

مروری کلی بر روش های تقسیم بندی خودکار در تصاویر بافت آسیب شناسی

نویسندگان: زهرا خرمیان ، مرضیه علیدادی ، مریم کاظمی



دانشگاه صنعتی اصفهان

1. چکیده :

در دنیای امروزی، خیلی از بیماری ها به دلیل اینکه سرعت تشخیص و درمان پایین است، منجر به مرگ بیماران می شود ؛ ولی روش های نوینی همچون دیجیتالی کردن، روند درمان را تشخیص و بهبود بخشیده است.

در این گزارش ما سعی می کنیم که ابعاد مختلف تقسیم بندی دیجیتالی در تصاویر بافت آسیب شناسی را مورد بررسی قرار داده و به بررسی برخی از این روش ها و چالش های پیش رو بپردازیم .

2. مقدمه :

درجه بندی سرطان به معنای تعیین میزان وخامت سرطان، یکی از مهم ترین مراحل پیش بینی و برنامه ریزی برای درمان سرطان به شمار می رود.

تحلیل بخش های حاوی تومور در تصاویر بافت شناسی، از روش های مهم و پر کاربرد در شناسایی و تشخیص سرطان هست. آسیب شناسان برای بررسی میزان بد خیمی بافت ها معمولا از ریخت شناسی (علم شناخت شکل و ساختار دقیق مواد) غدد استفاده می کنند ؛ که برای به دست آوردن آمار ریخت شناسی قابل اعتماد، لازم است تا از روی تصاویر بافت شناسی، غدد به صورت دقیق تقسیم شوند.

با وجود روش آسیب شناسی دیجیتالی، یعنی اسکن و ذخیره دیجیتالی بخش های بافت بیمار، می توان ابزارهایی را برای تجزیه و تحلیل کمی و اتوماتیک این داده های تصویری پیچیده که حاوی اطلاعات مفیدی هستند ساخت که در نتیجه، فرصت مناسبی را برای ما فراهم می کند تا به طور خودکار، مقادیر زیادی از این داده های تصویری پیچیده را تجزیه و تحلیل کنیم تا نتیجه گیری های زیست شناسی از آنها گرفته شود و مدل های ساختاری سلولی و بافتی را در مقیاس بزرگ مورد مطالعه قرار دهیم .

به همین دلایل، تجزیه و تحلیل داده های هیستوپاتولوژی در سال های گذشته مورد توجه بسیاری قرار گرفته است. از جمله کار هایی که در این نوع تحلیل صورت می گیرد، طبقه بندی و تقسیم غدد، توسط تجزیه و تحلیل خودکار مقادیر زیادی از داده های تصویری پیچیده است.

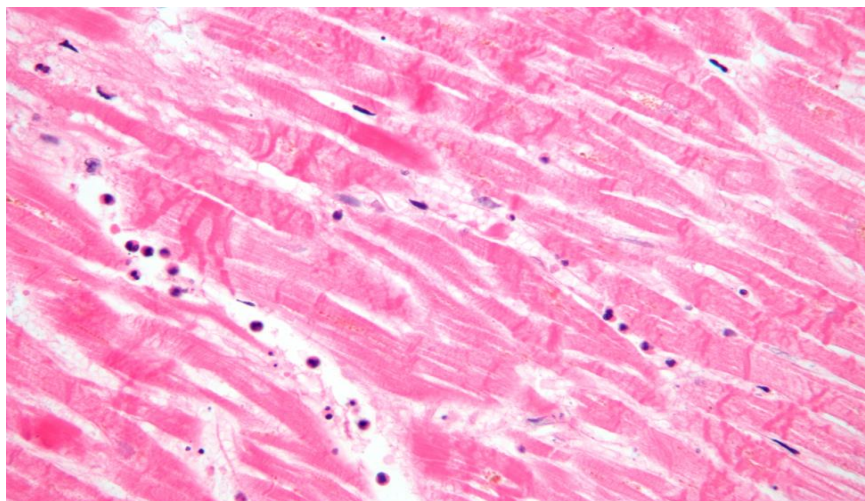
ارائه ی روش های خودکار برای این کار باعث کاهش بار کاری آسیب شناسان، افزایش دقت و شناسایی انواع مختلف تومور و سرطان می شود.

