

بهبود تشخیص خودکار توده‌های خوش‌خیم و بدخیم در تصاویر فراصوت پستان با استفاده از یک روش قطعه‌بندی بهینه

وجیهه سادات وطن‌پور^۱، سید محمود سخایی^۲، یاسر بالغی^۳، الهام سادات بنی‌مصطفوی^۴

چکیده

سرطان پستان شایع‌ترین نوع سرطان در جمعیت زنان جهان است. تشخیص زودهنگام و درمان مؤثر با هدف کاهش مرگ و میر ناشی از این بیماری، از طریق روش‌های غربالگری انجام می‌شود. تصویربرداری فراصوت (اولتراسوند) به دلیل ماهیت غیرتهاجمی و مزایای آن نسبت به سایر روش‌های تشخیصی، از مهم‌ترین و مؤثرترین روش‌های شناسایی و تشخیص این بیماری است. به منظور بهبود عملکرد تشخیص، سیستم‌های تشخیص کمک رایانه‌ای معرفی شدند. در این پژوهش یک سیستم تمام خودکار شناسایی و تشخیص سرطان پستان ارائه شده است که متشکل از چهار مرحله اصلی است: پیش‌پردازش تصویر در دو گام جهت برجسته‌تر کردن نواحی کم‌پژواک (هایپواک) و با هدف انتخاب نقطه بذر و ناحیه مطلوب و همچنین حذف نویز با استفاده از فیلتر میانگین‌های غیر محلی، قطعه‌بندی با بهره‌گیری از اطلاعات مکانی و فرکانسی تصویر و بهبود آن با استفاده از الگوریتم بهینه‌سازی ژنتیک، استخراج مجموعه‌ای شامل ۲۱ ویژگی مبتنی بر شکل و مرز و در نهایت طبقه‌بندی با استفاده از طبقه‌بند ماشین بردار پشتیبان به منظور تفکیک توده‌ها به دو گروه خوش‌خیم و بدخیم. نتایج ارزیابی‌های صورت گرفته بر روی تصاویر پایگاه‌های داده مختلف میزان دقت بیش از ۹۵/۵٪ را نشان داد.

کلیدواژه‌ها

سرطان پستان، فراصوت، خوش‌خیم و بدخیم، ویژگی شکل و مرز، طبقه‌بندی

پستان از طریق شناسایی زودهنگام و درمان مؤثر به کاهش مرگ و میر ناشی از این بیماری کمک می‌کند. سازمان بهداشت جهانی، غربالگری ماموگرافی را برای زنان در معرض خطر متوسط ابتلا به سرطان پستان توصیه می‌کند [۱]. ماموگرافی به عنوان روش استاندارد تشخیص مورد توجه است که یونیزان بودن اشعه مورد استفاده، مهم‌ترین محدودیت ماموگرافی است. دقیق‌ترین روش برای تشخیص قطعی نوع توده، نمونه‌برداری است، اما این روش تهاجمی و پرهزینه است [۲، ۳].

از بین روش‌های تصویربرداری، تصویربرداری فراصوت روش مکمل ایده‌آلی محسوب می‌شود و در کنار ماموگرافی می‌تواند سبب بهبود حساسیت و نیز بهبود شناسایی تومور در مراحل اولیه شود [۴]. ماهیت نویزی و تضاد پایین تصاویر فراصوت و فقدان مقادیر کمی برای ویژگی‌های استفاده شده در تحلیل تصویر باعث متفاوت بودن تفسیر رادیولوژیست‌ها از یک تصویر و در نتیجه تشخیص متفاوت می‌شود. لذا به منظور سهولت در تشخیص و

۱- مقدمه

طبق آمار جهانی، از بین ۱۰ نوع سرطان شایع در سال ۲۰۲۰، سرطان پستان به عنوان فراگیرترین و کشنده‌ترین نوع سرطان در جمعیت زنان جهان اعلام شده است. برنامه‌های غربالگری سرطان

این مقاله در آبان‌ماه ۱۴۰۰ دریافت شد در اردیبهشت‌ماه ۱۴۰۱ بازنگری و سپس پذیرفته شد.

^۱ دانش‌آموخته کارشناسی‌ارشد دانشگاه صنعتی نوشیروانی، دانشکده برق و کامپیوتر، گروه مهندسی پزشکی.

رایانامه: vs.vatanpour@gmail.com

^۲ دانشگاه صنعتی نوشیروانی، دانشکده برق و کامپیوتر، گروه مهندسی پزشکی.

رایانامه: smsakhaei@nit.ac.ir

^۳ دانشگاه صنعتی نوشیروانی، دانشکده برق و کامپیوتر، گروه الکترونیک

رایانامه: y.baleghi@nit.ac.ir

^۴ دانشگاه علوم پزشکی مازندران، دانشکده پزشکی، گروه رادیولوژی

رایانامه: razavi109@yahoo.com

نویسنده مسئول: سیدمحمود سخایی

طبقه‌بندی، متعلق به مجموعه ویژگی‌های ریخت‌شناسی با ۵ ویژگی با سطح زیر منحنی ۰/۹۴ است که عملکرد مؤثرتری نسبت به بهترین مجموعه از ویژگی‌های بافت با ۴ ویژگی و سطح زیر منحنی ۰/۸۹ را از خود نشان داد. به بیان دیگر، تنها با بهره‌مندی از ویژگی‌های ریخت‌شناسی نیز می‌توان توده‌ها را به گروه‌های خوش-خیم و بدخیم تفکیک کرد.

در پژوهش حاضر رویکردی جهت بهبود قطعه‌بندی خودکار توده‌ها به کمک الگوریتم بهینه‌سازی ژنتیک و همچنین استخراج ویژگی‌های مناسب به منظور ارتقاء کارایی یک سیستم تشخیص رایانه‌ای سرطان پستان مورد تأکید قرار گرفته است. در سیستم تشخیص خودکار پیشنهادی ابتدا پیش‌پردازش تصویر با هدف تعیین نقطه بذر و ناحیه مطلوب و همچنین اعمال فیلتر (NLM) جهت حذف نویز و عوامل ناخواسته انجام شد. در گام بعد، قطعه‌بندی با کمک مشخصه‌های تصویر در حوزه مکان و فرکانس و بهینه‌سازی پارامترهای آن با استفاده از الگوریتم ژنتیک صورت گرفت. سپس ویژگی‌هایی مبتنی بر شکل و مرز از توده استخراج گردید و در نهایت طبقه‌بندی ماشین بردار پشتیبان جهت تفکیک توده‌ها به دو گروه خوش‌خیم و بدخیم مورد استفاده قرار گرفت. ارزیابی‌ها نشان می‌دهد که سیستم تمام‌خودکار ارائه شده، قادر به تفکیک توده‌های بدخیم و خوش‌خیم با دقت قابل قبولی است. بعلاوه با بررسی عملکرد سیستم معرفی شده روی تصاویر حاصل از تجهیزات متفاوت، مشخص شد که سیستم در مقابل تغییرات کیفیت و ویژگی‌های تصویر مقاوم است.

ساختار مقاله حاضر به صورت زیر تنظیم شده است. در بخش ۲ به شرح روش‌های پیش‌پردازش، قطعه‌بندی، استخراج و انتخاب ویژگی و طبقه‌بندی پرداخته شده است. در بخش ۳ و ۴ نیز به ترتیب ارزیابی الگوریتم پیشنهادی و جمع‌بندی مطالب ارائه شده است.

۲- روش پیشنهادی

شکل (۱) روند نمای الگوریتم مورد استفاده را نشان می‌دهد. در ادامه جزئیات روش بیان می‌شود.

۲-۱ تعیین خودکار نقطه بذر و ناحیه مطلوب

تعیین صحیح ناحیه مطلوب بر بهبود عملکرد قطعه‌بندی تاثیرگذار است. یک ناحیه مطلوب مناسب باید ناحیه توده را به طور کامل پوشش دهد و تا حد ممکن شامل بافت‌های اطراف نباشد. در این پژوهش ناحیه مطلوب، ناحیه‌ای مستطیلی و شامل توده و بخشی از بافت اطراف آن است و به صورت کاملاً خودکار و بر اساس روش پیشنهاد شده در [۱۲] استخراج شد. این روش شامل سه مرحله پیش‌پردازش تصویر، تعیین نقطه بذر و تعیین ناحیه مطلوب است.

کاهش خطای انسانی ناشی از وابستگی به دستگاه و تجربه و تخصص پزشک، استفاده از سیستم‌های رایانه‌ای مطرح می‌گردد. این سیستم‌ها با رویکردهای رایانه‌ای مبتنی بر پردازش تصویر و یادگیری ماشین، خصوصیات از تصویر را ارائه می‌دهند تا نظر ثانویه‌ای را برای پزشک ایجاد کند و با تشخیص قطعی‌تر نوع توده تعداد نمونه‌برداری‌های غیرضروری کاهش یابد [۶، ۵، ۲].

