





دانشکده مهندسی کامپیوتر و فناوری اطلاعات
پایان نامه کارشناسی ارشد مهندسی هوش مصنوعی

تشخیص سرطان ریه به کمک یادگیری عمیق

نگارنده: محمد امین زاده

استاد راهنما

دکتر منصور فاتح

اساتید مشاور

دکتر مرتضی زاهدی

دکتر هدی مشایخی

شهریور ۱۴۰۰

شماره: ۷۹۵

تاریخ: ۱۴۰۳/۰۶/۰۳

ویرایش:

باسمه تعالی

فرمهای ارزشیابی پایان نامه کارشناسی ارشد
مربوط به ورودی‌های ۹۴ به بعد

مدیریت تحصیلات تکمیلی



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 آقای محمد امین زاده با شماره دانشجویی ۹۸۱۳۳۶۴ رشته مهندسی کامپیوتر گرایش هوش مصنوعی و رباتیکز تحت عنوان تشخیص سرطان ریه به کمک یادگیری عمیق که در تاریخ ۱۴۰۳/۰۶/۰۳ با حضور هیأت داوران در دانشگاه صنعتی شاهرود برگزار شد به شرح ذیل اعلام می‌گردد:

☒ الف) درجه عالی: نمره ۱۹-۲۰	☒ ب) درجه خیلی خوب: نمره ۱۸/۹۹-۱۸
☐ ج) درجه خوب: نمره ۱۶-۱۷/۹۹	☐ د) درجه متوسط: نمره ۱۴-۱۵/۹۹
☐ 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	منصور فاتح	دانشیار	
۳- استاد مشاور اول	مرتضی زاهدی	استادیار	
۴- استاد مشاور دوم	هدی مشایخی	استادیار	
۴- استاد داور اول	محسن رضوانی	دانشیار	
۵- استاد داور دوم	اسماعیل طحانیان	استادیار	
۶- نماینده تحصیلات تکمیلی	محسن فرهادی	مریی	

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 دید.

پدر و مادرم

کہ از نگاهشان صلابت

از رفتارشان محبت

و از صبرشان ایستادگی را آموختم.

خواہرم

کہ وجودش شادی بخش و صفتش مایہ آرامش من است.

محمد امین زاده

تیر ۱۴۰۰

سپاس، همواره زینده اوست که در آدمی میل به آموختن نهاد و پس سپاس شایسته نام آمان است که در لباسی از تار مهر و پود علم به من آموختند چگونه پندیشتم. سر تواضع بر آستان استادان گرانقدر که من را در این راه یاری کردند فرو می آورم و صمیمانه مراتب امتنان خود را به آمان تقدیم می دارم.

حال که با لطف خدای بزرگ کار تحقیق و نگارش این پایان نامه به اتمام رسیده است بر خود واجب می دانم من باب حدیث «من لم یشکر المخلوق لم یشکر الخالق» از کلمه عزیزان و بزرگوارانی که در طول دوران تحصیل و به ویژه در اجرای پایان نامه نهایت مساعدت را با اینجانب داشته اند، مراتب تقدیر و تشکر خود را اعلام نمایم.

استاد گرانقدر، جناب آقای دکتر منصور فاتح، استاد فرزانه ای که با درایت علمی خویش راهنمایی این پایان نامه را به عهده گرفتند و استاد راهنمای اینجانب بوده اند و با صبوری و دلسوزی در تمامی مراحل مرایای نموده اند. توفیق روزافزون این بزرگواران از درگاه ایزد منان مسألت دارم.

از دو استاد برجسته ی دانشکده، جناب آقای دکتر مرتضی زاهدی و خانم دکتر مهدی مشایخی که اینجانب را در به پایان رساندن این پایان نامه یاری نموده اند، کمال تشکر را می کنم.

از دوستان عزیزم آقایان امیررضا فاتح و مصطفی طاهری که در به پایان رساندن این اثر مرایای نمودند، صمیمانه قدردانی و تشکر می نمایم و خود را مدیون زحمات خالصانه شان می دانم.

تعمدنامه

اینجانب ... محمد امین زاده.... دانشجوی دوره کارشناسی ارشد (دکتری) رشته ... هوش مصنوعی... دانشکده مهندسی کامپیوتر و فناوری اطلاعات.... دانشگاه صنعتی شاهرود نویسنده پایان نامه تشخیص سرطان 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

سرطان یکی از بیماری‌های کشنده جهان است. یکی از دلایل تلفات زیاد این بیماری، تشخیص اشتباه پزشک یا تاخیر در تشخیص این بیماری بوده است. در بین تمام سرطان‌ها، تشخیص سرطان ریه به علت داشتن ساختار غیریکسان و ناهم‌هنگ، به سختی انجام می‌پذیرد. با پیشرفت علم و تکنولوژی، تشخیص این نوع بیماری می‌تواند از حالت سنتی خارج شود و با سیستم‌های خودکار و نیمه‌خودکار انجام شود. برای داشتن یک سیستم که بتواند بدون دخالت انسان بیماری را تشخیص دهد، می‌توان از تکنیک‌های یادگیری عمیق بهره برد. به طور کلی، سیستم‌های تشخیص سرطان ریه، شامل دو بخش اصلی بخش‌بندی و طبقه‌بندی است. خروجی قسمت بخش‌بندی، ورودی قسمت استخراج ویژگی و طبقه‌بندی است. مهم‌ترین مشکل قسمت بخش‌بندی ساختار غیریکسان و ناهم‌هنگ ریه است. برای از بین بردن این مشکل، روش‌های زیادی مطرح شده است که ایراد اغلب این روش‌ها، بالا بودن میزان مثبت کاذب است. در این تحقیق، برای تشخیص سرطان ریه، ابتدا بر روی تصاویر پیش‌پردازش‌هایی انجام شده است. این پیش‌پردازش‌ها به منظور فراهم کردن تصاویر خام ورودی و افزایش کیفیت آن‌ها برای آموزش شبکه یادگیری عمیق است. سپس با استفاده از شبکه عصبی BCDU-Net، تصاویر ریه بخش‌بندی شده‌اند. سپس با استفاده از شبکه عصبی کانولوشنی سه بعدی و با تغییرات بر روی لایه‌های این شبکه، استخراج ویژگی انجام شده است. در مرحله آخر، لایه تماماً متصل به شبکه XGBOOST متصل شده است تا نهایتاً طبقه‌بندی با دقت بالایی انجام پذیرد. با استفاده از این شبکه‌ها و تغییراتی که بر روی آن‌ها ایجاد کرده‌ایم، به دقت ۹۴/۹۷ درصدی دست پیدا کرده‌ایم و نسبت به کارهای پیشین، دقت کار افزایش داشته است.

کلمات کلیدی: بخش‌بندی، شبکه عصبی کانولوشنی، ریه، U-Net، ۳D CNN، تشخیص سرطان، یادگیری عمیق

فهرست مطالب

ج	فهرست جداول
ح	فهرست اشکال
۱	فصل ۱: مقدمه
۲	۱-۱ پیشگفتار.....
۳	۲-۱ هدف از تحقیق.....
۳	۳-۱ چالش‌ها و راه‌حل‌های آن.....
۵	۴-۱ سوال‌های تحقیق.....
۵	۵-۱ فرضیه‌های تحقیق.....
۶	۶-۱ نوآوری‌های تحقیق.....
۷	۷-۱ ساختار پایان‌نامه.....
۹	فصل ۲ ادبیات موضوع و کارهای پیشین
۱۰	۱-۲ مقدمه.....
۱۰	۲-۲ ساختار ریه.....
۱۰	۱-۲-۲ ریه.....
۱۰	۲-۲-۲ آناتومی ریه.....
۱۱	۳-۲-۲ درخت برونش.....
۱۳	۴-۲-۲ عملکرد ریه‌ها.....
۱۴	۳-۲ سرطان ریه.....
۱۴	۱-۳-۲ سرطان.....

۴-۲	انواع مختلف تصویربرداری پزشکی.....	۱۵
۴-۲-۱	سی تی اسکن یا توموگرافی کامپیوتری.....	۱۶
۵-۲	مقیاس ارزیابی تصاویر سی تی اسکن و قالب ذخیره ساز.....	۲۰
۵-۲-۱	مقیاس ارزیابی.....	۲۲
۶-۲	پیش پردازش.....	۲۳
۶-۲-۱	فیلترهای ریختشناسی ریاضی.....	۲۳
۷-۲	بخش بندی ریه.....	۲۷
۸-۲	مروری بر روشهای بخش بندی (کارهای پیشین).....	۳۰
۸-۲-۱	روشهای بخش بندی سنتی.....	۳۰
۹-۲	جمع بندی روشهای سنتی بخش بندی.....	۴۱
۱۰-۲	استخراج ویژگی و طبقه بندی.....	۴۲
۱۰-۲-۱	یادگیری ماشین.....	۴۲
۱۰-۲-۲	مقدمه ای بر یادگیری عمیق.....	۴۴
۱۱-۲	جمع بندی.....	۴۵
۴۷	فصل ۳ ساختارهای یادگیری عمیق	
۱-۳	مقدمه.....	۴۸
۲-۳	شبکههای عصبی کانولوشنی.....	۴۸
۳-۳	معماری شبکه عصبی کانولوشنی.....	۴۹
۱-۳-۳	لایه کانولوشن.....	۵۰
۲-۳-۳	لایه غیرخطی.....	۵۰
۳-۳-۳	لایه تجميع.....	۵۱
۳-۳-۴	لایه تماما متصل.....	۵۱

۴-۳ مفاهیم و الگوریتم‌ها در حوزه شبکه‌های عصبی کانولوشن ۵۲

۴-۳-۱ پیش‌آموزش ۵۲

۴-۳-۲ الگوریتم ResNet ۵۳

۴-۳-۳ الگوریتم U-Net ۵۴

۳-۵ بخش‌بندی تصاویر پزشکی مبتنی بر یادگیری عمیق ۵۶

۳-۵-۱ ساختار RU-Net ۵۷

۳-۵-۲ ساختار Attention U-Net ۵۸

۳-۶ استخراج ویژگی و طبقه‌بندی ۵۸

۳-۶-۱ شبکه عصبی کانولوشنی سه بعدی ۵۹

۳-۶-۲ تقویت گرادیان ۶۰

۳-۷ جمع‌بندی ۶۲

۶۳ فصل ۴ روش پیشنهادی

۴-۱ مقدمه ۶۴

۴-۲ پیش‌پردازش و آماده‌سازی داده‌ها ۶۴

۴-۲-۱ دودویی کردن تصویر ۶۴

۴-۲-۲ عملگر لبه‌یابی ۶۵

۴-۲-۳ عملیات گسترش ۶۵

۴-۳ مدل یادگیری عمیق پیشنهادی برای مرحله بخش‌بندی ۶۶

۴-۳-۱ ساختار شبکه BCDU-Net ۶۷

۴-۴ مدل یادگیری عمیق پیشنهادی برای مرحله طبقه‌بندی ۷۰

۴-۴-۱ لایه اول (ورودی تصویر) ۷۰

۴-۴-۲ لایه دوم (لایه کانولوشنی اول) ۷۱

۷۲	۳-۴-۴ لایه‌های سوم تا هفتم.....
۷۳	۴-۴-۴ لایه هشتم (لایه تمام متصل).....
۷۴	۵-۴-۴ لایه نهم (شبکه XGBOOST).....
۷۵	۵-۴ جمع‌بندی.....
۷۷	فصل ۵ پیاده‌سازی و نتایج
۷۸	۱-۵ مقدمه.....
۷۸	۲-۵ مجموعه‌دادگان استفاده شده در این پایان‌نامه.....
۷۸	۱-۲-۵ مجموعه دادگان LIDC-IDRI.....
۷۹	۲-۲-۵ مجموعه دادگان LUNA-۱۶.....
۸۰	۳-۵ معیارهای ارزیابی.....
۸۱	۱-۳-۵ حساست و تشخیص‌پذیری.....
۸۲	۲-۳-۵ ماتریس اختلاط یا درهم ریختگی.....
۸۷	۴-۵ نتایج آزمایش‌ها.....
۸۸	۱-۴-۵ دقت مدل پیشنهادی در مرحله آموزش و آزمایش.....
۸۹	۲-۴-۵ ماتریس درهم ریختگی مدل پیشنهادی.....
۹۰	۳-۴-۵ سطح زیر نمودار ROC برای مدل پیشنهادی.....
۹۱	۵-۵ مقایسه روش پیشنهادی با روش‌های دیگر.....
۹۴	۶-۵ جمع‌بندی.....
۹۵	فصل ۶ نتیجه‌گیری و پیشنهادها برای کارهای آتی
۹۶	۱-۶ مقدمه.....
۹۶	۲-۶ جمع‌بندی و نتیجه‌گیری.....
۹۷	۳-۶ پیشنهاد برای کارهای آتی.....

فهرست جداول

- جدول ۱-۲. مقادیر سی تی متناظر با بافت های گوناگون در واحدهانسفیلد [۲۶] ۲۳
- جدول ۲-۲. مقایسه روش های سنتی بیان شده برای بخش بندی ۴۱
- جدول ۱-۴. جزئیات هشت لایه شبکه کانولوشنی ۷۳
- جدول ۱-۵. نمونه برجسب گذاری (برجسب ۱ دارای سرطان و ۰ بدون سرطان) [۸۴] ۸۰
- جدول ۲-۵. مقایسه روش پیشنهادی با سایر روش ها ۹۲

فهرست اشکال

- شکل ۱-۱. نمونه‌ای از ریه‌ای سرطانی ۴
- شکل ۲-۱. ساختار ریه ۱۱
- شکل ۲-۲. آناتومی درخت برونش ۱۲
- شکل ۲-۳. ساختار دیافراگم ۱۳
- شکل ۲-۴. یک توده سرطانی ۱۴
- شکل ۲-۵. دستگاه اسکنر سی‌تی‌اسکن اسپیرال ۱۷
- شکل ۲-۶. دستگاه سی‌تی‌اسکن مدرن، که شامل محفظه تونلی، منبع اشعه ایکس و حسگرها می‌باشد. ۱۹
- شکل ۲-۶. نمایشی از ساختار یک فایل DICOM [۲۴] ۲۲
- شکل ۲-۸. مثالی از اعمال عملیات مورفولوژی در بخش‌بندی ریه [۲۹] ۲۷
- شکل ۲-۹. تصاویری از برش‌های قسمت‌های مختلف ریه مربوط به یک بیمار ۲۹
- شکل ۲-۱۰. شماتیک روش‌های رشد ناحیه‌ای [۵۰]. (الف): انتخاب بذر اولیه (ب): بررسی نقاط مجاور و انجام فرآیند بخش‌بندی ۳۵
- شکل ۲-۱۱. بخش‌بندی ریه به روش رشد ناحیه (الف): تصویر سی‌تی اولیه (ب): تصویر قطعه‌بندی شده به روش رشد ناحیه (توده چسبیده به دیواره ریه که در روش رشد ناحیه‌ای، در نظر گرفته نمی‌شوند، با فلش در تصاویر نشان داده شده‌اند). ۳۷
- شکل ۲-۱۲. شکل سمت راست ماسک سوبل افقی و شکل سمت چپ ماسک سوبل عمودی ۳۹
- شکل ۲-۱۳. نتیجه عملگر سوبل بر روی تصویر ۳۹
- شکل ۲-۱۳. نتیجه کنی بر روی تصویر ۴۰
- شکل ۳-۱. معماری ساده شبکه عصبی کانولوشنی [۷۷] ۴۹
- شکل ۳-۲. توابع غیرخطی در شبکه عصبی کانولوشن ۵۱
- شکل ۳-۳. معماری ساده شبکه تماماً متصل ۵۲
- شکل ۳-۴. معماری ساده شبکه resnet ۵۴
- شکل ۳-۵. معماری U-Net [۸۰] ۵۵
- شکل ۳-۶. ساختار بخش‌بندی تصویر ۵۶
- شکل ۳-۷. معماری شبکه RU-Net [۸۲] ۵۷
- شکل ۳-۸. معماری شبکه Attention U-Net [۸۳] ۵۸
- شکل ۳-۹. معماری cnn سه بعدی [۸۴] ۶۰
- شکل ۳-۱۰. معماری CNN- XGBoost ۶۲

- شکل ۴-۱: تصویر (الف)، تصویر اصلی و تصویر (ب)، تصویر دودویی شده است. ۶۵
- شکل ۴-۲: خروجی پیش‌پردازش عملگر لبه‌یابی. ۶۵
- شکل ۴-۳: تصویر (الف)، تصویر اصلی و تصویر (ب)، خروجی پیش‌پردازش عملیات گسترش. ۶۶
- شکل ۴-۴: تصویر (الف) بلوک‌های شبکه ResNet و تصویر (ب) بلوک‌های شبکه DenseNet. ۶۸
- شکل ۴-۵: شبکه BCDU-Net [۸۸]. ۶۹
- شکل ۴-۶: فیلترهای لایه کانولوشن سه بعدی. ۷۲
- شکل ۵-۱: تصاویر سی‌تی مربوط به یک بیمار از مجموعه داده LUNA ۱۶. ۸۰
- شکل ۵-۲: شکل کلی ماتریس درهم ریختگی [۹۲]. ۸۲
- شکل ۵-۳: منحنی مشخصه عملکرد سیستم. ۸۶
- شکل ۵-۴: معیارهای ارزیابی الگوریتم‌های هوش مصنوعی. ۸۷
- شکل ۵-۵: نمودار دقت در مرحله آموزش و آزمایش بر اساس تعداد تکرار. ۸۹
- شکل ۵-۶: ماتریس درهم ریختگی مدل پیشنهادی. ۹۰
- شکل ۵-۴: نمودار ROC مدل پیشنهادی. ۹۱
- شکل ۵-۸: تصویر ریه از پایگاه داده. ۹۳
- شکل ۵-۹: تصویر قطعه‌بندی شده توسط U-Net. ۹۳
- شکل ۵-۱۰: تصویر قطعه‌بندی شده توسط BCDU-Net. ۹۴

فصل ۱ : مقدمه

۱-۱ پیشگفتار

در جهان بیماری‌های خطرناک زیادی وجود دارند و مردم جهان با آن روبه‌رو هستند. هر کدام از این بیماری‌ها به شیوه خاص، بدن بیمار را درگیر می‌کنند و در بدترین حالت باعث مرگ بیمار می‌شوند. در بین این بیماری‌ها، سرطان یکی از خطرناک‌ترین بیماری‌ها است. سرطان با دلایل مختلف از جمله تاخیر در مراجعه به پزشک، تشخیص اشتباه پزشک معالج و غیره باعث مرگ بیمار می‌شود [۱]. تشخیص اشتباه پزشک، باعث مشکلاتی از جمله تاخیر در آغاز درمان، فراگیری بدن انسان از بیماری و مرگ بیمار می‌شود. با توجه به این مشکلات، تشخیص سریع و دقیق سرطان یکی از بزرگترین چالش‌های پیش‌روی پزشکان است. با توجه به این توضیحات، برای بیمار بسیار حیاتی است که بیماری وی به درستی و با سرعت تشخیص داده شود که این سرعت و دقت فقط متعلق به سیستمی خودکار است. پزشک و رادیولوژیست از تصاویر پرتونگاری کامپیوتری (CT)^۱ برای تحلیل، تفسیر و تشخیص سرطان ریه از بافت‌های ریه استفاده می‌کنند. با این وجود، در بیشتر موارد، دستیابی به نتیجه دقیق تشخیص بدون استفاده از ابزار پزشکی کار بسیار سخت و وقت‌گیر است. به همین دلیل استفاده از سیستم‌های کمک کامپیوتری (CAD)^۲ یک امر طبیعی و الزامی است. این نوع سیستم‌ها برای کمک به متخصصان طراحی شده است. هر چه این سیستم‌ها به صورت خودکار عمل کنند و عامل انسانی در آن دخیل نباشد، دقت بالاتری خواهند داشت و کمک بسیار بالاتری به تشخیص پزشکان می‌کنند [۲ و ۳].

در سیستم‌های کمک کامپیوتری تشخیص سرطان ریه، با پیمودن چندین مرحله، خروجی لازم برای پزشک به دست می‌آید. این مراحل به طور معمول پیش‌پردازش داده‌ها، تقسیم‌بندی ریه‌ها، تشخیص گره‌های نامزد و طبقه‌بندی می‌باشند. به طور خلاصه در مرحله پیش‌پردازش داده‌ها، کارهای لازم برای یکسان و یک‌دست کردن داده‌ها انجام خواهد شد. در قسمت بخش‌بندی ریه‌ها بافت ریه از قسمت‌های دیگر ریه جدا و مشخص می‌شود و بقیه تصویر حذف می‌شود. در مرحله تشخیص گره‌های نامزد از

^۱ Computed Tomography
^۲ CAD: Computer Aided Design

مدل‌های مختلف موجود استفاده شده است. در مرحله آخر از الگوریتم‌های یادگیری عمیق^۱ برای طبقه‌بندی استفاده می‌کنیم که در ادامه به طور کامل شرح داده خواهند شد [۴].

۱-۲ هدف از تحقیق

همان‌طور که در بخش‌های قبل گفته شد، سرعت و دقت در تشخیص سرطان ریه یکی از بزرگترین چالش‌ها در این حوزه می‌باشد. روش‌های بسیاری برای تشخیص سرطان ریه به کار برده شده است که یکی از آن‌ها، استفاده از شبکه‌های عصبی عمیق است. روش‌های یادگیری عمیق به تازگی توانسته‌اند نقش بسیار بزرگ و پررنگی در مباحث هوش مصنوعی، بینایی ماشین و قطعه‌بندی نواحی مختلف یک تصویر داشته باشند.

تکنیک‌های یادگیری عمیق به دلیل توانایی بالا در پردازش مقادیر حجیم داده‌ها به طور سریع و دقیق، کاربرد بسیار فراوانی در حوزه‌ها پزشکی از جمله قطعه‌بندی تصاویر پزشکی و طبقه‌بندی دارند [۵]. هدف کلی این تحقیق، پیاده‌سازی یک روش کاملاً خودکار با استفاده از دو شبکه عصبی عمیق است. الگوریتم اول مبتنی بر U-Net (شبکه‌ای کانولوشنی مختص بخش‌بندی تصاویر پزشکی) است که در آن تصاویر سی‌تی‌اسکن ریه به طور خودکار بخش‌بندی می‌شوند. الگوریتم دوم CNN است که وظیفه طبقه‌بندی و تشخیص نوع سرطان را دارد.

۱-۳ چالش‌ها و راه‌حل‌های آن

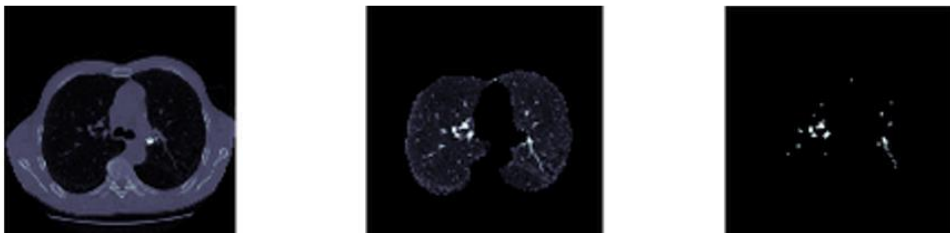
➤ پیچیدگی عکس‌های رادیوگرافی به دلیل داشتن ساختارهای اضافی

عکس‌های رادیوگرافی یا توموگرافی به دلیل داشتن ساختارهای اضافی (خون، آب، هوا و غیره) و تداخل برخی از این ساختارها با هم، پیچیده هستند. این موضوع باعث شده است که رادیولوژیست‌هایی که

سال‌های سال در این عرصه فعالیت دارند هم نتوانند به راحتی بخش‌های مختلف از جمله ندول‌های کوچک ریوی را تشخیص دهند و برای تشخیص دچار مشکل می‌شوند [۶]. شکل ۱-۱ این مشکل را واضح‌تر نشان می‌دهد.

➤ روشنایی بسیار کم توده‌های ریوی، شکل نامنظم و مکان پیچیده آناتومیک آن‌ها

یکی دیگر از مشکلاتی که در روش سنتی باعث به وجود آمدن خطا در تشخیص می‌شود، روشنایی بسیار کم توده‌های ریوی، شکل نامنظم و مکان پیچیده آناتومیک آن‌ها است. موارد ذکر شده، باعث به وجود آمدن مشکلات اساسی در تشخیص، جداسازی و آشکارسازی می‌شود. به دلیل حجم بالای تصاویر و تداخل برخی از لایه‌های ریه با هم، استفاده از روش سنتی یا تفسیر دستی آن‌ها بسیار سخت و طولانی است. این مشکلات باعث به وجود آمدن خطای انسانی در تشخیص بیماری می‌شود [۷].



شکل ۱-۰. نمونه‌ای از ریه‌ای سرطانی

➤ وجود فضای جستجو بسیار بزرگ در ریه

به‌طور کلی به دلیل وجود فضای جستجو بسیار بزرگ در ریه، باید مناطق اضافی ریوی حذف شوند تا تشخیص با دقت بالاتری صورت پذیرد. به دلیل محدودیت محاسبه، نمی‌توان کل تصویر را مستقیماً با استفاده از CNN‌های سه بعدی طبقه‌بندی کرد. برای طبقه‌بندی، باید مناطق احتمالی سرطان پیدا شوند

و سپس آن‌ها را طبقه‌بندی کرد. به همین دلایل از آستانه‌گذاری^۱ برای فیلتر کردن مناطق تاریک استفاده می‌شود.

با توجه به این چالش‌ها، نیاز به طراحی یک سیستم CAD است که بتواند تصاویر پزشکی را پردازش کند. این سیستم علاوه بر پردازش، قسمت‌های مشکوک به سرطان را شناسایی و خروجی لازم را برای مرحله‌ی بعد یعنی تشخیص نوع سرطان فراهم می‌کند. در نتیجه پزشک از سیستم CAD به منظور دستیابی به تشخیص بهتر و دقیق‌تر استفاده می‌کند [۷]. به طور کلی، دو مرحله مهم و حیاتی برای تشخیص نوع سرطان وجود دارد. اولین مرحله که در ابتدای کار قرار دارد، بحث بخش‌بندی تصاویر است. دومین مرحله که از خروجی مرحله قبل استفاده می‌کند، تشخیص نوع سرطان و طبقه‌بندی آن است [۸].

۱-۴ سوال‌های تحقیق

- آیا الگوریتم مورد استفاده در قسمت بخش‌بندی تصویر، قادر به تشخیص ندول‌های متصل به دیواره ریه می‌باشد؟
- آیا طبقه‌بندی انجام شده در پایان کار، می‌تواند به درستی نوع سرطان را تشخیص دهد؟
- آیا می‌توان شبکه‌ی عصبی دیگری را با لایه تماماً متصل پیوند داد؟

۱-۵ فرضیه‌های تحقیق

- تمام تصاویر سی‌تی‌اسکن که به عنوان ورودی و داده‌های این تحقیق می‌باشند، تصاویر سه بعدی هستند.
- تمام داده‌های استفاده شده در این تحقیق، دارای برچسب هستند و لیبل‌گذاری شده‌اند.
- برچسب ۱ به معنی دارای سرطان و ۰ به معنی بدون سرطان است.

^۱ thresholding

۶-۱ نوآوری‌های تحقیق

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

(۱) تغییر بر روی بخش‌های مختلف شبکه عصبی کانولوشنی سه بعدی.

در این قسمت، تغییراتی بر روی شبکه اصلی کانولوشنی سه بعدی در جهت بهبود عملکرد این شبکه

داده شده است و تعداد لایه‌های این شبکه به عدد ۹ رسانده شده است.

(۲) ترکیب دو الگوریتم یادگیری عمیق.

باتوجه به این که در جامعه امروزی، تلاش بر این است که سیستم‌ها خودکار شوند، ما در این تحقیق از

دو شبکه عصبی کاملاً خودکار استفاده کرده‌ایم. شبکه اول برای بخش‌بندی تصویر و شبکه دوم برای

تشخیص ویژگی و طبقه‌بندی و در مرحله آخر برای بالا بردن دقت، از یک شبکه دیگر در آخرین مرحله

استفاده کرده‌ایم.

(۳) استفاده از تقویت گرادیان و پیوند دادن آن با لایه تماماً متصل.

همان‌طور که در مورد قبلی ذکر شد، برای بالا بردن دقت طبقه‌بندی، از تقویت گرادیان و یکی از بهترین

شبکه‌های آن، استفاده کرده‌ایم.

(۴) افزایش دقت نسبت به روش‌های قبلی.

با توجه به تغییراتی که بر روی شبکه‌های اصلی داده شده است و با استفاده از تقویت گرادیان، دقت

تشخیص نسبت به تحقیقات دیگر بالاتر رفته است.

