



دانشگاه صنعتی شریف
دانشکده مهندسی کامپیوتر

پایان نامه کارشناسی
مهندسی کامپیوتر

بهبود قطعه بندی جراحات در تصاویر ماموگرافی سینه از طریق روش های نیمه نظارتی

نگارش

امیرحسین باقری

استاد راهنما

حمیدرضا ربیعی

خرداد ۱۴۰۳

به نام خدا
دانشگاه صنعتی شریف
دانشکده مهندسی کامپیوتر

پایان نامه کارشناسی

این پایان نامه به عنوان تحقق بخشی از شرایط دریافت درجه کارشناسی است.

عنوان: بهبود قطعه بندی جراحات در تصاویر ماموگرافی سینه از طریق روش های نیمه نظارتی

نگارش: امیرحسین باقری

چکیده

شبکه‌های عصبی عمیق پیشرفت‌های قابل توجهی در زمینه‌های مختلف داشته‌اند، از مهمترین و تاثیرگذارترین پیشرفت‌ها، حوزه پزشکی است. این شبکه‌ها به دلیل سرعت بالا، مقاومت در خطا، و همچنین توانایی خدمت دهی به طور مداوم در شبانه‌روز بدون خستگی یا قطعی به ابزارهای ارزشمندی در دستیاری پزشکان تبدیل شده‌اند. یکی از چالش‌برانگیزترین مسائل در حوزه علوم پزشکی، تفسیر تصاویر ماموگرافی است. دو بعدی بودن تصاویر، تنوع بالای بافت سینه، تشابه بافت جراحات با برخی از بافت‌های عادی سینه و همچنین دشواری تشخیص بافت جراحات از بافت‌های اطراف از جمله چالش‌های پیش‌رو در این امر هستند. آموزش مدل‌های هوش مصنوعی برای این کار نیازمند داده‌های فراوان و متنوع به همراه جزییات تفسیرشان است که امری بسیار هزینه‌بر و زمان‌بر است. برای این مهم نیازمند متخصصین زیادی هستیم تا تفاسیر تصاویر ماموگرافی را برای آموزش مدل‌ها تهیه کنند که با توجه به تعداد داده‌ها امری تقریباً محال است. در این پروژه بررسی می‌کنیم که آیا ادغام اطلاعات از مخازن داده بزرگ بدون تفسیر و مخازن داده کوچک با تفسیر می‌تواند در عملکرد مدل‌ها تاثیر مثبتی بگذارد. همچنین یک خط لوله جدید برای تفسیر داده‌های بدون تفسیر بصورت اتوماتیک و استفاده از این داده‌ها و روش‌های مختلف یادگیری نیمه نظارتی برای اهرم کردن داده‌های حجیم بدون تفسیر برای انجام کار تشخیص نواحی جراحات در عکس‌های ماموگرافی است. در نهایت نشان می‌دهیم با ترکیب تفاسیر بدست آمده از خط لوله و یک شیوه نوین در استفاده از آنها به صورت پیکسل-آگاه به بهبود تقریباً ۴ درصدی نسبت به خط پایه منجر می‌شود. در این پروژه همچنین شیوه انتطابق مدل‌های مطرح در مطالعات پیشین به شیوه نوین ما نیز بررسی می‌شوند.

کلیدواژه‌ها: یادگیری نیمه نظارتی، پردازش تصاویر پزشکی، پردازش تصاویر ماموگرافی، قطعه بندی تصاویر، قطعه بندی تصاویر پزشکی

فهرست مطالب

۱	مقدمه	۱
۲	۱-۱ مفاهیم اولیه	۲
۲	۱-۱-۱ ماموگرم	۲
۶	۱-۱-۲ مفاهیم یادگیری ماشین	۶
۸	۱-۱-۳ قطعه بندی	۸
۹	۲-۱ تعریف مساله	۹
۹	۳-۱ اهداف پژوهش	۹
۱۰	۴-۱ اهمیت موضوع	۱۰
۱۰	۵-۱ چالش ها	۱۰
۱۱	۶-۱ ساختار پایان نامه	۱۱
۱۳	۲ کارهای پیشین	۱۳
۱۳	۱-۲ روش های نیمه نظارتی	۱۳
۱۴	۲-۲ روش های نیمه نظارتی با شبه برچسب	۱۴
۱۴	۳-۲ روش های نیمه نظارتی مبتنی بر تابع سازگار ساز	۱۴
۱۶	۴-۲ روش های نیمه نظارتی ترکیبی	۱۶
۱۷	۵-۲ بررسی شبکه ها و توابع مورد استفاده این پژوهش	۱۷
۱۷	۲-۵-۱ توابع هزینه	۱۷

۱۹	۲-۵-۲ شبکه های پایه
۲۲	۶-۲ جمع بندی
۲۴	۳ روش پیشنهادی
۲۴	۱-۳ قاعده سازی مساله
۲۵	۲-۳ پیش پردازش داده
۲۸	۳-۳ یادگیری نیمه نظارتی
۲۹	۴-۳ جمع بندی
۳۱	۴ نتایج تجربی
۳۱	۱-۴ دیتاست و معیارهای ارزیابی
۳۲	۲-۴ جزئیات پیاده سازی و نتایج شبیه سازی
۳۲	۱-۲-۴ پیدا کردن بهترین تابع هزینه
۳۲	۲-۲-۴ پیدا کردن بهترین شبکه پایه
۳۳	۳-۲-۴ انطباق روش های نیمه نظارتی
۳۴	۳-۴ مقایسه با آموزش ساده
۳۴	۴-۴ جمع بندی
۳۶	۵ جمع بندی و کارهای آتی
۳۶	۱-۵ جمع بندی
۳۶	۲-۵ کارهای آتی
۳۸	مراجع
۴۳	واژه نامه

فهرست جدول‌ها

۱۷	۱-۲ مقایسه روش های پیشین در یک نگاه
۳۲	۱-۴ عملکرد توابع هزینه
۳۳	۲-۴ عملکرد معماری های مختلف
۳۳	۳-۴ عملکرد روش های مختلف یادگیری نیمه نظارتی
۳۴	۴-۴ مقایسه عملکرد یادگیری انتقالی

فهرست شکل‌ها

۳	۲-۱ یک نمونه از تصویر FFDMM
۳	۱-۱ یک نمونه از تصویر SFM
۴	۴-۱ یک نمونه از تصویر CC
۴	۳-۱ یک نمونه از تصویر MLO
۵	۵-۱ بافت‌ها سینه
۶	۷-۱ نمونه‌ای از یک گزارش پزشک
۶	۶-۱ خروجی مدل در نمای سیستم
۷	۹-۱ ویژگی‌های استخراج شده
۷	۸-۱ یک شبکه عصبی
۹	۱۰-۱ قطعه‌بندی معنایی و موردی
۱۹	۱-۲ تصویر بلاک‌ها ResUNet
۲۰	۲-۲ نمونه‌ای از کانولوشن آتروس
۲۱	۳-۲ معماری بلوک ConvUnext
۲۲	۴-۲ معماری TransUnet
۲۶	۱-۳ مدل معماری Unet
۲۷	۲-۳ معماری مدل اجماعی
۲۸	۳-۳ خط لوله

فصل ۱

مقدمه

شبکه‌های عصبی عمیق در سال‌های اخیر در بسیاری از حوزه‌ها به عملکرد بالایی دست یافته‌اند. در نتیجه، استفاده از آنها به حوزه‌های حیاتی مانند پزشکی گسترش یافته است. برخلاف انسان‌ها، ماشین‌ها در محاسبات سریع هستند، در تصمیم‌گیری‌های خود پایدار هستند و تحت تأثیر حالت روانی یا خستگی قرار نمی‌گیرند. این ویژگی‌ها، ماشین‌ها را برای کمک به پزشکان مناسب می‌کند. آنها می‌توانند زمان تشخیص را کاهش دهند، اندازه‌گیری‌های دقیق را آسان کنند، بیماران را اولویت‌بندی کنند، موارد اضطراری را جدا کنند و از خطاهای پزشکی بکاهند. تفسیر ماموگرام‌ها یکی از چالش‌برانگیزترین حوزه‌ها در علوم پزشکی است. این کار نیاز به شناسایی همه جراحات در سینه و تعیین خطرات آنها است. ماموگرام‌ها تصاویر دوبعدی از یک ساختار سه‌بعدی هستند که تداخل آنها در فضا ممکن است باعث سردرگمی در یافتن جراحات شود. چالش دیگر، تنوع گسترده بافت طبیعی سینه است. چیزی که در یک سینه به عنوان جراحات در نظر گرفته می‌شود، ممکن است در سینه دیگر به عنوان بافت طبیعی تلقی شود. علاوه بر این، تشخیص برخی از جراحات از بافت طبیعی اطراف بسیار دشوار است. در این حوزه، یک مدل هوش مصنوعی برای قطعه‌بندی جراحات در ماموگرام‌ها می‌تواند به رادیولوژیست‌های تازه‌کار کمک کند و از خطاهایشان بکاهد.

با وجود چند مجموعه داده ماموگرام با برچسب، مجموعه داده‌های خصوصی بیمارستان‌ها با تعدد زیاد تصاویر فرصت خوبی برای توسعه مدل‌های هوش مصنوعی به ارمغان می‌آورد. اگرچه تصاویر بیمارستان‌ها دقیقاً برچسب‌گذاری نشده‌اند، اما برای هر تصویر گزارش تخصصی وجود دارد که نوع، تعداد و مکان تقریبی جراحات را در خود جای داده است. با این حال، در سال‌های اخیر بسیاری از مطالعات بر آموزش مدل‌های هوش مصنوعی با استفاده از یک مجموعه داده کوچک و دقیق برچسب‌گذاری شده همراه با یک مجموعه داده بزرگ و یا بدون برچسب یا نادرست برچسب‌گذاری شده برای دستیابی به عملکرد و تعمیم‌پذیری با هزینه کم متمرکز شده‌اند.

در ادامه به صورت مختصری بر مفاهیم اولیه استفاده شده در این پژوهش می‌پردازیم و ارتباط آنها به یکدیگر را بسط می‌دهیم.

۱-۱ مفاهیم اولیه

۱-۱-۱ ماموگرم

در این بخش به شرح و توضیح تعدادی از اصطلاحات و مفاهیم مورد استفاده در این پژوهش می‌پردازیم.

ماموگرافی

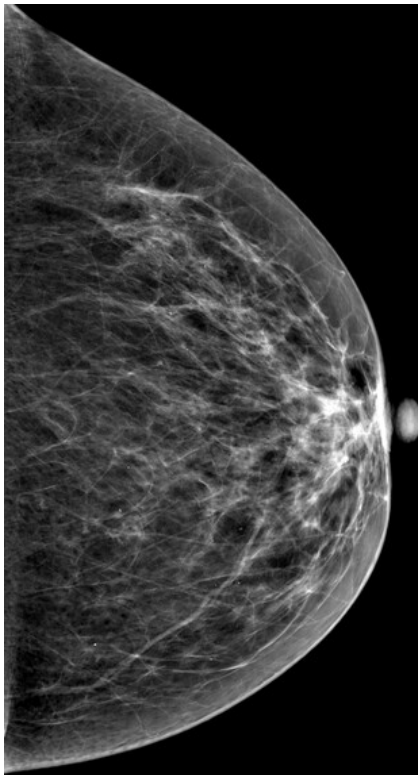
ماموگرافی یک روش تصویربرداری پزشکی است که از اشعه ایکس برای تهیه تصاویر از بافت‌های داخلی سینه استفاده می‌کند. این تصاویر، که به عنوان تصاویر ماموگرافی شناخته می‌شوند، برای تشخیص و ارزیابی بیماری‌های سینه مانند سرطان سینه، کیست‌ها، توده‌ها و دیگر ناهنجاری‌های بافت سینه به کار می‌روند.

ماموگرم

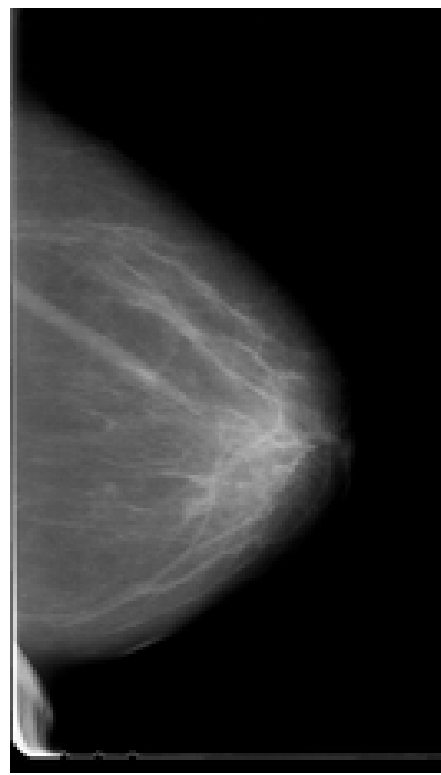
تصویر بدست آمده از فرایند ماموگرافی، ماموگرم نامیده می‌شود. تصاویر ماموگرافی دیجیتال کامل (FFDM)^۱ از تکنولوژی دیجیتال برای ثبت و ذخیره تصاویر استفاده می‌کنند. این نوع ماموگرافی از سنسورهای دیجیتال به جای فیلم‌های سنتی استفاده می‌کند، که به دریافت تصاویر با کیفیت بالاتر و وضوح بیشتر منجر می‌شود. تصاویر دیجیتال به راحتی قابل بزرگ‌نمایی و بررسی دقیق‌تر هستند و می‌توانند به سرعت به کامپیوترها منتقل و ذخیره شوند. این امر به پزشکان اجازه می‌دهد تا تغییرات کوچک و ناهنجاری‌های ریز را با دقت بیشتری شناسایی کنند. نمونه این عکس در تصویر ۱-۲ قابل مشاهده است. در مقابل، ماموگرافی فیلم-صفحه (SFM)^۲ از فیلم‌های حساس به نور برای ثبت تصاویر استفاده می‌کند. این روش سنتی‌تر است و کیفیت تصاویر ممکن است به اندازه تصاویر دیجیتال نباشد. در این نوع ماموگرافی، تصاویر بر روی فیلم‌های فیزیکی ذخیره می‌شوند و این می‌تواند به مشکلاتی نظیر سختی در ذخیره‌سازی و انتقال تصاویر منجر شود. همچنین، فیلم‌ها نیاز به توسعه و پردازش دارند که می‌تواند زمان‌بر باشد. نمونه این عکس در تصویر ۱-۱ قابل مشاهده است.

