



دانشگاه صنعتی شریف
دانشکده‌ی مهندسی کامپیوتر

پایان‌نامه‌ی کارشناسی
مهندسی نرم‌افزار

عنوان:

تشخیص سرطان تیروئید با استفاده از شبکه‌های عصبی از
روی اسلاید‌های سایتولوژی

نگارش:

امیرحسین محسن نژاد

استاد راهنما:

پروفسور حسین اسدی
پروفسور حمیدرضا ربیعی

شهریور ۱۴۰۱

اللهُ أَكْبَرُ

به نام خدا
دانشگاه صنعتی شریف
دانشکده‌ی مهندسی کامپیوتر

پایان‌نامه‌ی کارشناسی

عنوان: تشخیص سرطان تیروئید با استفاده از شبکه‌های عصبی از روی اسلایدهای سایتولوژی

نگارش: امیرحسین محسن نژاد

کمیته‌ی ممتحنین

استاد راهنمای اول: پروفسور حسین اسدی
امضاء:

استاد راهنمای دوم: پروفسور حمیدرضا ربیعی
امضاء:

استاد مشاور:
امضاء:

استاد مدعو:
امضاء:

تاریخ:

سپاس

از جناب دکتر اسدی و جناب دکتر ربیعی بسیار سپاس گزارم که چنین فرصت مطالعاتی را به من دادند. همچنین قدردان و متشرک از راهنمایی‌های سرکار خانم قوامی هستم که در طول مسیر زمان زیادی را برای مشاوره به بنده گذاشتند.

چکیده

سرطان تیروئید یکی از انواع شایع سرطان در بدن انسان‌هاست که در آن غده تیروئید درگیر می‌شود. یکی از راه‌های اصلی تشخیص این بیماری، استفاده از اسلایدهای دیجیتال ثبت شده از نمونه‌های سلولی است. به این صورت که متخصصان با استفاده از ویژگی‌هایی که سلول‌های سرطانی به خود می‌گیرند، می‌توانند آن‌ها را از سلول‌های عادی تیروئید تشخیص دهند و مقدار درگیری نمونه به سرطان را تشخیص دهند. هدف از این پروژه این است که با به کارگیری معماری‌های ResNet و Inception مبتنی شبکه‌های عصبی عمیق پیچشی، مدلی آموزش دهیم که با استفاده از اسلایدها، تشخیصی با دقت بالا انجام دهد. از آنجایی که تشخیص این سرطان به این روش توسط نیروی متخصص بسیار زمان بر است، این مدل می‌تواند دستیار و کمک بسیار خوبی برای آن‌ها باشد. در آخر، نتایج ارزیابی بدست آمده از آموزش مدل‌ها، دقت ۹۴/۹۸ را برای طبقه بند بدخیم-خوش‌خیم و دقت ۹۹/۷۲ را برای طبقه بند سالم-توموری نشان می‌دهند که گویای ظرفیت این روش‌ها برای حل مسائل این چنینی است.

کلیدواژه‌ها: یادگیری ماشین، پردازش تصویر، سرطان تیروئید، بدخیم، اسلایدهای دیجیتال، ResNet، Inception

فهرست مطالب

| | | |
|----|-------|--------------------------------|
| ۱۰ | ۱ | مقدمه |
| ۱۱ | ۱-۱ | تعریف مسئله |
| ۱۱ | ۲-۱ | چالش‌های مسئله |
| ۱۱ | ۳-۱ | اهمیت موضوع |
| ۱۲ | ۴-۱ | ساختار پایان‌نامه |
| ۱۳ | ۲ | مفاهیم اولیه |
| ۱۳ | ۱-۲ | سرطان تیروئید |
| ۱۳ | ۱-۱-۲ | انواع سرطان تیروئید |
| ۱۴ | ۲-۱-۲ | روش‌های تشخیص سرطان تیروئید |
| ۱۶ | ۳-۱-۲ | اسلایدهای دیجیتال |
| ۱۷ | ۲-۲ | شبکه عصبی پیچشی و یادگیری عمیق |
| ۱۸ | ۳-۲ | جمع‌بندی |
| ۱۹ | ۳ | کارهای پیشین |
| ۲۲ | ۱-۳ | جمع‌بندی |
| ۲۴ | ۴ | روش پیشنهادی |

| | | |
|----|--|-------|
| ۲۴ | مجموع داده | ۱-۴ |
| ۲۵ | اطلس ژنوم سرطان تیروئید | ۱-۱-۴ |
| ۲۶ | پاپ سوسایتی اطلس تیروئید | ۲-۱-۴ |
| ۲۶ | پایگاه داده ریزآرایه بافت استنفورد | ۳-۱-۴ |
| ۲۶ | استخراج قسمت‌های حاوی اطلاعات از اسلامید | ۲-۴ |
| ۳۳ | داده افزایی | ۳-۴ |
| ۳۷ | آموزش مدل | ۴-۴ |
| ۳۷ | مدل‌های مورد بررسی | ۱-۴-۴ |
| ۳۸ | روش آموزش مدل‌ها | ۲-۴-۴ |
| ۴۰ | جمع‌بندی | ۵-۴ |
| ۴۱ | نتایج آزمایش‌ها | ۵ |
| ۴۱ | بخش اول، طبقه‌بند بدخیم و خوش‌خیم | ۱-۵ |
| ۴۲ | بخش دوم، طبقه‌بند توموری و سالم | ۲-۵ |
| ۴۶ | جمع‌بندی | ۶ |

فهرست تصاویر

| | | |
|-----|---|----|
| ۱-۲ | نمونه اسلاید اسکن شده از غده تیروئید | ۱۶ |
| ۲-۲ | نمونه اسلاید اسکن شده از غده تیروئید با رنگ امیزی های متفاوت | ۱۷ |
| ۳-۲ | ساختر پایه معماری شبکه عصبی پیچشی | ۱۸ |
| ۱-۴ | نمودار روند انجام کار | ۲۵ |
| ۲-۴ | نمونه ای از مجموع داده ریزآرایه استنفورد تهیه شده از مجموع داده [۱] | ۲۷ |
| ۳-۴ | نمودارهای توزیع تعداد قطعه عکس های تولید شده و درصد قطعه عکس های تولید شده نسبت به تعداد کل | ۲۸ |
| ۴-۴ | تغیرات تابع هدف و آستانه برای فیلتر عکس های فاقد اطلاعات | ۳۰ |
| ۵-۴ | نواحی استخراج شده و ماسک های مرتبط | ۳۱ |
| ۶-۴ | نمونه ای از فیلتر تصاویر تار توسط روش گاوین | ۳۳ |
| ۷-۴ | نمونه ای از داده افزایی تغییر رنگ | ۳۵ |
| ۸-۴ | نمونه ای از داده افزایی تطبیق دامنه فوریه | ۳۶ |
| ۹-۴ | نمونه ای از داده افزایی ادغام تصویر | ۳۷ |
| ۱-۵ | نمودار دقت در هر دوره آموزش مدل بر روی اطلس ژنوم سرطان تیروئید | ۴۳ |

فهرست جداول

| | | |
|-----|---|----|
| ۱-۳ | مروری بر تعدادی از کارهای پیشین به صورت خلاصه | ۲۳ |
| ۱-۴ | دقت روش لایلسانین برای تشخیص نواحی پس زمینه از غیر پس زمینه | ۳۲ |
| ۱-۵ | نتایج آموزش با مدل های مختلف و روش های داده افزایی متفاوت بر روی مجموع دو مجموع داده پاپسوسایتی و ریزآرایه بافت استنفورد | ۴۴ |
| ۲-۵ | خلاصه ای از نتایج آموزش با تنظیمات منتخب شده بر اساس دقیق بر روی دو مجموع داده پاپسوسایتی و ریزآرایه بافت استنفورد | ۴۵ |
| ۳-۵ | خلاصه ای از بهترین نتایج کسب شده توسط دو بخش آموزش | ۴۵ |

فصل ۱

مقدمه

یادگیری ماشین و به خصوص پردازش تصویر امروزه یکی از پر کاربردترین تکنولوژی‌های روز دنیا است که در حوزه‌های مختلف صنعتی، پزشکی، شهری، امنیتی، علمی و فنی کاربرد فراوانی دارد. پردازش تصویر امروزه بیشتر به موضوع پردازش تصویر دیجیتال گفته می‌شود که شاخه‌ای از پردازش سیگنال است که با پردازش سیگنال دیجیتال که نماینده تصاویر برداشته شده با دوربین دیجیتال یا اسکن شده توسط اسکنر هستند سروکار دارد.

تشخیص سرطان تیروئید با استفاده از پردازش تصویر یکی از موضوعات تحقیقاتی است که از مدت‌ها پیش مورد توجه پژوهشگران بوده است. تشخیص این سرطان به معنی تشخیص وجود سلول‌های سرطانی در نمونه مورد بررسی است که ویژگی‌های متفاوتی از سلول‌های عادی دارند.

قبل از ادامه کار باید با دو کلمه در ادبیات پزشکی مربوط به سرطان آشنا شد. بیوپسی^۱ یک آزمایش پزشکی است که معمولاً توسط جراح، رادیولوژیست یا متخصص قلب و عروق انجام می‌شود. این فرآیند شامل استخراج سلول‌ها یا بافت‌های نمونه برای بررسی و تعیین وجود یا وسعت بیماری است. کلمه‌ی دیگر، پتالوژی^۲ علم علل و عوارض بیماری‌ها، به شاخه‌ای از پزشکی می‌گویند که به بررسی آزمایشگاهی نمونه‌های بافت بدن برای اهداف تشخیصی یا پزشکی قانونی می‌پردازد. یکی از راه‌های تشخیصی سرطان تیروئید، بیوپسی غده و بعد از آن بررسی نمونه توسط پتالوژیست^۳ است. پزشک متخصص نمونه را در زیر میکروسکپ مورد بررسی قرار می‌دهد و ناحیه‌های مشکوک را تشخیص می‌دهد. با توجه به اندازه

¹Biopsy

²Pathology

³Pathologist

نمونه و اسلایدها، این فرآیند ممکن است ساعت‌ها به طول بیان‌جامد و فرآیندیست که به دقت بالایی نیازمند است. تصویربرداری کل اسلاید^۴، به عنوان مثال، اسکن و دیجیتالی کردن اسلایدهای بافت شناسی، اکنون در سراسر جهان در آزمایشگاه‌های آسیب شناسی مورد استفاده قرار می‌گیرد. در گذشته راهکارهای ارائه‌شده برای حل این مسئله توسط کامپیوترها و با استفاده از تصویربرداری کل اسلاید، مبتنی بر روش‌های کلاسیک طبقه‌بندی و پیدا کردن ویژگی‌های سلولی به صورت دستی بوده است، اما با روی کار آمدن شبکه‌های عمیق، نتایج بهتری قابل کسب است که به دقت تشخیص و تسريع آن کمک می‌کند.

۱-۱ تعریف مسئله

در این پژوهش، هدف ارائه‌ی راه حلی برای تشخیص سرطان تیروئید و طبقه‌بندی وضعیت آن از روی اسلایدهای دیجیتال کل سایتولوژی است. این موضوع نیازمند یافتن مجموع داده مناسب و آموزش یک مدل بر روی آن است.

۲-۱ چالش‌های مسئله

این پژوهش در مراحل مختلف کار با چالش‌هایی از جمله، نبود داده عمومی کافی، اندازه‌ی بزرگ اسلایدهای کل، ناهمگونی در تصاویر، انواع روش‌های رنگ آمیزی اسلایدها و پیچیدگی ویژگی‌ها روبراست.

۳-۱ اهمیت موضوع

سرطان تیروئید یکی از سرطان‌های شایع بین انسان‌ها است که تشخیص سریع و زودهنگام آن بسیار در درمان غده موثر است. همانطور که پیشتر نیز گفته شد، فرآیند تشخیص سرطان از روی اسلایدهای کل، به دلیل اندازه بزرگ این اسلایدها، فرآیندی بسیار وقت‌گیر است، از این روی، استفاده از فرآیندهای

⁴Whole Slide Image

خودکار مبتنی بر یادگیری ماشین که دقت بالایی را نشان دهنده، از هدر رفت وقت متخصصان جلوگیری می‌کند و از سوی دیگر فرآیند تشخیص را تسريع می‌بخشدند.

۱-۴ ساختار پایان نامه

در فصل اول این پایان نامه به مسئله و مقدمات آن پرداخته شد. فصل دوم شامل مفاهیم اولیه مورد نیاز برای مسئله و راه حل پیشنهادی ارائه شده است. در فصل سوم به کارهای پیشین که در این زمینه انجام گرفته است، اشاره شده است. در فصل چهارم، مجموع دادگان، چالش‌ها، پیش‌پردازش داده‌ها، شبکه‌های عمیق استفاده شده و روش‌های بهبود مورد بحث قرار می‌گیرند. در فصل پنجم، نتایج آموزش مدل‌ها مورد بررسی قرار گرفتند و در آخر مطالب بیان شده در فصل پنجم جمع‌بندی شده است.

فصل ۲

مفاهیم اولیه

در این فصل به بررسی برخی مفاهیم اولیه مورد نیاز در این پایان نامه پرداخته می‌شود.

۱-۱ سرطان تیروئید

غده تیروئید یک غده پروانه‌ای شکل است که در قسمت تحتانی گردن قرار دارد. این غده وظیفه کنترل متابولیسم و سوخت و ساز بدن را به عهده دارد. علاوه بر این، تیروئید هورمون‌هایی تولید می‌کند که وظیفه تنظیم دمای بدن، میزان سوخت و ساز و مصرف اکسیژن را انجام می‌دهد. سرطان تیروئید زمانی رخ می‌دهد که سلول‌های غده تیروئید بر اثر عواملی، دچار تغییر می‌شوند و بعد از آن سلول‌های دارای ناهنجاری شروع به تکثیر می‌کنند و به تدریج تومور تیروئید را تشکیل می‌دهند. سرطان تیروئید در صورتی که زودهنگام تشخیص داده شود یکی از قابل درمان‌ترین انواع سرطان است.

۱-۱-۲ انواع سرطان تیروئید

سرطان تیروئید انواع مختلفی دارد که در هر کدام، ویژگی‌ها سلول متفاوت است.

- سرطان پاپیلاری^۱: این سرطان در ناحیه‌های کوچکی از بدن پخش می‌شود و حدود ۸۰ درصد از نمونه سرطان‌های تیروئید رو تشکیل می‌دهد و در سلول‌های فولیکولی تولید کننده تیروگلوبولین

¹Papillary Carcinomas

در تیروئید شروع می‌شود.