۱-۷ ساختار پایان نامه

این پایان نامه در شش فصل ارائه خواهد شد. در مقدمه توضیحاتی در مورد کلیت کار مطرح شد. در فصل بعدی، مروری بر کارهای پیشین در این حوزه انجام خواهد شد. بیشتر کارها اغلب در حوزه بخش بندی تصاویر بوده است. در فصل سوم، به معرفی یادگیری عمیق و الگوریتم های بخش بندی و طبقه بندی پرداخته خواهد شد. در این بخش چند نمونه از الگوریتم های یادگیری عمیق تشریح خواهد شد و در ادامه معماری اصلی شبکه عصبی کانولوشنی سه بعدی به طور کامل مورد بررسی قرار خواهد گرفت. در فصل چهارم، روش پیشنهادی تحقیق به طور کامل مطرح خواهد شد. در این فصل در ابتدا شبکه عصبی مورد استفاده در بخش بندی مورد بررسی قرار خواهد گرفت و در ادامه به توضیح تغییراتی لحاظ شده بر روی شبکه عصبی کانولوشنی سه بعدی پرداخته خواهد شد. در فصل پنجم پیاده سازی روش پیشنهادی و آزمایش ها و تفسیر نتایج آن ها به طور کامل بیان خواهد شد. در فصل ششم، نتیجه گیری و پیشنهادات آتی بیان خواهد شد.

فصل ۲ ادبیات موضوع و کارهای پیشین

۱-۲ مقدمه

در این فصل، در ابتدای کار توضیحی در مورد آناتومی ریه و سرطان داده می‌شود و در ادامه، انواع تصویربرداری پزشکی مورد بررسی قرار خواهند گرفت. در مرحله بعد، قالب ذخیره‌سازی و مقیاس ارزیابی تصاویر سی‌تی‌اسکن مطرح خواهد شد. در آخر یک جمع‌بندی از این فصل خواهیم داشت.

۲-۲ ساختار ریه

در این قسمت یکی از مهمترین قسمت‌های بدن انسان (ریه) مورد بررسی قرار می‌گیرد. این عضو در بدن وظیفه دریافت اکسیژن از محیط و انتقال آن به جریان خون را برعهده دارد. این اندام با بیش از ۶ میلیون تنفس در سال، بر عملکرد صحیح اندام‌های مختلف و سلامتی بدن ما بسیار موثر است [۹]. در ادامه به معرفی ریه، ساختار، عملکرد و بیماری‌هایی که بر ریه‌ها تأثیر می‌گذارند، پرداخته خواهد شد.

۱-۲-۲ ریه

ریه‌ها^۱ مرکز سیستم تنفسی در بدن هستند. هر سلول بدن برای زنده ماندن و عملکرد صحیح نیاز به اکسیژن دارد. از سویی دیگر سلول‌های بدن باید دی‌اکسید کربن تولیدی خود را از سلول‌ها خارج کنند. گاز دی‌اکسید کربن یک ماده دفعی است که از نتیجه واکنش‌های درون سلولی ایجاد می‌شود. ریه‌ها به طور اختصاصی برای تبادل این گازها در هنگام تنفس (دم و بازدم) طراحی شده‌اند [۹].

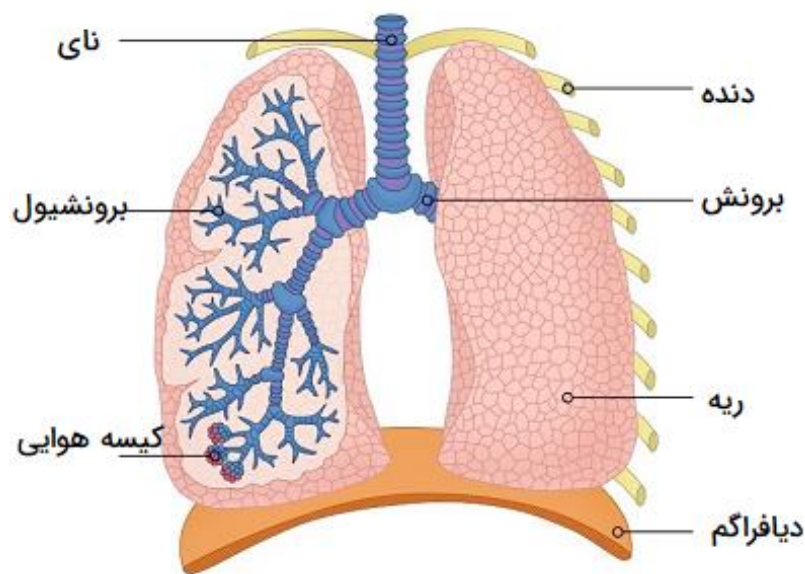
۲-۲-۲ آناتومی ریه

ریه‌ها یا شش‌ها در قفسه سینه، پشت دنده‌ها و در دو طرف قلب قرار دارند. آن‌ها تقریباً مخروطی شکل هستند و پایه‌های مسطح این مخروط‌ها، به دیافراگم متصل می‌شوند. اگرچه ریه‌ها به صورت جفت در

^۱ lungs

بدن ما وجود دارند، اما هر کدام از آن‌ها از نظر اندازه و شکل با هم یکسان نیستند. ریه راست از سه لوب تشکیل شده است در حالی که ریه چپ فقط دو لوب دارد.

ریه چپ دارای یک تورفتگی است که در این تورفتگی قلب جای می‌گیرد. به این تورفتگی شکاف قلبی نیز می‌گویند. ریه راست از نظر اندازه کوتاه‌تر است تا فضای کافی برای قرارگیری کبد در زیر آن فراهم شود. به طور کلی، ریه چپ دارای وزن و ظرفیت کمتری نسبت به ریه سمت راست می‌باشد [۱۰].



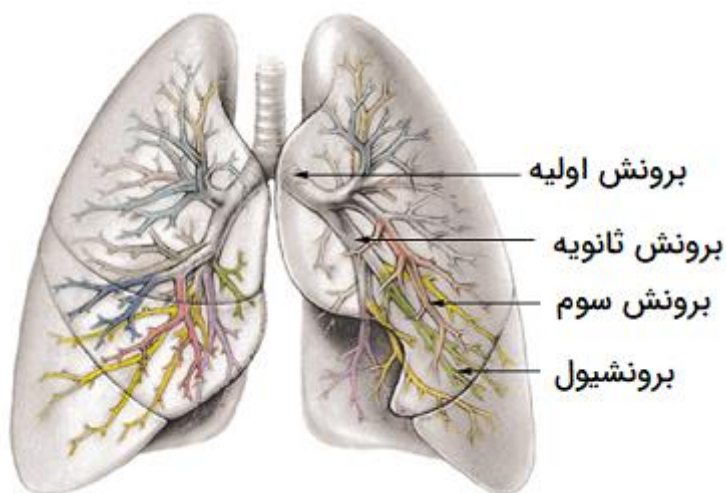
شکل ۱-۲. ساختار ریه

ریه‌ها توسط دو غشا احاطه شده‌اند که به عنوان پلور ریوی شناخته می‌شوند. لایه داخلی مستقیماً سطح بیرونی ریه‌ها را می‌پوشاند و لایه بیرونی به دیواره داخلی قفسه سینه متصل می‌شود. فضای بین این دو غشا نیز با مایع پلور پر شده است.

۳-۲-۲ درخت برونش

ریه‌ها از قسمت انتهایی نای شروع می‌شوند. نای لوله‌ای است که هوا را به داخل و خارج از ریه‌ها منتقل می‌کند. هر ریه، لوله‌ای به نام برونش یا نایژه دارد که به نای متصل می‌شود. ادامه مجاری هوایی

نای، یک ساختار Y وارونه را به نام برونش^۱ در سینه ایجاد می‌کند. این Y وارونه معمولاً درخت برونشی^۲ نامیده می‌شود.



شکل ۲-۲. آناتومی درخت برونش

برونش‌ها به برونش‌های کوچکتر و حتی لوله‌های کوچکتر به نام برونشیول^۳ منشعب می‌شوند و همانند شاخه‌های یک درخت، این لوله‌های ریز در تمام بخش‌های ریه‌ها کشیده می‌شوند. برخی از آن‌ها به قدری ریز هستند که ضخامت آن‌ها به اندازه ضخامت یک تار مو است. تقریباً در هر ریه سه هزار برونشیول وجود دارد [۱۰].

هر لوله برونشیول با خوشه‌ای از کیسه‌های هوایی کوچک به نام آلوئول^۴ (که به صورت جداگانه به عنوان آلوئولوس یا حبابک نامیده می‌شوند) به پایان می‌رسد. آن‌ها مانند دسته‌های انگور ریز یا بادکنک بسیار کوچک به نظر می‌رسند. در ریه‌ها حدود ۷۰۰ میلیون کیسه هوایی یا آلوئول وجود دارد [۱۰].

^۱ Bronchi

^۲ Bronchial Tree

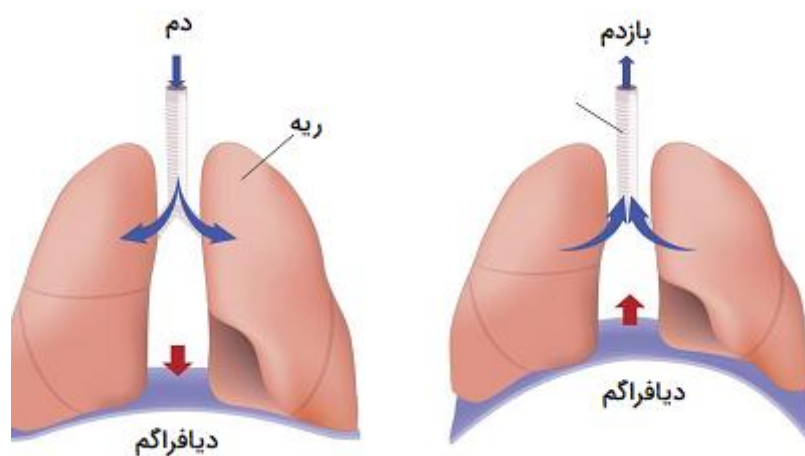
^۳ Bronchioles

^۴ Alveoli

۴-۲-۲ عملکرد ریه‌ها

نقش اصلی ریه‌ها دریافت اکسیژن از محیط بیرون و انتقال اکسیژن به جریان خون است. اکسیژن از ریه‌ها وارد خون شده و از آنجا به سایر بخش‌های بدن منتقل می‌شود. به منظور انجام صحیح مکانیسم تنفس، ریه‌ها می‌توانند از ساختارهای اطراف خود کمک بگیرند. از جمله این بخش‌ها می‌توان به عضله دیافراگم، عضلات بین دنده‌ای در قفسه سینه، عضلات شکم و حتی گاهی عضلات گردن اشاره کرد [۱۰].

دیافراگم^۱ ماهیچه‌ای است که در بخش بالایی خود حالت گنبدی دارد و در زیر ریه‌ها قرار گرفته است. بیشتر انرژی مورد نیاز برای مکانیسم تنفس توسط این ماهیچه تامین می‌شود. زمانی که دیافراگم منقبض می‌شود، به سمت پایین حرکت می‌کند و در این حالت فضای بیشتری در حفره قفسه سینه ایجاد شده و ظرفیت ریه‌ها را افزایش می‌دهد. با افزایش حجم حفره قفسه سینه، فشار داخل آن کاهش می‌یابد و هوا از طریق بینی، دهان و نای وارد ریه‌ها می‌شود [۱۰].



شکل ۳-۲. ساختار دیافراگم

زمانی که دیافراگم منبسط شده و به حالت استراحت خود باز می‌گردد، حجم ریه کاهش می‌یابد، زیرا فشار داخل حفره قفسه سینه بالا می‌رود و ریه‌ها هوا را خارج می‌کنند.

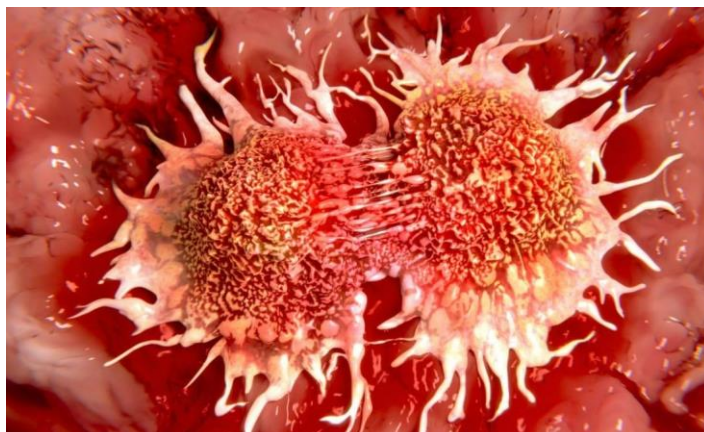
^۱ Diaphragm

۳-۲ سرطان ریه

در این بخش تعریف کوچکی از سرطان داده می‌شود و به طور خلاصه به سرطان ریه پرداخته می‌شود.

۱-۳-۲ سرطان

واژه سرطان به خانواده بزرگی از بیماری‌ها اطلاق می‌شود که شامل رشد غیر طبیعی سلول‌هاست. این سلول‌ها این قابلیت را دارند که به سایر نقاط بدن گسترش یابند و به آن حمله کنند [۱۱]. طی این بیماری، زیرمجموعه‌ای از نئوپلاسم‌ها شکل می‌گیرند. یک نئوپلاسم یا تومور درواقع دسته‌ای از سلول‌ها هستند که طی یک رشد غیر قابل کنترل به وجود آمده‌اند. به طور معمول، به تشکیل یک توده می‌انجامند، اما این توده می‌تواند در سرتاسر بدن گسترش یابد [۱۲].



شکل ۴-۲. یک توده سرطانی

این بیماری در هر قسمت از بدن مانند خون، پوست، سیستم گوارش، سیستم ادراری تناسلی، استخوان، ماهیچه، چشم و غیره ممکن است آغاز شود و در مراحل پیشرفته به سایر نواحی بدن گسترش یابد. بسته به این که توده تشکیل شده در کدام بخش از بدن باشد، علائم ظاهر شده نیز متفاوت خواهد بود. عوامل مختلفی در ابتلا به سرطان نقش دارند، از جمله عوامل ژنتیکی، مواد شیمیایی، پرتوها و غیره.

ریه یک عضو بزرگ بدن است. از این رو، توده سرطانی می‌تواند بدون این که علامتی به وجود آورد، مدت زیادی در ریه رشد کند. به همین دلیل احتمال زیادی دارد که دیر تشخیص داده شود. خستگی و سرفه دو علامتی هستند که ممکن است ناشی از سرطان ریه باشند. به طور معمول، در بیشتر موارد افراد از آن به سادگی می‌گذرند و آن را به علل دیگر ربط می‌دهند [۱۳].

گزارشات زیادی در راستای مرگ و میر سرطان ریه منتشر می‌شود، اما سازمان بهداشت جهانی در آخرین گزارشی که منتشر کرده است، خبر از کشته شدن بیش از ۳ میلیون نفر به دلیل داشتن سرطان ریه در سال را می‌دهد [۱۴]. ایالات متحده (US)^۱ گزارش‌هایی در سال‌های ۲۰۱۹ و ۲۰۲۰ منتشر کرد که خبر تکان دهنده‌ای در جهان بود. طبق این گزارش‌ها در سال ۲۰۱۹ بیش از ۱۱۰ هزار نفر زن و بیش از ۱۱۶ هزار نفر مرد بر اثر ابتلا به سرطان ریه جان خود را از دست داده‌اند [۱۵]. در سال ۲۰۲۰ این سرطان موجب مرگ بیش از ۱۳۵ هزار نفر شده است [۱۶].

۲-۴ انواع مختلف تصویربرداری پزشکی

در پزشکی مدرن، تکنیک‌های مربوط به تصویربرداری پزشکی پیشرفت چشمگیری داشته‌اند. امروزه توانایی دستیابی به اطلاعات مرتبط با بدن انسان و معاینات بالینی بسیار آسان‌تر شده است. در طول سال‌های گذشته، انواع مختلف تصویربرداری پزشکی به دنیا معرفی شده‌اند. هر کدام از این روش‌ها، مزایا و معایب خاص خود را دارند.

روش‌های تصویربرداری پزشکی مبتنی بر اشعه ایکس، شامل اشعه ایکس معمولی، توموگرافی کامپیوتری و ماموگرافی هستند.

^۱United States

تصویربرداری مولکولی، روش دیگری است که از آن معمولاً در پزشکی هسته‌ای استفاده می‌شود. تصویربرداری مولکولی بر پایه روش‌های مختلف جهت به تصویر کشیدن فرایندهای بیولوژیکی در سلول‌های زنده استوار است.

انواع دیگر تصویربرداری پزشکی شامل تصویربرداری رزونانس مغناطیسی (MRI) و سونوگرافی است. بر خلاف اشعه X، CT و تصویربرداری مولکولی، روش‌های MRI و سونوگرافی بدون تابش پرتوهای یونیزه کار می‌کنند. MRI از میدان‌های مغناطیسی قوی استفاده می‌کند که هیچ اثر بیولوژیکی برگشت ناپذیری در انسان باقی نمی‌گذارد. در سیستم‌های سونوگرافی تشخیصی نیز از امواج صوتی با فرکانس بالا برای تولید تصاویری از بافت‌های نرم و اندام‌های داخلی بدن استفاده می‌شود [۱۷].

۲-۴-۱ سی تی اسکن یا توموگرافی کامپیوتری

توموگرافی کامپیوتری (CT) از تکنولوژی پیشرفته اشعه ایکس برای کمک به تشخیص انواع بیماری‌ها و شرایط استفاده می‌کند. روش تصویربرداری سی تی اسکن، سریع، بدون درد، غیر تهاجمی و دقیق است. در موارد اضطراری، می‌تواند آسیب‌های داخلی و خونریزی را به اندازه کافی برای کمک به نجات جان افراد نشان دهد. [۱۸].

۲-۴-۱-۱ سی تی اسکن

توموگرافی کامپیوتری معمولاً به عنوان CT یا CAT شناخته می‌شود. سی تی اسکن، آزمایش پزشکی تشخیصی است که مانند اشعه ایکس سنتی، چندین تصویر یا تصاویر از داخل بدن تولید می‌کند. تصاویر مقطعی ایجاد شده در طی یک سی تی اسکن می‌تواند در چندین صفحه بازسازی شود. این نوع تصویر برداری، حتی می‌تواند تصاویر سه بعدی ایجاد کند. خروجی سی تی اسکن به صورت تصویر است و می‌توان بر روی یک مانیتور کامپیوتر مشاهده، بر روی فیلم یا چاپگر سه بعدی چاپ، یا به سی دی یا دی وی منتقل کرد [۱۹].

سی‌تی‌اسکن تصاویر اندام‌های داخلی، استخوان، بافت نرم و رگ‌های خونی را به طور کامل نشان می‌دهد. سی‌تی‌اسکن جزئیات بیشتری نسبت به اشعه X سنتی را فراهم می‌کند که به خصوص می‌توان به بافت‌های نرم و عروق خونی اشاره کرد [۲۰].

با استفاده از تجهیزات تخصصی برای ایجاد و تفسیر سی‌تی‌اسکن، رادیولوژیست‌ها می‌توانند به راحتی مشکلات مربوط به سرطان، بیماری‌های قلبی عروقی، بیماری‌های عفونی، آپاندیس، تروما و اختلالات اسکلتی عضلانی را تشخیص دهند [۲۱].



شکل ۵-۲. دستگاه اسکنر سی‌تی‌اسکن اسپیرال

۲-۴-۱- مزایای استفاده از سی تی اسکن

(۱) یکی از سریع‌ترین و با دقیق‌ترین ابزار برای بررسی قفسه سینه، شکم و لگن است، زیرا جزئیات دقیق و مقدماتی را در مورد انواع بافت ارائه می‌دهد.

(۲) برای بررسی بیماران مبتلا به آسیب‌های ناشی از تروما مانند یک تصادف خودرو موثر است.

(۳) در شرایط عادی، بهترین روش برای تشخیص سرطان در قفسه سینه، شکم، لگن، ریه، کبد، کلیه، تخمدان و پانکراس است. پزشک می‌تواند از تصویر سی‌تی‌اسکن، به منظور تایید حضور

تومور، اندازه‌گیری اندازه، تعیین محل دقیق، تعیین میزان درگیری آن با سایر بافت‌های نزدیک استفاده کند.

۴) سی‌تی‌اسکن، بدون درد، غیر تهاجمی و دقیق است.

۵) توانایی در تصویربرداری استخوان، بافت نرم و عروق خونی به طور همزمان.

۶) معاینات سی‌تی‌اسکن سریع و ساده هستند، به طوری که در موارد اضطراری، می‌تواند آسیب‌های داخلی و خونریزی را به طور دقیق در تصویر نشان دهد.

۲-۴-۱-۳ مزایا و معایب روش‌های دیگر تصویربرداری

رادیوگرافی، پرستفاده‌ترین آزمایش تصویربرداری برای تشخیص بسیاری از بیماری‌ها در کل جهان است. مقادیر زیاد اطلاعات، سادگی، تشعشع و هزینه پایین، مزایای این روش تصویربرداری هستند. این نوع تصویربرداری ناتوان در تشخیص دیافراگم، مرز دیواره ریه و محل درگیرشده توسط ندول است [۲۲].

در تصویربرداری توموگرافی کامپیوتری، تصاویر به صورت سه بعدی نشان داده می‌شوند. این امر با گرفتن عکسهای متعدد توسط اشعه‌ی ایکس در ارتباط با الگوریتم‌ها و محاسبات رایانه‌ای انجام می‌شود. در این نوع تصویربرداری، شخص بیمار بر روی تختی متحرک قرار می‌گیرد و همان طور که در شکل ۲-۶ مشاهده می‌کنید، وارد محفظه تونلی می‌شود. این محفظه، شامل منبع اشعه ایکس و تعدادی حس‌گر است. اطلاعات با توجه به شدت اشعه ایکس دریافت شده و توسط آشکارسازها دریافت می‌شود. این اطلاعات به سیستم مرکزی سی‌تی‌اسکن فرستاده می‌شود و در نهایت، سیستم اطلاعاتی را که در مرحله قبل به دست آورده است را مورد پردازش قرار می‌دهد. [۲۲].



شکل ۶-۲. دستگاه سی‌تی‌اسکن مدرن، که شامل محفظه تونلی، منبع اشعه ایکس و حسگرها می‌باشد.

تصویربرداری رزونانس مغناطیسی یا ام‌آر‌آی (MRI) یک روش پیشرفته تصویربرداری است که با ایجاد میدان مغناطیسی به تولید تصویر می‌پردازد. در این روش با استفاده از یک آهنربای بسیار قوی، امواج رادیویی و یک آبر رایانه، تصاویری با جزئیات دقیق از اندام‌ها و بافت‌های بدن ایجاد می‌شود. زمان تقریبی آزمایش ام‌آر‌آی حدود ۲۰ الی ۶۰ دقیقه است و به اندام مورد بررسی بستگی دارد. در ام‌آر‌آی، برخلاف سی‌تی‌اسکن و رادیوگرافی ساده، از اشعه ایکس برای تصویربرداری استفاده نمی‌شود. این موضوع باعث شده است تا ام‌آر‌آی نسبت به سایر روش‌های تصویربرداری از ایمنی بیشتری برخوردار باشد.

ام‌آر‌آی بهترین روش تشخیصی برای تصویربرداری از بافت‌های نرم مانند غضروف، تاندون، رگ و عصب‌ها است. ام‌آر‌آی نسبت به سی‌تی‌اسکن اطلاعات دقیق‌تری در مورد اندام‌های داخلی (بافت‌های نرم) مانند مغز، رباط‌ها، اندام‌های تولیدمثل و سایر سیستم‌های داخل بدن ارائه می‌دهد. ام‌آر‌آی یک روش کاملاً بدون درد است و تاکنون هیچ عارضه‌ای از آن گزارش نشده است.

هر دو روش ام‌آر‌آی و سی‌تی‌اسکن، می‌توانند تصویر مناسب از ریه را به پزشک ارائه دهند. با این حال، روش اول و مورد بررسی ما در این جا یعنی سی‌تی‌اسکن سریع‌تر بوده و کمتر از پنج دقیقه طول می‌کشد. روش دوم یعنی ام‌آر‌آی، همان‌طور که در بالا گفته شد، معمولاً ۲۰ الی ۶۰ دقیقه طول می‌کشد. هم‌چنین سی‌تی‌اسکن می‌تواند تصاویری از اندام‌ها، بافت‌ها و ساختار اسکلتی را ارائه دهد.

در آخر اگر بخواهیم این دو روش را با هم مقایسه کنیم، به طور کلی روش تصویربرداری سی تی اسکن استفاده بیشتری از آم آر آی دارد. یکی از دلایل آن زمان کمتری است که می گیرد که در مواقع ضروری و اورژانسی به شدت کاربرد بالاتری دارد. هزینه این روش بسیار پایین تر از روش آم آر آی است. تفاوت اصلی در روش تصویری برداری آن ها است که در سی تی اسکن از اشعه ایکس استفاده می شود اما در روش تصویربرداری آم آر آی از امواج رادیویی و مگنت استفاده می شود. هر دو روش بدون درد هستند اما روش دوم صدای زیادی در حین عکس برداری دارد.

با توجه به توضیحاتی که تا اینجا کار داده شد، برای تصویر برداری و تامین داده های مورد نیاز برای انجام کار، از روش سی تی اسکن استفاده می شود [۲۳].

۵-۲ مقیاس ارزیابی تصاویر سی تی اسکن و قالب ذخیره ساز

قالبی است که به طور معمول در ذخیره سازی تصاویر پزشکی استفاده می شود و با پسوند “ . dam” نوشته می شود. یک فایل DICOM که به صورت تصویری می باشد، به دست آمده از یک استاندارد تصویربرداری دیجیتال و ارتباطات در پزشکی است. در تاریخچه ای که در بیشتر منابع ذکر شده است، در سال ۱۹۸۳ حالت استاندارد توسط یک کمیته مشترکی که متشکل از دانشکده رادیولوژی آمری (ACR)^۱ و انجمن تولیدکنندگان بین المللی الکتریکی (NEMA)^۲ بود، شکل گرفت. هدف از تشکیل این کمیته ایجاد استقلال بین تصاویر دیجیتال پزشکی و دستگاه های تصویر برداری بود [۲۴].

اولین استاندارد در حوزه ثبت تصاویر بر روی نوارهای مغناطیسی توسط انجمن فیزیک پزشکی آمریکا (AAPM)^۳ تولید شد. نسخه اولی که توسط این استاندارد در سال ۱۹۸۵ منتشر شد، ACR_NEM نام گرفت و با نام ۱۰۰ ACR_NEMA منتشر شد و همان طور که انتظار می رفت، دارای خطاهایی بود.

^۱ American College of Radiology

^۲ National Electrical Manufacturers Association

^۳ American Association of Physicists in Medicine

این گروه برای به روزرسانی این انتشار، نسخه دوم که ۲٫۰ ACR_NEMA نام داشت را در سال ۱۹۸۸ را منتشر کرد. این نسخه در ارتباطات پزشکی بسیار محدود بود و برای رفع این مشکل، نسخه سوم را در سال ۱۹۹۲ منتشر کرد. این نسخه را NEMA DICOM یا ۳٫۰ DICOM نامید که نام اختصاری آن DICOM می‌باشد و به این اسم معروف شده است [۲۴].

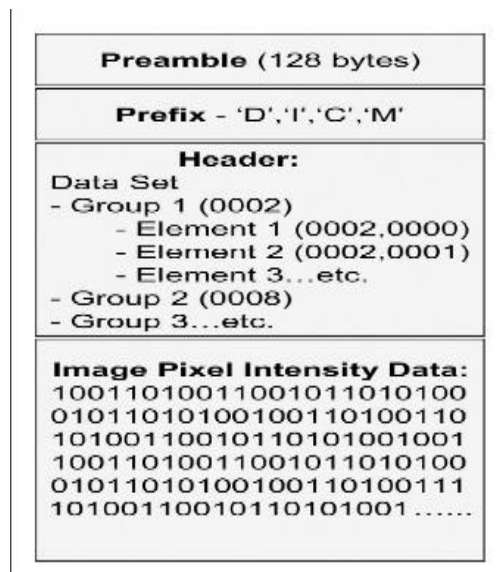
فایل DICOM از مجموعه داده تصویر و یک سرصفحه^۱ تشکیل شده است و در یک فایل واحد است. همان‌طور که در شکل ۲-۷ مشاهده می‌کنید، اطلاعات داخل یک سرصفحه که متشکل از یک سری ثابت و استاندارد از برچسب‌ها^۲ است، قرار دارد [۲۴].

استخراج داده‌ها از این برچسب‌ها اطلاعات مهمی درمورد بیمار، فاصله‌ی بین پیکسل‌ها در تصویر، نوع تصویربرداری، ضخامت برش اسلایس‌ها، و ابعاد تصویر می‌دهد [۲۴]. فایل‌های DICOM منفرد توسط ویندوز^۳ فایل‌های تصویر شناخته نمی‌شوند و این باعث می‌شود که شخص با کلیک بر روی این فایل، نتواند محتویات تصویر را مشاهده کند. [۲۵].

Header^۱

Tags^۲

Windows^۳



شکل ۷-۲. نمایشی از ساختار یک فایل DICOM [۲۴].

