

Breast Abnormality Segmentation Using the Wavelet Transform Coefficients Aggregation

Niroomandfam B^{1*}, Nickravan Shalmani A-R¹, Khalilian M¹

¹Department of Computer, College of Mechatronics, Karaj Branch, Islamic Azad University, Karaj, Iran

Receive: 2019/04/27
Accepted: 2019/07/09

*Corresponding Author:
Behrouz Niroomandfam
behrouz.niroomandfam@kiaiu.ac.ir

Ethics Approval:
Not applicable

Abstract

Introduction: Breast cancer is the most common cancer among women in the world. The automatic detection of masses in digital mammograms is a challenging task and a major step in the development of breast cancer CAD systems. In this study, we introduce a new method for automatic detection of suspicious mass candidate (SMC) regions in a mammogram.

Methods: Mammography is widely used for the early detection and diagnosis of breast cancer. Extracting the region of interest (ROI) helps to locate the abnormal areas, which may be analyzed further by a radiologist or a CAD system. In this study, we propose a new method for ROI detection in mammography images. After preprocessing the mammogram, an aggregation of discrete wavelet coefficients based on the lifting scheme and the texture characteristics of the mammogram was created. Then, the coefficients were optimized through noise removal and morphological operations, and a canny edge detector was used to segment the mammogram. Finally, to overcome the problem of over segmentation or under segmentation, reduce the false-negative rate, and enhance the detected regions, we used splitting and merging method. The proposed method was evaluated using images from the DDSM database.

Results: Sensitivity, C_{ROI} , $C_{G.T}$, and FPI were calculated to be 100%, 86.5%, 56%, and 5.4, respectively.

Conclusion: Experimental results indicate that the proposed method is able to detect and identify the abnormal regions of the mammogram that are candidates for breast masses. This technique could potentially improve the performance of CAD systems and diagnosis accuracy in mammograms and can be useful for medical staff and students.

Keywords: Mammogram, Segmentation, Discrete Wavelet Transform Lifting Scheme, Wavelet Coefficients Aggregation

قطعه‌بندی ناهنجاری‌های پستان با استفاده از تجمیع ضرایب موجک

بهروز نیرومندفام^{۱*}، علیرضا نیکروان شلمانی^۱، مجید خلیلیان^۱

^۱گروه کامپیوتر، دانشکده مکترونیک، واحد کرج، دانشگاه آزاد اسلامی، کرج، ایران

چکیده

تاریخ ارسال: ۹۸/۰۲/۰۷

تاریخ پذیرش: ۹۸/۰۴/۱۸

نویسنده مسئول:

بهروز نیرومندفام

behrouz.niroomandfam@kia.ac.ir

مقدمه: سرطان پستان شایع‌ترین سرطان در بین زنان بوده و بیشترین مرگ و میر ناشی از سرطان را در میان زنان ایجاد می‌کند. کشف و مرزبندی خودکار توده‌ها در ماموگرام کار چالش برانگیز و یک گام اساسی سیستم‌های CAD سرطان پستان است. در این مطالعه، روش جدیدی برای کشف خودکار نواحی کاندیدهای مشکوک به توده (SMC^1) در ماموگرام را معرفی نموده‌ایم.

روش بررسی: مطالعه حاضر بر روی کشف نواحی مشکوک به توده در تصاویر ماموگرام و استخراج نواحی مورد نظر (ROI) از تصاویر ماموگرام است. در این مطالعه، ابتدا پیش پردازش بر روی تصویر ماموگرام صورت گرفته و سپس تجمیع ضرایب موجک گسسته براساس ویژگی‌های بافتی تصویر محاسبه و ایجاد می‌شود. بعد از حذف نویز از این تجمیع و با اعمال عملیات مورفولوژی ضرایب تجمیع بهبود یافته و در نهایت با استفاده از آشکارساز لبه Canny، تقسیم‌بندی ماموگرام انجام می‌شود. برای غلبه بر مشکل over-segmentation و under-segmentation و کاهش نرخ منفی و مثبت کاذب و بهبود نواحی کشف شده از روش تقسیم و ادغام برای بهبود نواحی کشف شده استفاده می‌شود. ارزیابی روش پیشنهادی با استفاده از تصاویر پایگاه داده DDSM انجام شده است.

یافته‌ها: میزان حساسیت، C_{ROI} و $C_{G.T}$ به ترتیب برابر با ۱۰۰، ۸۶/۵ و ۵۶ درصد و مقدار FPI برابر با ۵/۴ برای هر تصویر گزارش شده‌اند.

نتیجه‌گیری: ارزیابی روش پیشنهادی نشان‌دهنده قابلیت کشف و تعیین خودکار نواحی ناهنجاری ماموگرام که کاندید توده پستانی هستند را نشان می‌دهد. از روش پیشنهادی می‌تواند به عنوان گام اول همه سیستم‌های CAD سرطان پستان استفاده نمود و یا به رادیولوژیست در تشخیص بهتر کمک کند.

واژه‌های کلیدی: ماموگرام، تقسیم‌بندی، موجک گسسته، تجمیع ضرایب موجک

^۱ suspicious mass candidates

مقدمه

سرطان پستان یکی از شایع‌ترین نوع سرطان در بین زنان و دومین عامل مرگ و میر زنان است. بنابر اعلام سازمان بهداشت جهانی (WHO)، تعداد موارد سرطان در سال ۲۰۲۵ به حدود ۱۹/۳ میلیون مورد خواهد رسید. و در سال ۲۰۱۸ تخمین زده می‌شود که ۶۲۷۰۰۰ زن از سرطان سینه فوت کرده‌اند که تقریباً ۱۵٪ از کل مرگ و میرهای ناشی از سرطان در میان زنان را شامل می‌شود. تشخیص زود هنگام احتمال بقای طولانی مدت بیمار را افزایش می‌دهد. در حال حاضر ماموگرافی یکی از روش‌های مهم تشخیص سرطان پستان در مراحل اولیه و موثرترین روش غربالگری برای تشخیص توده‌ها در ناحیه پستان است، اما رادیولوژیست‌ها با چالش تعداد زیاد تصاویر پستان و مشکل خواندن تصاویر ماموگرام و نرخ بالای مثبت‌های کاذب مواجه هستند. برای غلبه بر این مشکل و تشخیص زودرس سلول‌های تومور در پستان و کاهش میزان مرگ و میر ناشی از آن نیاز به ایجاد یک سیستم تشخیص کامپیوتری (CAD) برای کمک به متخصصان رادیولوژیست است (۱). کشف خودکار نواحی مورد نظر (ROI^۲) می‌تواند به محلی‌سازی بهتر بافت سرطانی به جای استخراج مستقیم ویژگی‌ها از تصویر اصلی ماموگرافی کمک کند.

توده یکی از ناهنجاری‌های شایع پستان است و کشف آن، اولین وظیفه مهم سیستم‌های CAD از بافت‌های پستان است. توده‌ها دارای تنوع زیادی در شکل، اندازه و کنتراست کم نسبت به بافت‌های اطراف خود هستند که منجر به تشخیص مشکل توده می‌شود (۲، ۳). در بیشتر سیستم‌های CAD سنتی نواحی مشکوک ابتدا به صورت دستی مشخص شده و سپس با استفاده از تکنیک‌های یادگیری سنتی و براساس ویژگی‌های استخراج شده، طبقه‌بندی می‌شوند (۴، ۵). در این سیستم‌ها به مجموعه‌ای از ویژگی‌های دستی یا نیمه اتوماتیک برای توصیف ناهنجاری، با قابلیت تمایز مناسب در تشخیص ناهنجاری‌های نیاز است. در واقع، سیستم‌های CAD متعارف مبتنی بر ویژگی، از تشابه بین توده و غیرتوده، برای دسته‌بندی بافت‌های خوش‌خیم و بدخیم پستان استفاده می‌کنند (۶، ۷).