نما

-
1. Full-Field Digital Mammography
 2. Screen Film Mammography



شکل ۱-۲: یک نمونه از تصویر FFD [۱]



شکل ۱-۱: یک نمونه از تصویر SFM [۱]

به طور کلی، دو نما برای ثبت جزئیات بیشتر از سینه وجود دارد: کرانیوکودال (CC) ^۳ و مدیولاترال اُبلیک (MLO) ^۴. تصاویر CC از نمای بالا و تصاویر MLO از نمای چپ سینه گرفته می‌شوند. در واقع، این دو نما تقریباً عمود بر هم هستند. استفاده از این دو نما باعث می‌شود که بخش‌های مختلف سینه با وضوح بیشتری دیده شوند و احتمال تشخیص زودهنگام افزایش یابد. در تصاویر ۱-۳ و ۱-۴ می‌توانید این دو نما را ببینید.

بافت سینه

در ویرایش ۲۰۱۳ بی‌آی-رَدز (BI-RADS)، تقسیم‌بندی ترکیب سینه به دسته‌های a, b, c, d تغییر کرده است و هر کدام با توصیفی همراه است:

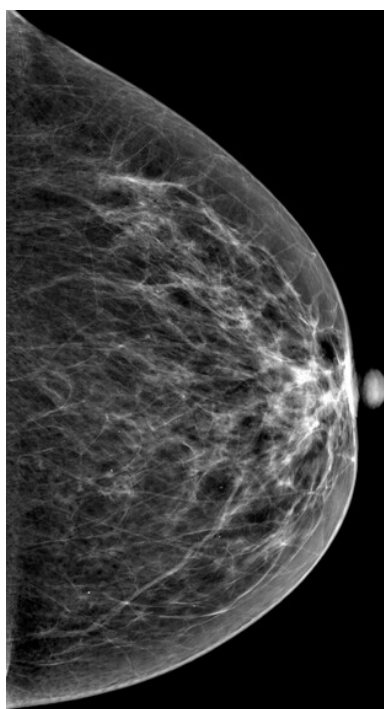
a تقریباً به‌طور کامل چرب هستند. حساسیت ماموگرافی در این حالت بسیار بالا است.

b نواحی پراکنده‌ای از تراکم فیبروگده‌ای وجود دارد. اصطلاح تراکم درجه تضعیف اشعه ایکس توسط بافت سینه را توصیف می‌کند، اما به معنای بافت‌های ماموگرافی مجزا نیست.

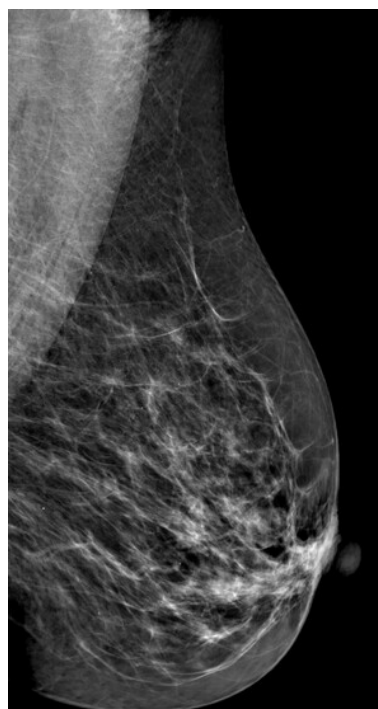
c سینه به‌طور ناهمگونی متراکم است که ممکن است توده‌های کوچک را پنهان کند. برخی نواحی در

3. Craniocaudal

4. Mediolateral Oblique



شکل ۱-۴: یک نمونه از تصویر CC [۱]



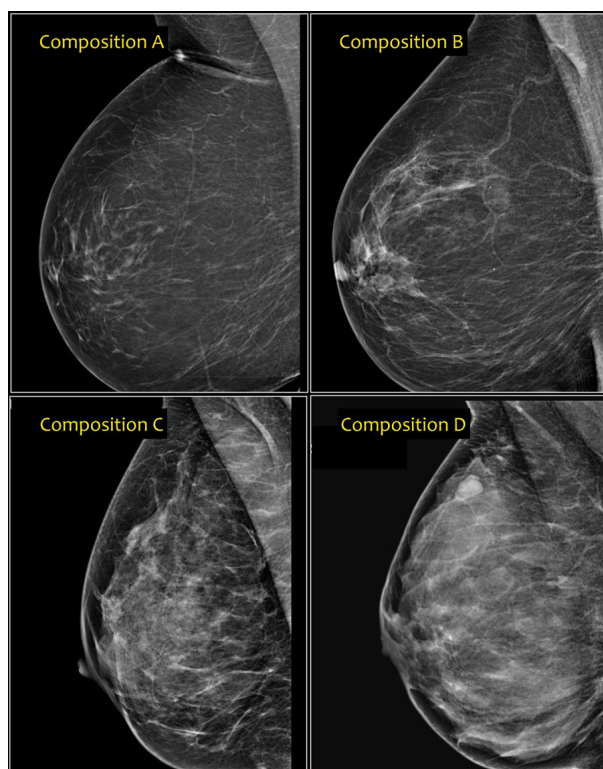
شکل ۱-۳: یک نمونه از تصویر MLO [۱]

سینه‌ها به اندازه کافی متراکم هستند که توده‌های کوچک را پنهان کنند.
d سینه‌ها بسیار متراکم هستند که حساسیت ماموگرافی را کاهش می‌دهد.

در تصویر ۱-۵ ماموگرافی از این تنوع بافتی را مشاهده می‌کنید.

جراحی: در این پژوهش عمدتاً توده‌های جراحات همان توده‌های سرطانی هستند که درون بافت سینه قرار دارند. هرچند برخی از این جراحات ممکن است کیست‌ها یا آماس‌ها و غیره باشند، ولی در این پژوهش منظور ما از جراحات غدد سرطانی هستند. بافت‌های متفاوتی از سرطان وجود دارند که شرح و تفصیل آنها از حوصله این پژوهش خارج است. در اینجا به طول خلاصه به اصطلاحات و مفاهیم مورد استفاده در این زمینه می‌پردازیم. ابتدا برای توصیف غدد انواع مختلفی از BI-RAD وجود دارند که هر کدام معنای خاص خود را دارا می‌باشند:

- BI-RADS 0: ارزیابی ناقص، نیاز به تصویر برداری بیشتر یا اطلاعات قدیمی‌تر برای تصمیم‌گیری نهایی.
- BI-RADS 1: منفی، هیچ‌گونه ناهنجاری مشاهده نشده است.
- BI-RADS 2: یافته‌های خوش‌خیم، ناهنجاری‌های خوش‌خیم مشاهده شده است که نیاز به پیگیری خاصی ندارند.



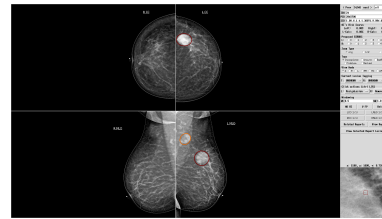
شکل ۱-۵: در شکل بالا یک ماموگرم از بافت های مختلف سینه را مشاهده می کنید [۱]

- BI-RADS 3: احتمالاً خوش خیم، ناهنجاری هایی که با احتمال بالا خوش خیم هستند، اما نیاز به پیگیری کوتاه مدت دارند (معمولاً شش ماهه).
- BI-RADS 4: ناهنجاری مشکوک، ناهنجاری هایی که نیاز به ^۵ بافت برداری دارند.
- BI-RADS 5: احتمالاً بدخیم، ناهنجاری هایی که احتمال بالایی بدخیمی دارند (بیش از ۹۵٪) و نیاز به اقدام فوری دارند.
- BI-RADS 6: سرطان اثبات شده، ناهنجاری هایی که به عنوان سرطان تأیید شده اند.

بافت های غدد جراحات نیز انواع مختلفی به همراه شکل های مختلفی دارند که برای پزشکان دانستن آنها ضروریست. اما در این پژوهش، تنها به این نکته اشاره می کنیم که از بین بافت های موجود بافت کلسیم دار از بافت های مورد بررسی ما نیست. برای مطالعات بیشتر برای انواع بافت سینه و غدد سینه و شکل آنها و نوع آنها می توانید به [۱] مراجعه کنید.

گزارش: مجموعه ای از متون نوشته شده توسط دکترها برای هر ماموگرام که نظر دکتر را بیان می کند. دقت کنید که درون گزارش ممکن است ویژگی های بسیاری از بیمار مانند سوابق بیمار و وزن و قد و دیگر صفت

A 60 years old patient has referred.
This is baseline screening mammogram for a 60 years old lady with positive family history of breast cancer in her first relative. She complains of palpable mass of left breast. Scattered areas of fibroglandular density are seen in both breasts (Breast composition: type B) . A large ill defined mass in UOQ of left breast are seen . There are also multiple suspicious lymph node in left axillary . These findings are highly suggestive of malignant mass with lymph node metastatic . In right breast a well defined mass in lateral part of right breast is seen which is correlative with a cyst in sonography .
Conclusion:
LEFT BREAST :BIRADS V : (highly suspicious for malignancy; biopsy is recommended).
RIGHT BREAST: BIRADS II : (benign findings)



شکل ۱-۶: این عکس خروجی مدل را در یک سیستم تحلیل نشان می‌دهد.

شکل ۱-۷: یک نمونه از گزارش دکتر

های بیمار قرار گرفته باشد. اما آنچه برای ما حائز اهمیت است نوع غده تشخیص داده شده و همچنین تعداد و مکان تقریبی غده است که دکتر می‌نویسد. دقت کنید که این با برچسب متفاوت است زیرا پزشک بر روی عکس محل غده را مشخص نمی‌کند بلکه تنها بر روی کاغذ به طور تقریبی پزشک بیان می‌کند که جراحی به کدام یک از چهار منطقه اصلی (بیرونی/داخلی بالا/پایین) سینه تعلق دارد. در تصویر ۱-۷ می‌توانید یک نمونه از گزارش داخل دیتا را به همراه مکانی که مدل ما به عنوان غده گزارش داده در ۱-۶ مشاهده کنید.

۱-۱-۲ مفاهیم یادگیری ماشین

داده: داده یا همان Data مهمترین رکن یادگیری ماشین است. داده‌ها می‌توانند به اشکال گوناگون مانند صوت و عکس و ویدئو و یا داده‌های جدولی باشند. البته در نظر بگیرید که در این پژوهش داده‌های مورد بررسی عکس‌های دوبعدی است.

برچسب: برچسب یا Label/Mask بخشی از داده است که بر داده‌های خام اضافه می‌شوند. این برچسب‌ها می‌توانند شامل دسته‌بندی داده‌ها یا قطعه‌بندی آنها و یا انواع و اقسام موارد مرتبط با وظایف گوناگون مدل ترتیب داده شده باشند.

آموزش: آموزش یا همان Training فرایندی است که طی آن یک مدل مبتنی بر هوش مصنوعی با الگوریتم‌های متفاوت بهینه‌سازی آموزش می‌یابد.

یادگیری نظارت شده: یادگیری نظارت شده یا همان Supervised Learning به نوعی از یادگیری اطلاق می‌شود که تمامی داده‌های مورد استفاده در زمان آموزش برچسب داشته باشند.

یادگیری نیمه نظارتی: یادگیری نظارت شده یا همان Semi-Supervised Learning به نوعی از یادگیری اطلاق می‌شود که بخشی از داده‌های برچسب داشته و بخشی دیگر بدون برچسب باشند.

یادگیری بدون نظارت: یادگیری بدون نظارت که بعضاً ممکن است به آن یادگیری خودآگاه اطلاق شود



شکل ۱-۸: یک شبکه عصبی که به همراه وزن های شبکه نشان داده شده اند [۲].
شکل ۱-۹: در این تصویر ویژگی ها استخراج شده از یک شبکه عصبی برای تشخیص دسته تصویر را نشان می دهد [۳].

نوعی از یادگیری است که در آن تمامی داده ها در طول آموزش بدون برچسب هستند. به این یادگیری (Self-Supervised / Unsupervised Learning) نیز گفته می شود.

توابع هزینه: توابع هزینه (خطا/زیان) یا Loss function ها یکی از اجزای مهم فرایند آموزش هستند. در حقیقت این توابع می توانند بیانگر میزان تفاوت خروجی مد نظر ما با خروجی مدل باشند که بصورت های مختلفی قابل استفاده هستند. این توابع می توانند برای تنظیم سازی (Regularization Loss) یا سازگارسازی (Consistency Loss/Regularization) بکار روند.

مدل یادگیری: مدل یادگیری، در زمینه یادگیری ماشین، به یک نمایش ریاضی یا محاسباتی از یک فرآیند واقعی یا پدیده اشاره دارد که الگوریتم های یادگیری ماشین از داده ها به آن یاد می گیرند. اساساً یک مدل یادگیری، یک چارچوب یا طرح بندی است که مشخص می کند چگونه الگوریتم باید الگوها را از داده ها تفسیر کند و به طور کلی از آنها برای پیش بینی یا تصمیم گیری استفاده کند.

وزن های مدل: در یادگیری ماشین، وزن های مدل به پارامترهایی اطلاق می شود که الگوریتم یادگیری آنها را در طول فرآیند آموزش تنظیم می کند تا مدل را به داده های آموزش بسازد. این وزن ها مقادیر عددی هستند که نمایانگر قدرت ارتباط بین نوروها در شبکه های عصبی یا ضرایب در مدل های خطی می باشند. آنها رابطه ریاضی بین ویژگی های ورودی و خروجی پیش بینی شده را تعریف می کنند. در شکل ۱-۸ می توانید یک نمونه از مدل ساده به همراه وزن ها را مشاهده کنید.

ویژگی: در مفهوم یادگیری ماشین، ویژگی (Feature) به معنای خصوصیت های قابل استخراج و تفکیک از داده ها است که برای توصیف یا نمایش آنها استفاده می شود. این ویژگی ها معمولاً به صورت عددی یا دسته ای مشخص می شوند در شکل ۱-۹ می توانید یک نمونه آنرا مشاهده کنید.

بهینه ساز: بهینه ساز یا Optimizer برای یادگیری داده های است. ابزاری است که در فرآیند آموزش مدل به کار می رود تا مقدار وزن ها یا پارامترهای مدل را بهینه کند. هدف اصلی بهینه سازها این است که با کمینه کردن تابع هزینه یا تابع هدفی که مدل باید به آن همگرا شود، مدل را به بهترین حالت ممکن برسانند.