۲-۵-۱ مقیاس ارزیابی

در سی‌تی‌اسکن پس از آن که اشعه ایکس از بدن بیمار عبور کرد، براساس شدت تضعیف اشعه توسط هر قسمت از اندام مورد تابش اشعه عبوری، به آشکارسازها (detectors) برخورد می‌کند. این اشعه طی مراحل تبدیل به یک سیگنال الکتریکی شده و به مرکز پردازش کامپیوتر، برای انجام محاسبات جهت بازسازی تصویر فرستاده می‌شود. در قسمت سیستمی، هریک از این سیگنال‌ها، معرف اعداد بسیار نزدیک به هم هستند. این اعداد در یک شبکه ماتریس جایگذاری می‌شود که توسط روش‌های ریاضی پیچیده نسبت ضرایب تضعیف در هر وکسل محاسبه می‌گردد. اعداد سی‌تی‌اسکن به‌دست آمده از این روش نزدیک به هم هستند و به‌کارگیری آن‌ها برای بیان خصوصیات جسم مشکل است. با این اعداد نمی‌توان تصویر مقیاس خاکستری (gray scale) به‌وجود آورد. برای رفع این مشکل، هر یک از اعداد نسبت به ضریب خطی آن سنجیده می‌شوند و این اعداد همان اعداد سی‌تی‌اسکن هستند. این اعداد در یک مقیاس جدید به نام هانسفیلد^۱ نام‌گذاری می‌شوند. این مقیاس توسط Godfrey Newbold Hounsfield یکی از مهندسين اصلی و مخترعين سی‌تی‌اسکن، به وجود آمد. صفر واحد هانسفیلد را شدت اشعه آب

^۱Hounsfield

مقطر در دما و فشار استاندارد (STP) ۱ در نظر گرفته‌اند. به طور کلی مقادیر سی‌تی‌اسکن در بازه‌ی ۱۰۰۰ تا ۳۰۰۰ واحد هانسفیلد هستند. در جدول ۱-۲ مقادیر سی‌تی‌اسکن متناظر با بافت‌های مختلف در واحد هانسفیلد ذکر شده است [۲۶].

جدول ۱-۲. مقادیر سی‌تی متناظر با بافت‌های گوناگون در واحد هانسفیلد [۲۶]

نوع بافت	مقدار سی‌تی 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
استخوان	+۴۰ تا +۱۰۰۰

۲-۲ پیش پردازش

حذف نویز و بهبود لبه‌ها در پیش‌پردازش تصاویر بسیار مهم هستند. با توجه به این‌که کار این پایان‌نامه بر مبنای دو قسمت مهم پیش‌پردازش و استفاده از یادگیری عمیق است، در ادامه توضیحاتی راجب فیلترهای ریخت‌شناسی ریاضی^۲ که از مهم‌ترین روش‌های پیش‌پردازش تصاویر هستند، داده خواهد شد.

۱-۲-۲ فیلترهای ریخت‌شناسی ریاضی

همان‌طور که می‌دانید، بسیاری از تصاویر دارای نویز هستند. این یکی از عواملی است که بر روی کیفیت تصویر تاثیر می‌گذارد. برای از بین بردن این مشکل، از فیلتر کردن تصاویر استفاده می‌شود. فیلتر بر اساس مورفولوژی ریاضی قادر است به جزییات و اطلاعات ویژگی هندسی اتکا پیدا کند. این فیلتر از

^۱ Standard Temperature and Ppressure

^۲ Mathematical Morphology Filters

عملگرهای مورفولوژیکی برای نگهداری اطلاعات اصلی تصویر و فیلتر کردن موثر موانع استفاده می‌کند. عملیات مورفولوژیک، اطلاعات تصویر را ساده می‌کند، به همین دلیل مشخصات شکل‌های تصویر را حفظ کرده و تمامی اطلاعات اضافی را پیدا و حذف می‌کند. مورفولوژی ریاضی یک تکنیک جدید در مطالعه اکوسیستم‌ها و علوم محیطی است. این تکنیک بر اساس تئوری مجموعه‌ها و عملیات جبری مینکوسکی پابرجا است و از این عملیات به‌طور معمول برای تشخیص لبه‌ها و رفع نویز تصویر استفاده می‌گردد [۲۷]. با انتخاب عناصر سازنده مناسب از لحاظ اندازه، شکل و جهت با توجه به ویژگی‌های بافتی تصویر، می‌توان لبه‌های تصاویر پزشکی را تشخیص داد. عملیات مورفولوژی دارای چهار دسته کلی به نام‌های دسته فرسایش، گسترش، باز کردن و بستن است [۲۸].

۱-۱-۶-۲ فرسایش و گسترش

اگر A و B را به عنوان دو مجموعه در نظر بگیریم، عملگر فرسایش با نماد $A \ominus B$ نمایش داده می‌شود. در این نماد، نمایش تصویری ورودی را A و نمایش عنصر ساختمانی را B در نظر می‌گیرند. فرسایش در اصل دو اثر متفاوت بر روی تصاویر خاکستری می‌گذارد. اثر اول به این صورت است که اگر تمام درایه‌های عنصر ساختمانی مثبت باشد، تصویر خروجی تیره‌تر از تصویر ورودی است. اثر دوم به این شکل است که اثر جزئیات روشن تصویر ورودی که سطحی کوچک‌تر از عنصر ساختمانی دارند، کاهش می‌یابد. به طور معمول در این عملیات، نقاط ناخواسته تصویر بانیری حذف می‌شوند و سایر اجزای تصویر یک یا چند پیکسل نازک خواهند شد. در مجموع اگر در کنار تمام نقاط سفید تصویر بررسی شده حداقل یکی از همسایگان انتخابی آن سیاه باشد، آن هم سیاه خواهد شد.

از طرفی برای افزایش ابعاد اجزاء داخل تصویر به اندازه یک یا چند پیکسل باید از عملیاتی به نام گسترش یا افزایش استفاده شود که برای مجموعه‌های A و B با نماد $A \oplus B$ نمایش داده می‌شود. با انجام و صورت گرفتن این عملیات، نقاطی که از یک تصویر بانیری در اثر عواملی چون تاثیر نویز یا اعمال حد آستانه نامطلوب جا افتاده است، تصحیح می‌شوند. این الگوریتم که به الگوریتم اعمال فیلتر افزایش

معروف است به این صورت کار می‌کند که تمامی نقاط سیاه تصویر را بررسی کرده و در صورتی که حداقل یکی از همسایگان انتخابی نقطه مورد بررسی سفید باشد، نقطه مورد نظر هم سفید خواهد شد و در این غیر این صورت سیاه می‌ماند.

۲-۱-۶-۲ باز کردن و بستن

عملگر باز کردن، باعث هموارسازی تصویر می‌شود. این عملگر کار حذف بیرون‌زدگی‌های باریک را بر عهده دارد و در آخر اتصالات ضعیف را می‌شکند. این عملگر از ترکیب‌های مختلف دو عمل فرسایش و گسترش ایجاد می‌شود.

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:

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

همان‌طور که در رابطه مشاهده می‌کنید، عملگر باز کردن در ابتدای کار عمل فرسایش را بر روی تصویر لحاظ می‌کند. این کار باعث می‌شود که اندازه عناصر موجود در تصویر کاهش یابد و باعث حذف عناصر کوچک‌تر از عنصر سازنده B می‌شود. در نهایت عملگر گسترش روی نتیجه حاصل‌شده اعمال می‌شود که این کار در آخر باعث بازگشت اندازه عناصر باقی‌مانده در تصویر به اندازه اولیه آن‌ها می‌شود. یکی دیگر از عملگرهایی که باعث هموارسازی تصویر می‌شود، عملگر بستن می‌باشد. این عملگر، معمولاً تورفتگی‌های باریک و انفصالات باریک و بلند را پیوند می‌زند. انفصال‌های موجود در منحنی‌ها را پر و حفره‌های کوچک را حذف 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:

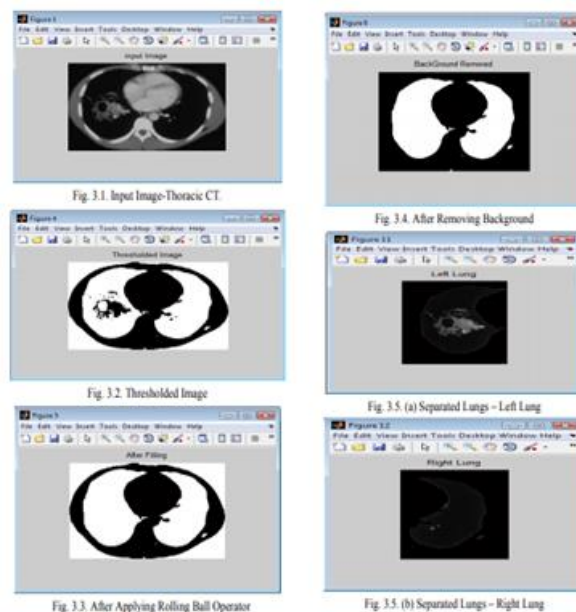
$$A \bullet B = (A \oplus B) \ominus B \quad (2-2)$$

با توجه به رابطه (۲-۲)، عملگر باز کردن ابتدا عمل گسترش را روی تصویر اعمال کرده و باعث می‌شود حفره‌ها و انفصالات کوچک از تصویر حذف شوند. در ادامه باعث بزرگ شدن اندازه عناصر موجود در تصویر می‌شود. در نهایت با اعمال عملگر فرسایش روی تصویر حاصل، اندازه عناصر بزرگ‌شده در تصویر به اندازه اولیه‌شان برمی‌گردند.

در ادامه به یک نمونه کامل از موارد استفاده از عملگرهای مورفولوژی در پردازش تصاویر پزشکی می‌پردازیم:

در مقاله [۲۹]، یک الگوریتم برای جداسازی دو ریه چپ و راست از تصاویر سی‌تی اسکن ارائه شده است، که به صورت کاملاً خودکار عمل می‌کند. در ابتدای این مقاله یک الگوریتم خوشه‌بندی فازی برای بخش‌بندی بافت ریه به کار می‌رود. بعد از این مرحله یعنی بخش‌بندی، امکان ظاهر شدن حفره‌هایی در نزدیکی مرز پارانشیم ریه وجود دارد. برای چسباندن قسمت‌های مختلف ریه و پرکردن این حفره‌ها، از عملگر مورفولوژی بستن استفاده شده. پس از آن، پس‌زمینه تصویر حذف می‌شود. در نهایت و مرحله آخر بر روی تصویر حاصل از مرحله قبل، عملگر مورفولوژی بستن و سپس باز کردن اعمال می‌شود. این امر منجر به صاف^۱ شدن مرزهای نامنظم موجود می‌شود و دو ریه چپ و راست از تصویر جدا می‌شوند. در شکل ۲-۸، چگونگی استخراج ریه نمایش داده شده است.

^۱ Smooth



شکل ۸-۲. مثالی از اعمال عملیات مورفولوژی در بخش‌بندی ریه [۲۹].

۷-۲ بخش‌بندی ریه

قطعه‌بندی یا بخش‌بندی تصویر، به فرآیند تفکیک اجزای تصویر گفت می‌شود. برای فهم بیشتر، فرآیندی است که تصویر را به بخش‌های معنی‌دار تقسیم می‌کند، به صورتی که هر بخش ویژگی‌های بصری مشابه داشته باشد [۳۰]. همان‌طور که گفته شد، برای تحلیل و بررسی تصاویر پزشکی به استخراج اطلاعات تصویر و نواحی مجزای هر تصویر نیاز است. برای این امر، نیاز به قطعه‌بندی تصویر وجود دارد که یکی از مهم‌ترین مراحل در پردازش و تحلیل می‌باشد. در تصاویر پزشکی، قطعه‌بندی برای استخراج، نمایش تصاویر و اندازه‌گیری بسیار اهمیت دارد. اکثر مواقع از کاربردهای پردازش تصاویر پزشکی برای تفکیک و جداسازی قسمت‌های اصلی بدن مانند قلب، مغز، ریه و غیره در تصاویر سی‌تی یا ام‌آر‌آی استفاده می‌شود. قطعه‌بندی برای تحلیل و سپس تشخیص نواحی پاتولوژیکی مثل تومورها سودمند است. الگوریتم‌های بخش‌بندی در سیستم‌های CAD، اولین گام در پردازش خودکار تصاویر جهت ارزیابی و تحلیل تصاویر پزشکی و نیز به‌عنوان مرحله‌ی پیش‌نیاز برای انجام عملیات تحلیل محتوای تصویر

و تشخیص الگوهای مورد جستجو در آن‌ها محسوب می‌شود؛ به همین علت عملکرد نهایی سیستم‌های CAD بسیار زیاد به مرحله‌ی بخش‌بندی وابسته می‌شود.

در قطعه‌بندی تصاویر پزشکی، هدف اصلی، جدا نمودن اجزای مطلوب و سپس تقسیم‌بندی آن‌ها به ساختارهای آناتومیکی مختلف است. همان‌طور که گفته شد، ناحیه ریه از همه بافت‌های درون سطح ریه یعنی بافت اصلی ریه، راه‌های هوایی و رگ‌ها^۱ و ساختارهای غیر طبیعی تشکیل شده‌است. به همین علت در قطعه‌بندی ریه، هدف اصلی جداکردن لوب‌های ریه از سایر نواحی تصویر سی‌تی‌اسکن می‌باشد. این هدف برای این است که تصاویر سی‌تی‌اسکن دارای نواحی می‌باشند که ممکن است که ندول‌های ریوی در آن وجود داشته باشند. به دست آوردن آن‌ها بسیار مهم است چرا که اولین و مهم‌ترین مرحله برای تمامی کارها از جمله آشکارسازی^۲ و تشخیص^۳ وجود این ندول‌ها می‌باشد.

هدف از تحقیقات و بررسی بر قطعه‌بندی تصاویر، می‌تواند افزایش اعتبار، دقت، صحت و کاهش هزینه‌های محاسباتی باشد. این امر با تعیین و شناسایی یک الگوریتم مناسب که بتواند با دقت بالا، ریه را قطعه‌بندی و مدل ریه را ایجاد کند، تحقق می‌یابد. بخش‌بندی ناحیه ریه همیشه به علت دارا بودن ساختار غیریکسان و پیچیده در طول و عرض، یک مسئله مهم و پیچیده بوده است. می‌توان این طور بیان کرد که، بخش‌بندی ناحیه ریه، به دلیل عدم وجود هماهنگی در این ناحیه که برای مثال می‌توان به وجود ساختارهای با شدت مشابه مثل عروق خونی^۴، رگ‌ها، نایچه‌ها و ریزنایچه‌ها به صورت مداوم به چالش کشیده می‌شود. به منظور داشتن درک بهتر از انواع الگوهای بیماری‌های ریوی و درمان و تشخیص سریع-تر آن، ایجاد یک سیستم تشخیص خودکار برای مکان‌یابی و بخش‌بندی ناحیه ریه مسئله‌ای مهم و بسیار با اهمیت است.

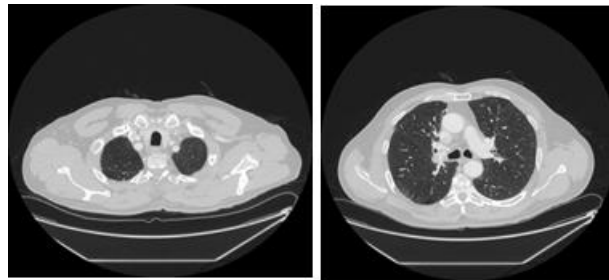
^۱ Veins

^۲ Detection

^۳ Recognition

^۴ Blood vessels

در آخر و برای جمع‌بندی باید بافت اصلی ریه از نواحی غیر ریه مانند پس‌زمینه جدا شود. مزیت این کار افزایش دقت است و در ضمن ناحیه مورد بررسی را محدود کرده و سرعت الگوریتم را افزایش می‌دهد. همچنین امکان بروز خطا و تغییرات نویز در مراحل بعدی تحلیل و بررسی این تصاویر کاهش یافته و نواحی اضافه را هم حذف می‌کند.



شکل ۹-۲. تصاویری از برش‌های قسمت‌های مختلف ریه مربوط به یک بیمار

۸-۲ مروری بر روش‌های بخش‌بندی (کارهای پیشین)

روش‌های زیادی برای بخش‌بندی پیشنهاد شده است. بسیاری از این روش‌ها به صورت خودکار و بسیاری از این روش‌ها نیمه‌خودکار هستند و بسته به نوع کار یکی از این روش‌ها انتخاب می‌شود [۲۸]. به طور کلی روش خاصی برای بخش‌بندی ریه از تصاویر سی‌تی‌اسکن وجود ندارد و محققان، سیستم‌های بخش‌بندی مختلفی را برای ریه پیشنهاد کرده‌اند. همان‌طور که ذکر شد، تعدادی از روش‌ها موجود، نیمه-خودکار هستند که دقت بالایی ندارند و در ضمن به‌دخالت عامل انسانی نیاز دارند. در این پایان‌نامه و در فصل دوم به بخش‌بندی‌های سنتی پرداخته می‌شود و فصل سوم روش‌های بخش‌بندی مبتنی بر یادگیری عمیق مفصل مورد بررسی قرار خواهند گرفت.

۱-۸-۲ روش‌های بخش‌بندی سنتی

روش‌های سنتی بخش‌بندی، خود به سه زیردسته به تقسیم‌بندی می‌شوند. در ادامه نام آن‌ها خواهد برده شد و به طور کامل بررسی خواهند شد و مزایا و معایبشان توضیح داده خواهد شد.

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

^۱ Threshold Based

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. به دلیل آن که در زمان تصویربرداری سی‌تی‌اسکن، ریه‌ها پر از هوا هستند، به‌صورت نواحی تیره در تصویر سی‌تی‌اسکن قابل رویت هستند و مشخص می‌شوند. به همین علت تباین بین ریه و بافت اطراف آن، اساس بسیاری از روش‌های بخش‌بندی ریه می‌باشد و به همین دلیل روش‌های بخش‌بندی زیادی بر مبنای آستانه‌گذاری عمل می‌کنند [۲۹-۳۷].

مهمترین اصل در روش‌های مبتنی بر آستانه‌گذاری برای جداسازی بافت ریه این است که همواره روشنایی بافت ریه از سایر نواحی اطراف، کمتر است. با توجه به همین امر، جداسازی ناحیه ریه صورت می‌گیرد [۳۸]. سه روش از مجموعه روش‌های آستانه‌گذاری که در بالا گفته شد و در مراجع قرار دادیم را در ادامه مورد بررسی قرار خواهیم داد.

Gao و هم‌کاران روشی را بر مبنای آستانه‌گذاری و در چهار مرحله در کارهای خود استفاده کرده‌اند [۳۶]. روش کار به این صورت بود که در ابتدای کار راه‌های هوایی با روش رشد ناحیه حذف می‌شود که در ادامه بیشتر توضیح داده خواهد شد. در ادامه از یک آستانه بهینه برای حذف رگ‌ها استفاده می‌شود. در مرحله‌ی بعد که مرحله‌ی سوم می‌باشد، ریه راست و چپ با استفاده از آستانه‌گذاری جدا می‌شود. در آخر از عملگرهای مورفولوژی که در قسمت‌های قبل توضیح داده شد، برای یک‌دست‌سازی مرز ریه استفاده شده‌است.

هو و همکاران [۳۲]، یک روش دیگر که بر مبنای مرحله به مرحله^۱ بود را پیشنهاد کردند. شیوه کار آن‌ها به این صورت بود که در ابتدای کار برای به‌دست آوردن ناحیه تقسیم‌بندی اولیه از یک آستانه تکراری استفاده کردند. در ادامه توسط یک عملگر مورفولوژی بازکردن^۲، ناحیه به‌دست آمده بسته می‌-

^۱ Stepwise

^۲ Opening Morphological Operation

شود. مثالی دیگر که از این روش استفاده کرده است، این است که با استفاده از تبدیل موجک^۱ و یک آستانه بهینه، بخش‌بندی اولیه را انجام داده و در ادامه این ناحیه را با استفاده از عملیت مورفولوژی تصحیح می‌کند [۳۹].

(۱) آستانه‌گذاری سراسری

در روش‌های آستانه‌گذاری سراسری، تلاش بر این است که فقط یک مقدار آستانه برای تمام پیکسل‌های تصویر یافت شود. یعنی این روش بر این پایه است که هر پیکسل در پیش‌زمینه یا پس‌زمینه صفحه بر پایه‌ی مقدار خاکستری‌اش، علامت‌گذاری می‌شود [۴۰ و ۴۱]. روش‌های بسیاری برای انتخاب آستانه سراسری موجود است. به عنوان مثال در مقاله [۴۲] روشی که وجود دارد، بر مبنای مدل دسته‌بندی است که احتمال خطا را حداقل می‌کند. این مدل به این منظور است که اگر تصویر فقط شامل پیش-زمینه و پس‌زمینه باشد، بر اساس این روش، خطا به‌عنوان تعداد مجموع پیکسل‌هایی از پیش‌زمینه که به‌اشتباه به‌عنوان پس‌زمینه و پیکسل‌هایی از پس‌زمینه که اشتباهی به‌عنوان پیش‌زمینه در نظر گرفته شده است، محاسبه می‌شود. جانسون و هم‌کارانش [۴۳] یک نسخه نیمه‌خودکار از این روش را ارائه داده‌اند که کاربرد آن در اندازه‌گیری توده‌های بطنی در تصاویر ام‌آر‌آی بوده است. روشی که به کار برده‌اند به این شکل صورت گرفته است که، دو عدد پیکسل از ناحیه پیش‌زمینه و پس‌زمینه به وسیله‌ی عملگر انتخاب می‌شود. میزان آستانه به‌شکل خودکار وقتی که شدت‌های پیکسل در نواحی مدور اطراف پیکسل‌های انتخاب‌شده مقایسه‌شد، انتخاب شده و این کار، با پایین‌ترین تعداد پیکسل‌های اشتباه دسته-بندی‌شده، بین دو توزیع می‌باشد.

آستانه‌گذاری سراسری از لحاظ محاسباتی سریع و ساده می‌باشد. این روش روی تصاویری که شی‌ها با مقادیر شدت یکنواخت روی کنتراست پیش‌زمینه قابل قبولی دارد، عملکرد مناسبی دارد. نقطه ضعف آن

^۱Wavelet Transform

به این شکل است که اگر یک کنتراست پایین بین شیء و پیش‌زمینه موجود باشد و در ضمن اگر شدت پیش‌زمینه به‌طور عمده در سراسر تصویر تغییر کند، با خطا و شکست روبه‌رو می‌شود.

۲) آستانه‌گذاری متغیر یا محلی

روش قبلی که به طور کامل شرح داده شد را نمی‌توان برای موارد خاص، استفاده کرد و بسط داد. به عنوان نمونه، زمانی که کنتراست اشیا در سراسر تصویر تغییر کنند و پیش‌زمینه ثابت نباشد، آستانه‌گذاری سراسری ممکن است در یک بخش از تصویر به خوبی عمل کند اما در همین زمان ممکن است نتایج ضعیفی را در نواحی دیگر به دنبال داشته باشد. برای حل این مشکل از روش دیگر که آستانه‌گذاری محلی است، استفاده می‌شود [۴۱ و ۴۴ و ۴۵].

آستانه‌گذاری محلی شامل دو روش می‌باشد که روش اول به این صورت عمل می‌کند که در ابتدا تصویر را به چندین زیرتصویر تقسیم می‌کند. در ادامه مقدار آستانه‌ای به‌صورت جدا از سایر نواحی، برای هر یک از این زیرتصویرها انتخاب می‌کند. در این روش نکته بسیار مهم این است که باید دقت شود که زیرتصویرها باید هر کدام، به‌اندازه‌ای باشند که قابل تفکیک به دو بخش پیش‌زمینه و پس‌زمینه باشند. روش دوم از این آستانه‌گذاری به این شکل است که با آزمایش کردن شدت‌های تصویر در همسایه هر پیکسل، مقدار آستانه قابل تعیین است. برای مثال و در مقاله‌ای که مورد بررسی قرار گرفت [۴۶]، ابتدا یک تصویر را به زیرتصاویر مستطیل‌شکل که هم دیگر هم‌پوشانی دارند تقسیم می‌کند. بعد از آن بافت‌نگار^۱هایی برای هر زیرتصویر محاسبه می‌شود. آنگاه اگر یک زیرتصویر، فقط شامل پیش‌زمینه و پس‌زمینه باشد، مقدار حداقل بین قله‌های بافت‌نگار را به‌عنوان آستانه محلی در نظر می‌گیرد. این روش به طور کلی از لحاظ محاسباتی هزینه‌برتر از روش قبلی است. کاربرد این روش اغلب در ناحیه‌بندی تصاویر از پیش‌زمینه‌های متفاوت، است.

^۱Histogram

در مجموع، روش‌های آستانه‌گذاری، زمانی که تصویر دارای نویز و ناهنجاری‌های ریه نباشد، عملکرد قابل قبولی دارد و بسیار ساده و سریع هستند. اما با این حال مشکلات خود را دارند که در ادامه بعضی از این مشکلات گفته خواهد شد:

(۱) به علت وجود سطوح خاکستری نزدیک به هم و مرزهای بسیار، یافتن یک حد آستانه مناسب دشوار است [۴۷].

(۲) ندول‌های ریوی متصل به دیواره ریه را نمی‌تواند حذف کند.

(۳) نای و نایژه‌ها را هم نمی‌تواند حذف کند.

(۳) روش‌های رشد ناحیه‌ای

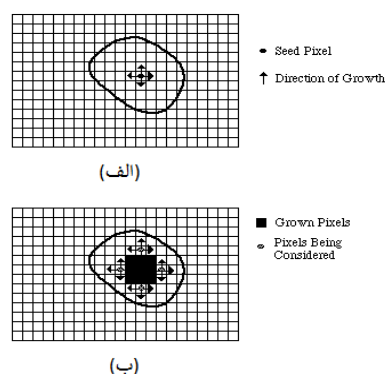
برخلاف روش آستانه‌گذاری که روی تفاوت شدت‌های پیکسل کار می‌کند، روش‌های رشد ناحیه در جستجوی گروهی از پیکسل‌ها با شدت‌های مشابه می‌باشند. در این روش، کار به این صورت پیش می‌رود که پیکسل‌ها یا زیرناحیه‌ها را در نواحی بزرگتر گروه‌بندی می‌کند. یکی از ساده‌ترین راه‌های موجود برای این کار، پیوستن پیکسل‌ها بر اساس یک سری خواص مشابه است. برای درک بیشتر، 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 به طور معمول به صورت دستی قرار داده می‌شود.

تعیین یک یا چند نقطه‌کار روی تصویر به عنوان بذر یا دانه اولیه، نخستین گام برای شروع بخش‌بندی تصاویر به کمک الگوریتم‌های رشد ناحیه‌ای است. در مرحله دوم، آزمایش شباهت بر روی پیکسل‌های همسایه برای اضافه شدن به ناحیه‌ی رشد صورت می‌گیرد. آن پیکسل‌هایی از تصویر که بیشترین شباهت به دانه را دارند تعیین و انتخاب می‌شوند و با برآورده کردن شرایط، شامل ناحیه‌ی رشد مورد نظر می‌شود. شرط خاتمه این الگوریتم به صورتی است که دیگر هیچ پیکسلی برای اضافه کردن وجود نداشته

^۱Seed Point

باشد. در نهایت به وسیله‌ی تمام پیکسل‌هایی که در این رویه رشد پذیرفته شده‌اند، ناحیه بخش‌بندی-شده‌ی مطلوب نمایش داده می‌شود [۴۸و۴۹].

همان‌طور که در شکل ۱۰-۲ قابل مشاهده است، کار ابتدا از یک پیکسل دانه شروع می‌شود و تمام نقاطی که در همسایگی این پیکسل بوده و بیشترین شباهت را به آن دارند، کم‌کم به ناحیه‌ی مورد نظر اضافه می‌شود. ناحیه مورد نظر بزرگ‌تر می‌شود و در نهایت، پیکسل‌هایی که از نظر ظاهر شبیه به هم هستند، جزء یک مجموعه محسوب می‌شوند و قرار خواهند گرفت.



شکل ۱۰-۲. شماتیک روش‌های رشد ناحیه‌ای [۵۰]. (الف): انتخاب بذر اولیه (ب): بررسی نقاط مجاور و انجام فرآیند بخش‌بندی.