روش‌های مختلفی برای استخراج ویژگی و کشف نواحی مورد نظر در مطالعات قبلی پیشنهاد شده است. از جمله رویکردهای مبتنی بر بافت مانند: (۸) Haralick، (۹) HOG، (۱۰) موجک و GLCM، (۱۱، ۱۲) GLCM. به تازگی، روش‌های تشخیص جدید مبتنی بر یادگیری عمیق برای غلبه بر وظایف چالش برانگیز کشف توده از ماموگرام معرفی شده است. در این مطالعات استخراج خودکار ویژگی‌های سلسله مراتبی سطح بالای عمیق برای تقسیم‌بندی^۴ توده به طور مستقیم از داده‌های خام ورودی بدون درگیر شدن در مشکلات ویژگی‌های استخراج شده دستی معرفی شده است (۱۳، ۱۴). اکثر این سیستم‌ها عملکرد طبقه‌بندی بهتری نسبت به روش‌های سنتی دارند. با این حال، تشخیص توده به صورت خودکار همچنان کار چالش برانگیزی است و در مطالعات متعددی نیز مورد توجه قرار گرفته است (۱۵، ۱۶).

هدف تقسیم‌بندی استخراج نواحی مورد نظر است، طوری که شامل تمام کاندیدهای مشکوک توده (SCM) در تصویر ماموگرام باشد. تقسیم‌بندی نواحی مشکوک در یک تصویر ماموگرافی با حساسیت بسیار بالایی صورت می‌گیرد. در نتیجه تعداد زیادی از مثبت‌های کاذب تولید می‌شود که قابل قبول است، زیرا انتظار می‌رود در مرحله بعد سیستم‌های CAD با تشخیص نوع آن، حذف شوند. کشف توده^۵ مرحله مهمی و مشکل چالش برانگیزی در سیستم‌های CAD تشخیص سرطان پستان است که تاکنون نیز کاملاً حل نشده است (۱۷).

به طور معمول در سیستم‌های CAD پستان، ابتدا ROIها کشف می‌شوند، سپس ناحیه و مرز توده با استفاده از روش‌های متعددی مشخص شده و در نهایت طبقه‌بندی^۶ (تعیین نوع توده) این نواحی صورت می‌گیرد. محاسبات و تحلیل‌های بعدی سیستم‌های CAD به طور قابل توجهی به ROIهای تعیین شده بستگی دارد. نتیجه‌ای که لی و همکاران (۱۸) از بررسی بر روی ماموگرام داشتند نشان داد که با تغییر موقعیت ROI، عملکرد طبقه‌بندی به طور قابل ملاحظه‌ای کاهش می‌یابد. اما مطالعه آنها هیچ نتیجه‌گیری در مورد تاثیر اندازه و تعداد ROI بر روی طبقه‌بندی ROIها نشان نمی‌دهد.

^۴ Segmentation

^۵ Mass detection

^۶ Classification

^۲ Region of Interest

^۳ Detect

حذف مصنوعات رادیولوژی و انتخاب ناحیه پستان^۸:
با استفاده از الگوریتم زیر مصنوعات رادیولوژی از ماموگرام حذف شده و ناحیه پستان انتخاب می شود:

۱- تبدیل باینری تصویر ماموگرام با آستانه $t=0.0784$
۲- حذف نویز با استفاده از فیلتر دو بعدی Median (2D) با اندازه 3×3 (شکل ۲، a)

۳- برچسب گذاری تصویر باینری و انتخاب بزرگترین شیء به عنوان ناحیه پستان و حذف بقیه اشیا تصویر ماموگرام (شکل ۴، b, c)

۴- برش پستان ناحیه ای از ماموگرام که شامل پستان است از پس زمینه تصویر برش داده می شود (شکل ۲، d)

بهبود تصویر ماموگرام: بهبود ماموگرام منجر به بهبود نویز نیز می شود در نتیجه می تواند عملکرد سیستم CAD پستان را تحت تاثیر قرار دهد. لذا با در نظر گرفتن این چالش، الگوریتم زیر تصویر اصلی ماموگرام (M_{org}) را بعنوان ورودی گرفته و $M_{enhanced}$ را بعنوان بهبود یافته تصویر ورودی تولید می کند.

EnhancedMammogram Enhancement (M_{org})

```
{
    Input: mammogram
    Output: enhanced of mammogram
     $M_{unsharped} = \text{unsharp\_masking}(M_{org}, \text{Threshold}=0.7);$ 
     $M_{filtered} = \text{gaussian\_filter}(M_{unsharped}, [3, 3]);$ 
     $M_{erod} = \text{eroded}(M_{filtered}, \text{strElement}('disk', 5));$ 
     $M_{dilate} = \text{dilated}(M_{erod}, \text{strElement}('disk', 5));$ 
     $M_{ero\_dil\_Dis} = \text{img\_subtract}(M_{filtered}, M_{dilate});$ 
     $M_{dilate} = \text{dilated}(M_{filtered}, \text{strElement}('disk', 5));$ 
     $M_{erod} = \text{eroded}(M_{dilate}, \text{strElement}('disk', 5));$ 
     $M_{dil\_ero\_Dis} = \text{imgsubtract}(M_{erod}, M_{filtered});$ 
     $M_{enhanced} = M_{org} + M_{ero\_dil\_Dis} - M_{dil\_ero\_Dis};$ 
    return  $M_{enhanced}$ ;
}
```

حذف نویز با استفاده از موجک^۹: حفظ ویژگی های بافتی ماموگرام مهم ترین اصل در حذف نویز است. حذف نویز با استفاده از روش های خطی باعث تار و حذف ناقص نویز از ماموگرام می شود (۲۱).

Vipul Sharma بر روی اندازه بهینه ROI ها به عنوان نماینده ماموگرام مطالعه نموده است (۱۹). نتایج بدست آمده نشان داد که یک ROI مربعی با ابعاد 200×200 پیکسل ناحیه بهینه ای است که می تواند باعث کاهش هزینه محاسبات استخراج ویژگی های بافت شود. عملکرد طبقه بندی پستان به طور معنی داری به اندازه ROI، محل و روش های استخراج ویژگی ها از این ناحیه بستگی دارد (۱۹). اخیرا نیز بسیاری از محققان از ROI برای تحلیل بافت ناهنجاری های مختلف پستان استفاده می کنند (۲۰). در این مقاله روش جدیدی برای کشف خودکار نواحی مشکوک ماموگرام ارائه شده است. از روش پیشنهادی می توان به عنوان مرحله اول همه سیستم های CAD پستان استفاده کرد و یا به رادیولوژیست ها در جهت تشخیص بهتر توده می تواند کمک کننده باشد. مهم ترین کار این مطالعه (۱) کشف خودکار نواحی مشکوک به توده در ماموگرام (۲) کشف موارد چالش برانگیز مانند نواحی کاندید توده در عضله سینه ای یا داخل ناحیه متراکم پستان (۳) کم بودن زمان اجرا و حافظه مورد نیاز در مقایسه با سایر مدل ها است (۴) حساسیت بالای روش پیشنهادی و (۵) پایین بودن نرخ مثبت های کاذب نسبت به حساسیت.

مواد و روش ها

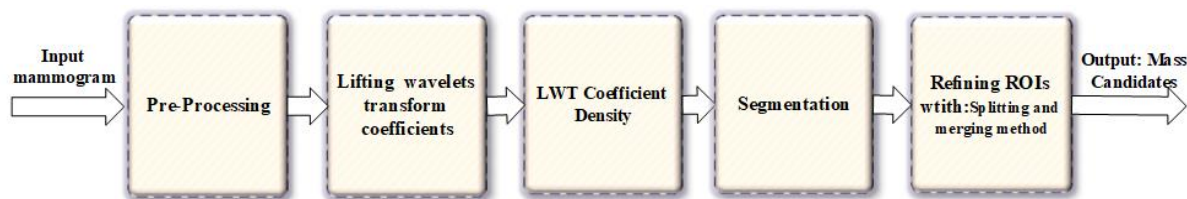
روش پیشنهادی شامل دو وظیفه اصلی است: تقسیم بندی نواحی مشکوک ماموگرام با استفاده از ضرایب موجک و بهبود تقسیم بندی. شکل ۱ خلاصه مراحل روش پیشنهادی را نشان می دهد. تقسیم بندی ماموگرام در روش پیشنهادی به صورت محافظه کارانه بوده و تمامی نواحی که نسبت به بافت اطراف خود متفاوت هستند را در حوزه فرکانس، به عنوان ROI تقسیم بندی می کند.

پیش پردازش: تصویر ماموگرام معمولا دارای نویز و مصنوعات رادیولوژی است. لذا ابتدا در فاز پیش پردازش حذف نویز، برچسب های رادیولوژی، رگ های خونی و بافت های غده ای انجام شده و سپس تصویر ماموگرام بهبود^۷ می یابد. مراحل پیش پردازش در ادامه توضیح داده شده است.