دوره: دوره یا همان Epoch به مرحله‌ای از فرآیند آموزش مدل اشاره دارد که در آن مدل با تمام داده‌های آموزشی خود یکبار آموزش داده می‌شود.

بسته‌های داده: بسته یا همان Batch به مجموعه‌ای از داده‌ها اشاره دارد که به صورت همزمان وارد فرآیند آموزش مدل می‌شوند. به آن بیچ نیز در فارسی گفته می‌شود.

۳-۱-۱ قطعه‌بندی

یکی از وظایف حوزه بینایی ماشین قطعه‌بندی تصاویر است که به طور کلی به دو شکل انجام می‌شود. قطعه‌بندی معنایی^۶ و قطعه‌بندی موردی^۷. در قطعه‌بندی معنایی، هر پیکسل در یک تصویر بر اساس دسته یا کلاس خود طبقه‌بندی می‌شود. نکات کلیدی عبارتند از:

- **قطعه‌بندی مبتنی بر دسته:** تمام اشیاء در یک دسته به عنوان یک موجودیت در نظر گرفته می‌شوند. برای مثال، در یک تصویر با چندین خودرو، همه پیکسل‌های خودرو به عنوان خودرو برچسب‌گذاری می‌شوند بدون اینکه تفاوتی بین خودروهای مختلف قائل شوند.

- **برچسب‌گذاری سطح پیکسل:** هر پیکسل در تصویر یک برچسب مربوط به دسته‌ای که به آن تعلق دارد دریافت می‌کند.

- **موارد استفاده:** معمولاً در کاربردهایی که درک کلی از محتوای تصویر مهم است استفاده می‌شود، مانند فهم صحنه، رانندگی خودکار (برای جاده، آسمان، درختان)، و تصویربرداری پزشکی (شناسایی انواع مختلف بافت‌ها).

در مقابل، قطعه‌بندی موردی گامی فراتر می‌گذارد و نه تنها هر پیکسل را قطعه‌بندی می‌کند، بلکه بین نمونه‌های مختلف از یک کلاس نیز تمایز قائل می‌شود. نکات کلیدی عبارتند از:

- **قطعه‌بندی مبتنی بر نمونه:** نمونه‌های مختلف از یک دسته به طور جداگانه در نظر گرفته می‌شوند. برای مثال، هر خودرو در یک تصویر به طور مجزا شناسایی می‌شود و برای هر خودرو برچسب‌های جداگانه‌ای اختصاص داده می‌شود.

- **شناسایی دقیق اشیاء:** این روش امکان شناسایی و تفکیک دقیق‌تر اشیاء را فراهم می‌کند، شناسایی هر شیء به عنوان یک واحد منحصربه‌فرد.

6. Semantic Segmentation

7. Instance Segmentation

- **برچسب‌گذاری سطح پیکسل و نمونه:** مانند تقسیم‌بندی معنایی، هر پیکسل برچسب‌گذاری می‌شود، اما علاوه بر آن، پیکسل‌ها بر اساس اشیاء منفرد گروه‌بندی می‌شوند.
- **موارد استفاده:** در کاربردهایی که نیاز به شناسایی دقیق و مکان‌یابی اشیاء دارند مفید است، مانند رباتیک (دستکاری اشیاء منفرد)، واقعیت افزوده (اضافه کردن اطلاعات دیجیتال بر روی اشیاء خاص)، و تصویربرداری پزشکی پیشرفته (شناسایی ضایعات یا ناهنجاری‌های مختلف).



شکل ۱-۱۰: نمونه‌ای از هرکدام از قطعه‌بندی‌های معنایی و موردی [۴]. در سمت راست قطعه‌بندی موردی و در سمت چپ قطعه‌بندی معنایی انجام گرفته است.

۲-۱ تعریف مساله

در این پژوهش قصد داریم تا با کمک داده‌های برچسب‌گذاری شده و تعداد زیادی داده‌های بدون برچسب به همراه گزارش پزشکان به یک مدل بهبود یافته برای کمک به پزشکان در تشخیص نواحی جراحات در تصاویر ماموگرافی دست‌یابیم. ورودی مساله، تصاویر X-ray ماموگرم به همراه برچسب‌هایشان (برای آن دسته داده‌هایی که وجود دارند). تعداد زیادی داده‌های بدون برچسب و همچنین گزارش نوشته‌شده پزشکان است. این ورودی‌ها وارد خط‌لوله ما می‌شود و مدلی با این ورودی‌ها آموزش می‌بیند که خروجی آن قطعه‌بندی تصاویر X-ray ورودیست. در ادامه این مدل می‌تواند در زمان تست به دکترها در قطعه‌بندی تصاویر و نشان دادن محل جراحات در داده‌های هدف کمک کند.

۳-۱ اهداف پژوهش

در این مطالعه، قصد داریم ارزیابی کنیم که آیا استفاده از اطلاعات گزارش‌های پزشکی برای یک مجموعه داده خصوصی بزرگ در کنار مجموعه‌های داده کوچک برچسب‌گذاری شده می‌تواند منجر به بهبود عملکرد مدل شود. ابتدا یک خط‌لوله معرفی کرده‌ایم تا از تصاویر خصوصی و گزارش‌های تخصصی مرتبط آن‌ها، یک مجموعه داده بزرگ با برچسب نادقیق برای انجام وظیفه شناسایی آسیب در ماموگرام‌ها به دست آوریم.

سپس چندین روش مختلف را برای استفاده از برچسب‌های نادقیق در آموزش مدل‌های هوش مصنوعی ارزیابی کرده‌ایم و نتایج را با زمانی که مدل تنها با استفاده از یک مجموعه داده کوچک دقیق آموزش داده شده است، مقایسه کرده‌ایم. این پژوهش نهایت دستاورد های زیر را به همراه دارد.

- معرفی یک خط‌لوله نوین برای برچسب‌گذاری ماموگرام‌ها به منظور شناسایی آسیب بدون نیاز به وجود برچسب در داده های آموزش اولیه.
- این مطالعه یک تلاش جامع برای ارزیابی اثربخشی راه‌حل‌های مختلف برای استفاده از مجموعه‌های داده با برچسب‌گذاری نادقیق در قطعه‌بندی جراحی در ماموگرام‌ها است.

۴-۱ اهمیت موضوع

تفسیر ماموگرام‌ها یکی از چالش‌های بزرگ در علوم پزشکی است. در چنین حوزه‌ای، یک مجموعه داده بزرگ و متنوع برای دستیابی به کلیت ضروری است. با وجود چندین مجموعه داده ماموگرام با برچسب، فرصتی برای آموزش بهتر مدل‌ها در مجموعه داده‌های خصوصی بیمارستان‌ها وجود دارد. اگرچه تصاویر بیمارستان‌ها به طور دقیق برچسب‌گذاری نشده‌اند، اما گزارش تخصصی برای هر تصویر وجود دارد که نوع، تعداد و موقعیت تقریبی آسیب‌ها در سینه را نشان می‌دهد. با این حال، یک مجموعه داده مناسب برای شناسایی آسیب‌ها نیازمند برچسب‌گذاری دقیق در سطح پیکسل است که بسیار زمان‌بر، هزینه‌بر و نیازمند دانش تخصصی است. با این حال، در سال‌های اخیر بسیاری از مطالعات بر روی آموزش مدل‌های هوش مصنوعی با یک مجموعه داده کوچک دقیق همراه با یک مجموعه داده بزرگ بدون برچسب یا با برچسب‌گذاری نادقیق تمرکز کرده‌اند تا عملکرد و کلیت را با هزینه کمتر بهبود بخشند.

۵-۱ چالش‌ها

در این پژوهش به برخی مشکلات و چالش‌ها برخوردیم که آن‌ها را برمی‌شماریم و سپس در فصول بعدی با ارائه راه‌حل خویش به شیوه‌ای نوین برای فائق آمدن بر آنها دست می‌یابیم.

- کثرت حجم دادگان بدون برچسب توانایی ما برای دستیابی به مخزن داده کاملاً برچسب گذاری شده را عملاً غیر ممکن می‌سازد. با توجه به اینکه برچسب زدن این داده‌ها نیازمند یک فرد متخصص و با تجربه است این مساله خودش بر پیچیدگی و سختی فرایند برچسب زدن افزوده می‌کند. بنابراین راه حل ما باید با ارائه یا استفاده از یک روش مشخص بر این مشکل غلبه می‌کرد.

- از دیگر مسائل پیش‌رو این است که تصاویر ماموگرم‌ها ابعاد بزرگی دارند 4096×4096 و ما به توان محاسباتی کافی برای آموزش یک مدل با این حجم بزرگ ورودی را دارا نیستیم. بنابراین باید با حجم کمتری از ابعاد کار کنیم، اما این مساله موجب می‌شود تا دقت و کیفیت حاصل از تصاویر با رزلوشن بالا را از دست بدهیم. بنابراین باید یکی روشی (که در این پژوهش روش مدل اجماعی پیشنهاد می‌شود) بکار ببریم که هم از مزیت تصاویر با رزلوشن بالا بهره‌برد هم در عین حال با ابعاد بالا دست و پنجه نرم نکند.
- با وجود کاهش ابعاد تصاویر به ۵۱۲، همچنان فضای مورد نیاز برای آموزش دادگان با توجه به توان پردازشی ما بالاست. بنابراین باید بین بعضی از فاکتورهای آموزش مانند تعداد اپیاک‌ها و حجم بچ‌ها مصالحه صورت بگیرد و بدست آوردن حالتی که بتوان به یک آموزش پایدار رسید یکی از چالش‌های مهم ما در این پژوهش می‌باشد.
- با توجه به اینکه در این پژوهش اطلاعات داخل گزارش پزشکان به سطح دانشی پیکسلی انتقال می‌یابند، تعمیم و تطابق‌پذیری روش‌های قبل برای یادگیری نیمه نظارتی یکی از مسائل مهمی است که باید در زمان ارائه شیوه‌نویس ما مورد توجه قرار گیرد. روش ما باید بتواند انواع متفاوتی از روش‌های نیمه نظارتی را به سادگی به خودش تطبیق دهد.

۱-۶ ساختار پایان‌نامه

• فصل اول: مقدمه

این فصل شامل مقدمه‌ای است که به بحث در مورد موضوعاتی که در پایان‌نامه بررسی می‌شوند، می‌پردازد. در این بخش، مسئله مورد بررسی معرفی و مفاهیم مورد استفاده در آن توضیح داده می‌شوند. سپس اهداف پژوهش شرح داده می‌شود و در ادامه به اهمیت موضوع و چالش‌ها موجود می‌پردازیم. همچنین، ساختار کلی پایان‌نامه نیز معرفی می‌شود تا خواننده بتواند به راحتی با جریان کلی مطالب آشنا شود.

• فصل دوم: کارهای پیشین

در این فصل به بررسی پیشینه تحقیق در حوزه‌های مرتبط پرداخته می‌شود تا زمینه علمی مناسبی برای درک بهتر مطالب فراهم آید.

• فصل سوم: روش پیشنهادی

در این فصل، پیکربندی برای اجرای روش پیشنهادی داده شده است. نتایج بدست آمده با استفاده از روش‌های سابق و همچنین روش پیشنهادی بررسی شده و پس از آن تحلیل می‌شوند. این فصل به تشریح جزئیات پیاده‌سازی و اجرای روش‌ها، ابزارها و داده‌های استفاده شده می‌پردازد.

• **فصل چهارم: نتایج تجربی** در این فصل به جزئیات پیاده‌سازی شبیه‌سازی‌ها و نتایج آنها پرداخته می‌شود. همچنین، نتایج تجربی به دست آمده ارائه و با روش‌های قبلی مقایسه می‌شوند تا مزایا و معایب روش پیشنهادی مشخص گردد.

• **فصل پنجم: جمع‌بندی و پیشنهادات**

این فصل به جمع‌بندی کارهای انجام شده در این پایان‌نامه و ارائه پیشنهادهایی برای کارهای آتی در حوزه‌ی قطعه‌بندی تصاویر ماموگرافی با استفاده از یادگیری شبه‌نظارتی می‌پردازد. در این بخش، یافته‌های اصلی این پایان‌نامه به طور خلاصه مورد بررسی قرار می‌گیرند و سپس پیشنهاداتی برای کارهای آینده در این حوزه ارائه می‌شوند. این پیشنهادات به منظور توسعه و بهبود روش‌ها و الگوریتم‌های مربوطه ارائه شده‌اند.

فصل ۲

کارهای پیشین

۱-۲ روش های نیمه نظارتی

یادگیری نیمه نظارتی (SSL) : یک نوع یادگیری ماشین است که به هدف بهره برداری از گروه های بزرگی از داده های بدون برچسب همراه با مجموعه های داده برچسب گذاری شده برای دستیابی به عملکرد بالاتر می پردازد [۵]، [۶]. روش های نیمه نظارتی مورد استفاده برای بهبود یادگیری بر روی داده های بدون برچسب می توانند همچنین بر روی مجموعه های داده با برچسب نادقیق به منظور آموزش مقاوم تری اعمال شوند. این روش ها عموماً دو روش کلی پیروی می کنند: روش ها با شبه برچسب (PS) ^۱ و روش های مبتنی بر تابع هزینه سازگار ساز (CR) ^۲. برخی از مطالعات همچنین این دو رویکرد را ترکیب کرده اند. پیش از بررسی هر رویکرد بهتر است یک نگاه اجمالی به فروضی که در مساله یادگیری نیمه نظاتی وجود دارد داشته باشیم.

• **فرض همواری:** اگر دو نقطه x_1, x_2 در یک منطقه با چگالی بالا نزدیک به هم باشند، خروجی های آنها y_1, y_2 نیز باید نزدیک به هم باشند. این فرضیه در وظایف طبقه بندی با اطمینان از اینکه ورودی های مشابه، خروجی های مشابهی دارند، کمک کننده است [۷].

• **فرض خوشه بندی:** نقاطی که در یک خوشه قرار دارند احتمالاً از یک کلاس هستند. مرزهای تصمیم باید در مناطق با چگالی پایین قرار گیرند تا از تقسیم خوشه ها به کلاس های مختلف جلوگیری شود [۷].

1. Pseudo-label
2. Consistency Regularization

- فرض منیفولد: داده‌های با ابعاد بالا روی یک منیفولد با ابعاد پایین قرار دارند. این امر با یافتن یک نمایش با ابعاد پایین از داده‌ها، وظایف را ساده می‌کند [۷].