خوشبختانه همه گیر شدن استفاده از تصاویر دیجیتالی از بافت ها، زیر ساخت لازم برای ایجاد یک روش دیجیتال برای آنالیز بافت ها را فراهم کرده است. در نتیجه ارائه ی الگوریتمی که بتواند ویژگی های ساختاری غدد را به طور خودکار به دست آورد، پیشرفتی تاثیر گذار در روند درمان بیماری محسوب می شود.

جالب تر این که، پیشرفت در شبکه های پیچیده عصبی عمیق (CNN) و به ویژه در بهینه سازی آنها باعث شده است تا آنها به الگوی پیشرفته ترین روش برای تشخیص اشیا تبدیل شوند. در واقع، شبکه های عصبی عمیق برای بومی سازی و طبقه بندی هسته ها در داده های بافت شناسی سرطان های پستان و روده بزرگ استفاده شده است.

3. آشنایی با تعاریف :

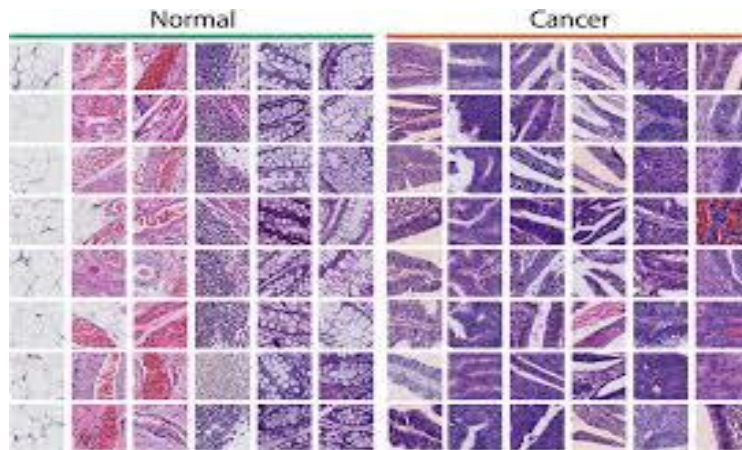
آسیب شناسی بافتی (به انگلیسی Histopathology) از دو کلمه ی Histo به معنای بافت و Pathology به معنای آسیب شناسی تشکیل می شود. در histopathology به بررسی میکروسکوپی بافت پرداخته می شود. این کار به منظور یافتن علائم و علت بیماری در بافت انجام می گیرد. آسیب شناسی دیجیتال: یعنی اسکن و ذخیره دیجیتالی بخش هایی از بافت بیمار برای بررسی و آسیب شناسی.



مثالی از تصویر هیستوپاتولوژی [مرجع 10]

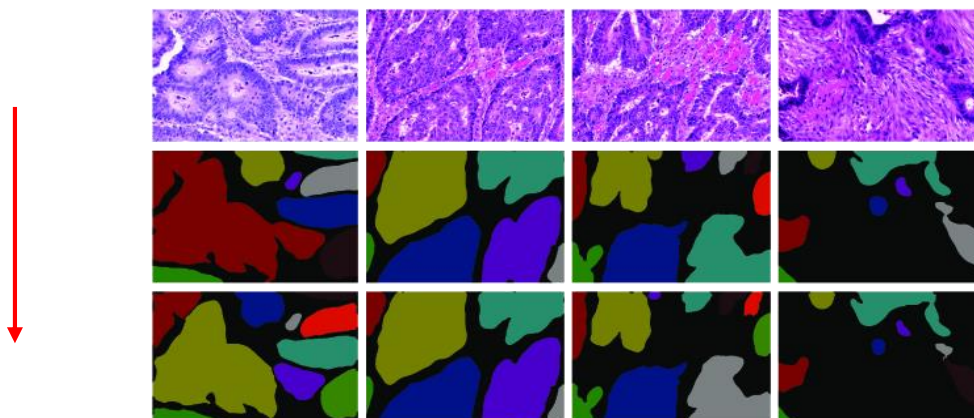
این مقاله به مبحث image segmentation پرداخته که نباید با classification اشتباه گرفته شود. دسته بندی (classification): یعنی تمام تصاویر به دو دسته تقسیم شوند؛