⁸ Breast Profile

⁹ Wavelet Denoising

⁷ Enhancement



شکل ۱: بلوک دیاگرام روش کشف ROI پیشنهادی

(به‌عنوان کاندید) استخراج می‌شوند. تقسیم‌بندی تصویر ماموگرام براساس خواص اولیه‌ی تبدیل موجک استوار است. یکی از این خواص، خاصیت فشرده‌سازی انرژی سیگنال است. این خاصیت تصریح می‌کند که عمده انرژی سیگنال تنها در چند ضرایب از ضرایب موجک متمرکز بوده و بقیه ضرایب عمدتاً دارای مقدار ناچیزی است. هر تصویر دیجیتال X را می‌تواند به عنوان یک سیگنال دو بعدی با ماتریس $[n, m]$ که در آن $x[n, m]$ مقدار پیکسل واقع در ستون n و سطر m است، نشان داد. تبدیل عمل ریاضی است که یک تابع را دریافت نموده و آنرا به یک تابع دیگر انتقال می‌دهد. اهمیت تبدیل در آشکارسازی اطلاعات مخفی شده در تابع ورودی است. عبارت ۱- تبدیل فوریه یک تابع را نشان می‌دهد.

$$F[k] = \int f(t) e^{-j2\pi kt/N} dt \quad (1)$$

$$f(t) = \frac{1}{2\pi} \sum_{k=0}^{N-1} F[k] e^{j2\pi kt/N}$$

برای شناسایی مرزهای نواحی مشکوک در تصویر ماموگرافی لازم است مشخص شود چه فرکانسی (رزولوشن فرکانسی) در کجای (رزولوشن زمانی/مکانی) ماموگرام وجود دارد. تبدیل فوریه تصویر ماموگرام را به حوزه فرکانس انتقال داده و قادر به ارایه اطلاعات اجزای ماموگرام است، اما در خصوص مکان (زمان) این اجزای اطلاعاتی ارایه نمی‌کند. برای شناسایی نواحی مشکوک در ماموگرام اطلاع از مکان (زمان) آن بسیار حایز اهمیت است. لذا تبدیل فوریه برای شناسایی نواحی مشکوک ماموگرام که به‌عنوان سیگنال با محتوای غیرایستا و طیف متغیر با زمان (مکان) است، مناسب نیست (۲۱). با استفاده از یک پنجره با طول متغیر می‌توان به مشکل تعیین مکان فرکانس غلبه نمود. از پنجره‌های با طول متغیر برای فرکانس‌های مختلف استفاده می‌شود. به تابع پنجره‌ای کردن سیگنال، موجک گفته می‌شود. موجک گسسته به‌صورت زیر بیان می‌شود (عبارت ۲):

برای غلبه بر محدودیت‌های روش‌های خطی، از فیلتر غیرخطی مبتنی بر موجک استفاده نموده‌ایم. فرایند حذف نویز با آستانه‌سازی در دامنه موجک، نویز را حذف می‌کند. حذف نویز مطابق با روش Donoho انجام می‌شود (۲۲). در روش آستانه Donoho اگر مقادیر ضرایب موجک زیر آستانه باشد مقدار آنها صفر می‌شود. زیرا این ضرایب بیشتر مربوط به نویز است. ضرایب مربوط به لبه سیگنال، معمولاً بالاتر از آستانه قرار می‌گیرد.

کنتراست محدود تطبیقی هیستوگرام سازگار: یکی از مشکلات بهبود تصویر، انتخاب الگوریتم بهبود مناسب است (۲۳). بهبود در ماموگرافی، به ویژه در پستان‌های متراکم، با افزایش کنتراست به‌دست می‌آید. کنتراست بین بافت بدخیم و بافت پستان در یک ماموگرام وجود دارد اما مقدار آن کمتر از آستانه‌ی قابل درک انسانی است (۲۴). کنتراست محدود تطبیقی هیستوگرام سازگار (CLAHE^{۱۰}) یکی از تکنیک‌های رایجی است که در آن محدوده کنتراست به گونه‌ای تغییر می‌کند که هیستوگرام براساس تابع توزیع تجمعی شکل مورد نیاز باشد. ما از CLAHE برای بهبود کنتراست تصویر و افزایش کنتراست بین توده‌ها و بافت‌های اطراف آن استفاده کرده‌ایم (۲۵) (شکل ۳).

فیلتر هدایتی (Guided Image Filtering): در میان فیلترهای حفظ لبه، هزینه‌ی محاسبات فیلتر هدایتی مستقل از اندازه فیلتر است. تصویر ماموگرافی مرحله قبل، با استفاده از فیلتر هدایتی هموار می‌شود. طبیعی بودن حفظ لبه فیلتر هدایتی امکان داشتن نتایج بهتر را فراهم می‌کند (۲۶). (شکل ۴)

تقسیم‌بندی ماموگرام و کشف ناحیه‌های مورد نظر: تصویر ماموگرام حاصل از خروجی فاز پیش‌پردازش، پردازش شده و نواحی مورد نظر و مشکوک ماموگرام

¹⁰ Contrast Limited Adaptive Histogram Equalization

$$f(t) = \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} C(j, k) \psi_{j,k}(t) \quad (۴)$$

تبدیل موجک گسسته (DWT) Discrete Wavelet Transform دو بعدی شامل نمونه برداری و بانک‌های فیلتر دیجیتال پایین‌گذر $f(n)$ و فیلتر بالا گذر $g(n)$ است. ستون‌ها و ردیف‌های تصویر به طور واضح بر روی تبدیل موجک یک بعدی جهت ایجاد DWT دو بعدی پردازش می‌شود.

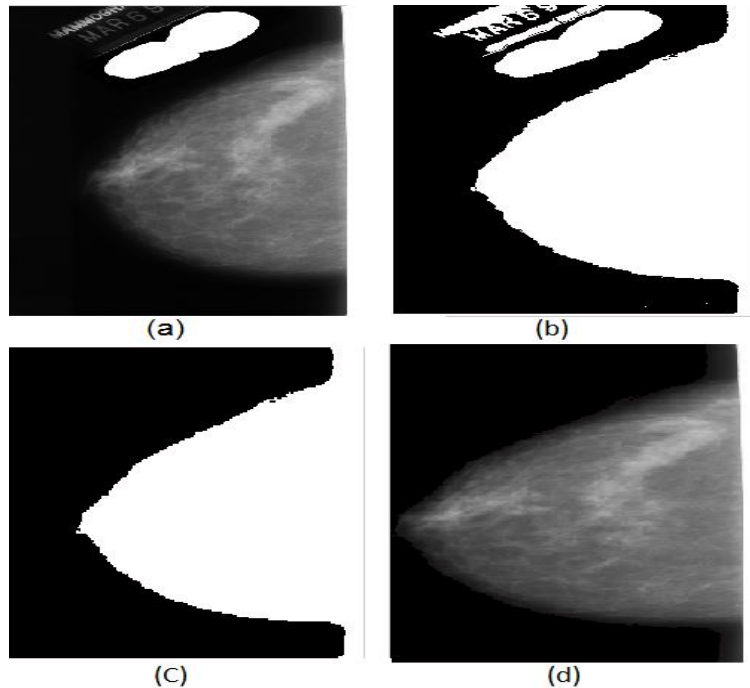
$$\psi_{j,k}(t) = 2^{-\frac{j}{2}} \psi(2^{-j}t - k) \quad (۲)$$

تبدیل گسسته موجک تابع $f(n)$ مطابق رابطه ۳ است.

$$C(j, k) = \sum_{n \in \mathbb{Z}} f(n) \psi_{j,k}(n) \quad (۳)$$

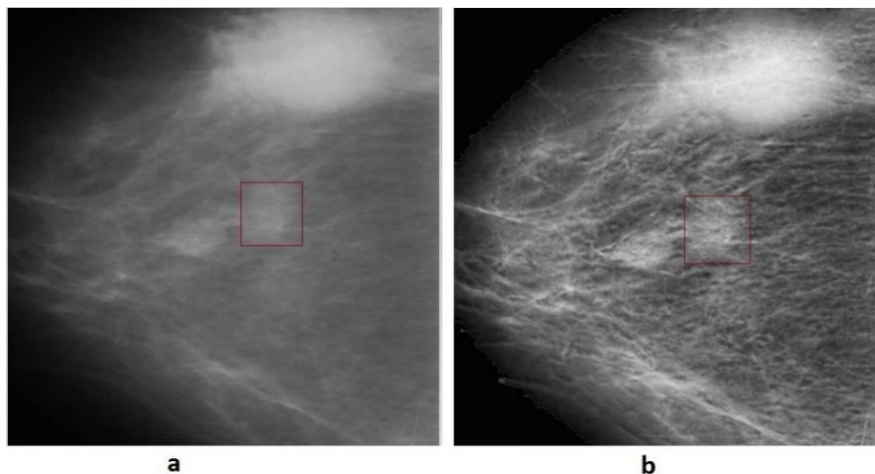
بازسازی سیگنال تجزیه شده با استفاده از رابطه ۴ به‌دست می‌آید.