طبق مطالعه انجام شده در بین مقالات موجود، این روش، در اغلب کاربردهای پزشکی مثل بخش-بندی حفره‌ها در تصاویر کاردیاک [۵۱]، بخش‌بندی کلیه [۵۲]، استخراج رگ‌های خونی در داده‌های آنژیوگرافی [۵۳]، استخراج سطح مغز [۵۴] بخش‌بندی تصاویر سی‌تی‌اسکن ریه [۵۵-۵۸] استفاده شده است. از آنجا که قسمت ابتدایی 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 [۵۵] برای استخراج ریه در تصاویر سی‌تی از روش رشد ناحیه استفاده کردند و یک روش کاملاً خودکار را ارائه دادند. روشی که آن‌ها پیشنهاد دادند به این شکل عمل می‌کرد که در

ابتدای کار تصاویر را خاکستری می‌کند. در ادامه برای تولید ماسک‌های اولیه از الگوریتم آستانه‌گذاری بر روی تصاویر سی‌تی ریه استفاده شده است. در مرحله‌ی بعدی نوبت به شروع الگوریتم رشد ناحیه با پیکسل دانه و پیکسل‌های دیگر که دانه را احاطه کرده‌اند می‌باشد. بعد از آن نوبت به تعیین پیکسل‌هایی که بیشترین شباهت به پیکسل دانه را دارند می‌رسد و اگر شرایط مورد نیاز مسئله را برآورده سازند، شامل ناحیه‌ی موردنظر نیز می‌شوند. شرط خاتمه این الگوریتم این است که، پیکسلی برای اضافه کردن موجود نباشد. ناحیه‌ی ایجاد شده، مکرراً توسط بررسی تمام غیرهمسایگی ناحیه رشد کرده و زمان پایان این فرآیند زمانی است که تفاوت شدت بین میانگین ناحیه و پیکسل جدید بیشتر از آستانه معین باشد. بعد از پایان یافتن روند الگوریتم رشد ناحیه، نوبت به لحاظ تشخیص‌گر لبه بر روی تصویر است تا یک ماسک برای استخراج ریه درست شود. در انتهای کار تصویر نهایی با ترکیب ماسک باینری لبه و تصویر اولیه به‌دست می‌آید. مشکلی که در انتهای کار به وجود می‌آید، در قسمت استخراج ریه شکل می‌گیرد. این مشکل به این گونه است که در بخش‌بندی ندول‌های متصل به دیواره ریه با شکست مواجه می‌شود که این مشکل در شکل ۲-۱۱ قابل دیدن است.

در مقاله‌ای که مورد بررسی قرار گرفت، الگوریتمی پنج مرحله‌ای برپایه رشد ناحیه استفاده شده است [۵۶]. در مرحله اول این الگوریتم، به این صورت عمل می‌شود که یک پیکسل اولیه درنظر گرفته می‌شود. در ادامه طی دو مرحله‌ی بخش‌بندی انطباقی محلی^۱ و پرکردن حفره، پیش‌زمینه و پس‌زمینه از یک‌دیگر جدا می‌شوند. در مرحله بعدی، در اول کار اشیای پیش‌زمینه غیرمتصل حذف شده و سپس نوبت به دوبرابر شدن ناحیه پیش‌زمینه بسط داده شده، می‌رسد. برای به دست آوردن پارامترهای لازم برای مراحل بعدی که شامل دو الگوریتم رشد ناحیه‌ای است. همچنین به عنوان محدودکننده حجم برای ناحیه رو به رشد، لازم است ماسک استفاده شود. در مرحله چهارم به وسیله روش رشد ناحیه اتصال فازی، نقشه اتصال فازی ایجاد می‌شود. شرط خاتمه این است که ماسک کامل که از مرحله قبل

به دست آمده ایجاد شود و در مرحله آخر، ناحیه رشد یافته کروی شکل بر روی نقشه فازی برای ایجاد نتایج بخش بندی چندتایی تطبیق داده می شود.



شکل ۱۱-۲. بخش بندی ریه به روش رشد ناحیه (الف) : تصویر سی تی اولیه (ب) : تصویر قطعه بندی شده به روش رشد ناحیه (توده چسبیده به دیواره ریه که در روش رشد ناحیه ای، در نظر گرفته نمی شوند، با فلش در تضاویر نشان داده شده اند).

➤ مزایای روش های رشد ناحیه ای، شامل موارد زیر هستند :

(۱) مقاومت زیاد در برابر نویز

(۲) آسان بودن پیاده سازی

(۳) استفاده چند معیار به طور هم زمان برای ناحیه بندی

(۴) دقت نسبتا بالا با شرط وضوح لبه های تصویر

➤ معایب روش های رشد ناحیه ای، شامل موارد زیر هستند :

(۱) وابستگی عملکرد الگوریتم به بذرهای اولیه

(۲) زمان بر بودن و پیچیدگی محاسباتی زیاد

(۳) ناتوانی در بخش بندی ندول های متصل به دیواره ریه

۲-۸-۱-۲ روش‌های مبتنی بر لبه

در این دسته از روش‌های قطعه‌بندی، برای بخش‌بندی ناحیه مورد نظر در تصویر، از تشخیص لبه^۱ استفاده می‌شود. در این روش به نقاطی از تصویر که در آن شدت روشنایی به طور ناگهانی تغییر کرده است، لبه تصویر گفته می‌شود. با توجه به این که به وسیله لبه قسمت‌های مختلف یک تصویر از هم مجزا می‌شوند، قطعه‌بندی تصویر با استخراج لبه‌های آن صورت می‌گیرد. این روش به دو دسته تقسیم می‌شود که دسته اول روش‌های تشخیص لبه در حوزه مکان و روش دوم روش‌های تشخیص لبه در حوزه تبدیل است. روش‌های زیادی در تشخیص لبه در حوزه مکان وجود دارد که برای مثال می‌توان به سوبل^۲، کنی^۳، رابرتز^۴، پرویت^۵ و لاپلاس^۶ اشاره کرد که تنها دو الگوریتم کاربردی و مهم سوبل و کنی مورد بررسی قرار خواهد گرفت.

➤ لبه‌یاب سوبل

این متد لبه‌ها را با استفاده از تخمین زدن مشتق پیدا می‌کند، که لبه‌ها را در آن نقاطی بر می‌گرداند که گرادیان تصویر I ، $\max$ است. در فیلتر سوبل دو ماسک به صورت زیر وجود دارد که ماسک سوبل افقی بیشتر لبه‌های افقی را مشخص می‌کند و ماسک سوبل عمودی، لبه‌های عمودی را مشخص می‌کند [۵۹].

^۱ Edge Detection

^۲ Sobel

^۳ Canny

^۴ Roberts

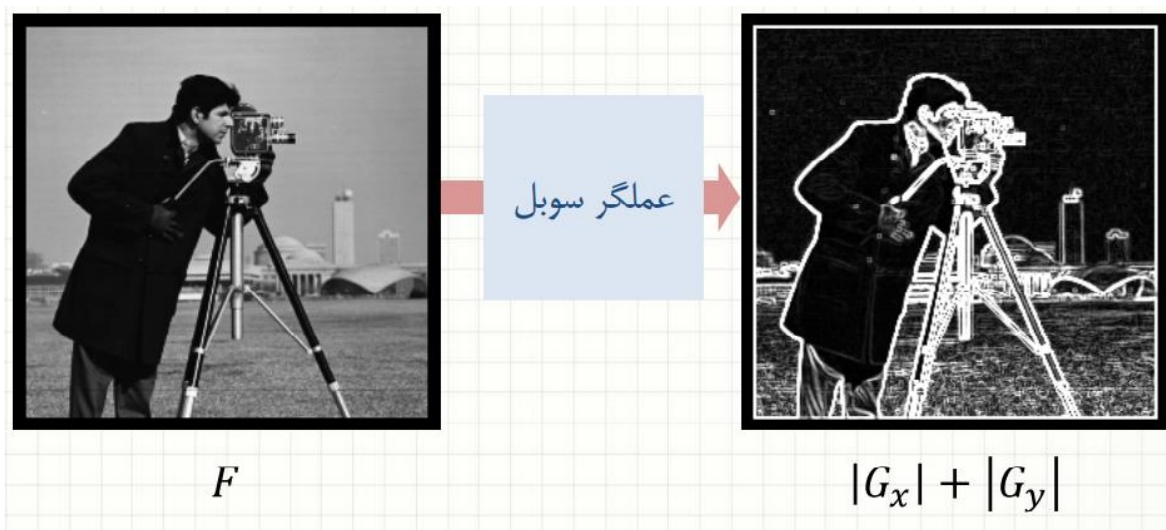
^۵ Prewitt

^۶ Laplace

$$G_y = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad G_x = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

شکل ۱۲-۲. شکل سمت راست ماسک سو بل افقی و شکل سمت چپ ماسک سو بل عمودی

برای مشخص شدن کلیه لبه ها اگر G_x و G_y تصاویر فیلتر شده به وسیله ماسک افقی و عمودی باشند، آنگاه تصویر $\sqrt{[G_x]^2 + [G_y]^2}$ لبه های تصویر را بهتر نشان می دهد. روال فوق به عملگر یا الگوریتم سو بل موسوم است و نتیجه این لبه بر روی تصویر را در شکل ۱۳-۲ می توان مشاهده کرد.



شکل ۱۳-۲. نتیجه عملگر سو بل بر روی تصویر

➤ لبه یاب کنی

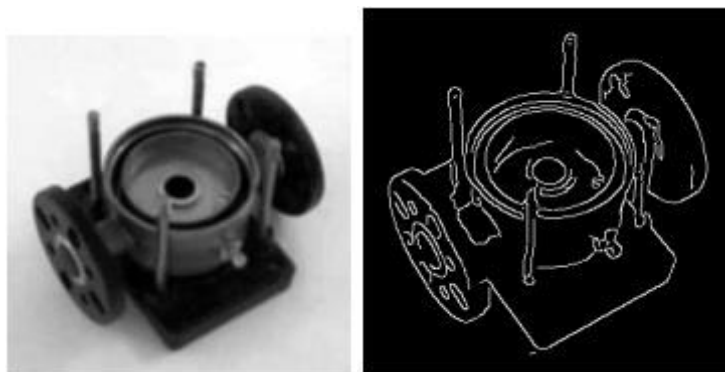
لبه‌یاب‌کنی توسط جان‌اف‌کنی^۱ در سال ۱۹۸۶ ایجاد شد و هنوز یک لبه‌یاب استاندارد، با دقت و کیفیت بالا است [۶۰]. الگوریتم لبه‌یاب‌کنی یکی از بهترین لبه‌یاب‌ها تا به امروز است. این الگوریتم لبه‌یابی از سه بخش اصلی زیر تشکیل شده است [۶۱]:

(۱) تضعیف نویز

(۲) پیدا کردن نقاطی که بتوان آن‌ها را به عنوان لبه در نظر گرفت.

(۳) حذف نقاطی که احتمال لبه بودن آن‌ها کم است.

در لبه‌یابی‌کنی تا حد امکان هیچ لبه‌ای در تصویر نباید گم شود و هم چنین هیچ چیزی که لبه نیست نباید به جای لبه فرض شود که این نکته اولین معیار این روش است. قرار گرفتن لبه‌ها در جای اصلی خود دومین معیار است. داشتن کمترین ضخامت برای لبه‌ها و داشتن یک پاسخ برای هر لبه به ترتیب معیار سوم و چهارم این روش است. این روش بیشتر به کشف لبه‌های ضعیف به درستی می‌پردازد و کمتر فریب نویز را می‌خورد و از بقیه روش‌ها بهتر است. همان‌طور که در شکل ۲-۱۴ می‌بیند، یک نمونه خروجی از این روش نشان داده شده است.



شکل ۲-۱۴. نتیجه کنی بر روی تصویر

^۱John F. Canny

این روش ساختار اشیای طبقه‌بندی شده را تا حدود زیادی حفظ می‌کند. به خوبی کار کردن بر روی پس‌زمینه و اشیا تصویر، دیگر مزیت این روش است [۶۲]. با توجه به مطالب گفته شده در مورد این روش لبه‌یاب‌کنی دارای بهترین عملکرد در بین لبه‌یابی‌ها می‌باشد اما اگر مرزهای ضعیفی در تصویر وجود داشته باشند یا ناحیه‌های تصویر دارای بافت متراکم باشند، این روش به درستی کار نمی‌کند و خروجی مطلوب را برای ادامه کار فراهم نمی‌سازد [۶۳].

۹-۲ جمع‌بندی روش‌های سنتی بخش‌بندی

در قسمت‌های قبل روش‌های بخش‌بندی سنتی را بررسی کردیم و به طور خلاصه آن‌ها را شرح دادیم. مزایا و معایب روش‌ها را بیان کردیم که در جدول قابل مشاهده است. در ادامه این فصل به بخش دوم تحقیق یعنی قسمت تشخیص نوع سرطان خواهیم پرداخت.

جدول ۲-۲. مقایسه روش‌های سنتی بیان‌شده برای بخش‌بندی

نام روش شماره مراجع	مزایا	معایب
آستانه‌گذاری	- سادگی در محاسبات - سرعت نسبتاً بالا - حافظه نسبتاً 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	- پیاده‌سازی ساده - مقاومت نسبتاً بالا در برابر نویز	- زمان‌بری بالا - وابستگی نسبتاً زیاد عملکرد الگوریتم به بذره‌های اولیه
مبتنی بر لبه	- شباهت نسبتاً زیاد به روش طبقه‌بندی انسانی - حفظ بیش از حد اشیای بخش‌بندی شده	- مقاومت نسبتاً کم در برابر نویز - شکست در زمان برخورد با بافت‌های متراکم تصویر

۱۰-۲ استخراج ویژگی و طبقه‌بندی

این مرحله نسبت به مرحله قبلی نسبتاً آسان‌تر دنبال می‌شود چون ورودی این مرحله برای پردازش، از خروجی مرحله قبل می‌باشد. در این مرحله از تحقیق، پیچدگی تصویر از بین رفته و تصاویر بخش‌بندی شده هستند. الگوریتم‌های زیادی برای طبقه‌بندی تصاویر وجود دارد که به طور خلاصه چند از این الگوریتم‌ها را مورد بررسی قرار داده و در ادامه جمع‌بندی نهایی این قسمت هم انجام خواهد شد.

۱-۱۰-۲ یادگیری ماشین

به طور کلی یادگیری ماشین به سیستم‌ها اجازه می‌دهد تا به‌طور خودگردان و بدون هیچ پشتیبانی خارجی، تصمیم‌گیری کنند. این تصمیمات در شرایطی اتخاذ می‌شود که دستگاه قادر به یادگیری از داده باشد و الگوهای اساسی موجود در آن را درک کند. سپس، از طریق تطبیق الگو و تحلیل بیشتر، نتیجه را که می‌تواند یک کلاسه بندی یا یک پیش‌بینی باشد، برمی‌گردانند [۶۴]. سه نوع الگوریتم مهم یادگیری ماشین وجود دارد که در ادامه به آن‌ها اشاره شده است.

(۱) یادگیری نظارت شده

(۲) یادگیری بدون نظارت

(۳) یادگیری تقویتی

۱-۱۰-۲-۱ یادگیری نظارت شده

یادگیری نظارت شده محبوب‌ترین الگو برای انجام عملیات یادگیری ماشین است. به طور گسترده برای داده‌هایی استفاده می‌شود که نگاشت دقیقی بین داده‌های ورودی - خروجی وجود دارد. در این حالت، مجموعه داده‌ها، برچسب زده شده‌اند، به این معنی که الگوریتم، ویژگی‌ها را به طور واضح مشخص کرده و براساس آن، پیش‌بینی‌ها یا کلاسه‌بندی را انجام می‌دهد. با انجام مرحله آموزش، الگوریتم قادر به

شناسایی روابط بین دو متغیر است به گونه‌ای که می‌توان نتیجه جدیدی را پیش‌بینی کرد. در ادامه به بعضی از الگوریتم‌های این یادگیری اشاره می‌شود. با توجه به این که این تحقیق در حوزه تشخیص می‌باشد، روش یادگیری با نظارت مورد استفاده قرار می‌گیرد، چون تمامی تصاویر ریه دارای برچسب هستند [۶۵].

۱) رگرسیون خطی

در رگرسیون خطی، رابطه خطی بین دو یا چند متغیر اندازه‌گیری می‌شود و براساس این رابطه، پیش‌بینی‌هایی را انجام می‌دهد که از این الگوی خطی پیروی می‌کنند [۶۶].

۲) تقویت گرادیان

تقویت گرادیان، یک روش یادگیری گروهی است که مجموعه‌ای از چندین درخت تصمیم‌گیری ضعیف می‌باشد که به یک کلاسه‌بند قدرتمند منتهی می‌شود [۶۷].

۳) ماشین بردار پشتیبان

SVMها کلاسه‌بندهای قدرتمندی هستند که با کمک ابرصفحه‌ها، برای طبقه‌بندی مجموعه داده‌های دودویی (باینری) به دو کلاس، استفاده می‌شوند [۶۸].

۴) رگرسیون لجستیک

این روش با استفاده از منحنی S زنگوله‌ای شکل ایجاد می‌شود که با کمک تابع logit برای دسته‌بندی داده‌ها در کلاس مربوطه‌شان تولید می‌شود [۶۹].

۵) شبکه‌های عصبی مصنوعی

شبکه‌های عصبی مصنوعی به تقلید از مغز انسان مدل‌سازی می‌شوند و در طول زمان از داده‌ها یاد می‌گیرند. آن‌ها بخش وسیع‌تری از یادگیری ماشین را تشکیل می‌دهند که یادگیری عمیق نامیده می‌شود.

۲-۱۰-۱-۲ یادگیری بدون نظارت

در مورد الگوریتم یادگیری بدون نظارت، داده‌ها صریحاً در کلاس‌های مختلف برچسب‌گذاری نمی‌شوند، یعنی هیچ برچسبی وجود ندارد. این مدل با یافتن الگوهای ضمنی قادر به یادگیری داده است. الگوریتم‌های یادگیری بدون نظارت، داده‌ها را براساس تراکم، ساختار، بخش‌های مشابه و دیگر ویژگی‌های مشابه، شناسایی می‌کنند [۷۰].

۲-۱۰-۱-۳ یادگیری تقویتی

یادگیری تقویتی، حوزه‌های بیشتری از هوش مصنوعی را در بر می‌گیرد که به ماشین‌ها اجازه می‌دهد تا برای رسیدن به اهداف خود با محیط پویای اطراف خود در تعامل باشند. با این کار، ماشین‌ها و عامل‌های نرم افزاری می‌توانند رفتار ایده آل را در یک زمینه خاص ارزیابی کنند. با کمک این بازخورد پاداش، عامل‌ها قادر به یادگیری رفتار و بهبود آن در دراز مدت هستند. این پاداش بازخورد ساده به عنوان یک سیگنال تقویتی شناخته می‌شود [۷۱].

۲-۱۰-۲ مقدمه ای بر یادگیری عمیق

همان‌طور که گفته شد، یادگیری عمیق نوعی از یادگیری ماشین و هوش مصنوعی است که در واقع از روشی که ذهن انسان برای یادگیری موضوع خاصی به کار می‌گیرد، تقلید می‌کند. این نوع از یادگیری یکی از عناصر مهم در علم داده می‌باشد که شامل آمار و مدل‌سازی است. یادگیری عمیق برای دانشمندان که وظیفه جمع‌آوری، تجزیه، تحلیل و تفسیر مقادیر زیادی از داده‌ها را دارند، بسیار مفید

است و این روند را سریع‌تر و آسان‌تر می‌کند [۷۲]. یادگیری عمیق کاربرد فزاینده در تمامی عرصه‌ها مانند پردازش تصویر، رباتیک، داده‌کاوی و پردازش گفتار دارد. یکی از دلایل استفاده مکرر آن می‌تواند وجود واحدهای پردازش مرکزی سریع (CPUs)^۱ و واحدهای پردازش گرافیکی (GPUs)^۲ باشد که باعث ایجاد سرعت در روند دسترسی به مجموعه‌های زیادی از داده‌ها شده‌اند و زمان آموزش و اجرا را کاهش داده است [۷۳].

چهار ساختار مهم از شبکه‌های عمیق را می‌توان به صورت زیر تقسیم کرد [۷۴]:

- شبکه‌های عصبی کانولوشنی^۳

- شبکه‌های از پیش آموزش دیده بدون ناظر^۴

- شبکه‌های عصبی مکرر^۵

- شبکه‌های عصبی بازگشتی^۶

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

۱۱-۲ جمع‌بندی

مفاهیم مورد نیاز در این پایان‌نامه، در این فصل مورد بررسی قرار گرفت و کارهای گذشته انجام شده در قسمت بخش‌بندی و طبقه‌بندی به اختصار شرح داده شد. همان‌طور که مشاهده شد، روش‌های بخش‌بندی به دو دسته مبتنی بر یادگیری عمیق و سنتی تقسیم شد و دریافتیم که اشکال مهم روش‌های سنتی، اغلب به وجود عامل انسانی و کند بودن فرآیند پردازش برمی‌گردد. این مشکل هم در

^۱ Central Processing Units CPUs:

^۲ GPUs: Graphics Processing Units

^۳ Convolutional neural networks (CNNs)

^۴ Unsupervised Pretrained Networks

^۵ Recurrent Neural Networks (RNNs)

^۶ Recursive Neural Network

قسمت طبقه‌بندی وجود داشت. در ادامه شرح مختصری از یادگیری ماشین و مقدمه‌ای بر یادگیری عمیق ارائه کردیم. دریافتیم تمامی مراحل این تحقیق می‌تواند بر اساس یادگیری عمیق پیش برود تا مشکلات ذکر شده در بالا را، از بین ببرد. در فصل بعد یعنی فصل سوم، به توضیح تعدادی از معماری‌های یادگیری عمیق پرداخته و روش‌های مبتنی بر آن را برای بخش‌بندی تصاویر و تشخیص و طبقه‌بندی، تشریح خواهیم کرد.

فصل ۳. ساختارهای یادگیری عمیق

۱-۳ مقدمه

در این فصل، ابتدا شبکه‌های عصبی کانولوشنی که مهم‌ترین و کاربردی‌ترین ساختار یادگیری عمیق هستند را مورد بررسی قرار داده و در ادامه مروری بر معماری‌های مبتنی بر این ساختار و روش‌های بخش‌بندی تصاویر و طبقه‌بندی در یادگیری عمیق خواهیم داشت.

۲-۳ شبکه‌های عصبی کانولوشنی

یادگیری عمیق ساختارهای متفاوتی دارد اما در بین آن‌ها شبکه‌های عصبی کانولوشنی یکی از پرکاربردترین و مهم‌ترین ساختارها هستند. به این شبکه به اختصار CNN یا ConvNet گفته می‌شود. شبکه CNN در سال ۱۹۹۰ با الهام‌گیری از آزمایش‌های انجام‌شده توسط Hubel و Wiesel روی قشر بینایی^۱ معرفی شد. یکی از اولین پروژه‌های انجام‌شده با این شبکه، پروژه معروف شناسایی ارقام دست‌نویس MNIST در سال ۱۹۹۸ توسط Yann Lecun بود که نتایج امیدوارکننده‌ای به همراه داشت [۷۵]. همان‌طور که در فصل دوم اشاره شد، کاربرد اصلی این شبکه‌ها، در مسائل طبقه‌بندی است. درخشش اصلی شبکه کانولوشنی در سال ۲۰۱۲ در مسابقه ImageNet اتفاق افتاد. رویداد ImageNet یک چالش دسته‌بندی با ۱/۲ میلیون تصویر در ۱۰۰۰ کلاس بود و شبکه AlexNet با خطای ۱۶/۴٪ به رتبه اول مسابقه رسید. این شبکه‌ها به‌گونه‌ای طراحی شده‌اند که برای ورودی‌های با ساختار ماتریسی (دوبعدی و سه‌بعدی) به‌خوبی کار می‌کنند. شبکه پرسپترون چند لایه، ساختار داده‌های ورودی را عوض می‌کند و یک ماتریس دوبعدی 100×100 را تبدیل به یک بردار به ابعاد ۱۰۰۰۰ می‌کند. اما شبکه CNN ساختار ورودی را عوض نمی‌کند و به ارتباط بین پیکسل‌های همسایه اهمیت می‌دهد. این شبکه علاوه بر تصاویر دوبعدی می‌تواند برای تصاویر زیر هم مناسب باشد.

➤ داده‌های یک بعدی : سیگنال

➤ داده‌های دوبعدی : تصویر و طیف صوت

^۱ Visual Cortex

➤ داده‌های سه بعدی : ویدئو و تصاویر حجمی

➤ داده‌های چهاربعدی : تصاویر حجمی همراه با زمان

۳-۳ معماری شبکه عصبی کانولوشنی

شبکه کانولوشنی یک شبکه عصبی عمیق است و الگوریتم آن به صورت سلسله‌مراتبی می‌باشد. در شکل ۱-۳ یک مثال ساده از معماری شبکه عصبی کانولوشنی نشان داده شده است. شبکه کانولوشنی معمولاً از بلوک‌های مختلفی تشکیل شده که لایه‌های آن به صورت یکی در میان می‌باشد. در آخرین مرحله این شبکه تعدادی لایه تمام‌متصل وجود دارد [۷۶]. لایه‌ها یا بلوک‌های مختلف در شبکه CNN عبارتند از:

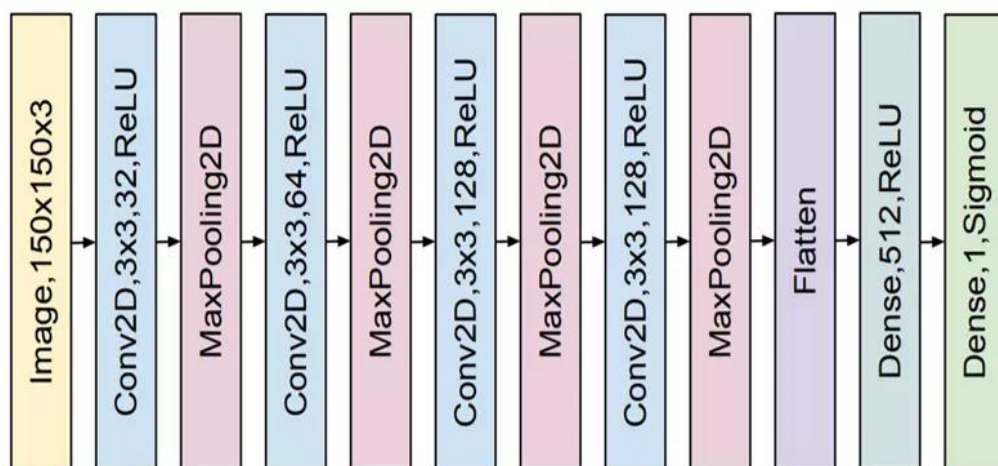
➤ لایه ورودی (بلوک زرد در شکل)

➤ لایه کانولوشنی (بلوک‌های آبی در شکل)

➤ لایه غیرخطی (بلوک‌های آبی در شکل - معمولاً تابع غیرخطی را همراه با لایه کانولوشنی یک‌جا نشان می‌دهند)

➤ لایه تجمیع (بلوک‌های قرمز در شکل)

➤ لایه تماماً متصل (بلوک‌های سبز در شکل)



شکل ۱-۳. معماری ساده شبکه عصبی کانولوشنی [۷۷].

۳-۱-۱ لایه کانولوشن

هسته اصلی شبکه CNN لایه کانولوشنی است که درصد اعظم محاسبات شبکه عصبی کانولوشنی را به خود اختصاص داده است. هر لایه کانولوشنی در شبکه عصبی کانولوشن، شامل مجموعه‌ای فیلترها است. از کانولوشن بین فیلترها و لایه ورودی، خروجی ساخته می‌شود. به خروجی لایه کانولوشنی، Feature Map گفته می‌شود. این لایه‌ها، الگوهای محلی را فرا می‌گیرند. به این صورت که اولین لایه آن، الگوهای کوچکی مانند لبه‌ها را شامل می‌شود. همان‌طور با بالا رفتن شماره لایه‌ها، فراگیری بالاتر می‌رود به طوری که الگوهای بزرگتری مانند چشم، گوش، اشکال هندسی و غیره را فرا می‌گیرد. به طور کلی کانولوشن پنجره‌هایی با اندازه 3×3 یا 5×5 را روی نقشه ویژگی^۱ ورودی می‌لغزاند و نسبت به ورودی‌اش تکه‌هایی را از نقشه ویژگی آن بیرون کشیده و سعی می‌کند به هدف شبکه برسد.

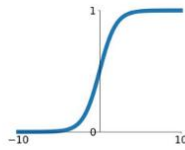
۳-۲-۱ لایه غیرخطی

مشابه با سایر شبکه‌های عصبی، شبکه عصبی کانولوشن هم از تابع تحریک^۲ بعد از لایه کانولوشنی استفاده می‌کند. استفاده از تابع غیرخطی باعث ایجاد خاصیت غیرخطی در شبکه عصبی می‌شود که خیلی مهم است. بعضی از فریمورک‌ها در همان تعریف لایه کانولوشنی، به شما این امکان را می‌دهند که نوع تابع غیرخطی را هم مشخص کنید. در بعضی از فریمورک‌ها هم باید یک لایه جداگانه بسازید. تعریف تابع غیرخطی به صورت جدا از لایه کانولوشنی انعطاف پذیری بیشتری ایجاد می‌کند. محاسبات موجود در تابع ReLU ساده و صرفاً یک مقایسه است. به همین دلیل محاسبات در بخش تابع غیرخطی با استفاده از ReLU نسبت به سایر توابع غیرخطی با سرعت بیشتری انجام می‌شود. از طرفی فرآیند آموزش با ReLU نسبت به سایر توابع غیرخطی سریع‌تر است. در شکل ۳-۲ توابع غیرخطی در شبکه عصبی کانولوشن را مشاهده می‌کنید.