مثلا: (1) سرطانی و (2) غیر سرطانی



مثالی از classification [مرجع 3]

اما تقسیم بندی (segmentation): یعنی قطعه بندی کردن یک تصویر دیجیتال به چند بخش. تقسیم بندی تصویر معمولاً برای پیدا کردن محل اشیا موردنظر و مرزها در تصویر استفاده می شود. خروجی این فرایند، مجموعه ای از بخش ها است که اجتماع آن ها، کل تصویر را شامل می شود.



مثالی از segmentation [مرجع 3]

4. روش های موجود:

4-1. ساختار کلی روش ها:

تقسیم بندی خودکار تصاویر توسط یک کامپیوتر، یکی از مسائل مربوط به حوزه ی بینایی کامپیوتر (Computer Vision) است.

هدف از بینایی کامپیوتر، قادر ساختن کامپیوترها به درک محیط، با استفاده از داده های بصری است. برای مثال با استفاده از الگوریتم تشخیص چهره و الگوریتم های مبتنی بر رنگ برای تشخیص پوشش و مو، می توان تعداد اشخاص موجود در یک تصویر را به طورخودکار به دست آورد.

استفاده از شبکه های عصبی پیچشی در حل مسائل مربوط به حوزه ی بینایی

کامپیوتر، نتایج موفقیت آمیزی را با خود به همراه داشته است. از طرفی تقسیم بندی تصاویر بافت آسیب شناسی نیز زیر مجموعه ی همین حوزه می باشد و استفاده از این شبکه ها در این مورد هم مورد استقبال محققان قرار گرفته است و تقریباً تمام روش های پیشنهادی اخیر مبتنی بر همین شبکه ها می باشد.

4-1-1. شبکه های عصبی (مصنوعی) :

در علوم کامپیوتر، شبکه های عصبی مصنوعی (Artificial Neural Networks) مجموعه ای از الگوریتم ها هستند که برای حل مسائل از مدل پردازشی مغز انسان الهام می گیرند.

در مغز انسان لایه های متعددی از ساختارهای عصبی وجود دارد. ابتدا داده ها توسط لایه های عصبی اولیه مانند سلول های مربوط به حس لامسه و یا بینایی ، جمع آوری شده و سپس با عبور از لایه های درونی مغز، پردازش می شوند و به این شکل شخص متوجه تماس با جسم داغ و یا مشاهده ی یک شیء خاص می شود.

در شبکه های عصبی مصنوعی، با پیاده سازی لایه های مختلف سعی می شود مانند فرایند یادگیری در مغز انسان که یک فرایند تدریجی است، از داده هایی که به عنوان ورودی به آنها داده می شود آموزش دیده و خروجی خواسته شده را به دست آورند.

4-1-2. شبکه های عصبی پیچشی :

شبکه های عصبی پیچشی (Convolutional Neural Networks / CNNs) یک نوع از شبکه های عصبی هستند که در همه یا بعضی از لایه های خود از عملیات کانولوشن ریاضی استفاده می کنند.

عملگر کانولوشن ریاضی (*) ، یک عملگر دوتایی است و مدل گسسته ی آن به شکل زیر تعریف می شود که در آن f و g دو تابع مختلط ، m و n اعداد صحیح هستند:

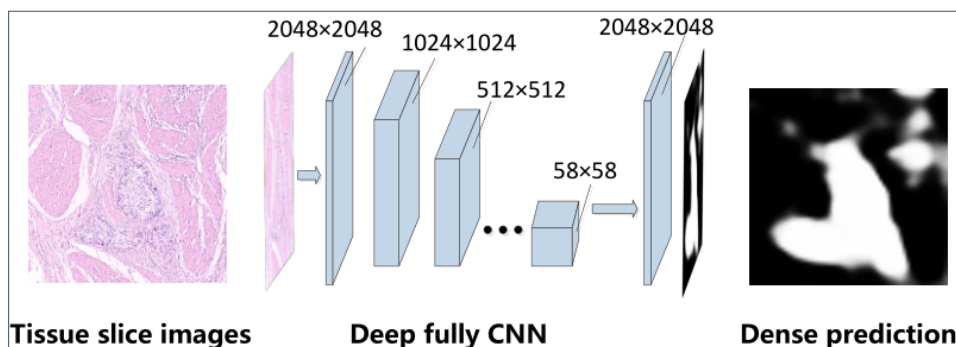
$$f(x * g)[n] = \sum_{m=-\infty}^{\infty} f[m]g[n - m]$$

