



دانشگاه تهران



پردیس دانشکده‌های فنی

دانشکده مهندسی برق و کامپیوتر

تشخیص خودکار سرطان سینه و میزان پیشرفت آن با استفاده از تصاویر ماموگرافی

پایان‌نامه برای دریافت وجه کارشناسی

در رشته مهندسی برق گرایش مخابرات

فرزاد مهری

۸۱۰۱۹۴۴۱۰

استاد راهنما:

دکتر محمد علی اخائی

شهریور ۱۳۹۸

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

تعهدنامه اصالت اثر

باسمه تعالی

اینجانب فرزاد مهری تأیید می‌کنم که مطالب مندرج در این پایان نامه حاصل کار پژوهشی اینجانب است و به دستاوردهای پژوهشی دیگران که در این نوشته از آنها استفاده شده است مطابق مقررات ارجاع گردیده است. این پایان نامه قبلاً برای احراز هیچ مدرک هم سطح یا بالاتر ارائه نشده است.

کلیه حقوق مادی و معنوی این اثر متعلق به دانشکده فنی دانشگاه تهران می‌باشد.

نام و نام خانوادگی دانشجو: فرزاد مهری

امضای دانشجو:



چکیده^۱

از تصاویر پزشکی برای تشخیص بیماری‌ها و روش درمان آن‌ها استفاده می‌شود. یکی از کاربردهای مهم تصاویر پزشکی، تشخیص ابتدایی وجود سرطان و میزان گسترش آن است. برای این منظور تصاویر پزشکی که به روش‌های مختلفی تهیه می‌شود توسط پزشک متخصص بررسی شده و ویژگی‌های تصویر بر پایه دانش پزشکی مورد بررسی قرار می‌گیرد. با توجه به اهمیت بالای تحلیل دقیق تصاویر پزشکی، استفاده از کامپیوترها برای تحلیل تصاویر همواره مد نظر بوده است. در تشخیص به کمک کامپیوتر با به‌کارگیری دانش هوش مصنوعی، بینایی کامپیوتری، پاتولوژی و رادیولوژی به پزشک در تشخیص بیماری کمک می‌شود. با توجه به موفقیت‌های چند سال اخیر پردازش کامپیوتری تصاویر با استفاده از روش‌های یادگیری عمیق، در این پروژه تلاش شده است الگوریتم‌ها و شبکه‌های مبتنی بر یادگیری عمیق بر تصاویر ماموگرافی اعمال شده و انواع سرطان سینه (رایج‌ترین سرطان در بین زنان) و در صورت امکان بدخیم یا خوشخیم بودن توده سرطانی احتمالی به طور کاملاً خودکار تشخیص داده شود. در مرحله‌ی اول به بررسی مسئله‌ی تشخیص توده‌های سرطانی^۲ پرداخته شده است. تشخیص محل این ناهنجاری‌ها مسئله‌ی مهمی است و می‌توان از نتیجه‌ی این مرحله برای پردازش‌ها و تشخیص‌های بعدی از جمله تشخیص خوشخیم یا بدخیم بودن و نیز تقسیم‌بندی^۳ ناهنجاری‌ها استفاده کرد. در قسمت بعدی به تقسیم‌بندی توده‌ها پراخته شد و در قسمت آخر، ناهنجاری‌ها به دو نوع خوش‌خیم و بدخیم دسته‌بندی شدند.

کلمات کلیدی: تشخیص به کمک کامپیوتر، تصویر ماموگرافی، سرطان سینه، یادگیری عمیق، یادگیری انتقالی، شبکه‌های کانوولوشنی

¹ Abstract

² Mass

³ Segmentation

فهرست مطالب

فصل ۱: مقدمه و بیان مساله	۱
۱-۱- مقدمه	۱
۱-۲- تاریخچه‌ای از موضوع تحقیق	۱
۱-۳- شرح مسئله تحقیق	۲
۱-۴- اهداف و آرمان‌های کلی تحقیق	۳
۱-۵- روش انجام تحقیق	۴
۱-۶- ساختار پایان‌نامه	۵
فصل ۲: مفاهیم اولیه	۶
۲-۱- شبکه‌های عصبی مصنوعی و یادگیری عمیق	۷
۲-۲- شبکه‌های نورونی کانولوشنی (CNN)	۸
۲-۲-۱- VGGNet	۱۰
۲-۲-۲- ResNet	۱۰
۲-۳- شبکه Faster RCNN	۱۱
۲-۴- شبکه Mask RCNN	۱۲
۲-۵- شبکه YOLO	۱۳
۲-۶- شبکه RetinaNet	۱۳
فصل ۳: آموزش و تست شبکه‌ها	۱۵
۳-۱- مقدمه	۱۶
۳-۲- دیتاست	۱۶
۳-۳- یادگیری انتقالی	۱۷

۱۷	۳-۴- کشف توده‌های سرطانی.....
۱۷	۳-۴-۱ RetinaNet.....
۱۸	۳-۴-۲ YOLO.....
۱۹	۳-۴-۳ Mask RCNN.....
۱۹	۳-۵- تشخیص خوشخیم و یا بدخیم بودن توده‌ها.....
۱۹	۳-۶- دسته‌بندی پیکسل به پیکسل توده‌ها.....
۱۹	۳-۶-۱ Mask RCNN استفاده از.....
۲۰	۳-۷- معیار ارزیابی.....
۲۰	۳-۷-۱ تعاریف اولیه.....
۲۲	۳-۷-۲ معیارهای اصلی.....
۲۲	۳-۸- نتایج کشف توده‌ها.....
۲۵	۳-۹- نتایج دسته‌بندی پیکسل به پیکسل توده‌ها.....
۲۷	۳-۱۰- نتایج دسته‌بندی توده‌های خوشخیم و بدخیم.....
۲۹	۳-۱۱- تحلیل نتایج.....
۲۹	۳-۱۲- خلاصه و جمع‌بندی.....
۳۰	فصل ۴: جمع‌بندی، نتیجه‌گیری و پیشنهادها.....
۳۱	۴-۱- جمع‌بندی.....
۳۱	۴-۲- نتیجه‌گیری.....
۳۱	۴-۲-۱ محدودیت‌ها.....
۳۳	فصل ۵: مراجع.....

فهرست شکل‌ها

- شکل ۲-۱، مقایسه دقت روش یادگیری عمیق در مقایسه با سایر روشهای یادگیری ماشین [۸] ۷
- شکل ۲-۲، مقایسه‌ی وزنهای پراکنده (بالا) و وزنهای کامل (پایین) [۹] ۹
- شکل ۲-۳، در بالا تنها سه پارامتر بین تمام ورودیها به اشتراک گذاشته شده است. اما در پایین هیچ پارامتر مشترکی وجود ندارد [۹] ۹
- شکل ۲-۴، ساختار شبکه VGGNet ۱۰
- شکل ۲-۵، ساختار شبکه ResNet ۱۱
- شکل ۲-۶، Faster RCNN [۱۰] ۱۲
- شکل ۲-۷، ساختار شبکه‌ی رتینانت ۱۴
- شکل ۳-۱، نمونه‌ای از تشخیص‌های صحیح توده‌ها، Mask RCNN (شکل راست)، رتینانت (شکل وسط) و YOLO (شکل چپ) ۲۲
- شکل ۳-۲، نمونه‌ای از تشخیص‌های اشتباه، Mask RCNN (شکل راست)، رتینانت (شکل وسط) و YOLO (شکل چپ) ۲۲
- شکل ۳-۳، توده‌ای که توسط هیچ‌کدام از شبکه‌ها کشف نشده است ۲۳
- شکل ۳-۴، نمودار TPR بر حسب IoU ۲۴
- شکل ۳-۵، نمودار FROC ۲۵
- شکل ۳-۶، مسک ارائه شده توسط Mask RCNN، قسمت رنگی مسک پیشبینی شده و قسمت تیره مسک حقیقی است ۲۵
- شکل ۳-۷، نمودار TPR بر حسب IoU در شبکه‌ی Mask RCNN ۲۶
- شکل ۳-۸، نمودار FROC شبکه‌ی Mask RCNN برای دسته‌بندی پیکس به پیکسل تصویر ۲۷
- شکل ۳-۹، نمودار TPR بر حسب IoU شبکه‌ی رتینانت با دسته‌بندی خوش‌خیم و بدخیم ۲۸
- شکل ۳-۱۰، نمودار FROC شبکه‌ی رتینانت با دسته‌بندی خوش‌خیم و بدخیم ۲۸

فهرست علائم اختصاری

CNN	Convolutional Neural Network
TP	True Positive
FP	False Positive
FN	False Negative
IoU	Intersection over Union
FPPI	False Positive per Image
TPR	True Positive Rate
CAD	Computer Aided Diagnosis

فصل ۱

فصل ۱:

مقدمه و بیان مساله

۱-۱- مقدمه

سرطان سینه شایع‌ترین سرطان در بین زنان است. در ایران میزان بروز این نوع سرطان بین ۱۱.۹ تا ۲۳.۵ نفر در هر ۱۰۰۰۰۰ نفر ذکر شده است [۱]. علی‌رغم اینکه تنها روش قطعی تشخیص سرطان سینه نمونه برداری از بافت است، استفاده از تصاویر ماموگرافی برای تشخیص زود هنگام سرطان حتی قبل از بروز علائم و نیز برای شناسایی موارد احتمالا سرطانی بسیار حائز اهمیت است و منجر به کاهش هزینه‌ها، کاهش نمونه‌برداری‌های غیرضروری از بافت و در نتیجه کاهش میزان اضطرابی که بیماران تحمل می‌کنند، می‌شود. با توجه به اینکه تشخیص زودهنگام سرطان سینه به میزان قابل توجهی احتمال درمان بیماری را افزایش می‌دهد [۳]، ارائه روشی برای تشخیص دقیق سرطان با استفاده از تصاویر ماموگرافی ضروری به نظر می‌رسد.