شکل ۱: بلوک روش کشف ROI پیشنهادی

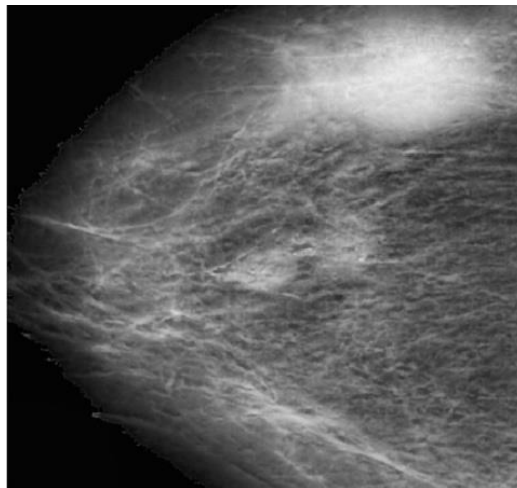


شکل ۲: حذف مصنوعات و پرچسب‌های ماموگرافی و انتخاب پستان

(a) تبدیل باینری شده تصویر اصلی با آستانه $t=0.0784$ که فیلتر میانه 3×3 بر روی آن اعمال شده است. (b) اشیاء موجود در ماموگرام. (c) انتخاب بزرگترین شیء و حذف بقیه اشیاء از ماموگرام. (d) تصویر نهایی پیش‌پردازش که پستان در ماموگرام برش داده شده است.



شکل ۳: ماموگرام (a) تصویر اصلی (b) تصویر بعد از اعمال CLAHE



شکل ۴: خروجی فیلتر هدایتی (Guided Image Filtering)

نمایش می‌دهند. تجمیع ضرایب تبدیل، براساس رابطه ۱۲ و ۱۳ محاسبه می‌شود.

$$A_{n \times n}^+ = \sum \max(0, (c_{n \times n}^m(i, j))) \quad (12)$$

$$A_{n \times n}^- = \sum \min(0, (c_{n \times n}^m(i, j))) \quad (13)$$

که در آن برای محاسبه تجمیع ضرایب مثبت از یک عملکرد غیرخطی همانند عبارت ۱۴ برای حذف ضرایب منفی استفاده شده است. و $c_{n \times n}^m$ ضرایب تبدیل تکرار m ام است. محاسبه تجمیع ضرایب منفی نیز بهمین ترتیب انجام شده و سپس هر دو تجمیع نرمال‌سازی می‌شوند. با توجه به اینکه توده یک ضایع پاتولوژی است که در دامنه خاکستری متوسط تا سایه‌های سفید در ماموگرافی ظاهر می‌شود. لذا از تجمیع ضرایب روشن‌تر برای تعیین نواحی مشکوک مطابق رابطه ۱۶ استفاده می‌کنیم.

$$f(x) = \max(0, \alpha) \iff f(x) = \begin{cases} \alpha & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases} \quad (14)$$

$$f(x) = \min(0, \alpha) \iff f(x) = \begin{cases} 0 & \text{if } x > 0 \\ \alpha & \text{otherwise} \end{cases} \quad (15)$$

$$A_{n,n}(i,j) = \begin{cases} 1 & \text{if } A^+(i,j) > A^-(i,j) \\ 0 & \text{otherwise} \end{cases} \quad (16)$$

تقسیم‌بندی و تعیین نواحی مشکوک ماموگرام براساس تجمیع محاسبه شده $(A_{n,n}(i,j))$ انجام می‌شود. قبل از تقسیم‌بندی، ابتدا پیش‌پردازشی بر روی این تجمیع انجام می‌شود. اول حفره‌های موجود پر شده و سپس با استفاده از عملیات مورفولوژیکی اشیای نویزی حذف می‌شوند.

در نتیجه تجزیه موجک، تصویر I به چهار مولفه زیرباند، تقسیم می‌شود که شامل زیرباند تقریب حاوی اطلاعات صاف تصویر (پس زمینه)، و سه زیرباند (به نام D, H, V) به ترتیب جزئیات عمودی، افقی و قطری) حاوی اطلاعات جزئیات تصویر (لبه) است (۲۲).

DWT یک تصویر نیاز به محاسبات ریاضی پیچیده‌ای داشته و وابسته به تبدیل‌های فوریه است. Sweldens از بهبود تبدیل موجک برای ساختن موجک‌های دو گوشه^{۱۱} استفاده نموده است. برخلاف روش‌های متداول، این طرح از همه ساختارهای حوزه تصویر مشتق شده و علاوه بر سادگی به محاسبات پیچیده ریاضی و به تبدیل‌های فوریه نیز وابسته نیست (۲۳). ویژگی اصلی طرح DWT مبتنی بر لغزاندن (Lifting Scheme) در تفکیک فیلترهای موجک بالا-گذر و پایین‌گذر به یک توالی از ماتریس‌های بالا و پایین مثلثی و تبدیل پیاده‌سازی فیلتر، به ضرایب ماتریس باندی است (۲۳). این طرح اغلب نیاز به محاسبات بسیار کمتر در مقایسه با DWT مبتنی بر کانولوشن دارد.

تقسیم‌بندی ماموگرام و انتخاب ناحیه مشکوک، براساس ویژگی‌های بافتی ماموگرام و با استفاده از تجمیع^{۱۲} ضرایب تبدیل موجک Lifting در حوزه فرکانس انجام می‌شود. ضرایب حاصل از هر مرحله تبدیل، شامل مقادیر مثبت و منفی است که جزئیات بافت ماموگرام است. ضرایب منفی، نقاط تاریک تصویر و ضرایب مثبت نقاط روشن تصویر را

^{۱۱} Biorthogonal

^{۱۲} Aggregation

که در آن $S(s_i)$ اندازه ناحیه s_i ، و δ_1 و δ_2 مقادیر آستانه است و $H(s_i)$ آنتروپی هیستوگرام ناحیه i و H آنتروپی شانون است:

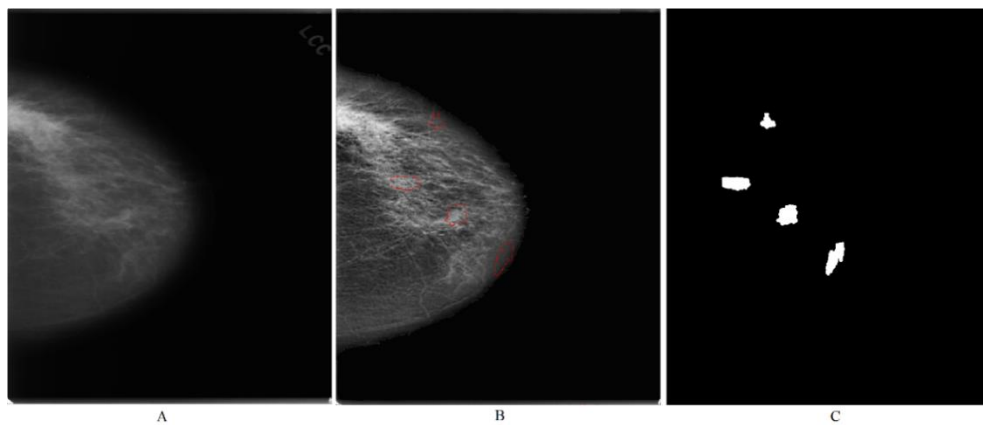
$$H = - \sum_{h(\bar{s}_i) \neq 0} h(\bar{s}_i) \log h(\bar{s}_i) \quad (۱۸)$$

و $h(\bar{s}_i)$ هیستوگرام نرمال شده ناحیه i است. هیستوگرام یک ناحیه با چند قله معمولاً آنتروپی کمی (عدم اطمینان کم) دارد و تصاویر با پراکندگی، دارای آنتروپی هیستوگرام بزرگتری است. رویکرد تقسیم‌بندی ناحیه با استفاده از خوشه‌بندی Meng-Hee و با انتخاب دو نقطه x_p و x_q در داخل قطعه s_i که از لحاظ فضایی فاصله بیشتری از هم دارند به عنوان میانگین خوشه اولیه صورت می‌گیرد. با تقسیم نواحی بزرگ و ایجاد نواحی کوچک‌تر، هنوز مشکل بیش‌تقسیم‌بندی حل نشده است، لذا نیاز به ادغام ناحیه‌ها برای به دست آوردن تقسیم‌بندی مناسب وجود دارد. ادغام دو ناحیه مجاور با حذف مرز مشترک آنها به دست می‌آید. مناطق مجاور که دارای معیار مشابهت هستند، می‌توانند کاندید ادغام باشند. اگر دو ناحیه مجاور به اندازه کافی مشابه باشند، با هم ادغام خواهند شد. فرایند ادغام تا زمانی که هیچ ادغام جدید رخ ندهد، تکرار می‌شود.