4-1-3. مزیت استفاده از شبکه های عصبی پیچشی :

مزیت عملی این شبکه ها این است که برای یادگیری به داده های آموزشی و زمان کمتری نیاز دارند. یعنی داده های محلی یک تصویر کفایت می کند.

به این صورت که مطابق تصویر زیر، با تکه تکه کردن هر تصویر به قطعات کوچک تر، شبکه به بررسی همه ی قطعه ها می پردازد.

این کار درست مانند خواندن کتاب با استفاده از یک ذره بین است؛ سرانجام شما کل صفحه را خوانده اید ، اما در هر زمان فقط به یک بخش کوچک از صفحه را نگاه می کنید.



تقسیم بندی یک تصویر بافت آسیب شناسی با استفاده از شبکه عصبی پیچشی [مرجع 5]

4-2. معرفی یک روش :

در این بخش به معرفی یک مثال از روش های جدید تقسیم بندی در تصاویر بافت آسیب شناسی (از مرجع چهارم این مقاله) پرداخته می شود.

این روش از زیرمجموعه ی خاصی از شبکه های عصبی پیچشی (CNNs) ، یعنی شبکه های عصبی کاملاً پیچشی (FCNs) ، استفاده می کند. همچنین ساختار هدف برای تقسیم بندی تصاویر، غده ها هستند.

این روش از مورد دیگری که در سال های قبل ارائه شده است الهام می گیرد . در روش قبلی یک شبکه عصبی کاملاً پیچشی طوری پیاده سازی شده است که با بررسی پیکسل به پیکسل تصویر ورودی، تصویرتقسیم بندی شده ی آنها را به عنوان خروجی ارائه می دهد.

اما در این روش قدیمی، در مواردی که به علت وخامت بیماری، شکل غده تغییر یافته باشد، الگوریتم نمی تواند با دقتی مناسب خروجی را نمایش دهد. برای حل این مشکل از یک شبکه ی دیگر نیز استفاده می شود که در آن مرزهای اطراف غدد در تصویری دیگر مشخص می شوند.

در نتیجه مطابق تصویر زیر با ادغام این دو شبکه در یک شبکه ی اصلی، نتیجه ی خروجی اشتراک این دو تصویر است و حتی اگر شکل غده تغییر یافته باشد، به علت تشخیص مرزها توسط زیر شبکه ی دوم، تصویر خروجی نتیجه ای با دقت بیشتر خواهد بود.

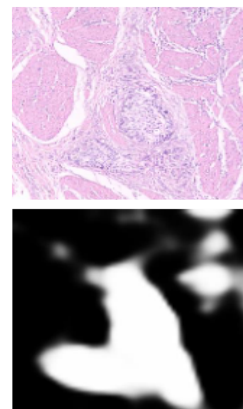
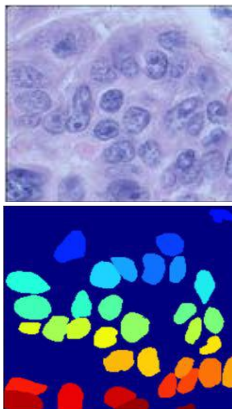
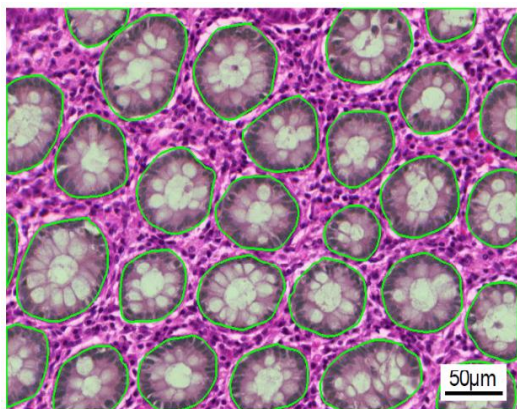
4-3. انواع روش ها :