۱-۲- تاریخچه‌ای از موضوع تحقیق

به طور کلی تحقیقات در زمینه سرطان سینه متمرکز بر توسعه سیستم‌های تشخیص به کمک کامپیوتر^۱

^۱ CAD

بوده است تا ابزاری برای کمک به پزشکان در تشخیص فراهم شود. تا قبل از موفقیت روش‌های مبتنی بر شبکه‌های عصبی عمیق در مسائل پردازش تصویر، اغلب برای تشخیص سرطان سینه ابتدا تعدادی ویژگی^۱ به صورت دستی استخراج می‌شد و در نهایت دسته‌بندی تصاویر با استفاده از این ویژگی‌ها انجام می‌شد. با توجه به پیچیدگی توده‌های سرطانی و تنوع آن‌ها این روش‌ها کارایی پایینی دارد. برای مثال رانگیان در [۴] با استفاده از ویژگی‌های مورفولوژیک برای تشخیص میزان ناهمواری مرزهای توده‌های سرطانی روشی برای دسته‌بندی ناهنجاری‌های تصویر ماموگرافی ارائه داد. ورما و دیگران در [۵] ویژگی‌های مختلفی مانند تراکم و مورفولوژی و... استخراج کرده و با استفاده از شبکه نورونی، تصاویر را دسته‌بندی کردند. بوئرا در [۶] با استفاده از DWT^۲ و GLCM^۳ ماتریسی از ویژگی‌ها تولید کرده و از آن برای دسته‌بندی تصاویر استفاده کرد. با توجه به محدودیت استخراج دستی ویژگی‌ها، این روش‌ها کارایی محدودی دارند. در سال‌های اخیر روش‌های یادگیری عمیق بر این مسئله اعمال شده و نتایج بسیار مناسبی بدست آمده است. برای مثال دانگل در [۷] با استفاده از شبکه‌های کانولوشنی عمیق به حساسیت ۹۰٪ به ازای FPPI برابر ۱ دست یافته است.

۳-۱- شرح مسئله تحقیق

دو نمود اصلی سرطان سینه در تصاویر ماموگرافی، توده‌های سرطانی و کلسیفیکاسیون است. توده‌های سرطانی مرزهای ساده و مشخصی ندارد و در اندازه‌ها و شکل‌های مختلف در تصاویر ماموگرافی وجود دارد. توده‌های سرطانی در تصاویر ماموگرافی سفید دیده می‌شود و این مسئله باعث می‌شود در مواردی که بافت سینه متراکم است (که اغلب در افراد جوان‌تر اینگونه است)، کارایی روش‌های تشخیص کاهش یابد زیرا بافت‌های متراکم هم سفید دیده می‌شود همچنین در بافت متراکم جزئیات تصویر نیز کمتر است که مسئله را دشوارتر می‌کند. در مقابل تشخیص کلسیفیکاسیون مسئله‌ی نسبتاً ساده‌تری است. کلسیفیکاسیون به صورت نقاط درخشان در تصویر ماموگرافی مشخص می‌شود و اغلب تشخیص کلسیفیکاسیون خوش‌خیم و بدخیم با قوانین ساده‌ای قابل تشخیص است (برای مثال اغلب کلسیفیکاسیون پراکنده و بدون الگو خوش‌خیم

¹ Feature

² Discrete Wavelet Transform

³ Gray Level Co-occurrence Matrix

است و اگر کلسیفاکسیون الگوی خاصی داشته باشد، بدخیم است). تمرکز این پروژه تشخیص توده‌های سرطانی است.

در این پروژه مسئله تشخیص سرطان سینه در سه بخش بررسی می‌شود. مسئله‌ی اول تحلیل تصویر ماموگرافی و کشف محل ناهنجاری‌ها و مشخص کردن محدوده‌ی مستطیلی شکلی است که این ناهنجاری‌ها را به طور کامل در بر می‌گیرد. مسئله‌ی دوم دسته‌بندی پیکسل به پیکسل محدوده‌ی شامل ناهنجاری است به طوری که مرز ناهنجاری به طور دقیق مشخص شود. و مسئله‌ی سوم تشخیص خوش‌خیم و یا بدخیم بودن ناهنجاری کشف شده است.

۴-۱- اهداف و آرمان‌های کلی تحقیق

به دلیل دشوار بودن تحلیل تصاویر ماموگرافی، تحلیل دقیق نیاز به پزشک متخصص دارد که در بسیاری از مناطق در دسترس نیست. همچنین گاهی تشخیص پزشک‌ها متفاوت است و پزشک‌ها از معیارهای مختلفی برای تحلیل استفاده می‌کنند (علیرغم تلاش برای استانداردسازی تحلیل). با استفاده از کامپیوترها، تحلیل می‌تواند با سرعت بیشتری انجام شود و با حذف خطای انسانی، نیاز به آزمایش و بررسی‌های بیشتر کاهش یابد. زمان تشخیص سرطان در درمان آن نقش مهمی دارد. با تشخیص کامپیوتری، امکان تشخیص سریع‌تر و کم‌هزینه‌تر سرطان سینه در تمام مناطقی که به دلیل نبود نیروی انسانی امکان انجام بررسی توسط پزشک وجود ندارد، فراهم می‌شود. مهم‌ترین موضوع در مسئله تشخیص سرطان، تشخیص درست تمامی ناهنجاری‌ها (TP^۱) است و البته مطلوب است که تعداد تشخیص‌های اشتباه (انتخاب مناطقی که در واقع سرطانی نیستند، FP^۲) حداقل باشد. در این پروژه تلاش بر این است که محدوده‌ی گزارش شده برای ناهنجاری (چه به صورت محدوده‌ی مستطیلی و چه به صورت پیکسل به پیکسل) بیشترین اشتراک را با محدوده‌ی صحیح مشخص شده^۳ در داده‌ها داشته باشد. یکی از مشکلات اساسی در تحلیل خودکار تصاویر

¹ True Positive

² False Positive

³ Ground Truth

پزشکی، تنوع ابزار استفاده شده برای تصویربرداری است که تصاویر با کنتراست، رزولوشن، روشنایی و... مختلفی تولید می کنند. از طرفی محدود بودن دیتاست های پزشکی باعث می شود کارایی روش هایی که روی هر کدام از دیتاست ها آموزش داده می شوند، در دیتاست های دیگر کاهش یابد. از اهداف پروژه کاهش این اثر است.

۵-۱- روش انجام تحقیق

۵-۱-۱ دیتاست:

دیتاست مورد استفاده در این پروژه دیتاست CBIS-DDSM [۸] است. این دیتاست در واقع حاصل انجام اصلاحاتی بر روی دیتاست DDSM است از جمله حذف تصاویری که هیچ ناهنجاری ای ندارند، تبدیل تصاویر از فرمت با پشتیبانی محدود LJPG به فرمت های به روز تر و تقسیم داده ها به داده های آموزش^۱ و تست^۲ با نسبت ۸۰٪ و ۲۰٪ تا امکان مقایسه ی روش های مختلف بر روی دیتاست استاندارد فراهم شود. دیتاست CBIS-DDSM شامل ۱۵۹۲ مورد تصویر ماموگرافی شامل توده ی سرطانی است که برای هر تصویر، دسته بندی پیکسل به پیکسل ذکر شده است و همچنین نوع توده ی مورد نظر نیز مشخص شده است (خوش خیم، بدخیم).

۵-۱-۲ کشف^۳ توده های سرطانی:

در مرحله ی اول این پروژه تلاش شده است تمامی توده ها صرف نظر از خوش خیم یا بدخیم بودن کشف شده و محدوده ی شامل آن گزارش داده شود. برای این منظور شبکه های کانوولوشنی مختلف رتینانت^۴، YOLO و Mask RCNN مورد استفاده قرار گرفته و نتیجه ی نهایی حاصل ترکیب پیش بینی این شبکه ها است.

¹ Training set

² Test set

³ Mass detection

⁴ Retinanet

۵-۱-۳ تشخیص^۱ نوع توده‌های سرطانی:

برای تشخیص نوع توده‌های سرطانی از شبکه‌های ذکر شده در قسمت ۵-۱-۲ استفاده شده است. با این تفاوت که در این حالت توده‌های خوش‌خیم و بدخیم از هم متمایز در نظر گرفته شده است.

۵-۱-۴ دسته‌بندی پیکسل به پیکسل تصویر:

برای دسته‌بندی پیکسل به پیکسل و ارائه‌ی مسک^۲ از شبکه‌ی Mask RCNN استفاده شده است.

۵-۱-۵ رابط کاربری گرافیکی:

در نهایت رابط گرافیکی تهیه شده است تا استفاده از نرم‌افزار به سادگی امکانپذیر باشد.

۶-۱- ساختار پایان‌نامه

فصل دوم، شامل بررسی تعاریف اساسی مربوط به حوزه‌ی یادگیری عمیق و شبکه‌های CNN است و مفاهیم اولیه و اجزای اساسی شبکه‌های استفاده شده در این تحقیق معرفی و بررسی می‌شود.

فصل سوم در برگیرنده‌ی توضیح مربوط به نحوه‌ی پیاده‌سازی شبکه‌های CNN و نیز جزئیات آموزش این شبکه‌هاست. همچنین در این فصل در مورد معیار ارزیابی نتایج بدست آمده صحبت خواهیم کرد و نتایج این پروژه را ارائه خواهیم داد.

در نهایت، در فصل چهارم، نتیجه‌گیری‌های کلی حاصل شده در این تحقیق، پیاده‌سازی‌ها انجام شده و محدودیت‌ها مورد بحث قرار می‌گیرد.

¹ Classification

² Mask

فصل ۲

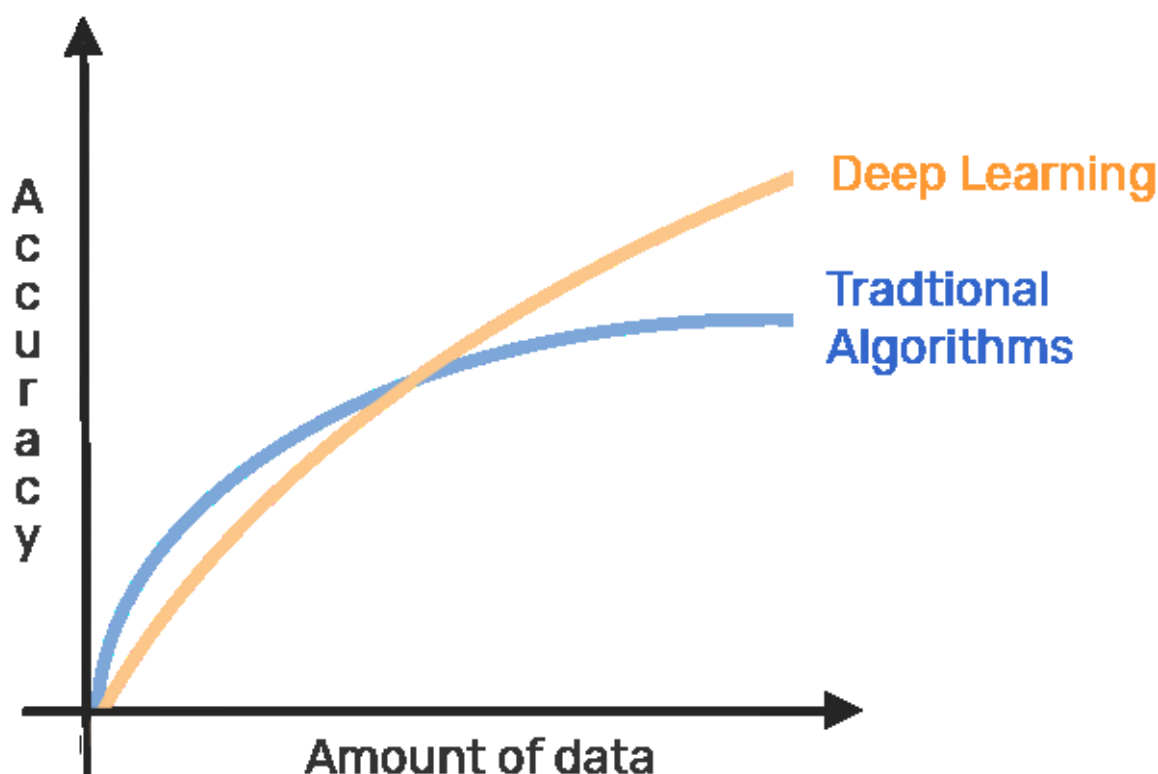
فصل ۲:

