

ارایه یک سیستم کمکی کامپیوتری خودکار جهت تشخیص خوش خیمی و بدخیمی توده‌ها در تصاویر ماموگرافی

امید رحمانی سرباست: گروه مهندسی برق و الکترونیک، دانشگاه حکیم سبزواری، سبزوار، ایران
 جواد حدادنیا*: گروه مهندسی پزشکی، دانشگاه حکیم سبزواری، سبزوار، ایران
 حسین قیومی‌زاده: گروه مهندسی پزشکی، دانشگاه حکیم سبزواری، سبزوار، ایران

چکیده

مقدمه: ماموگرافی رایج‌ترین روش غربالگری برای تشخیص سرطان پستان است. در این مقاله یک سیستم کمکی برای تشخیص خوش خیمی و بدخیمی توده‌ها ارایه شده است.

روش بررسی: در این روش ابتدا با استفاده از حذف نویز و افزایش وضوح تصویر، توده‌ها را آماده بخش‌بندی می‌کنیم. سپس با استفاده از یک الگوریتم جدید بر مبنای رشد ناحیه، توده مورد نظر را بخش‌بندی کرده و ویژگی‌های از بافت و مرز آن استخراج می‌کنیم. در ادامه با استفاده از یک معماری نوین رده‌بندهایی را با یکدیگر ترکیب کرده و از این رده‌بند مرکب جهت دسته‌بندی خوش خیمی و بدخیمی توده‌ها بهره می‌گیریم.

یافته‌ها: این سیستم بر روی دو پایگاه داده مورد ارزیابی قرار گرفته است و نتایج حاصل نشان‌دهنده این است که سیستم کمکی پیشنهادی از لحاظ دقت کاملاً قابل رقابت با سیستم‌های مدرن موجود می‌باشد. دقت حاصل شده در پایگاه داده MIAS برابر با ۹۳٪ و در پایگاه داده DDSM برابر با ۹۰٪ بوده است.

نتیجه‌گیری: به طور کلی می‌توان نوآوری‌های موجود در این مقاله را در معرفی یک سیستم خودکار تطبیقی بر مبنای الگوریتم رشد ناحیه جهت بخش‌بندی توده‌ها، استفاده از توصیف‌گرهای مبتنی بر تجزیه مود ذاتی و معرفی یک معماری جدید برای ترکیب رده‌بندها استفاده کرد.

واژه‌های کلیدی: سرطان پستان، ماموگرافی، خوش خیمی و بدخیمی.

* نشانی نویسنده پاسخگو: خراسان رضوی، سبزوار، دانشگاه حکیم سبزواری، گروه مهندسی پزشکی، جواد حدادنیا.
 نشانی الکترونیک: haddadnia@hsu.ac.ir

مقدمه

سرطان پستان یک بیماری مهم در کشورهای مختلف (خصوصاً کشورهای غربی) محسوب می‌شود. طبق آمار، سرطان پستان بعد از سرطان ریه رایج‌ترین سرطان (۱۰٫۹٪) از سرطان‌ها در زنان و مردان ناشی از سرطان پستان است) و پنجمین عامل مرگ‌های ناشی از سرطان محسوب می‌شود (۱). گرچه ابتلا به سرطان پستان در دهه گذشته نرخ صعودی داشته است. اما مرگ و میر ناشی از این سرطان در بین زنان (در تمامی سنین) کاهش یافته است. این سیر نزولی به استفاده گسترده از تصویربرداری پزشکی مربوط می‌شود که با استفاده از آن رادیولوژیست‌ها می‌توانند سرطان را در مراحل ابتدایی شناسایی کنند (۲). شناسایی زودهنگام بهترین روش مقابله با بیماری سرطان پستان می‌باشد. چون آمار نشان می‌دهد که ۹۶٪ از سرطان‌هایی که در مرحله مقدماتی باشند قابل درمان هستند (۱). ماموگرافی رایج‌ترین روش تصویربرداری از پستان است. این روش به تنهایی نرخ مرگ و میر ناشی از سرطان پستان را به اندازه ۲۵ تا ۳۰ درصد کاهش داده است (۲).

یکی از مهم‌ترین علایم سرطان پستان وجود توده در تصاویر ماموگرافی است. در طول تقریباً ۳ دهه اخیر، دانشمندان کامپیوتر بسیار تلاش کردند تا به رادیولوژیست‌ها در شناسایی و تشخیص نوع توده‌ها کمک کنند. هدف از طراحی سیستم‌های CAD کمک به رادیولوژیست‌ها در شناسایی و تشخیص اختلالات موجود در تصاویر می‌باشد. آمار نشان می‌دهد که بین ۶۵ تا ۹۰ درصد از مواردی که جهت انجام بیوپسی به پاتولوژی ارجاع داده می‌شوند مربوط به نمونه‌های خوش‌خیم می‌باشند (۳).

تصاویر ماموگرافی مانند بسیاری دیگر از تصاویر پزشکی به علت ویژگی‌های خاصی که دارند به راحتی قابل تحلیل نیستند. به همین دلیل یکی از اقداماتی که در طراحی یک سیستم تشخیص پزشکی می‌تواند بسیار موثر باشد بهبود کیفیت تصویر است (۴). روش‌هایی که برای بهبود کیفیت تصاویر ماموگرافی به کار گرفته می‌شوند بسیار متنوع‌اند. بعضی از روش‌ها بر مبنای پردازش سراسری هیستوگرام استوار هستند. روش‌های دیگری مانند یک‌نواختی