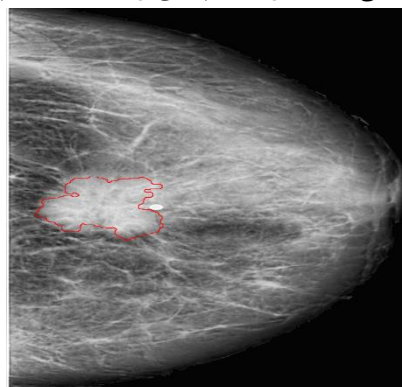
برای کاهش یا حذف نویز موجود در تصویر باینری حاصل، از فیلتر میانگین استفاده شده است. در مرحله بعد با برچسب گذاری، همه اشیای موجود در $(A_{n,n}(i,j))$ شناسایی شده و در نهایت با استفاده از روش Canny لبه‌های این اشیا نیز شناسایی می‌شود (شکل ۵ و ۶).

بهبود نواحی مشکوک پیدا شده: در مرحله قبل همه کاندیدهای که امکان توده را دارند، پیدا شده‌اند. جهت کاهش نرخ منفی و مثبت‌های کاذب و غلبه بر مشکل بیش تقسیم‌بندی (Oversegment) (تعداد نواحی خیلی زیاد، نواحی کوچک) و undersegment (تعداد نواحی خیلی کم، نواحی بزرگ)، نیاز است که نواحی یافته شده در مرحله قبل بازسازی شوند. ما از رویکرد تقسیم و ادغام جهت بهبود تقسیم‌بندی اولیه نواحی استفاده می‌کنیم. ادغام دو ناحیه سراسر است، اما تقسیم یک ناحیه نیاز به ایجاد زیر ناحیه‌های مناسب ناحیه‌ی قابل تقسیم را دارد. ناحیه‌ی تقسیم‌بندی دوباره می‌شود اگر: اندازه آن ناحیه از یک آستانه بزرگتر باشد و آنتروپی هیستوگرام آن نیز از یک آستانه‌ی کوچک‌تر باشد (عبارت ۱۷).

$$H(s_i) < \delta_1 \text{ And } S(s_i) > \delta_2 \quad (۱۷)$$



شکل ۵: (A) ماموگرام اصلی (B) بعد از تقسیم‌بندی اولیه (C) تقسیم‌بندی نهایی و ROIs



شکل ۶: ROI تعیین شده منطبق بر روی ضایعه

یافته‌ها

در این پژوهش، مدلی جهت کشف و مرزبندی ناحیه‌های مشکوک به توده در تصویر ماموگرام معرفی شده است. برای نشان دادن کارایی روش پیشنهادی از پایگاه داده دیجیتال برای ماموگرافی غربالگری DDMS استفاده کرده‌ایم (۲۷). پیاده‌سازی روش پیشنهادی با نرم‌افزار متلب و بر روی یک پردازنده Core i7 2.3Ghz با ۶GB حافظه انجام شده است. ارزیابی سیستم‌های تشخیص به کمک کامپیوتر با معیارهای حساسیت (Sensitivity) یا نرخ مثبت‌های صحیح^{۱۳}، تشخیص (Specificity) یا نرخ منفی‌های صحیح^{۱۴}، دقت (Accuracy) و صحت (Precision) انجام می‌شود (۳). (۲۹، ۲۸، ۱۴). از آنجایی که هدف ما در این مطالعه ارایه سیستم کشف به کمک کامپیوتر (CADE^{۱۵}) توده در تصاویر ماموگرافی است، پارامتر حساسیت مهم‌ترین پارامتر ارزیابی است زیرا میزان شناسایی و پوشش توده روش پیشنهادی در تصویر ماموگرافی را نشان می‌دهد. محاسبه حساسیت با مقایسه ناحیه واقعی ضایعه^{۱۶} (G.T) با ناحیه پیدا شده (ROI) توسط مدل پیشنهادی است. اگر بخشی از ناحیه G.T توسط یکی از ناحیه‌های پیدا شده پوشش داده شود ((True Positive (TP))، به‌عنوان موفقیت روش پیشنهادی در نظر گرفته می‌شود و اگر هیچ بخشی از G.T توسط هیچکدام از نواحی یافته شده پوشش داده نشود ((False Negative (FN)) به‌عنوان شکست روش پیشنهادی در نظر گرفته می‌شود. لذا حساسیت روش پیشنهادی مطابق با رابطه ۱۹ تعریف می‌شود.

$$\text{sensitivity} = \frac{TP}{FN + TP} \quad (19)$$

مقدار حساسیت روش پیشنهادی ۱۰۰٪ محاسبه شده است. حساسیت ۱۰۰ درصدی به‌معنی کشف و شناسایی قطعی محل ناهنجاری در ماموگرام (در صورت وجود) است. اما این پارامتر تعداد ROI‌های پیدا شده به ازای هر تصویر را مشخص نمی‌کند. به‌عنوان مثال اگر تعداد زیادی ROI در تصویر ماموگرام تعیین شود، به احتمال زیاد این

نواحی توده موجود در ماموگرام را پوشش خواهند داد، که در این صورت مسئله کشف و تعیین محل ناهنجاری کماکان به قوت خود باقی است و کارکرد مدل پیشنهادی زیر سوال می‌رود. بنابراین برای نشان دادن میزان نرخ مثبت‌های کاذب هر تصویر^{۱۷} از پارامتر FPI مطابق با رابطه ۲۲ استفاده می‌شود. شکل ۷ نمودار^{۱۸} FROC را نشان می‌دهد که بیانگر میزان حساسیت نسبت به نرخ مثبت‌های کاذب هر تصویر است.

$$FPI = \frac{\text{Number of ROI that is FP}}{\text{Number of Images}} \quad (22)$$

برای ارزیابی میزان پوشش G.T توسط ROI از پارامتر $C_{G.T}$ استفاده کرده‌ایم. این پارامتر نشان‌دهنده درصد درست تقسیم‌بندی (شناسایی) G.T یا درصد مثبت‌های صحیح است. پارامتر $C_{G.T}$ مطابق رابطه ۲۰ تعریف می‌شود.

$$C_{G.T} = \frac{S_{G.T} \cap S_{ROI}}{S_{G.T}} \quad (20)$$

که در آن $S_{G.T}$ مساحت ناحیه G.T و S_{ROI} مساحت ناحیه‌ی ROI است. مقدار $C_{G.T}$ بین صفر و یک است. صفر به این معنی است که ضایعه در هیچکدام از ROI‌ها قرار ندارد و یک به‌معنی همپوشانی کامل ناحیه ضایعه با یکی از ROI‌ها است. متوسط مقدار $C_{G.T}$ ۵۶٪ است. با وجود اینکه $C_{G.T}$ پارامتر خوبی برای ارزیابی روش پیشنهادی است، اما به تنهای برای نمایش دقت روش پیشنهادی کافی نیست، زیرا به بررسی اندازه ROI نمی‌پردازد. لذا ما از پارامتر ارزیابی رابطه ۲۱ برای محاسبه اندازه مفید^{۱۹} ROI استفاده می‌نمایم.

$$C_{ROI} = \frac{S_{G.T} \cap S_{ROI}}{S_{ROI}} \quad (21)$$

پارامتر C_{ROI} نشان‌دهنده درصد مثبت‌های صحیح ROI است به‌عبارت دیگر بیانگر درصدی از ROI که شامل G.T است. متوسط مقدار C_{ROI} عبارت است از ۸۶/۵٪ و حداکثر آن ۱۰۰٪ و حداقل آن نیز ۶۹٪ است (شکل‌های ۸ و ۹).