^۱ Feature Map
^۲ activation function

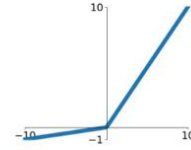
Sigmoid

$$\sigma(x) = \frac{1}{1+e^{-x}}$$



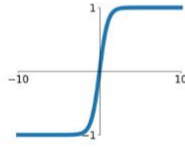
Leaky ReLU

$$\max(0.1x, x)$$



tanh

$$\tanh(x)$$

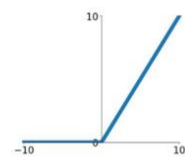


Maxout

$$\max(w_1^T x + b_1, w_2^T x + b_2)$$

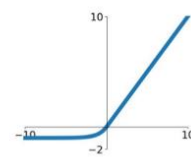
ReLU

$$\max(0, x)$$



ELU

$$\begin{cases} x & x \geq 0 \\ \alpha(e^x - 1) & x < 0 \end{cases}$$



شکل ۲-۳. توابع غیرخطی در شبکه عصبی کانولوشن

۳-۳-۳ لایه تجمیع

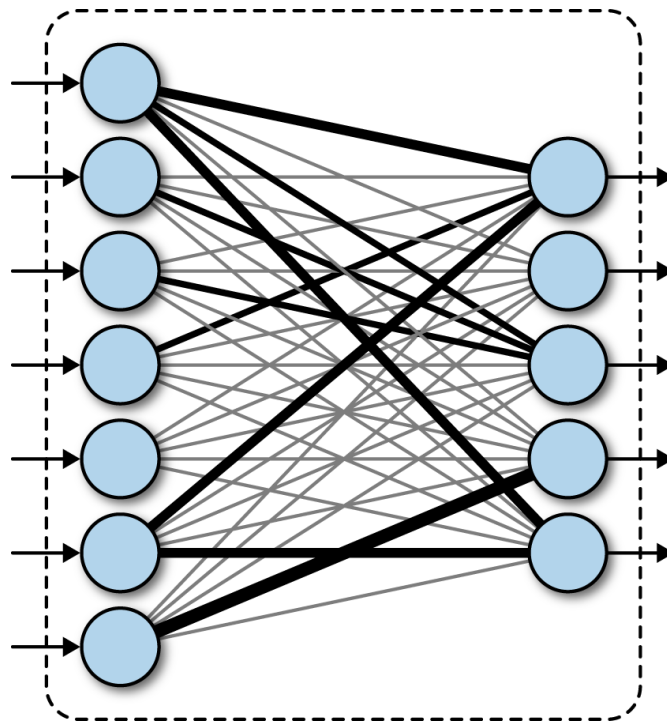
لایه تجمیع یکی دیگر از لایه‌های مهم در شبکه عصبی کانولوشن است. هدف لایه تجمیع کاهش اندازه مکانی Feature Map بدست آمده با استفاده از لایه کانولوشنی است. لایه تجمیع پارامتر قابل آموزش ندارد و صرفاً یک نمونه برداری ساده و موثر را انجام می‌دهد. تجمیع عملکردی شبیه کانولوشن دارد و یک پنجره روی تصویر حرکت می‌کند. به‌طور کلی بعد از هر لایه کانولوشن، در حالت عادی یک لایه تجمیع قرار دارد و ویژگی‌های مطلوب را حفظ کرده و موارد کم‌اهمیت‌تر از میان این ویژگی‌ها را حذف می‌کند. رایج‌ترین نمونه تجمیع max pooling و average pooling است.

۳-۳-۴ لایه تماماً متصل

به‌طور معمول آخرین لایه یک شبکه عصبی کانولوشن برای طبقه‌بندی را، لایه تماماً متصل تشکیل می‌دهد. یکی از کاربردهای اصلی این لایه در شبکه کانولوشن، استفاده به عنوان طبقه‌بند یا کلاسیفایر^۱ است. به‌طور معمول مجموعه ویژگی‌های استخراج شده با استفاده از لایه‌های کانولوشنی تبدیل به یک بردار می‌شوند. درنهایت این بردار ویژگی به یک طبقه‌بند تماماً متصل داده می‌شود تا

^۱ Classifier

کلاس درست را شناسایی کند. در شکل ۳-۳ یک نمونه لایه تماما متصل را می‌توانید مشاهده کنید.



شکل ۳-۳. معماری ساده شبکه تماما متصل

۳-۴ مفاهیم و الگوریتم‌ها در حوزه شبکه‌های عصبی کانولوشن

شبکه‌های عصبی کانولوشن امروزه نیروی پیش‌برنده در پس هر پیشرفتی در حوزه تحقیقات بینایی رایانه محسوب می‌شود. در این بخش، با توجه به این که کار موردنظر این پایان‌نامه در دو مرحله بخش‌بندی و تشخیص می‌باشد، بعد از معرفی مفهومی به‌نام پیش‌آموزش، مهم‌ترین الگوریتم‌ها بر پایه شبکه‌های کانولوشنی را معرفی می‌کنیم.

۳-۴-۱ پیش‌آموزش

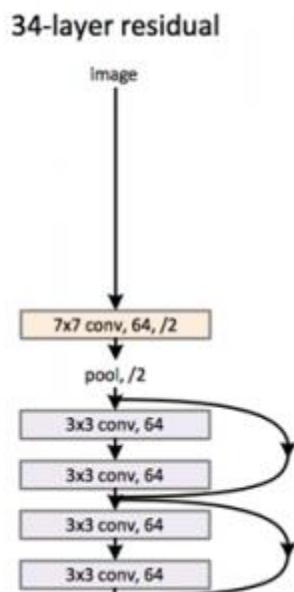
به دلیل وجود تعداد بالای پارامترها و الزام به داده‌های آموزش زیاد، آموزش شبکه‌های عمیق، زمان و سخت‌افزار بالا و مناسبی را نیاز دارد. این در حالی است که وقتی شبکه بر روی یک مجموعه داده خاص مثل ImageNet آموزش ببیند، اطلاعات کلی راجع به جنس داده یاد گرفته و فقط اطلاعات مربوط به

آن مجموعه را یاد نمی‌گیرد [۷۸]. این خصوصیت در این شبکه‌ها، باعث این شده است که بتوانیم از یک شبکه که برای یک وظیفه منحصر به فرد مثل تشخیص برچسب در داده‌های ImageNet آموزش داده شده، داخل کارهای دیگر مانند بخش‌بندی تصاویر پزشکی استفاده کنیم. این روش در مدل‌هایی که بر اساس CNN هستند بسیار محبوب است چون قابلیت تعمیم مدل را افزایش و فرآیند یادگیری را تسریع می‌بخشد. یکی از انواع این شبکه‌ها، شبکه ResNet می‌باشد که در ادامه توضیح مختصری درباره آن داده می‌شود.

۳-۴-۲ الگوریتم ResNet

ResNet^۱ یکی از معماری‌های بزرگ است که نشان می‌دهد یک معماری یادگیری عمیق تا چه حد می‌تواند عمیق باشد. در سال ۲۰۱۵ تعدادی از محققین مایکروسافت معماری ResNet را ارائه دادند که خطا را به ۳٫۶ درصد کاهش داد. این معماری از ۱۵۲ لایه تشکیل شده است که یک رکورد در هر سه بخش تشخیص اشیاء، کلاسه‌بندی و localization با استفاده از یک مدل به جا گذاشت [۷۹]. این مدل در حقیقت قدرت انسان در تشخیص اشیاء را پشت سر گذاشت و با دقت به مراتب بالاتر از انسان‌ها، اشیاء را تشخیص داد. ایده اصلی یک بلاک residual این است که ورودی توسط یک لایه کانولوشن، لایه ReLU و یک لایه کانولوشن پردازش می‌شود. نتیجه این تبدیلات یک تابع $F(x)$ تولید می‌کند و نتیجه این تابع با ورودی لایه قبل جمع می‌شود.

^۱ Residual Networks



شکل ۴-۳. معماری ساده شبکه resnet

به طور کلی همه بر این باورند که با افزایش عمق شبکه، دقت شبکه نیز افزایش می‌یابد اما وقتی عمق شبکه از یک حد خاصی بیشتر شود، خطا شروع به بالا رفتن می‌کند. بلوک‌های باقی‌مانده^۱ در این الگوریتم ورودی‌های لایه قبلی را بدون واسطه به لایه بعدی منتقل کرده و باعث می‌شود خطای هر لایه به لایه قبلی انتقال یابد که این کار باعث بالا رفتن سرعت و عمیق‌تر شدن شبکه می‌شود.

۳-۴-۳ الگوریتم U-Net

معماری U-Net یک شبکه عصبی پیچشی است که دراصل برای بخش‌بندی تصاویر در حوزه پزشکی طراحی شده است. شکل این معماری شبیه حرف U انگلیسی است و به همین دلیل نیز U-Net یا شبکه U شکل نامیده می‌شود. این الگوریتم در سال ۲۰۱۵ در دانشگاه فرایبورگ آلمان توسط اولاف‌رنه‌برگر^۲ و همکارانش به منظور بخش‌بندی تصویر طراحی شده که در این شبکه سرعت و دقت پردازش افزایش یافت. این شبکه، هیچ لایه تماماً متصل ندارد و همین نیاز شبکه به داده‌های آموزشی زیاد را از بین می‌برد. این ویژگی باعث قطعه‌بندی بسیار دقیق اجزای تصویر می‌شود. معماری این شبکه از دو بخش

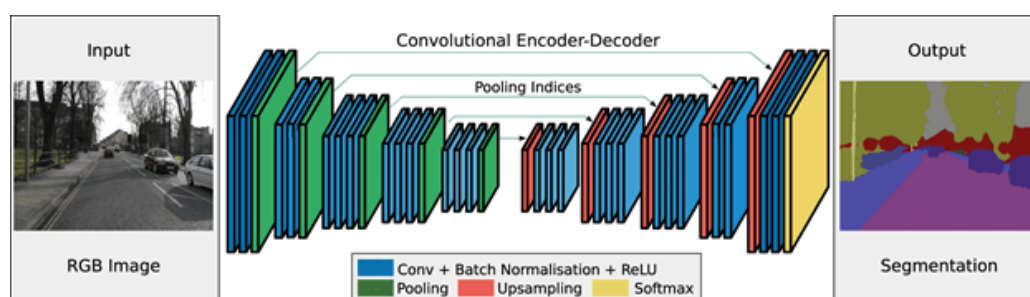
^۱ Residual

^۲ Olaf Ronneberger

همان‌طور که در شکل بالا ملاحظه می‌کنید، هر مستطیل آبی رنگ متناظر با یک خروجی چندکاناله است. تعداد کانال‌های خروجی در بالای هر کادر مستطیل‌شکل و ابعاد آن نیز در گوشه پایین و سمت چپ مستطیل‌ها درج شده است. کادرهای مستطیلی سفید رنگ نسخه کپی شده خروجی‌ها هستند. هر پیکان نیز نمایان‌گر یک عمل خاص می‌باشد.

۳-۵ بخش‌بندی تصاویر پزشکی مبتنی بر یادگیری عمیق

در فصل دوم، به طور مختصر روش‌های بخش‌بندی سنتی را بررسی کردیم. در این فصل روش‌های مبتنی بر یادگیری عمیق در حیطه بخش‌بندی تصاویر پزشکی را مورد بررسی قرار خواهیم داد. معماری و ساختار اصلی در بخش‌بندی تصویر شامل یک کدگذار و یک کدگشا است. بخش کدگذار ویژگی‌های خاص هر تصویر را به کمک فیلترها استخراج می‌کند و بخش کدگشا نیز مسئولیت تولید خروجی نهایی را برعهده دارد که به‌طور معمول در آن یک ماسک بخش‌بندی‌کننده، طرح کلی شیء را مشخص می‌کند. همان‌طور که در شکل ۳-۶ مشاهده می‌کنید، معماری فرآیندهای بخش‌بندی تصویر، در اکثر مواقع، مشابه این معماری هستند.



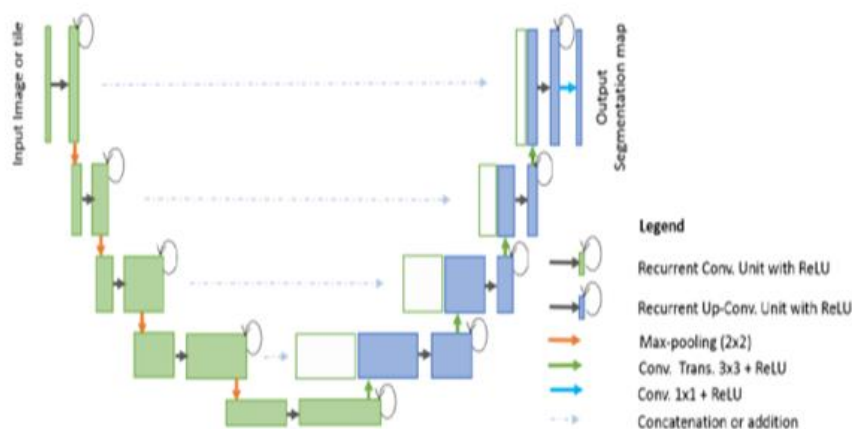
شکل ۳-۶. ساختار بخش‌بندی تصویر

همان‌طور که در بخش‌های قبلی گفته شد، شبکه U-Net در میان معماری‌های مختلف از خانواده شبکه‌های کانولوشن، متداول‌ترین و مناسب‌ترین شبکه‌ی طراحی‌شده برای بخش‌بندی تصاویر پزشکی به حساب می‌آید. این شبکه توانسته است در حوزه بخش‌بندی تصاویر پزشکی، به لبه تکنولوژی دست یابد. همان‌طور که گفته شد، این شبکه کاربردهای فراوانی در بخش‌بندی تصویر دارد که برای مثال، در مرجع

[۸۱] از این شبکه یعنی U-Net به منظور تقسیم غضروف و منیسک^۱ از تصاویر ام‌آر‌آی زانو بهره برده شده‌است. با توجه به پرکاربرد بودن این الگوریتم، در سال‌های اخیر، در راستای بهبود معماری این شبکه، با دست‌کاری و ایجاد تغییراتی در آن، نسخه‌های مختلفی ارائه شده‌است که در ادامه به بررسی چند نسخه تغییر یافته از شبکه U-Net خواهیم پرداخت.

۳-۵-۱ ساختار RU-Net

یکی از نسخه‌های تغییر یافته U-Net، با تلفیق ساختار کانولوشنی RCNN^۲ و U-Net به دست آمده است [۸۲]. نوآوری این شبکه، در مسیر انقباض، قبل از لایه‌های DownSampling و در مسیر انبساط، قبل از لایه‌های UpSampling است و از لایه‌های کانولوشنی بیشتری استفاده کرده است. استفاده از لایه‌های کانولوشنی مکرر، مزیت این شبکه است که موجب استخراج ویژگی‌های مفیدتر در مسیر انبساطی شبکه می‌شود. معماری کلی این شبکه در شکل ۳-۷ قابل مشاهده می‌باشد.



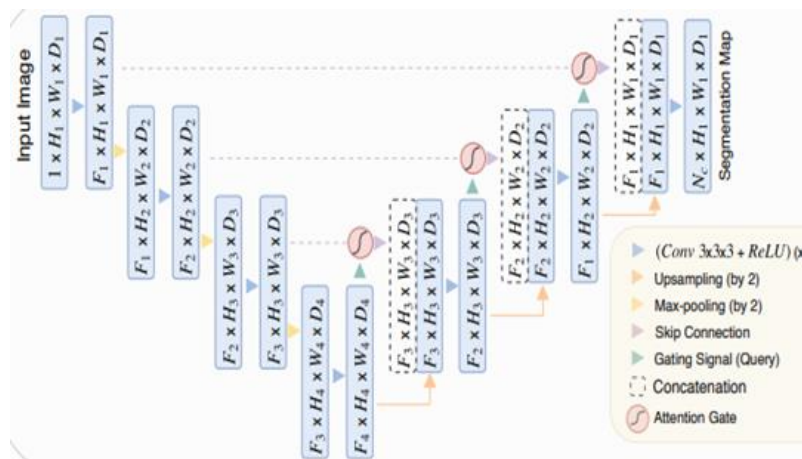
شکل ۳-۷. معماری شبکه RU-Net [۸۲].

^۱ Meniscus

^۲ Recurrent Convolutional Neural Network

۳-۵-۲ ساختار Attention U-Net

در ساختار شبکه U-Net، در مسیر انبساط، محتوای لایه‌ی قبلی با محتوای لایه‌ی معادل از مسیر انقباض، الحاق می‌شود که این کار برای بازیابی اطلاعات مکانی تصویر در هر لایه می‌باشد. این رویه باعث استخراج ویژگی‌های سطح پایین و اضافی می‌شود. ساختار Attention U-Net با استفاده از دروازه‌هایی به نام Attention در اتصالات پرش این مشکل را حل کرد. یعنی به این گونه عمل می‌کرد که نقشه‌ویژگی‌های^۱ استخراج‌شده در اتصالات پرش، پس از پردازش توسط دروازه‌های Attention، با محتوای لایه‌ی قبلی در مسیر انبساط، الحاق می‌شوند. این تغییر باعث قدرت تعمیم بهتر و کاهش منابع محاسباتی هدر رفته می‌شود. شکل ۸-۳، معماری این شبکه را نشان می‌دهد [۸۳].



شکل ۸-۳. معماری شبکه Attention U-Net [۸۳].

۳-۶ استخراج ویژگی و طبقه‌بندی

در فصل دوم و در بخش‌های قبل ccn را به طور کامل توضیح دادیم و بخش‌های مختلف آن را به طور کامل بررسی کردیم. همان‌طور که در بخش‌های قبلی گفته شد، cnn شامل لایه‌ی کانولوشنی، لایه تجمیع و لایه تمام متصل است. هر کدام از این لایه‌ها وظیفه منحصر به فرد خود را دارند. ما برای

^۱ Feature Maps

استخراج ویژگی و طبقه‌بندی از حالت سه بعدی cnn استفاده کرده‌ایم که در ادامه به بررسی این شبکه خواهیم پرداخت و دلیل استفاده از این شبکه را خواهیم گفت.

۳-۶-۱ شبکه عصبی کانولوشنی سه بعدی

شبکه عصبی کانولوشنی به طور معمول دارای سه بعد می‌باشد که هر کدام از این شبکه‌ها در جای به خصوص خود استفاده می‌شوند. در ادامه بررسی مختصری از این سه دسته شده است.

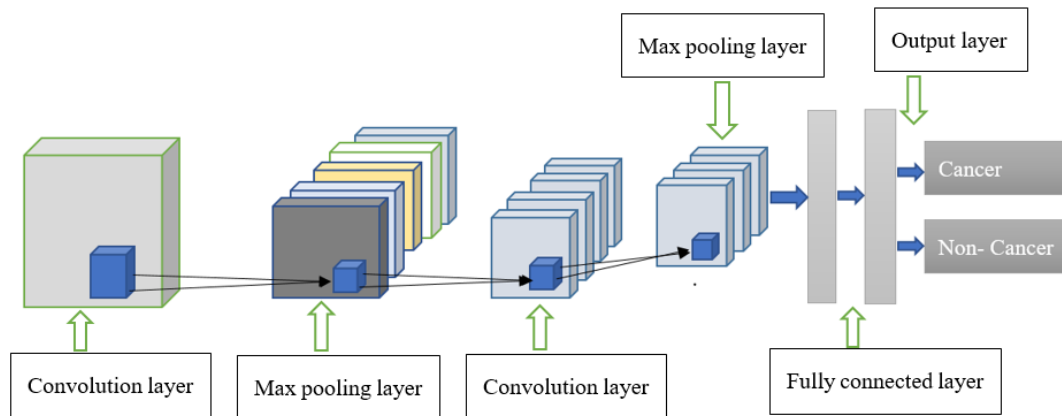
(۱) در شبکه عصبی کانولوشنی یک بعدی، هسته در یک جهت حرکت می‌کند و داده‌های ورودی و خروجی دو بعدی هستند. از این شبکه عمدتاً برای داده‌های سری زمانی استفاده می‌شود.

(۲) در شبکه عصبی کانولوشنی دو بعدی، هسته در دو جهت حرکت می‌کند و داده‌های ورودی و خروجی سه بعدی هستند. از این شبکه عمدتاً برای داده‌های تصویر استفاده می‌شود.

(۳) در شبکه عصبی کانولوشنی سه بعدی، هسته در سه جهت حرکت می‌کند و داده‌های ورودی و خروجی چهار بعدی هستند. از این شبکه معمولاً برای داده‌های تصویر سه بعدی (ام‌آر‌آی و سی‌تی‌اسکن) استفاده می‌شود.

با توجه به کار این تحقیق و این که اندازه هر گره سرطانی متفاوت است و گره بسیار کوچک است، برای شناسایی آن‌ها، از شبکه عصبی کانولوشن سه بعدی استفاده می‌شود. شبکه عصبی کانولوشنی حالت دو بعدی هم می‌تواند در این تحقیق استفاده شود اما با استفاده از این شبکه، ممکن است اطلاعات مهم و ارزشمند در مورد گره از دست برود. به همین دلیل و با توجه به تحقیقاتی که انجام شده است، از یک شبکه عصبی کانولوشنی سه بعدی استفاده می‌شود. در این شبکه تصاویر ورودی سه بعدی هستند. همان‌طور که در فصل قبل گفته شد، شبکه کانولوشنی از لایه‌های مختلفی تشکیل شده‌اند که هر کدام

وظیفه خود را دارند. شکل ۳-۹ یک نمونه از آن را نشان می‌دهد که تمامی لایه‌های این شبکه بررسی شده است. در فصل چهارم به طور دقیق‌تر آن را موشکافی خواهیم کرد [۸۴].



شکل ۳-۹. معماری ۳D cnn سه بعدی [۸۴]

۳-۶-۲ تقویت گرادیان

تقویت گرادیان یک روش یادگیری ماشین برای مسائل رگرسیون و طبقه‌بندی است، که یک مدل پیش‌بینی‌کننده را در قالب مجموعه‌ای از مدل‌های پیش‌بینی‌کننده ضعیف ایجاد می‌کند. همان‌طور که در معماری شبکه عصبی کانولوشنی قابل ملاحظه است، برای طبقه‌بندی از لایه تماماً متصل استفاده شده است. برای بهبود خروجی در مرحله طبقه‌بندی، از روش تقویت گرادیان استفاده خواهیم کرد که در ادامه به توضیح آن پرداخته می‌شود [۸۵]. به طور کلی دقت یک مدل پیش‌بینی‌کننده را می‌توان به دو روش تقویت کرد:

الف) با استفاده از مهندسی ویژگی

ب) با اعمال مستقیم الگوریتم‌های تقویت کننده

الگوریتم‌های تقویت کننده زیادی وجود دارند که باعث بهبود خروجی در این مرحله می‌شود، که در زیر قابل مشاهده می‌باشند. در این تحقیق از روش XGBoost برای بهبود طبقه‌بندی استفاده می‌شود.

- Gradient Boosting
- XGBoost
- AdaBoost

- Gentle Boost

۱-۲-۶-۳ شبکه XGBoost

XGBoost الگوریتمی است که اخیراً در حوزه یادگیری ماشین به کار گرفته می‌شود. الگوریتم XGBoost یک پیاده‌سازی از تقویت گرادیان درخت تصمیم‌گیری است. این شبکه برای سرعت و کارایی بالا طراحی شده است. اصولاً XGBoosting نوعی کتابخانه نرم‌افزاری است که می‌توان آن را دانلود و روی ماشین نصب و سپس باید از رابط‌های مختلفی به آن دسترسی پیدا کرد. دو ویژگی بزرگ این شبکه که باعث شده است امروزه کارایی بسیار زیادی در پروژه‌های یادگیری ماشین داشته باشد، سرعت اجرا و عملکرد مناسب آن در طبقه‌بندی می‌باشد [۸۵].

➤ سرعت اجرای XGBoost

XGBoosting در مقایسه با پیاده‌سازی تقویت گرادیان، بسیار سریع است. الگوریتم XGBoost با سایر پیاده‌سازی‌های تقویت گرادیان و درختان تصمیم‌گیری مقایسه شده است و دارای سرعت بالاتری نسبت به بقیه روش‌ها می‌باشد.

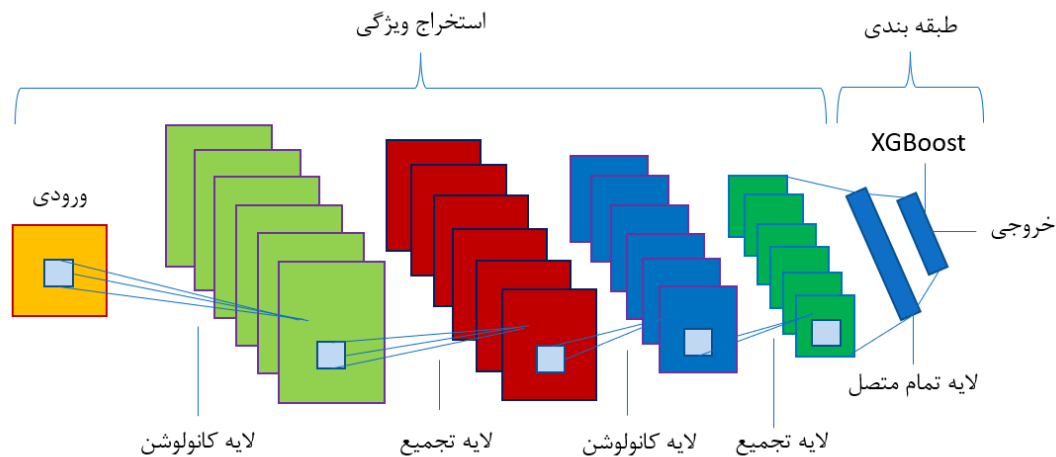
➤ عملکرد مدل XGBoost

این مدل بر مجموعه داده‌های ساخت‌یافته در مسائل مدلسازی پیشگویانه طبقه‌بندی و رگرسیون حاکم است. این الگوریتم در مسابقات Kaggle رتبه اول را به دست آورده است و به همین دلایل در این تحقیق از این الگوریتم استفاده است.

با توجه به تحقیقاتی که صورت گرفته و خروجی‌هایی که موجود است، از شبکه XGBoost برای طبقه‌بندی نوع سرطان‌ها استفاده می‌شود. این شبکه در آخرین لایه و بعد از لایه‌ی تماماً متصل قرار

دارد. به این صورت که خروجی لایه تماماً متصل به این شبکه داده می‌شود و طبقه‌بندی صورت می‌گیرد.

معماری کلی شبکه کانولوشنی سه بعدی و XGBoost در شکل ۳-۱۰ قابل مشاهده می‌باشد.



شکل ۳-۱۰: معماری CNN- XGBoost

۳-۷ جمع‌بندی

در ابتدای این فصل، یادگیری عمیق توضیح داده شده و در ادامه یکی از مهم‌ترین ساختارهای آن یعنی شبکه‌های عصبی کانولوشنی را مورد بررسی قرار دادیم. با توجه به این که این تحقیق شامل دو بخش مهم بوده است. در بخش اول به بخش‌بندی تصاویر پزشکی با تکیه بر معماری‌های مبتنی بر شبکه‌های کانولوشنی پرداخته و به این نتیجه رسیدیم که مهم‌ترین شبکه‌ی کانولوشنی جهت بخش‌بندی تصاویر پزشکی شبکه U-Net است. در بخش دوم، به استخراج ویژگی و طبقه‌بندی مبتنی بر یادگیری عمیق پرداخته و دریافتیم که بهترین الگوریتم برای استخراج ویژگی، cnn سه بعدی است. در ادامه برای طبقه‌بندی، لایه تماماً متصل است. در فصل چهارم، با استفاده از مدل U-Net تصاویر را بخش‌بندی کرده و به وسیله cnn سه بعدی و لایه تماماً متصل، ویژگی‌ها را استخراج می‌کنیم. در این پایان‌نامه بر خلاف دیگر تحقیقات، شبکه cnn سه بعدی و لایه تماماً متصل طبقه‌بندی را انجام نمی‌دهند، بلکه خروجی این مرحله را به شبکه xgboost داده و این لایه که آخرین لایه شبکه است، طبقه‌بندی را انجام می‌دهد.

فصل ۴ روش پیشنهادی

۴-۱ مقدمه

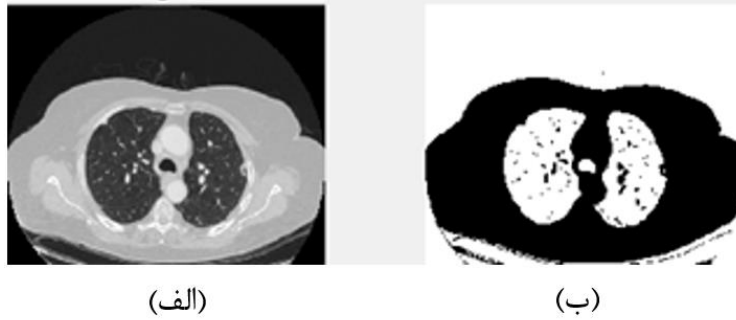
هدف اصلی این رساله، تشخیص سرطان ریه با استفاده از یادگیری عمیق است. همان‌طور که در فصل‌های گذشته گفته شد، این تحقیق از دو بخش اصلی بخش‌بندی تصویر و طبقه‌بندی تشکیل شده است. در این بخش ابتدا پیش‌پردازش‌ها و آماده‌سازی داده‌ها شرح داده خواهد شد و در آخر دو بخش اصلی تحقیق یعنی مدل یادگیری عمیق پیشنهادی برای مرحله بخش‌بندی و یادگیری عمیق پیشنهادی برای مرحله طبقه‌بندی به طور کامل مورد بررسی قرار خواهد گرفت و در انتها یک جمع‌بندی از این فصل ارائه خواهد شد.