هیستوگرام چندقله‌ای^۱ و بسط هیستوگرام^۲ نیز وجود دارند (۵). در (۶) روشی ارایه شده که با استفاده از خصوصیات آماری محلی و همچنین تصویر گرادینتی^۳ کیفیت تصویر را بهبود می‌بخشد. روش‌های دیگری نیز وجود دارند که بر پایه پردازش تصویر در چند مقیاس مانند تبدیل موجک استوار هستند. برای دسته‌بندی توده‌ها لازم است ویژگی‌هایی از توده‌ها استخراج شوند که بتوانند بین دو دسته مورد نظر تمایز ایجاد کنند. این ویژگی‌ها مربوط به روشنایی توده، بافت توده، مرز توده و همچنین ویژگی‌های هندسی توده می‌باشند. ویژگی‌های مربوط به روشنایی توده از اطلاعات آماری مربوط به روشنایی پیکسل‌های موجود در ناحیه‌ی تومور استخراج می‌شوند. اطلاعاتی از قبیل مقادیر کمینه و بیشینه، میانگین، واریانس، چولگی^۴ و برجستگی^۵ را می‌توان در این دسته قرار داد (۷). معمولاً در کارهای تشخیص پزشکی با استفاده از سیستم‌های کامپیوتری و هوش مصنوعی، ویژگی‌های متنوع و متعددی استخراج می‌شود. بدیهی است که از بین این ویژگی‌ها تعدادی از آنها ممکن است جزء ویژگی‌های مازاد و یا نامربوط محسوب شوند و بدین ترتیب سبب افت دقت سیستم شوند. بنابراین می‌توان برای بهبود دقت رده‌بند از متدهای انتخاب ویژگی بهره گرفت. به عنوان مثال در (۸) از روش انتخاب ویژگی گام به گام استفاده شده است. در روش دیگری که توسط (۹) ارایه شده است، روش SVM-RFE را با تکنیکی به نام NMIFS ترکیب کردند.

در این مقاله یک استراتژی جدید برای رده‌بندی نمونه‌ها پیشنهاد شده است که به نوع رده‌بند مورد استفاده چندان وابسته نیست و در واقع یک چارچوب جدید برای ترکیب رده‌بندها است.

¹ Multi-Peak Histogram Equalization

² Histogram Stretching

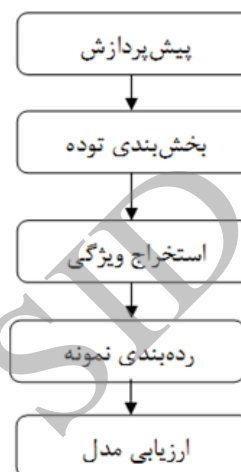
³ Gradient Image

⁴ Skewness

⁵ Kurtosis

مواد و روش‌ها

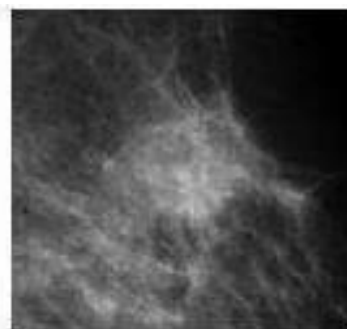
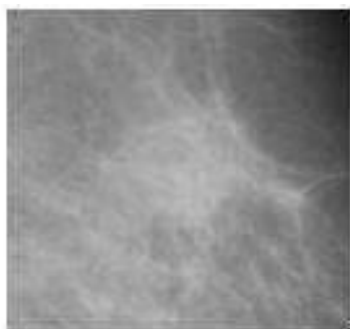
در این قسمت به شرح روش پیشنهاد شده در این مقاله به منظور تشخیص نوع توده در تصاویر ماموگرام می‌پردازیم. بلوک دیاگرام روش پیشنهاد شده در این مقاله از فلوجارت شکل ۱ پیروی می‌کند.



شکل ۱: بلوک دیاگرام کلی روش‌های کمکی

استفاده کردیم. پایگاه داده Mini-MIAS که این پایگاه داده شامل ۳۳۰ تصاویری با ابعاد 1024×1024 است. این ۳۳۰ تصویر فقط شامل نمای MLO از بیمار هستند. در این ۳۳۰ تصویر، ۲۰۹ تصویر نرمال هستند، یعنی هیچ‌گونه *distortion*، *mass*، *asymmetry* و *micro calcification* ای در آنها دیده نمی‌شود. از میان مابقی تصاویر، ۵۶ تصویر شامل حداقل یک توده هستند. یکی از این ۵۶ تصویر (mdb059) در آزمایش‌های ما کنار گذاشته شده‌اند زیرا در این تصویر هیچ‌گونه اطلاعاتی از موقعیت توده وجود ندارد. در این ۵۵ تصویر که ما آزمایش خود را بر روی آنها انجام دادیم، ۵۸ توده وجود دارد که از آن تعداد، ۳۸ مورد از آنها خوش‌خیم و ۲۰ مورد بدخیم هستند. و پایگاه داده DDSM که این پایگاه داده یک پایگاه داده عمومی با تعداد تصاویر زیاد است که معمولاً محققان یک زیرمجموعه از تصاویر آن را انتخاب می‌کنند و آزمایش‌های خود را بر روی آن انجام می‌دهند. در این پایگاه داده برای هر بیمار، هم نمای MLO وجود دارد و هم نمای CC. این پایگاه داده توسط اسکنرهای مختفل و با عمق بیت‌های مختلفی جمع‌آوری شده است. ما در این مقاله از این پایگاه داده ۱۲۰ تصویر دارای توده خوش‌خیم و ۱۲۰ تصویر دارای توده بدخیم را انتخاب کردیم. در ادامه هر یک از ماژول‌های موجود در را با جزئیات بیشتری مورد بررسی قرار می‌دهیم.

یک روش مؤثر برای بهبود تصاویر ماموگرام باید هدفش این باشد که ابتدا نویز موجود در آن ناحیه را کم کرده و سپس لبه‌های مربوط به توده را تقویت کنیم. در این مقاله برای حذف نویز موجود در تصاویر از فیلتر میانه با اندازه 3×3 بهره گرفتیم. سپس برای افزایش کنتراست موجود در تصویر از رابطه زیر استفاده کردیم. در شکل ۲ تأثیر روش پیشنهادی بر روی ناحیه مورد نشان داده شده است.