¹⁷ False Positive per Image (FPI)

¹⁸ Free-response Receiver Operating Characteristic (FROC)

¹⁹ درصد پیکسل‌های ROI که با G.T همپوشانی دارد.

¹³ True Positive Rate

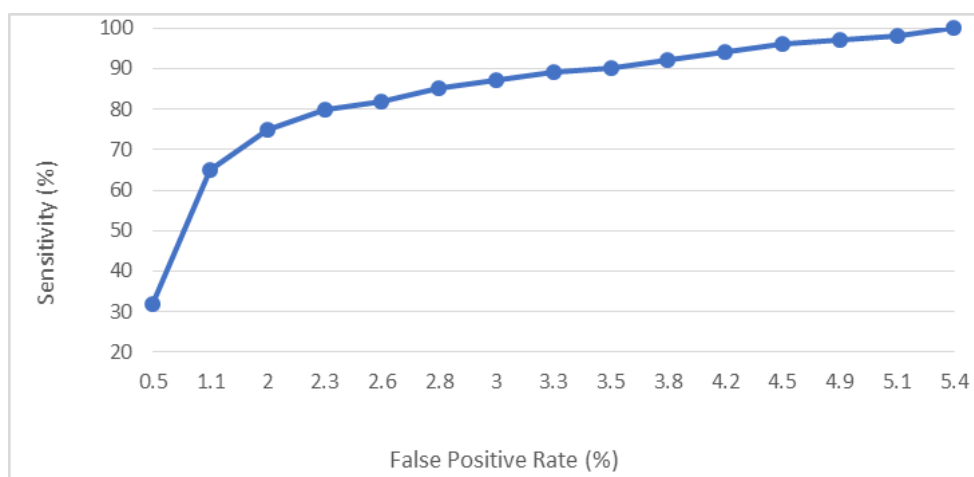
¹⁴ True Negative Rate

¹⁵ Computer aided Detection

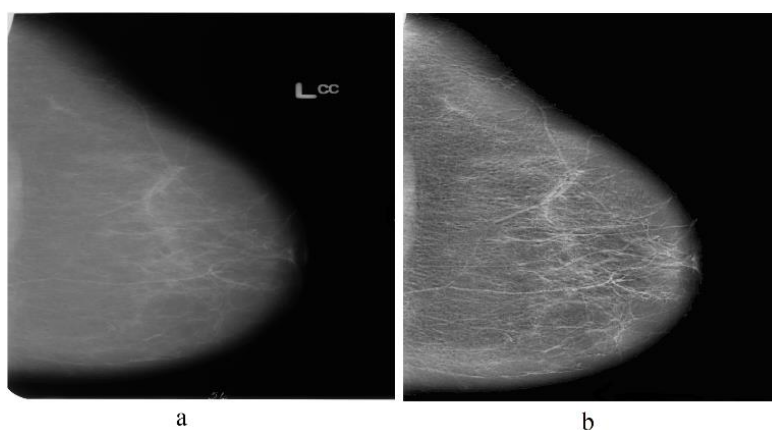
¹⁶ Ground Truth

کاهش محاسبات مورد نیاز تقسیم بندی ROI ها تصویر ماموگرام است.

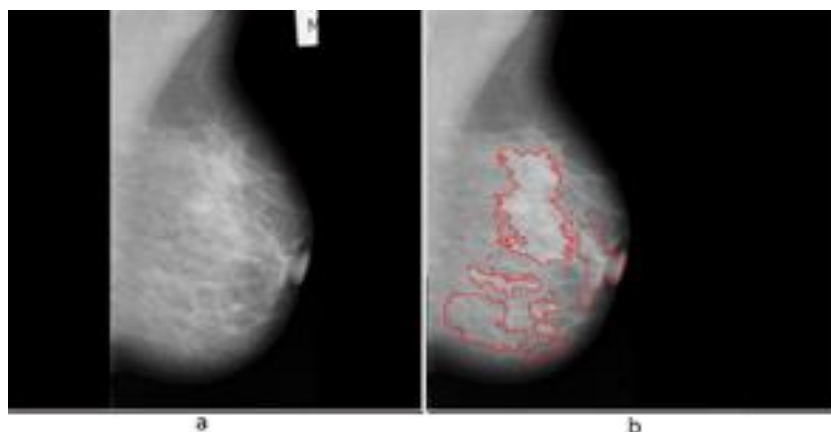
در جدول ۱ مقایسه ی بین روش پیشنهادی با سایر مطالعات صورت گرفته است. این جدول نشان دهنده بهبود عملکرد روش پیشنهادی در مقایسه با سایر روش ها و



شکل ۷: منحنی FROC روش پیشنهادی



شکل ۸: ماموگرام نرمال (a) تصویر اصلی و (b) نتیجه تعیین نواحی ROI با مدل پیشنهادی. (هیچ ROI پیدا نشده است)



شکل ۹: ماموگرام غیر نرمال. (a) تصویر اصلی و (b) نتیجه تعیین نواحی ROI با مدل پیشنهادی با ۷ ناحیه مشکوک شناسائی شده

جدول ۱: مقایسه روش پیشنهادی با روش‌های تقسیم‌بندی ROI در مطالعات قبلی

Methods	Database	Sensitivity (%)	FP-rate (%)	FN-rate (%)
Nunes et al.[31]	DDSM	83.24	0.55	0.17
Wang et al.[32]	DDSM	90.6	3.6	-
Norbert et al[33]	DDSM	95.1	4.3	-
Serhat et al[34]	MIAS	81	0.33	-
Hassanien[35]	MIAS	-	-	-
Oliver A et al. [36]	MIAS	72	10.50	-
Proposed method	DDSM	100	5.4	-

بحث

در سال‌های اخیر، مطالعات زیادی بر روی توسعه سیستم‌های کشف و تشخیص به کمک کامپیوتر (CADE) و (CADx) سرطان پستان که براساس تحلیل تصویر ماموگرام و در راستای کمک به تشخیص دقیق‌تر رادیولوژیست است، انجام گرفته است. کشف و جداسازی دقیق نواحی مشکوک در پستان (CADE) تاثیر مستقیمی بر دقت فرایند تشخیص نوع ضایعه پستان (توسط پزشک و سیستم CAD) دارد. روش‌های متداول سیستم‌های CAD سرطان پستان وابسته به ویژگی‌های دستی است که باعث کاهش دقت آنها در کشف نواحی مشکوک در تصویر ماموگرام می‌شود (۳۰). علاوه براین، روش‌های متداول کشف توده تعداد زیادی از مثبت‌های کاذب را تولید می‌نمایند (۳۱، ۳۲) که باعث از دست رفتن تعداد زیادی از مثبت‌های واقعی (TP) می‌شود. و همین‌طور مرز شناسایی شده ROI در اغلب این روش‌ها به طور دقیق با توده همسو نیست، که تاثیر منفی بر تقسیم‌بندی و طبقه‌بندی مراحل بعدی می‌گذارد (۱۴). در این مطالعه برای غلبه بر این چالش‌ها، روش جدیدی برای کشف و استخراج دقیق نواحی مشکوک به توده (ROI) در تصویر ماموگرام معرفی شده است. روش پیشنهادی محافظه کارانه عمل کرده و همه نواحی تصویر ماموگرافی که نسبت به بافت اطراف خود متفاوت است را به‌عنوان ROI شناسایی و مرز دقیق آنرا نیز رسم می‌نماید. مهم‌ترین چالش این مطالعه یافتن همه نواحی مشکوک در ماموگرام است، طوری‌که ضایعات توده را به‌طور کامل، با حساسیت بالا پوشش و در عین حال نرخ مثبت‌های کاذب قابل قبولی نیز داشته باشد. در حالت ایده‌آل تعداد و اندازه نواحی استخراج شده باید مطابق با تعداد و اندازه ضایعات موجود (G.T) در ماموگرام باشد.