۴-۲ پیش‌پردازش و آماده‌سازی داده‌ها

این مرحله یکی از مراحل اصلی است که در اغلب تحقیقات با هم متفاوت است. این مغایرت ناشی از تفاوت در مجموعه داده‌ها و ورودی تصویر است. هدف این مرحله، فراهم کردن تصاویر خام ورودی برای آموزش شبکه یادگیری عمیق است. در این تحقیق از سه پیش‌پردازش استفاده شده است که هر کدام وظیفه به خصوصی دارند که به طور کامل ذکر خواهد شد.

۴-۲-۱ دودویی کردن تصویر

یکی از پیش‌پردازش‌هایی که بر روی تصاویر مجموعه داده‌ها این تحقیق انجام شده است، دودویی کردن تصویر است. در این پیش‌پردازش تصویر به حالت دودویی در می‌آید. یعنی تصویر دارای دو مقدار در سطح خاکستری (سیاه و سفید) می‌شود [۸۷]. این فرآیند برای جدا کردن قسمت‌های ریه از هم صورت می‌گیرد و بافت‌های ریوی، شامل رنگ سیاه با مقدار منطقی صفر هستند. در شکل ۴-۱ فرآیند دودویی کردن تصویر نشان داده شده است.



شکل ۴-۱: تصویر (الف)، تصویر اصلی و تصویر (ب)، تصویر دودویی شده است.

۴-۲-۲ عملگر لبه‌یابی

فرآیند تشخیص لبه، رئوس یک جسم و مرزهای بین اشیا 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 کرده‌ایم.



شکل ۴-۲: خروجی پیش‌پردازش عملگر لبه‌یابی

۴-۲-۳ 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. 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 ۳-۴: تصویر (الف)، تصویر اصلی و تصویر (ب)، خروجی پیش پردازش عملیات گسترش

همان‌طور که در شکل ۳-۴ مشاهده می‌شود، پس از اعمال عملگر گسترش بر روی تصویر باینری، قسمت سفید اطراف ریه در بخش شکل سمت راست که با فلش مشخص شده‌است، حذف گردیده و حفره‌های سیاه ریز داخل دو ریه چپ و راست، نیز محو شدند.

۳-۴ مدل یادگیری عمیق پیشنهادی برای مرحله بخش‌بندی

هدف اصلی این پایان‌نامه، تشخیص سرطان ریه است. یکی از مراحل که نیاز است به طور دقیق شرح داده شود و پیش‌نیاز قسمت طبقه‌بندی است، مرحله‌ی بخش‌بندی تصویر است. همان‌طور که در فصل قبل بررسی شد، شبکه عمیق U-Net یکی از بهترین معماری‌های مبتنی بر یادگیری عمیق برای بخش‌بندی تصویر است. این شبکه مزیت‌هایی زیادی نسبت به دیگر شبکه‌ها دارد که در ادامه به دو نمونه از برتری‌های این شبکه نسبت به شبکه‌های دیگر اشاره می‌شود.

(۱) در یادگیری عمیق، هر چه تعداد داده‌ها بیشتر باشد، آموزش بهتر انجام می‌شود. به دست آوردن

یک مجموعه کامل و بزرگ از تصاویر پزشکی کار بسیار سختی می‌باشد. این شبکه قادر است با

مجموعه داده نسبتاً کوچک، به خوبی آموزش ببیند و نتایج خوبی را به ارمغان بیاورد.

(۲) این شبکه می‌تواند به طور هم‌زمان از ویژگی‌های محتوایی و مکانی تصویر استفاده کند.

همان‌طور که در فصل سوم گفته شد، در سال‌های اخیر نسخه‌های مختلفی از این شبکه ارائه شده

است که تمام این نسخه‌ها برای ایجاد تغییرات و بهبود عملکرد این شبکه بوده‌اند. یکی از ساختار مهم و

به روز این شبکه، BCDU-Net است. این نسخه از سایر نسخه‌های تغییر یافته U-Net، بهتر عمل می‌کند.

ما در این تحقیق از این شبکه استفاده خواهیم کرد.

۴-۳-۱ ساختار شبکه BCDU-Net

در سال ۲۰۱۹ یکی از بهترین نسخه‌های شبکه U-Net ارائه شد و طبق آزمایش‌های انجام شده، این

شبکه نسبت به سایر نسخه‌ها، عملکرد بهتری داشت [۸۸]. در این شبکه، مسیر کدگذاری دارای چهار

مرحله است. هر مرحله دارای دو فیلتر کانولوشنی 3×3 و بعد از آن، یک تابع Pooling 2×2 و فعال‌ساز

Relu است. در این شبکه، به تدریج، بازنمایی تصاویر استخراج شده و در ادامه بعدهای لایه‌های بازنمایی

لایه به لایه افزایش می‌یابند. در آخر، در مسیر کدگذاری، یک نمایش تصویر ابعادی بالا با اطلاعات

معنایی بسیار بالا و قابل قبول تولید می‌شود. در حالت عادی شبکه U-Net، ویژگی‌های اضافی را در

کانولوشن‌های پیوسته، یاد می‌گیرد که به دلیل وجود دنباله‌ای از لایه‌های کانولوشنی است. در این نسخه،

این مشکل با استفاده از مکانیزم لایه‌های کانولوشنی متراکم^۱ در آخرین لایه کانولوشنی مسیر کدگذاری،

حل شده است. این مکانیزم برای اولین بار در سال ۲۰۱۷ و در مرجع [۸۹] مطرح شد و با توجه به بخش

۳-۴-۲، این ایده از روی شبکه ResNet گرفته شده است. این روش نسب به شبکه ResNet، تفاوت دارد

^۱ Densely connected convolutions

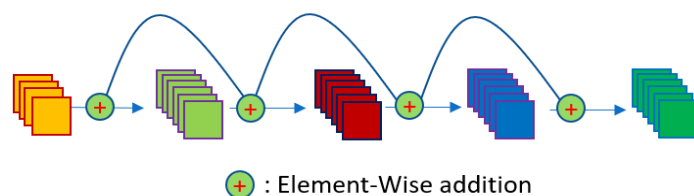
و به دلیل همین تفاوت، برتری این روش به وجود آمده است. در این روش بر خلاف ResNet، ترکیب ویژگی‌ها از طریق الحاق ویژگی‌ها صورت می‌گیرد. در صورتی که در شبکه ResNet ترکیب ویژگی‌ها از طریق جمع کردن، صورت می‌پذیرد. استفاده از لایه‌های کانولوشنی متراکم به جای لایه‌های کانولوشنی معمولی، مزایایی دارد که در ادامه و به طور مختصر به آن‌ها اشاره می‌شود.

(۱) جلوگیری از اثر محوشدگی گرادیان.

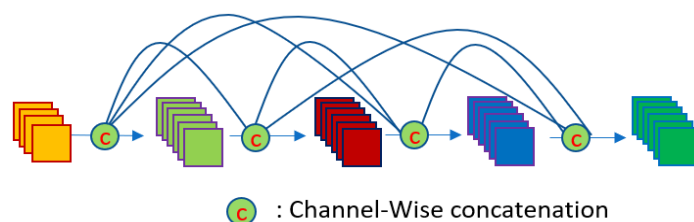
(۲) بالا رفتن قدرت بازنمایی شبکه، به وسیله‌ی انتشار جریان اطلاعات از طریق شبکه و استفاده مجدد از ویژگی‌ها.

(۳) یادگیری مجموعه متنوعی از نقشه‌های مشخصه.

به طور کلی و با استفاده از این این مکانیزم در آخرین لایه کانولوشنی مسیر کدگذاری، ویژگی‌های یاد گرفته شده در هر بلوک، به بلوک بعدی منتقل می‌شوند. این رویه باعث می‌شود که شبکه از یادگیری ویژگی‌های اضافی اجتناب کند. شکل ۴-۴، نشان‌دهنده‌ی بلوک Res و بلوک Dense است.



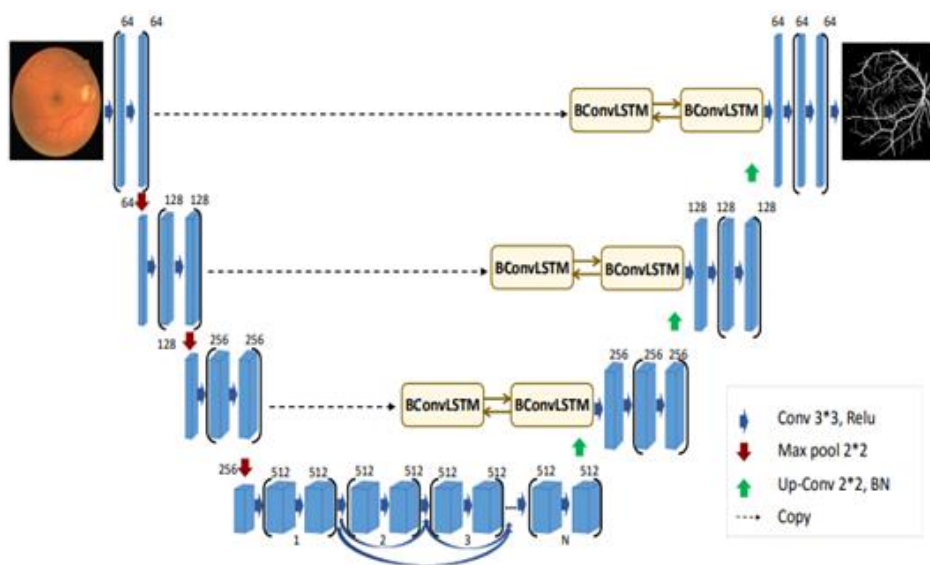
(الف)



(ب)

شکل ۴-۴. تصویر (الف) بلوک‌های شبکه ResNet و تصویر (ب) بلوک‌های شبکه DenseNet

در مسیر رمزگشایی یا همان مسیر انبساط، جهت بازیابی اطلاعات مکانی تصویر در هر لایه، محتوای لایه قبلی در مسیر انبساط با محتوای لایه معادل از مسیر انقباض، الحاق می‌شود که اصطلاحاً به آن، پرش اتصال^۱ گفته می‌شود. برای ترکیب نقشه‌های ویژگی، از ConvLSTM دو جهته (BConvLSTM) استفاده می‌شود که پردازش پیچیده‌تری دارد [۹۱]. در شکل ۴-۵، معماری این شبکه قابل مشاهده است.



شکل ۴-۵. شبکه BCDU-Net [۸۸].

به طور کلی در این شبکه، چند نوآوری مهم به کار رفته است که در ادامه آمده است. این مزایا باعث استفاده این شبکه در این تحقیق شده است.

➤ در این شبکه از مکانیزم لایه‌های کانولوشنی متراکم در آخرین لایه کانولوشنی مسیر کدگذاری استفاده شده است. این تغییر باعث می‌شود که از یادگیری ویژگی‌های اضافی در مسیر کانولوشن‌های پیوسته، جلوگیری شود.

➤ به طور کلی نرمال‌سازی دسته‌ای به بهبود عملکرد، پایداری و افزایش سرعت شبکه‌های عصبی کمک می‌کند. در این شبکه و در مسیر کدگذاری بعد از هر لایه up-sample، از

^۱ Bi-Directional ConvLSTM

^۲ Skip connection

نرمال‌سازی دسته‌ای استفاده شده است، به‌طوری که خروجی حاصل از تابع نرمال‌سازی دسته‌ای، به یک لایه BConvLSTM داده می‌شود. این فرآیند به این معنا است که در قسمت اتصال پرش، به وسیله‌ی لایه BConvLSTM، از پردازشی پیچیده‌تر نسبت به حالت عادی U-Net جهت الحاق نقشه ویژگی‌ها استفاده شده است. لازم به ذکر است که این لایه، از دو ConvLSTM برای پردازش کلیه داده‌ها در قسمت ورودی به مسیرهای رو به جلو و رو به عقب استفاده می‌کند.

با توجه به توضیحاتی که در این قسمت و فصل‌های پیشین داده شد، برای بخش‌بندی تصویر، از این روش استفاده خواهد شد.

۴-۴ مدل یادگیری عمیق پیشنهادی برای مرحله طبقه‌بندی

همان‌طور که در بخش ۳-۶-۱ بررسی شد، برای طبقه‌بندی و تشخیص سرطان، از شبکه عصبی کانولوشنی سه بعدی استفاده خواهد شد. این شبکه عصبی عمیق، بهترین شبکه در راستای طبقه‌بندی و استخراج ویژگی است. ما در این تحقیق بر روی این شبکه تغییراتی را به وجود آورده‌ایم تا خروجی بهتری در اختیار ما قرار دهد که در ادامه به طور کامل شرح می‌دهیم.

شبکه عصبی کانولوشنی سه بعدی پیشنهادی دارای سه لایه کانولوشن و به دنبال آن سه لایه تجمیع و یک لایه کاملاً متصل است. سیستم پیشنهادی، در کل شامل نه لایه است که شامل لایه‌های ورودی و خروجی است. در ادامه و به طور کامل، تمام لایه‌های این شبکه مورد بررسی قرار خواهند گرفت.

۴-۴-۱ لایه اول (ورودی تصویر)

این مدل به عنوان ورودی، تصویر سی‌تی‌اسکن با اندازه $20 \times 96 \times 96$ را می‌پذیرد. در بخش‌های بعدی و برای درک بیشتر از مراحل بعد، بررسی کلی بر روی خروجی لایه کانولوشن و لایه تجمیع می‌شود. سپس، تک‌تک لایه‌های شبکه استفاده شده مورد بررسی قرار می‌گیرد.

۴-۴-۲ لایه دوم (لایه کانولوشنی اول)

لایه کانولوشنی، اندازه حجم $W \times H \times D \times C$ را می‌پذیرد که در آن، W اندازه عرض تصویر، H ارتفاع تصویر، D عمق تصویر است و C کانال تصویر را نشان می‌دهد. اگر کانال رنگی باشد، عدد سه و اگر خاکستری باشد، عدد یک را می‌پذیرد. در این تحقیق از تصویر مقیاس خاکستری استفاده شده است و مقدار ۱ است.

به طور کلی در لایه کانولوشن، ورودی به صورت سه‌تایی $W1 \times H1 \times D1$ می‌باشد. در این رابطه $W1$ عرض تصویر ورودی، $H1$ ارتفاع تصویر ورودی و $D1$ عمق تصویر ورودی در لایه کانولوشن می‌باشند. در ضمن، در لایه کانولوشن چهار ابر پارامتر وجود دارد که در ادامه معرفی شده‌اند.

F : تعداد هسته.

K : اندازه هسته.

S : اندازه stride که در هر گام مشخص می‌شود.

P : لایه‌گذاری صفر به منظور اضافه کردن سطر و ستون صفر به صورت متقارن به دور ماتریس

ورودی.

با توجه به این پارامترها، رابطه کلی مورد استفاده برای یافتن اندازه خروجی تنسور (تصویر) f در لایه

کانولوشن، به صورت رابطه (۴-۱) می‌باشد.

$$\text{Output Tensor Size (O)} = \frac{W - K + 2P}{S} + 1 \quad (4 - 1)$$

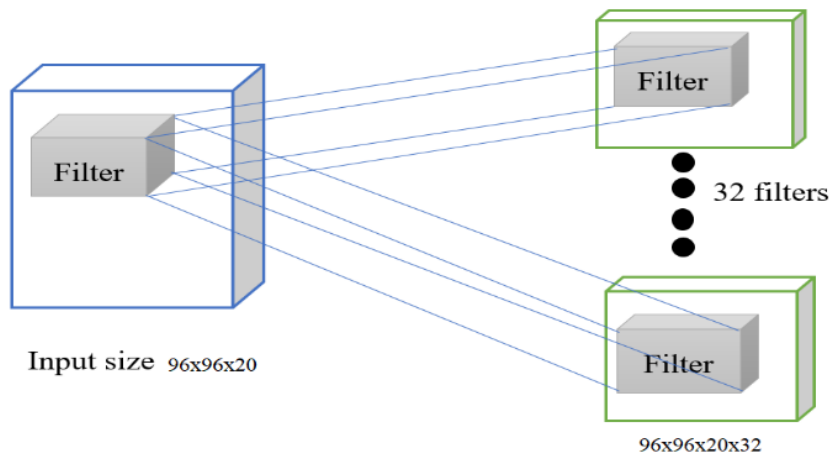
در مرحله اول، اندازه داده ورودی به لایه کانولوشن $32 \times 64 \times 64$ است. در این کانولوشن، یک لایه

با اندازه هسته $(3, 3, 3)$ و اندازه stride $(1, 1, 1)$ استفاده شده است. این لایه از تابع فعال سازی

$(ReLU)$ استفاده می‌کند و این تابع فعال سازی به صورت رابطه (۴-۲) تعریف می‌شود.

$$F(x) = \max(0, x) \quad (2)$$

در این رابطه، $F(x)$ تابع فعال سازی واحد خطی اصلاح شده است و x داده ورودی است. در این تابع، اگر مقدار x بزرگتر از صفر باشد، خروجی x است. اگر مقدار x کوچکتر یا مساوی صفر باشد، خروجی صفر است. این تابع باعث می‌شود مدل ما در محاسبات کارآمدتر باشد و به یادگیری شبکه سرعت می‌بخشد.



شکل ۴-۶. فیلترهای لایه کانولوشن سه بعدی

۴-۴-۳ لایه‌های سوم تا هفتم

به طور کلی لایه تجمیع اولین لایه از این شبکه در لایه سوم است. ورودی این لایه، به صورت سه‌تایی $W3 \times H3 \times D3$ می‌باشد. در این رابطه $W3$ عرض تصویر ورودی، $H3$ ارتفاع تصویر ورودی و $D3$ عمق تصویر ورودی در لایه کانولوشن می‌باشند. بر خلاف لایه کانولوشنی، در این لایه دو ابر پارامتر وجود دارد، که این دو پارامتر، اندازه تجمیع K و اندازه گام S می‌باشند. با توجه به این پارامترها، رابطه کلی مورد استفاده برای یافتن اندازه خروجی تانسور (تصویر) f در لایه تجمیع، به صورت (۴-۳) می‌باشد.

$$\text{Output Tensor Size (O)} = \frac{W - K}{S} + 1 \quad (4 - 3)$$

بعد از لایه اول کانولوشنی، یک لایه تجمیع در شبکه می‌آید. در این لایه اندازه هسته (۳،۳،۳) و اندازه گام (۲،۲،۲) است. از لایه تجمیع برای کاهش محاسبات به وسیله‌ی کاهش ابعاد تصاویر استفاده می‌شود.

لایه‌های چهارم تا هفتم این شبکه از لایه‌های کانولوشنی و تجمیع تشکیل شده است. این لایه‌ها دارای یک نظم هستند و به صورت یکی در میان تکرار می‌شوند. لایه‌های چهارم و ششم، از نوع لایه کانولوشنی و لایه‌های پنجم و هفتم، از نوع لایه تجمیع می‌باشند.

در لایه‌های چهارم و ششم، خروجی اولین لایه تجمیع مستقیماً و به طور مجزا به هر کدام از لایه‌ها وارد می‌شود. در این دو بخش، اندازه هسته (۳،۳،۳) و اندازه stride (۱،۱،۱) است و در انتهای آن‌ها از تابع ReLu استفاده شده است.

در لایه‌های پنجم و هفتم، خروجی لایه دوم کانولوشنی مستقیماً و به طور مجزا به هر کدام از لایه‌ها وارد می‌شود. در این لایه‌ها اندازه هسته (۳،۳،۳) و اندازه گام (۲،۲،۲) است. از لایه تجمیع برای کاهش محاسبات به وسیله‌ی کاهش ابعاد تصاویر استفاده می‌شود.

۴-۴-۴ لایه هشتم (لایه تمام متصل)

پس از آخرین لایه تجمیع، خروجی از بردار سه بعدی به یک بعدی تبدیل می‌شود. همان‌طور که در فصل قبل اشاره شد، لایه تماماً متصل برای طبقه بندی وضعیت سرطان بیمار استفاده می‌شود. اما برای بهبود طبقه بندی، خروجی این لایه به شبکه XGBOOST داده می‌شود. این لایه، طبقه بندی را انجام می‌دهد. در جدول ۴-۱، تمام جزئیات لایه‌ها تا پایان لایه هشتم آورده شده است.

جدول ۴-۱. جزئیات هشت لایه شبکه کانولوشنی

خروجی	فعال ساز	نام قلم	لایه‌ها
۹۶x۹۶x۲۰			ورودی
۹۶x۹۶x۲۰x۳۲	ReLu	گام: ۱x۱x۱ اندازه کرنل: ۳x۳x۳	کانولوشن اول
۴۸x۴۸x۱۰x۳۲		گام: ۲x۲x۲ اندازه کرنل: ۳x۳x۳	لایه تجمیع اول
۴۸x۴۸x۱۰x۶۴	ReLu	گام: ۱x۱x۱ اندازه کرنل: ۳x۳x۳	کانولوشن دوم
۲۴x۲۴x۵x۶۴		گام: ۲x۲x۲ اندازه کرنل: ۳x۳x۳	لایه تجمیع دوم

کانولوشن سوم	گام: $1x1x1$ اندازه کرنل: $3x3x3$	ReLU	$24x24x5x128$
لایه تجمیع سوم	گام: $2x2x2$ اندازه کرنل: $3x3x3$		$12x12x3x128$
لایه تماما متصل	ورودی نورون ها : $12x12x3x128 = 55,296$		نورون های خروجی : (1024)

۴-۴-۵ لایه نهم (شبکه XGBOOST)

همان طور که در فصل قبل و این فصل مورد بررسی قرار گرفت، از این شبکه برای بهبود طبقه بندی استفاده می شود. ویژگی به وسیله ی لایه تماما متصل استخراج می شود و از اعمال شبکه XGBOOST بر روی این ویژگی ها، بهبود طبقه بندی حاصل می شود. تعداد کلاس خروجی، دو کلاس می باشد که به عنوان طبقه بندی باینری نیز شناخته می شود.

همان طور که به طور کامل شرح داده شد، طراحی شبکه عصبی کانولوشنی سه بعدی ما شامل نه لایه است که دو لایه آن ورودی و خروجی و یک لایه آن هم، لایه شبکه تقویت گرادیان است. شش لایه بعدی هم شامل سه لایه کانولوشن و سه لایه تجمیع است. در تحقیق های مختلف، شبکه عصبی کانولوشنی سه بعدی دارای تعداد لایه های متفاوتی است و این تفاوت به دلیل طرز طراحی و نحوه به کارگیری این شبکه است.

ما در این تحقیق تعداد لایه های این شبکه را چندین بار تغییر داده ایم و با این کار، دقت شبکه پایین تر آمد. در تست اول، به جای استفاده از شش لایه، از چهار لایه که شامل دو لایه کانولوشن و دو لایه تجمیع بود، استفاده کردیم و دقت کار پایین تر آمد. در تست دوم، تعداد لایه ها را بالاتر بردیم و از هشت لایه که شامل چهار لایه کانولوشن و چهار لایه تجمیع بود، استفاده کردیم، که این شبکه هم دقت پایین تری نسبت به حالت شش لایه ای داشت. با توجه به این تست های انجام شده و مقالات خوانده شده پیرامون این حوزه، شبکه کانولوشنی نه لایه ای را استفاده کرده ایم.

۴-۵ جمع‌بندی

در این فصل، روش پیشنهادی به طور کامل شرح داده شد و گفتیم که روش پیشنهادی از سه بخش پیش‌پردازش تصاویر، بخش‌بندی تصاویر و استخراج ویژگی و طبقه‌بندی تشکیل شده است. با توجه به این داده‌های یک دست هستند، پیش‌پردازش خاصی بر روی داده‌ها انجام نشده است و پیش‌پردازش‌های معمول بر روی داده‌ها لحاظ شده است. برای بخش‌بندی تصاویر از شبکه‌ی عصبی عمیق BCDU-Net استفاده شده است. همان طور که در بخش ۴-۳-۱ گفته شد، این شبکه مزایای زیادی نسبت به دیگر شبکه‌ها در این حوزه دارد. در فصل بعد، خروجی حاصل از این شبکه با شبکه اصلی U-Net، مقایسه و بررسی خواهد شد. برای استخراج ویژگی از شبکه عصبی کانولوشنی سه بعدی استفاده شده است. این شبکه در این حوزه بهترین نتیجه و خروجی را دارد. به طور کلی، با تغییرات در تعداد و طرز چینش این لایه‌ها، می‌توان خروجی این شبکه را بهبود بخشید. در این شبکه و در مرحله آخر و برای طبقه‌بندی، از شبکه‌ی XGBOOST استفاده شده است. به طور کل، از لایه تماماً متصل برای طبقه‌بندی استفاده می‌شود، اما ما در این تحقیق و به عنوان یک نوآوری، این شبکه را به انتهای لایه تماماً متصل، پیوند دادیم تا درصد دقت را بالا ببریم.

فصل ۵ پیاده‌سازی و نتایج

۱-۵ مقدمه

همان‌طور که در فصل‌های قبلی گفته شد، این تحقیق، از سه بخش پیش‌پردازش، بخش‌بندی و طبقه‌بندی تشکیل شده است. در قسمت بخش‌بندی تصاویر، از شبکه‌ی عمیق استفاده کرده‌ایم تا بهترین خروجی را برای ما داشته باشد. در این بخش از بهترین نسخه شبکه UNET یعنی BCDU-Net استفاده کرده‌ایم. این شبکه، پیش‌پردازش‌هایی را بر روی تصاویر انجام می‌دهد. در آخرین مرحله و با توجه به برچسب‌دار بودن تصاویر، از یادگیری عمیق نظارت شده استفاده کرده‌ایم. یکی از بهترین و بزرگترین الگوریتم‌های این حوزه، شبکه عصبی کانولوشنی سه بعدی است و ما تغییراتی را بر روی این شبکه اعمال کرده‌ایم که خروجی بهتری را در اختیار ما بگذارد. علاوه بر این تغییرات، برای طبقه‌بندی نهایی، از شبکه XGBOOST استفاده کرده‌ایم که طبقه‌بندی را بهبود می‌بخشد. در این فصل ابتدا به معرفی مجموعه داده‌های استفاده شده در این پایان‌نامه خواهیم پرداخت و به طور کامل آن را بررسی خواهیم کرد. در ادامه توضیحات کاملی پیرامون معیارهای ارزیابی در این حوزه داده خواهد شد. در انتهای این فصل، نتایج آزمایش به طور کامل شرح داده خواهد شد و در بخش بعدی، این نتایج را با نتایج روش‌های دیگر مقایسه خواهیم کرد.

۲-۵ مجموعه داده‌گان استفاده شده در این پایان‌نامه

در این تحقیق باید از مجموعه داده‌های قفسه سینه استفاده کرد. مجموعه‌های مختلفی از قفسه سینه موجود است که یکی از آن‌ها، مجموعه داده LUNA-۱۶ است. این مجموعه داده یک بخش کوچکتري از مجموعه داده LIDC-IDRI است. در ادامه توضیح مختصری از پایگاه داده LIDC-IDRI و LUNA-۱۶ داده خواهد شد.

۱-۲-۵ مجموعه داده‌گان LIDC-IDRI

این مجموعه، شامل ۱۰۱۸ تصویر سی‌تی‌اسکن قفسه سینه از همین تعداد بیمار مختلف می‌باشد. این

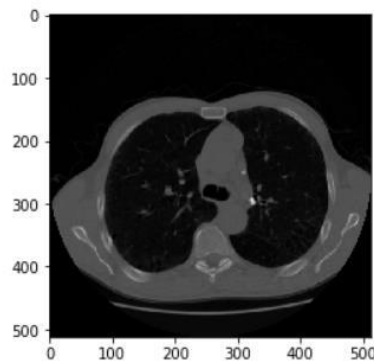
مجموعه به عنوان منبع تحقیق تصویربرداری پزشکی به کار می‌رود. به‌طور کلی، این مجموعه دادگان را با توجه به ضخامت برش به زیر دسته‌های مختلف تقسیم می‌کنند. یکی از زیر دسته‌های این مجموعه دادگان، LUNA-۱۶ است که دارای ۸۸۸ تصویر است. همان‌طور که گفته شد، تقسیم‌بندی‌ها بر اساس ضخامت برش صورت می‌گیرد. از مجموعه دادگان اصلی، تصاویری با ضخامت بیشتر از ۲/۵ میلی متر، حذف شده‌اند و مجموعه دادگان LUNA-۱۶ به وجود آمده است که در ادامه به بررسی آن خواهیم پرداخت.

