hu, Xiaoping, Noam Alperin, David N. Levin, Kim K. Tan, and Michel Menegeot. "Visualization of MR angiographic data with segmentation and volume- rendering techniques." *Journal of Magnetic Resonance Imaging* 1, no. 5 (1991): 539-546.

- [63] Tang, Jinshan, Steven Millington, Scott T. Acton, Jeff Crandall, and Shepard Hurwitz. "Surface extraction and thickness measurement of the articular cartilage from MR images using directional gradient vector flow snakes." *IEEE Transactions on Biomedical Engineering* 53, no. 5 (2006): 896-907.
- [64] Cline, Harvey E., C. L. Dumoulin, H. R. Hart Jr, William E. Lorensen, and S. Ludke. "3D reconstruction of the brain from magnetic resonance images using a connectivity algorithm." *Magnetic resonance imaging* 5, no. 5 (1987): 345-352.
- [65] Mesanovic, Nihad, Mislav Grgic, Haris Huseinagic, Matija Males, Emir Skejic, and Muamer Smajlovic. "Automatic CT image segmentation of the lungs with region growing algorithm." In *18th International Conference on Systems, Signals and Image Processing-IWSSIP*, pp. 395-400. 2011.
- [66] da Silva Felix, John Hebert, P. C. Cortez, M. A. Holanda, and R. C. S. Costa. "Automatic Segmentation and Measurement of the Lungs in healthy persons and in patients with Chronic Obstructive Pulmonary Disease in CT Images." In *IV Latin American Congress on Biomedical Engineering 2007, Bioengineering Solutions for Latin America Health*, pp. 370-373. Springer, Berlin, Heidelberg, 2007.
- [67] Sluimer, Ingrid, Mathias Prokop, and Bram Van Ginneken. "Toward automated segmentation of the pathological lung in CT." *IEEE transactions on medical imaging* 24, no. 8 (2005): 1025-1038.
- [68] Dehmeshki, Jamshid, Hamdan Amin, Manlio Valdivieso, and Xujiong Ye. "Segmentation of pulmonary nodules in thoracic CT scans: a region growing approach." *IEEE transactions on medical imaging* 27, no. 4 (2008): 467-480.
- [69] Gonzalez, Rafael C., Richard E. Woods, and R. Barry. "Masters." *Digital Image Processing, Third Edition. Journal of Biomedical Optics* 14, no. 2 (2009): 029901.
- [70] Canny, John. "A computational approach to edge detection." *IEEE Transactions on pattern analysis and machine intelligence* 6 (1986): 679-698.
- [71] Maini, Raman, and Himanshu Aggarwal. "Study and comparison of various image edge detection techniques." *International journal of image processing (IJIP)* 3, no. 1 (2009): 1-11.
- [72] Shin, Min C., Dmitry B. Goldgof, Kevin W. Bowyer, and Savvas Nikiforou. "Comparison of edge detection algorithms using a structure from motion task." *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 31, no. 4 (2001): 589-601.
- [73] Jain, Geetanjali, and Supreet Kaur. "A Review on Various Edge Detection Techniques in Distorted Images." *International Journal* 7, no. 5 (2017).
- [74] Mendonça, Ana Maria, Jorge Alves da Silva, and Aurélio Campilho. "Automatic delimitation of lung fields on chest radiographs." In *2004 2nd IEEE International Symposium on Biomedical Imaging: Nano to Macro (IEEE Cat No. 04EX821)*, pp. 1287-1290. IEEE, 2004.
- [75] Campadelli, Paola, Elena Casiraghi, and Diana Artioli. "A fully automated method for lung nodule detection from postero-anterior chest radiographs." *IEEE transactions on medical imaging* 25, no. 12 (2006): 1588-1603.
- [76] Hedberg, Hugo. "A survey of various image segmentation techniques." Dept. of Electrosience, Box 118 (2010).
- [77] Hedberg, Hugo. "A survey of various image segmentation techniques." Dept. of Electrosience, Box 118 (2010).
- [78] Kass, Michael, Andrew Witkin, and Demetri Terzopoulos. "Snakes: Active contour models." *International journal of computer vision* 1, no. 4 (1988): 321-331.
- [79] Makhanov, Stanislav S. "Active contours in medical image processing. Theory and applications." In *2013 5th International Conference on Knowledge and Smart Technology (KST)*, pp. xviii-xxv. IEEE, 2013.
- [80] Khalifa, Fahmi, Ahmed Elnakib, Garth M. Beache, Georgy Gimel'farb, Mohamed Abo El-Ghar, Rosemary Ouseph, Guela Sokhadze, Samantha Manning, Patrick McClure, and Ayman El-Baz. "3D kidney segmentation from CT images using a level set approach guided by a novel stochastic speed function." In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 587-594. Springer, Berlin, Heidelberg, 2011.
- [81] Itai, Yoshinori, Hyoungseop Kim, Seiji Ishikawa, Akiyoshi Yamamoto, and Katsumi Nakamura. "A segmentation method of lung areas by using snakes and automatic detection of abnormal shadow on the areas." *International Journal of Innovative Computing Information and Control* 3, no. 2 (2007): 277-84.
- [82] Mirderikvand, Negar, Marjan Naderan, and Amir Jamshidnezhad. "Accurate automatic localisation of lung nodules using Graph Cut and snakes algorithms." In *2016 6th International Conference on Computer and Knowledge Engineering (ICCKE)*, pp. 194-199. IEEE, 2016.