مفاهیم اولیه

در فصل پیش رو مقدمات، مفاهیم اولیه و پیش‌زمینه‌هایی را که جهت درک هر چه بهتر موضوع‌های مطرح شده در این پایان‌نامه مورد نیاز است، از جمله شبکه‌های CNN، رتینانت و سایر شبکه‌های استفاده شده توضیح می‌دهیم.

۲-۱- شبکه‌های عصبی مصنوعی و یادگیری عمیق

یادگیری ماشین به طور کلی به روش‌هایی گفته می‌شود که با استفاده از آن‌ها، کامپیوتر بدون اینکه مستقیماً برای انجام فعالیتی برنامه‌نویسی شود، با استفاده از داده‌های موجود آموزش دیده و نحوه انجام آن فعالیت را فرا می‌گیرد. یادگیری عمیق زیرمجموعه‌ای از یادگیری ماشین^۱ است که در آن از شبکه‌های عصبی مصنوعی^۲ با تعداد لایه‌های میانی زیاد استفاده می‌شود. لفظ "عمیق" به دلیل زیاد بودن تعداد لایه‌های شبکه عصبی به کار می‌رود. مزیت یادگیری عمیق در مقایسه با سایر روش‌های یادگیری ماشین این است که کارایی سایر روش‌ها با افزایش حجم داده بیشتر از یک حد آستانه دیگر بهبود چندانی نمی‌یابد اما در یادگیری عمیق با افزایش داده‌ها کارایی همواره افزایش می‌یابد (شکل ۲-۱).



شکل ۲-۱، مقایسه دقت روش یادگیری عمیق در مقایسه با سایر روش‌های یادگیری ماشین [۸]

^۱ Machine learning

^۲ Artificial Neural Networks

یادگیری عمیق در حوزه‌های مختلف دستیار مجازی، ترجمه‌ی ماشینی، تشخیص چهره و اشیا و بسیاری موارد دیگر با موفقیت اعمال شده و حتی در مواردی به دقت‌هایی بالاتر از دقت انسانی دست یافته است.

۲-۲- شبکه‌های نرونی کانولوشنی (CNN)

شبکه‌های CNN نوعی از شبکه‌های نرونی مصنوعی هستند که حداقاً در یکی از لایه‌های آن‌ها عمل ریاضی کانولوشن استفاده می‌شود. عمل کانولوشن به صورت زیر تعریف می‌شود:

$$s[n] = (x * w)[n] = \sum x[m]w[n - m] \quad \text{رابطه (۱)}$$

در شبکه‌های کانولوشنی به x ، ورودی^۱ و به w ، کرنل^۲ گفته می‌شود. البته رابطه‌ی کانولوشن را میتوان برای بیش‌تر از یک متغیر هم به کار برد، برای مثال در حالت دو متغیره:

$$s[i, j] = (I * J)[i, j] = \sum_m \sum_n I(m, n)J(i - m, j - n) \quad \text{رابطه (۲)}$$

استفاده‌ای کانولوشن مزایایی دارد از جمله وزن‌های پراکنده^۳، به اشتراک گذاری پارامترها و هموردایی نسبت به انتقال^۴[۹]. وزن‌های پراکنده در صورت استفاده از کرنل با ابعاد کوچک‌تر از ابعاد ورودی محقق می‌شود و برای مثال می‌توان با استفاده از کرنل‌های خیلی کوچک مواردی مانند گوشه‌های^۵ موجود در تصویر را تشخیص دهیم. این مسئله منجر به کاهش تعداد پارامترهای مدل می‌شود و در نتیجه خروجی با اعمال کمتری بدست می‌آید. شکل (۲-۲). اشتراک گذاری پارامترها به معنی استفاده‌ی از یک پارامتر در چند تابع در شبکه است. در شبکه‌ی CNN، هر عضو کرنل در تمامی نقاط ورودی (بجز احتمالاً در برخی گوشه‌ها) استفاده می‌شود (شکل ۲-۳). این موارد باعث می‌شود که حافظه‌ی مورد نیاز شبکه‌های CNN، تا حد زیادی کاهش یابد.

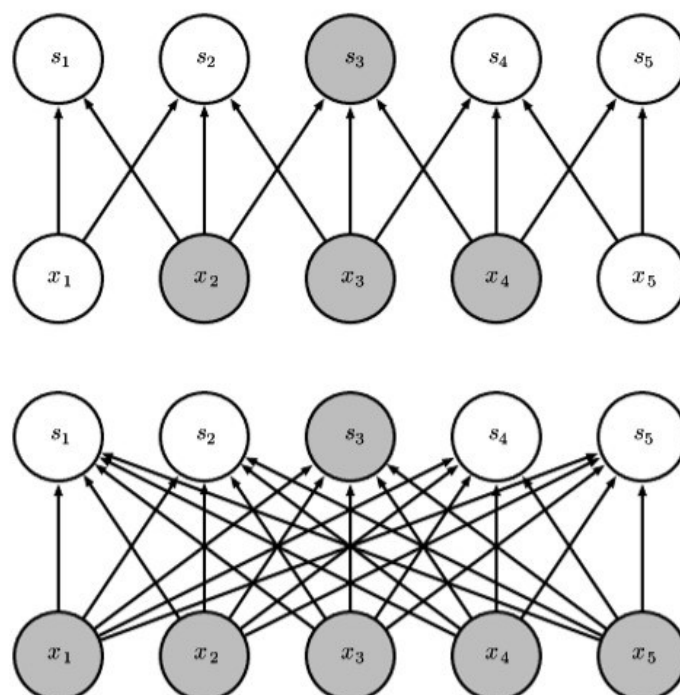
^۱ Input

^۲ Kernel

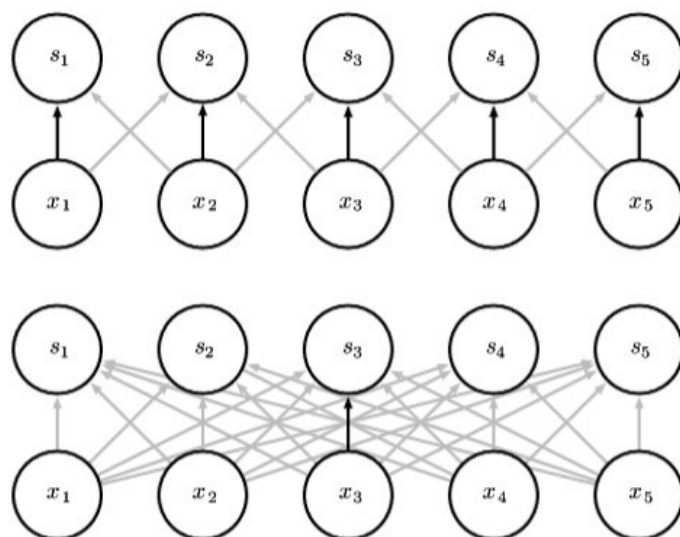
^۳ Sparse weights

^۴ Equivariance to translation

^۵ Edge



شکل ۲-۲، مقایسه‌ی وزن‌های پراکنده (بالا) و وزن‌های کامل (پایین) [۹]

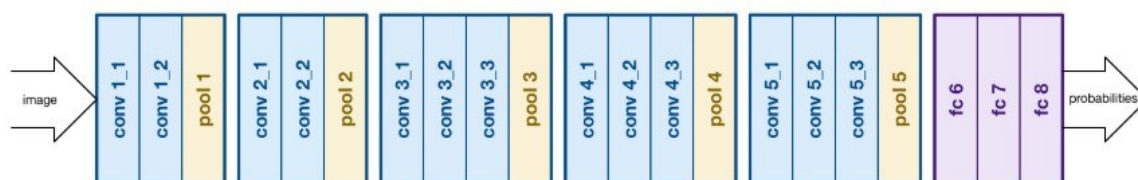


شکل ۲-۳، در بالا تنها سه پارامتر بین تمام ورودی‌ها به اشتراک گذاشته شده است. اما در پایین هیچ پارامتر مشترکی

وجود ندارد [۹]

۱-۲-۲-۱ VGGNet

VGGNet یک شبکه CNN است که نسخه‌ی اولیه‌ی آن از ۱۳۸ میلیون پارامتر تشکیل شده است. VGG به طور ساده از پشت سر هم قرار گرفتن تعدادی لایه‌ی کانولوشنی با کرنلی به ابعاد 3×3 تشکیل شده است. بین هر چند لایه کانولوشنی، یک لایه‌ی پولینگ^۱ قرار دارد. لایه‌ی پولینگ برای کاهش ابعاد شبکه و نیز ایجاد ناوردایی نسبت به برخی تغییرات در ورودی به کار می‌رود. انواع مختلف لایه‌ی پولینگ وجود دارد مانند پولینگ میانگین و پولینگ بزرگترین عدد. پولینگ میانگین از بین هر چند داده‌ی ورودی، میانگین آن‌ها را انتخاب می‌کند و پولینگ بیشترین عدد از بین چند داده‌ی ورودی، فقط بیشترین عدد را حفظ می‌کند. ساختار نمونه‌ی اولیه‌ی شبکه VGGNet در شکل ۲-۴ آمده است.



شکل ۲-۴، ساختار شبکه VGGNet

در نهایت، خروجی لایه‌های CNN و پولینگ، به سه لایه‌ی پشت سر هم تماماً متصل^۲ داده شده است تا خروجی تولید شود.