۵-۲-۲ مجموعه دادگان LUNA-۱۶

با توجه به توضیحات بخش‌های قبلی، پایگاه داده مورد استفاده در این تحقیق LUNA-۱۶ می‌باشد که این پایگاه داده یک زیر مجموعه از مجموعه داده LIDC-IDRI است. همان‌طور که گفته شد، ملاک و معیار تقسیم‌بندی‌ها و ایجاد زیردسته در مجموعه داده LIDC-IDRI، ضخامت برش در نظر گرفته شده است. سی‌تی‌اسکن‌هایی با ضخامت برش بیش از ۲/۵ میلی‌متر، حذف شده‌اند. از کل داده‌های موجود، ۸۸۸ تصویر باقی ماندند که پایگاه داده LUNA-۱۶ را تشکیل داده‌اند. پردازش‌های مختلفی توسط چهار رادیولوژیست بر روی این مجموعه داده انجام شده است که در ادامه به بررسی آن‌ها پرداخته خواهد شد. این چهار رادیولوژیست، تمام تصاویر سی‌تی‌اسکن را مورد بررسی قرار دادند و سپس تمام نتایج بررسی را جمع‌آوری کردند. در ادامه و در بررسی دوم، رادیولوژیست‌ها گزارش‌های خود را با گزارش‌های دیگر رادیولوژیست‌ها مقایسه کردند. در ادامه، هر رادیولوژیست ضایعات تشخیص داده شده به عنوان ندول را مشخص کرد و در دو دسته بزرگتر از سه میلی‌متر و کوچکتر از سه میلی‌متر قرار داد. استاندارد مرجع چالش ما، شامل همه‌ی ندول‌های بزرگتر مساوی سه میلی‌متر است که حداقل از هر ۴ رادیولوژیست ۳ نفر آن را پذیرفته‌اند [۸۶].

هر سی‌تی‌اسکن کلینیکی دارای یک فایل XML است که در این فایل نتایج تفسیر تصویر دو مرحله‌ای توسط چهار رادیولوژیست قفسه سینه وجود دارد. قالب تمام تصاویر این مجموعه، DICOM

است. این تصاویر دارای سه کانال بوده و در اندازه 512×512 پیکسل هستند. در شکل ۵-۱، یک نمونه از تصاویر این مجموعه داده قابل مشاهده است.



شکل ۵-۱. تصاویر سی تی مربوط به یک بیمار از مجموعه داده LUNA ۱۶

تمام تصاویر مجموعه داده LUNA-۱۶، دارای برچسب و نیز دارای یک برچسب گره می‌باشند. خود این برچسب‌ها شامل لیست مختصات و قطر مرکز گره می‌باشند. یک نمونه برچسب گذاری در جدول ۱-۵ قابل مشاهده است [۸۴].

جدول ۵-۱. نمونه برچسب‌گذاری (برچسب ۱ دارای سرطان و ۰ بدون سرطان) [۸۴]

شماره تصویر	کد منحصر به فرد	تشخیص در سرطان
۱	۰۰۱۵ceb۸۵۱d۷۲۵۱b۸f۳۹۹e۳۹۷۷۹d۱e۷d	۱
۲	۰۰۳۰a۱۶۰d۵۸۷۲۳ff۳۶d۷۳ff۱b۱۷۰ec۲۱	۰
۳	۰۰۳f۴۱c۷۸e۶acfa۹۲۴۳۰a۰۵۷ac۰b۳۰۶e	۰
۴	۰۰۶b۹۶۳۱۰a۳۷b۳۶cccb۲ab۴۸d۱۰b۴۹a۳	۱
۵	۰۰۸۴۶۴bb۸۵۲۱d۰۹a۴۲۹۸۵dd۸add۳d۰d۲	۱

۳-۵ معیارهای ارزیابی

یکی از مهم‌ترین مراحل پس از طراحی و ساخت یک مدل یا یک الگوریتم، ارزیابی کارایی^۱، دقت و

^۱ performance

صحت آن است. در ادامه روش‌هایی برای ارزیابی مدل پیشنهادی ارائه می‌شود. معیارهای متنوعی برای ارزیابی کارایی الگوریتم‌ها وجود دارد و در انتخاب معیارهای ارزیابی کارایی می‌بایست دلایل قانع کننده وجود داشته باشد، چرا که چگونگی اندازه‌گیری و مقایسه کارایی الگوریتم‌ها، کاملاً به معیارهایی که انتخاب شده وابسته است. در ابتدا به معرفی چندین اصطلاح پرداخته خواهد شد، تا درک بهتری از ارزیابی صورت گیرد.

۵-۳-۱ حساسیت و تشخیص پذیری

حساسیت و تشخیص‌پذیری دو شاخص مهم برای ارزیابی آماری عملکرد نتایج آزمون‌های طبقه‌بندی دودویی یا دوحالته هستند که در آمار به عنوان توابع طبقه‌بندی شناخته می‌شوند. به طور کل نتایج را بعد از تحلیل می‌توان به دو گروه داده مثبت و منفی تقسیم کرد که روش‌های آزمون و ارزیابی، این نتایج را به این دو دسته تقسیم می‌کند و سپس کیفیت الگوریتم با استفاده از شاخص‌های حساسیت و تشخیص‌پذیری، قابل اندازه‌گیری و توصیف می‌شوند. بعد از تحلیل صورت گرفته بر روی داده‌ها، دسته‌بندی‌ها داده به صورت زیر انجام می‌پذیرد.

۱) مثبت صحیح (True Positive) که یک معیار مثبت است. در این دسته، الگوریتم نمونه را در دسته مثبت طبقه‌بندی می‌کند و نمونه هم مثبت است.

۲) مثبت کاذب (False Positive) که یک معیار منفی است. در این دسته، الگوریتم نمونه را در دسته مثبت طبقه‌بندی می‌کند، اما نمونه منفی است.

۳) منفی صحیح (True Negative) که یک معیار مثبت است. در این دسته، الگوریتم نمونه را در دسته منفی طبقه‌بندی می‌کند و نمونه هم منفی است.

۴) منفی کاذب (False Negative) که یک معیار منفی است. در این دسته، الگوریتم نمونه را در دسته منفی طبقه‌بندی می‌کند، اما نمونه مثبت است.

به بیان دیگر، زمانی که الگوریتم کلاس نمونه را اشتباه پیش‌بینی می‌کند، نتیجه FN یا FP خواهد بود و در مقابل، زمانی که الگوریتم کلاس نمونه را به درستی پیش‌بینی می‌کند، نتیجه TN یا TP خواهد بود.

۵-۳-۲ ماتریس اختلاط یا درهم ریختگی

در حوزه الگوریتم‌های هوش مصنوعی، ماتریس در هم ریختگی به ماتریسی گفته می‌شود که نتیجه ارزیابی عملکرد الگوریتم‌ها در آن نمایش داده می‌شود. معمولاً چنین نمایشی برای الگوریتم‌های یادگیری با ناظر استفاده می‌شود. اگرچه در یادگیری بدون ناظر نیز کاربرد دارد و معمولاً به کاربرد این ماتریس در الگوریتم‌های بدون ناظر ماتریس تطابق گفته می‌شود. هر ستون از این ماتریس، نمونه‌ای از مقدار پیش‌بینی شده را نشان می‌دهد و هر سطر نمونه‌ای واقعی را در بر دارد [۹۲].

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

شکل ۵-۲. شکل کلی ماتریس درهم ریختگی [۹۲].

به طور کلی بعد از هر تحلیلی، نتایج به‌دست‌آمده را باید در چهار گروه مثبت صحیح، مثبت کاذب، منفی صحیح و منفی کاذب، دسته‌بندی کرد تا بتوان کیفیت تحلیل را مورد ارزیابی قرار داد. همچنین کارایی الگوریتم را برای استفاده‌های مختلف، تفسیر و توصیف نمود. در ارزیابی نتایج، معمولاً بعد از تشکیل ماتریس در هم ریختگی، معیارهایی متصور هستند که در ادامه به بررسی آن‌ها پرداخته شده

است. این معیارها، بیشترین استفاده را نسبت به معیارهای ارزیابی دیگر دارند.

➤ دقت (Accuracy)

به طور کلی، دقت به این معناست که مدل تا چه اندازه خروجی را درست پیش‌بینی می‌کند. با نگاه کردن به دقت، بلافاصله می‌توان دریافت که آیا مدل، درست آموزش دیده است یا خیر و کارایی آن به طور کلی چگونه است. این معیار اطلاعات جزئی در مورد کارایی مدل ارائه نمی‌دهد. رابطه (۵-۱) دقت را محاسبه می‌کند [۹۳].

$$\text{Accuracy} = (TP+TN) / (TP+FN+FP+TN) \quad (۵ - ۱)$$

➤ حساسیت (Sensitivity) یا یادآوری (recall)

پارامترهای دیگری نیز علاوه بر معیار دقت وجود دارد که می‌توان به سادگی از ماتریس درهم‌ریختگی استخراج کرد. یکی از متداول‌ترین آن‌ها، معیار حساسیت است که آن را نرخ پاسخ‌های مثبت درست^۱ نیز می‌گویند. حساسیت به معنی نسبتی از موارد مثبت است که آزمایش، آن‌ها را به درستی به عنوان نمونه مثبت تشخیص داده است. این معیار مشخص می‌کند دسته‌بند، به چه اندازه در تشخیص تمام افراد مبتلا به بیماری موفق بوده‌است. همانگونه که از رابطه (۵-۲) مشخص است، تعداد افراد سالمی که توسط دسته‌بند به اشتباه به عنوان فرد بیمار تشخیص داده شده‌اند، هیچ تاثیری در محاسبه این پارامتر ندارد. در واقع زمانی که پژوهشگر از این پارامتر به عنوان پارامتر ارزیابی برای دسته‌بند خود استفاده می‌کند، هدفش دستیابی به نهایت صحت در تشخیص نمونه‌های کلاس مثبت است [۹۳]. این پارامتر به صورت رابطه (۵-۲) محاسبه می‌شود.

$$\text{Sensitivity (TPR)} = TP / (TP+FN) \quad (۵ - ۲)$$

^۱ True Positive Rate

➤ تشخیص‌پذیری (Specificity)

در نقطه مقابل پارامتر حساسیت، ممکن است در مواقعی صحت تشخیص کلاس منفی حائز اهمیت باشد. به همین علت از متداول‌ترین پارامترها که معمولاً در کنار حساسیت بررسی می‌شود، پارامتر خاصیت است که به آن نرخ پاسخ‌های منفی درست^۱ نیز می‌گویند. خاصیت به معنی نسبتی از موارد منفی است که آزمایش، آن‌ها را به درستی به عنوان نمونه منفی تشخیص داده است [۹۳]. این پارامتر به صورت رابطه (۵-۳) محاسبه می‌شود.

$$\text{Specificity (TNR)} = \text{TN} / (\text{TN} + \text{FP}) \quad \text{➤ (۵ - ۳)}$$

➤ صحت (Precision)

به طور کلی زمانی که ارزش مثبت کاذب بالا باشد، معیار صحت، معیار مناسبی خواهد بود. برای مثال اگر در این تحقیق مدل پیشنهادی صحت پایینی داشته باشد، نتیجه امر این است که این مدل، بیماری بسیاری از افراد را به اشتباه سرطان تشخیص می‌دهد و در نتیجه استرس زیاد، آزمایش‌های فراوان و هزینه‌های گزافی را برای بیمار به دنبال خواهد داشت [۹۴].

در واقع صحت، نسبت موارد صحیح طبقه‌بندی شده توسط الگوریتم از یک کلاس مشخص، به کل تعداد مواردی که الگوریتم در آن کلاس طبقه‌بندی کرده، است که در رابطه (۵-۴) قابل مشاهده است.

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

➤ معیار (F-Measure)

این پارامتر برای ارزیابی عملکرد دسته‌بندها بسیار مورد استفاده قرار می‌گیرد و از ترکیب دو پارامتر حساسیت و صحت حاصل می‌شود [۹۴]. معیار F-Measure به صورت رابطه (۵-۵) تعریف می‌شود.

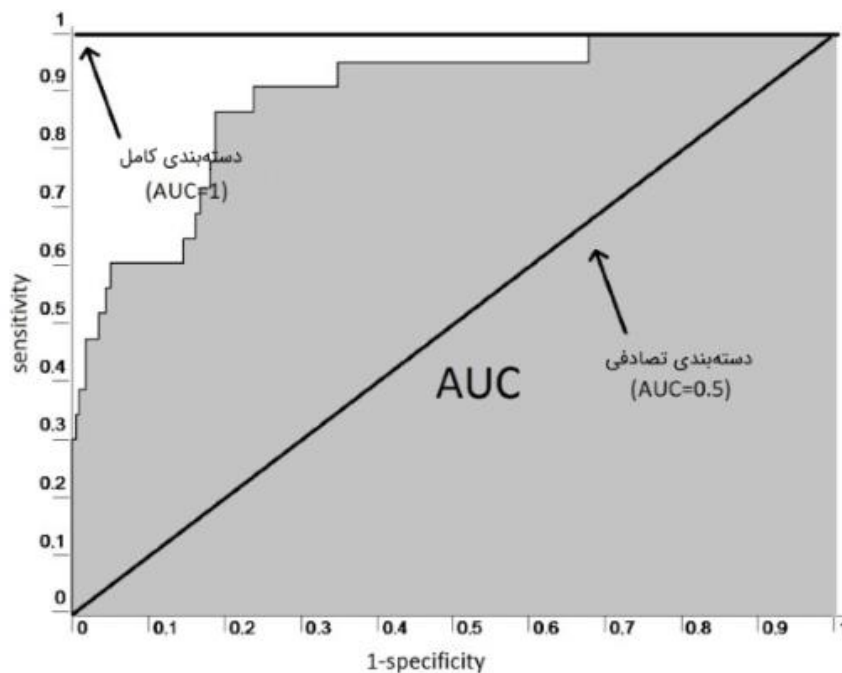
$$\text{F-measure} = 2 * (\text{Recall} * \text{Precision}) / (\text{Recall} + \text{Precision}) \quad \text{➤ (۵ - ۵)}$$

^۱ True Negative Rate

➤ سطح زیر نمودار ROC برای مدل پیشنهادی

دو پارامتر حساسیت و خاصیت نیز مشابه معیار صحت، معمولاً به صورت درصد بیان می‌شوند. واضح است که پیش‌بینی عالی، پیش‌بینی است که مقادیر حساسیت و خاصیت مربوط به آن، هر دو صد درصد باشند؛ اما احتمال وقوع این اتفاق در واقعیت بسیار کم است و همیشه یک حداقل خطایی وجود دارد. پارامترهای حساسیت و خاصیت، بنابر ماهیتی که دارند همواره در رقابت با یکدیگر هستند. یعنی افزایش یکی با کاهش دیگری همراه است. همین وضعیت منجر به تولید ابزاری دیگر برای ارزیابی کیفیت دسته‌بندی شده است. منحنی مشخصه عملکرد سیستم^۱ عبارت است از منحنی که ارتباط بین دو پارامتر حساسیت و خاصیت را بیان می‌کند.

همان‌طور که در شکل ۳-۵ مشاهده می‌کنید، محور عمودی این نمودار نشان‌دهنده نرخ مثبت صحیح و محور افقی نشان‌دهنده مقدار نرخ مثبت غلط است. نتایج مختلف دسته‌بندی نشانگر نقاط مختلف بر روی این نمودار هستند و در نهایت یک منحنی را تشکیل می‌دهند. با توجه به شکل ۳-۵، در بهترین حالت و با فرض طبقه‌بندی صد درصد صحیح در هر دو دسته، نقطه مربوطه عبارت است از نقطه گوشه بالای سمت چپ، یعنی نقطه (۰,۱) و نیز با فرض دسته‌بندی به صورت تصادفی، نقطه متناظر در منحنی، یکی از نقاط موجود روی خط واصل نقطه (۰,۰) و نقطه (۱,۱) خواهد بود. در واقعیت، منحنی حاصل از یک دسته‌بندی، منحنی بین این دو حالت است [۹۵].



شکل ۳-۵. منحنی مشخصه عملکرد سیستم

مساحت زیر این نمودار، به عنوان یک معیار برای ارزیابی عملکرد دسته‌بند مورد استفاده قرار می‌گیرد. با توجه به توضیحاتی که پیش‌تر ارائه شد، بدیهی است که در حالت ایده‌آل، مساحت زیر منحنی برابر با بیشترین مقدار خود، یعنی یک است. بنابراین، هر چه مساحت زیر نمودار به عدد یک نزدیکتر باشد، به معنای بهتر بودن عملکرد دسته‌بند است. علاوه بر دو پارامتر حساسیت و خاصیت، پارامترهای دیگری هم از ماتریس درهم‌ریختگی استخراج می‌شوند که هر یک بیان‌کننده مفهومی هستند و کاربردهای متفاوتی دارند. از جمله آن‌ها، دو پارامتر مقدار پیش‌بینی شده مثبت^۱ و مقدار پیش‌بینی شده منفی^۲ هستند، که برای بیان نسبت پاسخ‌های درست در هر دسته استفاده می‌شوند. ارزش اخباری مثبت، بیان‌کننده این است که چند درصد از الگوهایی که دسته‌بند آن‌ها را مثبت تشخیص داده، در واقعیت هم مثبت هستند. به همین ترتیب، ارزش اخباری منفی نشان می‌دهد که چند درصد از نمونه‌هایی که عضو دسته منفی تشخیص داده شده‌اند، در واقعیت هم عضو همین دسته هستند. این دو پارامتر نیز به سادگی از روی ماتریس درهم‌ریختگی با رابطه‌های (۵-۶) و (۵-۷) تعریف می‌شوند.

^۱ Positive Prediction Value

^۲ Negative Predictive Values

$$PPV = TP / (TP + FP) \quad (6 - 5)$$

$$NPV = TN / (TN + FN) \quad (7 - 5)$$

در شکل ۴-۵ روابط معیارهای ارزیابی به طور خلاصه آورده شده است.

		Predicted Class		
		Positive	Negative	
Actual Class	Positive	True Positive (TP)	False Negative (FN) Type II Error	Sensitivity $\frac{TP}{(TP + FN)}$
	Negative	False Positive (FP) Type I Error	True Negative (TN)	Specificity $\frac{TN}{(TN + FP)}$
		Precision $\frac{TP}{(TP + FP)}$	Negative Predictive Value $\frac{TN}{(TN + FN)}$	Accuracy $\frac{TP + TN}{(TP + TN + FP + FN)}$

شکل ۴-۵. معیارهای ارزیابی الگوریتم‌های هوش مصنوعی

۴-۵ نتایج آزمایش‌ها

در این قسمت، مدل پیشنهاد شده را مورد بررسی قرار خواهیم داد و نتایج حاصل را بیان خواهیم کرد. در این بخش، ماتریس درهم ریختگی و نمودار ROC برای بررسی روش پیشنهادی به طور کامل تحلیل می‌شود و با روش‌های دیگر مقایسه می‌شود.

در این پایان‌نامه، برای ارزیابی نتایج از پایگاه داده LUNA-۱۶ استفاده شده است. ۹۰ درصد از کل داده‌ها برای مرحله آموزش و ۱۰ درصد از داده‌ها برای قسمت آزمون در نظر گرفته شده است. تمام پیاده‌سازی‌های این پایان‌نامه، با استفاده از چارچوب کراس^۱ در زبان برنامه‌نویسی پایتون می‌باشد و

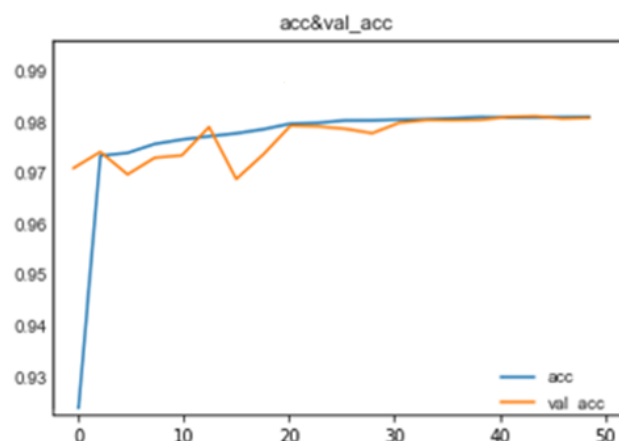
^۱ Keras

از نسخه ۳/۶ تانسورفلو^۱ استفاده شده است. با توجه به این که این پایان‌نامه در حوزه یادگیری عمیق می‌باشد. برای اجرای کدها به GPU قوی نیاز است. در شرایط محیط آزمایش گرافیک: NVIDIA GTX ۱۰۵۰، پردازنده: Intel(R) Core(TM) i۷-۹۷۵۰H و حافظه رم: ۱۶ گیگابایت است.

۵-۴-۱ دقت مدل پیشنهادی در مرحله آموزش و آزمایش

در مرحله آموزش، ورودی شبکه پیشنهادی شامل تصاویر مورد نظر و برچسب‌های آن‌ها در فایل‌های جداگانه است. در شبکه‌های عصبی کانولوشن، هدف اصلی کاهش مقدار تابع $loss$ از طریق اصلاح وزن‌ها است. این اصلاح وزن از طریق تابع بهینه‌ساز Adam انجام می‌شود که یک الگوریتم بهینه‌سازی بر اساس شیب است. هدفی که این بهینه‌سازی دنبال می‌کند، تعیین متغیرهای طراحی برای کمینه یا بیشینه کردن تابع هدف است. تابع انرژی شبکه توسط یک تابع $softmax$ در سطح پیکسل که در ترکیب با تابع هزینه $cross-entropy$ به نقشه خروجی اعمال می‌شود، محاسبه می‌شود. در مرحله آزمایش، ما یک تصویر به عنوان ورودی به شبکه داده و در خروجی آن، انتظار طبقه‌بندی صحیح را داریم. پس از مرحله آموزش، مرحله آزمایش انجام می‌شود و دقت مدل در پیش بینی برچسب‌ها ارزیابی می‌شود. در مرحله آزمایش یک تصویر جدید به شبکه تغذیه شده و همه کارهایی که در فاز آموزش انجام می‌شود، در این مرحله هم انجام می‌شود. تنها تفاوت این مرحله با مرحله آموزش این است که دیگر فاز پس انتشار خطا وجود ندارد. صرفاً ورودی در مقادیر وزن‌ها ضرب می‌شوند. نمودار دقت شبکه در فاز آموزش و آزمایش در شکل ۵-۵ نشان داده شده است.

^۱ Tensorflow



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

همان‌طور که شکل ۵-۵ 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 ۹۷ 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.

۵-۴-۲ ماتریس درهم ریختگی مدل پیشنهادی

در مرحله آموزش، دو فایل به عنوان ورودی مدل وجود دارد. فایل اول شامل تصاویر مورد نظر و فایل دوم برچسب‌های آن‌ها است. هدف اصلی شبکه‌های عصبی کانولوشنی، کاهش مقدار تابع هزینه است که این امر از طریق اصلاح وزن‌ها صورت می‌گیرد. اصلاح وزن نیز به‌وسیله‌ی تابع بهینه‌ساز Adam انجام می‌شود. ساز و کار این الگوریتم بهینه‌سازی، بر پایه‌ی شیب است و هدف از این بهینه‌سازی، کمینه یا بیشینه کردن تابع هدف است. با توجه به توضیحات داده شده در ابتدای این فصل، برای ارزیابی مدل پیشنهادی، نیاز به تشکیل ماتریس درهم ریختگی است و با تشکیل این ماتریس، می‌توان معیارهای صحت، دقت و

یادآوری را به دست آورد. در شکل ۵-۶، ماتریس درهم ریختگی نشان داده شده است که هر ستون از ماتریس، نمونه‌ای از مقدار پیش‌بینی شده و هر سطر، نمونه‌ای واقعی را دربر دارد. همان‌طور که ذکر شد، برای به دست آوردن مقادیر دقت، صحت و یادآوری از این ماتریس استفاده می‌شود و طبق این ماتریس، این سه معیار به ترتیب برابر با $97/94\%$ ، $99/93\%$ ، $97/59\%$ می‌باشند و در ضمن مقدار F-Measure، برابر $95/3\%$ است.

Confusion matrix

True lable	۰	۱
	۰	۱
۰	۲۸۴۲۹۵۴	۷۰۱۲۰
۱	۱۹۹۰	۵۸۶۷۹۸

شکل ۵-۶. ماتریس درهم ریختگی مدل پیشنهادی

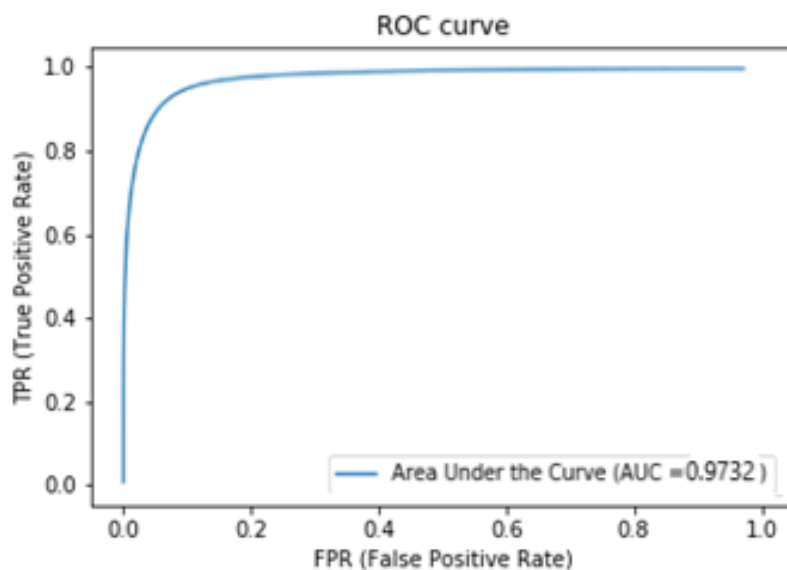
۵-۴-۳ سطح زیر نمودار ROC برای مدل پیشنهادی

در بخش‌های قبل یکی از مهمترین معیارهای ارزیابی یعنی میزان AUC^1 مورد بررسی قرار گرفت و گفته شد که نشان‌دهنده سطح زیر نمودار ROC^2 است و هر چه مقدار آن نزدیک به یک باشد، دسته‌بند مطلوب‌تر ارزیابی می‌شود [۹۵]. این معیار مستقل از آستانه تصمیم‌گیری دسته‌بند است و میزان قابل

^۱ Area Under Curve

^۲ Receiver Operating Characteristic

اعتماد بودن خروجی یک دسته‌بند مشخص به ازای مجموعه داده‌های متفاوت را نشان می‌دهد. نمودار ROC برای مدل پیشنهادی در شکل ۷-۵ نشان داده شده است. همان‌طور که در شکل ۷-۵ قابل مشاهده است، سطح زیر نمودار به یک بسیار نزدیک است و این نشان‌دهنده کارا بودن روش پیشنهادی می‌باشد.



شکل ۷-۵. نمودار ROC مدل پیشنهادی

مطابق شکل ۷-۵، 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.

۵-۵ مقایسه روش پیشنهادی با روش‌های دیگر

همان‌طور که در فصل چهارم به طور کامل بیان و بررسی شد، روش پیشنهادی از دو بخش اصلی بخش‌بندی و طبقه‌بندی تشکیل شده است. در قسمت بخش‌بندی تصویر، ما از شبکه عصبی BCD-UNET استفاده کرده‌ایم که این شبکه نسبت به شبکه‌ها دیگر دو برتری محسوس دارد. اولین برتری آن، جلوگیری از یادگیری ویژگی‌های اضافی در مسیر کانولوشن‌های پیوسته با استفاده از مکانیزم لایه‌های

کانولوشنی متراکم در آخرین لایه کانولوشنی مسیر کدگذاری است. دومین برتری آن بهبود عملکرد، پایداری و افزایش سرعت شبکه عصبی است که این امر با استفاده از نرمال‌سازی دسته‌ای صورت می‌پذیرد. قسمت دوم روش پیشنهادی، طبقه‌بندی و استخراج ویژگی است که از یک کانولوشن سه بعدی استفاده کرده‌ایم. تغییراتی نسبت به شبکه کانولوشن سه بعدی دارد که این تغییرات باعث بهبود عملکرد این شبکه می‌شود. در بخش آخر و برای طبقه‌بندی، از XGBOOST برای بهبود طبقه‌بندی استفاده کرده‌ایم و این شبکه را به لایه تماماً متصل پیوند داده‌ایم.

جدول ۲-۵. مقایسه روش پیشنهادی با سایر روش‌ها

روش‌ها و مراجع	دقت	یادآوری	صحت	امتیاز F1
Deep Convolutional Neural Networks for Lung Cancer Detection[۹۶]	۷۵.۱	۷۷	۷۴.۱	۷۵.۵۲
Lung cancer detection using Convolutional Neural Network (CNN)[۸۴]	۷۶.۷	-	-	-
Lung cancer detection and classification using ۳D-CNN[۴]	۸۶.۶	۸۵.۳	۸۸.۱	۸۶.۶۷
روش پیشنهادی	۹۷.۹۴	۹۷.۵۹	۹۹.۹۳	۹۸.۷۴

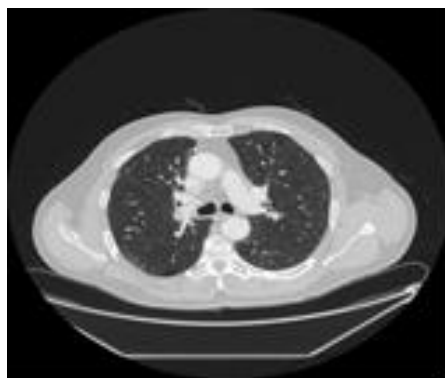
همان‌طور که در جدول ۲-۵ قابل ملاحظه است، روش پیشنهادی نسبت به روش‌های دیگر برتری دارد و خروجی بهتری را در تشخیص سرطان می‌دهد. همچنین می‌توان نتایج زیر را از جدول ۲-۵ و خروجی روش پیشنهادی گرفت.