سیستم‌های موجود بر اساس میزان دخالت کاربر در فرآیند تشخیص به گروه‌های نیمه‌خودکار و کاملاً خودکار تقسیم می‌شوند [۷]. در [۸] SHAN و همکاران رویکردی مبتنی بر کمی سازی ۱۰ ویژگی از BIRADS^۱ برای طبقه‌بندی توده‌ها در تصاویر فراصوت پستان ارائه کردند که در آن چهار روش یادگیری ماشین شامل درخت تصمیم، شبکه عصبی مصنوعی، ماشین بردار پشتیبان (SVM) و جنگل تصادفی برای تشخیص توده‌های بدخیم از خوش‌خیم به کار گرفته شد. برای یافتن بهترین ترکیب که بهترین عملکرد را در هر روش یادگیری ماشین داشته باشد از یک روش انتخاب ویژگی پایین به بالا^۲ استفاده شد. بیشترین دقت به میزان ۸۴/۲ درصد با استفاده از ۵ ویژگی منتخب توسط ماشین بردار پشتیبان حاصل شد.

استفاده از تحلیل بافت و رویکرد فرا تفکیک‌پذیری برای محاسبه درجه عضویت بدخیم بودن هر تصویر در [۹] پیشنهاد شد. سیستم تشخیصی مورد استفاده متشکل از دو گام آموزش و آزمایش و چهار مرحله محاسبات فراتفکیک‌پذیری^۳، استخراج ناحیه مطلوب (ROI)^۴، استخراج ویژگی و طبقه‌بندی است. در این روش، ابتدا مجموعه‌ای از تصاویر وضوح کم از توالی تصاویر ورودی استخراج می‌شود. سپس با استفاده از الگوریتم فراتفکیک‌پذیری، یک تصویر با وضوح بالا بازسازی می‌شود. سپس ۵ نوع ویژگی بافت استخراج و در آخر به یک طبقه‌بند برای ساخت یک مدل وارد می‌شوند. این مدل برای تمایز بین ROI‌های خوش‌خیم و بدخیم آموزش می‌بیند. باتوجه به نتایج، ویژگی‌های بافت بدست آمده از روش هیستوگرام گرادینان-های جهت‌دار با مقدار سطح زیر منحنی ۰/۹۹ بیشترین کارایی را در تشخیص توده‌های بدخیم نشان داد.

در [۱۰] ویژگی‌های ریخت‌شناسی و بافتی متنوعی که توسط محققین مختلف برای سیستم‌های تشخیص رایانه‌ای مبتنی بر تصاویر فراصوت پستان پیشنهاد شده بود، بررسی شد. گام پیش-پردازش توسط فیلتر انتشار ناهمسانگرد و گام قطعه‌بندی نیز به کمک روشی مبتنی بر تبدیل آب‌پخش‌انجام شد. از تصویر باینری حاصل از مرحله قطعه‌بندی، ۲۶ ویژگی ریخت‌شناسی و از تصویر سطح خاکستری تعداد ۱۴۶۵ ویژگی بافت استخراج شد. طبقه‌بندی داده‌ها به وسیله طبقه‌بند تحلیل تمایز فیشر محلی^۵ (LFDA) اجرا شد. نتایج ارزیابی‌ها حاکی از آن است که بهترین عملکرد

¹ Breast Imaging Reporting and Data System

² Bottom-up

³ Super-resolution

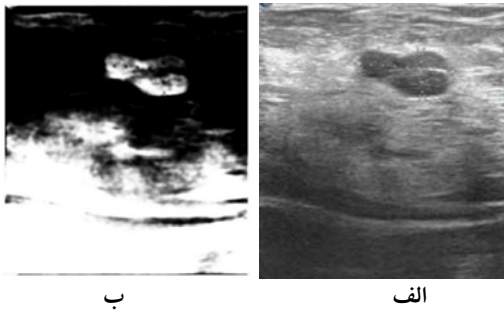
⁴ Region of interest

⁵ Linear Fisher Discriminant Analysis

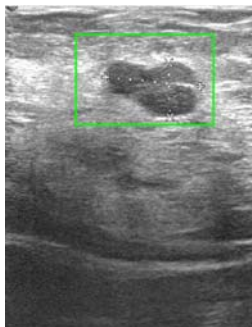
⁶ Non-Local Means

خروجی الگوریتم پیش پردازش را بر روی یک تصویر نشان می-دهد.

برای انتخاب نقطه مرجع برای تولید بذر، از یک روش کاملاً خودکار و فقی استفاده شده است. در این روش، به این نکته توجه شده است که ساختمان پستان دارای لایه های متفاوتی با شدت روشنای های متفاوت است. به منظور جلوگیری از قرار گرفتن نقطه بذر انتخابی در لایه چربی متصل به قفسه سینه یا لایه نزدیک پوست، استفاده همزمان از تابع مرکز چگالی وزن دار^۴ و تابع وزن به صورت انطباق پذیر و براساس آناتومی پستان (که شدت روشنایی نواحی اطرف را تضعیف می کند) مورد توجه قرار گرفت. در نهایت ناحیه مطلوب با اعمال عملگرهای مورفولوژیکی و عملیات آستانه گذاری خودکار Otsu [۱۳] و با کمک نقطه بذر، از تصویر اصلی استخراج گردید. خروجی این الگوریتم در شکل (۳) مشاهده می شود.



شکل (۲): الف، تصویر اصلی و ب، خروجی حاصل از الگوریتم پیش پردازش

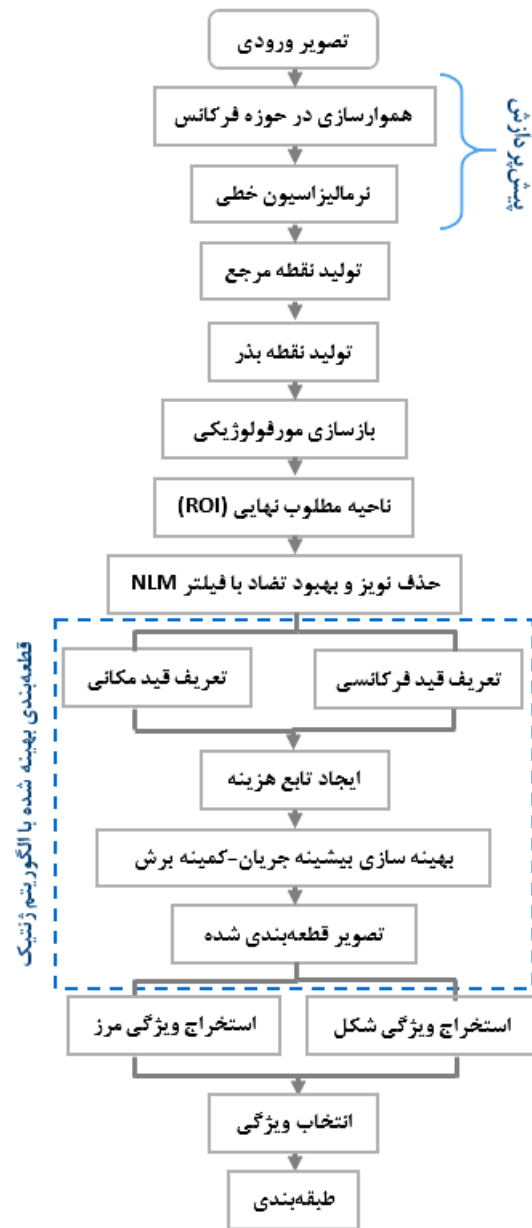


شکل (۳): کادر سبز رنگ، ناحیه مطلوب نهایی است

۲-۲ فیلتر کردن تصویر

همان طور که ملاحظه شد با اعمال مراحل بالا بر روی تصویر، ناحیه مطلوب که شامل توده و قسمتی از بافت سالم اطراف آن است از تصویر اصلی جدا گردید. بنابراین، تصویر بدست آمده همچنان نویزی بوده و تضاد پایینی دارد و ضروری است به منظور بهبود کیفیت تصویر، پیش از گام قطعه بندی، یک مرحله پیش-پردازش بر روی تصویر اعمال شود. از میان بسیاری از روش هایی که تا کنون برای این کار ارائه شده، فیلتر میانگین های غیر محلی [۱۴] در پژوهش حاضر مورد استفاده قرار گرفته است. در

مرحله پیش پردازش با هدف بهبود تصویر و برجسته کردن نواحی کم پژواک^۱ (نواحی با اکوی کمتر نسبت به بافت اطراف) که از مشخصه های توده در تصاویر فراصوت پستان است، انجام شد.