۲-۲ روش‌های نیمه نظارتی با شبه برچسب

کمینه‌سازی آنتروپی [۸] یکی از اولین راه‌حل‌های پیشنهادی بود که در آن از داده‌های بدون برچسب برای کمینه‌سازی آنتروپی استفاده می‌شود، که منجر به مدل تیزتر می‌شود. در [۹] به طور تجربی اثبات کرده‌اند که استفاده از شبه برچسب که برای نمونه‌های بدون برچسب بر اساس بیشینه احتمال اختصاص داده می‌شود، موثر است. [۱۰] روش‌های مبتنی بر گراف را پیشنهاد داده‌اند که در آن شبه‌برچسب‌ها از نمونه‌های برچسب‌گذاری شده به نمونه‌های بدون برچسب منتقل می‌شوند. [۱۱] پیشنهاد استفاده از شبه برچسب‌های مطمئن‌تر (افزودن یک مرز برای احتمال) در آموزش مدل را می‌دهد که در عمل به بهبودی عملکرد مدل منجر می‌شود. شبه برچسب زنی بسیار به مفهوم خودآموزی [۱۲]، [۱۳] شباهت دارد که در آن مدل از دانش خود برای برچسب‌گذاری داده‌های بدون برچسب برای استفاده بیشتر در فرآیند آموزش استفاده می‌کند. [۱۴] از یک مدل معلم-دانش‌آموز^۳ برای تولید شبه‌برچسب استفاده کند و مدل دانش‌آموز با آن‌ها را آموزش دهد و نتیجه آن این است که دانش مدل معلم به دانش‌آموز منتقل می‌شود. در ادامه این متدها، [۱۵] تخمین عدم اطمینان را در CR برای اندازه‌گیری سازگاری بین مدل معلم و دانش‌آموز گنجانده است که منجر به انتقال بهتر دانش از معلم به دانش‌آموز می‌شود. علاوه بر این، یکی دیگر از متدهای مشهور استفاده از روش‌های خودنظارتی با تابع هزینه متناقض محاسبه گر^۴ برای پیش‌آموزش و بهینه‌سازی مدل برای وظایف نیمه‌نظارتی در آینده است ([۱۶] [۱۷] [۱۸] [۱۹] [۲۰] [۲۱]).

۳-۲ روش‌های نیمه نظارتی مبتنی بر تابع سازگار ساز

توافق شبه‌برچسب اجماعی^۵ (PEA) [۲۲] پیشنهاد کرد که از واریانس خروجی چندین مدل به عنوان تابع هزینه تنظیم کننده برای هر لایه شبکه برای نمونه‌های بدون برچسب استفاده شود. [۲۳] یک تابع هزینه بدون نظارت پیشنهاد کرد که دو گذر با دگرگونی تصادفی ورودی و سپس تفاوت خروجی را به عنوان جریمه برای مدل محاسبه می‌کند تا کالیبره شود. مدل II توسط [۲۴] معرفی شد، که تابع هزینه CR را بر روی میانگین متحرک نمایی (EMA)^۶ پیش‌بینی مدل در طول زمان برای هر نمونه آموزشی اعمال می‌کند تا از دو

3. Teacher-Student Model

4. Contrastive Loss

5. Pseudo-Ensemble Agreement

6. Exponential Moving Average

گذر برای هر نمونه در مقایسه با [۲۳]^۷ جلوگیری کند. معلم میانگین متحرک نمایی وزن‌های مدل را به جای به اشتراک گذاری وزن‌ها با مدل دانش آموز انجام می‌دهد؛ مدل معلم از وزن‌های EMA مدل دانش آموز استفاده می‌کند و اطلاعات را بعد از هر مرحله به جای هر دوره جمع‌آوری می‌کند. با دنبال کردن ایده Mix-Up [۲۶]، که داده‌ها را با ترکیب محدب نمونه‌های بدون برچسب تولید می‌کند و CR را بر روی پیش‌بینی این ترکیب و ترکیب محدب خروجی هر نمونه اعمال می‌کند، سازگاری درون‌ترازی^۸ (ICT) [۲۷] از ایده Mix-Up برای فرآیند قطعه‌بندی استفاده می‌کند. در مقابل، سازگاری متقابل (CCT)^۹ [۲۸] پیشنهاد داد که رمزگذار را با اعمال CR بین خروجی رمزگشای اصلی و چندین رمزگشای کمکی که نسخه‌های مختلف افزوده شده خروجی رمزگذار را دریافت می‌کنند، مقاوم‌تر کند. افزودگی بدون نظارت داده‌ها (UDA)^{۱۰} [۲۹] از CR بین خروجی‌های ورودی اصلی و نسخه افزوده ورودی استفاده می‌کند. سازگاری هرم اصلاح‌شده بر اساس عدم قطعیت (URPC)^{۱۱} [۳۰] از پیش‌بینی‌های تولیدشده از هر لایه شبکه برای ارائه یک تابع هزینه سازگاری چندمقیاسی برای نمونه‌های بدون برچسب استفاده می‌کند. استفاده از مدل مقاوم در برابر حملات خصمانه [۳۱] [۳۲] [۳۳] نیز برای روش‌های مختلف قابل اعمال در CR نیمه‌نظارتی در نظر گرفته شد، مانند [۳۴] که به جای برچسب‌ها از خروجی مدل با حملات خصمانه با تابع هزینه خصمانه استفاده می‌کند تا مدل در برابر نویز ورودی مقاوم شود. [۳۵] تابع هزینه خصمانه با ترکیبی از آنتروپی را برای این منظور پیشنهاد می‌کند. [۳۶] مدل‌های معلم-دانش آموز را برای اعمال CR به کار گرفته است. در این روش، CR بین خروجی مدل معلم بر روی ورودی افزوده شده و خروجی افزوده شده مدل دانش آموز بر روی ورودی اصلی محاسبه شده است. [۳۷] پیشنهاد کرد که CR بین مناطق جایگزین شده تصاویر ورودی برای مدل معلم و دانش آموز استفاده شود، که در آن‌ها قسمت‌های مختلف تصویر را کپی و جای‌گذاری می‌کنند و از خروجی مدل معلم برای خروجی دانش آموز به عنوان تابع هزینه CR استفاده می‌کنند.

[۳۸] به استفاده از تابع هزینه سازگار بین تو نسخه دگرگون شده از ورودی با استفاده از دگرگونی سطح پیکسلی و دگرگونی سطحی تر با توابع سطح توصیه می‌کند.

7. Mean-Teacher

8. Interpolation Consistency Training

9. Cross-Consistency Training

10. Unsupervised Data Augmentatio

11. Uncertainty Rectified Pyramid Consistency

۴-۲ روش های نیمه نظارتی ترکیبی

FixMatch [۳۹] از پیش‌بینی مدل بر روی یک تصویر ضعیف-افزوده شده با اطمینان بالا برای نظارت بر پیش‌بینی نسخه قوی-افزوده شده همان تصویر استفاده می‌کند. UniMatch [۴۰] که از FixMatch الهام گرفته شده است، به جای یک تصویر، از دو تصویر قوی-افزوده شده استفاده می‌کند و علاوه بر این، یک اختلال ویژگی را اعمال می‌کند که افزودن نویز به نقشه ویژگی قبل از رمزگشا به خروجی رمزگذار شبکه است. در [۴۱]، نویسندگان MixMatch را پیشنهاد می‌کنند که در آن یک MixUp از دیدهای افزوده‌شده هر دو نمونه برچسب‌دار و بدون برچسب در آموزش استفاده می‌شود. علاوه بر این، در DivideMix [۴۲]، روش پیشنهادی از یک مدل مخلوط برای تقسیم نمونه‌ها به گروه‌های تمیز و نویزی استفاده می‌کند تا داده‌های تمیز را به‌طور تکراری فیلتر کرده و از مجموعه داده‌های نویزی با روش Mix-Match استفاده کند. روش‌های جدید نگرانی‌های مربوط به عدم توازن داده‌های برچسب‌دار در آموزش را مطرح می‌کنند، در حالی که توزیع واقعی متوازن است. [۴۳] برای روش FixMatch، آستانه‌های مختلفی برای کلاس‌های مختلف پیشنهاد می‌کند، در حالی که روش‌های قبلی از یک آستانه ثابت برای همه نمونه‌ها استفاده می‌کردند، این روش از نسبت‌نمایی نمونه شبه‌برچسب‌ها برای آستانه استفاده می‌کند. [۴۴] پیشنهاد می‌کند که بایاس شبکه را، که همان تمایل شبکه به برچسب‌گذاری نمونه‌ها به سمت کلاس‌های فراوان‌تر است، بر روی یک لایه MLP اضافی در انتهای شبکه اعمال کند، لایه اضافی بر روی دسته‌های متوازن کلاس‌ها آموزش داده می‌شود در حالی که سایر پارامترها ثابت هستند. در نظارت متقابل شبه‌ای (CPS) [۴۵]^{۱۲}، دو شبکه برای آموزش مشترک استفاده می‌شوند. در هر تکرار، برچسب‌های سخت حاصل از هر دو شبکه مختل شده، و برای نظارت بر دیگری استفاده می‌شوند.

در جدول ۱-۲ می‌توانید به مقایسه بین روش‌های مهم و نکات کلیدی قوت و ضعف هر کدام نگاه کنید.

12. Cross Pseudo Supervision

جدول ۲-۱: مقایسه روش های مهم در کارهای پیشین

دسته	روش	مزایا	معایب
شبه برچسب	کمینه سازی آنتروپی	هزینه کم + تعمیم پذیری	بهبود محدود + حساس به فراپارامتر
سازگار ساز	مدل II	تعمیم پذیری	سربار اضافی محاسباتی + وابستگی به دیتا
	Mix-Up	سادگی	تنها برای داده های تصویری
	ICT	سادگی + مقیاس پذیری	وابستگی به داده + سربار اضافی محاسباتی
	CCT	تعمیم پذیری	وابستگی به دیتا + پیچیدگی بالا
	UDA	بهره وری داده ای بالا	خطر افزونی مخرب به داده
	URPC	مقیاس پذیری	پیچیدگی + وابستگی به ساختار هرمی
	ADVENT	افزایش اطمینان مدل	وابستگی به شیفت توزیع
ترکیبی	UA-MT	توجه به عدم اطمینان	وابستگی به تخمین گر عدم اطمینان
	CPS	بهره وری داده ای بالا	وابسته به شیفت توزیع
	FixMatch	مقیاس پذیری + سادگی	وابسته به استراتژی شبه برچسب
	UniMatch	توجه به ویژگی های لایه ای	پیچیدگی + وابستگی به افزونی
	MixMatch	استفاده از مدل اجماعی	وابسته به فراپارامتر
	Adaptive Bias	عدم وابستگی به ناتوانی	پیچیدگی + وابسته به تخمین عدم اطمینان

۲-۵ بررسی شبکه ها و توابع مورد استفاده این پژوهش

در این قسمت به بررسی شبکه ها و توابع هزینه مورد استفاده در این پژوهش از کارهای قبلی می پردازیم.

۲-۵-۱ توابع هزینه

Cross entropy loss: تابع هزینه انتروپی متقاطع به صورت زیر تعریف می شود:

$$\text{تابع هزینه انتروپی متقاطع} = - \sum_{j=1}^C y_{ij} \log(\hat{y}_{ij})$$

که در آن:

- C تعداد کلاس ها است.

- y_{ij} برابر با یک است اگر برچسب نمونه i برابر با کلاس j باشد و صفر در غیر این صورت.

- \hat{y}_{ij} احتمال پیش‌بینی شده برای کلاس j توسط مدل برای نمونه i است.

این تابع هزینه برای اندازه‌گیری تفاوت بین توزیع احتمال پیش‌بینی شده توسط مدل و توزیع واقعی برچسب‌ها در مسائل طبقه‌بندی استفاده می‌شود. با استفاده از لگاریتم در تابع هزینه، پیش‌بینی دقیق‌تر (یعنی احتمال بالاتر برای کلاس صحیح) تشویق می‌شود و خطاهای بزرگ‌تر (که با احتمالات پیش‌بینی پایین تر همراه هستند) بیشتر مجازات می‌شوند.

Tversky: تابع هزینه تورسکی یکی از توابع معمول استفاده شده در وظایف قطعه بندی تصاویر است، به ویژه در شبکه‌های عمیق کانولوشنی کاملاً سه‌بعدی. این تابع برای مقابله با عدم تعادل کلاس‌ها طراحی شده است و تعادل بین مثبت‌های غلط و منفی‌های غلط را تنظیم می‌کند. تابع هزینه تورسکی به صورت زیر تعریف می‌شود:

$$\text{تابع هزینه تورسکی} = \frac{\sum_i y_i \hat{y}_i}{\sum_i y_i \hat{y}_i + \alpha \sum_i y_i (1 - \hat{y}_i) + \beta \sum_i (1 - y_i) \hat{y}_i}$$

که در آن:

- y_i و \hat{y}_i مقادیر باینری برچسب‌های واقعی و پیش‌بینی شده برای هر واکسل یا پیکسل i هستند.

- α و β پارامترهایی هستند که وزن‌دهی مثبت‌های غلط و منفی‌های غلط را کنترل می‌کنند.

Dice:

تابع هزینه دایس (Dice) به صورت زیر تعریف می‌شود:

$$\text{تابع هزینه GDO} = 1 - \frac{2 \sum_{c=1}^C w_c \sum_{i=1}^N p_{ic} g_{ic}}{\sum_{c=1}^C w_c \left(\sum_{i=1}^N p_{ic} + \sum_{i=1}^N g_{ic} \right)}$$

که در آن:

- C تعداد کلاس‌ها است.

- N تعداد نمونه‌ها یا واکسل‌ها است.

- p_{ic} احتمال یا امتیاز پیش‌بینی شده برای کلاس c در نمونه i است.

- g_{ic} برچسب واقعی باینری برای کلاس c در نمونه i است (اگر نمونه i به کلاس c تعلق داشته باشد، $g_{ic} = 1$ و در غیر این صورت $g_{ic} = 0$).

- $w_c = \frac{1}{(\sum_{i=1}^N g_{ic})^2}$ وزنی است که به هر کلاس اختصاص داده می‌شود و معکوس مربع اندازه هر کلاس در مجموعه داده را در نظر می‌گیرد.

این تابع هزینه به طور معمول در شرایطی استفاده می‌شود که عدم تعادل کلاسی بسیار زیاد است، مانند وظایف قطعه بندی تصاویر پزشکی. می‌دهد.