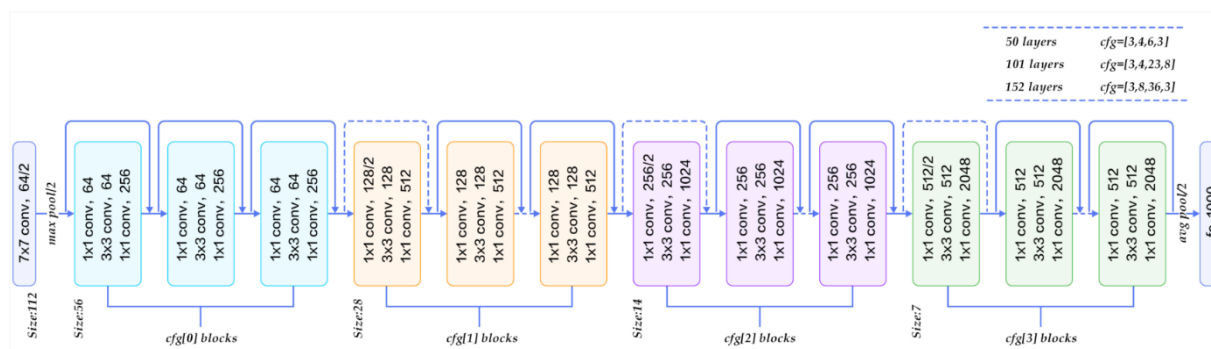
۲-۲-۲ ResNet

با افزایش هر چه بیشتر تعداد لایه‌های شبکه‌ی عصبی، آموزش شبکه با مشکلاتی روبرو می‌شود برای مثال در شبکه‌های با تعداد لایه‌های زیاد، گرادیان تغییرات برای لایه‌های ابتدایی صفر شده و وزن‌های این لایه‌ها تغییر نمی‌کند. شبکه‌ی ResNet برای حل این مشکل طراحی شده است. ایده اصلی ResNet معرفی

^۱ Pooling layer

^۲ Fully connected layer

ارتباطات میانبری همانی^۱ است به این معنی که ورودی برخی لایه‌ها، به طور مستقیم به ورودی لایه‌های بعدی اضافه می‌شود. با این روش امکان آموزش شبکه‌هایی با تعداد ۱۵۲ لایه و در عین حال با پیچیدگی کمتر از VGGNet فراهم شده است. ساختار این شبکه در شکل ۵-۲ آمده است.



شکل ۵-۲، ساختار شبکه ResNet

۲-۳- شبکه Faster RCNN

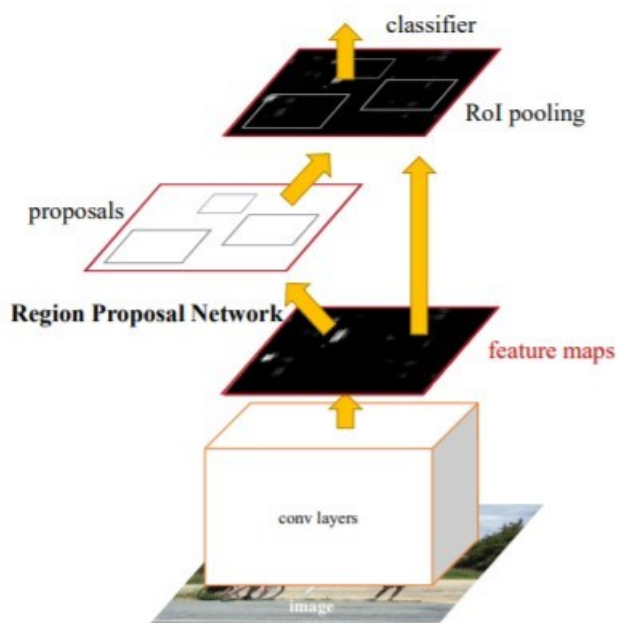
استفاده مستقیم از شبکه‌های یاد شده برای کشف اشیا در یک تصویر کامل مشکلاتی را به همراه دارد. برای مثال ممکن است در یک تصویر بیش از یک نمونه از شی مورد نظر وجود داشته باشد و یا چند شی مختلف در تصویر وجود داشته باشد. راه‌حلی که در ابتدا برای این مسئله پیشنهاد شد استخراج تعداد زیادی تصویر از عکس اولیه و دسته‌بندی جداگانه هر کدام از این بخش‌ها و در نهایت ترکیب نتایج بود. این روش به دلیل اینکه لازم است تعداد زیادی بخش از هر تصویر استخراج شود، محاسبات بسیار سنگینی نیاز دارد. Faster RCNN یکی از شبکه‌هایی است که برای حل این مشکل پیشنهاد شده است. Faster RCNN از دو بخش تشکیل شده است. بخش اول یک شبکه‌ی پیشنهاد ناحیه^۲ است و بخش دوم شامل یک شبکه‌ی تشخیص دهنده‌ی اشیا و یک شبکه‌ی محلی‌سازی^۳ (به معنی تعیین محل شیء کشف شده در تصویر) است. شبکه‌ی پیشنهاد ناحیه یک شبکه‌ی تماماً

^۱ Identity shortcut connections

^۲ Region proposal network

^۳ localization

کانوولوشنی است که تعدادی ناحیه^۱ که احتمالا شیئی در آن‌ها وجود دارد استخراج می‌کند و به بخش دوم می‌دهد. بخش دوم بر روی هر کدام از ناحیه‌های پیشنهادی اجرا می‌شود و شیء موجود در ناحیه را در صورت وجود تشخیص می‌دهد (شکل ۶-۲).



شکل ۶-۲، Faster RCNN [۱۰]

۴-۲- شبکه Mask RCNN

Mask RCNN، بر مبنای Faster RCNN طراحی شده است و با اضافه کردن بخشی به شبکه، دسته‌بندی پیکسل به پیکسل تصویر را نیز انجام می‌دهد.

¹ Region

۵-۲- شبکه YOLO

نام YOLO از حرف‌های اول عبارت شما فقط یکبار نگاه می‌کنید^۱ گرفته شده است. YOLO بر خلاف Mask RCNN فقط از یک بخش تشکیل شده است. ابتدا تصویر را به $s*s$ بخش تقسیم می‌کند و سپس در هر کدام از بخش‌ها، m محدوده انتخاب می‌کند و شیء موجود در هر کدام از این محدوده‌ها را تشخیص می‌دهد. این شبکه سرعت بسیار بالاتری از شبکه‌ی Faster RCNN دارد اما دقت آن کمتر است.

۶-۲- شبکه RetinaNet

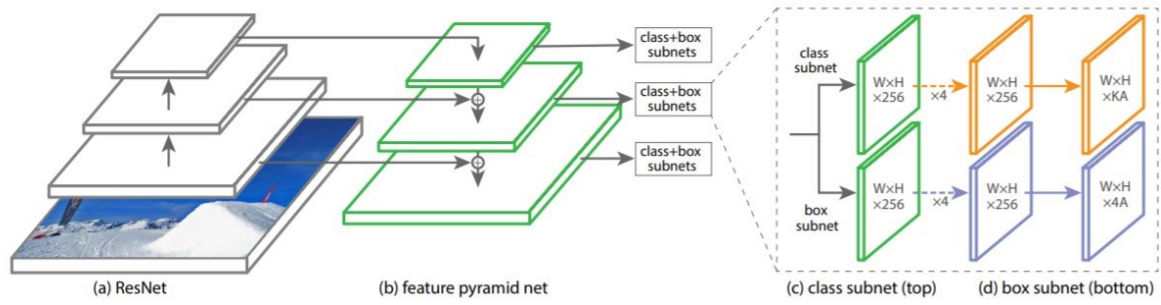
RetinaNet نیز یک شبکه تک‌بخشی است که طرح آن با تاکید بر اینکه علت اصلی دقت پایین شبکه‌های تک‌بخشی عدم تعادل در داده‌های کلاس‌ها^۲ هنگام آموزش شبکه است و با اصلاح تابع خطا^۳ به گونه‌ای که این مسئله را در نظر بگیرد ارائه شده است. در واقع در شبکه‌های دو بخشی، حدود ۲۰۰۰ ناحیه به قسمت دوم شبکه ارائه می‌شود اما در شبکه‌های تک بخشی، تعداد محدوده‌های مشخص شده به حدود ۱۰۰۰۰۰ ناحیه هم می‌رسد و روش‌هایی که در شبکه‌های دو بخشی برای مدیریت عدم تعادل کلاس استفاده می‌شود کارایی خود را تا حد زیادی از دست می‌دهد [۱۰]. در طراحی شبکه‌ی RetinaNet، از تابع خطای کراس آنترپی که به طور پویا مقیاس می‌شود^۴ استفاده شده است. در واقع با استفاده از یک ضریب مقیاس، اثر مثال‌های آسان کاهش داده می‌شود و شبکه روی مثال‌های سخت‌تر تمرکز می‌کند. ساختار کلی RetinaNet در شکل ۷-۲ آمده است.

¹ You Only Look Once

² Class imbalance

³ Loss function

⁴ Dynamically scaled cross entropy loss function



شکل ۷-۲، ساختار شبکه‌ی رتینانت

فصل ۳

فصل ۳: آموزش شبکه‌ها

فصل سوم در برگیرنده‌ی توضیح مربوط به آماده‌سازی دیتاست و نیز جزئیات آموزش شبکه‌ها است. در پایان فصل نتایج تست هر کدام از شبکه‌ها گزارش داده شده است.

۱-۳- مقدمه

در این فصل نخست به معرفی نحوه تقسیم‌بندی دیتاست اشاره می‌شود. سپس نحوه آموزش شبکه برای انجام هر کدام از تسک‌های تشخیص توده‌های سرطانی، تشخیص خوش خیم و یا بدخیم بودن توده‌ها و نیز دسته‌بندی پیکسل به پیکسل تصویر شرح داده می‌شود و در انتها رابط گرافیکی آماده شده معرفی می‌شود.