شکل (۱): روندنمای الگوریتم مورد استفاده

در این روش سه مرحله ای ابتدا هموارسازی تصویر در حوزه فرکانس توسط یک فیلتر پایین گذر گاوسی با ابعاد 300×300 و فرکانس قطعه ۳۰ صورت گرفت. سپس به منظور حذف تغییرات دامنه شدت روشنایی، یک مقیاس خطی به کران های بالا و پایین تابع توزیع تجمعی اعمال گردید. در نهایت، یک تابع Z شکل^۲ مبتنی بر چولگی^۳ توزیع شدت روشنایی به تصویر اعمال شد تا مقادیر شدت روشنایی آن معکوس و به بازه [۰، ۱] تبدیل شود. این کار همچنین موجب بهبود نواحی کم پژواک می شود. شکل (۲)

^۱ Hypoecho

^۲ Z-shaped Function

^۳ Skewness

^۴ Weighted Density Center

ها است. رئوس متناظر با یک پیکسل یا یک ناحیه از تصویر هستند. ضلع‌ها نیز هر جفت از رأس‌های همسایه را طبق یک سیستم همسایگی از پیش تعریف شده به هم متصل می‌کنند. هر ضلع وزن $W(v_i, v_j)$ مربوط به خود را دارد که معمولاً به عنوان هزینه جدا کردن دو رأس متصل به کلاس‌های مختلف تعریف می‌شود. در تعریف نظریه گراف، میزان عدم تشابه بین دو قطعه می‌تواند در قالب یک برش گراف محاسبه شود. یک برش به مجموعه‌ای از ضلع‌ها گفته می‌شود که توسط آن گراف به دو مجموعه ناهمپوشان A و B تقسیم می‌شود. بنابراین قطعه‌بندی تصویر می‌تواند مطابق رابطه ۲ به عنوان برش گراف تعریف شود.

$$\text{cut}(A, B) = \sum_{u \in A, v \in B} w(u, v) \quad (2)$$

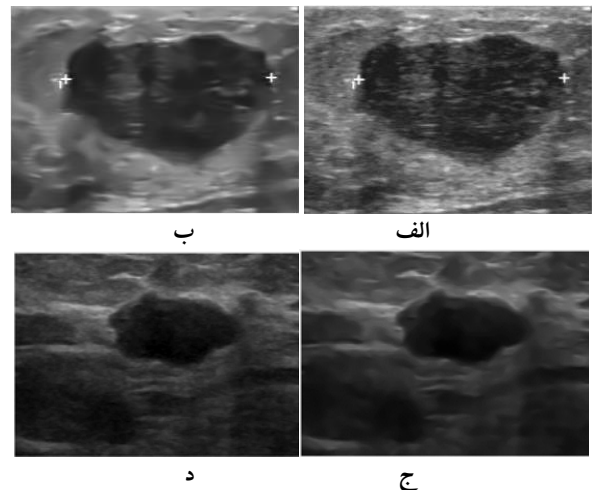
در رابطه بالا u و v رأس‌های موجود در دو قسمت هستند. با کمیته کردن تابع cut، ضلع‌های با وزن کمتر قطع می‌شوند و ضلع‌های با وزن بیشتر در داخل یک قسمت باقی می‌مانند و به این ترتیب، خروجی عملیات کمترین برش، قطعه‌بندی بهینه را فراهم می‌کند.

۲-۴ تنظیم پارامتر با الگوریتم ژنتیک

در بخش قبل شرح روش قطعه‌بندی ارائه گردید. در این روش تعداد قابل توجهی از پارامترها مورد استفاده قرار گرفت که هر یک به نوعی بر عملکرد الگوریتم تاثیرگذارند. از میان تمامی پارامترهای تعریف شده، برخی از آن‌ها نقش مستقیم و تعیین کننده‌ای بر خروجی دارند و لذا انتخاب صحیح مقادیر آن‌ها منجر به بهبود عملکرد سیستم خواهد شد. اما با توجه به این مسئله که تعیین این مقادیر به صورت دستی و برحسب سعی و خطا علاوه بر زمان‌بر بودن، اغلب منجر به نتیجه دقیقی نیز نخواهد شد، از این رو و به منظور کاهش محدودیت‌های حاصل از تعیین دستی مقادیر پارامترهای تنظیمی مورد نظر، استفاده از الگوریتم بهینه‌سازی ژنتیک مورد توجه قرار گرفته است.

در این روش ابتدا تعداد ۲۰ تصویر آموزش از هر دو کلاس داده خوش‌خیم و بدخیم از پایگاه داده دارای تصاویر قطعه‌بندی دستی انتخاب شد و براساس میزان خطای حاصل از اختلاف خروجی الگوریتم قطعه‌بندی و قطعه‌بندی دستی، بهترین مقادیر برای ۶ پارامتر δ (پارامتری که میزان تمایز دامنه روشنایی تصویر در داخل و خارج توده را تعریف می‌کند)، λ (پارامتر متعادل کننده هزینه فرکانسی و مکانی در رابطه (۱))، آستانه باینری کردن تصویر و سه پارامتر مربوط به خصوصیات آماری تغییرات دامنه در داخل توده و خارج آن به کمک الگوریتم ژنتیک بدست آمد.

شکل (۴) نتیجه حاصل از اعمال فیلتر NLM بر روی نمونه‌ای از تصاویر قابل مشاهده است.



شکل (۴): الف و ج تصاویر اصلی، ب و د تصاویر فیلتر شده با فیلتر NLM

۲-۳ قطعه‌بندی

پس از اتمام روند پیش‌پردازش و استخراج ناحیه مطلوب، لازم است ناحیه توده از بافت‌های سالم اطراف آن تفکیک شود. در این مرحله از روش ارائه شده در [۱۲] استفاده شد. مطابق روندنمای شکل (۱) ابتدا یک تابع هزینه براساس قیود فرکانسی (E^{fc}) و مکانی (E^{sc}) تعریف شد. قید فرکانسی در واقع یک لبه‌یاب در حوزه فرکانس است که از دو شاخص همگامی فاز^۱ و عبور از صفر^۲ استفاده می‌کند. قید مکانی نیز با کمک اطلاعات حوزه مکان تصویر شامل حالت، موقعیت و توزیع شدت روشنایی توده ایجاد شد. دو قید اول با استفاده از تابع گوسی بیضوی دو بعدی، و مدلسازی توزیع شدت روشنایی با کمک توزیع رایلی انجام شد. در نهایت قطعه‌بندی بهینه با کمیته کردن این تابع هزینه که در رابطه ۱ آمده است، حاصل گردید.

$$E(f|y_{ROI}) = \lambda \cdot E^{sc}(f|y_{ROI}) + E^{fc}(f|y_{ROI}), \quad \lambda \geq 0 \quad (1)$$

در رابطه بالا، yROI پیکسل‌های ROI در تصویر است. E^{fc} و E^{sc} به ترتیب تابع هزینه مکانی و تابع هزینه فرکانسی هستند و پارامتر λ نیز میزان اثرگذاری این دو را کنترل می‌کند. بهینه‌سازی برای رسیدن به کمترین مقدار تابع هزینه (۱) توسط الگوریتم پیشینه جریان-کمیته برش^۳ صورت پذیرفت. این الگوریتم که توسط Boykov و Kolmogorov برای قطعه‌بندی تصاویر در [۱۵] پیشنهاد شد، روشی مؤثر برای بدست آوردن بهینه سراسری است. در این الگوریتم، ابتدا تصویر به صورت یک گراف در نظر گرفته می‌شود که شامل مجموعه‌ای از رئوس و ضلع-

¹ Phase Congruency

² Zero Crossing

³ Max-Flow Min-Cut

۲-۵ استخراج ویژگی

۶. وسعت [۱۷،۱۰]: Bounding_Rectangle مساحت کوچک ترین مستطیل در برگیرنده توده است.

$$Ext = \frac{Area}{Bounding_Rectangle}$$

۷. فشردگی [۲۱،۱۰]:

$$Cmp = \frac{perimeter^2}{Area}$$

۸. دوری از مرکز [۲۰،۱۰]:

$$Ecc = \frac{\sqrt{Major_Axis^2 - Minor_Axis^2}}{Major_Axis}$$

۹. کشیدگی [۱۰]: نسبت بین طول و عرض کوچک ترین مستطیل در برگیرنده توده است.

$$Elng = \frac{Length}{Width}$$

۱۰. اختلاف مساحت [۲۰،۱۰]:

$$Area_Diff = Convex_Area - Area$$

۱۱. مقدار باقی مانده نرمال شده [۲۱،۱۰]:

$$Nrv = \frac{Area_Difference}{Convex_perimeter}$$

۱۲. جهت [۱۰]: با اندازه گیری زاویه بین محور اصلی بیضی معادل با ضایعه و محور X محاسبه می شود. قرارگیری در جهت موازی با پوست یکی از مؤلفه های خوش خیمی است و اغلب توده های بدخیم در جهت عمود بر پوست قرار می گیرند. دامنه این ویژگی بین ۹۰- و ۹۰+ است.

• ویژگی مبتنی بر مرز

❖ توصیف گر فوریه بیضوی نامتغیر (IEFD) [۲۲]

در این پژوهش از توصیف گر بیضوی نامتغیر به منظور بررسی اطلاعات فرکانسی مرز استفاده شد.

در تقریب سری فوریه، نقاط در فضای مختصات (x,y) در دو بعد به یک جفت معادله تبدیل می شوند که به عنوان تابعی از یک متغیر سوم (t) نوشته شده اند. سپس ضرایب فوریه براساس تقریب گسسته سری فوریه محاسبه می شود. تقریب سری فوریه بیضوی برای مرز بسته بر روی محور x و y را می توان به صورت رابطه ۳ تعریف کرد:

$$X_N(t) = A_0 + \sum_{n=1}^N a_n \cos\left(\frac{2n\pi t}{T}\right) + b_n \sin\left(\frac{2n\pi t}{T}\right)$$

بیان کمی مشخصه های متمایز کننده توده های خوش خیم و بدخیم، هدفی است که در این گام محقق می شود. اطلاعاتی که پس از فرآیند قطعه بندی از شکل توده حاصل می شوند، نقش تعیین کننده ای در تشخیص نوع ضایعه دارند. به این منظور در این پژوهش از ویژگی های شکل و مرز توده که برخلاف ویژگی های بافت مستقل از کاربرد و تنظیمات دستگاه هستند، استفاده شده است.

• ویژگی های مبتنی بر شکل.

از هر توده قطعه بندی شده، ۱۲ ویژگی مبتنی بر شکل محاسبه شد. در ادامه هر یک از ویژگی ها به همراه روابط آن ها آمده است.

۱. ضریب شکل [۱۷،۱۰]: Area و Perimeter به ترتیب محیط و مساحت توده اند.