تبدیل موجک گسسته یکی از بهترین ابزار تجزیه و تحلیل تصویر ماموگرافی است که یک چارچوب سلسله مراتبی ساده برای تفسیر اطلاعات تصویر را فراهم می‌کند. این تبدیل قادر است اطلاعات قابل توجهی از تصویر با تفکیک‌پذیری و سطوح^{۲۰} مختلف را به شکل متعامد و دو طرفه ارایه نماید. در سایر مطالعات از ضرایب موجک و ویژگی‌های استخراج شده از این ضرایب برای تقسیم و طبقه‌بندی ماموگرام استفاده شده است (۳۳، ۳۴). در این روش‌ها نیاز است که ویژگی‌های دستی بر روی ضرایب موجک برای کشف و طبقه‌بندی توده تعریف شود. در نتیجه عملکرد این روش‌ها قابل قبول نبوده و علاوه بر محاسبات زیاد (۲۳)، تعداد زیادی مثبت کاذب نیز تولید می‌کنند (۳۵، ۳۶). در این مطالعه، با ایجاد تراکم ضرایب موجک (wavelet transform coefficients) بر پایه فیلترهای بالاگذر و پایین‌گذر و بدون نیاز به استخراج ویژگی، تقسیم‌بندی و استخراج ROI انجام شده است. برخلاف روش‌های متداول، این روش از همه ساختارهای حوزه تصویر ماموگرام مشتق شده و علاوه بر سادگی به محاسبات پیچیده ریاضی و به تبدیل‌های فوری نیز وابسته نیست.

سیستم‌های تشخیص کامپیوتری، بدنال به حداکثر رساندن مقدار حساسیت خود هستند. میزان حساسیت (ناحیه‌های نرمالی که به‌عنوان ناحیه مشکوک (ROI) تقسیم‌بندی شده است، مثبت‌های کاذب یا False Positive (FP) نامیده می‌شوند). برای نمایش میزان حساسیت نسبت به نرخ مثبت‌های کاذب هر تصویر از نمودار FROC استفاده نموده‌ایم. با توجه به شکل ۷ میزان حساسیت به نرخ مثبت‌های کاذب روش پیشنهادی نسبت به سایر مطالعات بهتر است.

²⁰ Resolutions & levels

شناسایی همه نواحی مشکوک تصویر ماموگرام، در مواردی اندازه و تعداد ROI بهینه نیست، که می تواند منجر به افزایش هزینه محاسبات در مراحل بعدی تشخیص سیستم CAD شود. اما این مسئله اساسی نبوده و عملکرد روش پیشنهادی را تحت تاثیر قرار نمی دهد، زیرا در مراحل بعدی تشخیص (توسط رادیولوژیست یا سیستم CAD) نوع این ناحیه تعیین خواهد شد. آنچه در این مطالعه اهمیت دارد پوشش همه ضایعات G.T تصویر ماموگرام است، طوریکه بر چالش سیستم های CAD در تعیین ROI، که تاثیر مستقیمی بر دقت و عملکرد این سیستم ها دارد غلبه نماید. البته باید توجه داشت که بهینه بودن تعداد و اندازه ROI کاملاً وابسته به هدف سیستم CADx است.

یکی از چالش های مرتبط با غربالگری سرطان پستان محدودیت های مربوط به میزان حساسیت غربالگری (۳۷) و نرخ مثبت های کاذب آن است. طوری که حدود ۱۰٪ از زن های که به غربالگری مراجعه می نمایند به بررسی های بیشتر ارجاع داده شده (Recall) و از این بین تنها ۵٪ سرطان پستان تشخیص داده می شوند (۳۷-۳۹). مثبت های کاذب، منجر به فراوانی تست های غیرضروری، هزینه های مراقبت های بهداشتی و مشکلات بهداشت روانی می شود. با توجه به حساسیت و FPI مناسب روش پیشنهادی نسبت به سایر مطالعات (جدول ۱)، یکی از کاربردهای روش پیشنهادی، کمک به کاهش نرخ مثبت های کاذب ماموگرافی است. زیرا با توجه به حساسیت سیستم پیشنهادی، این سیستم به طور متوسط حدود ۵ ناحیه با مرزهای مشخص به عنوان کاندیدهای احتمالی ناهنجاری در تصویر ماموگرام را (برای تحلیل و بررسی بیشتر) در اختیار رادیولوژیست قرار می دهد و نیازی به بررسی کل ماموگرام نیست.

نشان داده شده است که برای گروه های خاص جمعیت (زنان جوان یا زنان با سینه های متراکم) میزان حساسیت کشف توده کاهش می یابد (۴۰)، در نتیجه می تواند به از دست رفتن سرطان های پستان بیشتری منجر شود. کشف توده در پستان های متراکم یا در غلظت سینه ای نیز کاری چالش برانگیز است (۴۰). در این مطالعه ما با انتقال تصویر به حوزه فرکانس، تحلیل تصویر ماموگرام و تعیین

پارامتر حساسیت به بررسی اندازه و میزان مثبت های کاذب (بخش های نرمال ناحیه ای که به عنوان ROI تقسیم بندی شده است) و منفی های کاذب (بخش های از G.T که در ROI قرار ندارد) ناحیه ROI حاوی G.T نمی پردازد. مشخص است که خطای FN نسبت به خطای FP اهمیت بیشتری داشته و بحرانی تر نیز است، زیرا در خطای FP ناحیه نرمال به عنوان ناحیه مشکوک به توده تقسیم بندی شده است که در مراحل بعدی تشخیص (توسط پزشک و یا سیستم CADx) نرمال بودن آن می تواند مشخص شود. اما خطای FN دقت و حساسیت سیستم CADx و یا پزشک را می تواند شدیداً تحت تاثیر قرار داده و به طور معنی داری میزان آنرا کاهش دهد. لذا ما برای ارزیابی روش پیشنهادی از پارامتر C_{ROI} و $C_{G.T}$ نیز استفاده کردیم. پارامتر C_{ROI} درصد درست تقسیم بندی ناحیه ROI حاوی G.T را نشان می دهد. هر چقدر مقدار این پارامتر به ۱ نزدیک تر باشد میزان FP ناحیه ROI کمتر است. و در صورتی که مقدار پارامتر ۱ باشد، میزان FP در تصویر صفر بوده و ROI دقیقاً همان G.T است. پارامتر $C_{G.T}$ نیز نشان دهنده درصد شناسایی درست ناحیه G.T است. مقدار این پارامتر هر چقدر به ۱ نزدیک تر باشد یعنی میزان FN به همان میزان کمتر است.

جدول ۱ مقایسه میان نتایج روش های توسعه یافته تشخیص توده ها در تصاویر ماموگرافی را نشان می دهد. تعیین پارامتر ارزیابی مناسب جهت مقایسه عملکرد مدل پیشنهادی با سایر روش های موجود بعلاوه تنوع پایگاه داده ها و پارامترهای ارزیابی استفاده شده، سخت است. به عنوان مثال از پارامتر specificity جهت مقایسه عملکرد بین روش های مختلف نمی توان استفاده کرد، زیرا این پارامتر تعداد بافت های سالمی که بعنوان ROI استخراج شده اند را مشخص نمی کند. در مطالعات قبلی از پارامترهای حساسیت و FPI به عنوان پارامترهای ارزیابی مبتنی بر ناحیه استفاده شده است (شکل ۷). ما هم برای ارزیابی روش پیشنهادی با سایر روش ها از آن استفاده نموده ایم. همانطوریکه در جدول ۱ مشاهده می شود، روش پیشنهادی دارای حساسیت ۱۰۰ درصدی با نرخ مثبت های کاذب به ازای هر تصویر ۵،۴ است. لذا با وجود

نتیجه‌گیری

نتایج ارزیابی روش پیشنهادی نشان‌دهنده توانایی آن در تشخیص خودکار و تقسیم ناهنجاری‌های تصویر ماموگرام دارد. این روش، محافظه‌کارانه عمل کرده و همه موارد مشکوک و مرز آنها را پیدا می‌کند. حساسیت روش پیشنهادی ۱۰۰٪ بوده و با وجود اینکه تعداد موارد مشکوک پیدا شده و اندازه آنها همواره بهینه نیست، اما با توجه به پوشش کامل ضایعه‌های ماموگرام، در فاز بعدی دسته‌بندی می‌تواند نوع نواحی پیدا شده را تعیین کرد.

تقدیر و تشکر

این مقاله مستخرج از پایان‌نامه دکتری با عنوان "دسته‌بندی توده‌های ماموگرام بر اساس تحلیل ویژگی‌های پستان و استخراج قوانین انجمنی ویژگی‌ها" است. که با کد تصویب ۴۹۴۳۵ در مورخ ۱۳۹۶/۰۵/۱۴ در گروه کامپیوتر دانشکده مکترونیک دانشگاه آزاد اسلامی واحد کرج تصویب شده است.