Focal:

تابع هزینه فوکال به صورت زیر تعریف می‌شود:

$$FL(p_t) = -(1 - p_t)^\gamma \log(p_t)$$

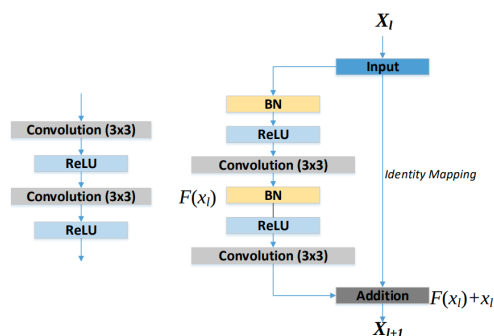
که در آن p_t احتمال پیش‌بینی شده کلاس واقعی است و γ پارامتری است که میزان تمرکز را بر روی نمونه‌های دشوارتر تنظیم می‌کند.

تابع هزینه فوکال به مشکل عدم تعادل کلاس‌ها در تشخیص اشیاء پرداخته و راه‌حلی برای این موضوع ارائه می‌دهد. در وظایف تشخیص اشیاء، اکثر موقعیت‌های تصویر شامل اشیاء مورد نظر نیستند و اکثر تصویر را کلاس پس زمینه تشکیل می‌دهد، در این صورت تابع هزینه‌های پیشین معمولاً بیشتر بر روی کلاس غالب (پس زمینه) تمرکز دارند و نمی‌توانند عملکرد تابع فوکال را داشته باشند.

۲-۵-۲ شبکه‌های پایه

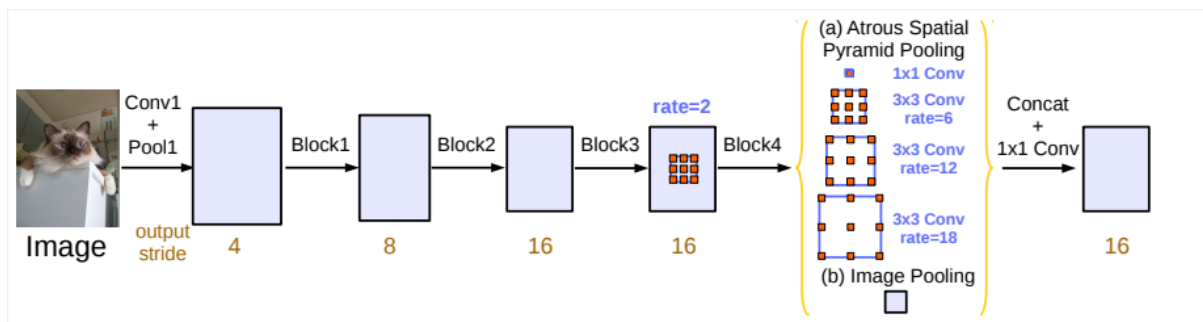
در این بخش به بررسی کوتاهی از معماری شبکه‌های پایه روشنمان می‌پردازیم. دقت کنید که بیشتر مباحث این بخش به صورت تصویری ارائه می‌شود.

ResUNet:



این معماری با کمک بلاک‌های مقابل و در کنار هم قرار دادنشان ساخته می‌شود. سپس با استفاده از آپ‌سَمپلر ها و ایجاد ارتباط هر لایه از لایه متناسب کدگذارش بخش کدگشا ساخته می‌شود.

شکل ۲-۱: تصویر بلاک‌ها ResUNet [۴۶]



شکل ۲-۲: نمونه ای از فرایند کانولوشن و پولینگ آتروس برای یک تصویر [۴۷]

DeepLabv3: معماری این شبکه به شرح زیر است

شبکه اصلی:

- **ResNet-۱۰۱**: شبکه اصلی نسخه‌ی اصلاح شده‌ای از ResNet-۱۰۱ است که پیش‌تر بر روی ImageNet آموزش دیده و برای استخراج ویژگی‌ها استفاده می‌شود.
- **کانولوشن آتروس**: کانولوشن‌های آتروس جایگزین برخی از کانولوشن‌های استاندارد در لایه‌های انتهایی ResNet-۱۰۱ می‌شوند تا میدان دید گیرنده را بدون کاهش وضوح فضایی گسترش دهند. نمونه‌ای از این فرایند در تصویر ۲-۲ قابل مشاهده است.

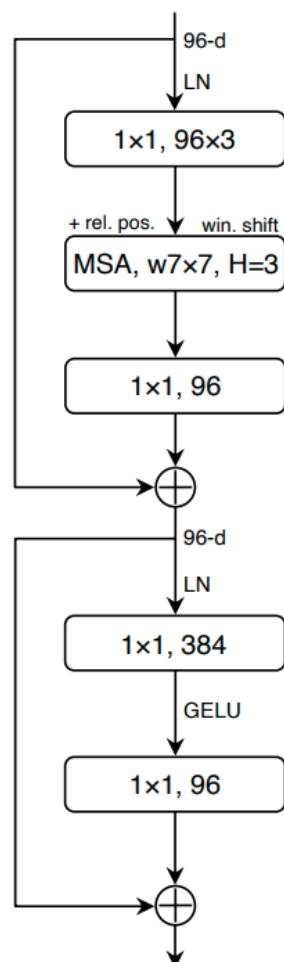
ماژول رمزگشا:

- **تصفیه**: رمزگشا نقشه تفکیک درشت بدست آمده از ماژول آتروس را تصفیه می‌کند و وضوح و دقت را بهبود می‌بخشد.
- **بزرگ‌نمایی**: نقشه‌های ویژگی با استفاده از تفسیر دوخطی بزرگ‌نمایی می‌شوند تا با وضوح اصلی تصویر ورودی مطابقت داشته باشند.

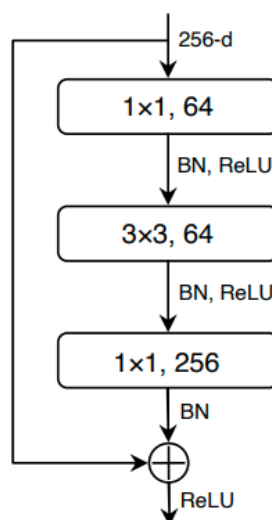
در طی سال‌های اخیر توجه تحقیقات به سمت مبدل‌های بینایی‌ها جلب شده و برخی کارها برای قطعه‌بندی تصاویر با این ترند جدید ترکیب شده‌اند که از چشم این پژوهش‌جا نماندند. ما در ادامه به توضیح کوتاهی از دو روش مبتنی بر مبدل‌های بینایی صحبت می‌کنیم که در روش اصلی به عنوان شبکه پایه استفاده شده‌اند.

ConvUnet: تغییرات حاصل شده در کانولوشن معمولی به شرح زیر است.

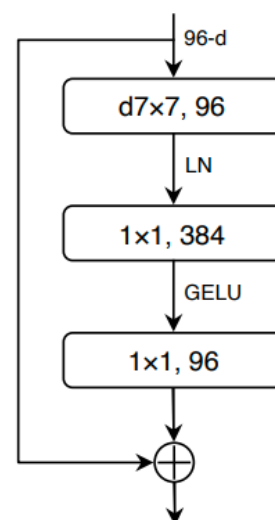
Swin Transformer Block



ResNet Block



ConvNeXt Block



شکل ۲-۳: نمونه از نشان دادن الهامات برگرفته از مبدل بینایی Swin در طراحی این شبکه. برای سادگی، لایه‌های خطی در بلوک‌های MLP ترانسفورمر را نیز به عنوان کانولوشن‌های 1×1 در نظر می‌گیریم می‌کنیم زیرا معادل هستند [۴۷].

- بازبینی طراحی شبکه‌های عصبی کانولوشن استاندارد: این تحقیق طراحی شبکه‌های عصبی کانولوشن استاندارد را بازبینی کرده و چندین تغییر معماری الهام گرفته از مبدل‌های بینایی را در آن‌ها ادغام کرده‌اند. این تغییرات شامل اندازه‌های بزرگ کرنل، بلوک‌های معکوس و نرمال‌سازی لایه هستند.
- عملیات کانولوشن استاندارد: ConvNeXt از عملیات کانولوشن استاندارد استفاده می‌کند و از عملیات پیچیده مانند خودتوجهی در مبدل‌ها اجتناب می‌کند. معماری به گونه‌ای طراحی شده است که ساده و کارآمد باشد و در عین حال عملکرد بالایی در وظایف شناسایی تصویر داشته باشد.

:TransUnet

TransUNet مشابه U-Net و شامل یک کدگذار و یک کدگشا است که اطلاعات تصویر را کدگذاری

که اخیراً دو مساله جدید در این حوزه مطرح شده است و در های جدیدی باز شده اند. اولین مساله زمانی است که داده ها در واقعیت توزیع متوازن دارند اما در داده های با برچسب این توزیع نامتوازن است و این سبب می شود که زمان استفاده از داده های بدون برچسب بایاس موجود در داده های با برچسب به آنها انتقال یابد و توزیع اصلی یادگرفته نشود. مساله دوم مساله Confirmation Bias است که بیانگر این است که توزیع اولیه که مدل یاد می گیرد بسیار مهم است و بایاس موجود در این مرحله بر باقی فرایند آموزش تاثیر بسزایی دارد. دقت کنید که این پژوهش نیم نگاهی به یکی از روش هایی که به عدم توازن توزیع توجه می کند دارد و نتایج آن در بخش شبیه سازی قابل مشاهده است.

فصل ۳

روش پیشنهادی

در این بخش به شرح روش خودمان می‌پردازیم. ابتدا مساله را با ابزارهای ریاضی قاعده مند می‌کنیم و سپس روش ما که شامل دو بخش است را به تفصیل توضیح می‌دهیم. بخش اول شامل یک خط لوله است که ورودی گزارش ها و عکس ها را گرفته و در خروجی یک برچسب تولید شده و یک نقشه اطمینان می‌دهد که جریان اصلی روش ماست و آنرا در شکل ۳-۳ مشاهده می‌کنید. سپس به شرح چگونگی انطباق روش های نیمه نظارتی بر این دو خروجی قسمت دوم می‌پردازیم.

۳-۱ قاعده سازی مساله

ما در نظر داریم که ناحیه جراحی دیده در تصویر ماموگرافی را با استفاده از جفت تصاویر ورودی و برچسب ها (تصاویر قطعه‌بندی شده) و با بهره‌گیری از تصاویر بدون برچسب فراوان، قطعه‌بندی کنیم. برای این منظور مسئله را به صورت زیر فرموله می‌کنیم. (x_l^i, y_l^i) به عنوان i^{th} جفت تصاویر ورودی و برچسب حقیقی از داده‌های برچسب‌دار در نظر گرفته می‌شود. فرمول زیر تابع هزینه که یک تابع هزینه مرسوم است را ارائه می‌دهد.

$$\mathcal{L}_l = \frac{1}{N} \sum_{i=1}^N \ell_l(f_x(h(x_l^i; \theta)), f_y(y_l^i)) \quad (۳-۱)$$

اکنون، به روش نیمه نظارتی می پردازیم. (x_u^i, y_u^i) به عنوان i^{th} جفت داده های ورودی برای داده های بدون برچسب و شبه برچسب های آن داده ها در نظر گرفته می شود. تابع هدف به صورت زیر است:

$$\mathcal{L}_u = \frac{1}{M} \sum_{i=1}^M \ell_u(f_x(h(x_u^i; \theta)), f_y(\hat{y}_u^i)) \quad (2-3)$$

اکنون، تابع هدف نهایی برای هر دو مجموعه داده به صورت زیر خواهد بود:

$$\mathcal{L}_{total} = \mathcal{L}_l + \lambda_u \mathcal{L}_u \quad (3-3)$$

که در آن \hat{y}_u^i می تواند از روش های مختلفی استخراج شود و λ_u یک فرایارامتر^۱ است که اهمیت تابع هزینه داده های بدون برچسب را نشان می دهد. تابع هزینه داده های برچسب دار، \mathcal{L}_l ، که اختلاف بین برچسب های پیش بینی شده و برچسب های واقعی را اندازه گیری می کند، با استفاده از تکنیک های یادگیری نظارتی محاسبه می شود. در حالی که تابع هزینه داده های بدون برچسب، \mathcal{L}_u ، تلاش دارد از اطلاعات موجود در داده های بدون برچسب برای بهبود عملکرد مدل بهره برداری کند. f_x و f_y در اکثر موارد توابع یکسانی هستند، اما گاهی اوقات می توانند خروجی های متعددی از لایه های قبلی باشند، و این بسته به روش آموزشی که اعمال می کنید می تواند متفاوت باشد. برای سادگی از این پس فرض می کنیم هر دو f_x و f_y توابع همانی در فرمول بندی ریاضی ما هستند.

۲-۳ پیش پردازش داده

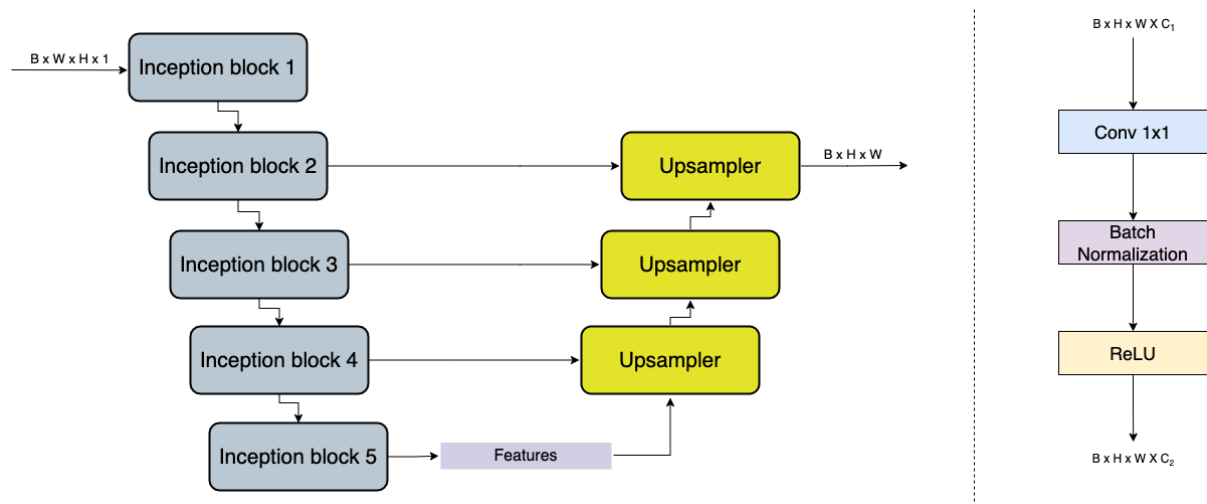
این مرحله جزوی از خطلوله ابدایی این پژوهش است که در ادامه ابتدا با معرفی اجزای آن و سپس طریقه اجرای فرایند آنرا تشریح می کنیم.