- سرطان فولیکولار^۲: این نوع از سرطان در ناحیه بزرگتری از بدن و در زمان کمتری پخش می‌شود و حدود ۱۵ درصد از نمونه سرطان‌های تیروئید را تشکیل می‌دهد.
- سرطان مدولار^۳: این نوع از سرطان حدود ۳ درصد از جمعیت کل بیماران سرطان غده تیروئید را تشکیل می‌دهد که از سلول‌های تولید کننده کلسی تونین در تیروئید ناشی می‌شود.
- انواع نادر دیگر

هر کدام از انواع سرطان اشاره شده در قسمت اخیر را می‌توان با توجه به سطح خطری که آن‌ها برای بیماران دارند نیز دسته‌بندی کرد. این دسته‌بندی مربوط به نوع سرطان تیروئید نیست و می‌توان سرطان‌های دیگری را نیز با این معیار دسته‌بندی کرد.

- سرطان بدخیم^۴: تومور بدخیم دارای مرزهای نامنظم است، سریعتر از یک تومور خوش‌خیم رشد می‌کند و می‌تواند به سایر قسمت‌های بدن شما نیز سراپت کند.
- سرطان خوش‌خیم^۵: تومور خوش‌خیم دارای مرزهای مشخص، صاف و منظم است، یک تومور خوش‌خیم می‌تواند بسیار بزرگ شود اما به بافت مجاور حمله نمی‌کند.

۲-۱-۲ روش‌های تشخیص سرطان تیروئید

برای تشخیص این سرطان روش‌های بسیاری وجود دارد. از جمله این روش‌ها می‌توان به موارد زیر اشاره کرد.

- معاينه فيزيكى: در اين روش پزشك با معاينه بيمار به صورت فيزيكى و بررسى ناحيه تيروئيد ممکن است متوجه غده سرطاني شود.
- سى تى اسکن^۶: اين آزمایش با استفاده از پرتوهای اشعه ايکس و به منظور تشخيص سايز و ميزان انتشار تومورهای سرطاني تيروئيد در سایر بخش‌های بدن انجام می‌شود. به طور مثال، اگر فردي

²Follicular Carcinomas

³Medullary Carcinoma

⁴Malignant Carcinomas

⁵Benign Carcinomas

⁶Computed Tomography (CT) scan

مشکوک به سرطان تیروئید باشد، سی تی اسکن از ناحیه گردن و قفسه سینه یا قسمت دیگری از بدن مانند شکم نیز انجام می شود.

- بیوپسی: در صورت وجود توده در ناحیه گردن، پزشک با استفاده از یک سوزن نازک، اقدام به نمونه برداری از بافت توده می کند تا وجود احتمالی سرطان را با استفاده از انجام آزمایشات روی این نمونه تشخیص دهد.

- روش های دیگر از جمله آزمایش های خون و ژنتیک، سونوگرافی گردن، اسکن رادیو یودین^۷ و ...

همانطور که اشاره شد، یکی از روش های تشخیص سرطان تیروئید، بیوپسی است. در ادامه به جزیيات بیشتری از این روش پرداخته می شود زیرا در مرحله اول روش استاندارد طلایی تشخیص این سرطان است و همچنین در این پژوهش، تمرکز کلی و ارایه روش های هوش مصنوعی مبتنی بر این روش خواهند بود.

تشخیص سرطان تیروئید به روش بیوپسی

روش اصلی تشخیص این سرطان بیوپسی است. در صورتی که پزشک تشخیص دهد که بیمار نیاز به بیوپسی دارد، اسانترین راه برای تشخیص سرطانی بودن توده، نمونه برداری سوزنی^۸ است. این نوع بیوپسی در بعضی مواقع می تواند، در کلینیک و دفتر پزشک انجام شود.

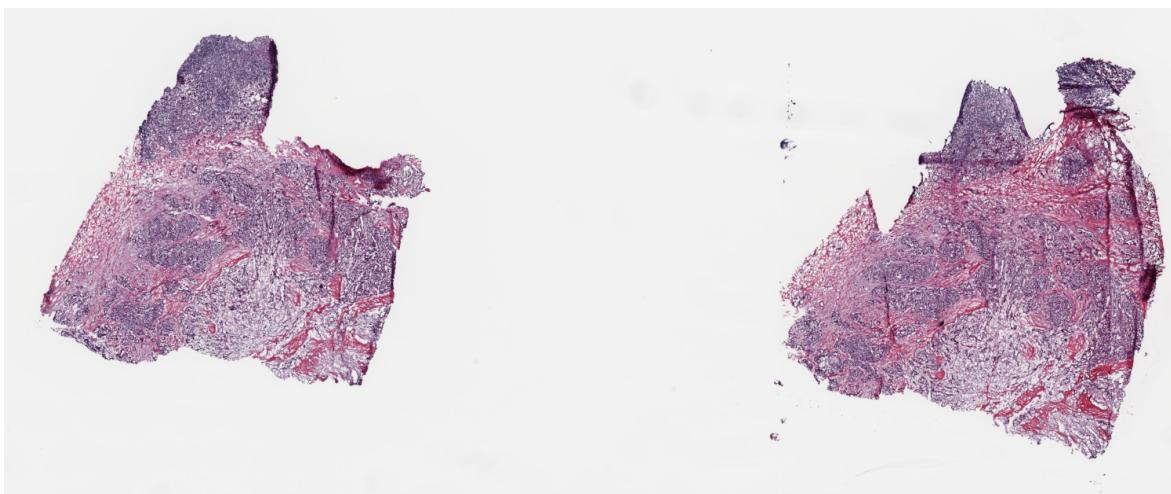
پزشک یک سوزن نازک و توخالی را مستقیماً در غده قرار می دهد تا تعدادی سلول و چند قطره مایع را داخل سرنگ ببریزد. دکتر ممکن است این فرآیند را دو یا سه بار تکرار کند تا از مکان های مختلف غده نمونه برداری کند. نمونه های گرفته شده، سپس در زیر میکروسکوپ توسط متخصصین مورد بررسی قرار می گیرد تا از روی ویژگی های سلول های سرطانی از جمله شکل، اندازه، رنگ، هسته سلول، ساختار سلول ها و دیگر موارد فرآیند تشخیص کامل شود.

برای راحتی کار یا انجام فرآیند بررسی یه صورت ریموت، می توان این نمونه ها را توسط دستگاه هایی اسکن کرد تا اسلاید دیجیتالی از آنها ایجاد شود. به این روش تصویر برداری تمام لغزشی^۹ می گویند. از مزایای دیگر این روش این است که اسلایدها به صورت فایل های دیجیتال در می آیند، که می توانند

⁷ Radioiodine scans

⁸ Fine Needle Aspiration (FNA)

⁹ Whole Slide Imaging



شکل ۲-۱: نمونه اسلايد اسکن شده از غده تیروئید تهیه شده از مجموع داده [۲]

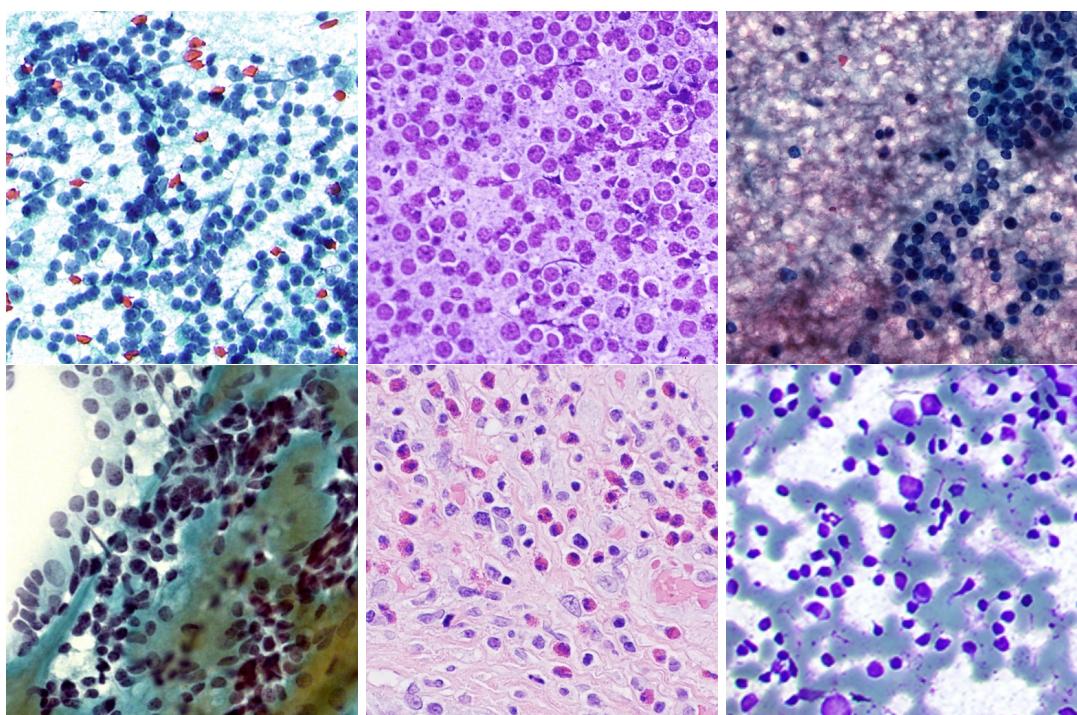
به اشتراک گذاشته شوند و در فرآيندهای آموزشی و تحصیلی و حتی يادگیری ماشین مورد استفاده قرار گيرند.

۳-۱-۲ اسلايدهای ديجيتال

همانطور که در قسمت قبل گفته شد، نمونه های تهیه شده از ناحیه های مشکوک بیمار از غده تیروئید، توسط دستگاه های اسکن می شوند و به صورت تصاویری ديجيتال در می آیند که برای کامپیوتراهای قابل خواندن هستند. قبل از اينکه اين نمونه ها توسط دستگاه اسکن شوند ممکن است فرآيندي روی نمونه ها انجام شود تا سلول ها در زير ميكروسکپ و يا در تصوير ديجيتال به خود رنگ بگيرند. اين کار به متخصصين کمک می کنند تا سلول ها را راحت تر از ديگر مواد داخل نمونه تشخيص دهند و اين باعث افزایش دقت تشخيص می شود. به همين دليل اسلايدهای تهیه شده به اين روش، ممکن است رنگ های مختلفی به خود بگيرند.

نمونه ای از اين اسلايدها در شکل شکل ۱-۲ آمده است که از مجموع داده [۲] برداشته شده که اندازهی تصوير اصلی آن برابر با 30077×14240 پикسل است.

عکس های ديگر از اين نمونه ها که توسط روش های مختلف رنگ آميزي شده اند نيز در شکل شکل ۲-۲ آمده است. منبع اين شش تصوير نيز مجموع داده [۳] است. اين اسلايدها عموماً بعد بسيار بزرگی دارند و ممکن است تا بزرگنمایي ۴۰ برابر را پشتيبانی کنند. به همين دليل ممکن است



شکل ۲-۲: نمونه اسلايد از غده تیروئید با رنگ امیزی های متفاوت تهیه شده از مجموع داده [۳]

حجمی بین چند مگابایت تا چند گیگابایت را به خود بگیرند.

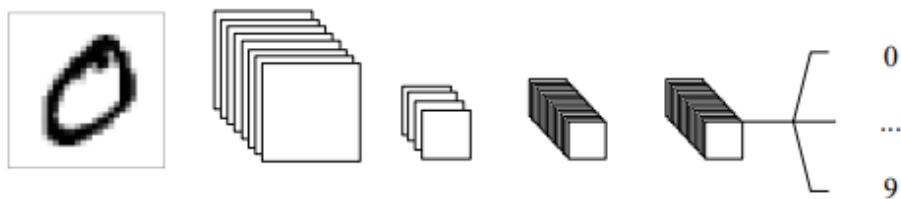
۲-۲ شبکه عصبی پیچشی و یادگیری عمیق

شبکه عصبی پیچشی^{۱۰} یک معماری یادگیری عمیق^{۱۱} است که می‌تواند یک تصویر ورودی را گرفته، به جنبه‌ها یا اشیای مختلف آن تصویر از طریق وزن دهی قابل یادگیری، اهمیت بخشد و بتواند یکی را از دیگری متمایز کند. معماری شبکه‌های عصبی پیچشی مشابه الگوی اتصال نورون‌ها در مغز انسان است و از ساختار قشر بصری الهام گرفته است.

معماری شبکه عصبی پیچشی ابتدا در سال ۲۰۱۵ و در مقاله [۴] معرفی شد که در تصویر ۲-۳ آمده است. ساختار کلی آن، از تعدادی لایه‌های پیچشی تشکیل شده است که در هر لایه تعدادی کanal وجود دارد که هر کدام به طور مستقل از ماتریس‌های هسته آن لایه و خروجی لایه قبل به وجود می‌آیند.

¹⁰Convolutional Neural Network

¹¹Deep Learning



شکل ۲-۳: ساختار پایه معماری شبکه عصبی پیچشی تهیه شده از [۴]

عملکرد این معماری برای گروه‌بندی تصاویر در بسیاری از موارد بالا است و شبکه‌هایی با این معماری وجود دارند که در پردازش تصویر به دقت انسان‌ها رسیده و حتی در مواردی دقیقی بیشتر از انسان را داشته‌اند. به همین دلیل در سال‌های اخیر، معماری پایه‌ی بسیاری از پژوهش‌ها است. این معماری پایه‌ی بسیاری از شبکه‌های عصبی دیگر برای پردازش تصویر است.

۳-۲ جمع‌بندی

در این فصل، به بررسی مفاهیم اولیه در رابطه با سرطان تیروئید و روش‌های مختلف تشخیص آن پرداخته شد. سپس اسلاید‌های دیجیتال معرفی شدند که در این پژوهش به عنوان ورودی اصلی برای آموزش مدل‌ها استفاده می‌شود. در آخر شبکه عصبی پیچشی عمیق معرفی شدند که پایه‌ی بسیاری از شبکه‌های عصبی دیگر برای پردازش تصویر است.

فصل ۳

کارهای پیشین

در سال‌های گذشته برای تشخیص سرطان تیروئید و گروه‌بندی‌های آن کارهایی مبتنی بر یادگیری ماشین انجام گرفته است. این پژوهش‌ها بر روی مجموع داده‌های مختلف و با اهداف متفاوتی شکل گرفته‌اند. در دسته‌ای از این پژوهش‌ها هدف تشخیص کلی وجود یا عدم وجود سرطان تیروئید بوده است. در گروهی دیگر هدف تشخیص نوع سرطان (پاپیلاری، فولیکولار، مدولاری) و زیرگروه‌های هر کدام بوده است. گروهی تمرکز خود را بر روی نوعی خاص از انواع سرطان تیروئید گذاشته و وجود یا عدم وجود آن را در داده‌ها بررسی می‌کنند. هدف تعدادی از این پژوهش‌ها پیشینی معیار TBSR^۱ بوده است. این معیار وضعیت غده تیروئید را به شش گروه در ادامه این فصل، به بررسی تعدادی از کارهای پیشین انجام شده در این زمینه می‌پردازیم.