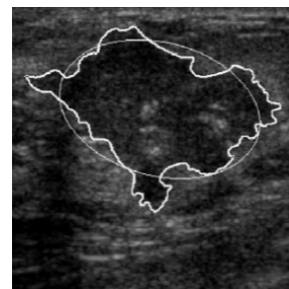
$$FF = \frac{4\pi \cdot Area}{Perimeter^2}$$

۲. دایروی بودن [۱۸،۱۷،۱۰]: Max_Diameter طول قطر بزرگ توده است.

$$Rnd = \frac{4 \cdot Area}{\pi \cdot Max_Diameter^2}$$

۳. نسبت محور بزرگ به محور کوچک [۱۹،۱۸،۶]: نسبت طول محور اصلی به محور فرعی بیضی معادل با توده است. مطابق شکل (۵) بیضی معادل، یک بیضی است هم مرکز با ضایعه با همان مساحت که ممان مرتبه دوم آن با ممان مرتبه دوم ضایعه برابر است.

$$L_s = \frac{Major_Axis}{Minor_Axis}$$



شکل (۵): نمایش بیضی معادل و هم مرکز با یک توده بدخیم [۵]

۴. ضریب تحدب [۱۸،۱۷،۱۰]: Convex_Perimeter محیط پوسته محدب است.

$$Cnv = \frac{Convex_perimeter}{Perimeter}$$

۵. جمود [۲۰،۱۸،۱۷،۱۰]: Convex_Area مساحت پوسته محدب است.

$$Sld = \frac{Area}{Convex_Area}$$

¹ Invariant elliptic Fourier descriptors

$$\psi = \arctan \left[\frac{c_1^*}{a_1^*} \right] \quad (0 \leq \psi \leq 2\pi)$$

$$IEFD = \text{abs} \left(\frac{B_x}{B_x(1)} \right) + \text{abs} \left(\frac{B_y}{B_y(1)} \right) \quad (6)$$

که در آن A_0 و C_0 به ترتیب اولین ضرایب توصیف‌گر فوریه در سیستم مختصات x و y و $B_x(1)$ و $B_y(1)$ اولین عناصر ماتریس سطری B_x و B_y هستند. در این پژوهش از چهار ضریب اول $IEFD$ به عنوان ویژگی جهت طبقه‌بندی توده‌ها استفاده شد ($N=4$).

❖ فاصله مرز توده و پوسته محدب

فاصله اقلیدسی بین مرز توده و پوسته محدب در هر نقطه محاسبه و سپس ۵ ویژگی میانگین، انرژي، آنتروپی، کشیدگی و چولگی از آن استخراج شد.

در مجموع تعداد ۲۱ ویژگی از هر توده استخراج شد که عبارتند از ۱۲ ویژگی مبتنی بر شکل، ۴ ویژگی از ضرایب توصیف‌گر فوریه بیضوی نامتغیر و ۵ ویژگی از فاصله مرز توده و پوسته محدب.

۲-۶ انتخاب ویژگی

الگوریتم بیشترین ارتباط-کمترین افزونگی یکی از روش‌های انتخاب ویژگی مبتنی بر اطلاعات متقابل و نوعی از الگوریتم جستجوی متوالی است. با توجه به مفهوم اطلاعات متقابل، هدف از انتخاب ویژگی یافتن مجموعه ویژگی است که به طور مشترک بیشترین وابستگی را به کلاس هدف دارند. این طرح با عنوان بیشترین وابستگی تعریف می‌شود. در این حالت احتمالاً ویژگی-های انتخاب شده براساس بیشترین ارتباط، افزونگی زیادی دارند. به این معنی که ممکن است وابستگی بین این ویژگی‌ها زیاد باشد. وقتی دو ویژگی به شدت به یکدیگر وابسته باشند در صورت حذف یکی از آن‌ها، قدرت تمایز طبقه‌بندی تغییر چندانی نخواهد کرد. بنابراین، شرط حداقل افزونگی که توسط Peng و همکاران [۲۴] ارائه شد، شرایط بیان شده را برآورده می‌کند [۲۳، ۲۴]. با توجه به توضیحات بیان شده، از روش mRMR برای انتخاب مؤثرترین ویژگی‌ها به منظور طبقه‌بندی استفاده شد.

۲-۷ طبقه‌بندی

در پژوهشی که توسط Yassin و همکاران انجام شد [۲۵]، به مروری بر الگوریتم‌های مختلف تشخیص سرطان پستان با روش‌های مختلف تصویربرداری پرداخته شد. در این تحقیق مشخص شد که از میان انواع مختلفی از الگوریتم‌های نظارت شده مورد استفاده بین سال‌های ۲۰۱۳ تا ۲۰۱۷، طبقه‌بند ماشین بردار پشتیبان عملکرد مطلوب و بیشترین استفاده را در گام طبقه‌بندی

$$Y_N(t) = C_0 + \sum_{n=1}^N c_n \cos \left(\frac{2n\pi t}{T} \right) + d_n \sin \left(\frac{2n\pi t}{T} \right) \quad (3)$$

در اینجا t ، گام مورد نیاز یک پیکسل واحد در امتداد مرز است. به این ترتیب که $t_{p-1} < t < t_p$ برای مقادیر p در محدوده $1 \leq p \leq K$ است. N تعداد هارمونیک‌های فوریه و K تعداد کل نقاط است. تعداد هارمونیک‌های مورد نیاز از طیف فوریه متوسط برآورده می‌شود. A_0 و C_0 ضرایب مربوط به فرکانس صفر هستند. عبارت‌های a_n, b_n, c_n, d_n نشان دهنده ضرایب فوریه بیضوی هارمونیک n هستند. با فرض درونیابی خطی بین نقاط همسایه، EFD ها در معادله ۳ از هارمونیک (a_n, b_n, c_n, d_n) را می‌توان با استفاده از تقریب فوریه گسسته محاسبه کرد.

ضرایب فوریه‌ای که با استفاده از رابطه ۳ تعیین می‌شوند متغیر با اندازه و جهت توده‌ها هستند. در نتیجه نمی‌توان آن‌ها را به عنوان شاخص تغییرپذیر برای مقایسه شکل توده‌ها در نظر گرفت. بنابراین، ضرایب فوریه براساس بیضی هارمونیک اول نرمال‌سازی می‌شوند. اگر ضرایب نرمال شده هارمونیک n م $a_n^{**}, b_n^{**}, c_n^{**}, d_n^{**}$ باشند، آنگاه:

$$\begin{bmatrix} a_n^{**} & b_n^{**} \\ c_n^{**} & d_n^{**} \end{bmatrix} = \frac{1}{E} \begin{bmatrix} \cos \psi & \sin \psi \\ -\sin \psi & \cos \psi \end{bmatrix} \begin{bmatrix} a_n & b_n \\ c_n & d_n \end{bmatrix} \begin{bmatrix} \cos n\theta & -\sin n\theta \\ \sin n\theta & \cos n\theta \end{bmatrix}$$

$$A_x = \sqrt{(a_n^{**2} + b_n^{**2})}$$

$$A_y = \sqrt{(c_n^{**2} + d_n^{**2})}$$

(۴)

$$B_x = \frac{A_x}{A_0} \quad B_y = \frac{A_y}{C_0}$$

(۵)

این روش استاندارد سازی براساس بیضی هارمونیک اول است. به این معنی که ضرایب بگونه‌ای اصلاح می‌شود که محور اصلی بیضی هارمونیک اول در راستای محور افقی قرار بگیرد و نیز اندازه ضرایب نسبت به اندازه محور اصلی بیضی، نرمال‌سازی می‌شود. Ψ و E^* به ترتیب طول و جهت محور اصلی بیضی هارمونیک اول هستند. θ فاز بردار شعاعی بیضی هارمونیک اول در $p=0$ (نقطه شروع کانتور) است. $a_n^*, b_n^*, c_n^*, d_n^*$ ضرایب فوریه استاندارد شده و $a_1^*, b_1^*, c_1^*, d_1^*$ به ترتیب ضرایب فوریه بیضوی شیفت یافته و اولین ضرایب فوریه هستند. ضرایب استاندارد بدست آمده، مستقل از اندازه، چرخش و نقطه شروع مرز هستند.

$$E^* = \sqrt{(a_1^{*2} + c_1^{*2})}, \quad \begin{bmatrix} a_1^* \\ c_1^* \end{bmatrix} = \begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix} \begin{bmatrix} \cos \theta \\ \sin \theta \end{bmatrix}$$

$$\theta = \frac{1}{2} \arctan \left[\frac{2(a_1 b_1 + c_1 d_1)}{(a_1^2 + c_1^2 - b_1^2 - d_1^2)} \right] \quad (0 \leq \theta \leq \pi)$$