(۱) به دلیل استفاده از BConvLSTM و نیز مکانیزم لایه‌های کانولوشنی متراکم، در مجموع مدل BCDU-Net، بهتر از نسخه اصلی شبکه U-Net عمل می‌کند. در مقاله [۴] از شبکه اصلی U-Net در قسمت بخش‌بندی استفاده شده است و باعث پایین آمدن دقت در این تحقیق شده است.

(۲) تغییرات به وجود آمده در ساختار اصلی شبکه‌ی کانولوشنی سه بعدی، دقت را به طور قابل ملاحظه‌ای افزایش داده است.

۳) همان‌طور که پیش‌تر گفته شد، در حالت عادی از لایه تماماً متصل برای طبقه‌بندی استفاده می‌شود. اما در روش پیشنهادی با استفاده از XGBOOST، طبقه‌بندی به طور دقیق‌تری انجام شده است.

برای یک نتیجه‌گیری کامل، یک مقایسه کوچک در قسمت بخش‌بندی صورت گرفته است. همان‌طور که در شکل ۵-۸ مشاهده می‌شود، تصویر یک ریه از پایگاه داده با استفاده از کتابخانه‌های موجود رسم شده است.



شکل ۵-۸. تصویر ریه از پایگاه داده

همان‌طور که در شکل ۵-۹ مشاهده می‌کنید، ریه توسط شبکه U-Net بخش‌بندی شده است و خیلی موفق عمل نکرده است. یکی از مشکلات آن این است که ندول‌های متصل به دیواره‌ی ریه را حذف کرده و نمی‌تواند حفظ کند.



شکل ۵-۹. تصویر قطعه‌بندی شده توسط U-Net

همان‌طور که گفته شد مدل BCDU-Net، بسیاری از ضعف‌ها و کاستی‌های موجود در تصویر قطعه‌بندی شده توسط U-Net را از بین برده و به درستی و کامل عمل می‌کند. در این روش، مشکل ذکر شده در بالا هم تا حد زیادی رفع شده و تصویر قطعه‌بندی شده نهایی، شباهت بسیار زیادی به تصویر ماسک متناظرش دارد. در شکل ۵-۱۰ ریه توسط شبکه BCDU-Net بخش‌بندی شده است.



شکل ۵-۱۰. تصویر قطعه‌بندی شده توسط BCDU-Net

۶-۵ جمع‌بندی

در این فصل، نتایج به دست آمده از مراحل مختلف روش پیشنهادی به طور کامل مورد بررسی قرار گرفت و خروجی مرحله بخش‌بندی و طبقه‌بندی با ذکر جزئیات بیان شد. با انجام پیاده‌سازی و آزمایش‌ها، دیدیم که در قسمت بخش‌بندی، استفاده از بهترین نسخه U-NET یعنی BCDU-Net، بهترین خروجی را برای مرحله طبقه‌بندی می‌دهد. در بخش طبقه‌بندی، با ایجاد تغییرات بر روی شبکه اصلی کانولوشنی سه بعدی، طبقه‌بندی با دقت بسیار بالایی صورت گرفته است. برای بالا بردن این دقت، از تقویت گرادیان و روش XGBoost استفاده کردیم و دقت طبقه‌بندی بالاتر رفت و بهترین خروجی و بالاترین دقت را در بر داشت. به طور کلی با تغییرات بر روی این شبکه‌ها و ترکیب آن‌ها با هم و با داشتن پایگاه داده مناسب، دقت به طور چشم‌گیری افزایش یافته است و در مقایسه با روش‌های دیگر بهتر عمل کرده است.

فصل ۶ نتیجه‌گیری و پیشنهادها برای کارهای آتی

۶-۱ 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.

۶-۲ جمع بندی و نتیجه گیری

سرطان یکی از کشنده ترین بیماری های جهان است و اشتباه پزشک در تشخیص این بیماری، یکی از بزرگترین مشکلات این حوزه است. برای کمک به پزشکان و برای بالا بردن قدرت تشخیص، می توان از سیستم های کاملاً خودکار و نیمه خودکار در این حوزه استفاده نمود. همان طور که گفته شد، این سیستم ها بخش های مختلفی دارند که کاملاً به یکدیگر متصل هستند و خروجی هر مرحله، ورودی مرحله بعدی است. در این پایان نامه، پیش پردازش هایی بر روی تصاویر انجام داده ایم که تصاویر، در بهترین حالت به مرحله بخش بندی برسند. مرحله بعدی، مرحله بخش بندی تصویر است که از اهمیت ویژه ای برخوردار است. با توجه به دقت و حساسیت این مرحله، از یکی از بهترین نسخه های U-NET یعنی BCDU-Net استفاده کرده ایم. یکی از چالش های بزرگی که در قسمت بخش بندی وجود دارد، وجود بافت هایی در اطراف ناحیه ریه است. این بافت ها جزو بافت ریوی است اما اغلب روش های بخش بندی، آن را لحاظ نمی کنند. این نسخه این مشکل بزرگ را برطرف کرده است. مرحله بعدی که از خروجی قسمت بخش بندی تغذیه می کند، قسمت استخراج ویژگی و طبقه بندی است. در این قسمت ما از شبکه عصبی کانولوشنی سه بعدی استفاده کرده ایم و با تغییراتی که بر روی الگوریتم آن به وجود آورده ایم، خروجی بسیار بهتری را در اختیار ما قرار می دهد. در آخرین مرحله، برای بالا بردن دقت طبقه بندی، از تقویت گرادیان استفاده کرده ایم. در این مرحله از شبکه XGBOOST استفاده کرده ایم و خروجی لایه تماماً متصل را به این شبکه متصل کرده ایم. با توجه به درصد دقتی که در فصل پنجم به طور کامل شرح داده شد، با ترکیب این شبکه ها و با تغییرات بر روی بدنه ی این الگوریتم ها به یک

خروجی کامل رسیدیم و درصد دقت ۹۷/۹۴ را به دست آورده‌ایم. یکی از معایبی که این تحقیق دارد، دخالت جزئی در بخش برچسب‌گذاری تصاویر است. به طور کلی از خطای انسانی نمی‌توان چشم‌پوشی کرد و همین امر می‌تواند باعث پایین آمدن دقت در سیستم شود.

۳-۶ پیشنهاد برای کارهای آتی

با توجه به این که این تحقیق دارای بخش‌های مختلفی است، می‌توان با تغییرات روی مراحل مختلف آن، خروجی را ارتقا داد. در مرحله‌ی بخش‌بندی، می‌توان از مدل‌های دیگر استفاده کرد و بخش‌بندی را بهبود بخشید. در مرحله استخراج ویژگی، می‌توان بر روی شبکه کانولوشنی، تغییراتی دیگری به وجود آورد و تشخیص را قوی‌تر کرد. یکی از راه‌های دیگر که می‌توان روی آن تمرکز کرد، استفاده از روش‌های افزایش داده است که باعث بالارفتن تعداد تصاویر شده و هر چه تصاویر بیشتر باشد، خروجی بهتر و تشخیص قوی‌تر خواهد بود. همان‌طور که در فصل‌های قبل گفته شد، تصاویر دارای برچسب هستند و این برچسب‌ها توسط متخصصان بر روی عکس‌ها زده شده‌اند. به همین علت، ما از روش یادگیری عمیق نظارت شده در این تحقیق استفاده کرده‌ایم. یکی از پیشنهادهایی که می‌توان از آن در این تحقیق استفاده کرد، استفاده از یادگیری عمیق بدون نظارت است. به این صورت که تصاویر دارای برچسب نیستند و خود سیستم یاد می‌گیرد که کدام تصویر دارای سرطان است و کدام نیست. این باعث از بین رفتن خطای انسانی در زمان برچسب‌گذاری می‌شود.

فهرست واژگان

عنوان فارسی	عنوان انگلیسی
اندازه	Size
برچسب	Label
تقسیم‌بندی	Segmentation
تقسیم‌بندی مبتنی بر پایه آستانه‌گذاری	Region based segmentation
تقسیم‌بندی مبتنی بر خوشه‌بندی	Thresholding based segmentation
تقسیم‌بندی مبتنی بر لبه‌های	Edge based segmentation
تقسیم‌بندی مبتنی بر ناحیه	Edge based segmentation
ساختارها	Structure
گسترش	Dilation
لبه‌یاب سوبل	Prewitt edge detector
لبه‌یاب کنی	Sobel edge detector
یادگیری عمیق	Deep learning
آستانه‌گذاری	thresholding
ندول	Nodule
ریه	lungs
درخت برونشی	Bronchial Tree
برونش	Bronchi
برونشیول	Bronchioles
آلوئول	Alveoli
دیافراگم	Diaphragm
برچسب	Tag
سرصفحه	Header
فرسایش	Erosion
گسترش	Dilation
باز کردن	Opening
بستن	Closing
صاف	Smooth
رگ‌ها	Veins

Recognition	آشکارسازی
Detection	تشخیص
Blood vessels	عروق خونی

فهرست اختصارات

عنوان اصلی	عنوان مختصر شده
Computer Aided Design	CAD
United States	US
DICOM	Digital Imaging and Communication in Medicine
ACR	American College of Radiology
NEMA	National Electrical Manufacturers Association
AAPM	American Association of Physicists in Medicine
Computed Tomography	CT
Standard Temperature and Ppressure	STP

فهرست مراجع

- [۱] 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. ۱۰۹-۱۲۱. Springer, Cham, ۲۰۱۴.
- [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
- [۳] 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 ۲۰, no. ۶ (۲۰۰۱): ۴۹۰-۴۹۸.
- [۴] Wafaa Alakwaa, Mohammad Nassef, Amr Badr "Lung Cancer Detection and Classification with ۳D Convolutional Neural Network (۳D-CNN)" (IJACSA) International Journal of Advanced Computer Science and Applications, ۲۰۱۷
- [۵] Skourt, Brahim Ait, Abdelhamid El Hassani, and Aicha Majda. "Lung CT image segmentation using deep neural networks." Procedia Computer Science ۱۲۷ (۲۰۱۸): ۱۰۹-۱۱۳.
- [۶] Skourt, Brahim Ait, Abdelhamid El Hassani, and Aicha Majda. "Lung CT image segmentation using deep neural networks." Procedia Computer Science ۱۲۷ (۲۰۱۸): ۱۰۹-۱۱۳.
- [۷] Kulkarni, Anjali, and Anagha Panditrao. "Classification of lung cancer stages on CT scan images using image processing." In ۲۰۱۴ IEEE International Conference on Advanced Communications, Control and Computing Technologies, pp. ۱۳۸۴-۱۳۸۸, IEEE, ۲۰۱۴.
- [۸] Jithy, P. K., and Philomina Simon. "Review on Automatic Segmentation Techniques in Medical Images." (۲۰۱۷).
- [۹] 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 ۲۰۱۹ IEEE ۱۶th International Symposium on Biomedical Imaging (ISBI ۲۰۱۹), pp. ۱۳۵۳-۱۳۵۷. IEEE, ۲۰۱۹.
- [۱۰] 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. ۱-۲۰. Academic Press, ۲۰۱۹.
- [۱۱] Nasser, Ibrahim M., and Samy S. Abu-Naser. "Lung Cancer Detection Using Artificial Neural Network." International Journal of Engineering and Information Systems (IJEAIS) ۳, no. ۲ (۲۰۱۹): ۱۷-۲۳.
- [۱۲] 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 ۱۸, no. ۱۰ (۲۰۰۳): ۱۲۰۹-۱۲۱۸.
- [۱۳] R. Nall, "Medical News Today," ۱۲ ۱۱ ۲۰۱۸. [Online]. Available: <https://www.medicalnewstoday.com/articles/۳۲۳۶۴۸.php>. [Accessed ۱۴ ۳ ۲۰۱۹]
- [۱۴] Sun, Shanhui, Christian Bauer, and Reinhard Beichel. "Automated ۳-D segmentation of lungs with lung cancer in CT data using a novel robust active shape model approach." IEEE transactions on medical imaging ۳۱, no. ۲ (۲۰۱۱): ۴۴۹-۴۶۰.
- [۱۵] Jameson, J. Larry. Harrison's principles of internal medicine. McGraw-Hill Education,, ۲۰۱۸.
- [۱۶] American Cancer Society's publication, Cancer Facts & Figures ۲۰۲۰, Available: <http://www.cancer.org/research/cancerfactsstatistics/cancerfacts figures۲۰۲۰>. /accessed on ۲nd November, ۲۰۲۰
- [۱۷] MacDonald, Sharyn LS, and David M. Hansell. "Staging of non-small cell lung cancer: imaging of intrathoracic disease." European journal of radiology ۴۵, no. ۱ (۲۰۰۳): ۱۸-۳۰.
- [۱۸] 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 ۲۵, no. ۲ (۱۹۹۴): ۶۷-۸۰.
- [۱۹] 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 ۶۵ (۲۰۱۸): ۵۸-۶۸.
- [۲۰] 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. ۲۸۳-۲۸۹. Springer, Cham, ۲۰۱۷.

- [୧୧] 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* ୪୫, no. ୪ (୨୦୧୮): ୨୦୦୧-୨୦୧୨.
- [୧୨] Schaefer-Prokop, C., and M. Prokop. "New imaging techniques in the treatment guidelines for lung cancer." *European Respiratory Journal* ୨୭, no. ୨ suppl (୨୦୦୮): ୪୧୫-୪୨୫.
- [୧୩] 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* ୧୨୭, no. ୩ (୨୦୧୮): ୫୨୭-୫୩୦.
- [୧୪] Varma, Dandu Ravi. "Managing DICOM images: Tips and tricks for the radiologist." *The Indian journal of radiology & imaging* ୨୨, no. ୧ (୨୦୧୮): ୫.
- [୧୫] 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* ୬୦, no. ୧୧ (୨୦୦୫): ୧୧୩୩-୧୧୪୦.
- [୧୬] 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* ୫, no. ୭ (୨୦୧୪).
- [୧୭] 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* ୧, no. ୧ (୨୦୧୫): ୧୩୧-୧୩୬.
- [୧୮] Huidrom, Ratishchandra, Yambem Jina Chanu, and Kh Manglem Singh. "A Study on Lung Segmentation for the Detection and Diagnosis of Lung Cancer." *International Journal* ୬, no. ୧୧ (୨୦୧୬).
- [୧୯] 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* ୫, no. ୨ (୨୦୧୮): ୧୧୦-୧୨୨.
- [୨୦] Armato III, Samuel G., and William F. Sensakovic. "Automated lung segmentation for thoracic CT: Impact on computer-aided diagnosis." *Academic Radiology* ୧୧, no. ୭ (୨୦୦୫): ୧୦୧୧-୧୦୨୧.
- [୨୧] 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 scheme." *Academic radiology* ୧୦, no. ୧୧ (୨୦୦୩): ୧୨୧୫-୧୨୨୬.
- [୨୨] 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* ୨୦, no. ୬ (୨୦୦୧): ୫୭୦-୫୭୮.
- [୨୩] Sun, Xuejun, Haibo Zhang, and Huichuan Duan. "3D computerized segmentation of lung volume with computed tomography." *Academic radiology* ୧୩, no. ୬ (୨୦୦୬): ୬୮୦-୬୮୮.
- [୨୪] 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* ୧୬, no. ୬ (୧୯୯୮): ୮୨୮-୮୩୭.
- [୨୫] 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* ୨୫, no. ୧୨ (୧୯୯୮): ୨୫୩୨-୨୫୩୭.
- [୨୬] 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 ୨୦୦୮)*, vol. ୨, pp. ୨୮୦-୨୮୭. IEEE, ୨୦୦୮.
- [୨୭] Ukil, Soumik, and Joseph M. Reinhardt. "Smoothing lung segmentation surfaces in three-dimensional x-ray ct images using anatomic guidance." *Academic radiology* ୧୮, no. ୧୨ (୨୦୦୫): ୧୦୦୮-୧୦୧୧.
- [୨୮] 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. ୧୫, pp. ୨୨୮-୨୩୩. ୨୦୦୬.
- [୨୯] Korfiatis, Panagiotis, S. Skiadopoulos, P. Sakellaropoulos, C. Kalogeropoulou, and Lena Costaridou. "Combining 3D wavelet edge highlighting and 3D thresholding for lung segmentation in thin-slice CT." *The British journal of radiology* ୮୦, no. ୯୬୦ (୨୦୦୮): ୯୭୬-୯୮୫.
- [୩୦] Sonka, Milan, Vaclav Hlavac, and Roger Boyle. *Image processing, analysis, and machine vision*. Cengage Learning, ୨୦୧୫.
- [୩୧] Davies, E. Roy. *Computer and machine vision: theory, algorithms, practicalities*. Academic Press, ୨୦୧୮.

- [୧୪] Mitiche, Amar, and Jake K. Aggarwal. "Image segmentation by conventional and information-integrating techniques: a synopsis." *Image and Vision Computing* ୩, no. ୪ (୧୯୮୦): ୦୦-୧୪.
- [୧୫] 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* ୧୧, no. ୧ (୧୯୯୩): ୧୩୪୩-୧୩୪୮.
- [୧୬] Haralick, Robert M., and Linda G. Shapiro. "Image segmentation techniques." *Computer vision, graphics, and image processing* ୪୯, no. ୧ (୧୯୮୦): ୧୦୦-୧୩୪.
- [୧୭] Kundu, Amlan. "Local segmentation of biomedical images." *Computerized Medical Imaging and Graphics* ୧୧, no. ୩ (୧୯୯୦): ୧୪୩-୧୪୪.
- [୧୮] Chow, C. K., and T. Kaneko. "Automatic boundary detection of the left ventricle from cineangiograms." *Computers and biomedical research* ୦, no. ୧ (୧୯୪୪): ୩୮୮-୧୯୦.
- [୧୯] Talakoub, Omid, Javad Alirezaie, and Paul Babyn. "Lung segmentation in pulmonary ct images using wavelet transform." In ୪୦୦୪ IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'୦୪, vol. ୧, pp. ୧-୧୦୩. IEEE, ୪୦୦୪.
- [୨୦] 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* ୨୦, no. ୧୨ (୨୦୦୧): ୧୨୧୨-୧୨୦୦.
- [୨୧] 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* ୧୯୪, no. ୨ (୧୯୯୦): ୩୯୪-୧୯୬.
- [୨୨] 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 ୨୦୧୧ international conference on engineering and technology (ICET), pp. ୧-୧. IEEE, ୨୦୧୧.
- [୨୩] 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* ୩୪, no. ୩ (୧୯୯୪): ୧୧୮-୧୨୧.
- [୨୪] 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* ୦୩, no. ୦ (୨୦୦୬): ୮୯୧-୯୦୪.
- [୨୫] Hu, Xiaoping, Noam Alperin, David N. Levin, Kim K. Tan, and Michel Mengeot. "Visualization of MR angiographic data with segmentation and volume-rendering techniques." *Journal of Magnetic Resonance Imaging* ୧, no. ୦ (୧୯୯୧): ୦୩୯-୦୧୧.
- [୨୬] 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* ୦, no. ୦ (୧୯୮୪): ୩୧୦-୩୦୪.
- [୨୭] 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 ୧୮th International Conference on Systems, Signals and Image Processing-IWSSIP, pp. ୩୯୦-୧୯୦. ୨୦୧୧.
- [୨୮] 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 ୨୦୦୪, Bioengineering Solutions for Latin America Health, pp. ୩୪୦-୩୪୩. Springer, Berlin, Heidelberg, ୨୦୦୪.
- [୨୯] Sluimer, Ingrid, Mathias Prokop, and Bram Van Ginneken. "Toward automated segmentation of the pathological lung in CT." *IEEE transactions on medical imaging* ୨୧, no. ୮ (୨୦୦୦): ୧୦୨୦-୧୦୩୮.
- [୩୦] 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* ୨୪, no. ୧ (୨୦୦୫): ୧୧୪-୧୧୮.
- [୩୧] Gonzalez, Rafael C., Richard E. Woods, and R. Barry. "Masters." *Digital Image Processing, Third Edition. Journal of Biomedical Optics* ୧୧, no. ୨ (୨୦୦୬): ୦୨୯୯୦୧.
- [୩୨] Canny, John. "A computational approach to edge detection." *IEEE Transactions on pattern analysis and machine intelligence* ୧ (୧୯୮୬): ୧୪୧-୧୫୮.
- [୩୩] Maini, Raman, and Himanshu Aggarwal. "Study and comparison of various image edge detection techniques." *International journal of image processing (IJIP)* ୩, no. ୧ (୨୦୦୬): ୧-୧୧.
- [୩୪] "A survey of various image segmentation techniques." Dept. of Electrosience, Box ୧୧୮ (୨୦୧୦).
- [୩୫] Hedberg, Hugo. "A survey of various image segmentation techniques." Dept. of Electrosience, Box ୧୧୮ (୨୦୧୧).
- [୩୬] How the machine 'thinks': Understanding opacity in machine learning algorithms." Jenna Burrell"

- [۱۵] A Comparative Study of Training Algorithms for Supervised Machine Learning." Hetal Bhavsar, Amit Ganatra ".International Journal of Soft Computing and Engineering (IJSCE) ISSN: ۲۲۳۱-۲۳۰۷, Volume-۲, Issue-۴, September ۲۰۱۲
- [۱۶] A Multiple Linear Regression Approach for the Analysis of Stress Factors of Faculty in Higher Educational Institutions" P.Ramesh Reddy^۱ and Dr.K.L.A.P.Sarma^۲".International Journal of Mathematics And its Applications , Volume ۳, Issue ۴{A (۲۰۱۵), ۹۵{۱۰۳. , ISSN: ۲۳۴۷-۱۵۵۷
- [۱۷] A novel high-sulfur fine coal clean desulfurization pretreatment : Microwave magnetic separation, high-gradient effect and magnetic strengthen. "BoZhang, GuangqingZhu.
- [۱۸] Use of a support vector machine for categorizing free-text notes: assessment of accuracy across two institutions . "Adam Wrigh, Allison B McCoy, Stanislav Henkin, Abhivyakti Kale, Dean F Sittig.
- [۱۹] Excellence networks in science: A Web-based application based on Bayesian multilevel logistic regression (BMLR) for the identification of institutions collaborating successfully . "MoritzStefaner,Felixde Moya AnegóncRüdigerMutzd. Journal of Informetrics Volume ۱۰, Issue ۱, February ۲۰۱۶, Pages ۳۱۲-۳۲۷
- [۲۰] Introduction: The Challenge of Reinforcement Learning . "Part of the The Springer International Series in Engineering and Computer Science book series (SECS, volume ۱۷۳)
- [۲۱] Unsupervised and Semi-supervised Multi-class Support Vector Machines . "Linli Xu, Dale Schuurmans "
- [۲۲] Silva, Flávio Henrique Schuindt da. "Deep learning for Corpus Callosum segmentation in brain magnetic resonance images." (۲۰۱۸).
- [۲۳] Shen, Dinggang, Guorong Wu, and Heung-Il Suk. "Deep learning in medical image analysis." Annual review of biomedical engineering ۱۹ (۲۰۱۷): ۲۲۱-۲۴۸.
- [۲۴] Patterson, Josh, and Adam Gibson. Deep learning: A practitioner's approach. " O'Reilly Media, Inc.", ۲۰۱۷.
- [۲۵] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." In Advances in neural information processing systems, pp. ۱۰۹۷-۱۱۰۵. ۲۰۱۲.
- [۲۶] 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 ۳۶, no. ۲ (۲۰۱۸): ۶۳-۷۱.
- [۲۷] LeCun, Yann, Léon Bottou, Yoshua Bengio, and Patrick Haffner. "Gradient-based learning applied to document recognition." Proceedings of the IEEE ۸۶, no. ۱۱ (۱۹۹۸): ۲۲۷۸-۲۳۲۴.
- [۲۸] Menegola, Afonso, Michel Fornaciali, Ramon Pires, Sandra Avila, and Eduardo Valle. "Towards automated melanoma screening: Exploring transfer learning schemes." arXiv preprint arXiv:۱۶۰۹.۰۱۲۲۸ (۲۰۱۶).
- [۲۹] 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. ۷۷۸-۷۷۸. ۲۰۱۶.
- [۳۰] 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. ۲۳۴-۲۴۱. Springer, Cham, ۲۰۱۵.
- [۳۱] Norman, Berk, Valentina Podoia, and Sharmila Majumdar. "Use of ۲D U-Net convolutional neural networks for automated cartilage and meniscus segmentation of knee MR imaging data to determine relaxometry and morphometry." Radiology ۲۸۸, no. ۱ (۲۰۱۸): ۱۷۷-۱۸۵.

[۸۲] جلالی یگانه؛ منصور فاتح و محسن رضوانی، ۱۳۹۹، " بخش‌بندی خودکار تصاویر سی‌تی‌اسکن ریه به کمک یادگیری عمیق. " پایان‌نامه

کارشناسی ارشد، شاهرود، دانشگاه صنعتی شاهرود، صفحه ۵۱

- [۸۳] 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:۱۸۰۴.۰۳۹۹۹ (۲۰۱۸).
- [۸۴] Simie Endalew, Kaur Mandeep; International Journal of Advance Research, Ideas and Innovations in Technology , ۲۰۱۹.
- [۸۵] Qingfang Yang ,Yunfeng Fang ,Lili Zheng ,Xiangyu Zhou ,Bo Peng" An expressway traffic incident detection method based on Convolutional Neural Network and Extreme Gradient Boosting "۲۰۲۰.
- [۸۶] <https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI>

- [⁸⁷] Shivashankar, T. M., and S. B. Shivakumar. "Insights on Security Improvements and Implications of Artificial Intelligence in MANET." *Communications* 7 (2019): 1-9.
- [⁸⁸] Azad, Reza, Maryam Asadi-Aghbolaghi, Mahmood Fathy, and Sergio Escalera. "Bi-directional ConvLSTM U-net with Densely connected convolutions." In *Proceedings of the IEEE International Conference on Computer Vision Workshops*, pp. 1-11. 2019.
- [⁸⁹] 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. 4711-4718. 2017.
- [⁹⁰] Tsang, Sik-Ho. "Review: Densenet-dense convolutional network (image classification)." en línea.[consulta: 7 abril 2019]. Disponible en: <https://towardsdatascience.com/review-densenetimage-classification-b6631a8ef802> (2019).
- [⁹¹] 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. 710-721. 2018.
- [⁹²] Townsend, James T. "Theoretical analysis of an alphabetic confusion matrix." *Perception & Psychophysics* 9, no. 1 (1971): 41-50.
- [⁹³] Flach, Peter, and Meelis Kull. "Precision-recall-gain curves: PR analysis done right." In *Advances in neural information processing systems*, pp. 838-846. 2010.
- [⁹⁴] Chicco, Davide, and Giuseppe Jurman. "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation." *BMC genomics* 19, no. 1 (2018): 6.
- [⁹⁵] Muschelli, John. "Roc and auc with a binary predictor: a potentially misleading metric." *Journal of Classification* (2019): 1-13.
- [⁹⁶] A. Chon, N. Balachandar, and P. Lu, "Deep convolutional neural networks for lung cancer detection," tech. rep., Stanford University, 2017.

Abstract

Cancer is one of the deadliest diseases in the world. One of the reasons for the high mortality of this disease has been the wrong diagnosis of the doctor or the delay in diagnosing this disease. Among all cancers, lung cancer is difficult to diagnose due to its uneven and uncoordinated structure. With the advancement of science and technology, the diagnosis of this type of disease can be taken out of the traditional mode and done with automatic and semi-automatic systems. In order to have a system that can diagnose the disease without human intervention, deep learning techniques must be used. This implementation consists of two very important parts: segmentation and classification. The output of the segmentation part is the input of the feature extraction and classification part. There are many problems with segmentation, the most important of which is the uneven structure of the lungs. To eliminate this problem, many methods have been proposed, and the problem with most of these methods is the high level of false positives. In this research, a new U-Net network structure has been used for segmentation, which has had the best segmentation result to date. After this step, to extract the feature, the convolutional neural network is used, which is a suitable structure for extracting the visual features. Then, gradient amplification is used to increase the classification accuracy and the output of the convulsive neural network is connected to another network. In summary, to diagnose lung cancer, pre-processing images were performed and then lung images were segmented using the BCDU-Net neural network. Then, using three-dimensional convolutional neural network and with changes on the layers of this network, feature extraction is performed and in the last step, the fully connected layer is connected to the XGBOOST network to finally classify with high accuracy.

Keywords: segmentation, convulsive neural network, lung, U-Net, CNN
‡D, cancer diagnosis, deep learning



Shahrood University of Technology

Faculty of Computer Engineering

M.Sc Thesis in Artificial Intelligence Engineering

Lung cancer detection using Deep Learning

By:

Mohammad Aminzadeh

Supervisor:

Dr. Mansoor Fateh

Advisor:

Dr. Morteza Zahedi

Dr. Hoda Mashayekhi