در سال ۲۰۱۹ پژوهش [۵] در باب این مسئله چاپ شد. هدف اصلی این پژوهش، تشخیص سرطان تیروئید با استفاده از تصاویر سونوگرافیک و مبتنی بر روش‌های شبکه‌های عصبی پیچشی عمیق^۲ بوده است. ورودی مدل، تصاویر معمولی سونوگرافی از غده و خروجی مدل وجود یا عدم وجود ناحیه‌های سرطانی بوده است.

در همان سال در پژوهش آزمایشی [۶] گوان و همکاران، هدف خود را ابتدا طبقه‌بندی سرطان تیروئید به سه گروه فولیکولار ادنوما^۳ و فولیکولار کارسینوما^۴ و پاپیلاری کارسینوما^۵ قرار داده و در مرحله بعد

¹The Bethesda System for Reporting

²Deep Convolutional Neural Network - DCNN

³Follicular Adenoma

⁴Follicular Carcinoma

⁵Papillary Carcinoma

سعی کردند دسته فولیکولار کارسینوما را به چهار زیرگروه دیگر طبقه‌بندی کنند. این تیم، از اسلامیدهای سیتوولوژی دیجیتال استفاده کردند و نتایج خود را بروی هر دو مدل VGG16 و InceptionV3 گزارش کردند. در این پژوهش، از اسلامیدهای تصاویر کوچکتری استخراج شد و از این تصاویر برای آموزش مدل‌ها استفاده شد. در نهایت بهترین نتیجه را با مدل VGG16 و با دقت ۶۶.۹۷ بروی تصاویر استخراج شده کسب کردند.

باز در همین سال، مقاله [۴] منتشر شد که هدف آن تشخیص بدخیم و یا خوشخیم بودن غله تیروئید بوده است. در این پژوهش از تصاویر سونوگرافی استفاده کردند، اما اینبار هر دو دامنه فضایی عکس و دامنه فوریه عکس را مورد بررسی قرار دادند. معماری سیستم توسعه داده شده بدین صورت است که، تصاویر ابتدا وارد طبقه‌بند مبتنی بر فوریه می‌شوند. در این قسمت توزیع مولفه‌های فویر مورد بررسی قرار می‌گیرد و با توجه به آن پیش‌طبقه‌بندی بر روی تصویر انجام می‌شود. در نهایت خروجی قسمت قبل، وارد شبکه عمیقی می‌شود که طبقه‌بندی نهایی را انجام می‌دهد. شبکه عمیق به کار رفته در این قسمت از نوع ResNet تغییر یافته بوده که جزیيات بیشتری در مورد معماری آن در فصل‌های بعدی آمده است.

در پژوهش دیگر [۵] که در سال ۲۰۲۰ میلادی منتشر شد، دنیل و همکاران خود سیستمی را توسعه دادند تا معیار TBSR را از روی اسلامیدهای دیجیتال پیشیبینی کنند. این تیم در ابتدا با استفاده از شبکه VGG11 مدلی را آموزش دادند تا ناحیه‌های مورد نظر^۶ یک اسلامید را تشخیص دهد. برای این کار اسلامیدهای را به صورت تصادفی به ناحیه‌های مختلف تقسیم کرند و مدل را بر روی این تصاویر آموزش دادند. سپس باز با استفاده از شبکه VGG11 شبکه‌ی مستقل دیگری را آموزش دادند تا با به کارگیری ناحیه‌های مورد نظر بدست آمده، TBSR نمونه را پیشیبینی کند.

در سال ۲۰۲۱ نیز در پژوهش [۶] سعی بر این شد مدلی برای تشخیص سرطان تیروئید از نوع پاپیلاری توسعه داده شود. در این مسیر، روش‌های مختلف و مدل‌های مختلفی مورد بررسی قرار گرفتند و مقایسه‌ای از عملکرد هر یک گزارش شد. دو روشی که در این پژوهش مورد بررسی قرار گرفتند به صورت زیر هستند.

- طبقه‌بندی بر اساس ویژگی

قدم اول این روش، مرزبندی^۷ است که در آن، هدف مشخص کردن ناحیه‌های مربوط به هسته و سلول است. برای بدست آوردن مرزبندی مناسب از شبکه UNet استفاده شد. این شبکه ابتدا در

⁶Region Of Interest

⁷Segmentation

سال ۲۰۱۵ و در مقاله [۱۰] معرفی شد. این شبکه، یک معماری رمزگذار-رمزگشای U شکل است که از چهار بلوک رمزگذار و چهار بلوک رمزگشا تشکیل شده است. پلهایی بین رمزگذار و رمزگشا وجود دارد که آنها را به هم متصل می‌کند و جریان اطلاعات را کامل می‌کند. این اتصالات منجر به الحق ویژگی‌های متفاوت از نظر معنایی می‌شود که عملکرد شبکه را برای مرزبندی بهبود می‌دهد. برای فرآیند استخراج ویژگی، تصاویر تیروئید، مرزبندی‌های مرتبط و ماسک‌های تهیه شده از سلول‌ها استفاده شده است تا ویژگی‌های مختلف سلول بحسب بیانند. در مرحله بعد، ویژگی‌های بدست آمده جمع می‌شوند و روی آنها پیش‌پردازش انجام می‌گیرد تا مورد استفاده مدل برای طبقه‌بندی قرار گیرند. الگوریتم‌های متفاوت و کلاسیکی مانند طبقه‌بندی بردار پشتیبانی، نزدیک‌ترین همسایه، بیز ساده گوسی و دیگر موارد برای این کار مورد آزمایش قرار گرفت که در نهایت بهترین آنها انتخاب گردید.

• طبقه‌بندی مبتنی بر یادگیری عمیق

این روش، روش مستقیمی برای طبقه‌بندی به حساب می‌آید، به همین دلیل نیاز به استخراج ویژگی به صورت دستی نیست و قدم‌ها مورد نیاز برای طبقه‌بندی به شدت کاهش پیدا می‌کنند. باید توجه شود که در روش اول، از یادگیری عمیق تنها برای مرزبندی تصاویر استفاده شد و برای طیقه‌بندی نهایی، الگوریتم‌های کلاسیک به کار گرفته شدند، از این جهت پیچیدگی مرزبندی روش اول به تنها ی با پیچیدگی روش دوم برابر می‌کند. برای کاهش زمان آموزش مدل‌ها نیز، از مدل‌های از پیش آموزش داده شده بر روی مجموع داده ImageNet [۱۱] استفاده شد. در نهایت قبل از استفاده از داده‌ها، روش‌های مختلف از جمله چرخش تصویر، آینه کردن، متعادل سازی هسیتوگرام تصویر^۸، تطبیق دامنه فوریه و دیگر موارد برای داده‌افزایی به کار گرفته شد.

بهترین نتیجه‌ای که توسط این پژوهش گزارش شد، از روش مبتنی بر ویژگی و با دقت‌های ۸۳ و ۸۹ بر روی دو مجموع داده بدست آمد. در هر دو مجموع داده دقت روش مبتنی بر ویژگی بالاتر از روش یادگیری عمیق بوده، اما در پایان، پیشنهاد شد که برای مجموع داده‌های کوچک تر و مرزی تر (منظور از مرزی بودن داده در اینجا این است که به سادگی نمی‌توان به گروه‌های مورد نظر طبقه‌بندی کرد) از روش مبتنی بر ویژگی و برای مجموع داده‌های بزرگ تر و کلی تر از روش مبتنی بر یادگیری عمیق استفاده گردد.

برای کسب دیدگاه و مقایسه کلی کارهای پیشین، تعدادی از آنها در جدول ۱-۳ نیز ارائه شده‌اند.

^۸Contrast Limited Adaptive Histogram Equalization

۱-۳ جمع‌بندی

در این فصل، چندین نمونه از کارهای پیشین انجام شده در این زمینه مورد بررسی قرار گرفتند و توضیح مختصری درباره هدف اصلی، ایده، نکات مثبت و منفی و یا نتایج آنها داده شده است تا درک بهتری از سایر راه حل‌های پیشنهادی در حل این مسئله بدست آید.

جدول ۱-۳: مروری بر تعدادی از کارهای پیشین به صورت خلاصه

| نکات مثبت و منفی | ایده اصلی | هدف اصلی | سال | مرجع |
|---|--|---|------|------|
| - وجود داده‌هایی ناقص و تصاویر خارج از فوکس | آموزش مدل برای طبقه‌بندی چند نوع سرطان | تشخیص سرطان‌های مربوط به نواحی سر و سینه | ۲۰۱۹ | [۱۲] |
| نکات برجسته‌ای نداشته صرفا مجموعه داده خود را بر روی مدل خود آموزش داده و ارزیابی کرده است | روش‌های مبتنی بر شبکه‌های عمیق عصبی | تشخیص سرطان تیروئید با استفاده از تصاویر سونوگرافیک | ۲۰۱۹ | [۵] |
| + استفاده از اطلاعات تصاویر در دامنه فوریه | استفاده روش طبقه‌بندی مبتنی بر دامنه فوریه تصاویر و ادغام این روش با طبقه‌بندی با شبکه تغییر یافته ResNet | تشخیص بدخیم و یا خوشخیم بودن غده تیروئید با استفاده از تصاویر سونوگرافی | ۲۰۱۹ | [۷] |
| - استفاده از داده‌های همگون که توزیع نزدیک بهم دارند ممکن است در تعمیم پذیربودن مدل موثر است. | آموزش دو مدل از هم از هم مستقل از نوع VGG11 یکی برای پیدا کردن نواحی مورد نظر و دیگری برای پیش‌بینی معیار TBSR | پیش‌بینی معیار TBSR بر روی اسلامیدهای دیجیتال سیتولوژی | ۲۰۲۰ | [۸] |
| + استفاده از روش‌های متنوع داده‌افزایی از جمله تطبیق دامنه فوریه + استفاده از شبکه UNet برای مرزبندی | مقایسه دو روش ویژگی محور و روش مبتنی بر شبکه عمیق | تشخیص سرطان تیروئید از نوع پاپیلاری از روی اسلامیدهای دیجیتال | ۲۰۲۱ | [۹] |
| - استخراج بریده عکس‌های اسلامیدها، پیش از تقسیم آن‌ها به سه دسته آموزش، ارزیابی و تست می‌تواند یک نکته منفی این پژوهش باشد. | شبکه اول انواع سه‌گانه سرطان تیروئید را تشخیص می‌دهد و شبکه دوم وظیفه تشخیص فولیکولار کارسینوما را به عهده دارد. | تشخیص انواع مختلف سرطان تیروئید با استفاده از دو شبکه عصبی پیچشی آبشاری | ۲۰۲۱ | [۱۳] |

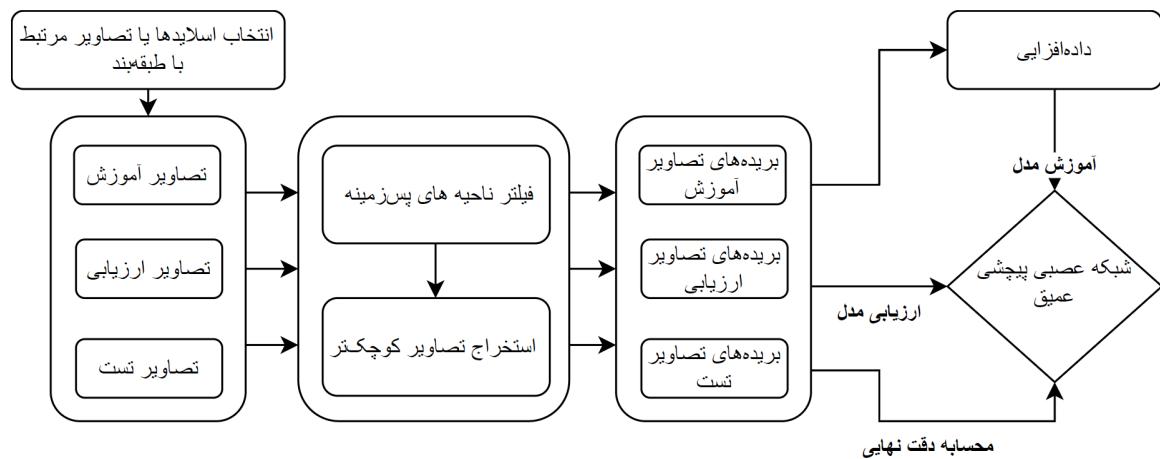
فصل ۴

روش پیشنهادی

روش استفاده شده برای حل مسئله در تصویر ۱-۴ ارائه شده است. در ادامه فصل به توضیح بیشتر قسمت‌های آن پرداخته می‌شود.

۱-۴ مجموعداده

یکی از چالش‌هایی که در حل این مسئله با آن روبرو هستیم، نبود مجموعداده مناسب و با دسترسی عمومی است و گام اول برای این پژوهش به حساب می‌آید. در سال ۲۰۱۰ مقاله [۱۲] منتشر شد که هدف آن تشخیص سرطان با به کارگیری اسلایدهای ناحیه‌های سر و گردن بوده است. این پژوهش، بعد از آموزش مدل بر روی مجموعداده خصوصی تهیه شده، سعی شد تا برای ارزیابی خارجی، مدل را بر روی مجموعداده عمومی ۲۰۱۶ CAMELYON تست کنند. این مجموعداده در چالشی در سال ۲۰۱۶ منتشر شد که شرکت کنندگان باید از آن برای حل مسئله استفاده می‌کردند. هدف اصلی این چالش تشخیص سرطان سینه بوده و مجموعداده نیز شامل اسلایدها و برچسب‌هایی مربوط به این سرطان بوده است. در مقاله‌ای دیگر [۹] مجموعداده عمومی از اسلایدهای دیجیتال و برچسب‌های آن تهیه شد. هدف آن تشخیص پاپیلاری بودن یا نبودن غده تیروئید بوده است که علیرغم اعلام نویسنده، موفق به تماس برای دریافت مجموعداده نشدیم. با توجه به گفته‌های اخیر، جمع آوری داده یکی از بخش‌های مهم این پروژه است تا با آن بتوان مدل را به خوبی آموزش داد و عملکرد خوبی را دریافت کرد. در ادامه به مجموعدادهایی که تهیه شد و درنهایت برای آموزش و ارزیابی مدل به کار گرفته شد اشاره می‌کنیم.