¹ Max-Relevance and Min-Redundancy

- بخش رادیولوژی و سونوگرافی کلینیک تخصصی و فوق تخصصی باغبان شهرستان ساری با استفاده از دستگاه Phillips (Affiniti 50G) با پروب خطی مدل L12-4 با پهنای باند Radius Curvature: 4-12 MHZ، 34mm، و ذخیره تصاویر به فرمت jpeg.
 - بخش رادیولوژی و سونوگرافی مجتمع پزشکی تخصصی و فوق تخصصی دکتر شهیدی شهرستان ساری با استفاده از دستگاه Samsung (WS80A) با پروب خطی مدل L3-12A با پهنای باند 3-12MHZ و ذخیره تصاویر به فرمت jpeg.
- پایگاه داده دوم که توسط Gomez و همکاران [۱۱] جمع-آوری شده است، شامل ۱۰۰ تصویر توده خوش خیم و ۱۰۰ تصویر توده بدخیم همراه با تصاویر دودویی قطعه‌بندی شده توسط پزشک متخصص است. این تصاویر از موسسه ملی سرطان ریودوژانیرو برزیل^۴ و از سه نوع دستگاه تصویربرداری سونوگرافی با مبدل‌های خطی بدست آمد:
- General Electric (Logiq 5) at 12 MHz
 - General Electric (Logiq 7) at 10 MHz
 - Siemens (Sonoline Sienna) at 7.5 MHz

۲-۳ ارزیابی الگوریتم قطعه‌بندی

به منظور بررسی عملکرد الگوریتم قطعه‌بندی، از مقایسه نتایج روش پیشنهادی با تصاویر قطعه‌بندی شده‌ی دستی موجود در پایگاه داده دوم استفاده شده است. شباهت سنجی قطعه‌بندی توسط دو شاخص Dice و Jaccard انجام شد. این شاخص‌ها، همپوشانی بین تصاویر Ground Truth و خروجی الگوریتم قطعه‌بندی ارائه شده را اندازه‌گیری می‌کنند. تعریف این دو شاخص در رابطه ۱۰ و ۱۱ آمده است [۲۷، ۲۶].

$$Dice(A, B) = \frac{2|A \cap B|}{|A| + |B|} \quad (10)$$

$$Jaccard = \frac{|A \cap B|}{|A \cup B|} \quad (11)$$

که در اینجا A و B به ترتیب تصویر Ground Truth و تصویر حاصل از روش قطعه‌بندی هستند. هرچه عدد بدست آمده به مقدار ۱ نزدیک‌تر باشد نشان‌دهنده همپوشانی بیشتر و عملکرد مطلوب قطعه‌بندی است. به منظور مقایسه تاثیر تنظیم دستی و بهینه پارامترهای قطعه‌بندی، ابتدا ۲۰ تصویر از مجموع ۲۰۰ تصویر موجود در پایگاه داده دوم انتخاب و جهت محاسبه پارامترهای بهینه قطعه‌بندی در الگوریتم ژنتیک استفاده شد. به این صورت که برای این ۲۰ تصویر منتخب، اختلاف بین تصاویر

سیستم‌های تشخیص بیماری‌های پستان داشته است. از همین رو در پژوهش حاضر طبقه‌بند ماشین بردار پشتیبان با هسته تابع پایه شعاعی برای طبقه‌بندی توده‌ها به دو گروه خوش خیم و بدخیم مورد استفاده قرار گرفت.

۳- ارزیابی

به منظور ارزیابی دقیق و منصفانه سیستم‌های مختلف تشخیص رایانه‌ای لازم است تا از پایگاه داده مشابه و مشترک استفاده شود. از آنجایی که مرجع عمومی و مشخصی برای این کار در دسترس نیست، با وجود استفاده از معیارهای ارزیابی مشترک، تصمیم‌گیری در مورد مطلوب بودن عملکرد یک سیستم دشوار است. در ادامه ابتدا پایگاه داده مورد استفاده معرفی خواهد شد و سپس عملکرد روش قطعه‌بندی با کمک معیارهای شباهت سنجی Dice و Jaccard مورد ارزیابی قرار می‌گیرد. پس از آن نتایج طبقه‌بندی داده‌ها بیان می‌شود. در این مرحله متداول‌ترین معیارهای سنجش و بررسی عملکرد سیستم‌های تشخیص خودکار که عبارتند از دقت (صحت)^۱، حساسیت^۲ و قطعیت^۳ طبق روابط ۷ تا ۹ مورد استفاده قرار می‌گیرند. در تشخیص سرطان پستان دو معیار دقت و حساسیت از اهمیت بیشتری برخوردارند. این ارزیابی‌ها با استفاده از تمام ویژگی‌ها و نیز ویژگی‌های منتخب بدست آمده از الگوریتم انتخاب ویژگی mRMR صورت می‌پذیرد. به منظور سنجش عملکرد طبقه‌بندی از روش ارزیابی متقابل 5-Fold Cross Validation استفاده و کارایی الگوریتم پیشنهادی در تشخیص توده‌ها به دو گروه خوش خیم و بدخیم بررسی می‌گردد.

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (7)$$

$$Sensitivity = \frac{TP}{TP+FN} \quad (8)$$

$$Specificity = \frac{TN}{TN+FP} \quad (9)$$

۳-۱ پایگاه داده

در این پژوهش دو پایگاه داده مورد استفاده قرار گرفت. پایگاه داده اول متشکل از ۱۹۰ تصویر فراصوت پستان و شامل ۱۰۰ تصویر توده خوش خیم و ۹۰ تصویر از توده‌های بدخیم است که بدخیم بودن تمام نمونه‌های مشکوک به سرطان بر اساس آزمایشات پاتولوژی و هیستولوژی به اثبات رسیده است. این تصاویر طی مدت دو سال و از دو مرکز تصویربرداری جمع‌آوری شده‌اند:

¹ Accuracy

² Sensitivity

³ Specificity

⁴ National Cancer Institute (INCa)

جدول (۱): مقادیر پارامترهای تعیین شده با الگوریتم ژنتیک

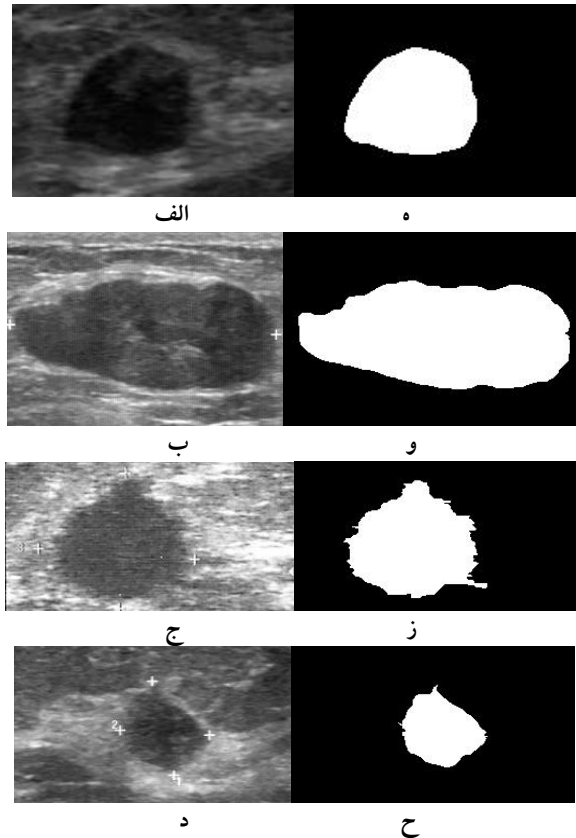
نام پارامتر	مقدار تعیین شده با الگوریتم ژنتیک
δ	۹۴/۱۸۸
λ	۶۶/۷۳
BW <u>indx</u> ^۱	۰/۸۴
Inside alpha	۲۲/۵۲
Outside mean	۴۰/۱۸
Outside standard deviation	۲۰/۳۷

۳-۳ ارزیابی طبقه‌بندی توده‌های خوش‌خیم و بدخیم

عملکرد طبقه‌بند ماشین بردار پشتیبان توسط روش ارزیابی متقابل 5-Fold Cross Validation مورد بررسی قرار گرفت. به این صورت که تمامی داده‌ها ابتدا به ۵ بخش تقسیم شده و در تکرارهای مختلف هر بار یک بخش به عنوان داده آزمایش و باقی داده‌ها به عنوان داده آموزش مورد استفاده قرار گرفتند. با این کار تمامی داده‌ها در گام آموزش و آزمایش به کار گرفته شدند. نتیجه نهایی از میانگین این تکرارها حاصل شد. نتایج حاصل از طبقه‌بندی داده‌ها با استفاده از تعداد متفاوتی از ویژگی‌های منتخب که توسط روش انتخاب ویژگی مبتنی بر اطلاعات متقابل حاصل شد، مورد تجزیه و تحلیل قرار گرفتند. همچنین با استخراج ویژگی از تصاویر طبقه‌بندی شده‌ی دستی موجود در پایگاه داده دوم و طبقه‌بندی آنها تاثیر دقت طبقه‌بندی بر عملکرد طبقه‌بند ارزیابی شد. نتایج نشان می‌دهد که تنها با استفاده از تعداد محدودی ویژگی منتخب، طبقه‌بند SVM عملکرد مطلوبی در تفکیک توده‌ها به دو گروه خوش‌خیم و بدخیم دارد. هرچند با استفاده از تعداد بیشتری از ویژگی‌ها نیز میتوان به دقت مشابه دست یافت اما این کار موجب افزایش هزینه محاسباتی و کاهش سرعت طبقه‌بند می‌گردد. همچنین نتایج نشان می‌دهد که استفاده از ویژگی‌های مبتنی بر شکل، عملکرد قابل قبولی در طبقه‌بندی دارد. جزئیات نتایج طبقه‌بندی و تاثیر طبقه‌بندی و استفاده از ویژگی‌های منتخب بر عملکرد آن در جدول (۳) قابل مشاهده است. بر اساس مقادیر این جدول، مشاهده می‌شود که ۵ ویژگی منتخب توسط روش بیشترین ارتباط - کمترین افزونگی عملکرد مطلوبی را در طبقه‌بندی ارائه می‌دهد. نکته قابل ذکر در این جدول این است که طبقه‌بند تنظیم شده با استفاده از ۲۰ تصویر از پایگاه دوم، برای طبقه‌بندی پایگاه داده اول نیز نتایج قابل قبولی نشان داده است.