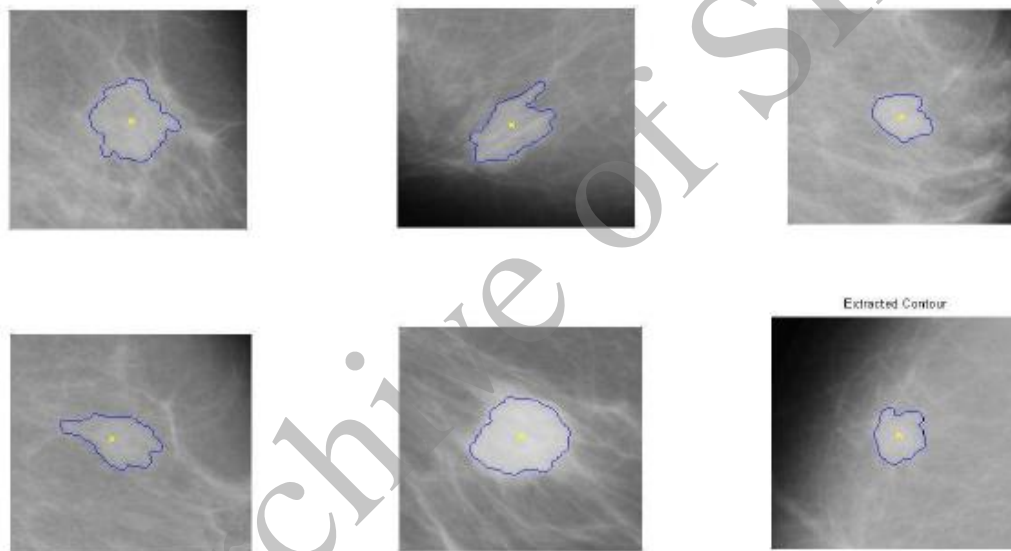
شکل ۲: مثالی از روش پیش‌پردازش پیشنهادی. شکل سمت چپ تصویر اصلی و شکل سمت راست تصویر بهبود یافته است.

سپس ویژگی sum average از روی ماتریس GLCM ساخته شده بر روی هر بلوک استخراج می‌گردد. برای اعمال الگوریتم رشد ناحیه باید به ۲ سوال اساسی پاسخ دهیم: (۱) نقطه شروع (دانه اولیه) کجا در نظر گرفته شود (۲) معیار توقف چه چیزی است. در روش پیشنهادی، نقطه شروع نقطه‌ای در نظر گرفته می‌شود که بالاترین sum average را دارد. برای معیار دوم نیز رشد ناحیه تا زمانی ادامه می‌یابد که شعاع تقریبی ناحیه حاصل به شعاع توده که توسط رادیولوژیست‌ها تخمین زده شده است (اطلاعات موجود در پایگاه داده) نزدیک باشد. نمونه‌ای از اعمال این الگوریتم در شکل ۳ نشان داده شده است.

$$I_{en}(i, j) = \left(\frac{I(i, j)}{I_{max}} \right)^k * I_{max}$$

$$k = 2, 3, 4, \dots$$

یکی از مهم‌ترین گام‌ها در رده‌بندی خوش‌خیمی و بدخیمی توده‌ها بخش‌بندی توده است. زیرا میزان نامنظم بودن مرز توده مبین میزان بدخیمی توده می‌تواند باشد. از این رو در این مقاله از یک روش رشد ناحیه تطبیقی و خودکار جهت بخش‌بندی توده‌ها بهره گرفتیم که اکنون به تشریح این روش می‌پردازیم. ابتدا ناحیه مورد نظر که توده در آن قرار گرفته است به تعدادی بلوک‌های ناهمپوشان با اندازه $32 * 32$ تقسیم می‌شود.



شکل ۳: نمونه‌هایی از اعمال الگوریتم بخش‌بندی پیشنهادی

نیاز به بیان مختصات آنها می‌باشد. از این رو دو معیار فاصله شعاعی و مختصات مرز تعریف می‌گردد. **فاصله شعاعی:** برای محاسبه طول شعاعی (NRL) ابتدا مختصات مرکز توده را بدست می‌آوریم، سپس فاصله اقلیدسی نقاط مرزی از مرکز توده اندازه‌گیری می‌شود و در نهایت نرمالیزه می‌گردد.

$$d(i) = \frac{\sqrt{(x(i) - X_0)^2 + (y(i) - Y_0)^2}}{\max(d(i))} \quad i = 1, 2, \dots, N$$

در رابطه بالا، (X_0, Y_0) مختصات مرکز توده، $(x(i), y(i))$ مختصات نقاط روی مرز در مکان i ام، N

استخراج ویژگی نقش بسیار مهمی در شناسایی الگو ایفا می‌نماید و انتخاب ویژگی‌های مناسب دقت رده‌بندی را بالا می‌برد. این ویژگی‌ها باید توانایی نشان دادن شباهت‌ها بین اشیا در یک دسته را داشته باشند و همزمان قادر به نمایش تفاوت آن اشیا با اشیا موجود در دسته دیگر باشند. عموماً ویژگی‌هایی که برای تشخیص نوع توده‌ها استخراج می‌شود را به دو گروه می‌توان تقسیم کرد:

(۱) ویژگی‌های بافت شناسی

(۲) ویژگی‌های ریخت شناسی

برای استخراج ویژگی‌های ریخت شناسی ما از توصیفگرهای مرز توده بهره گرفتیم که در ادامه به تشریح آنها می‌پردازیم. برای تعیین ویژگی‌های مرتبط با مرز توده

با گرفتن تبدیل فوریه از مختصات مرز، و نرمالیزه کردن آن توصیفگر شکل فوریه (FDM) بدست می‌آید که مولفه‌های فرکانسی نقاط مرزی را نشان داده و چگونگی تغییرات رفتار مرز را توصیف می‌نماید.

$$FD = \text{fft}(\text{boundingcoordinates})$$

$$FDM = \frac{FD}{\text{abs}(FD(1,1))}$$

بر اساس این تعریف با بررسی آماری آن ویژگی‌های زیر تعریف می‌شود:

• Mean_FDM: مقدار میانگین تابع توصیفگر شکل فوریه

• Variance_FDM: مقدار واریانس تابع توصیفگر شکل فوریه

• Energy_FDM: مقدار انرژی تابع توصیفگر شکل فوریه

• Entropy_FDM: مقدار بی‌نظمی تابع توصیفگر شکل فوریه

هرچقدر که یک توده دارای بافت ناهمگن‌تر و ناهموارتری باشد احتمال بدخیمی توده‌ها نیز افزایش می‌یابد. در این راستا ما تعدادی ویژگی مربوط به بافت را از هر توده استخراج می‌کنیم. در این مقاله از ویژگی‌های حاصل از تجزیه مود ذاتی (۸، ۹) برای توصیف بافت توده مورد نظر بهره گرفتیم. تجزیه مود ذاتی روشی است که هرگونه مجموعه داده‌های ترکیبی را به تعدادی محدود و اغلب کوچک از اجزاء، که به عنوان توابع مود ذاتی (IMF) شناخته می‌شوند، تجزیه می‌کند. این تجزیه می‌تواند برای تحلیل داده‌های غیرخطی و ناماننا استفاده شود. هنگامی که EMD برای تحلیل بافت تصویر استفاده می‌شود، به عنوان BEMD شناخته می‌شود. در این مقاله برای استخراج ویژگی‌ها، BEMD بر روی ROI‌های پیش پردازش شده ماموگرام‌ها اجرا و دو IMF برای هر ROI بدست آمد. از آنجا که IMFها ماتریس‌های دو بعدی هستند؛ اندازه بردار ویژگی برای همه تصاویر بسیار بزرگ خواهد بود. این امر، کار رده‌بند را بسیار پیچیده خواهد کرد و زمان محاسبه نیز افزایش خواهد یافت. به منظور کاهش اندازه بردار ویژگی، پنج پارامتر آماری، یعنی، میانگین، انحراف معیار، چولگی، کشیدگی و آنتروپی از این ضرایب استخراج شده است و به عنوان ویژگی‌های بافتی به ویژگی‌های مرز توده الحاق شده‌اند.

تعداد کل نقاط روی مرز و $\max(d(i))$ بزرگترین طول شعاعی استخراج شده است. ویژگی‌های تعریف شده بر اساس فاصله شعاعی عبارتند از:

۱. Average NRL:

میانگین طول شعاعی نرمالیزه شده بیانگر گستردگی توده در بافت می‌باشد و به نوعی به اندازه توده نیز اشاره می‌کند:

$$d_{avg} = \frac{1}{N} \sum_{i=1}^N d(i)$$

۲. Standard deviation of NRL:

انحراف استاندارد طول شعاعی معیار دیگری است که به نوعی بیانگر میزان تغییرات طول شعاعی در توده است.

$$\sigma = \frac{1}{N-1} \sqrt{\sum_{i=1}^N (d(i) - d_{avg})^2}$$

۳. Roughness Index:

شاخص ناهمواری معیاری جهت محاسبه ناهمواری مرزهای توده می‌باشد.

$$R = \frac{1}{N} \sum_{i=1}^N |d(i) - d(i+1)|$$

۴. Zero crossing:

میزان عبور از صفر بر اساس دفعاتی که فاصله شعاعی حاصل از نقاط مرزی مقداری برابر با فاصله شعاعی میانگین را دارد محاسبه می‌گردد؛ لذا علاوه بر شاخه شاخه بودن مرزها و بی‌نظمی آنها به تعداد لوب‌ها در مرز نیز اشاره می‌نماید.

۵. Moments of Boundary:

این معیار به بررسی ساختارهای ریزمرزها و میزان شاخه شاخه بودن آن می‌پردازد و از رابطه زیر بدست می‌آید که برای حساسیت کمتر به نویز می‌توان از ممان‌های مرتبه پایین‌تر نیز بهره جست.

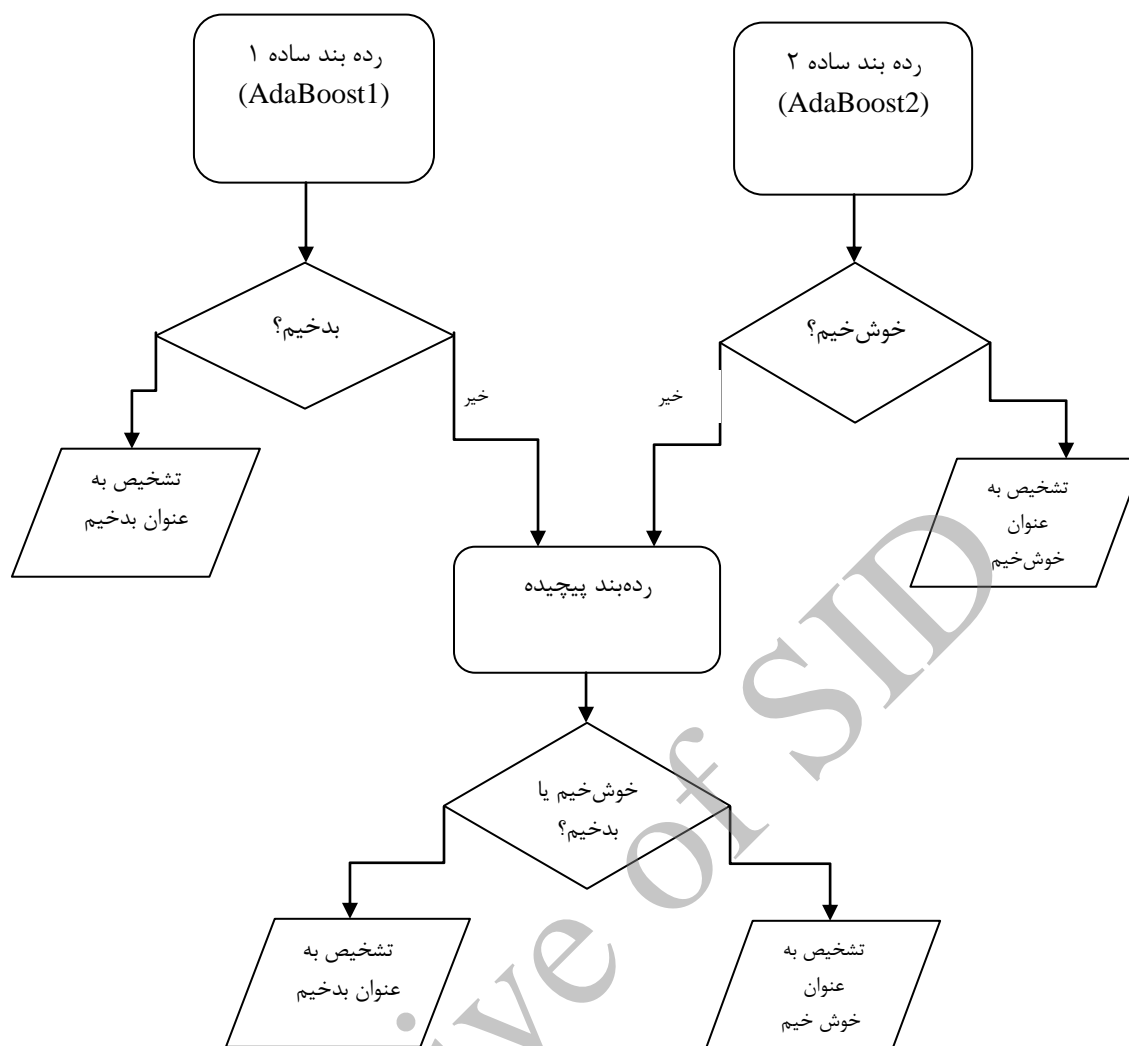
$$\text{moment_boundary}(i) = \frac{1}{N} \sum_{j=1}^N (d(j) - d_{avg})^i \quad i = 1, \dots, 5$$

$$\text{Low_moment}(i) = \frac{\sqrt[i+1]{\text{moment_boundary}(i)}}{d_{avg}} \quad i = 1, 2, 3$$