شکل ۴-۱: نمودار روند انجام کار در این شکل ارائه شده است. برای دو مجموع داده پاپ سوسایتی و ریزآرایه استنفورد تصاویر با برچسب های بد خیم و خوش خیم انتخاب شدند و برای مجموع داده اطلس ژنوم سرطان تیروئید اسلاید هایی که صدرصد سلول های توموری یا صدرصد سلول های سالم را داشتند انتخاب شدند.

۱-۱-۴ اطلس ژنوم سرطان تیروئید

اطلس ژنوم سرطان تیروئید^۱ بخشی از یک تلاش بزرگتر برای ایجاد یک جامعه تحقیقاتی متمرکز بر اتصال فنوتیپ های^۲ سرطان به ژنوتیپ ها^۳ با ارائه تصاویر بالینی منطبق با افراد از اطلس ژنوم سرطان [TCGA]^[۱۴] است. این پروژه شامل اطلاعات ۵۰۷ بیمار است که هدف اصلی آن، بررسی سرطان تیروئید در این بیماران است. این اطلاعات شامل توالی دی ان ای، جنسیت، سن، داده های بالینی از بررسی بیمار توسط پزشک، اسلاید های دیجیتال نمونه برداری شده از غده، درصد سلول های توموری و سالم و بسیاری موارد دیگر است. تعداد کل اسلاید های جمع آوری شده مورد استفاده ما از این پروژه ۶۳۹ عدد است که در مجموع حجمی بالغ بر ۱۰۵ گیگابایت دارند. برای هر کدام از این اسلاید ها سه برچسب درصد سلول های عادی، درصد سلول های سرطانی و درصد سلول های استورمال آمده است که ما نیز از این برچسب ها برای مسئله خود استفاده می کنیم. از بین این اسلاید ها و با توجه به درصد سلول های نرمال، توموری و استورمال، تعداد کل ۱۸۲ اسلاید انتخاب شد که در بین آن ها ۱۰۶ اسلاید صدرصد توموری

¹The Cancer Genome Atlas-Thyroid Cancer/TCGA-THCA

²Phenotypes

³Genotypes

و ۷۶ اسلاید آن، صد درصد عادی است. نمونه‌ای از اسلاید این مجموع داده در ۱-۲ و در فصل قبل آمده است. در صورت نیاز، جزئیات بیشتری از این مجموع داده در آدرس [۲] قرار دارد.

۲-۱-۴ پاپ سوسایتی اطلس تیروئید

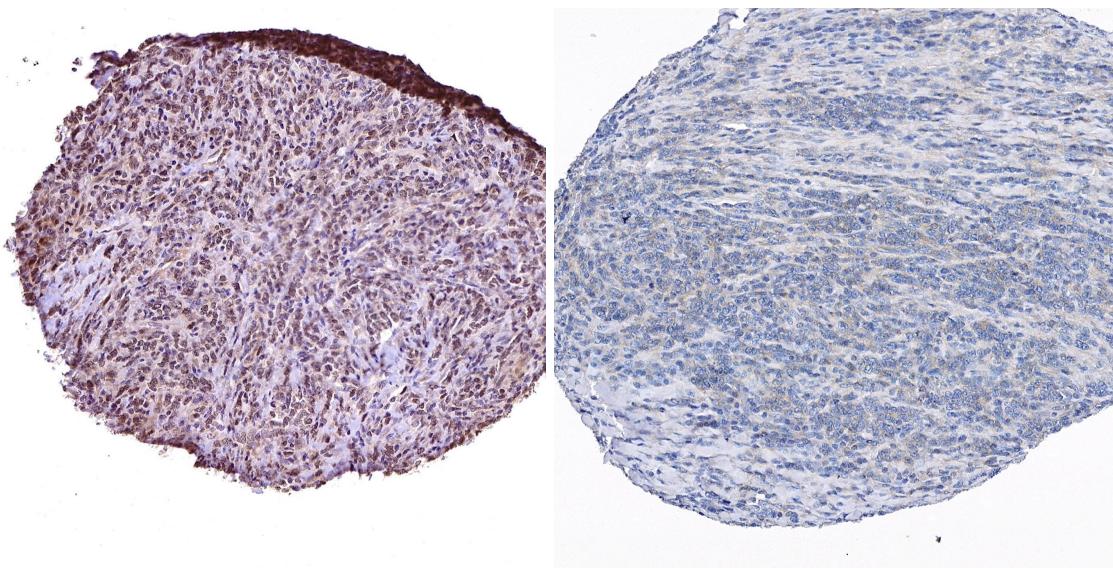
پاپ سوسایتی، انجمنی است که به عنوان یک شرکت اداره می‌شود و منحصراً به عنوان یک مؤسسه خیریه معاف از مالیات برای سازمان‌های علمی آموزشی، مدنی و خیریه سازماندهی شده است. اطلس تیروئید این انجمن [۳]، شامل نزدیک به ۳۰۰ تصویر از غده تیروئید و نواحی مختلف آن از اسمیر مستقیم یک نمونه برداری سوزنی است که تمامی تصاویر به طور مستقل توسط متخصصین بررسی و برچسب گذاری شده‌اند. تصاویر توسط اسکریپتی از صفحات وب این اطلس جمع آوری شد و که هر کدام از تصاویر ابعادی بین 1485×1752 و 360×525 را دارا هستند. مجموع تصاویر این مجموع داده حجمی ۱۲۵ ماگابایتی دارند که از بین این تصاویر، ۲۶۹ عکس به عنوان، بدخیم و خوشخیم بچسب زده شده‌اند. نمونه‌هایی از تصاویر این مجموع داده در ۲-۲ و در فصل قبل آمده است.

۳-۱-۴ پایگاه داده ریزآرایه بافت استنفورد

پایگاه داده ریزآرایه بافت استنفورد [۱] بخش تیروئید، شامل ۲۲۵ تصویر است که تصاویر و برچسب‌های آن توسط اسکریپتی از آدرس [۱۵] استخراج شدند. هر کدام از تصاویر ابعادی بین 2478×2525 و 1440×1504 را دارا هستند و در مجموع حجم این مجموع داده ۳۵۲ مگابایت است. دو نمونه از تصاویر این مجموع داده در ۲-۴ آمده است.

۲-۴ استخراج قسمت‌های حاوی اطلاعات از اسلاید

اسلایدهایی که متخصصان از آن برای تشخیص استفاده می‌کنند، ابعاد بزرگی در حد چندده هزار پیکسل را دارند، به طوری که به هیچ‌وجه در حافظه پردازنده‌های رایج جای نمی‌گیرد. برای حل این مشکل در این پروژه سعی بر این شد که اسلایدها به بریده‌های کوچک‌تر 512×512 عکس شکسته شوند به طوری که قابل پردازش باشند. برای پوشش سلول‌هایی که در ناحیه‌های مرزی عکس و به صورت ناقص قرار می‌گیرند، ناحیه‌ها همپوشانی ۱۰ درصد را با یکدیگر دارند. یکی از موضوعات مهمی که باید در



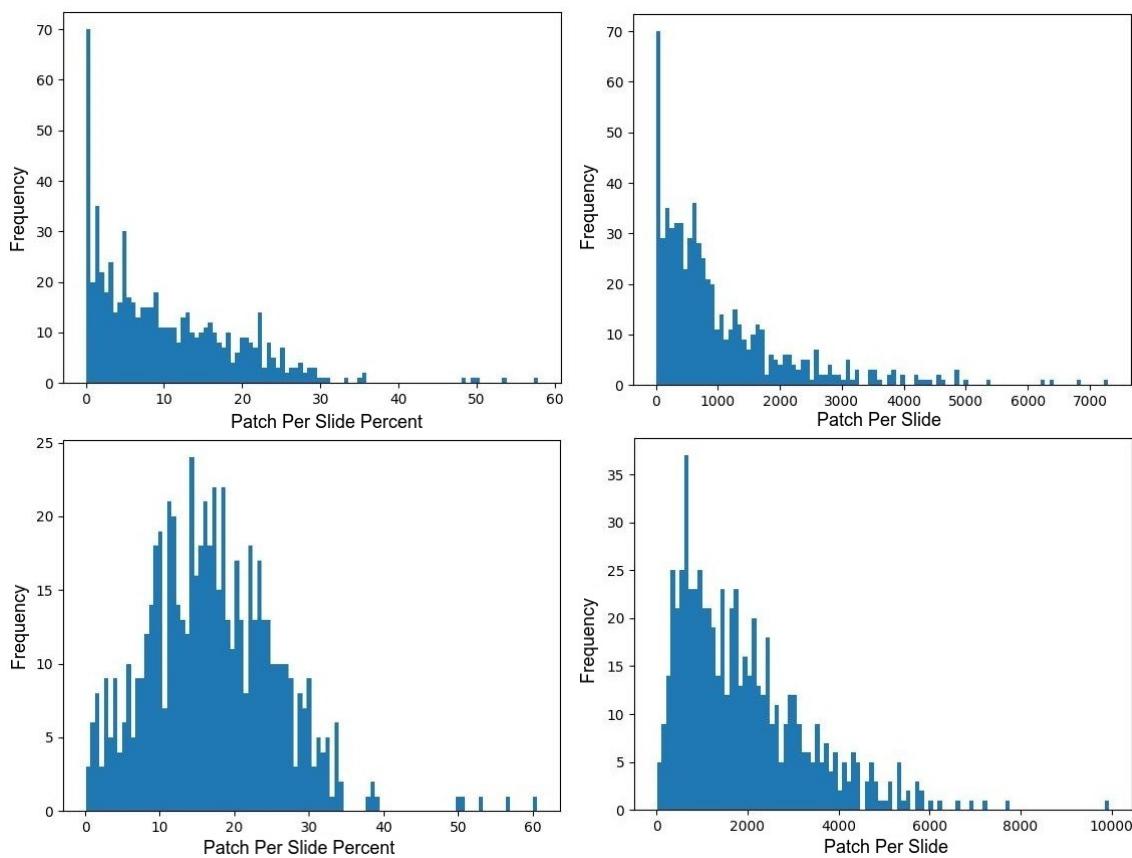
شکل ۲-۴: نمونه‌ای از مجموع داده ریزآرایه استنتفورد تهیه شده از مجموع داده [۱]

نظر گرفت، ناحیه‌های غیر مهم و فاقد اطلاعات است که باید در این فرآیند حذف گردند. این کار به دو دلیل انجام می‌شود. اولاً اینکه این قسمت‌ها، بخش بزرگی از اسلایدها را تشکیل می‌دهند و با حذف این موارد می‌توان در استفاده از منابع پردازشی صرفه جویی کرد و دوماً اگر تعداد زیادی از عکس‌هایی که هنگام آموزش مدل استفاده می‌کنیم، از این نوع باشند، دقت مدل کاهش پیدا می‌کند و مدل در فرآیند یادگیری با مشکل مواجه می‌شود. دلیل این امر هم این است که این عکس‌ها حاوی ویژگی‌های مورد نظر ما برای تشخیص سلول‌های سرطانی نیستند و مدل در حین فرآیند آموزش، ویژگی‌های نامرتبه را از روی این عکس‌ها یاد می‌گیرد.

برای حل این موضوع، روشی که به کار گرفته شد، استفاده از واریانس لابلسین ناحیه است. لابلسین یک عکس از محاسبه مشتق دوم روی شدت رنگ‌های پیکسل‌های آن محاسبه می‌شود و در نتیجه لبه و گوشه‌های عکس مقدار بیشتری می‌گیرد. برای محاسبه مشتق دوم برای یک عکس از رابطه ۱-۴ استفاده شده است که در آن $Im(x)$ مقدار شدت رنگ را در موقعیت x از تصویر نشان می‌دهد.

$$\begin{aligned} Im'(x) &= Im(x+1) - Im(x), \quad Im'(x+1) = Im(x+2) - Im(x+1) \\ Im''(x) &= Im(x+2) - 2 \times Im(x+1) + Im(x) \end{aligned} \quad (1-4)$$

بعد از محاسبه واریانس شدت رنگ پیکسل‌ها در لابلسین عکس و استفاده از یک آستانه عکس‌هایی که مقدار کمتری دارند فیلتر می‌شوند. به این ترتیب ناحیه‌های فاقد اطلاعات که مقدار لابلسین کمی



شکل ۳-۴: نمودارها به ترتیب از راست به چپ توزیع تعداد قطعه عکس‌های تولیدشده از هر اسلاید و درصد قطعه عکس‌های تولیدشده نسبت به تعداد کل را برای هر اسلاید به ازای آستانه اولیه ۵۰۰ و آستانه نهایی بدست آمده ۲۹۸ در قسمت پایانی را نشان می‌دهد.

نیز دارند حذف می‌شوند.

برای بدست آوردن مقدار بهینه آستانه و همچنین آزمودن این روش، به ترتیب زیر عمل شده است.

۱. ابتدا حدس اولیه ۵۰۰ برای آستانه انتخاب شد و اسلایدها با توجه به این آستانه به قطعه‌های کوچک عکس درآمدند. مقدار اولیه آستانه در این مرحله، نیاز به دقت بالایی ندارد، زیرا همانطور که در قسمت بعد نیز توضیح داده خواهد شد، از آن برای انتخاب اسلایدها استفاده می‌کنیم. این حدس اولیه نیز با مشاهده مقادیر به صورت عملی و از روی نواحی چند اسلاید انتخاب شد. توزیع تعداد و درصد قطعه عکس‌های مجموع اسلایدها در تصویر ۳-۴ آمده است.

۲. سپس سه اسلاید از پایین بازه و سه اسلاید از بالای بازه مشخص شده در نمودار توزیع قطعه

عکس‌ها انتخاب شدند. سه اسلاید بالای بازه اسلامیدهایی هستند که روش به کار رفته تعداد زیادی قطعه عکس از آن‌ها تولید کرده است و سه اسلاید پایین بازه اسلامیدهایی هستند که به دلیل نابهینه بودن آستانه، قطعه عکس‌های کمی را برای آن‌ها تولید کرده است.

۳. برای شش اسلاید انتخاب شده، با استفاده از نرم افزار [۱۶] GIMP ماسک‌هایی تولید شد. در این مرحله، بعد از بارگذاری اسلایدها در نرم افزار، با استفاده از ابزارهای فراهم شده در آن و به صورت دستی، نواحی مورد نظر انتخاب شدند و در آخر در قالب تصویری جدا ذخیره شدند.