اگر چه روش های پیشنهاد شده، در بسیاری از موارد از جمله کلیت الگوریتم های مورد استفاده، مشابه هم هستند اما در موارد زیادی، با یکدیگر متفاوت هستند. روش های موجود را می توان بر اساس عوامل متفاوتی دسته بندی کرد که در این مقاله به بررسی سه مورد از آنها پرداخته می شود. این سه مورد عبارت اند از: نوع خروجی، ساختار تقسیم شده و نوع عملکرد.

4-3-1. انواع روش ها از نظر نوع خروجی :

خروجی سیستم ها می تواند به شکل های مختلفی باشد که در این بند برای نمونه به سه مورد از آنها اشاره می کنیم.

خروجی باینری (سیاه و سفید) : بخش تقسیم شده با رنگ سفید از زمینه جدا می شود. خروجی رنگی: تصویر خروجی یک تصویر رنگی می شود. خروجی مرز دار: بخش تقسیم شده با خطوط رنگی از سایر بخش ها مجزا می شود.



خروجی مرز دار [مرجع 2]

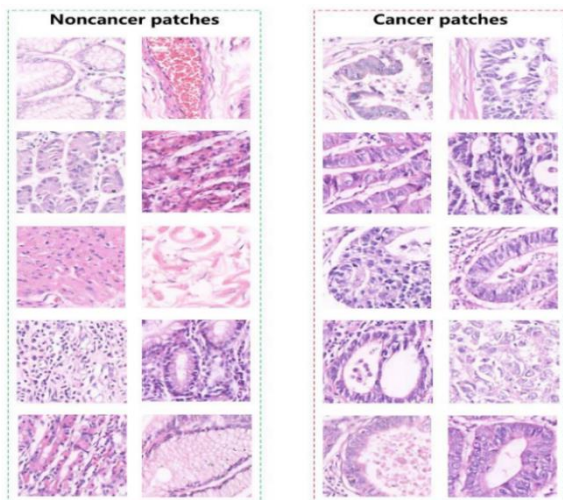
خروجی رنگی [مرجع 1]

خروجی باینری [مرجع 5]

4-3-2. انواع روش ها از نظر ساختار تقسیم شده :

هر کدام از روش های ارائه شده برای تقسیم بندی تصاویر، ممکن است به علت کاربردهای در بیماری های مختلف به دنبال تقسیم اجزای متفاوتی از بافت مانند غده ها، سلول ها و... باشند. مثلاً ممکن است ملاک، تشخیص سرطان باشد؛ که در اینصورت هر ساختار دلخواه از بافت را بر اساس سرطانی و غیر سرطانی بودن تقسیم بندی می کند.

مثلاً در شکل زیر، هدف روش استفاده شده، تشخیص سرطان در بافت بوده است. در نتیجه بافت ها بر اساس سرطانی و غیرسرطانی بودن تقسیم بندی شده اند، و قسمت های سرطانی بافت از قسمت های غیر سرطانی آن تشخیص داده و جدا شده اند.



مثالی از تقسیم بندی بافت مبنی بر سرطانی و غیر سرطانی بودن [مرجع 5]

4-3-3. انواع روش ها از نظر نوع عملکرد :

روش های ارائه شده برای تقسیم بندی خودکار تصاویر، هر کدام عملکرد خاص خودشان را دارند و تصاویر را مبتنی بر آن عملکرد خاص تقسیم بندی می کنند.