۶. Fourier Shape descriptor

ساده‌تر قابلیت تعمیم بیشتری دارند. در راستای کم کردن این شکاف یک معماری تطبیقی پیشنهاد شده است که مطابق آن تنها برای نمونه‌های مشکل از رده‌بندهای پیچیده‌تر استفاده خواهد شد نه برای نمونه‌های ساده. برای این منظور ما بلوک دیاگرام زیر را پیشنهاد می‌کنیم. برای تفهیم راحت‌تر، ابتدا از مرحله تست شروع می‌کنیم. یک نمونه جهت تعیین خوش‌خیم یا بدخیم بودن وارد سیستم ما می‌شود. این نمونه به AdaBoost1 داده خواهد شد. AdaBoost1 رده‌بندی است که اگر بگوید نمونه‌ای بدخیم است با اطمینان می‌گوید که بدخیم است و از طرف دیگر AdaBoost2 نمونه‌های خوش‌خیم را با اطمینان تشخیص می‌دهد. در واقع AdaBoost دارای یک آستانه است که در حالت پیش‌فرض این آستانه روی صفر است. اگر مقدار ساپورت حاصل برای یک نمونه، مثبت باشد آنگاه نمونه مربوطه به عنوان بدخیم و اگر منفی باشد به عنوان خوش‌خیم تشخیص داده می‌شود. برای اینکه یک رده‌بند بسازیم که به نظرش در مورد بدخیم بودن نمونه‌ها اعتماد کنیم (AdaBoost1)، باید آستانه مورد نظر را بالا ببریم ولی در مورد AdaBoost2 باید آستانه را پایین بیاوریم. زمانی که یک نمونه مشکلی باشد (نمونه‌ای که رده‌بندهای ساده نتوانند با اطمینان در مورد خوش‌خیم و بدخیم بودن آن نظر دهند)، آنگاه باید از راهکارهای دیگری برای افزایش قدرت تفکیک رده‌بند استفاده کرد. به عنوان مثال می‌توان ویژگی‌های بیشتری به ازای آن نمونه مشکل استخراج کرد (افزایش بُعد) و یا اینکه انعطاف‌پذیری رده‌بند را افزایش دهیم. مثلاً اگر از SVM استفاده می‌کنیم، می‌توانیم با بهره‌گیری از یک کرنل پیچیده‌تر، انعطاف‌پذیری بیشتری به رده‌بند بدهیم. لازم به ذکر است که به جای AdaBoost می‌توان از رده‌بندهای دیگری مانند Naive Bayes نیز استفاده کرده و با تغییر آستانه آن به اهداف خود برسیم. ما نیز در این مقاله از این استراتژی بهره می‌گیریم و از بین همه ویژگی‌ها تنها یک زیر مجموعه از آنها را برای رده‌بندهای ساده در نظر می‌گیریم اما در مورد رده‌بند پیچیده، نمونه‌ها با تمامی ویژگی‌هایشان را استفاده خواهیم کرد.

رده‌بندی: در این بخش به یکی از مهم‌ترین نوآوری‌های این مقاله می‌پردازیم و آن طراحی و پیاده‌سازی یک رده‌بند مرکب جدید است. اصولاً طراحی رده‌بند مرکب (Ensemble) بر مبنای اصلی به نام Diversity استوار است. بدین معنی که به نحوی باید در رده‌بندهای پایه مورد استفاده تنوع ایجاد کرد. این ایجاد تنوع می‌تواند در ۴ سطح صورت گیرد: (۱) سطح داده (۲) سطح ویژگی (۳) سطح رده‌بندها (۴) سطح ترکیب‌کننده. تنوع در سطح داده به این معنی است که رده‌بندهای پایه مورد استفاده از نمونه‌های مختلفی برای آموزش دیدن استفاده کنند مانند Bagging و AdaBoost. تنوع در سطح ویژگی به این معنی است که همه رده‌بندهای پایه روی مجموعه ثابتی از ویژگی‌ها آموزش نینند و از مجموعه ویژگی‌های متفاوتی برای آموزش استفاده کنند (مانند Rotation forest). سطح سوم بر مبنای این اصل استوار است که از متخصص‌های (رده‌بندها) مختلفی برای آموزش استفاده کنیم. در واقع از رده‌بندهایی استفاده شود که نحوه آموزششان با یکدیگر متفاوت است. سطح آخر مربوط به ترکیب‌کننده (combiner) است. به عنوان مثال می‌توان چندین نوع استراتژی مانند رأی اکثریت، رأی‌گیری وزن‌دار و استفاده از یک رده‌بند را با هم ترکیب کرد و بدین‌گونه تنوع ایجاد کرد. ما در این مقاله قصد داریم یک استراتژی جدید برای ترکیب ارایه دهیم. در این قسمت به تشریح معماری پیشنهادی جهت ایجاد یک رده‌بند ترکیبی جدید خواهیم پرداخت. هدف این است که نمونه‌هایی که به راحتی رده‌بندی می‌شوند توسط رده‌بندهای غیرپیچیده (دارای انعطاف‌پذیری کم و با تعداد ویژگی‌های کمتر) و نمونه‌های مشکل توسط رده‌بندهای پیچیده‌تر (با در نظر گرفتن ویژگی‌های بیشتر برای نمونه‌ها) تشخیص داده شوند. همان‌طور که می‌دانیم، هر چقدر از رده‌بند پیچیده‌تری استفاده کنیم خطای آموزش ممکن است پایین بیاید ولی خطای تست ممکن است افزایش یافته و قابلیت تعمیم کم شود. در واقع شکاف بین خطای آموزش و تست افزایش می‌یابد. در واقع طبق اصل تیغ اوکام (razor occum's) بین فرضیه‌های رقیب آن فرضیه‌ای که فرض‌های کمتری دارد بهتر است. بنابراین سادگی یک اصل مهم در تعمیم است. معمولاً سیستم‌های