تابع برچسب زن (AF)^۲:

ما از یک مدل اجماعی برای تولید برچسب های نیمه نظارتی مورد نیاز فرآیند آموزشی خود استفاده کردیم. این مدل شامل سه شبکه UNet [۴۹] است که هر کدام ماژول های Inception [۵۰] را در لایه های کدگذار خود دارند. شکل ۱-۳ ساختار کلی شبکه مذکور را نشان می دهد. ما به هر شبکه یک اندازه ورودی متفاوت می دهیم؛ یکی با وضوح ۱۰۲۴، دیگری با وضوح ۲۰۴۸، و دیگری با وضوح ۴۰۹۶. این به ما کمک می کند تا از ویژگی های بزرگ مقیاس و ویژگی های جزئی تصاویر استفاده کنیم. سپس از رأی گیری سختگیرانه برای تولید خروجی نهایی قطعه بندی مدل خود استفاده می کنیم.

1. Hyper-parameter

2. Annotator Function



شکل ۳-۱: معماری مدل UNet با ماژول‌های Inception (سمت چپ) و بلوک آپ‌سَمپلر (سمت راست).

ما از تابع هزینه فوکال [۵۱] روی خروجی‌های درونیابی شده^۳ لایه‌های کدگشا در هر شبکه استفاده می‌کنیم و سپس مجموع این توابع هزینه را به عنوان تابع هزینه نهایی خود استفاده می‌کنیم که به صورت زیر تعریف می‌شود:

$$\mathcal{L} = \sum_{i=1}^N -(1 - p_{d_i})^\gamma \log(p_{d_i}) \quad (۴-۳)$$

که در آن p_{d_i} نقشه احتمالاتی درونیابی شده^۴ لایه کدگشا i ام و N تعداد لایه‌های کدگشا در شبکه است. ما به طور خاص از تابع هزینه فوکال برای مقابله با عدم توازن تعداد کلاس در تصاویر ماموگرافی استفاده کردیم.

انگیزه اصلی ما برای استفاده از مدل اجماعی در این مرحله، بهره‌برداری و ترکیب دیدگاه‌ها و ویژگی‌های مختلف تصاویر ورودی برای تولید برچسب‌ها است. آموزش یک شبکه تکی با یک دیدگاه تصویر به عنوان ورودی، ما را از دستیابی به الگوهای پیچیده در بافت سینه باز می‌دارد. پس از آموزش مدل اجماعی خود بر روی یک مجموعه داده کاملاً برچسب‌دار، می‌توانیم با استفاده از مدل خود برای تولید برچسب‌های نیمه‌نظارتی برای تصاویر بدون برچسب، مجموعه داده خود را برای آموزش نیمه‌نظارتی آماده کنیم.

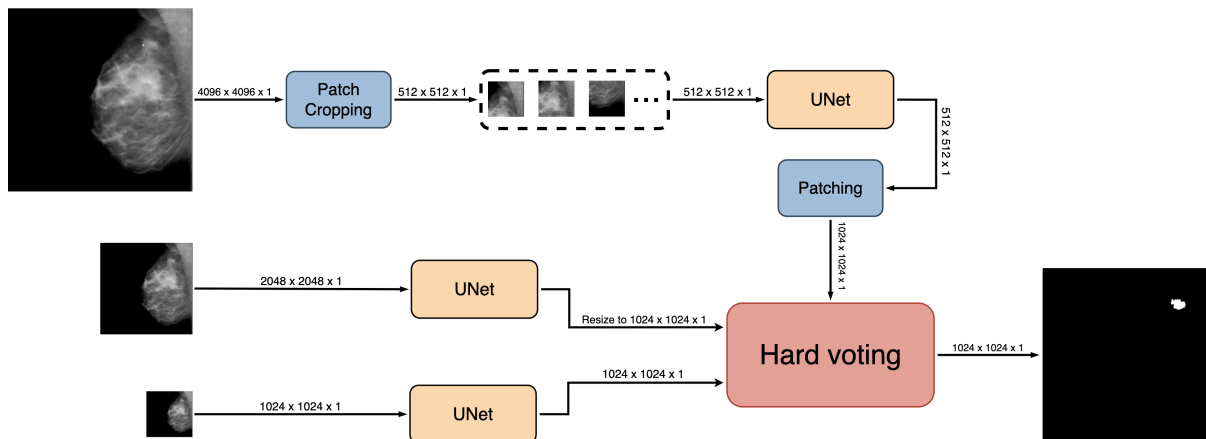
نقشه اطمینان غیرتصادفی (DCM)^۵

زمانی که تعداد جراحات پیش‌بینی شده توسط مدل اجماعی با تعداد آسیب‌های گزارش شده برابر باشد، ما به مدل اعتماد می‌کنیم. زمانی که تعداد پیش‌بینی‌های مدل کمتر از تعداد جراحات گزارش شده باشد، در آن صورت این جراحات را به برای جراحات واقعی در نظر می‌گیریم؛ بقیه را هم به عنوان پس‌زمینه در

3. Interpolated outputs

4. Interpolated Probability Map

5. Deterministic Confidence Map



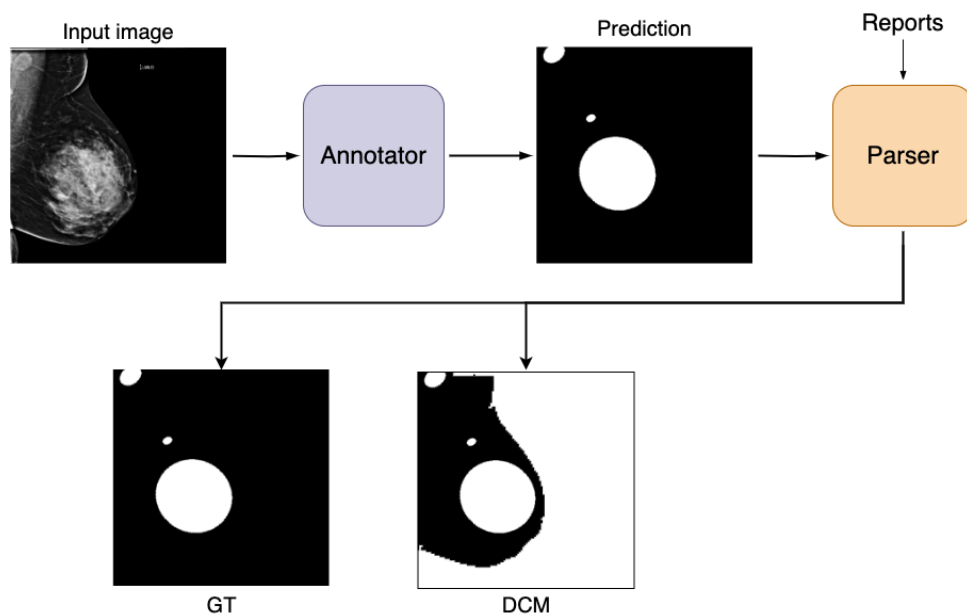
شکل ۳-۲: معماری مدل اجماعی: هر شبکه با یک اندازه ورودی منحصر به فرد تغذیه می‌شود و سپس ماسک نهایی با استفاده از رأی‌گیری سختگیرانه تولید می‌شود.

نظر می‌گیریم، اما باقی جراحات گزارش شده توسط پزشک را غیرقطعی می‌دانیم زیرا می‌دانیم حداقل یک جراحی دیگر در بقیه جراحات وجود دارد، اما نمی‌دانیم کجاست. وقتی وضعیت برعکس باشد و تعداد جراحات پیش‌بینی شده توسط مدل بیشتر از جراحات گزارش شده باشد، بقیه ناحیه را به عنوان پس‌زمینه در نظر می‌گیریم. با این حال، جراحات را به عنوان مناطق غیر قابل اعتماد در نظر می‌گیریم زیرا می‌دانیم حداقل یکی از آن‌ها جراحات نیست. تمامی GT و DCM را برای تمامی نواحی ادغام می‌کنیم و GT و DCM نهایی را ایجاد می‌کنیم. توجه داشته باشید که ما یک بیضی دور مناطق پیش‌بینی شده توسط مدل رسم می‌کنیم تا آن را به عنوان یک شیء پیوسته در تصویر نمایش دهیم. ممکن است در این فرآیند انتقال خطا وجود داشته باشد، اما این بهترین روش برای استفاده از اطلاعات گزارش‌هاست.

ما از گزارش‌های پزشکان برای بهبود اطمینان برچسب‌های تولید شده توسط هر تابع برچسب‌زن استفاده می‌کنیم که مکان آن در خط لوله در شکل ۳-۳ نشان داده شده است. اولین مرحله از این فرآیند شامل پارسر ما است که از برچسب‌های تولید شده توسط AF و گزارش‌های پزشکان بهره می‌برد. ما از خروجی پارسر برای تصمیم‌گیری در مورد اطمینان از پیش‌بینی برچسب پیکسل AF استفاده می‌کنیم. سپس دو خروجی متمایز، برچسب حقیقی (GT)^۶ و نقشه اطمینان غیرتصادفی (DCM) تولید می‌کنیم. هدف DCM شناسایی پیکسل‌هایی است که باید به خروجی AF اعتماد کنیم. ما از اصطلاحات غیرتصادفی استفاده می‌کنیم زیرا خروجی الگوریتم ما برای هر پیکسل بردار تنها یک است. می‌توانید فرآیند را در این بخش در شکل ۳-۳ مشاهده کنید.

شایان ذکر است که در تصاویر ماموگرافی، با دو نوع تصویر، جراحات میکرو و جراحات ماکرو، مواجه می‌شویم. در این مطالعه، ما فقط بر روی نوع ماکرو تمرکز می‌کنیم زیرا از نظر هزینه محاسباتی، توانایی

6. Ground Truth



شکل ۳-۳: روبه کلی تولید GT و DCM. مدل اجماعی پیش‌بینی‌های خام را برای پارسر فراهم می‌کند. سپس پارسر با استفاده از پیش‌بینی‌ها و گزارش‌ها، GT و DCM را تولید می‌کند.

آموزش تصاویر در شبکه‌های سبک‌تر با وضوح پایین‌تر از اهمیت ویژه‌ای برخوردار است. نوع ماکرو نیازمند دریافت نواحی وسیع‌تری از تصویر است تا قابل درک باشد که به ما امکان می‌دهد از تصاویر با وضوح پایین‌تر استفاده کنیم. از طرف دیگر، نوع میکرو نیاز به تحلیل با وضوح بالا دارد. هر تصویر به اندازه 512×512 تغییر اندازه داده می‌شود، نسبت ارتفاع/عرض اصلی حفظ شده و سپس به شبکه تغذیه می‌شود.

۳-۳ یادگیری نیمه نظارتی

اکنون این امر به صورت شهودی قابل فهم است که با استفاده از اطلاعات موجود در گزارش، می‌توانیم بگوییم که هر نمونه‌ای که در بخش ۳-۱ توصیف شد، نمی‌تواند به عنوان داده‌های برچسب‌دار یا بدون برچسب در نظر گرفته شود زیرا اکنون هر تصویر دارای نقشه‌ای به نام DCM است که نشان می‌دهد به کدام برچسب پیکسل اعتماد داریم و به کدام برچسب پیکسل اعتماد نداریم. به عبارت دیگر، ما هر تصویر را به پیکسل‌های برچسب‌دار و پیکسل‌های بدون برچسب تقسیم می‌کنیم. پیکسل‌های بدون برچسب آنهایی هستند که به برچسب‌هایشان اعتماد نداریم. می‌توانیم از این تنظیم استفاده کرده و آن را در مسئله نیمه نظارتی به کار بگیریم تا از پیکسل‌های غیرقابل اعتماد بهره‌مند شویم و از دانش مدل برای کمک به خود استفاده کنیم. اکنون باید نمادگذاری را تغییر دهیم تا با مسئله ما سازگار شود. اکنون تابع هزینه کلی ما به

صورت زیر خواهد بود.

$$\mathcal{L}_l = \left(\frac{1}{T} \sum_{t=1}^T \frac{1}{\sum_{i,j} W_{i,j}^H} \right) \left(\sum_{t=1}^T \sum_{i,j} 1_{DCM_t=1} \ell_l \left(h(x^t; \theta)_{i,j}, y_{i,j}^t \right) \right) \quad (5-3)$$

$$\mathcal{L}_u = \left(\frac{1}{T} \sum_{t=1}^T \frac{1}{\sum_{i,j} W_{i,j}^H} \right) \left(\sum_{t=1}^T \sum_{i,j} 1_{DCM_t \neq 1} \ell_u \left(h(x^t; \theta)_{i,j}, \hat{y}_{i,j}^t \right) \right) \quad (6-3)$$

$$\mathcal{L}_{total} = \mathcal{L}_l + \lambda_u \mathcal{L}_u \quad (7-3)$$

که همه فرض‌ها مانند معادله ۳-۳ هستند فقط توجه داشته باشید که اکنون T اندازه بسته داده است است. اکنون ساختار چارچوب^۷ روش ما دارای سه مرحله است. اول تاسیس یک تابع برچسب زن که در پژوهش ما این تابع یک تابع اجماعی بود، دوم تولید DCM و سوم استفاده از تابع هزینه نهایی به همراه روش‌های مختلفی که در یادگیری نظارتی اعمال می‌شوند که در بخش کارهای پیشین به تفصیل مورد بحث قرار گرفت.

ما ادعا می‌کنیم که دو دلیل برای استفاده از DCM وجود دارد. اول از همه، به وضوح نادیده گرفتن برخی از داده‌های پیکسلی تصمیم ایده‌آلی نیست زیرا داده‌ها یک دارایی مهم در وظایف یادگیری ماشین هستند. دوم اینکه ما با یک مسئله پیچیده‌تر مواجه هستیم که می‌توانیم از اطلاعات در سطح پیکسل استفاده کنیم؛ این مسئله ما را به یک مسئله دشوارتر تبدیل می‌کند زیرا در این وضعیت، برخی از فرضیات درباره روش نیمه نظارتی، فرضیه خوشه‌بندی^۸ [۵۲] [۷]، ضعیف‌تر هستند. مهم است که توجه داشته باشیم این فرضیات هیچ‌گاه در روش‌های نیمه نظارتی معمول به طور کامل برآورده نمی‌شوند. با این حال، آنها در این ساختار ضعیف‌تر هستند. بنابراین، استفاده از DCM به ما کمک می‌کند تا داده‌های بدون برچسب و برچسب‌دار را بهتر جدا کنیم، که این امر احتمال برآورده شدن فرضیه خوشه‌بندی را افزایش می‌دهد.