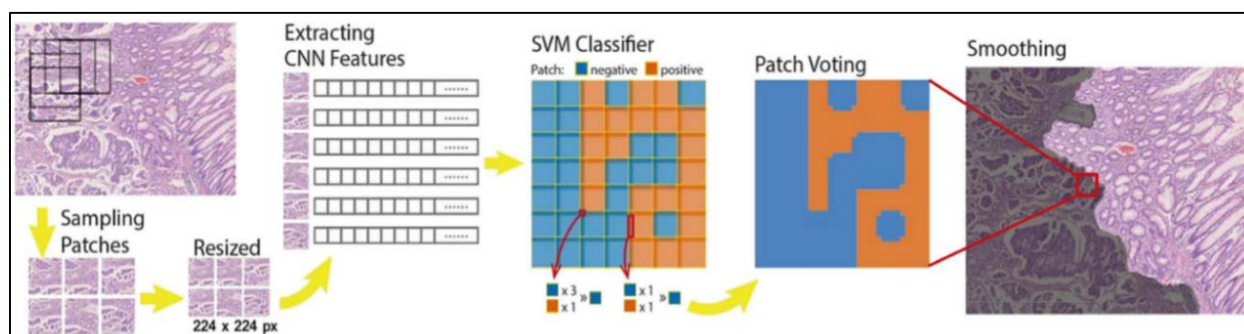
به طور عمده این روش ها از سه نوع عملکرد استفاده می کنند:

(1) روش مبتنی بر پیکسل: در این روش، ابتدا تمام پیکسل ها مورد بررسی قرار می گیرند، و سپس با توجه به انسجام پیکسل های مشخص و ساختار و شکل آنها، ساختاری که به دنبال تقسیم آن هستند، را می سازند.

(2) روش مبتنی بر ساختار: در این روش ابتدا ساختار های مختلف در تصویر تشخیص داده می شوند، و سپس بررسی می شود که آیا هر کدام از این ساختار ها متعلق به ساختار مد نظر هست یا نه.

(3) سایر روش ها: در برخی از روش های جدید، مانند مورد اول عمل می شود با این تفاوت که به جای پیکسل از مربع هایی با ابعاد کوچک استفاده می شود. به عبارتی، در این روش معیار دیگری را به جای پیکسل در نظر می گیرند، و تصاویر را بر اساس آن معیار تکه تکه می کنند. که معمولاً ابعاد این معیار از ابعاد یک پیکسل بزرگ تر است و مثلاً شامل چند پیکسل می شود.

ممکن است سوال شود "دلیل این کار چیست؟" دلیل در نظر گرفتن معیار بزرگ تر از پیکسل، این است که آن ها بر این باورند که یک پیکسل به قدری کوچک است که ارزش مطالعه ندارد و ساختار خاصی درون آن مشخص نیست. در نتیجه تکه های بزرگ تری از بافت را به عنوان معیار در نظر می گیرند؛ که این معیار، قابل مطالعه تر است و ساختار بافت را بهتر نشان می دهد. مانند شکل زیر:

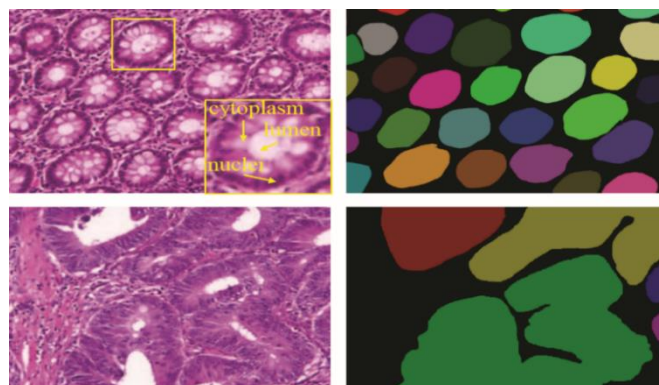


مثالی از تقسیم بندی با استفاده از patch هایی شامل چند pixel [مرجع 3]

5. چالش ها :

بسته به درجه های مختلف بافت شناسی و همچنین از یک بیماری به بیماری دیگر، تفاوت های زیادی در ساختار یک بافت وجود دارد. خصوصیات بافت غیر غده گاهی اوقات می تواند شبیه بافت غده باشد و در این صورت مجموعه داده ها می تواند الگوریتم های تقسیم بندی را به اشتباه بیندازد. به علاوه، در موارد بدخیم ممکن است ساختار بافت به طور جدی تخریب شده باشد.