شکل ۴: معماری پیشنهادی برای ترکیب رده بندها

در مورد مرحله آموزش نیز به این صورت عمل کردیم که بر روی مجموعه آموزش، از استراتژی 10-fold-crossvalidation برای تعیین آسان یا سخت بودن نمونه‌ها بهره گیریم. در واقع از روی مجموعه آموزش، مجموعه‌های آموزش و تست دیگری ساختیم. طبق این استراتژی در هر بار آزمایش، ۹۰٪ نمونه‌ها برای آموزش و مابقی برای تست (اعتبارسنجی) استفاده شدند. برای سادگی، تکرار اول از ۱۰ تکرار ممکن را در نظر بگیریم. در این تکرار ۱۰٪ اول از جموعه برای آزمایش و ۹۰٪ مابقی برای آموزش استفاده می‌گردد. رده بندهای ساده را با این ۹۰٪ آموزش می‌دهیم. حال ۱۰٪ مجموعه تست را در نظر بگیریم. این ۲ رده بند ساده را بر روی نمونه‌های این مجموعه آزمایش می‌کنیم. اگر پیش‌بینی آنها در مورد نمونه‌ای اشتباه باشد و یا با اطمینان کم پیش‌بینی درست داشته باشند، آنگاه این نمونه به مجموعه آموزش رده بند

پیچیده اضافه می‌شود. حال به سراغ تکرار دوم می‌رویم. ۱۰٪ دوم از مجموعه برای تست و مابقی برای آموزش استفاده می‌شوند و نمونه‌های دشوار از مجموعه تست دوم را نیز طبق گفته قبلی جدا می‌کنیم. این روند را تا تکرار دهم ادامه می‌دهیم. پس از تعیین نمونه‌های دشوار در کل مجموعه آموزش، مجموعه آموزش رده بند پیچیده آماده خواهد شد. بنابراین رده بند پیچیده را بر روی مجموعه آماده شده آموزش می‌دهیم.

یافته‌ها

در این مقاله برای گام آزمایش‌ها و نتایج، ۵ مرتبه از استراتژی 10-fold crossvalidation استفاده شد و تمامی نتایج مربوط به گام رده‌بندی توده‌ها بر اساس میانگین این ۵ آزمایش صورت گرفت. در مورد چگونگی تنظیم این پارامترها و پارامترهای دیگر در بخش یادگیری

است که چند درصد از توده‌های بدخیم توسط رده‌بند به اشتباه به عنوان خوش‌خیم رده‌بندی شده‌اند. ما پارامترهای خود را در validation set به گونه‌ای تنظیم کردیم که FNR کمتر از FPR باشد. در دنیای واقعی نیز این‌گونه است که هزینه این ۲ خطا با یکدیگر برابر نیست و FNR مهم‌تر از FPR است. زیرا اگر توده بدخیمی به اشتباه به عنوان خوش‌خیم رده‌بندی شود آنگاه بیمار در معرض مرگ قرار می‌گیرد ولی اگر توده خوش‌خیمی به عنوان بدخیم تشخیص داده شود آنگاه بیمار به پاتولوژی فرستاده می‌شود که در این صورت بیمار فشار روحی و هزینه بیشتری را متقبل خواهد شد. در واقع می‌توان گفت که مساله رده‌بندی توده یک مساله حساس به هزینه است. این حساس به هزینه بودن را می‌توان به شیوه‌های مختلف به روش رده‌بندی تزریق کرد اما از آنجایی این مساله خارج از هدف این مقاله است به آن نمی‌پردازیم. یکی از معیارهای معروفی که اکثر سیستمی‌های تشخیص پزشکی از آن برای ارزیابی کارایی رده‌بند خود استفاده می‌کنند سطح زیر منحنی ROC است. یک منحنی بر حسب sensitivity در برابر specificity که با تغییر آستانه تصمیم‌گیری به دست می‌آید. هر چقدر این عدد به ۱ نزدیک‌تر باشد نشان‌دهنده کارایی بالاتر رده‌بند است. در جدول ۲ مقایسه مختصری بین روش پیشنهاد شده در این مقاله و روش‌های دیگر انجام شده است. گرچه لازم به ذکر است که مقایسه دقیق در این مورد امکان‌پذیر نیست. زیرا روش‌های مختلف معمولاً از داده‌های یکسان و استراتژی یکسان برای آزمایش‌های خود استفاده نمی‌کنند و هیچ‌گونه benchmark استاندارد در این زمینه وجود ندارد. همان‌گونه که در جدول زیر ملاحظه می‌کنید در صورت استفاده از روش انتخاب ویژگی پیشنهاد شده، کارایی سیستم به طرز قابل توجهی بهبود می‌یابد. علاوه بر این سیستم پیشنهاد شده به false negative rate کمتری نسبت به false positive rate منتج می‌شود. علاوه بر این در روش پیشنهادی با استفاده از متد انتخاب ویژگی مربوطه به سطح زیر منحنی ROC بیشتری می‌رسیم. متأسفانه در اکثر سیستم‌ها false negative rate که مهم‌ترین معیار خطا می‌باشد گزارش نشده است اما خوشبختانه معیار خطای false negative rate در سیستم ما پایین است به گونه‌ای که بر روی پایگاه داده mini-MIAS به مقدار تنها ۱٪ رسیده است.