- [83] Ćurikovič, Roman, Kazufumi Kaneda, and Hideo Yamashita. "Dynamic contour: a texture approach and contour operations." *The Visual Computer* 11, no. 6 (1995): 277-289.
- [84] Caselles, Vicent, Francine Catté, Tomeu Coll, and Françoise Dibos. "A geometric model for active contours in image processing." *Numerische mathematik* 66, no. 1 (1993): 1-31.
- [85] Malladi, Ravi, James A. Sethian, and Baba C. Vemuri. "Shape modeling with front propagation: A level set approach." *IEEE transactions on pattern analysis and machine intelligence* 17, no. 2 (1995): 158-175.
- [86] Osher, Stanley, and James A. Sethian. "Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations." *Journal of computational physics* 79, no. 1 (1988): 12-49.
- [87] Sethian, James Albert. *Level set methods and fast marching methods: evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science*. Vol. 3. Cambridge university press, 1999.
- [88] Silva, Flávio Henrique Schuindt da. "Deep learning for Corpus Callosum segmentation in brain magnetic resonance images." (2018).
- [89] Shen, Dinggang, Guorong Wu, and Heung-Il Suk. "Deep learning in medical image analysis." *Annual review of biomedical engineering* 19 (2017): 221-248.
- [90] Patterson, Josh, and Adam Gibson. *Deep learning: A practitioner's approach*. " O'Reilly Media, Inc.", 2017.
- [91] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." In *Advances in neural information processing systems*, pp. 1097-1105. 2012.
- [92] Roth, Holger R., Chen Shen, Hirohisa Oda, Masahiro Oda, Yuichiro Hayashi, Kazunari Misawa, and Kensaku Mori. "Deep learning and its application to medical image segmentation." *Medical Imaging Technology* 36, no. 2 (2018): 63-71.
- [93] LeCun, Yann, Léon Bottou, Yoshua Bengio, and Patrick Haffner. "Gradient-based learning applied to document recognition." *Proceedings of the IEEE* 86, no. 11 (1998): 2278-2324.
- [94] Menegola, Afonso, Michel Fornaciali, Ramon Pires, Sandra Avila, and Eduardo Valle. "Towards automated melanoma screening: Exploring transfer learning schemes." *arXiv preprint arXiv:1609.01228* (2016).
- [95] He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep residual learning for image recognition." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770-778. 2016.
- [96] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." In *International Conference on Medical image computing and computer-assisted intervention*, pp. 234-241. Springer, Cham, 2015.
- [97] Sevastopolsky, Artem. "Optic disc and cup segmentation methods for glaucoma detection with modification of U-Net convolutional neural network." *Pattern Recognition and Image Analysis* 27, no. 3 (2017): 618-624.
- [98] Roy, Abhijit Guha, Sailesh Conjeti, Sri Phani Krishna Karri, Debodoot Sheet, Amin Katouzian, Christian Wachinger, and Nassir Navab. "ReLayNet: retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks." *Biomedical optics express* 8, no. 8 (2017): 3627-3642.
- [99] Norman, Berk, Valentina Pedoia, and Sharmila Majumdar. "Use of 2D U-Net convolutional neural networks for automated cartilage and meniscus segmentation of knee MR imaging data to determine relaxometry and morphometry." *Radiology* 288, no. 1 (2018): 177-185.
- [100] Apostolopoulos, Stefanos, Sandro De Zanet, Carlos Ciller, Sebastian Wolf, and Raphael Sznitman. "Pathological OCT retinal layer segmentation using branch residual u-shape networks." In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 294-301. Springer, Cham, 2017.
- [101] Gibson, Eli, Francesco Giganti, Yipeng Hu, Ester Bonmati, Steve Bandula, Kurinchi Gurusamy, Brian Davidson, Stephen P. Pereira, Matthew J. Clarkson, and Dean C. Barratt. "Automatic multi-organ segmentation on abdominal CT with dense v-networks." *IEEE transactions on medical imaging* 37, no. 8 (2018): 1822-1834.
- [102] Leclerc, Sarah, Erik Smistad, Thomas Grenier, Carole Lartizien, Andreas Ostvik, Frederic Cervenansky, Florian Espinosa et al. "RU-Net: A refining segmentation network for 2D echocardiography." In *2019 IEEE International Ultrasonics Symposium (IUS)*, pp. 1160-1163. IEEE, 2019.