در جدول (۴) مقایسه نتایج بدست آمده از طبقه‌بندی توده‌های پستان بین الگوریتم‌های پیشنهادی و چند پژوهش مرتبط دیگر بر اساس پایگاه داده مشترک و روش طبقه‌بندی مشابه به اختصار آمده است. شکل (۷) نیز نمودار منحنی مشخصه گیرنده^۱ برای بالاترین دقت بدست آمده بر روی تصاویر پایگاه داده دوم با ۵

قطعه‌بندی شده‌ی دستی و تصاویر حاصل از الگوریتم طبقه‌بندی توسط دو شاخص Dice و Jaccard محاسبه شد. شرط توقف این الگوریتم تولید ۶۰ نسل کروموزوم تعریف شد و بر این اساس مناسب‌ترین مقادیر برای پارامترها بدست آمد. مقادیر (زیر) بهینه پارامترهای بدست آمده از این روش در جدول (۱) بیان شده است. در ادامه کارایی مقادیر حاصل بر روی ۱۸۰ تصویری که در فاز



شکل (۶): الف تا د توده‌های خوش‌خیم و بدخیم و ه تا ح به ترتیب تصاویر طبقه‌بندی شده آن‌ها را نشان می‌دهد

آموزش مورد استفاده قرار نگرفتند، بررسی گردید. در شکل (۶) نتایج حاصل از اعمال الگوریتم طبقه‌بندی روی ۴ تصویر مختلف قابل مشاهده است.

نتیجه ارزیابی روش طبقه‌بندی بر روی پایگاه داده دوم برای هر دو کلاس در جدول (۲) بر اساس سه معیار Dice، Jaccard و Hausdorff آمده است. مشاهده می‌شود که با انتخاب بهینه پارامترها توسط الگوریتم ژنتیک، دقت طبقه‌بندی بین ۵ تا ۱۰ درصد بهبود یافته است. هرچند به نظر می‌رسد که دقت نتایج طبقه‌بندی، مطلوب نباشد، اما این نتایج بگونه‌ای است که عموماً اکثر ویژگی‌های توده‌های خوش‌خیم و بدخیم را حفظ می‌کند (مانند کشیدگی یا دایروی بودن توده و نامنظمی یا همواری مرز و ...). در نتیجه ویژگی‌های مورد نظر برای طبقه‌بندی کمتر تحت تاثیر عدم دقت طبقه‌بندی قرار می‌گیرد. نتایج طبقه‌بندی با توجه به آنچه در بخش بعدی آمده نیز این ادعا را تایید می‌کند.

^۱ Receiver Operating Characteristic(ROC)

جدول (۲) نتیجه ارزیابی الگوریتم قطعه بندی بر روی پایگاه داده دوم

گروه داده	معیار Dice		معیار Jaccard		معیار Hausdorff	
	تنظیم دستی پارامتر	تنظیم پارامتر با الگوریتم ژنتیک	تنظیم دستی پارامتر	تنظیم پارامتر با الگوریتم ژنتیک	تنظیم دستی پارامتر	تنظیم پارامتر با الگوریتم ژنتیک
۲۰ تصویر آموزش (خوش-خیم و بدخیم)	۰/۸۱	۰/۸۷	۰/۶۸	۰/۷۷	۱۵/۳۱	۱۳/۵۹
خوش خیم	۰/۷۹	۰/۸۵	۰/۶۵	۰/۷۴	۱۶/۰۳	۱۴/۶۶
بدخیم	۰/۷۷	۰/۸۱	۰/۶۳	۰/۶۸	۱۷/۷۹	۱۶/۶۸
۱۸۰ تصویر آزمایش پایگاه داده	۰/۷۸	۰/۸۳	۰/۷۱	۰/۷۷	۱۷/۱۰	۱۵/۳۵

جدول (۳) نتایج طبقه بندی توده های خوش خیم و بدخیم بر روی دو پایگاه داده و دو روش قطعه بندی خودکار و دستی

پایگاه داده	۲۱ ویژگی			۱۵ ویژگی			۱۰ ویژگی			۵ ویژگی		
	Acc %	Sen %	Spec %	Acc %	Sen %	Spec %	Acc %	Sen %	Spec %	Acc %	Sen %	Spec %
پایگاه داده اول	۹۴/۵	۹۶	۹۳	۹۵	۹۵	۹۵	۹۳/۵	۹۴	۹۳	۹۴/۵	۹۵	۹۴
پایگاه داده دوم	۹۴/۵	۹۵	۹۴	۹۴	۹۵	۹۳	۹۳	۹۵	۹۱	۹۶/۵	۹۹	۹۴
قطعه بندی دستی	۹۵/۵	۹۳	۹۸	۹۵/۵	۹۲	۹۹	۹۴	۹۱	۹۷	۹۵/۵	۹۳	۹۸

جدول (۴) مقایسه نتایج روش پیشنهادی با سایر مطالعات مرتبط جهت طبقه بندی ضایعات پستان

مقاله	تعداد تصاویر	پیش-پردازش	قطعه بندی	تعداد ویژگی ها	طبقه بند	نوع تشخیص	اعتبار سنجی	دقت	حساسیت	قطعیت
Rodríguez-Cristerna 2018 [28]	۱۱۲۸	×	تبدیل آب پخشان	۲۱۶	جنگل تصادفی	کلاس BIRADS	-	AUC ۰/۸۷	۰/۸۲	۰/۹۱
Flores WG 2015 [9]	۶۴۱	×	تبدیل آب پخشان	۱۴۹۱ (۵ ویژگی منتخب)	LFDA	خوش خیم و بدخیم	-	AUC ۰/۹۴	-	-
Rodríguez-Cristerna 2019 [29]	۲۰۵۰	-	-	۳۲۵	SVM	خوش خیم و بدخیم	10 Fold CV	AUC ۰/۸۲	۰/۸۰	۰/۸۵
González-Luna FA 2019 [30]	۲۰۳۲	-	دستی	۱۳۷	LDA	خوش خیم و بدخیم	10 Fold CV	۰/۸۹	۰/۸۲	۰/۹۳
Gómez-Flores W 2020 [31]	۲۰۵۴	-	دستی	۳۹	LDA	خوش خیم و بدخیم	10 Fold CV	۰/۸۸	۰/۸۱	۰/۹۱
روش پیشنهادی	۲۰۰	-	دستی	۲۱ (۵ ویژگی منتخب)	SVM	خوش خیم و بدخیم	5 Fold CV	۰/۹۵۵	۰/۹۳	۰/۹۸
روش پیشنهادی	۱۹۰	NLM	مبتنی بر برش گراف	۲۱ (۵ ویژگی منتخب)	SVM	خوش خیم و بدخیم	5 Fold CV	۰/۹۶۵	۰/۹۹	۰/۹۴

۴- بحث

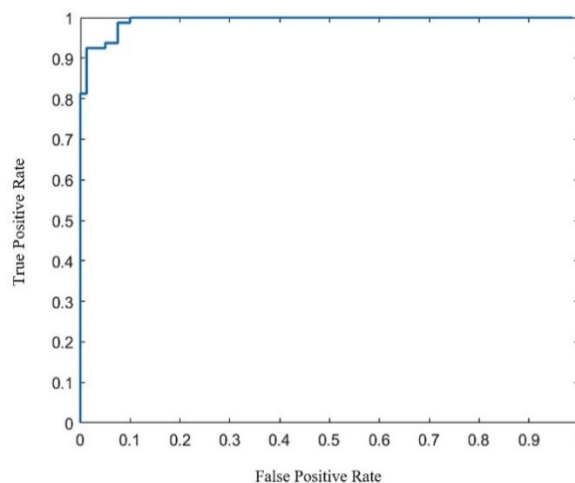
همانگونه که در بخش ۲ بیان شد، ۴ ویژگی از ۲۱ ویژگی مورد استفاده در طبقه‌بندی، ۴ ضریب اول تبدیل فوریه بیضوی نامتغیر است. واضح است که به ازای $N=4$ تقریب فوریه، جزئیات مرز را نمی‌تواند نشان دهد و بنابراین ضرایب فوریه، ویژگی‌های کافی برای طبقه‌بندی فراهم نمی‌کند و استفاده از سایر ویژگی‌ها ضرورت دارد. به عنوان مثال برای یکی از تصاویر نشان داده شده در شکل (۶) (تصویر ۶-الف)، کانتور حاصل از تقریب فوریه تا مرتبه ۴ و میزان شاخص Jaccard برای بیان کمی شباهت کانتور تقریبی با کانتور واقعی محاسبه و نتایج آن در شکل (۸) نمایش داده شده است. دیده می‌شود که با افزایش درجه تقریب، شباهت بیشتر می‌شود اما تغییرات آن کند می‌شود و برای نزدیک کردن آن به حداکثر مقدار ممکن یک، درجه تقریب باید خیلی بزرگ شود و در نتیجه تعداد ویژگی‌ها زیاد می‌شود. با اینحال بیان جزئیات در روی کانتور اصلی، بخاطر نویزهای روی تصویر و خطاهای قطعه‌بندی نمی‌تواند ویژگی‌های مناسبی را معرفی کند و بنابراین قادر به بهبود دقت طبقه‌بندی نخواهد بود. در عمل هم مشاهده شده که افزایش مقدار N بیشتر از ۴ تاثیری در دقت طبقه بندی ندارد. بنابراین استفاده از ۴ ضریب اول فوریه در حل مساله مورد نظر مورد توجه قرار گرفته است.