۴. در قدم بعد، باید معیارهایی را برای روش به کار رفته انتخاب کرد تا با استفاده از آن‌ها بتوان آستانه نابهینه را پیدا کرد و در نهایت آن را ارزیابی کرد. همانطور که در قسمت‌های قبل گفته شد، فیلتر ناحیه‌های فاقد اطلاعات، اهمیت زیادی برای ما دارد از این رو، در ارزیابی این روش علاوه بر دقت^۴، صحبت^۵ نیز عامل مهمی در کارکرد درست است. صحبت، با توجه به رابطه ۲-۴ هرچه به مقدار عددی ۱ نزدیک‌تر باشد به این معناست که ناحیه‌های بدست آمده از این روش، به احتمال بالاتری دارای اطلاعات هستند و در نتیجه ناحیه‌های فاقد اطلاعات کمتری تولید می‌شوند.

$$Precision = \frac{TP}{TP + FP}$$

روش به درستی عکس را غیرپس زمینه تشخیص داده : $TruePositive(TP)$

روش به درستی عکس را پس زمینه تشخیص داده : $TrueNegative(TN)$ (۲-۴)

روش اشتباه^۶ عکس را غیرپس زمینه تشخیص داده : $FalsePositive(FP)$

روش اشتباه^۷ عکس را پس زمینه تشخیص داده : $FalseNegative(FN)$

دقت روش نیز با توجه به رابطه ۳-۴ محاسبه می‌شود.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (۳-۴)$$

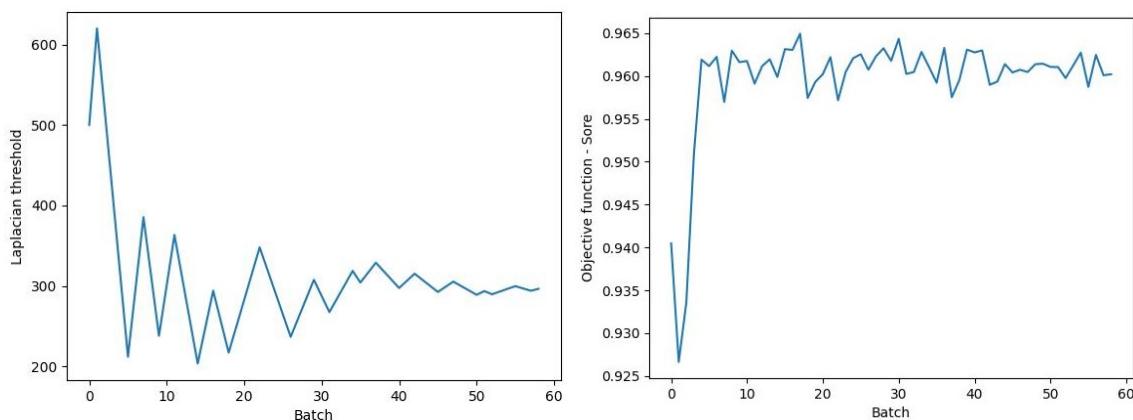
۵. حال با توجه به این دو معیار، قطعه کدی نوشته شد که شش اسلاید و ماسک‌های مربوط به آن‌ها را به عنوان ورودی می‌گیرد و با شروع از آستانه ۵۰۰ و محاسبه ماتریس درهم ریختگی^۸، در جهتی آستانه را تغییر می‌دهد تا دو معیار گفته شده بیشینه شوند. لازم به ذکر است، در نهایت از تابع هدف^۹ $0.25 \times Precision + 0.75 \times Accuracy$ آستانه استفاده شد، که

⁴Accuracy

⁵Precision

⁶Confusion Matrix

⁷Objective Function



شکل ۴-۴: به ترتیب از راست به چپ، تغییرات تابع هدف و آستانه قابل مشاهده است که در نهایت تابع هدف به مقدار ۹۶/۰ و آستانه به ۲۹۸ همگرا شده است.

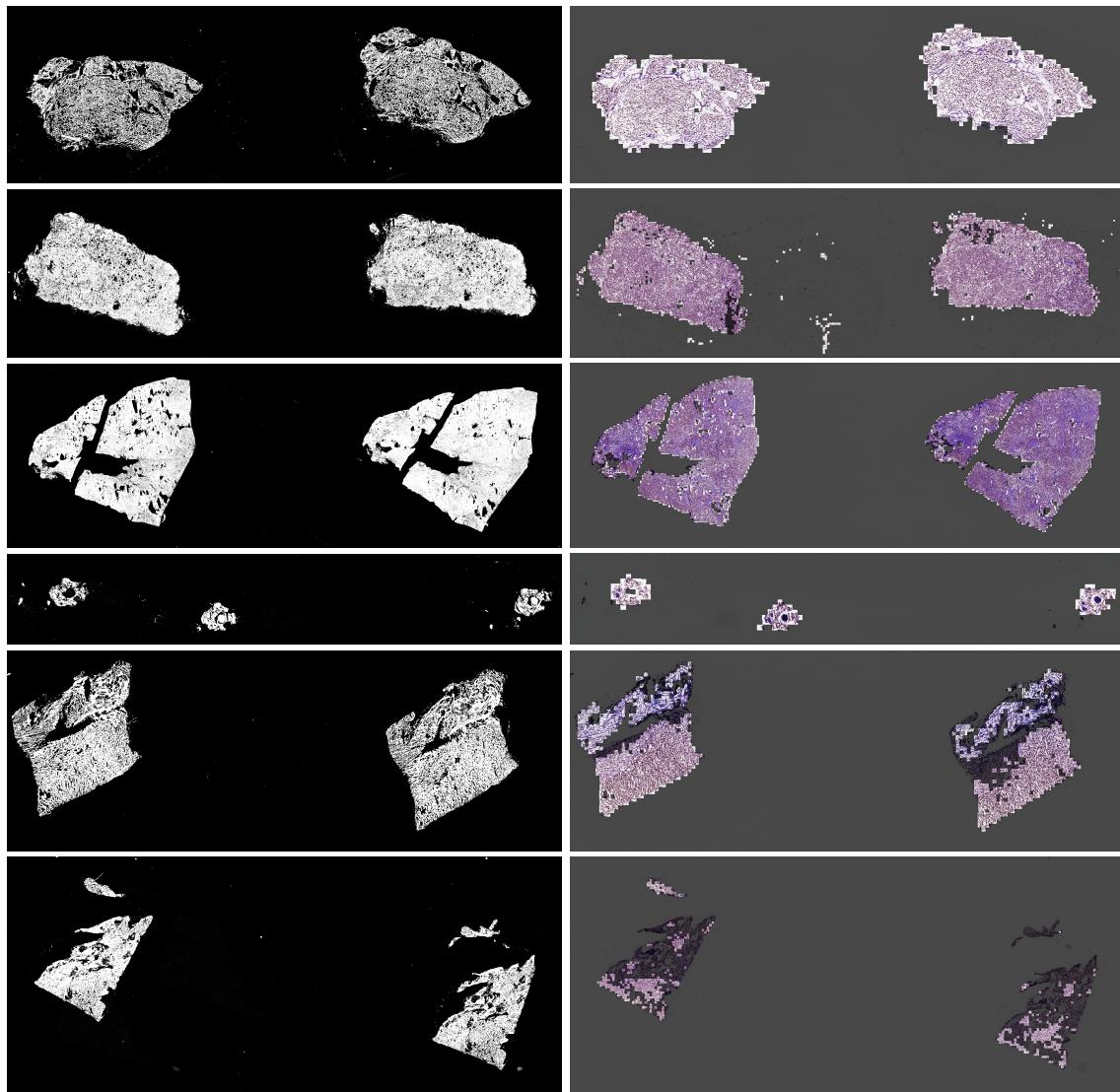
در آن ضرایب طوری انتخاب شدند که اولاً "مجموع آن‌ها یک شود و دوماً"، با توجه به اهمیت این دو معیار دقت و صحبت، آن‌ها را بشیئنه کند. در هر دوره از اجرای کد، از هر اسلاید ۲۰۰۰ (با توجه به اینکه درصد زیادی از اسلایدها را پس‌زمینه تشکیل می‌دهد، عدد بزرگی انتخاب شد تا آستانه در هر دور، در جهت درستی حرکت کند) و در مجموع 6×2000 قطعه عکس مورد بررسی قرار می‌گیرد به صورتی که آستانه در جهت افزایش تابع هدف و با اندازه پرش^۸ ۱۲۰ و نرخ نزولی^۹ ۸۵/۰، کاهش و یا افزایش پیدا می‌کند. نمودار تغییرات تابع هدف و آستانه در طی اجرای برنامه در تصویر ۴-۴ آمده است که در آن‌ها تابع هدف و آستانه به ترتیب به مقادیر ۹۶/۰ و ۲۹۸ همگرا شده‌اند. اسلایدها و ماسک‌های به کار رفته آن‌ها در تصویر ۵-۴ ارائه شده است و دقت و حساسیت روش بر روی این شش اسلاید نیز در جدول ۱-۴ آمده است.

بعد از یافتن آستانه، می‌توان اسلایدها را یکی پس از دیگری خواند و نواحی مورد نیاز برای آموزش مدل را استخراج کرد. با توجه به توضیحات داده شده برای این روش استخراج، در ادامه به دو تا از مزیت‌های آن اشاره می‌کنیم:

- از آنجایی که حجم داده‌ای که با آن‌ها کار داریم بسیار زیاد است و اسلایدهای زیادی نیاز به پیش‌پردازش دارند سرعت و دقت عامل مهمی در عملکرد است. در این روش، بعد از تهیه

⁸Jump Size

⁹Decay Rate



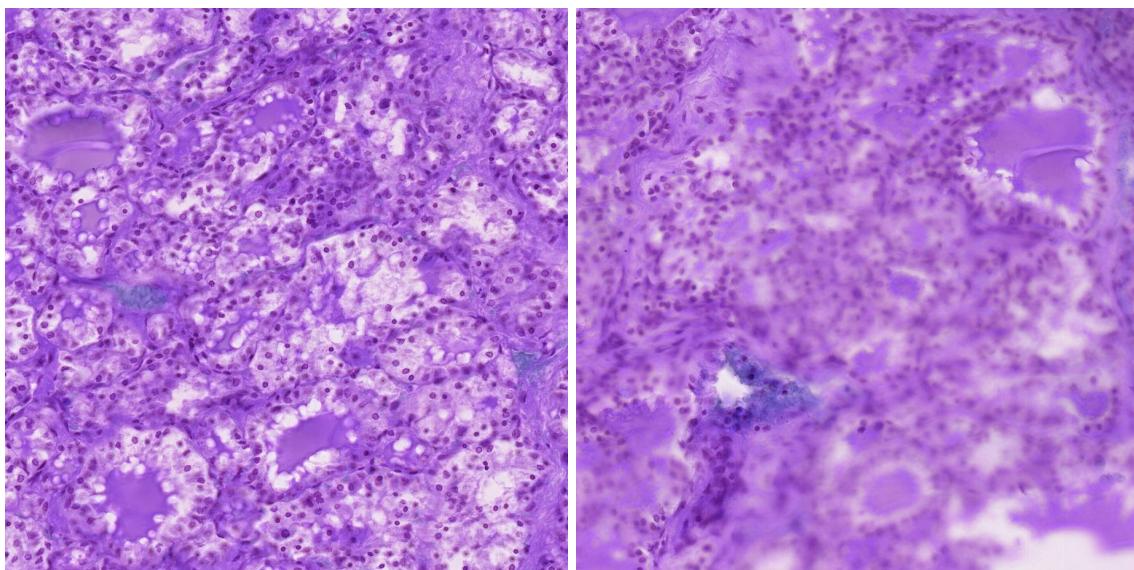
شکل ۴-۵: در این تصویر، سطر یکم تا ششم به ترتیب، در سمت راست، نواحی استخراج شده شش اسلاید ۱، ۲، ۳، ۴، ۵، ۶ توسط روش پیشنهادی در آستانه ۲۹۸ و در سمت چپ، ماسک‌های مرتبط با آنها را که به صورت دستی تهیه شده‌اند، نشان می‌دهد.

جدول ۱-۴: ستون اول این جدول، شماره اسلاید را مشخص می‌کند. در ستون دوم ماتریس درهم ریختگی روش لپلاسین بر روی اسلاید آن آمده است. ستون و چهارم به ترتیب دقت و صحت روش ذکر شده برای فیلتر کردن نواحی فاقد اطلاعات را نشان می‌دهد.

| اصالید | ماتریس درهم ریختگی | دقت | صحت |
|--------|--------------------------------------|------|------|
| ۱ | TP: ۴۶۲۴ FP: ۱۲۶ TN: ۱۴۱۰۰ FN: ۲۲۶ | ۰/۹۸ | ۰/۹۷ |
| ۲ | TP: ۱۱۰۳ FP: ۱۷۲ TN: ۳۷۵۸ FN: ۷ | ۰/۹۶ | ۰/۸۶ |
| ۳ | TP: ۷۶۱۵ FP: ۹۲ TN: ۲۰۸۷۱ FN: ۲۳۴ | ۰/۹۸ | ۰/۹۸ |
| ۴ | TP: ۷۸ FP: ۱۸ TN: ۱۸۸۰ FN: ۴ | ۰/۹۸ | ۰/۸۱ |
| ۵ | TP: ۱۱۳۸ FP: ۴ TN: ۶۶۷۱ FN: ۴۹۲ | ۰/۹۴ | ۰/۹۹ |
| ۶ | TP: ۴۶۰ FP: ۰ TN: ۱۹۶۱۸ FN: ۱۴۲۶ | ۰/۹۳ | ۰/۹۹ |
| روی هم | TP: ۱۵۰۱۸ FP: ۴۱۲ TN: ۶۶۸۹۸ FN: ۲۳۸۹ | ۰/۹۶ | ۰/۹۷ |

ماسک‌ها برای شش اسلاید ذکر شده به صورت دستی و تعیین آستانه، دیگر نیاز به دخالت انسان نداریم، از این روی، این روش دقت و سرعت بالایی را دارد.

- همانطور که پیشتر نیز گفته شد، فایل‌های دیجیتال اسلایدها، از طریق اسکن نمونه‌ها توسط دستگاه و در زیر میکروسکوپ تهیه می‌شوند. در بعضی مواقع به دلیل شرایط فیزیکی نمونه، خطای دستگاه و یا حتی خطای انسانی ممکن است نمونه و یا قسمتی از آن به درستی اسکن نشود. این نواحی نیز علاوه بر نواحی پس زمینه اطلاعات مورد نیاز ما را ندارند و یا از دست داده‌اند و حالت بلوری به خود گرفته‌اند. با توجه به ماهیت روش استخراج ذکر شده، این مشکل در این فرآیند حل می‌شود. لپلاسین یک تصویر حاوی اطلاعاتی نظیر گوشه‌ها و خطوط است از این روی، یکی از روش‌های اصلی تشخیص بلوری بودن تصویر است، زیرا در این تصاویر اطلاعاتی مانند خطوط اندکی محو می‌شوند. با توجه به اینکه مبنای اصلی روش استخراج ذکر شده نیز لپلاسین تصویر است، در طی این فرآیند تصاویر بلوری مقدار واریانس لپلاسین کمتری گرفته و خود به خود حذف می‌گردند. مثالی از این پدیده در تصویر ۴-۶ آمده است.