ماشین بدین صورت عمل می‌کنیم که از ۹۰٪ مجموعه آموزش در استراتژی 10-fold به عنوان validation set استفاده می‌کنیم. در واقع ما روی آن ۹۰٪، یک 10-fold cross validation دیگر برای تنظیم پارامترها انجام می‌دهیم. این استراتژی برای تنظیم پارامترهای الگوریتم مربوطه در یادگیری ماشین یک روش استاندارد محسوب می‌شود. برای ارزیابی کارایی رده‌بندی توده‌ها معمولاً از معیارهای زیر استفاده می‌شود که این معیارها از روی ماتریس confusion به دست می‌آید:

جدول ۱: ماتریس سردرهم ریختگی

		Predicted class	
		کلاس = بدخیم	کلاس = خوش‌خیم
Actual class	کلاس = بدخیم	a (TP)	d (FN)
	کلاس = خوش‌خیم	b (FP)	c (TN)

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

$$FNR = \frac{FN}{FN + TP}$$

$$FPR = \frac{FP}{FP + TN}$$

$$sensitivity = \frac{TP}{TP + FN}$$

$$specificity = 1 - FPR = \frac{TN}{TN + FP}$$

در روابط بالا، TP، TN، FP و FN به ترتیب مخفف true positive، true negative، false positive و false negative هستند. Accuracy نشان می‌دهد که رده‌بند مربوطه چند درصد از موارد را به درستی رده‌بندی کرده است. علاوه بر این، FPR و FNR نشان‌دهنده false positive rate و false negative rate هستند. در واقع این دو معیار معرف خطای سیستم هستند. False positive rate که گاهی false alarm rate هم گفته می‌شود نشان‌دهنده این است که چند درصد از توده‌های خوش‌خیم به اشتباه به عنوان بدخیم رده‌بندی شده‌اند. False negative rate هم مبین این

جدول ۱: مقایسه روش پیشنهادی با روش های رقیب

	database	Accuracy (%)	FPR (%)	FNR (%)	AUC
رده بند پیشنهادی	MIAS	93	14.5	1	0.94
رده بند پیشنهادی	DDSM	90	12	7.5	0.93
Mu et.al	MIAS	-	-	-	0.92
Rojas et al.	DDSM	81	-	-	-
Rangayan et al.	MIAS	-	-	-	0.82
Tahmasbi et al.	MIAS	93.6	8.2	3.2	-
RANGAYAN 2000	MIAS	81.5	-	-	0.70

خوشه‌بند *Means-K* را برای حذف نویز ضربه به کار بردند که پیچیدگی و زمان بر بودن مدت اجرا از نقاط ضعف این روش می‌باشد که در روش پیشنهادی در این مقاله برای حذف و بهبود نویز موجود در تصاویر از فیلتر میانه با اندازه 3×3 بهره گرفته شده است. سپس از تکنیک‌های افزایش کنتراست موجود در تصویر استفاده شده است. نتایج ارزیابی نشان‌دهنده عملکرد بهتر و برتری ریافت پیشنهادی در قیاس با آخرین روش‌های ارائه شده برای آشکارسازی و حذف نویز ضربه، تشخیص و بازیابی بافت و لبه‌های تصویر و ارتقای کیفیت تصویر نویزی است.

در (۱۲) نویسندگان با استفاده از آستانه‌گذاری‌های چند مرحله‌ای و تقویت تصویر توانستند توده‌های سرطانی در تصاویر ماموگرافی را جداسازی و تشخیص دهند، طرح ارائه شده اگرچه از یک الگوریتم ساده طبیعت می‌کند و دقت تشخیصی خوبی دارد اما از این جهت که دارای حجم و زمان محاسباتی بسیار شدید می‌باشد که آن هم به دلیل این است که بایستی تعداد بسیار زیادی از مناطق در تصویر ماموگرافی به صورت محلی پردازش شود نقطه ضعف این الگوریتم می‌باشد. که این مشکل در طرح پیشنهادی مقاله، رفع گردیده شده است. در واقع با استفاده از چارچوب پیشنهادی می‌توان درجه آزادی بیشتری در اختیار رادیولوژیست‌ها قرار داد. این درجه آزادی، اعمال آستانه‌گذاری بر روی خروجی الگوریتم‌های رده‌بندی است. از آنجایی که خروجی الگوریتم‌های رده‌بندی از جنس احتمال است، رادیولوژیست‌ها می‌توانند با آستانه‌گذاری بر روی این خروجی احتمالی به نرخ مثبت درست (true positive rate) و مثبت نادرست (false positive) دلخواه خود دست یابند. این آزادی عمل از ویژگی‌های مثبت این سیستم‌های کمکی محسوب می‌گردد. به عبارتی دیگر، رادیولوژیست‌ها با اعمال آستانه دلخواه خود می‌توانند مصالحه مورد نظر بین دو نرخ مذکور را برقرار سازند، زیرا همان‌طور که می‌دانیم نرخ مثبت درست و نرخ مثبت نادرست در تضاد با یکدیگر بوده و افزایش یک نرخ موجب کاهش نرخ دیگر می‌گردد و برعکس. در مقاله (۱۳) نویسندگان با ارائه ۱۴ ویژگی پیشنهادی و از طریق تکنیک ماشین بردار پشتیبان به کلاس‌بندی تصاویر پرداخته‌اند که در روش پیشنهادی این تعداد ویژگی کاهش یافته است. در مقاله (۱۴) نویسنده با بکارگیری فیلتر بانک گابور به تشخیص و حذف اعوجاج

بحث

یکی از موثرترین راه‌های تشخیص سرطان پستان، به خصوص در مراحل اولیه بیماری، انجام ماموگرافی است. از راه‌های مبارزه با این بیماری نیز تشخیص آن در مراحل اولیه پیدایش است. به عقیده اکثر پزشکان در صورت تشخیص به موقع سرطان در مراحل اولیه می‌توان درمان موثرتری را انجام داد و از میزان مرگ و میر احتمالی کاست. مطالعات نشان می‌دهند که تصویربرداری ماموگرافی دوره‌ای در زنان بدون نشانه، نرخ ابتلا به سرطان را به میزان زیادی کاهش می‌دهد.