۵- جمع‌بندی

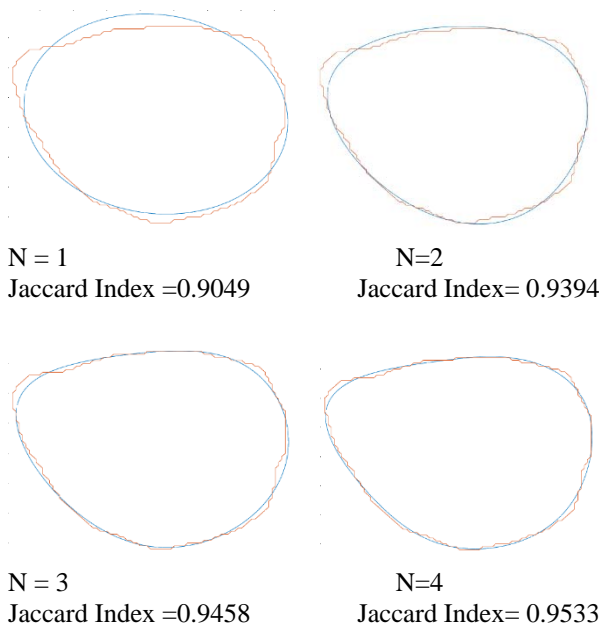
در این پژوهش، طراحی و توسعه یک سیستم رایانه‌ای تشخیص خودکار سرطان پستان از روی تصاویر فراصوت با هدف بهبود عملکرد تشخیص این بیماری، ارائه شد. پس از جمع‌آوری مجموعه‌ای شامل ۳۹۰ تصویر سونوگرافی پستان، گام پیش‌پردازش جهت بهبود نواحی توده، با استفاده از یک روش سه مرحله‌ای تحقق یافت. ابتدا تصویر با اعمال یک فیلتر پایین‌گذر گوسی دو بعدی هموار شد. در ادامه و با هدف حذف تغییرات محدوده شدت روشنایی، یک نرمالیزاسیون خطی مورد استفاده قرار گرفت. در آخرین مرحله از گام پیش‌پردازش، نواحی کم پژواک تصویر توسط یک تابع وفقی بهبود داده شد. پیش از قطعه‌بندی، ابتدا به کمک دو الگوریتم خودکار، نقطه بذر درون ناحیه توده انتخاب و سپس ناحیه مطلوب شامل بافت توده از تصویر اصلی جدا شد. در ادامه و با هدف حذف نویز و بهبود تضاد، تصویر حاصل از یک فیلتر میانگین‌های غیرمحملی عبور داده شد. در مرحله قطعه‌بندی با بهره‌گیری از اطلاعات فرکانسی و مکانی توده، یک تابع انرژی مبتنی بر این دو قید تعریف گردید. سپس الگوریتم بهینه‌سازی جریان بیشینه-برش کمینه، بهترین حالت را برای جدا کردن بافت توده از پس‌زمینه ایجاد کرد که باعث تشکیل تصویر دودویی نهایی گردید. همچنین به منظور تعیین دقیق‌تر و خودکار مقادیر پارامترهای قطعه‌بندی، از الگوریتم بهینه‌سازی ژنتیک بهره گرفته شد. در گام بعد، استخراج سه دسته ویژگی مبتنی بر شکل و

ویژگی منتخب را نشان می‌دهد. بیشترین میزان سطح زیر منحنی نیز برابر 0.98 بدست آمده است.

این پژوهش با استفاده از نرم افزار Matlab نسخه ۲۰۱۸ و بر روی سیستمی با حافظه (RAM) ۶ گیگا بایت و پردازنده مرکزی ۵ هسته‌ای (2.7GHz) انجام شد. زمان اجرای الگوریتم قطعه‌بندی بر روی تصویر سطح خاکستری با ابعاد 171×275 به طور متوسط برابر با $10/38$ ثانیه به ثبت رسیده است. همچنین زمان لازم جهت انجام فرآیند طبقه‌بندی به طور تقریبی برابر با 0.24 ثانیه است.



شکل (۷): نمودار منحنی مشخصه گیرنده برای بهترین نتیجه بدست آمده



شکل (۸): کانتورهای حاصل از تقریب فوریه با درجات مختلف برای تصویر ۶-الف به همراه شاخص شباهت Jaccard

مراجع

- مرز توده صورت گرفت. دسته اول شامل ۱۲ ویژگی از شکل و روابط هندسی حاکم بر آن محاسبه شد. در دسته دوم با استفاده از تعریف فاصله مرز توده و پوسته محدب، ۵ ویژگی آماری از مرز توده بدست آمد. ۴ ویژگی دسته سوم نیز از ضرایب توصیف‌گر فوریه بیضوی نامتغیر از مرز توده حاصل شد. به این ترتیب در مجموع ۲۱ ویژگی از هر تصویر پایگاه داده استخراج گردید. به منظور کاهش ابعاد فضای ویژگی و بهبود عملکرد طبقه‌بندی از روش انتخاب ویژگی مبتنی بر اطلاعات متقابل استفاده شد. در آخرین مرحله، از طبقه‌بند ماشین بردار پشتیبان برای تفکیک توده‌ها به دو گروه خوش‌خیم و بدخیم بهره گرفته شد.
- روش پیشنهادی بر روی تصاویر دو پایگاه داده اعمال گردید. بررسی عملکرد سیستم توسط معیارهای ارزیابی دقت، حساسیت و قطعیت انجام شد. همچنین کارایی الگوریتم طبقه‌بندی با دو شاخص شباهت‌سنجی Dice و Jaccard مورد ارزیابی قرار گرفت. با توجه به این نکته که تعداد قابل توجهی از تصاویر موجود در پایگاه داده دوم که به منظور ارزیابی گام طبقه‌بندی مورد استفاده شد، از کیفیت مطلوبی برخوردار نبودند با این وجود نتایج بدست آمده از بررسی عملکرد روش طبقه‌بندی در حالتی که تنظیم پارامترهای الگوریتم به صورت دستی انجام شد میزان ۰/۷۸ و زمانی که این مقادیر توسط الگوریتم بهینه‌سازی ژنتیک تعیین شده باشد به میزان ۰/۸۳ افزایش یافت. عملکرد طبقه‌بندی در تشخیص توده‌های خوش‌خیم از سرطانی بر روی تصاویر پایگاه داده جمع-آوری شده میزان دقت، حساسیت و قطعیت ۹۵٪ را با ۱۵ ویژگی منتخب نشان داد. این مقادیر برای تصاویر پایگاه داده دوم برابر است با دقت ۹۶/۵٪، حساسیت ۹۹٪ و قطعیت ۹۴٪ با استفاده از ۵ ویژگی بهینه. نتایج طبقه‌بندی با تصاویر موجود در پایگاه داده دوم که به صورت دستی طبقه‌بندی شده بودند با میزان دقت ۹۵/۵٪، حساسیت ۹۳٪ و قطعیت ۹۸٪ با ۵ ویژگی بهینه حاصل گردید. مقایسه سه حالت مختلف طبقه‌بندی، اهمیت گام طبقه‌بندی و استفاده از ویژگی‌های بهینه بر عملکرد سیستم رایبان می‌کند.
- در پژوهش حاضر تلاش بر این بود تا با کمک الگوریتم‌های مختلف از جمله الگوریتم بهینه‌سازی ژنتیک، تا حد امکان از دخالت کاربر در فرآیند طبقه‌بندی و تشخیص توده جلوگیری شود. در گام استخراج ویژگی نیز استفاده از ویژگی‌های مبتنی بر مرز، شامل فاصله مرز و پوسته محدب و توصیف‌گر فوریه بیضوی نا متغیر در کنار ویژگی‌های شکل نشان داد ترکیب این ویژگی‌ها نتایج قابل قبولی را ارائه می‌دهند. همچنین با توجه به اینکه پایگاه داده متشکل از تصاویر ۵ دستگاه مختلف است، نشان داده شد که روش پیشنهادی از قدرت تعمیم خوبی برخوردار بوده و نسبت به تنظیمات دستگاه و کیفیت تصویر مقاوم است.
- در آخر به منظور بهبود پژوهش انجام شده استفاده از تعداد بیشتری از تصاویر در گام آموزش فرآیند تنظیم پارامتر طبقه‌بندی و افزایش تعداد نسل‌های الگوریتم ژنتیک پیشنهاد می‌گردد.
- [1] Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries.
- [2] نعمت هدا، محلوجی فر علی، گویا علی، احمدی‌نژاد نسرن، "کلاس‌بندی تومورهای خوش‌خیم و بدخیم در تصاویر اولتراسوند پستان با کمک ویژگی‌های ریخت‌شناسی"، مجله ماشین بینایی و پردازش تصویر، دوره ۴، شماره ۲، صفحه ۷۵-۸۹، ۲۰۱۷.
- [3] Jesneck, J. L., Lo, J. Y., & Baker, J. A., "Breast mass lesions: computer-aided diagnosis models with mammographic and sonographic descriptors," Radiology, Vol. 244, No.2, pp. 390-398, 2007.
- [4] Ohuchi, N., Suzuki, A., Sobue, T., Kawai, M., Yamamoto, S., Zheng, Y.F., Shiono, Y.N., Saito, H., Kuriyama, S., Tohno, E. and Endo, T., "Sensitivity and specificity of mammography and adjunctive ultrasonography to screen for breast cancer in the Japan Strategic Anti-cancer Randomized Trial (J-START): a randomised controlled trial," The Lancet, Vol. 387, No.10016, pp. 341-348, 2016.
- [5] Jalalian, A., Mashohor, S. B., Mahmud, H. R., Saripan, M. I. B., Ramli, A. R. B., & Karasfi, B., "Computer-aided detection/diagnosis of breast cancer in mammography and ultrasound: a review," Clinical imaging, 37(3), 420-426, 2013.
- [6] Cheng, H. D., Shan, J., Ju, W., Guo, Y., & Zhang, L., "Automated breast cancer detection and classification using ultrasound images: A survey," Pattern recognition, Vol. 43, No. 1, pp. 299-317, 2010.
- [7] Xian, M., Zhang, Y., Cheng, H. D., Xu, F., Zhang, B., & Ding, J., "Automatic breast ultrasound image segmentation: A survey," Pattern Recognition, Vol. 79, pp. 340-355, 2018.
- [8] Shan, J., Alam, S. K., Garra, B., Zhang, Y., & Ahmed, T., "Computer-aided diagnosis for breast ultrasound using computerized BI-RADS features and machine learning methods," Ultrasound in medicine & biology, Vol. 42, No.4, pp. 980-989, 2016.
- [9] Abdel-Nasser, M., Melendez, J., Moreno, A., Omer, O. A., & Puig, D., "Breast tumor classification in ultrasound images using texture analysis and super-resolution methods," Engineering Applications of Artificial Intelligence, Vol. 59, pp. 84-99, 2017.
- [10] Flores, W. G., de Albuquerque Pereira, W. C., & Infantosi, A. F. C., "Improving classification performance of breast lesions on ultrasonography," Pattern Recognition, Vol. 84, No. 4, pp. 1125-1136, 2015.
- [11] Gomez-Flores, W., & Ruiz-Ortega, B. A., "New fully automated method for segmentation of breast lesions on ultrasound based on texture analysis," Ultrasound in