۲-۳- دیتاست

دیتاست مورد استفاده در این پروژه، CBIS-DDSM است. همچنین داده‌ها طبق جداسازی استاندارد ارائه شده در CBIS-DDSM به مجموعه‌های آموزش و تست تقسیم شده‌است. داده‌های تست شامل ۳۶۱ تصویر است و داده‌های آموزش شامل ۱۲۳۱ تصویر است. داده‌های تست خود به دو بخش اعتبارسنجی^۱ و تست تقسیم شده‌است تا داده‌های تست در روند آموزش نقشی نداشته باشد. به دلیل کوچک بودن دیتاست در مقایسه با حجم داده‌هایی که شبکه‌های عصبی برای عملکرد بهینه به آن نیاز دارد، از روش‌های مرسوم در این حوزه از جمله داده‌افزایی^۲ و یادگیری انتقالی^۳ برای افزایش دقت استفاده شده است. داده‌افزایی به معنی ایجاد برخی تغییرات در داده‌های موجود (به طوریکه این تغییرات تاثیری بر شناسایی اشیاء مورد نظر نداشته باشد) و نیز اضافه کردن این داده‌های تولیدی به دیتاست است. دو روش داده‌افزایی آنلاین و آفلاین برای اعمال داده‌افزایی وجود دارد. در داده‌افزایی آفلاین داده‌های اضافی قبل از آموزش شبکه تولید شده و به دیتاست افزوده می‌شود و حجم دیتاست چندین برابر می‌شود. در داده‌افزایی آنلاین در هنگام آموزش شبکه و در هر بار بارگیری^۴ تصویر تغییراتی روی آن اعمال می‌شود به طوریکه اندازه‌ی دیتاست ثابت می‌ماند اما در هر اپاک^۵ (یکبار آموزش شبکه

^۱ Validation

^۲ Data augmentation

^۳ Transfer learning

^۴ Load

^۵ Epoch

با استفاده از تمام داده‌ها) تصاویر ورودی شبکه تغییر می‌کند. در این پروژه از داده‌افزایی آنلاین استفاده شده است. تغییرات اعمال شده بر تصویر در این پروژه شامل وارونه کردن افقی و عمودی تصویر، مقیاس تصویر با ضریب تصادفی، برش تصادفی تصویر، و دوران تصادفی تصویر می‌باشد.

۳-۳- یادگیری انتقالی

یکی دیگر از روش‌هایی که برای غلبه بر مشکل کم بودن داده‌ها استفاده شده است، استفاده از یادگیری انتقالی است. یادگیری انتقالی به معنی ذخیره‌ی دانشی که شبکه در یک مسئله آموخته و اعمال این دانش به یک مسئله دیگر است. برای این منظور ابتدا شبکه‌ی مورد نظر بر روی دیتاست بزرگی آموزش داده می‌شود و سپس پارامترهای بدست آمده، به عنوان نقطه‌ی شروع برای مسئله‌ی جدید استفاده می‌شود. در این پروژه برای تمامی حالت‌ها، پارامترهای شبکه با وزن‌های بدست آمده از آموزش بر روی دیتاست ایمیجنت^۱ مقداردهی اولیه شده است. با وجود اینکه تصاویر دیتاست ایمیجنت با تصاویر استفاده شده در تشخیص سرطان بسیار متفاوت است، طبق [۱۲] استفاده از وزن‌های ایمیجنت در این مسئله مفید گزارش شده است.

۳-۴- کشف توده‌های سرطانی

۳-۴-۱- RetinaNet

برای آموزش شبکه RetinaNet، از پیاده سازی متن باز RetinaNet در کراس^۲ استفاده شده است^۳. کراس یک API سطح بالای شبکه‌های عصبی به زبان برنامه‌نویسی پایتون است که امکان پیاده‌سازی انواع شبکه‌های عصبی را به سادگی فراهم می‌کند. ابتدا رزولوشن داده‌های ورودی برای کاهش جافظی مورد نیاز و افزایش سرعت آموزش شبکه کاهش داده شده است به طوریکه با حفظ رزولوشن تصویر، بعد کوچکتر آن به ۸۰۰

^۱ Imagenet

^۲ Keras

^۳ <https://github.com/fizyr/keras-retinanet>

پیکسل تبدیل شده است. در آموزش شبکه تا جای ممکن از تصاویر با رزولوشن نسبتاً زیاد استفاده شده است تا دقت تشخیص افزایش یابد [۱۳]. همانطور که ذکر شد پارامترهای شبکه طبق آموزش روی دیتاست ایمیجنت مقداردهی اولیه شده است و البته لایه‌هایی که بین شبکه ایمیجنت و شبکه استفاده‌شده در اینجا تفاوت دارد (مانند لایه‌ی آخر که برای دسته‌بندی استفاده می‌شود، در اینجا فقط دو کلاس وجود دارد ولی در دیتاست ایمیجنت تعداد کلاس‌ها بسیار بیش‌تر است) با وزن‌های تصادفی مقداردهی شده است. با توجه به استفاده از یادگیری انتقالی، مقدار نرخ‌یادگیری^۱ اولیه مقدار پایین 10^{-5} انتخاب شده است. نرخ یادگیری در صورتی که دقت در ۵ اپیاک متوالی بهبود نیابد در ۰.۱ ضرب می‌شود. در نهایت پس از ۳۰ اپیاک آموزش شبکه، وزن‌هایی که بهترین عملکرد را بر روی دیتاست اعتبارسنجی داشت به عنوان پارامترهای نهایی انتخاب شد.

۲-۴-۳- YOLO

در آموزش شبکه YOLO از کتابخانه‌ی متن‌باز دارکنت^۲ استفاده شده است. دارکنت کتابخانه‌ای برای توسعه‌ی شبکه‌های عصبی است که به زبان C توسعه یافته. نسخه‌ی رسمی دارکنت از سیستم‌عامل ویندوز پشتیبانی نمی‌کند، در اینجا از نسخه‌ی تغییر یافته‌ی دارکنت استفاده شده است^۳. به دلیل محدودیت‌های حافظه رزولوشن تصاویر ورودی در این جا در بیش‌ترین حالت برابر 416×416 انتخاب شده است. به دلیل وجود توده‌های با سایزهای مختلف کوچک و بزرگ در دیتاست، از مدلی از YOLO استفاده شده است که همه‌ی این موارد را پشتیبانی کند. همچنین برای بهبود عملکرد شبکه، در هنگام آموزش شبکه از تصاویر با رزولوشن‌های مختلف استفاده شده است.

^۱ Learning rate

^۲ Darknet

^۳ <https://github.com/kriyeng/darknet>

۳-۴-۳ Mask RCNN

نسخه‌ی رسمی Mask RCNN در کتابخانه‌ی Caffe2 پیاده‌سازی شده است. در این پروژه به منظور سهولت استفاده، نسخه‌ی دیگری از این شبکه که با استفاده از کراس پیاده‌سازی شده است به کار گرفته شده است.^۱ این شبکه با تصاویر با رزولوشن 1024×1024 آموزش داده شده است. آموزش شبکه‌ی Mask RCNN در چند مرحله صورت گرفته است به طوریکه در مرحله‌ی اول تنها وزن‌های لایه‌های پایانی (که مقداردهی اولیه تصادفی داشته است) اصلاح شده و سپس در سه مرحله به تدریج تعداد لایه‌های آموزش داده‌شده افزایش یافت تا اینکه در نهایت تمامی لایه‌ها آموزش داده شد. نرخ یادگیری در اینجا از 10^{-3} شروع شد و در هر بار توقف کاهش خطا، به ۰.۱ مقدار کاهش یافت.

۳-۵-۳ تشخیص خوش خیم و یا بدخیم بودن توده‌ها

برای این مسئله دقیقاً مانند ۳-۴-۱ عمل کردیم و از شبکه‌ی رتینانت استفاده کردیم. تنها تفاوت نسبت به قسمت ۳-۴-۱ در تعداد کلاس‌های خروجی است و در این حالت سه کلاس داریم: توده خوش خیم، توده بدخیم و پس‌زمینه.

۳-۶-۳ دسته‌بندی پیکسل به پیکسل توده‌ها

۳-۶-۱-۳ استفاده از Mask RCNN

برای انجام دسته‌بندی پیکسل به پیکسل از مدل Mask RCNN مطابق ۳-۴-۳ استفاده شد زیرا همانطور که ذکر شده است این مدل مسک دسته‌بندی پیکس به پیکسل نیز ارائه می‌دهد.

^۱ https://github.com/matterport/Mask_RCNN

۷-۳- معیار ارزیابی^۱

برای معرفی معیارهای ارزیابی ابتدا چند تعریف را بیان کرده و سپس معیارهای ارزیابی را معرفی می‌کنیم.

۱-۷-۳- تعاریف اولیه

۱-۱-۷-۳- IoU^۲

IoU دو مجموعه برابر است با اندازه‌ی اشتراک آن‌ها تقسیم بر اندازه‌ی اجتماع آن‌ها. در این پروژه IoU برای بررسی میزان تطابق محدوده‌ی مستطیلی پیشبینی شده برای توده‌ی سرطانی (و یا مسک تولید شده برای تقسیم‌بندی تصویر) و محل حقیقی این توده استفاده می‌شود. در واقع در اینجا IoU برابر است:

$$IoU = \frac{\text{Area of intersection of predicted region and ground truth region}}{\text{Area of union of predicted region and ground truth region}} \quad \text{رابطه (۲)}$$

IoU مقداری بین صفر و یک دارد.

۲-۱-۷-۳- TP^۳

TP به معنای این است که شبکه به درستی وجود یک توده را در قسمتی از تصویر پیشبینی کرده و این پیشبینی بر محل حقیقی توده تا حدودی منطبق است.

۳-۱-۷-۳- FP^۴

FP به این معنی است که شبکه پیشبینی کرده است که در قسمتی از تصویر یک توده وجود دارد اما در واقع توده‌ای آنجا وجود نداشته باشد (و یا نوع توده اشتباه تشخیص داده شده باشد).

¹ Evaluation metric

² Intersection over Union

³ True Positive

⁴ False Positive

۳-۷-۱-۴ FN^۱

FN به معنی آن است که در قسمتی از تصویر یک توده وجود دارد اما شبکه نتوانسته وجود آن توده را پیشبینی کند.

۳-۷-۱-۵ IoU - آستانه‌ی IoU

اغلب ناحیه‌های پیشبینی شده همپوشانی کامل با ناحیه‌ی حقیقی ندارد به همین دلیل یک حد آستانه‌ی IoU تعیین می‌شود و در صورتی که IoU ناحیه‌ی پیشبینی شده با ناحیه‌ی حقیقی از این مقدار بیش‌تر باشد، پیشبینی TP در نظر گرفته می‌شود.