شکل ۴-۶: دو تصویر از اسلایدی از مجموعه داده اطلس ژنوم تیروئید^{۱۳} citeclark تهیه شده‌اند که در آن تصویر سمت راست به دلایلی مقداری تار شده است و توسط روش پیشنهادی فیلتر می‌شود و تصویر سمت چپ آن تصویری سالم است که از نواحی کناری تصویر دیگر بریده شده است.

۳-۴ داده افزایی

داده افزایی یکی از پیش‌پردازش‌هایی است که قبل از داده شدن داده‌ها به مدل، روی آن‌ها انجام می‌شود. در داده افزایی^{۱۰} از تصاویر موجود، تصاویر جدیدی بازآفرینی می‌شوند. معمولاً تصاویر جدید ایجاد شده، توزیع داده‌ی نزدیکی، به توزیع داده‌ی داده‌های اصلی و تست دارند و هدف از این کار این است که تعمیم یافتنگی مدل را افزایش دهیم و بیش برآش را به تعویق بیاندازیم تا دقیق‌تر مدل بر روی داده‌های تست افزایش پیدا کند. روش‌ها و الگوریتم‌های مختلف با پیچیدگی‌های مختلفی وجود دارد که در ادامه روش‌هایی که در این پژوهه مورد بررسی قرار گرفت، آمده است.

چرخش تصویر به صورت تصادفی: در این روش تصاویر به اندازه مقداری تصادفی چرخ داده می‌شوند^{۱۱}. از آنجایی که در این مسئله، با چرخش، تصاویر اطلاعاتی از دست نمی‌دهند و علاوه بر آن، مقدار خروجی مدل ما نیز نباید تغییر کند، لذا می‌توان از این روش برای داده افزایی استفاده کرد.

آینه کردن به صورت تصادفی: مانند قسمت قبل، از این روش نیز می‌توان برای داده افزایی استفاده

¹⁰Data Augmentation

¹¹Random Rotate

کرد. تصاویر در این روش چپ به راست یا بالا به پایین معکوس می‌شوند طوری که انگار جلوی آیینه قرار گرفته‌اند.^{۱۲}.

تغییر مقیاس به صورت تصادفی: معمولاً عکس‌هایی که از آن‌ها برای آموزش مدل استفاده می‌کنیم دارای ابعاد و رزولوشن متفاوتی هستند و در نتیجه اندازه سلول‌ها در آن‌ها متفاوت است. از آنجایی که مدل در نهایت تلاش می‌کند ویژگی‌های مرتبط با این سلول‌ها را بیابد و با توجه به آن‌ها، تشخیص درستی را ارایه دهد، باید پیش‌پردازشی انجام گیرد تا مدل نسبت به اندازه سلول‌ها، حساس نشود. از این روی می‌توان قبل از استفاده از عکس‌ها در آموزش مدل، مقیاس تصاویر را به صورت تصادفی تغییر داد.^{۱۳}

بلورتر کردن به صورت تصادفی و اضافه کردن نویه گوسین: ممکن است تصاویر اسکن شده، به خوبی تصاویر زمان آموزش مدل نباشد و عوامل محیطی، باعث به وجود آمدن نویه^{۱۴} در داده‌ها شده باشد، از این روی ما از قصد نویز و یا بلور رو به عکس زمان آموزش اضافه می‌کنیم تا مدل توانایی بیشتری برای پیش‌بینی درست روی عکس‌های نویزی داشته باشد.

تغییر رنگ: همانطور که پیشتر نیز گفته شد، عکس‌ها و اسلامیدها ممکن است با روش‌های متفاوتی رنگ شده باشند و رنگ‌های مختلفی را به خود بگیرند. از این روی باید عکس‌های زمان آموزش مدل نیز به اندازه کافی تنوع رنگ داشته باشد و مدل توانایی تشخیص درست در بازه رنگ‌های متفاوتی را داشته باشد. برای این کار می‌توان رنگ‌آمیزی عکس‌ها را پیش از دادن به مدل تغییر داد. روشی که در اینجا استفاده می‌شود به این صورت است که در ابتدا عکس‌ها را از حالت RGB به HSV تبدیل می‌کنیم. در دامنه HSV، کanal‌ها به ترتیب حاوی اطلاعات رنگ^{۱۵}، اشباع^{۱۶} و روشنایی^{۱۷} هستند. در اینجا کافیست مقدار کanal رنگ را که مقداری بین ۰ تا ۳۶۰ به خود می‌گیرد را صورت تصادفی تغییر دهیم. در تصویر ۷-۴ مثالی از این پیش‌پردازش آمده است.

تطبیق دامنه فوریه

در هنگام آموزش یک مدل، فرض می‌شود که داده‌های آموزشی (چه بزرگ یا کوچک) نماینده خوبی از توزیع کلی داده‌هاست. با این حال، اگر ورودی‌ها در زمان آزمون به طور قابل توجهی با داده‌های آموزشی

¹²Random Flip

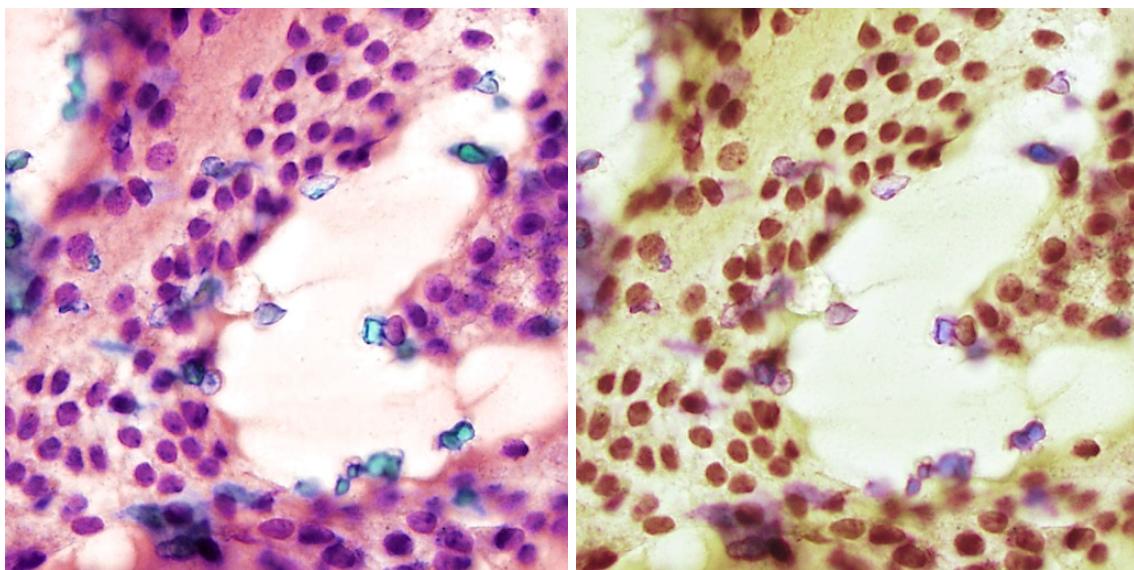
¹³Random Scale

¹⁴Noise

¹⁵Hue

¹⁶Saturation

¹⁷Brightness



شکل ۴-۷: به ترتیب از راست به چپ عکس اصلی و عکس تبدیل شده حاصل از داده‌افزایی تغییر رنگ آمده است.

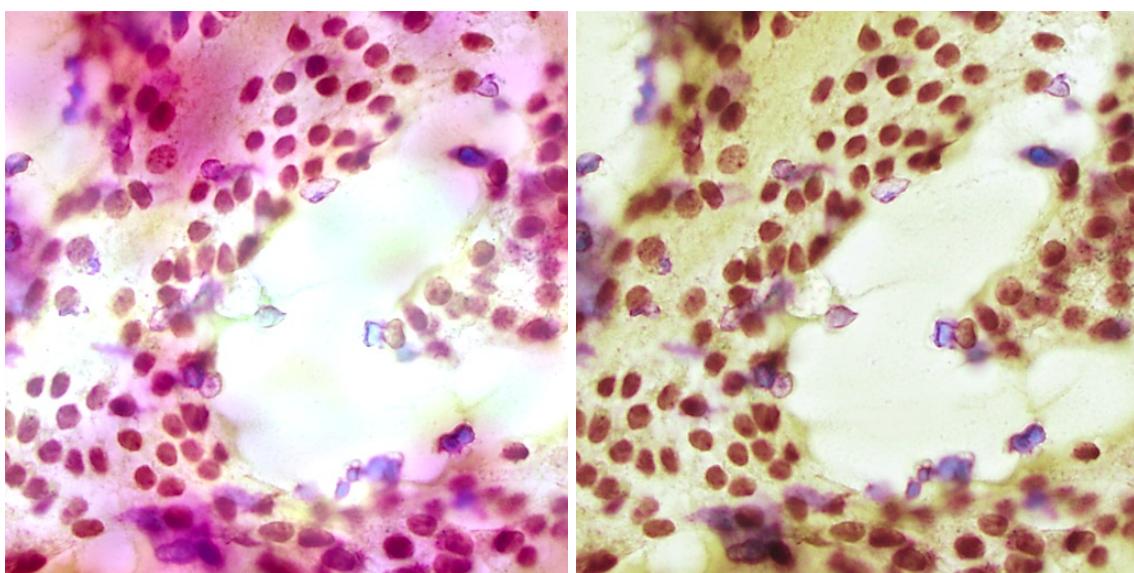
متفاوت باشد، مدل ممکن است عملکرد چندان خوبی نداشته باشد در حالی که برای یک انسان، بایادگیری مفهوم یک موضوع این مشکل کمتر وجود دارد. دلیل اینکه مدل شما در این سناریوها خیلی خوب عمل نمی‌کند این است که دامنه مسئله تغییر کرده است. در این مورد تطبیق دامنه به کمک شما می‌آید. تطبیق دامنه زیرشاخه‌ای از یادگیری ماشین است که به سناریوهایی می‌پردازد که در آن یک مدل آموزش‌دیده بر روی توزیع منبع در زمینه توزیع هدف متغراست استفاده می‌شود. به طور کلی، تطبیق دامنه^{۱۸} از داده‌های برچسب گذاری شده در یک یا چند دامنه منبع برای حل وظایف جدید در یک دامنه هدف استفاده می‌کند.

به طور مثال در مسئله این پروژه تنظیمات دستگاه اسکنر و نوع آن، روش اسکن، روش نمونه برداری و دیگر موارد ممکن است روی توزیع داده‌ها تاثیر بگذارد در حالی که این عوامل در بیمارستان‌های مختلف و مجموعه‌داده‌های مختلف متغراست است، لذا استفاده از تطبیق دامنه، گزینه خوبی برای افزایش عملکرد مدل است.

یکی از روش‌های تطبیق دامنه، تطبیق دامنه فوریه^{۱۹} است. این روش ابتدا در سال ۲۰۲۰ میلادی و در مقاله [۱۷] معرفی شد و به این صورت عمل می‌کند که ابتدا عکس‌های هدف و منبع را با تبدیل فوریه

¹⁸Domain adaptation

¹⁹Fourier domain adaptation



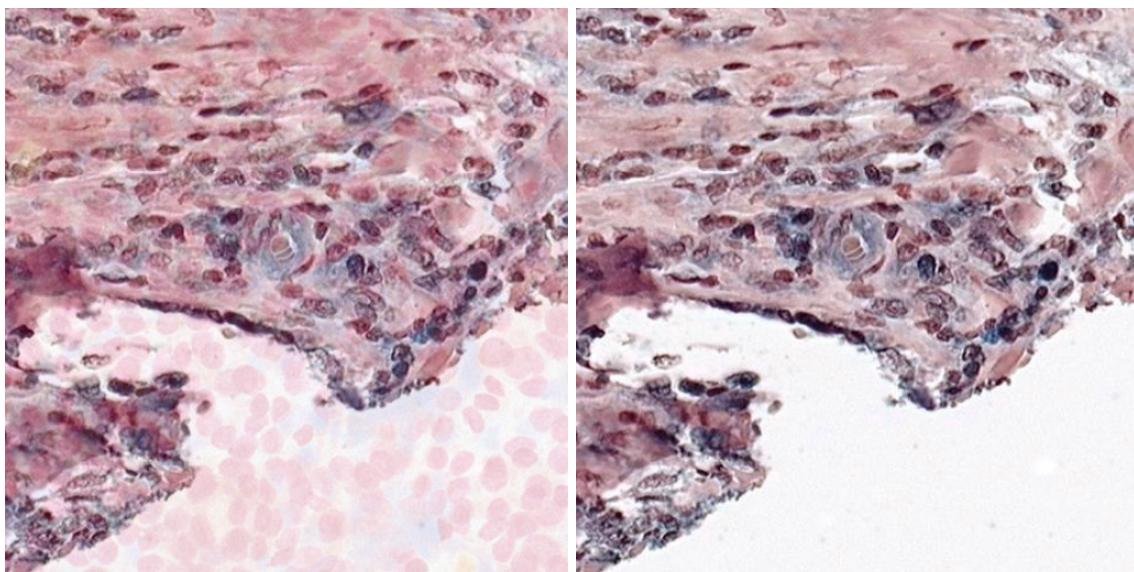
شکل ۴-۸: به ترتیب از راست به چپ عکس اصلی و عکس تبدیل شده حاصل از داده افزایی تطبیق دامنه فوریه آمده است.

به حوزه فرکانس می برد، سپس ناحیه با فرکانس پایین در داده های منبع را با ناحیه های هدف جاگزین می کند و بعد از آن فوریه معکوس را روی داده ها انجام داده و داده های به حالت قبل بر می گردند. دلیل و انگیزه این روش است که طیف دامنه سطح پایین حوزه فرکانس می تواند به طور قابل توجهی تغییر کند بدون اینکه بر درک معناشناسی سطح بالا تأثیر بگذارد و با این تغییر، توزیع تصویر نهایی به تصویر هدف نزدیک می شود. این روش با اینکه هزینه محاسباتی بسیار کمی دارد، اما در مواردی عملکرد خوبی را ارائه کرده است که نمونه از عملکرد این روش، در تصویر ۴-۸ آمده است.