مثلاً تفاوت زیادی در ساختارهای غده ای بین موارد خوش خیم و بدخیم از بافت روده بزرگ وجود دارد. در شکل زیر، شکل بالا سمت چپ یک نمونه از موارد خوش خیم است؛ شکل پایین سمت چپ یک نمونه از موارد بدخیم است. همانطور که مشخص است، ساختار بافت به کل متفاوت شده است و همچنین تشخیص قسمت های مختلف بافت در مورد بدخیم دشوار تر شده است.

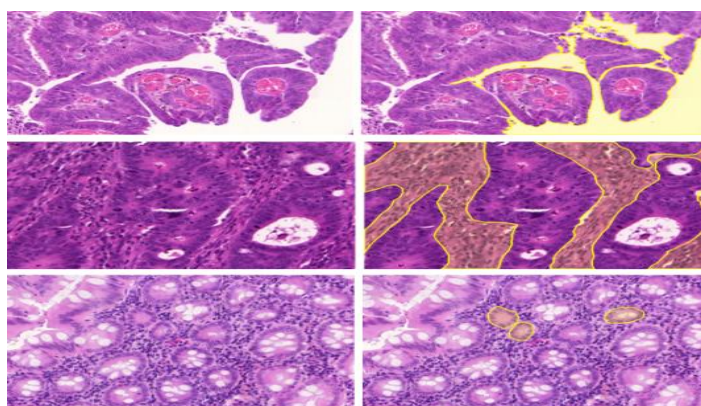


مثالی از تقسیم بندی غدد در موارد بدخیم و خوش خیم [مرجع 4]

به علاوه، ارائه ی ویژگی های بالینی مناسب و کافی برای انواع بیماری ها و انواع بافت ها در قسمت های مختلف بدن، یک مسئله ی مهم در تحلیل تصاویر بافت شناسی است.

تقسیم بندی هسته ها، کار بسیار پیچیده ای است: وجود انواع مختلف بافت ها، تفاوت در رنگ تصاویر و انواع مختلف هسته ها، ویژگی های متفاوتی را به وجود می آورد که باعث می شود الگوریتم های تقسیم بندی مبتنی بر روش های قدیمی، کار بسیار پیچیده ای شود. انواع مختلف سرطان ها، تفاوت های قابل توجهی در بافت، رنگ و مقیاس تصاویر دارند، که باعث می شود یافتن یک الگو که برای تشخیص سرطان در بافت های مختلف مثلاً هم سرطان روده و هم سرطان مغز مناسب باشد، کار دشواری شود. در صورتی که در بسیاری از روش ها، تمرکزشان را بر روی ویژگی های بافت ها و شکل شیء ها گذاشته اند.

مثلاً در شکل زیر: تصاویر سمت چپ، تصاویر اصلی هستند و تصاویر سمت راست، تصاویر تقسیم بندی شده هستند. به ترتیب از بالا مربوط به (1) دستگاه گوارش، (2) بافت با هسته های متراکم در لایه ی مخاطی، (3) غدد کوچک هستند. همانطور که مشخص است، ساختار بافت ها کاملاً متفاوت از هم هستند.



مثالی از ساختار متفاوت بافت ها در قسمت های مختلف بدن [مرجع 2]

همچنین، گاهی اشیاء در غدد با هم در تماس هستند، و تشخیص مرز بین آن ها با استفاده از روش های خودکار کار سختی می شود. به علاوه، داده های استخراج شده از تصاویر بافت شناسی بالینی، کم هستند.

استفاده از این داده ها مناسب نیست؛ به دو دلیل: یک اینکه با توجه به اینکه وقوع این بیماری ها زیاد نیست، تصاویر بافت شناسی کمتری از آن ها موجود است؛ در نتیجه، اطلاعات کمتری برای استخراج وجود دارد. و دوم اینکه نیروی کار زیادی برای حاشیه نویسی دستی اطلاعات حاصل از این تصاویر بافت شناسی مورد نیاز است.