۳-۷-۱-۶ TPR^۲

TPR که حساسیت^۳ نیز نامیده می‌شود به معنای نسبت کل پیشبینی‌های صحیح به کل موارد نمونه‌های صحیحی است که در دیتاست وجود دارد:

$$TPR = \frac{TP}{P} = \frac{TP}{TP+FN} \quad \text{رابطه (۳)}$$

۳-۷-۱-۷ FPPI^۴

FPPI برابر است با نسبت کل پیشبینی‌های اشتباه به کل تصاویر دیتاست:

$$FPPI = \frac{FP}{\text{Number of images}} \quad \text{رابطه (۴)}$$

^۱ False Negative

^۲ True Positive Rate

^۳ Sensitivity

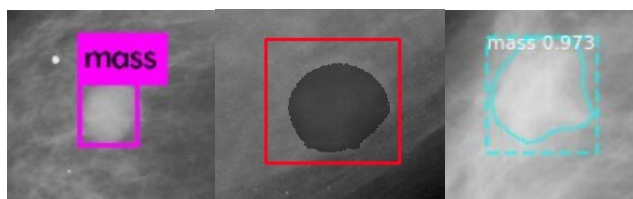
^۴ False Positive per Image

۲-۷-۳- معیارهای اصلی

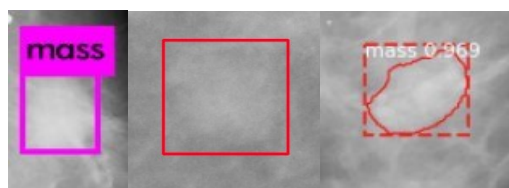
دو معیار اصلی ارزیابی با توجه به تعاریف ارائه شده عبارتند از نمودار TPR بر حسب آستانه‌ی IoU انتخابی و نیز نمودار TPR بر حسب FPPI (که به آن نمودار FROC^۱ هم گفته می‌شود). برای تولید نمودار اول، مدل آموزش داده شده شبکه روی دیتاست تست و به ازای آستانه‌ی IoU مختلف اجرا می‌شود و در هر حالت TPR بدست می‌آید. همچنین برای بدست آوردن نمودار FROC، در یک IoU ثابت با تغییر دادن آستانه‌ی اطمینان^۲ شبکه، نقاط TPR و FPR مختلف بدست می‌آید.

۸-۳- نتایج کشف توده‌ها

نمونه‌ای از تشخیص صحیح سه شبکه‌ی YOLO، رتینانت و Mask RCNN در شکل ۱-۳ و نمونه‌ای از تشخیص نادرست هر کدام از شبکه‌ها در شکل ۲-۳ آمده است:



شکل ۱-۳، نمونه‌ای از تشخیص‌های صحیح توده‌ها، Mask RCNN (راست)، رتینانت (وسط) و YOLO (چپ)



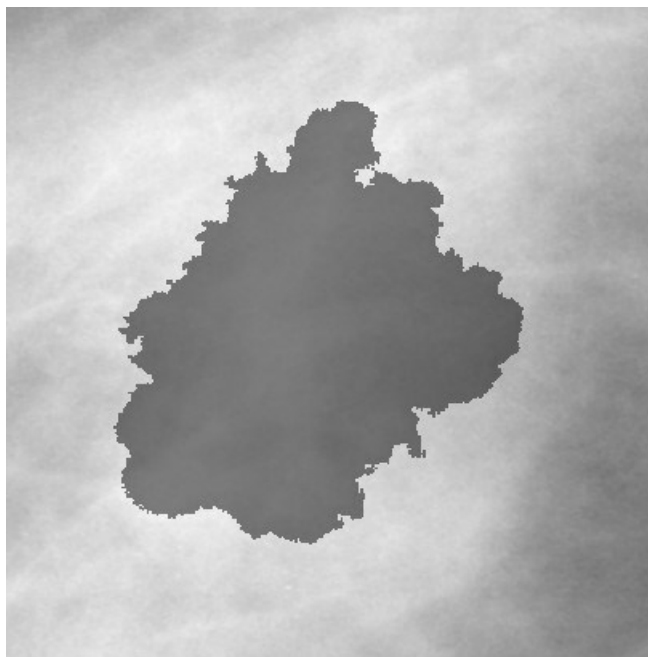
شکل ۲-۳، نمونه‌ای از تشخیص‌های اشتباه، Mask RCNN (راست)، رتینانت (وسط) و YOLO (چپ)

نکته‌ای که در مورد پیش‌بینی‌های اشتباه قابل مشاهده است، این است که بسیاری از پیش‌بینی‌ها غلط در بافت‌های متراکم است.

در شکل ۳-۳ نمونه‌ای از توده مشاهده می‌شود که توسط هیچ کدام از شبکه‌ها کشف نشده است.

^۱ Free Response Operating characteristic

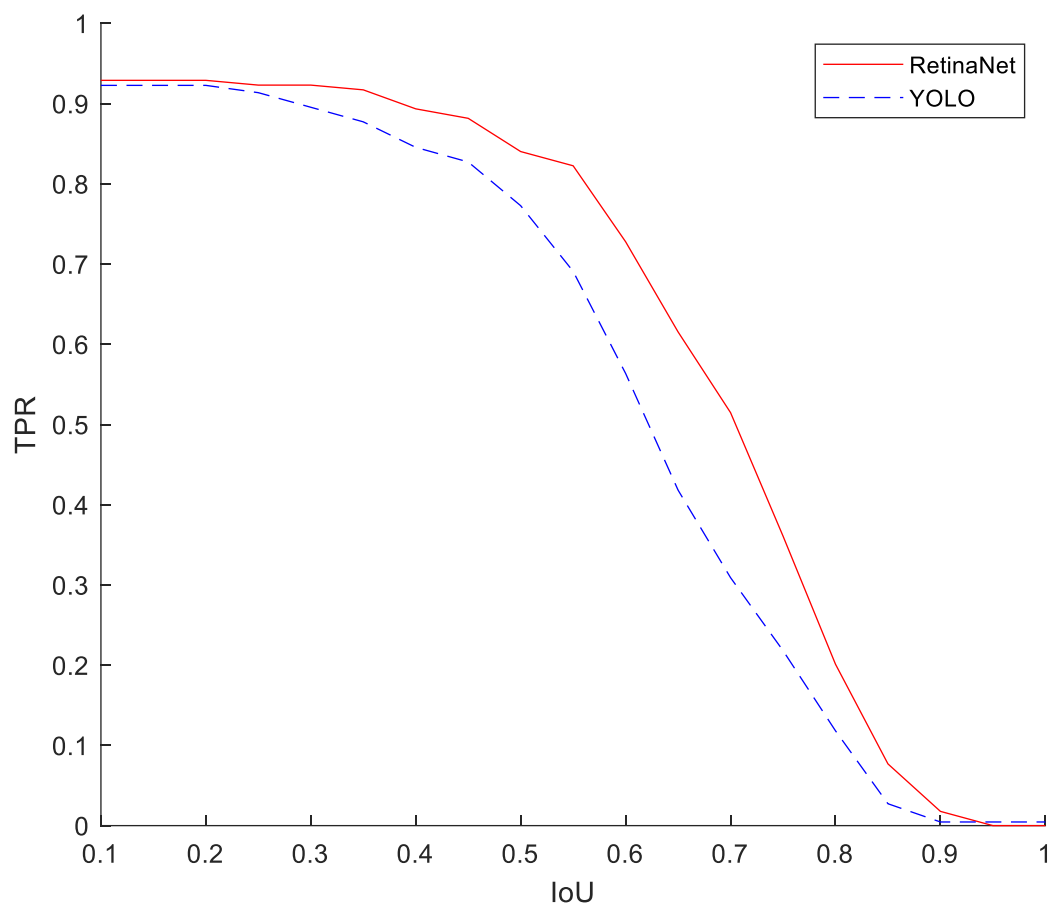
^۲ Confidence Threshold



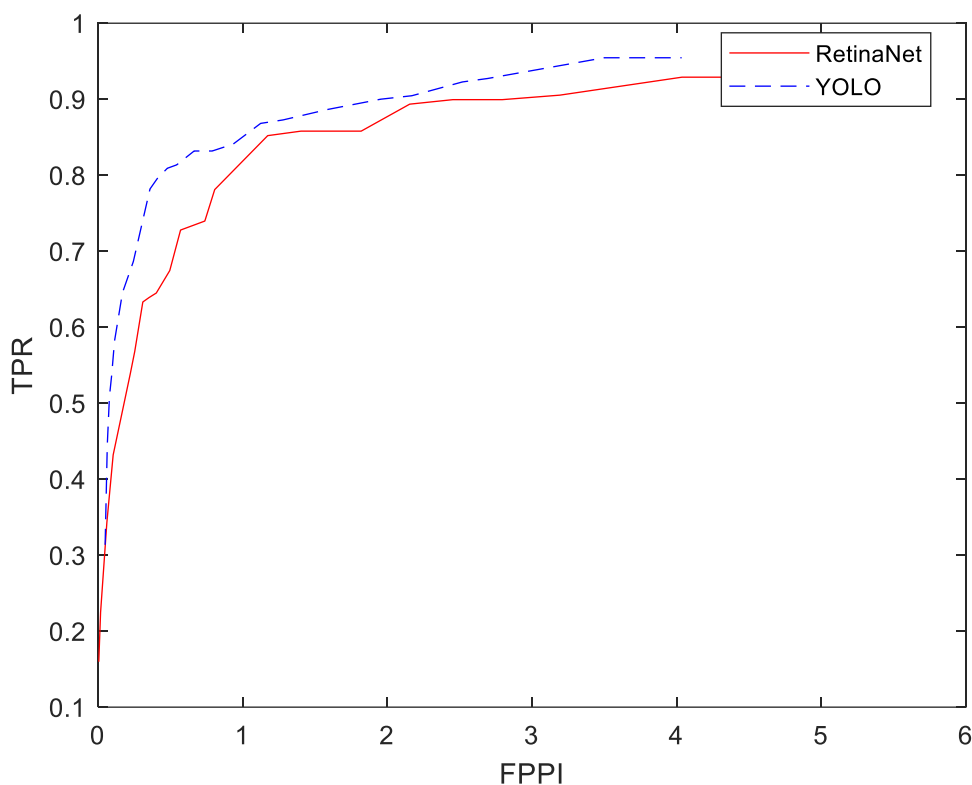
شکل ۳-۳، توده‌ای‌ای که توسط هیچ‌کدام از شبکه‌ها کشف نشده است.

نکته‌ی مهم در این تصویر باز هم تراکم بالای بافت است که تشخیص توده را دشوار کرده است.

نمودار TPR بر حسب IoU شبکه‌های YOLO و رتینانت در شکل ۳-۴ آمده است.