۴-۳ جمع بندی

با توسعه ابزار معرفی شده در این بخش DCM، که کمک می‌کند که در سطح پیکسل اطلاعات موجود در داده‌های بدون برچسب را به مدل انتقال دهیم توانستیم به یک روش و خط‌لوله دست‌یابیم که در طی این فرایند دو ماسک تولید می‌شود که یکی از آنها برچسب و دیگری نشان می‌دهد که کدامین پیکسل‌های

7. Framework Setup

8. Cluster Assumption

برچسب گذاری شده قابل اعتماد هستند. در انتها با معرفی تابع هزینه بخش بابرچسب و بدون برچسب مولفه های مورد نیاز برای آموزش یک مدل را با کمک داده های بابرچسب و بدون برچسب در اختیار داریم و در بخش بعدی از آنها در اجزای شبیه سازی هایمان استفاده می کنیم.

فصل ۴

نتایج تجربی

۴-۱ دیتاست و معیارهای ارزیابی

برای ارزیابی عملکرد هر مدل، نرخ ضبط^۱ را به کار گرفتیم. به طور خاص، از خروجی تقسیم‌بندی مدل استفاده کرده و یک بیضی در اطراف هر گروه از مناطق پیوسته تقسیم‌بندی شده رسم کردیم و هر بیضی را به عنوان یک شیء مستقل در نظر گرفتیم. برای گزارش موفقیت در ضبط یک شیء، از معیار اشتراک سطوح (IoU)^۲ استفاده کردیم. اگر IoU بیش از ۰.۵ برای مواردی که بیضی مدل کوچکتر از GT بود و بیش از ۰.۷۵ برای مواردی که بیضی مدل بزرگتر از GT بود، در نظر گرفته شد که یک ضبط موفقیت‌آمیز است. علاوه بر این، ما فقط نتایجی را که خطای مثبت کاذب^۳ کمتر از ۲ داشتند را در نظر میگیریم.

داده‌های اصلی ما که در بخش آموزش نیمه نظارتی استفاده شدند، از مجموعه بیمارستان امام خمینی جمع‌آوری شده و گزارش‌های ماموگرافی‌ها توسط شش رادیولوژیست با تجربه ارائه شدند. این مجموعه داده شامل ۴۵,۹۹۴ تصاویر آموزشی و ۵۰۰ تصاویر اعتبارسنجی می‌باشد. ما مجموعه اعتبارسنجی را به صورت دستی انتخاب کردیم تا صحت برچسب‌های تولید شده و تنوع در انواع جراحات را تضمین کنیم. رزولوشن اصلی ماموگرافی‌ها 4096×4096 بود، اما برای کاهش هزینه‌های محاسباتی، رزولوشن تصاویر را به 512×512 کاهش دادیم.

مجموعه داده CBIS-DDSM که برای آموزش مدل اجماعی استفاده شد، شامل ۲۱۲۱ تصویر آموزشی و ۲۶۷ تصویر اعتبارسنجی است. لازم به ذکر است که ما توده‌های کلسیم‌دار را از تقسیم‌بندی حذف کردیم.

-
1. Capture Rate
 2. Intersection over Union
 3. False Positive Rate

جدول ۴-۱: عملکرد ResUNet Light با توابع هزینه مختلف

زیان	نرخ ضبط	نرخ مثبت کاذب	حساسیت منفی
CE	۷.۰۳	۰.۳۰	۹۹.۸۷
Tversky [۵۵]	۰.۰۰	۰.۰۰	۱۰۰.۰۰
Dice [۵۶]	۰.۰۰	۰.۰۰	۱۰۰.۰۰
Focal-۴ [۵۱]	۵۳.۹۱	۱.۶۲	۹۷.۵۰

۴-۲ جزئیات پیاده‌سازی و نتایج شبیه‌سازی

برای اجرای مدل‌ها از فریمورک PyTorch ۱.۱۳.۱ [۵۳] و استفاده از GPU مدل RTX Quadro ۸۰۰۰ بهره می‌بریم. فرآیند آموزش با بهینه‌ساز ADAM [۵۴] با نرخ یادگیری ۰.۰۰۰۱ و کاهش اپسیلونی ۰.۹۹ بهینه‌سازی شد. هر فرآیند آموزشی شامل اندازه دسته ۳۲ و ۲۰۰ تکرار در هر دسته بود. این تنظیمات برای تمامی آزمایش‌های بعدی نیز یکسان است.

۴-۲-۱ پیدا کردن بهترین تابع هزینه

برای شروع پژوهش، به تحلیل مناسب‌ترین تابع هزینه پرداختیم. ما با استفاده از معماری سبک ResUNet شامل سه رمزگذار و رمزگشای عمقی، برای تعیین تابع هزینه مناسب برای فرآیند آموزشی خود آزمایش کردیم. شبکه به مدت ۵۰ دوره فقط بر روی پیکسل‌های مطمئن آموزش داده شد و نتایج در جدول ۴-۱ خلاصه شده است. بر اساس این یافته‌ها، تابع هزینه فوکال (با گاما برابر با ۴ و آلفا برابر با ۰.۷) را به عنوان مؤثرترین تابع هزینه برای تکرارهای آموزشی بعدی انتخاب کردیم. این انتخاب به دلیل عدم توازن در تعداد کلاس پیکسل‌ها بود.

۴-۲-۲ پیدا کردن بهترین شبکه پایه

ما برای تعیین مناسب‌ترین مدل پایه برای این وظیفه، پنج مدل پایه را فقط بر روی پیکسل‌های مورد اعتماد آموزش داده و عملکرد آن‌ها را برای ۲۰۰ دوره مقایسه کردیم. نتایج این ارزیابی در جدول ۴-۲ آمده است. از میان پنج مدل پایه‌ای که فقط بر روی پیکسل‌های مورد اعتماد آموزش دیده‌اند، مدل ResUNet [۴۶] با تنها سه رمزگذار و رمزگشا انتخاب شد. شما می‌توانید ببینید که این مدل از سایر مدل‌های پایه پیشی گرفته است.

جدول ۴-۲: مقایسه بین معماری‌های مختلف شبکه برای ارزیابی پایه در مجموعه داده ما. پایه‌ها فقط بر روی مناطق مورد اطمینان از ماسک‌های تولید شده آموزش دیده‌اند.

شبکه پایه	نرخ ضبط	نرخ مثبت کاذب	حساسیت منفی ^۴
ResUNet [۴۶]	۵۶.۲۵	۱.۸۹	۹۸.۰۲
TransUNet [۴۸]	۳۵.۱۶	۱.۲۱	۹۴.۲۰
DeepLabv۳ [۴۷]	۵۰.۰۰	۱.۵۹	۹۸.۳۹
ConvUnext [۵۷]	۴۹.۲۲	۱.۴۶	۹۸.۲۴

جدول ۴-۳: مقایسه بین روش‌های نیمه نظارتی در مجموعه داده ما.

روش	نرخ ضبط	نرخ مثبت کاذب	حساسیت منفی
ICT [۲۷]	۵۵.۴۷	۱.۷۹	۹۸.۶۳
URPC [۳۰]	۶۰.۱۶	۱.۹۲	۹۷.۵۰
CPS [۴۵]	۴۸.۴۴	۱.۷۴	۹۸.۴۳
UA-MT [۱۵]	۵۶.۲۵	۱.۶۵	۹۸.۹۷
FixMatch [۳۹]	۴۹.۲۲	۱.۶۵	۹۸.۲۷
UniMatch [۴۰]	۵۰.۰۰	۱.۶۳	۹۸.۴۴
ADVENT(MinEnt) [۳۵]	۵۷.۰۳	۱.۶۵	۹۸.۲۱
Bias Adaptive [۴۴]	۴۹.۲۱	۱.۹۶	۹۸.۱۰

۴-۲-۳ انطباق روش‌های نیمه نظارتی

با مدل پایه مشخص شده، به بررسی تأثیر روش‌های یادگیری نیمه نظارتی (SSL) بر فرآیند آموزشی خود پرداختیم. هدف ما بررسی این بود که کدام تکنیک SSL عملکرد را بهبود می‌بخشد و به نتایج بهتر کمک می‌کند. با ادغام روش‌های SSL در خط لوله آموزشی خود، انتظار داریم که بینش‌های عمیق‌تری درباره رفتار مدل و توانایی آن در استفاده مؤثر از داده‌های برچسب‌گذاری نشده کسب کنیم. این تکنیک‌های SSL می‌توانند قابلیت‌های تعمیم مدل را بهبود بخشیده و به مدل اجازه دهند که از داده‌های برچسب‌گذاری نشده به صورت معنی‌دارتری یاد بگیرد. نتایج را می‌توانید در جدول ۴-۳ ببینید.

جدول ۴-۴: عملکرد ResUNet بر روی هر دو مجموعه داده CBIS-DDSM و مجموعه داده ما. مدل بر روی مجموعه داده CBIS-DDSM در یک چارچوب نظارت شده آموزش دیده و سپس بر روی مجموعه داده ما ارزیابی شده است.

مجموعه داده	نرخ ضبط	نرخ مثبت کاذب	حساسیت منفی
CBIS-DDSM	۶۱.۰۶	۴.۰۱	۹۹.۵۱
داده های این پژوهش	۶.۲۶	۲.۸۷	۹۸.۴۱

۴-۳ مقایسه با آموزش ساده

برای نشان دادن اثربخشی و ضرورت آموزش نیمه نظارتی، ما به بررسی تأثیرات اتکا بر یک مدل از پیش آموزش دیده از یک مجموعه داده دیگر در زمینه یادگیری انتقالی پرداختیم. هدف ما ارزیابی پتانسیل آن برای دستیابی به عملکرد خوب بر روی مجموعه داده اصلی خود و مقایسه آن با عملکردی بود که از طریق تکنیک های SSL به دست می آید، جایی که دانش مدل به خودش اعمال می شود. علاوه بر این، ما به استفاده از اطلاعات حاصل از گزارش های پزشکان که به نقشه های اطمینان غیراحتمالی (DCM) تبدیل شده بودند، پرداختیم تا درک مدل و عملکرد کلی آن را بهبود بخشیم. نتایج مطالعه ما، که در جدول ۴-۴ ارائه شده است، شواهدی را ارائه می دهد که استفاده از انحصاری از یک مدل از پیش آموزش دیده از یک مجموعه داده دیگر، عملکرد رضایت بخشی را بر روی مجموعه داده هدف ما به دست نمی آورد. این امر بر ضرورت آموزش SSL تأکید می کند که امکان استفاده از دانش کسب شده مدل را به خودش می دهد. علاوه بر این، استفاده از اطلاعات موجود در DCM تبدیل شده از گزارش های پزشکان برای دستیابی به درک بهتر و عملکرد برتر حیاتی است. بنابراین، یافته های ما بر ضرورت آموزش SSL و استفاده از DCM برای به حداکثر رساندن پتانسیل یادگیری انتقالی و بهبود عملکرد در مجموعه داده اصلی ما تأکید می کند.

۴-۴ جمع بندی

در انتها به نتایج بدست آمده از شبیه سازی هایمان خواهیم پرداخت. اول آنکه مطابق جدول ۴-۲ مدل های سنگین تر پایه منجر به نتایج بهتر نمی شود. اعتقاد ما بر آن است که بدلیل وجود نویز و خطا موجود در داده ها مدل های سنگین توانایی تعمیم پذیریشان به آن نویز ها حساس می شود (Overfit). البته دقت کنید که در سال های اخیر به طور نظری اثبات شده است که نویز در مدل های بسیار سنگین می تواند یک منظم ساز بسیار عالی باشد اما برای اجرای شبیه سازی یک مدل به سنگینی مورد نیاز برای چنین منظم سازی نیازمند

توان محاسباتی فوق‌العاده داریم که متاسفانه پژوهشگران این تحقیق در اختیار نداشتند. برای مطالعه بیشتر می‌توانید به [۵۸] مراجعه کنید. در ادامه با توجه به جدول ۴-۳ می‌توان مشاهده نمود که استفاده از روش نیمه نظارتی URPC منجر به بهبود ۴ درصدی می‌شود. استفاده از DCM و انتطابق روش URPC با داده های ما فرصت استفاده از دانش موجود در داده های بدون برچسب را فراهم کرد و کمک کرد تا به بهبود ۴ درصدی برسیم.

فصل ۵

جمع‌بندی و کارهای آتی

۵-۱ جمع‌بندی

نتایج مدل‌سازی‌ها و شبیه‌سازی‌های که جداول ۴-۳ و ۴-۴ نشان داده شدند بیانگر آن اند که استفاده از گزارش‌های رادیولوژیست‌ها، به رغم اینکه عکس نیستند و متن‌اند، در قالب یک نقشه اطمینان غیر احتمالی در سطح پیکسل به ما این امکان را می‌دهد که بتوانیم روش‌های نیمه نظارتی را با اندکی تغییر استفاده کنیم که ارمغان آن استفاده از حجم دادگان زیاد بدون برچسب در دسترس و امکان انتقال دانش از گزارش‌ها به مدل در حال آموزش است که نتیجه آن بهبود عملکرد کلی مدل برای کمک به پزشکان نسبت به نتایج قبل است. آنچه در پایان این پایان این تحقیق قرار گرفت به طور خلاصه به شرح زیر است:

- ارائه یک خط‌لوله جدید برای استفاده از گزارش پزشکان و انتطابق داده‌ها با یادگیری نیمه نظراتی
- انتباط روش‌های نیمه نظارتی برای استفاده آنها در سطح پیکسل و رسیدن به بهبود ۴ درصدی (جدول ۴-۳) در حالت عدم استفاده از آنها.

۵-۲ کارهای آتی

یکی از مسائلی که اخیراً مورد توجه قرار گرفته است مبحث بایاس اعترافی یا Confirmation Bias در یادگیری نیمه نظراتی است که می‌توان با کمک آن اثر پس‌زمینه را کاهش داد. بنابراین یکی از جهت‌هایی که می‌توان این کار را ادامه داد در جهت در نظر گرفتن بایاس حاصل از توزیع نامتوازن کلاس‌های پیکسل‌ها است در فرایند آموزش است.