در مورد کنترل و نگه داری حجم اطلاعات ذخیره شده هم با چالش هایی مواجه هستیم. هر کدام از تصاویر بافت شناسی ابعاد بزرگی دارد، که حتی اگر تعداد این تصاویر محدود باشد، پایگاه داده ی این تصاویر، در مقیاس بزرگ تلقی خواهد شد.

همچنین، الگوریتم های پیشنهادی اکثراً وابسته به نوع ورودی هستند؛ در نتیجه در صورت استفاده در دنیای واقعی که دستگاه های تولید تصاویر دیجیتالی متفاوتی وجود دارد و تصاویر با تراکم داده ی متفاوت استفاده می شوند، انتظار می رود مشکل به وجود آید.

در آخر، هزینه ی محاسباتی بالا و زمان پردازش زیاد می تواند یک مشکل هنگام استفاده از آن ها در یک برنامه ی واقعی باشد؛ به خصوص برای تجزیه و تحلیل تصاویر بزرگ بافت آسیب شناسی.

6. نتیجه گیری :

در این گزارش پیشرفت ها، روش ها و چالش های موجود در فناوری تقسیم بندی خودکار تصاویر بافت آسیب شناسی، باهدف افزایش توجه به این حوزه، مورد بررسی قرار گرفت.

به طور کلی، بحث و بررسی های تحقیقاتی انجام شده در حول این موضوع، به پیشرفت های قابل توجهی دست یافته اند و با در نظر گرفتن چالش های موجود به اصلاحاتی پرداخته اند، تا روند تقسیم بندی را بهبود بخشند.

در عین حال، تا کنون هیچ کدام از این روش های پیشنهادی، آمادگی لازم برای اجرای گسترده در دنیای واقعی و با شرایط واقعی را کسب نکرده اند؛ و همچنان لازم است تا روش های برتری ارائه شود، که از لحاظ کمی و کیفی عملکرد بهتری داشته باشند و در مقیاس بزرگ عملی باشند، تا پتانسیل بهره برداری عمومی را داشته باشند.

7. مراجع :

1. Naylor, Peter, Marick Laé, Fabien Rey, and Thomas Walter. "Nuclei segmentation in histopathology images using deep neural networks." In 2017 IEEE 14th international symposium on biomedical imaging (ISBI 2017), pp. 933-936. IEEE, 2017.

2. Sirinukunwattana, Korsuk, Josien PW Pluim, Hao Chen, Xiaojuan Qi, Pheng-Ann Heng, Yun Bo Guo, Li Yang Wang et al. "Gland segmentation in colon histology images: The glas challenge contest." Medical image analysis 35 (2017): 451-502.

3. Xu, Yan, Zhipeng Jia, Liang-Bo Wang, Yuqing Ai, Fang Zhang, Maode Lai, I. Eric,

and Chao Chang. "Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features." *BMC Bioinformatics* 18, no. 1 (2017): 281.

4. Chen, Hao, Xiaojuan Qi, Lequan Yu, and Pheng-Ann Heng. "DCAN: deep contour-aware networks for accurate gland segmentation." In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, pp. 2487-2496. 2016.

5. Peng, Binbin, Lin Chen, Mingsheng Shang, and Jianjun Xu. "Fully Convolutional Neural Networks for Tissue Histopathology Image Classification and Segmentation." In *2015 25th IEEE International Conference on Image Processing (ICIP)*, pp. 1403-1407. IEEE, 2015.

6. James Le, "The 4 Convolutional Neural Network Models That Can Classify Your Fashion", [www. towardsdatascience.com](http://www.towardsdatascience.com) ,2018.

7. Shirai, Yoshiaki. *Three-dimensional computer vision*. Springer Science & Business Media, 2012.

8. Szeliski, Richard. *Computer vision: algorithms and applications*. Springer Science & Business Media, 2010.

9. Damelin, Steven B., and Willard Miller Jr. *The mathematics of signal processing*. Vol. 48. Cambridge University Press, 2012.

10. Michael Bonert ,Wikipedia page:"<https://commons.wikimedia.org/wiki/User:Nephron>" .Image is in direct link: "<https://www.wikipedia.org/wiki/Histopathology>" wikipedia page content edited on 21, July, 2020.