- [24] Peng, H., Long, F., & Ding, C. "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy," IEEE Transactions on pattern analysis and machine intelligence, Vol. 27, No. 8, pp. 1226-123, 2005.
- [25] Yassin, N. I., Omran, S., El Houby, E. M., & Allam, H. "Machine learning techniques for breast cancer computer aided diagnosis using different image modalities: A systematic review," Computer methods and programs in biomedicine, Vol. 156, pp. 25-45, 2018.
- [26] Kozegar, E., Soryani, M., Behnam, H., Salamati, M., & Tan, T. "Mass segmentation in automated 3-D breast ultrasound using adaptive region growing and supervised edge-based deformable model," IEEE transactions on medical imaging, Vol. 37, No. 4, pp. 918-928, 2017.
- [27] Rouhi, R., Jafari, M., Kasaei, S., & Keshavarzian, P. "Benign and malignant breast tumors classification based on region growing and CNN segmentation," Expert Systems with Applications, Vol. 42, No. 3, pp.990-1002, 2015.
- [28] Rodríguez-Cristerna, A., Gómez-Flores, W., & de Albuquerque Pereira, W. C. "A computer-aided diagnosis system for breast ultrasound based on weighted BI-RADS classes," Computer methods and programs in biomedicine, Vol. 153, pp. 33-40, 2018.
- [29] Gómez-Flores, W., Rodríguez-Cristerna, A., & de Albuquerque Pereira, W. C. "Texture analysis based on auto-mutual information for classifying breast lesions with ultrasound," Ultrasound in medicine & biology, Vol. 45, No. 8, pp. 2213-2225, 2019.
- [30] González-Luna, F. A., Hernández-López, J., & Gómez-Flores, W. "A performance evaluation of machine learning techniques for breast ultrasound classification," In 2019 16th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE) (pp. 1-5), IEEE, 2019.
- [31] Gómez-Flores, W., & Hernández-López, J. "Assessment of the invariance and discriminant power of morphological features under geometric transformations for breast tumor classification," Computer methods and programs in biomedicine, Vol. 185, pp. 105173, 2020.
- medicine & biology, Vol. 42, No. 7, pp. 1637-1650, 2016.
- [12] Xian, M., Zhang, Y., & Cheng, H. D. "Fully automatic segmentation of breast ultrasound images based on breast characteristics in space and frequency domains," Pattern Recognition, Vol. 48, No. 2, pp. 485-497, 2015.
- [13] Otsu, N., "A threshold selection method from gray-level histograms," IEEE transactions on systems, man, and cybernetics, Vol. 9, No. 1, pp. 62-66, 1979.
- [14] Buades, A., Coll, B., & Morel, J. M., "A review of image denoising algorithms, with a new one," Multiscale modeling & simulation, Vol. 4, No.2, pp.490-530, 2005.
- [15] Boykov, Y., & Kolmogorov, V., "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," IEEE transactions on pattern analysis and machine intelligence, Vol. 26, No.9, pp. 1124-1137, 2004.
- [16] Goyal, B., Dogra, A., Agrawal, S., Sohi, B. S., & Sharma, A. "Image denoising review: From classical to state-of-the-art approaches," Information fusion, Vol. 55, pp. 220-244, 2020.
- [17] Chang, R. F., Wu, W. J., Moon, W. K., & Chen, D. R., "Automatic ultrasound segmentation and morphology based diagnosis of solid breast tumors," Breast cancer research and treatment, Vol. 89, No. 2, pp. 179-185, 2005.
- [18] Yap, M. H., Edirisinghe, E. A., & Bez, H. E. "A comparative study in ultrasound breast imaging classification," In Medical Imaging 2009: Image Processing (Vol. 7259, p. 72591S). International Society for Optics and Photonics,
- [19] Claridge, E., Hall, P. N., Keefe, M., & Allen, J. P. "Shape analysis for classification of malignant melanoma," Journal of biomedical engineering, Vol. 14, No. 3, pp. 229-234, 1992.
- [20] Zakeri, F. S., Behnam, H., & Ahmadinejad, N., "Classification of benign and malignant breast masses based on shape and texture features in sonography images," Journal of medical systems, Vol. 36, No. 3, pp. 1621-1627, 2012.
- [21] Alvarenga, A. V., Infantosi, A. F. C., Pereira, W. C. A., & Azevedo, C. M., "Assessing the performance of morphological parameters in distinguishing breast tumors on ultrasound images," Medical engineering & physics, Vol. 32, No. 1, pp. 49-56.
- [22] Mebatsion, H. K., Paliwal, J., & Jayas, D. S., "A novel, invariant elliptic Fourier coefficient based classification of cereal grains," Biosystems engineering, Vol. 111, No. 4, pp. 422-428, 2012.
- [23] Chandrashekar, G., & Sahin, F. "A survey on feature selection methods," Computers & Electrical Engineering, Vol. 40, No. 1, pp. 16-28, 2014.

وجیهه سادات وطن پور تحصیلات کارشناسی خود را در سال ۱۳۹۰ در رشته مهندسی برق گرایش الکترونیک آغاز و در سال ۱۳۹۴ فارغ التحصیل شد. مدرک کارشناسی ارشد خود را در رشته مهندسی پزشکی گرایش بیوالکترونیک در سال ۱۴۰۰ از دانشگاه صنعتی نوشیروانی بابل دریافت کرد.

سید محمود سخایی مقاطع تحصیلی کارشناسی، کارشناسی ارشد و دکتری خود را بترتیب در دانشگاه تهران، صنعتی امیرکبیر و تربیت مدرس گذرانده است. وی از سال ۱۳۹۰ عضو هیات علمی دانشکده مهندسی برق و کامپیوتر دانشگاه صنعتی نوشیروانی بابل است. زمینه های پژوهشی وی شامل پردازش سیگنال های پزشکی و تصویربرداری اولتراسوند است.

یاسر بالغی مقاطع تحصیلی کارشناسی، کارشناسی ارشد و دکتری خود را در دانشگاه علم و صنعت ایران گذرانده است. وی از سال ۱۳۸۸ عضو هیات علمی دانشکده مهندسی برق و کامپیوتر دانشگاه صنعتی نوشیروانی بابل است. زمینه های پژوهشی وی شامل پردازش تصویر و الکترونیک دیجیتال است.

الهام سادات بنی مصطفوی فارغ التحصیل رشته پزشکی عمومی در سال ۱۳۸۵ از دانشگاه علوم پزشکی ایران می باشد. وی در سال ۱۳۹۰ موفق به اخذ درجه تخصصی در رشته رادیولوژی از دانشکده پزشکی دانشگاه علوم پزشکی مازندران گردید. دوره فوق تخصصی بیماری های پستان را در مجتمع تصویربرداری بیمارستان امام خمینی (ره) دانشگاه علوم پزشکی تهران گذرانده اند. ایشان اکنون به عنوان عضو هیئت علمی در بخش رادیولوژی بیمارستان امام خمینی (ره) ساری دانشگاه علوم پزشکی مازندران مشغول به کار می باشند