- [103] Oktay, Ozan, Jo Schlemper, Loic Le Folgoc, Matthew Lee, Mattias Heinrich, Kazunari Misawa, Kensaku Mori et al. "Attention u-net: Learning where to look for the pancreas." arXiv preprint arXiv:1804.03999 (2018).
- [104] Gu, Zaiwang, Jun Cheng, Huazhu Fu, Kang Zhou, Huaying Hao, Yitian Zhao, Tianyang Zhang, Shenghua Gao, and Jiang Liu. "Ce-net: Context encoder network for 2d medical image segmentation." IEEE transactions on medical imaging 38, no. 10 (2019): 2281-2292.
- [105] Lau, Stephen LH, Xin Wang, Yang Xu, and Edwin KP Chong. "Automated Pavement Crack Segmentation Using Fully Convolutional U-Net with a Pretrained ResNet-34 Encoder." arXiv preprint arXiv:2001.01912 (2020).
- [106] Azad, Reza, Maryam Asadi-Aghbolaghi, Mahmood Fathy, and Sergio Escalera. "Bi-directional ConvLSTM U-net with Densley connected convolutions." In Proceedings of the IEEE International Conference on Computer Vision Workshops, pp. 0-0. 2019.
- [107] Huang, Gao, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q. Weinberger. "Densely connected convolutional networks." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4700-4708. 2017.
- [108] Tsang, Sik-Ho. "Review: Densenet-dense convolutional network (image classification)." en línea.[consulta: 7 abril 2019]. Disponible en: <https://towardsdatascience.com/review-densenetimage-classification-b6631a8ef803> (2019).
- [109] Song, Hongmei, Wenguan Wang, Sanyuan Zhao, Jianbing Shen, and Kin-Man Lam. "Pyramid dilated deeper convlstm for video salient object detection." In Proceedings of the European conference on computer vision (ECCV), pp. 715-731. 2018.
- [110] Ioffe, Sergey, and Christian Szegedy. "Batch normalization: Accelerating deep network training by reducing internal covariate shift." arXiv preprint arXiv:1502.03167 (2015).
- [111] Xingjian, S. H. I., Zhourong Chen, Hao Wang, Dit-Yan Yeung, Wai-Kin Wong, and Wang-chun Woo. "Convolutional LSTM network: A machine learning approach for precipitation nowcasting." In Advances in neural information processing systems, pp. 802-810. 2015.
- [112] Cui, Zhiyong, Ruimin Ke, Ziyuan Pu, and Yinhai Wang. "Deep bidirectional and unidirectional LSTM recurrent neural network for network-wide traffic speed prediction." arXiv preprint arXiv:1801.02143 (2018).
- [113] <https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI>
- [114] Habib, M. S. L. "A computer aided diagnosis system (CAD) for the detection of pulmonary nodules on CT scans." Systems and biomedical engineering Department, Faculty of Engineering, Cairo University, Giza, Egypt (2009).
- [115] Shivashankar, T. M., and S. B. Shivakumar. "Insights on Security Improvements and Implications of Artificial Intelligence in MANET." Communications 7 (2019): 1-9.
- [116] Dice, Lee R. "Measures of the amount of ecologic association between species." Ecology 26, no. 3 (1945): 297-302.
- [117] Chicco, Davide, and Giuseppe Jurman. "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation." BMC genomics 21, no. 1 (2020): 6.
- [118] Flach, Peter, and Meelis Kull. "Precision-recall-gain curves: PR analysis done right." In Advances in neural information processing systems, pp. 838-846. 2015.
- [119] Townsend, James T. "Theoretical analysis of an alphabetic confusion matrix." Perception & Psychophysics 9, no. 1 (1971): 40-50.
- [120] Muschelli, John. "Roc and auc with a binary predictor: a potentially misleading metric." Journal of Classification (2019): 1-13.

Abstract

Lung CT image segmentation is a key process and the first most important step in many applications such as lung cancer detection systems. This is a very challenging problem because of similar densities in the pulmonary structures, different scanners and scanning protocols. Some of the existing segmentation methods are semi - automatic and dependent on human factor or lack proper accuracy. Next disadvantages to many of these methods is high false positive rate. In recent years, deep learning-based approaches achieved state-of-the-art performance in medical image segmentation. Among existing networks, the U-Net has been great success in this field. In this paper, we propose a deep neural network to perform automatic lung CT image segmentation process. In the proposed method, first several extensive preprocessing techniques are applied to raw CT images, then ground truths corresponding to these images, have extracted by morphological operation and a few manually re-form. Finally, all prepared images with their corresponding ground truth fed to the deep neural network. Neural network architecture is a modified U-Net (BCDU-Net) in which the encoder is replaced with a pre-trained ResNet-34 network. Instead of a simple concatenation in the skip connection, we employ BConvLSTM to combine the feature maps extracted from the corresponding encoding path and the previous decoding up-convolutional layer and also used densely connected convolutions in the last convolutional layer of the encoding path. Finally, the proposed method is tested on LIDC-IDRI dataset and achieved dice coefficient index of 97.31% that as compared to the best previous method (BCDU-Net), it has increased by one percentage.

Keywords: Segmentation, Lung, CT image, U-Net, ResNet-34, BConvLSTM



Shahrood University of
Technology

Faculty of Computer Engineering

M.Sc Thesis in Artificial Intelligence Engineering

Lung CT Image Segmentation Using Deep Learning

By: Yeganeh Jalali

Supervisor:
Dr. Mansoor Fateh

Advisor:
Dr. Mohsen Rezvani

October 2020