ادغام تصاویر

ادغام تصویر^{۲۰} یکی از روش های افزایش داده است که ابتدا در سال ۲۰۱۷ مقاله [۱۸] معرفی شد. این الگوریتم کاملاً نظاممند نامگذاری شده است که در آن به معنای واقعی کلمه ویژگی ها و برچسب های مربوط به آنها را با هم مخلوط می کنیم. شبکه های عصبی، مستعد به خاطر سپردن برچسب های اشتباہ هستند. رویه ذکر شده این کار را با ترکیب ویژگی های مختلف با یکدیگر کاهش می دهد (همین مورد برای برچسب ها نیز اتفاق می افتد) به طوری که یک شبکه در مورد رابطه بین ویژگی ها و برچسب های آنها

²⁰Mix-up



شکل ۹-۴: به ترتیب از راست به چپ عکس اصلی و عکس تبدیل شده حاصل از داده‌افزایی ادغام تصویر آمده است.

بیش از حد مطمئن نشود. همانطور که گفته شد در ادغام دو عکس را با ضرایب λ و $1-\lambda$ بین صفر و یک، پیکسل به پیکسل با هم جمع می‌کنند تا عکس جدیدی بدست آید. مقدار λ طوری انتخاب می‌شود که تصویر اصلی دچار تغییر معنا نشود. نمونه‌ای از نتایج این روش در تصویر ۹-۴ آمده است.

۴-۴ آموزش مدل

۱-۴-۴ مدل‌های مورد بررسی

در این پژوهش، از دو نوع مدل با ایده اصلی متفاوت و ظرفیت شبکه‌های متفاوت استفاده شد، تا بتوان نتیجه نهایی را برای هر کدام مقایسه کرد. این مدل‌ها [۲۰]InceptionV4، [۱۹]InceptionV3 و [۲۱]ResNet18 [۲۱]ResNet101 هستند. در این قسمت به معرفی این دو مدل می‌پردازیم.

معماری Inception

این معماری ابتدا در [۲۲] معرفی شد که یک شبکه عمیق مبتنی بر شبکه‌های عصبی پیچشی است. این شبکه از اجزای تکرارشونده با نام ماژول Inception تشکیل شده است. هر کدام از این ماژول‌ها، ورودی را با ماتریس‌های هسته با اندازه متفاوت پردازش می‌کنند. به این معنی که بجای اینکه در هر لایه از شبکه CNN از هسته با یک سایز استفاده شود در هر ماژول از هسته‌هایی با سایزهای متفاوت استفاده می‌شود و در نهایت خروجی‌ها را با هم ادغام می‌کنیم. یکی از اهداف این کار پیدا کردن فیچرها در مقیاس‌های متفاوت است. جزئیات بیشتر این معماری در مقاله اشاره شده آمده است.

معماری ResNet

مبناً این معماری نیز شبکه‌های عصبی پیچشی است که ابتدا در پژوهش [۲۱] معرفی شد. یکی از مشکلاتی که شبکه‌های عصبی به خصوص پیچشی با آن روپرتو است، نیاز به استفاده از تعداد لایه‌های زیاد برای یادگیری ویژگی‌های مناسب مسئله است. با افزایش تعداد لایه‌ها در این شبکه‌ها، قابلیت شبکه برای پیدا کردن ویژگی‌های پیچیده بیشتر می‌شود اما افزایش عمق شبکه، تنها با انباشتن لایه‌ها در کنار هم کار نمی‌کند. آموزش شبکه‌های عمیق این چنینی به دلیل مشکل ناپدید شدن گرادیان بسیار دشوار است زیرا، در حین فرآیند آموزش که گرادیان به لایه‌های قبلی انتشار می‌یابد، ضرب مکرر ممکن است گرادیان را بینهایت کوچک کند و در چنین حالتی لایه‌های اول به درستی آموزش نمی‌بینند. این معماری با رها کردن تصادفی و پرش از لایه‌های آن در حین آموزش و استفاده از شبکه کامل در زمان آزمایش، روشی ضد شهودی را برای آموزش یک شبکه بسیار عمیق پیشنهاد کرد. برای این کار در بعضی نقاط از شبکه، ارتباطات مستقیم بین دو لایه غیر مجاور وجود دارد. این ارتباطات، خروجی لایه قبل را به ورودی لایه پایان ارتباط اضافه می‌کند. با این کار از همگرا شدن گرادیانت به صفر جلوگیری می‌کند. جزئیات بیشتر این معماری در مقاله اصلی آمده است.

۲-۴-۴ روش آموزش مدل‌ها

برای فرآیند آموزش، مدل‌های معرفی شده در بخش قبل هر کدام به صورت جداگانه مورد و مستقل آموزش داده شدند. این مدل‌ها بر روی دو مجموع داده پاپسوسایتی و پایگاه داده ریزآرایه استنفورد و روش‌های مختلف داده افزایی مورد آزمایش قرار گرفتند. که در مجموع ۱۲۵۸ تصویر برای آموزش، ۱۷۴ تصویر

برای ارزیابی و ۳۷۸ تصویر نیز برای تست به کار گرفته شد. نکته‌ای که در هنگام تقسیم بندی تصاویر به سه گروه آموزش، ارزیابی و تست باید در نظر گرفت، این است که نباید تصویری از یک اسلاید در دو گروه قرار بگیرند. به عنوان مثال، اگر از یک اسلاید چند تصویر استخراج شود باید تصاویر استفاده شده تنها در یکی از این سه گروه قرار بگیرند زیرا این تصاویر توزیع بسیار نزدیکی بهم دارند در نتیجه، قرار دادن آن‌ها در یک گروه به ارزیابی بهتر مدل کمک می‌کند. نتایج نهایی اجرای این جدول ۱-۵ آمده است.

در ستون اول این جدول مدل‌های مورد آزمایش، در ستون دوم روش‌های داده افزایی مختلف، ستون سوم بهترین دور در حین فرآیند آموزش، ستون چهارم دقت نهایی در تشخیص بدхیم و یا خوشخیم بودن تصویر و ستون آخر نیز ویژگی عملکرد گیرنده^{۲۱} برای گروه بدخیم آمده است. دقت به صورت میانگین دقت مدل در دو گروه محاسبه شده است. روش‌های داده افزایی که در ستون دوم ذکر شده و به کار گرفته شد به چند گروه متفاوت دسته بندی می‌شوند. منظور از mixup همان روش ادغام تصاویر است. منظور از fda نیز روش تطبیق دامنه فوریه است و jit نیز تغییر تصادفی رنگ تصویر است و در نهایت منظور از all ترکیبی از هر سه روش است. در این ستون دو کلمه دیگر نیز دیده می‌شود. کلمه base اشاره به ترکیبی از روش‌های آئینه کردن، چرخاندن، تغییر مقیاس و اضافه کردن نویز گوسی دارد و کلمه base-nrs تنها تفاوتی که در روش داده افزایی دارد عدم وجود روش تغییر مقیاس است. در نهایت none اشاره به عدم استفاده از داده افزایی می‌کند.

در ادامه از مجموع داده اطلس ژنوم سرطان تیروئید برای آموزش مدل ResNet101 استفاده شد. همانطور که در فصل‌های پیش نیز اشاره شد، برچسب‌های این مجموع داده به صورت در صد سلول‌ها نرمال و توموری تهیه شده است. در ادامه سعی شده تا دسته بندی بر روی این مجموع داده آموزش داده شود تا با داشتن تصویر، وجود یا عدم وجود سلول‌های توموری را تشخیص دهد. تعداد تصاویری که در نهایت از این مجموع داده استخراج شد، ۲۶۵۴۱۸ است که از ۱۸۵۰۹ تصویر برای آموزش، ۲۵۹۰۹ تصویر برای ارزیابی و ۵۴۴۱۹ برای تست استفاده شد. باز باید روی این نکته تاکید کرد که تصاویر استخراج شده از یک اسلاید نباید در دو گروه قرار بگیرند. با توجه به حجم زیاد داده‌ها و زمان بر بودن پردازش، در این قسمت تنها از روش mixup و base برای پیش‌پردازش استفاده شد. مدل در ۲۰ دور^{۲۲} بر روی داده‌ها آموزش داده شد. دقت و نمودارهای مربوط به آن در تصویر ۱-۵ آمده است. دستگاه گرافیک به کار رفته مدل Quadro RTX 8000 بوده و زمان اجرا ۳۲ ساعت طول کشیده است.

²¹Receiver operating characteristic - ROC

²²Epoch

۵-۴ جمع‌بندی

ابتدای این فصل به معرفی مجموع دادگان به کاررفته در این پژوهش پرداخته شد و توضیحات مختصه‌ی درباره نکات و ویژگی‌های مربوط به آنها آمد. سپس راه حلی برای چالش استخراج بریده‌های حاوی اطلاعات پیشنهاد شد و به مراحل کار، مثال‌ها و مزیت‌های آن اشاره شد. بعد از آن، داده‌افزایی و روش‌های متنوع بکارگرفته شده آن معرفی شدند و در نهایت، آزمایش‌های مختلفی که در این پژوهش انجام شد به طور مفصل توضیح داده شدند.

فصل ۵

نتایج آزمایش‌ها

سه مجموع داده استفاده شده در این پژوهش توزیع‌های بسیار متفاوتی دارند و برای اهداف متفاوتی مورد استفاده قرار گرفتند. در ادامه نتایج بخش‌های اول و دوم آموزش آورده شده است که در آن به ترتیب طبقه‌بندی‌های بدخیم- خوش‌خیم و طبقه‌بند توموری- سالم آورده شده است.

۱-۱ بخش اول، طبقه‌بند بدخیم و خوش‌خیم

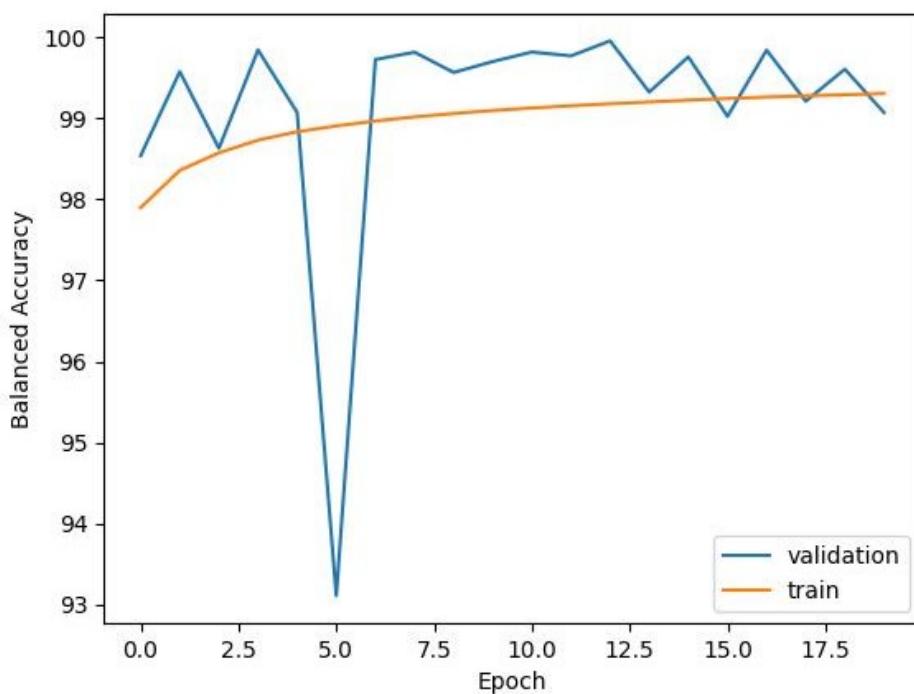
دو مجموع داده اطلس پاپسوسایتی و ریزآرایه استنفورد برای طبقه‌بندی بدخیم و خوش‌خیم بودن تصاویر نمونه مورد استفاده قرار گرفتند که تعداد کل تصاویر این دو مجموع داده روی هم ۱۸۱۰ عدد بوده و از بین آن‌ها به ترتیب ۱۲۵۸، ۱۷۴ و ۳۷۸ تصویر برای آموزش، ارزیابی و تست استفاده شد. بهترین دقیقی که بر روی آن‌ها بدست آمد ۹۶/۹۸ بوده است. از طرفی، تفاوت فاحشی بین روش‌های مختلف داده‌افزایی و شبکه‌های مختلف دیده نمی‌شود. این اتفاق می‌تواند به دلیل تعداد کم تصاویر آموزش باشد که وجود مجموع داده خارجی می‌توانست کمک شایانی را به حل این موضوع بکند. بهترین نتیجه ذکر شده، با روش داده‌افزایی jit بدست آمده است. تصاویر موجود در این دو مجموع داده، حالت مطالعه موردنی^۱ داشته و از تنوع زیادی در روش رنگ آمیزی شیمیایی سلول‌ها برخوردار بوده است، از این جهت، این روش داده‌افزایی، توانسته تنوع رنگ خوبی در داده‌های آموزش ایجاد کند، تا عملکرد مدل را بهبود ببخشد. استفاده از گروه داده‌افزایی base نیز به نظر تصمیم مناسبی بوده است، زیرا که در بیشتر موارد

¹Case Study

روش none دقت کمتری را نسبت به روش‌های دیگر کسب کرده است. روش داده‌افزایی all از آنجایی که تمام روش‌ها را باهم انجام می‌دهد، به نظر توزیع تصاویر آموزش را بیش از حد تغییر داده، از این جهت، نمودارها، دقت پایداری را نیز در هنگام آموزش نداشته‌اند دقت روش fda برروی دوره‌ها، در هنگام آموزش بسیار نوسانی بوده است، اما دقت‌های بالایی را در بهترین دوره خود نشان داده است. شاید از این جهت باشد که با جابجایی مولفه‌های پایین فوریه تصاویر، ویژگی‌های مهم از نظر طبقه‌بندی نیز جابجا می‌شوند و معنا و برچسب تصویر نهایی را تغییر می‌دهند. روش mixup نیز به نظر تفاوت چندانی در عملکرد با دیگر روش‌ها نداشته است اما در طی فرآیند آموزش نمودار پایداری را ثبت کرد. خلاصه از بهترین نتایج کسب شده بر روی مدل‌ها، در جدول ۲-۵ آمده است.