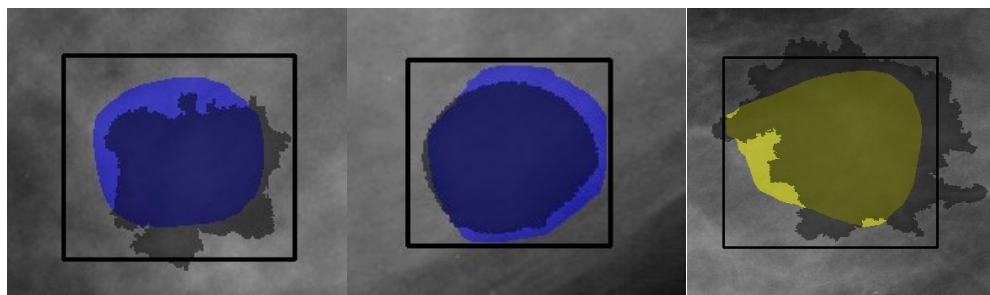
شکل ۳-۴، نمودار TPR بر حسب IoU
نمودار TPR بر حسب FPPI در شکل ۳-۵ آمده است.



شکل ۵-۳، نمودار FROC

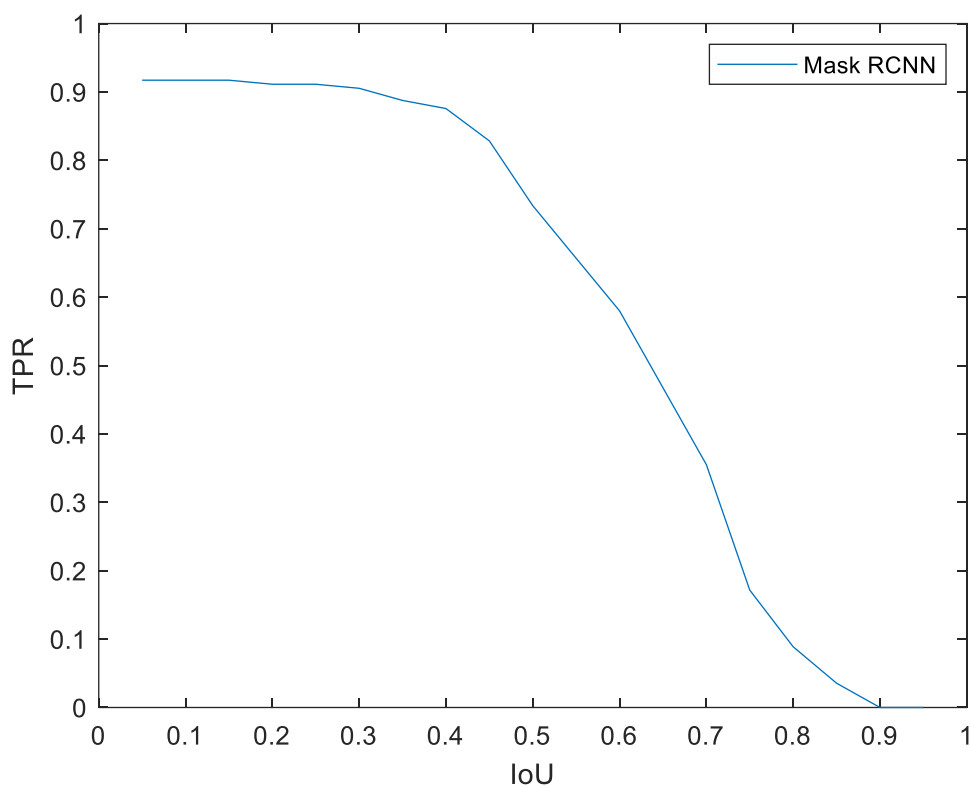
۹-۳- نتایج دسته‌بندی پیکسل به پیکسل توده‌ها

نتایج تعدادی از مسک‌های دسته‌بندی پیکسل به پیکسل صحیح که توسط شبکه‌ی Mask RCNN بدست آمده است در شکل ۶-۳ آمده است. همانطور که ذکر شد، Mask RCNN علاوه بر کشف توده‌ها، مسک نیز ارائه می‌دهد.

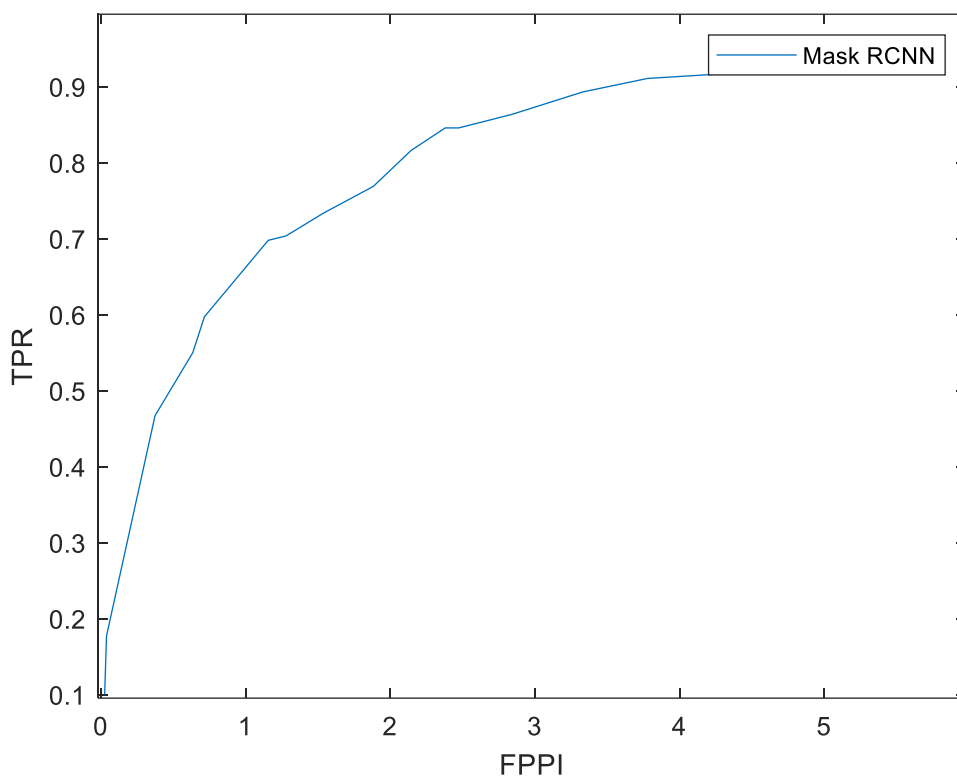


شکل ۶-۳، مسک ارائه شده توسط Mask RCNN، قسمت رنگی مسک پیش‌بینی شده و قسمت تیره مسک حقیقی است.

نمودار TPR بر حسب IoU در شکل ۳-۷ و نمودار TPR بر حسب FPPI در شکل ۳-۸ آمده است.



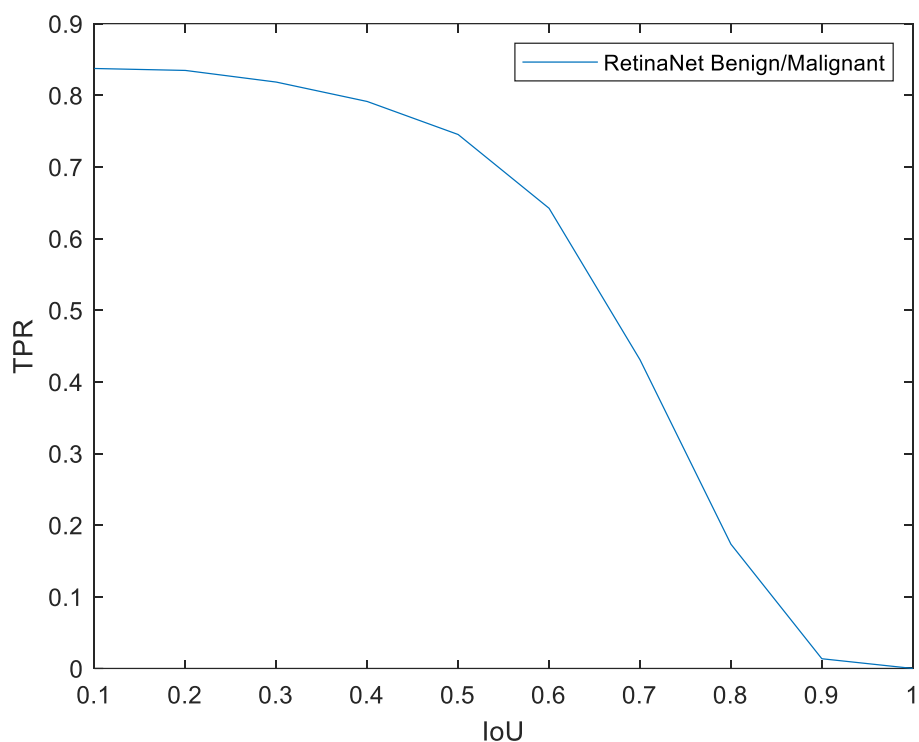
شکل ۳-۷، نمودار TPR بر حسب IoU در شبکه‌ی Mask RCNN



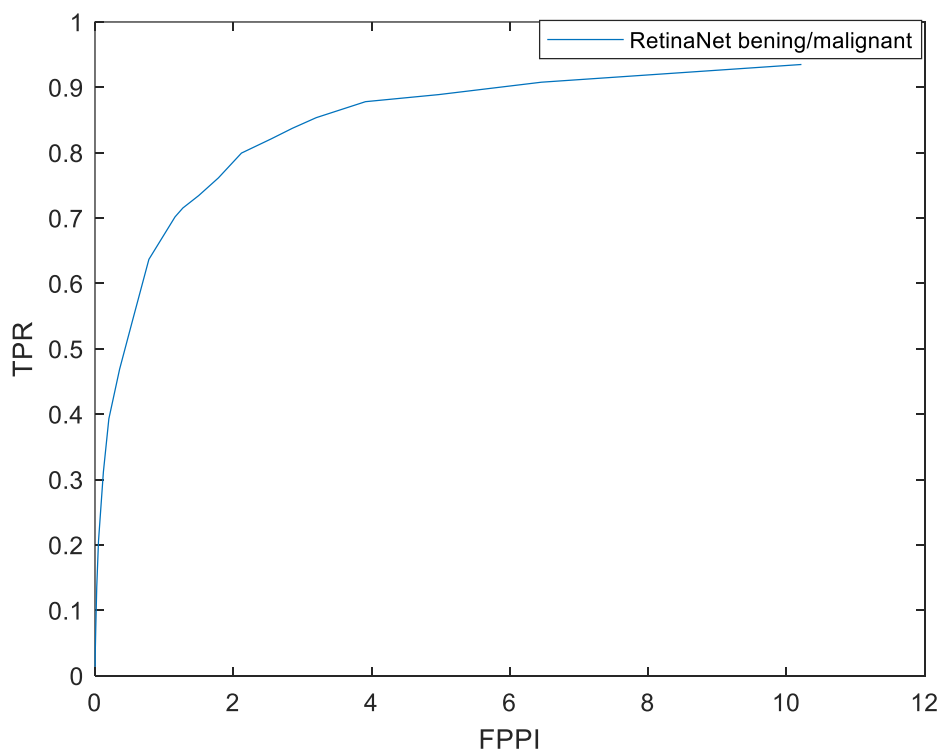
شکل ۸-۳، نمودار FROC شبکه‌ی Mask RCNN برای دسته‌بندی پیکس به پیکسل تصویر

۱۰-۳- نتایج دسته‌بندی توده‌های خوش‌خیم و بدخیم

نمودارهای TPR بر حسب IoU و نیز نمودار FROC برای تشخیص توده‌های خوش‌خیم و بدخیم به ترتیب در شکل ۹-۳ و ۱۰-۳ ارائه شده است.