یکی دیگر از راه‌های پیش‌رو تعمیم مقیاس قطعه‌بندی به چند کلاسه است که در آن بتوان نوع غدد را نیز مشخص کرد سپس یکی از راهکاری‌هایی که بنظر بسیار جذاب بنظر می‌رسد افزودن ایده Conformal Prediction به فرایند پیش‌بینی است که اخیراً در حوزه پزشکی توجه بسیاری را به خود جذب کرده است. با این کار می‌توان اثر نویز و خطای وارده از بخش مدل اجماعی را نیز تا حد خوبی خنثی کرد.

- [1] Radiology assistant. <https://radiologyassistant.nl/breast/bi-rads/bi-rads-for-mammography-and-ultrasound-2013>. Accessed: 2024-06-20.
- [2] Nn models. <https://www.geeksforgeeks.org/artificial-neural-networks-and-its-applications/>. Accessed: 2024-06-20.
- [3] J. R. Lee, S. Kim, I. Park, T. Eo, and D. Hwang. Relevance-cam: Your model already knows where to look. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 14944–14953, June 2021.
- [4] A. M. Hafiz and G. M. Bhat. A survey on instance segmentation: state of the art. *International Journal of Multimedia Information Retrieval*, 9(3):171–189, July 2020.
- [5] R. Jiao, Y. Zhang, L. Ding, R. Cai, and J. Zhang. Learning with limited annotations: A survey on deep semi-supervised learning for medical image segmentation, 2022.
- [6] A. Peláez-Vegas, P. Mesejo, and J. Luengo. A survey on semi-supervised semantic segmentation, 2023.
- [7] O. Chapelle, B. Scholkopf, and A. Zien, Eds. Semi-supervised learning (chapelle, o. et al., eds.; 2006) [book reviews]. *IEEE Transactions on Neural Networks*, 20(3):542–542, 2009.
- [8] Y. Grandvalet and Y. Bengio. Semi-supervised learning by entropy minimization. In L. Saul, Y. Weiss, and L. Bottou, editors, *Advances in Neural Information Processing Systems*, volume 17. MIT Press, 2004.
- [9] E. Arazo, D. Ortego, P. Albert, N. E. O’Connor, and K. McGuinness. Pseudo-labeling and confirmation bias in deep semi-supervised learning, 2020.

- [10] A. Iscen, G. Tolias, Y. Avrithis, and O. Chum. Label propagation for deep semi-supervised learning, 2019.
- [11] M. N. Rizve, K. Duarte, Y. S. Rawat, and M. Shah. In defense of pseudo-labeling: An uncertainty-aware pseudo-label selection framework for semi-supervised learning, 2021.
- [12] H. Scudder. Probability of error of some adaptive pattern-recognition machines. *IEEE Transactions on Information Theory*, 11(3):363–371, 1965.
- [13] K. Nigam and R. Ghani. Analyzing the effectiveness and applicability of co-training. In *Proceedings of the Ninth International Conference on Information and Knowledge Management, CIKM '00*, page 86–93, New York, NY, USA, 2000. Association for Computing Machinery.
- [14] Q. Xie, M.-T. Luong, E. Hovy, and Q. V. Le. Self-training with noisy student improves imagenet classification, 2020.
- [15] L. Yu, S. Wang, X. Li, C.-W. Fu, and P.-A. Heng. Uncertainty-aware self-ensembling model for semi-supervised 3d left atrium segmentation, 2019.
- [16] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton. A simple framework for contrastive learning of visual representations, 2020.
- [17] T. Chen, S. Kornblith, K. Swersky, M. Norouzi, and G. Hinton. Big self-supervised models are strong semi-supervised learners, 2020.
- [18] C. You, W. Dai, Y. Min, L. Staib, and J. S. Duncan. Bootstrapping semi-supervised medical image segmentation with anatomical-aware contrastive distillation, 2023.
- [19] K. Chaitanya, E. Erdil, N. Karani, and E. Konukoglu. Local contrastive loss with pseudo-label based self-training for semi-supervised medical image segmentation. *Medical Image Analysis*, 87:102792, 2023.
- [20] H. Wu, Z. Wang, Y. Song, L. Yang, and J. Qin. Cross-patch dense contrastive learning for semi-supervised segmentation of cellular nuclei in histopathologic images. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 11656–11665, 2022.
- [21] X. Zhao, C. Fang, D.-J. Fan, X. Lin, F. Gao, and G. Li. Cross-level contrastive learning and consistency constraint for semi-supervised medical image segmentation, 2022.

- [22] P. Bachman, O. Alsharif, and D. Precup. Learning with pseudo-ensembles, 2014.
- [23] M. Sajjadi, M. Javanmardi, and T. Tasdizen. Regularization with stochastic transformations and perturbations for deep semi-supervised learning, 2016.
- [24] S. Laine and T. Aila. Temporal ensembling for semi-supervised learning, 2017.
- [25] A. Tarvainen and H. Valpola. Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results, 2018.
- [26] H. Zhang, M. Cisse, Y. N. Dauphin, and D. Lopez-Paz. mixup: Beyond empirical risk minimization, 2018.
- [27] V. Verma, K. Kawaguchi, A. Lamb, J. Kannala, A. Solin, Y. Bengio, and D. Lopez-Paz. Interpolation consistency training for semi-supervised learning. *Neural Networks*, 145:90–106, jan 2022.
- [28] Y. Ouali, C. Hudelot, and M. Tami. Semi-supervised semantic segmentation with cross-consistency training, 2020.
- [29] Q. Xie, Z. Dai, E. Hovy, M.-T. Luong, and Q. V. Le. Unsupervised data augmentation for consistency training, 2020.
- [30] X. Luo, W. Liao, J. Chen, T. Song, Y. Chen, S. Zhang, N. Chen, G. Wang, and S. Zhang. Efficient semi-supervised gross target volume of nasopharyngeal carcinoma segmentation via uncertainty rectified pyramid consistency, 2021.
- [31] I. J. Goodfellow, J. Shlens, and C. Szegedy. Explaining and harnessing adversarial examples, 2015.
- [32] P. Wang, J. Peng, M. Pedersoli, Y. Zhou, C. Zhang, and C. Desrosiers. Cat: Constrained adversarial training for anatomically-plausible semi-supervised segmentation. *IEEE Transactions on Medical Imaging*, pages 1–1, 2023.
- [33] T. Lei, D. Zhang, X. Du, X. Wang, Y. Wan, and A. K. Nandi. Semi-supervised medical image segmentation using adversarial consistency learning and dynamic convolution network. *IEEE Transactions on Medical Imaging*, 42(5):1265–1277, 2023.
- [34] T. Miyato, S. ichi Maeda, M. Koyama, and S. Ishii. Virtual adversarial training: A regularization method for supervised and semi-supervised learning, 2018.
- [35] T.-H. Vu, H. Jain, M. Bucher, M. Cord, and P. Pérez. Advent: Adversarial entropy minimization for domain adaptation in semantic segmentation, 2019.

- [36] X. Li, L. Yu, H. Chen, C.-W. Fu, L. Xing, and P.-A. Heng. Transformation consistent self-ensembling model for semi-supervised medical image segmentation, 2020.
- [37] Y. Bai, D. Chen, Q. Li, W. Shen, and Y. Wang. Bidirectional copy-paste for semi-supervised medical image segmentation, 2023.
- [38] X. Luo, J. Chen, T. Song, and G. Wang. Semi-supervised medical image segmentation through dual-task consistency. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 8801–8809, 2021.
- [39] K. Sohn, D. Berthelot, C.-L. Li, Z. Zhang, N. Carlini, E. D. Cubuk, A. Kurakin, H. Zhang, and C. Raffel. Fixmatch: Simplifying semi-supervised learning with consistency and confidence, 2020.
- [40] L. Yang, L. Qi, L. Feng, W. Zhang, and Y. Shi. Revisiting weak-to-strong consistency in semi-supervised semantic segmentation, 2023.
- [41] D. Berthelot, N. Carlini, I. Goodfellow, N. Papernot, A. Oliver, and C. Raffel. Mixmatch: A holistic approach to semi-supervised learning, 2019.
- [42] J. Li, R. Socher, and S. C. H. Hoi. Dividemix: Learning with noisy labels as semi-supervised learning, 2020.
- [43] L.-Z. Guo and Y.-F. Li. Class-imbalanced semi-supervised learning with adaptive thresholding. In K. Chaudhuri, S. Jegelka, L. Song, C. Szepesvari, G. Niu, and S. Sabato, editors, *Proceedings of the 39th International Conference on Machine Learning*, volume 162 of *Proceedings of Machine Learning Research*, pages 8082–8094. PMLR, 17–23 Jul 2022.
- [44] R. Wang, X. Jia, Q. Wang, Y. Wu, and D. Meng. Imbalanced semi-supervised learning with bias adaptive classifier, 2023.
- [45] X. Chen, Y. Yuan, G. Zeng, and J. Wang. Semi-supervised semantic segmentation with cross pseudo supervision, 2021.
- [46] Z. Zhang, Q. Liu, and Y. Wang. Road extraction by deep residual u-net. *IEEE Geoscience and Remote Sensing Letters*, 15(5):749–753, may 2018.
- [47] L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam. Rethinking atrous convolution for semantic image segmentation, 2017.

- [48] J. Chen, Y. Lu, Q. Yu, X. Luo, E. Adeli, Y. Wang, L. Lu, A. L. Yuille, and Y. Zhou. Transunet: Transformers make strong encoders for medical image segmentation, 2021.
- [49] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation, 2015.
- [50] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna. Rethinking the inception architecture for computer vision, 2015.
- [51] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár. Focal loss for dense object detection, 2018.
- [52] J. E. Van Engelen and H. H. Hoos. A survey on semi-supervised learning. *Machine learning*, 109(2):373–440, 2020.
- [53] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer. Automatic differentiation in pytorch. 2017.
- [54] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization, 2017.
- [55] S. S. M. Salehi, D. Erdogmus, and A. Gholipour. Tversky loss function for image segmentation using 3d fully convolutional deep networks, 2017.
- [56] C. H. Sudre, W. Li, T. Vercauteren, S. Ourselin, and M. J. Cardoso. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pages 240–248. Springer International Publishing, 2017.
- [57] Z. Han, M. Jian, and G.-G. Wang. Convunext: An efficient convolution neural network for medical image segmentation. *Knowledge-Based Systems*, 253:109512, 2022.
- [58] A. Damian, T. Ma, and J. D. Lee. Label noise sgd provably prefers flat global minimizers, 2021.

واژه‌نامه

الف

Calibrate تیز سازی مدل
 Uncertainty Estimation تخمین عدم اطمینان
 Constrastive Loss ... تابع هزینه متناقض محاسبه‌گر
 Regularization Loss تابع هزینه تنظیم کننده
 Unsupervised Loss تابع هزینه بدون نظارت
 Convex Combination ترکیب محدب
 Lever Function توابع سطح
 Annotator Function تابع برچسب زد

Inference استنتاج
 Majority اکثریت
 Feature Perturbation اختلال ویژگی
 Hyper parameter فرایارامتر

ب

ح

Adversarial Attack حملات خصمانه

خ

Pipeline خط لوله
 Self-Training خودآموزی
 Self-Attention خودتوجهی

Biopsie بافت برداری
 Mask/Label برچسب
 Optimizer بهینه‌ساز
 Batch بسته های داده
 Hard Label برچسب سخت
 Ground Truth/ True Label برچسب حقیقی
 one-hot بردار تنها یک

پ

د

Data داده
 Epoch دوره
 Stochastic Transformation دگرگونی تصادفی

pixel-wise پیکسل-آگاه
 Pre-Training پیش آموزش

ت

Consistency Regularization تابع هزینه سازگار ساز
 Consistency تنظیم کننده سازگار ساز

ر

Regularization

Vision Transformer..... مبدل بینایی	Encoder..... رمز گذار
	Decoder..... رمز گشا
	Hard Voting..... رای گیری سخت گیرانه

ن

View..... نما

Feature Map..... نقشه ویژگی

Capture Rate..... نرخ ضبط

Learning Rate..... نرخ یادگیری

ش

Pseudo Label..... شبه برچسب

%dicalphabetغ

ض

Weakly Augmented..... ضعیف افزوده شده

ق

Segmentation..... قطعه بندی

Instance Segmentation..... قطعه بندی موردی

Semantic Segmentation..... قطعه بندی معنایی

Strongly Augmented..... قوی افزوده شده

و

Input/Output..... ورودی/خروجی

Feature..... ویژگی

Large-Scale Feature..... ویژگی بزرگ مقیاس

Detailed Feature..... ویژگی جزئی

ک

Craniocaudal..... کرانیوکودال

Calcification..... کلسیم دار

Application..... کاربرد

Epsilon Decay..... کاهش اپسیلونی

ه

Artificial Intelligence..... هوش مصنوعی

ي

Cell..... یافته

Supervised Learning..... یادگیری نظارت شده

Semi-Supervised Learning.. یادگیری نیمه نظارتی

Unsupervised Learning.... یادگیری بدون نظارت

Self-Supervised Learning..... یادگیری خودآگاه

م

Mammography..... ماموگرافی

Mammogram..... ماموگرام

Mediolateral Oblique..... مِدیولاترال اُبلیک

Robust..... مقاوم

Mixture Model..... مدل مخلوط

Ensemble Model..... مدل اجماعی

Abstract

Deep neural networks have made significant advancements in various fields, especially in medicine. Due to their high speed, error resistance, and ability to operate continuously without fatigue, these networks have become valuable tools for assisting physicians. One of the most challenging tasks in medical sciences is the interpretation of mammography images.

The 2D nature of the images, the high diversity of breast tissue, the similarity of lesion tissue to some normal breast tissues, and the difficulty of distinguishing lesion tissue from surrounding tissues are major challenges. Training AI models for this task requires abundant and diverse data with detailed annotations, which is costly and time-consuming. Numerous specialists are needed to interpret mammography images for model training, making it nearly impossible due to the data volume.

In this project, we explore whether integrating information from large unlabeled datasets with small labeled datasets can improve model performance. We introduce a pipeline for automatically interpreting unlabeled data and using these data with semi-supervised learning methods to detect lesion areas in mammography images. We show that combining pipeline interpretations with a novel pixel-wise approach results in an approximately 4% improvement over the baseline. This project also reviews the adaptation of prominent models from previous studies to our novel method.

Keywords: Semi-Supervised Learning, Image Processing, Mammography Image Processing, Image Segmentation, Medical Image Processing



Sharif University of Technology
Department of Computer Engineering

B.Sc. Thesis

Enhancing Injury Segmentation in Breast Mammograms Through Semi-Supervised Learning

By:

Amirhossein Bagheri

Supervisor:

Hamid R. Rabiee

January 2024