در این مطالعه، مجموعه‌های تصاویر ماموگرافی بر اساس یک پایگاه داده استاندارد مورد پردازش قرار گرفته است. در این پژوهش تعداد قابل قبولی از بیماران با تصاویر ماموگرافی دیجیتال با حداقل داده‌های از دست رفته به کار گرفته شده است. یکی از مهم‌ترین نوآوری‌های این مقاله ارائه یک چارچوب عمومی و کارا برای رده‌بندی نمونه‌های خوش‌خیم و بدخیم است. در این مقاله از الگوریتم بوستینگ به عنوان رده‌بندهای پایه استفاده شده است. حال آنکه می‌توان از الگوریتم‌های دیگر نیز بهره جست. به عنوان مثال می‌توان از جنگل چرخش، جنگل تصادفی و رده‌بندهای مرکب دیگر بهره جست. این شیوه رده‌بندی در واقع به عنوان یک مفسر دوم در اختیار رادیولوژیست‌ها قرار می‌گیرد و هدف این است که به رادیولوژیست‌ها در شناسایی خوش‌خیمی و بدخیمی کمک کند. هیچگاه هدف از توسعه چنین سیستم‌هایی جایگزین کردن آنها به جای متخصصان نیست، بلکه قرار گرفتن در کنار رادیولوژیست به عنوان یک پیشنهاد دهنده است.

ژیاو و همکاران در (۱۰) از فیلتر هارمونیک و فازی برای حذف نویز استفاده کردند. خان و همکاران در (۱۱)

محدودیت‌هایی که در مدل پیشنهادی می‌توان ذکر کرد عدم ارایه ویژگی‌های استخراجی جهت طبقه‌بندی می‌باشد که ممکن است در برخی از نمونه‌ها ویژگی‌های ارایه شده به درستی پاسخ مناسبی نداشته باشند که این امر نیز ممکن است به دلیل ماهیت تصویر ماموگرام در آن نمونه باشد. پیشنهادی که برای کارهای آینده می‌توان ارایه نمود استفاده از شبکه‌های عصبی عمیق جهت طبقه‌بندی تصاویر ماموگرافی به عنوان خوش‌خیم و بدخیم و یا سالم می‌باشد. شبکه‌های عصبی به دلیل قابلیت یادگیری ویژگی‌های ترکیبی و جدید در لایه‌ها به طور گسترده در همه زمینه‌های دسته‌بندی مورد استفاده قرار گرفته‌اند. مدل‌های شبکه عمیق توسعه یافته مدل‌های شبکه عصبی برای یادگیری تبدیل غیرخطی روی داده‌ها هستند. به عبارت دیگر تفاوت اساسی مدل‌های شبکه عمیق با شبکه عصبی این است که در هر لایه سعی می‌کنند قابلیت بازسازی داده را حفظ کنند که نکته متمایز جهت استفاده در تصاویر ماموگرافی می‌باشد.

قبل از مرحله کلاستر بندی در تصاویر ماموگرام پرداخته است با این حال حساسیت روش ارایه شده ۸۴٪ بوده است که در مدل پیشنهادی این مقدار به ۹۳٪ با انتخاب ویژگی‌های مناسب، رسیده است.

نتیجه‌گیری

در این مقاله یک سیستم کامپیوتری کمکی خودکار به عنوان یک تصمیم‌گیرنده ثانی برای کمک به رادیولوژیست‌ها جهت تشخیص خوش‌خیمی و بدخیمی توده‌ها معرفی شده است. با استفاده از ایده‌های مختلف و ترکیب تکنیک‌های پردازش تصویر و یادگیری ماشین توانستیم به سطح زیر منحنی ۰/۹۴ برای پایگاه داده MIAS و ۰/۹۳ برای پایگاه داده DDSM دست یابیم که نشان دهنده قابلیت تشخیص بالای سیستم پیشنهادی است. به طور کلی می‌توان نوآوری‌های موجود در این مقاله را در معرفی یک سیستم خودکار تطبیقی بر مبنای الگوریتم رشد ناحیه جهت بخش‌بندی توده‌ها، استفاده از توصیف‌گرهای مبتنی بر تجزیه مود ذاتی و معرفی یک معماری جدید برای ترکیب رده‌بندها خلاصه کرد.

References

1. de Oliveira JEE, de Albuquerque Araújo A, Deserno TM. Content-based image retrieval applied to BI-RADS tissue classification in screening mammography. *World journal of radiology* 2011; 3(1):24.
2. Oliver i Malagelada A. Automatic mass segmentation in mammographic images: Universitat de Girona. 2007.
3. Eltonsy NH, Tourassi GD, Elmaghraby AS. A concentric morphology model for the detection of masses in mammography. *IEEE Transactions on Medical Imaging* 2007; 26(6):880-9.
4. Zhu D, Hong T, Lei S, editors. Research on mammogram enhancement. *Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference on 2010: IEEE.*
5. Kim JK, Park JM, Song KS, Park HW. Adaptive mammographic image enhancement using first derivative and localstatistics. *IEEE Transactions on medical imaging* 1997; 16(5):495-502.
6. Cheng HD, Shi XJ. A simple and effective histogram equalization approach to image enhancement. *Digital signal processing.* 2004 Mar 31; 14(2):158-70.
7. Gurcan MN, Yardimci Y, Cetin AE, Ansari R. Detection of microcalcifications in mammograms using higher order statistics. *IEEE Signal Processing Letters.* 1997 Aug; 4(8):213-6.
8. Huang NE, Shen Z, Long SR, Wu MC, Shih HH, ZhengQ, et al., editors. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 1998: The Royal Society.
9. Rilling G, Flandrin P, Goncalves P, editors. On empirical mode decomposition and its algorithms. *IEEE-EURASIP*

- workshop on nonlinear signal and image processing 2003: IEEE.
10. Xiao L, Li C, Wu Z, Wang T. An enhancement method for X-ray image via fuzzynoise removal and homomorphic filtering. *Neurocomputing* 2016; 195:56-64.
 11. Khan A, Waqas M, Ali MR, Altalhi A, Alshomrani S, Shim S-O. Image denoising using noise ratio estimation, K-means clustering and non-local means-based estimator. *Computers & Electrical Engineering* 2016.
 12. Dominguez AR, Nandi AK, editors. Enhanced multi-level thresholding segmentation and rank based region selection for detection of masses in mammograms. 2007 IEEE International Conference on Acoustics, Speech and Signal Processing- ICASSP'07 2007: IEEE.
 13. Domínguez AR, Nandi AK. Toward breast cancer diagnosis based on automated segmentation of masses in mammograms. *Pattern Recognition* 2009; 42(6):1138-48.
 14. Ayres FJ, Rangayyan RM. Reduction of false positives in the detection of architectural distortion in mammograms by using a geometrically constrained phase portrait model. *International Journal of Computer Assisted Radiology and Surgery* 2007; 1(6):361-9.

Archive of SID