۲-۵ بخش دوم، طبقه‌بند توموری و سالم

در قسمت دوم آموزش نیز، مجموع داده بزرگ‌تر دیگری به کار رفت. اطلس ژنوم سرطان تیروئید، روی هم‌رفته دارای ۲۶۵۴۱۸ بریده تصویر است که از بین آن‌ها به ترتیب ۵۴۴۱۹، ۱۸۵۰۹۰ و ۲۵۹۰۹ تصویر برای آموزش، ارزیابی و تست استفاده شد. تصاویر بر روی مدل ResNet101 آموزش داده شدند که مدل توانست وجود یا عدم وجود سلول‌های توموری را در بهترین دوره با دقت ۹۹/۷۲ بر روی تصاویر تشخیص دهد. این مجموع داده، اسلایدهای همگونی داشته که عامل مهمی در عملکرد خوب مدل بوده است.



شکل ۱-۵: بهترین دوره انتخاب شده دوره ۱۲ بوده که دقت و ویژگی عملکرد گیرنده بر روی داده‌های تست در این دوره به ترتیب 99.72% و 99.0% است.

جدول ۱-۵: نتایج آموزش با مدل‌های مختلف و روش‌های داده افزایی متفاوت بر روی مجموع دو مجموع داده پاپسوسایتی و ریزآرایه بافت استنفورد

| مساحت زیر نمودار ROC | دقت | بهترین دور | روش‌های داده افزایی | معماری شبکه |
|----------------------|-------|------------|---------------------|--------------|
| ۰/۹۶۵ | ۹۲/۹۴ | ۱۷ | none | Inception V4 |
| | ۹۲/۱۶ | ۸۵ | base & mixup | Inception V4 |
| | ۹۱/۴۴ | ۸۵ | base & fda | Inception V4 |
| | ۹۳/۷۳ | ۵۸ | base & jit | Inception V4 |
| | ۹۳/۸۸ | ۷ | base & all | Inception V4 |
| | ۹۲/۹۹ | ۹ | base-nrs & jit | Inception V4 |
| | ۹۴/۸۷ | ۲۷ | base-nrs & all | Inception V4 |
| ۰/۹۳۷ | ۹۱/۲۵ | ۵۶ | none | Inception V3 |
| | ۹۳/۶۳ | ۸۱ | base & mixup | Inception V3 |
| | ۹۳/۷۳ | ۹۹ | base & fda | Inception V3 |
| | ۹۱/۳۴ | ۲ | base & jit | Inception V3 |
| | ۹۳/۴۱ | ۷۵ | base & all | Inception V3 |
| | ۹۲/۰۶ | ۴۰ | base-nrs & jit | Inception V3 |
| | ۹۳/۴۱ | ۷۵ | base-nrs & all | Inception V3 |
| ۰/۹۷۱ | ۹۳/۱۰ | ۱۴ | none | Resnet101 |
| | ۹۲/۲۲ | ۸۴ | base & mixup | Resnet101 |
| | ۹۴/۰۴ | ۴۰ | base & fda | Resnet101 |
| | ۹۲/۴۷ | ۹۹ | base & jit | Resnet101 |
| | ۹۲/۰۳ | ۲۵ | base & all | Resnet101 |
| | ۹۲/۲۸ | ۵۵ | base-nrs & jit | Resnet101 |
| | ۹۲/۰۳ | ۲۵ | base-nrs & all | Resnet101 |
| ۰/۹۲۲ | ۸۹/۷۵ | ۵۵ | none | Resnet18 |
| | ۹۱/۰۹ | ۷ | base & mixup | Resnet18 |
| | ۹۴/۰۱ | ۴۸ | base & fda | Resnet18 |
| | ۹۴/۹۸ | ۹۴ | base & jit | Resnet18 |
| | ۹۳/۸۸ | ۸۰ | base & all | Resnet18 |
| | ۹۰/۷۸ | ۹۰ | base-nrs & jit | Resnet18 |
| | ۹۴/۱۰ | ۱۳ | base-nrs & all | Resnet18 |

جدول ۲-۵: خلاصه‌ای از نتایج آموزش با تنظیمات منتخب شده براساس دقت بر روی دو مجموعه داده پاپسوسایتی و ریزآرایه بافت استنفورد

| مساحت زیر نمودار ROC | دقت | بهترین دور | روش‌های داده‌افزایی | معماری شبکه |
|----------------------|-------|------------|---------------------|--------------|
| ۰/۹۷۶ | ۹۳/۷۳ | ۵۸ | base & jit | Inception V4 |
| ۰/۹۶۶ | ۹۳/۸۸ | ۷ | base & all | Inception V4 |
| ۰/۹۵۷ | ۹۳/۶۳ | ۸۱ | base & mixup | Inception V3 |
| ۰/۹۷۴ | ۹۳/۷۳ | ۹۹ | base & fda | Inception V3 |
| ۰/۹۶۴ | ۹۴/۰۴ | ۴۰ | base & fda | Resnet101 |
| ۰/۹۴۳ | ۹۴/۰۱ | ۴۸ | base & fda | Resnet18 |
| ۰/۹۷۴ | ۹۴/۹۸ | ۹۴ | base & jit | Resnet18 |

جدول ۳-۵: خلاصه‌ای از بهترین نتایج کسب شده توسط دو بخش آموزش

| AUC | دقت | دوره | برچسب | تصاویر تست | تصاویر ارزیابی | تصاویر آموزش | تصاویر آموزش | بخش |
|-------|-------|------|------------------|------------|----------------|--------------|--------------|---------|
| ۰/۹۷۴ | ۹۴/۹۸ | ۹۴ | خوش‌خیم بدخیم | ۳۷۸ | ۱۷۴ | ۱۲۵۸ | | بخش اول |
| ۰/۹۹۹ | ۹۹/۷۲ | ۱۲ | سالم توموری | ۵۴۴۱۹ | ۲۵۹۹۰۹ | ۱۸۵۰۹۰ | | بخش دوم |

فصل ۶

جمع‌بندی

همانطور که پیشتر نیز گفته شد یکی از مشکلات اصلی که در این پژوهش با آن روبرو هستیم، نبود مجموع دادگان عمومی کافی است. در این پژوهش ابتدا سعی شد مجموع داده‌های متفاوتی جمع آوری و معرفی شوند. در مرحله بعد مدل‌های متفاوتی بر روی این دادگان آموزش داده شد و با وجود چالش‌هایی نظیر حجم عظیم داده‌ها، فیلتر کردن ناحیه‌های پس زمینه و دیگر موارد، با به کار بردن روش‌های پیش‌پردازشی و تنظیم ابرپارامترها، عملکرد خوبی بدست آمد.

نتایج نهایی مدل‌ها، ارقام قابل قبولی را برای دقت آن‌ها اعلام می‌کنند. با این حال، برای استفاده از آن‌ها و مفید واقع شدن آنها در دنیای امروزی، نیاز به کار بیشتر است. به عنوان مثال، برای اطمینان از عملکرد مدل‌ها می‌توان از مجموع داده‌های خارجی استفاده کرد که ارزیابی خارجی، معیار مناسبی و اطمینان بیشتری برای عملکرد کلی مدل است. جدای از این موضوع، برای ادامه کار و پیشبرد عملکرد مدل‌ها، تمرکز بیشتری را می‌تان بر روی توزیع‌های متفاوت دادگان گذاشت، که استفاده بیشتر از مفاهیمی چون تطبیق دامنه و یا انتقال یادگیری^۱ را می‌طلبد.

¹Transfer Learning

Bibliography

- [1] R. J. Marinelli, K. Montgomery, C. L. Liu, N. H. Shah, W. Prapong, M. Nitzberg, Z. K. Zachariah, G. J. Sherlock, Y. Natkunam, R. B. West, et al. The stanford tissue microarray database. *Nucleic acids research*, 36(suppl_1):D871–D877, 2007.
- [2] National cancer institute tcga-thca project. <https://portal.gdc.cancer.gov/projects/TCGA-THCA>. [Online; accessed July 5, 2022].
- [3] Papanicolaou society of cytopathology. <https://www.papsociety.org/image-atlas/>. [Online; accessed July 5, 2022].
- [4] K. O’Shea and R. Nash. An introduction to convolutional neural networks. *arXiv preprint arXiv:1511.08458*, 2015.
- [5] X. Li, S. Zhang, Q. Zhang, X. Wei, Y. Pan, J. Zhao, X. Xin, C. Qin, X. Wang, J. Li, et al. Diagnosis of thyroid cancer using deep convolutional neural network models applied to sonographic images: a retrospective, multicohort, diagnostic study. *The Lancet Oncology*, 20(2):193–201, 2019.
- [6] Q. Guan, Y. Wang, B. Ping, D. Li, J. Du, Y. Qin, H. Lu, X. Wan, and J. Xiang. Deep convolutional neural network vgg-16 model for differential diagnosing of papillary thyroid carcinomas in cytological images: a pilot study. *Journal of Cancer*, 10(20):4876, 2019.
- [7] D. T. Nguyen, T. D. Pham, G. Batchuluun, H. S. Yoon, and K. R. Park. Artificial intelligence-based thyroid nodule classification using information from spatial and frequency domains. *Journal of clinical medicine*, 8(11):1976, 2019.

- [8] D. D. Elliott Range, D. Dov, S. Z. Kovalsky, R. Henao, L. Carin, and J. Cohen. Application of a machine learning algorithm to predict malignancy in thyroid cytopathology. *Cancer cytopathology*, 128(4):287–295, 2020.
- [9] M. B ö hland, L. Tharun, T. Scherr, R. Mikut, V. Hagenmeyer, L. D. Thompson, S. Perner, and M. Reischl. Machine learning methods for automated classification of tumors with papillary thyroid carcinoma-like nuclei: A quantitative analysis. *Plos one*, 16(9):e0257635, 2021.
- [10] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.
- [11] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, pages 248–255. Ieee, 2009.
- [12] M. Halicek, M. Shahedi, J. V. Little, A. Y. Chen, L. L. Myers, B. D. Sumer, and B. Fei. Head and neck cancer detection in digitized whole-slide histology using convolutional neural networks. *Scientific reports*, 9(1):1–11, 2019.
- [13] A. S. El-Hossiny, W. Al-Atabany, O. Hassan, A. M. Soliman, and S. A. Sami. Classification of thyroid carcinoma in whole slide images using cascaded cnn. *IEEE Access*, 9:88429–88438, 2021.
- [14] K. Clark, B. Vendt, K. Smith, J. Freymann, J. Kirby, P. Koppel, S. Moore, S. Phillips, D. Maffitt, M. Pringle, et al. The cancer imaging archive (tcia): maintaining and operating a public information repository. *Journal of digital imaging*, 26(6):1045–1057, 2013.
- [15] Stanford tissue microarray. <https://tma.im/cgi-bin/selectImages.pl?organ=thyroid>. [Online; accessed July 5, 2022].
- [16] The GIMP Development Team. Gimp.
- [17] Y. Yang and S. Soatto. Fda: Fourier domain adaptation for semantic segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4085–4095, 2020.

- [18] H. Zhang, M. Cisse, Y. N. Dauphin, and D. Lopez-Paz. mixup: Beyond empirical risk minimization. *arXiv preprint arXiv:1710.09412*, 2017.
- [19] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2818–2826, 2016.
- [20] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi. Inception-v4, inception-resnet and the impact of residual connections on learning. In *Thirty-first AAAI conference on artificial intelligence*, 2017.
- [21] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
- [22] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1–9, 2015.

واژه‌نامه

۱

ت

- تابع هدف Objective Function
مقیاس تصادفی Random Scale
تطبیق دامنه Domain adaptation
تطبیق دامنه فوریه Fourier domain adaptation
تصویر برداری کل Whole-Slide Imaging

- اسلاید slide
اندازه پرش Jump Size
آینه کردن تصادفی Random Flip
اشباع Saturation
ادغام Mix up
اطلس ژنوم سرطان تیروئید The Cancer Genome Atlas-Thyroid Cancer

د

- دقت Accuracy
داده افزایی Data Augmentation
دور Epoch

- فولیکولار ادنوما Follicular Adenoma
انتقال یادگیری Transfer Learning

ب

- بیوپسی Biopsy

چ

- چرخش تصادفی Random Rotate

- پتالوژی Pathology

ر

- رنگ Hue

- پتالوژیست Pathologist

- سرطان پاپیلاری Papillary Carcinomas

- سرطان پاپیلاری Papillary Carcinomas

پ

ق

سرطان فولیکولار Follicular Carcinomas

روشنایی Brightness

رادیو یودین Radioiodine

ل

laplacian لاپلاسی

ژ

ژنوتیپ Genotype

م

derivative مشتق

س

mask ماسک

سرطان خوشخیم Benign Carcinoma

Confusion Matrix ماتریس درهم ریختگی

سرطان بدخیم Malignant Carcinoma

Medullary carcinoma سرطان مدولا

سی تی اسکن Computed Tomography scan

Segmentation مرزبندی

ش

Contrast Limited متعادل سازی هستیوگرام تصویر

شبکه عصبی پیچشی Convolutional Neural

Adaptive Histogram Equalization

Network

Case Study مطالعه موردی

شبکه‌های عصبی پیچشی عمیق Deep

ن

Convolutional Neural Network

Decay Rate نرخ نزولی

ص

Fine Needle Aspiration نمونه برداری سوزنی

صحت Precision

(FNA)

Region Of Interest ناحیه مورد نظر

ف

و

فنتوتیپ Phenotype

варіанс variance

فولیکولار کارسینوما Follicular Carcinoma

ویژگی feature

| | | |
|---------------------|-------------------------|---------------------------|
| ي | Receiver operating..... | ویژگی عملکرد گیرنده |
| Deep Learning | يادگيری عميق | characteristic |

Abstract

Thyroid cancer is one of the most common types of cancer in the human body that affects the thyroid gland. One of the main ways to diagnose this disease is to use digital slides recorded from a sample of cells. In this way, using the characteristics that cancer cells take on, experts can distinguish them from normal thyroid cells and determine the amount of involvement of the sample in cancer. The project aims to train a model that uses digital slides to perform a high-accuracy diagnosis by using ResNet and Inception architectures based on deep convolutional neural networks. Since it takes a lot of time for experts to diagnose this cancer in this way, the developed model can be a great assistant and help for them. Finally, the evaluation results obtained from training the models show the accuracy of 94.98 for the malignant-benign classification and the accuracy of 99.72 for the healthy-tumor classification, which shows the capacity of these methods to solve such problems.

Keywords: Machine Learning, Image Processing, Thyroid Carcinoma, Malignant, Digital Slide



Sharif University of Technology
Department of Computer Engineering

B.Sc. Thesis

Diagnosis of Thyroid Cancer Using Neural Networks from Cytology Slides

By:

Amir Hossein Mohsen Nezhad

Supervisor:

Prof. Hossein Asadi

Prof. Hamidreza Rabiee

August 2022