شکل ۹-۳، نمودار TPR بر حسب IoU شبکه‌ی رتینانت با دسته‌بندی خوش‌خیم و بدخیم



شکل ۱۰-۳، نمودار FROC شبکه‌ی رتینانت با دسته‌بندی خوش‌خیم و بدخیم

۳-۱۱- تحلیل نتایج

با توجه به شکل ۳-۴، مشخص است که هر دوی شبکه‌های استفاده شده برای کشف توده‌ها عملکرد نسبتاً مناسبی دارند و شبکه‌ی رتینانت توانایی بیش‌تری در کشف توده‌ها دارد. البته با دقت در شکل ۳-۵ می‌توان متوجه این نکته شد که در FPPI برابر، شبکه‌ی YOLO دقت بیش‌تری دارد. در واقع شبکه‌ی YOLO می‌تواند با تعداد پیش‌بینی‌های غلط کمتری به نتایج مناسب برسد اما در صورتی که تعداد پیش‌بینی‌های غلط را در نظر بگیریم، در یک IoU ثابت رتینانت تعداد بیش‌تری توده را کشف می‌کند. از آنجایی که IoU در شبکه‌ی Mask RCNN بر اساس دسته‌بندی پیکسل به پیکسل محاسبه شده است (که نسبت به ناحیه‌ی مستطیلی شکل مسئله‌ی سخت‌تری است) دقت این مدل از دو مدل دیگر کمتر بدست آمده است و حساسیت این مدل در IoU های بالا سریع‌تر افت می‌کند. با توجه به شکل ۳-۱۰ و ۳-۱۱ مشخص است که تشخیص خوش‌خیم و بدخیم بودن توده‌ها با شبکه‌ی رتینانت با حساسیت ۸۳٪ در IoU ۰.۱ انجام شده است و همانطور که انتظار می‌رفت حساسیت نسبت به حالت قبل کاهش یافته است. تمامی نمودارهای FROC در IoU برابر ۰.۱ محاسبه شده است. علت انتخاب IoU برابر ۰.۱ این است که در مسئله‌ی تشخیص توده‌ها در مقایسه با تشخیص اشیا که مرزها نسبتاً متمایز است، مرزهای توده پیچیده‌تر است. لازم به ذکر است که در تشخیص اشیا، اغلب IoU برابر ۰.۵ معیار انتخاب می‌شود. با توجه به شکل ۳-۶ مشخص است که شبکه‌ی Mask RCNN در تشخیص مرزهای پیچیده و ناهموار توده‌ها تا حدی ناتوان است و علی‌رغم اینکه محدوده‌ی توده را با دقت مناسبی مشخص می‌کند، مرزهای توده را به درستی دنبال نمی‌کند.

۳-۱۲- خلاصه و جمع‌بندی

در این فصل ابتدا روش کار از جمله نحوه‌ی آماده‌سازی دیتاست و نیز نحوه آموزش شبکه شرح داده شد و سپس با معرفی معیارهای ارزیابی روش‌های پیشنهادی، نتایج بررسی این روش‌ها ارائه و تحلیل شد.

فصل ۴

فصل ۴: جمع بندی و نتیجه گیری

۴-۱- جمع‌بندی

در این تحقیق ما تعدادی از شبکه‌هایی را که در کشف اشیا در تصاویر موفق بوده انتخاب کرده و ضمن ایجاد تغییرات لازم، از آن‌ها برای تشخیص وجود سرطان در تصاویر ماموگرافی استفاده کردیم. روش ارائه شده به طور کامل خودکار تصویر کامل ماموگرافی را بررسی می‌کند و نواحی احتمالا سرطانی را شناسایی می‌کند. از شبکه‌های YOLO و رتینانت برای تشخیص توده‌های سرطانی و نیز تشخیص خوش‌خیم و یا بدخیم بودن آن‌ها و از شبکه‌ی Mask RCNN برای دسته‌بندی پیکسل به پیکسل تصویر ماموگرافی استفاده کردیم.

۴-۲- نتیجه‌گیری

نتایج بدست آمده نشان می‌دهد استفاده از شبکه‌های عصبی و عمیق و به طور خاص استفاده از شبکه‌های کانوولوشنی برای تشخیص توده‌های سرطانی در تصاویر ماموگرافی مفید است. شبکه‌های تست شده با حساسیت بیشتر از ۹۰٪ توانایی تشخیص توده‌ها در داده‌های تست را دارد. لازم به ذکر است که دیتاست CBIS-DDSM به دلیل پایین بودن کیفیت تصاویر از دیتاست‌های دشوار برای تشخیص سرطان به شمار می‌رود و اغلب دقت مدل‌ها بر روی دیتاست‌های دیگر (مانند INBreast) بیش‌تر است (البته پس از آن که مدل بر روی دیتاست دیگر آموزش داده شود) [۱۳]. اصلی‌ترین ضعف روش ارائه شده، بالا بودن تعداد پیش‌بینی‌های اشتباه مدل است (FP).

۴-۲-۱- محدودیت‌ها

محدودیت اصلی که در این پروژه (و اغلب مسائل پزشکی دیگر) با آن روبرو هستیم پایین بودن حجم داده‌های موجود است. شبکه‌های عمیق برای عملکرد بهینه به داده‌های بسیار زیادی نیاز دارد که در حال حاضر تهیه این حجم از داده‌ها در مسائل پزشکی از جمله مسئله تشخیص سرطان سینه بسیار هزینه‌بر و زمان‌بر است. به علاوه دیتاست‌های آزاد در دسترس در این زمینه‌ها نیز محدود است. در صورتی که داده‌های بیشتری فراهم شود کارایی شبکه‌های عمیق بهبود می‌یابد. محدودیت مهم دیگر در به کارگیری شبکه‌های عمیق سخت‌افزار

مورد نیاز برای آموزش شبکه‌ها است. با توجه بالا بودن تعداد پارامترهای شبکه‌های عصبی عمیق و محاسبات زیادی که باید انجام شود برای آموزش شبکه‌های عصبی به کارت گرافیک نیازمند هستیم. در این پروژه تمامی شبکه‌ها را در بستر فضای اشتراکی گوگل^۱ که کارت گرافیک رایگان به محققان ارائه می‌دهد آموزش دادیم. با توجه به محدودیت‌های این روش، آموزش شبکه‌ها تا حد بهینه انجام نشده است و در صورت استفاده از سخت‌افزار محلی، میتوان انتظار داشت با انجام آموزش شبکه‌ها به مدت طولانی‌تر عملکرد شبکه‌ها بهبود یابد.

¹ <https://colab.research.google.com>

فصل ٥

فصل ٥ : مراجع

- [1] Abachizadeh K, Moradi Kouchi A, Ghanbari Motlagh A, Kousha A, Shekarritz-Foumani R, Erfani "Breast Cancer in Iran: Levels, Variations and Correlates," *Community Health*, vol. 5, no. 1, pp. 11-21, 2018
- [2] Azam Hamidinekoo, Erika Denton, Andrik Rampun, Kate Honnor, Reyer Zwiggelaar, "Deep Learning in Mammography and Breast Histology, an Overview and Future Trends," *Medical Image Analysis*, 2018
- [3] I. Schreer, "Dense breast tissue as an important risk factor for breast cancer and implications for early detections, *Breast Care*, vol. 4, pp. 89-92, 2009
- [4] R.M. Rangayyan, N.M. El-Faramawy, J.L. Desautels, et al., "Measures of acutance and shape for classification of breast tumors", *IEEE Trans. Med. Imaging* vol. 16, no. 6, pp. 799–810, 1997
- [5] B. Verma, P. McLeod, A. Klevansky, "A novel soft cluster neural network for the classification of suspicious areas in digital mammograms", *Pattern. Recognit.* vol. 42, no. 9, pp. 1845–1852, 2009
- [6] S. Beura, B. Majhi, R. Dash, "Mammogram classification using two dimensional discrete wavelet transform and gray-level co-occurrence matrix for detection of breast cancer", *Neurocomputing*, vol. 154, pp. 1-14, 2015
- [7] N. Dhungel, G. Carneiro, A. P. Bradley, "A deep learning approach for the analysis of masses in mammograms with minimal user intervention," *Medical Image Analysis*, vol. 37, pp. 114-128, 2017
- [8] MonkeyLearn, "Text Classification, A comprehensive guide to classifying text with machine learning," <https://monkeylearn.com/text-classification>, Aug. 31
- [9] I. Goodfellow, Y. Bengio, A. Courville (2016), "Deep learning book," *MIT Press*, pp. 334-337
- [10] S. Ren, K. He, R. Girshick, J. Sun, "Faster R-CNN: Towards Real-Time ObjectDetection with Region Proposal Networks," *Arxiv*, arXiv:1506.01497 [cs.CV], 2015
- [11] T. Lin, P. Goyal, R. Girshick, K. He, P. Dollár, "Focal Loss for Dense Object Detection," *Arxiv*, arXiv:1708.02002 [cs.CV], 2017
- [12] G. Carneiro, J. Nascimento, and A. P. Bradley, "Unregistered multiview mammogram analysis with pre-trained deep learning models," *Lect. Notes Comput. Sci.* no. 9351, pp. 652–660, 2015

- [13] D. Ribli, A. Horváth, Z. Unger, P. Pollner, I. Csabai, Detecting and classifying lesions in mammograms with Deep Learning, *Scientific Reports*, vol. 8, no. 4165, pp. 3, 2018