تعارض منافع

نویسندگان اعلان می‌دارند که هیچ تعارض منافعی در پژوهش حاضر وجود ندارد.

نواحی مشکوک را انجام می‌دهیم. لذا در صورت وجود ناحیه مشکوک در نواحی چالش برانگیز پستان و یا در گروه‌های خاص جمعیتی، این نواحی نیز کشف شده و مرزهای آن رسم می‌شود.

محدودیت‌های این مطالعه شامل: (۱) اندازه و قابلیت دسترسی پایگاه‌داده. به‌عنوان مثال تقسیم‌بندی دقیقی از نواحی ROI در DDSM وجود ندارد و حاشیه نویسی نواحی ROI این پایگاه داده تنها شامل موقعیت کلی ضایعات است (۲) در دسترس نبودن پزشک متخصص برای مشاوره‌ی تخصصی (۳) اندازه و تعداد ROI‌های کشف شده و نرخ مثبت و منفی‌ها کاذب، است. روش پیشنهادی شده در این مطالعه علاوه بر قابل استفاده بودن در سیستم CAD تشخیص سرطان پستان، برای پزشک و کارکنان خدمات پزشکی در تشخیص دقیق و تفسیر ماموگرام بویژه در مواجهه با تعداد زیاد تصویرهای ماموگرام (مانند غربالگری) نیز می‌تواند مفید واقع شود. همچنین روش پیشنهادی بر موارد چالش برانگیز تشخیص ناهنجاری در ماموگرام مانند: وجود توده در عضله سینه یا پستان‌های متراکم نیز غلبه می‌نماید. بهبود عملکرد کشف ROI در تصویر ماموگرافی در نهایت منجر به افزایش دقت و بهبود دسته‌بندی در مراحل بعدی شده و می‌توان از آن در آموزش دانشجویان و کادر پزشکی نیز استفاده کرد.

References

1. Tang J, Rangayyan RM, Xu J, El Naqa I, Yang Y. Computer-aided detection and diagnosis of breast cancer with mammography: recent advances. *IEEE Trans Inform Technol Biomed* 2009; 13: 236-51.
2. M. Al-masni, M. Al-antari, j. Park, G. Gi, T. Kim, P. Rivera, E. Valarezo, et al. Detection and classification of the breast abnormalities in digital mammograms via regional convolutional neural network: 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC'17), Jeju Island, South Korea, 2017: 1230-6.
3. Sena Busra Yengec Tasdemir, Kasim Tasdemir, Zafer Aydin. ROI Detection in Mammogram Images using Wavelet-Based Haralick and HOG Features. 17th IEEE International Conference on Machine Learning and Applications. 2018:105-9.
4. C. Muramatsu, T. Hara, T. Endo, H. Fujita, Breast mass classification on mammograms using radial local ternary patterns. *Comput. Biol. Med*, 2016; 72 (1): 43-53.
5. J. Virmani, N. Dey, V. Kumar, PCA-PNN and PCA-SVM Based CAD Systems for Breast Density Classification, Warsaw, Springer; International Publishing, Poland, 2016: 159-80
6. M.A. Al-antari, M.A. Al-masni, S.U. Park, J.H. Park, M.K. Metwally, Y.M. Kadah, S.M. Han, T.-S. Kim, An automatic computer-aided diagnosis system for breast cancer in digital mammograms via deep belief network, *J. Med. Biol. Eng.* 2017; 38 (3): 443-56.
7. M. Al-masni, M. Al-antari, j. Park, G. Gi, T. Kim, P. Rivera, E. Valarezo, S.-M. Han, T.-s. Kim, Detection and classification of the breast

- abnormalities in digital mammograms via regional convolutional neural network, 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC'17), Jeju Island, South Korea, 2017: 1230-6.
8. R. Ragayyan, T. Nyugen, F. Ayres and A. Nandi, Effect of Pixel Resolution on Texture Features of Breast Masses in Mammograms. *Journal of Digital Imaging*, 2010; 23: 547-53.
9. Oliver, A., et al., A review of automatic mass detection and segmentation in mammographic images, *Medical Image Analysis*, 2010; 4(2): 87-110.
10. S. Beura, B. Majhi and R. Dash, Mammogram Classification using two-dimensional discrete wavelet transform and gray-level co-occurrence matrix for detection of breast cancer, *Neurocomputing*, 2015; 154: 1-14.
11. M. Pratiwi, Alexander, J. Harefa and S. Nanda, Mammogram classification using gray-level co-occurrence matrix and radial basis function neural network. *Procedia Computer Scienc*, 2015; 59: 83-91.
12. K. Kanadam and S. Chereddy. Mammogram classification using sparse-ROI: A novel representation to arbitrary shaped masses. *Expert Systems with Applications*, 2016; 57: 204-13.
13. N. Dhungel, G. Carneiro, A.P. Bradley, deep structured learning segmentation from mammograms. *Image Processing (ICIP) IEEE International Conference*. 2015: 2950-4.
14. N. Dhungel, G. Carneiro, A.P. Bradley, A deep learning approach for the analysis of masses in mammograms with minimal user intervention, *Med. Image Anal.* 2017; 37(1): 114-28.
15. M.A. Al-masni, M. Al-antari, J.-m.P. Park, G. Gi, T.-Y.K. Kim, P. Rivera, et al. Simultaneous detection and classification of breast masses in digital mammograms via a deep learning YOLO-based CAD system. *Comput. Methods Prog. Biomed.* 2018; 157: 85-94.
16. T. Kooi, A. Gubern-Merida, J.-J. Mordang, R. Mann, R. Pijnappel, K. Schuur, Ad. Heeten, N. Karssemeije, A comparison between a deep convolutional neural network and radiologists for classifying regions of interest in mammography, *International Workshop on Digital Mammography*, Sweden. 2016: 51-6.
17. Ayelet Akselrod-Ballin, L. Karlinsky, S. Alpert, S. Hasoul, R. Ben-Ari, E. Barkan, A region based convolutional network for tumor detection and classification in breast mammography, *International Workshop on Large-Scale Annotation of Biomedical Data and Expert Label Synthesis*, Athens, Greece, Springer International Publishing, 2016: 197-205.
18. Li, Hui, et al, Computerized analysis of mammographic parenchymal patterns for assessing breast cancer risk: Effect of ROI size and location. *Medical Physics*, 2004; 31(3): 549-55.
19. Vipul Sharma, Comparative Analysis of Region of Interest of Different Sizes for Breast Density Classification, *International Journal of Medical Research & Health Sciences*, 2017; 6(3): 76-84.
20. N. M. Basheer and M. M. H. Mohammed, Segmentation of breast masses in digital mammograms using adaptive median filtering and texture analysis, *International Journal of Recent Technology and Engineering*, 2013; 2(1): 39-43.
21. Harikumar, R. Performance analysis of neural networks for classification of medical images with wavelets as a feature extractor. *International Journal of Imaging Systems and Technology*, 2015; 25(1): 33-40.
22. Varakorn Kidsumran, Werapon Chiracharit, Contrast enhancement mammograms using denoising in wavelet coefficients, 10th International Joint Conference on Computer Science and Software Engineering (JCSSE), 2013: 82-6.
23. W. Sweldens, The lifting scheme: a custom-design construction of bioorthogonal wavelets, *Appl. Comput. Harmon. Anal.* 1996; 3(2):186-200.
24. Rahmati, P, Adler, A, Hamarneh, G. Mammography segmentation with maximum likelihood active contours. *Med. Image Anal.* 2012; 16: 1167-86.
25. N. Dhungel, G. Carneiro, A.P. Bradley, Deep learning and structured prediction for the segmentation of mass in mammograms, *International Conference on Medical Image Computing and Computer-Assisted intervention*; 2015: 605-12.
26. Smriti Bhatnagar, R.C. Jain, Modified Algorithm for Denoising of Mammographic Images, *International Journal of Electronics Engineering Research*. 2017; 9(2): 217-32.
27. <http://marathon.csee.usf.edu/Mammography/Databse.html>
28. G. N. Sarage, S. Jambhorkar, Enhancement of Mammography Images for Breast Cancer Detection using Histogram Processing Techniques, *International Journal of Computer Science & Technology*. 2011; 2(4):177-9
29. Y. Li, H. Chen, L. Cao, and J. Ma, A survey of computer-aided detection of breast cancer with

- mammography, *J Health Medical Information*, 2016; 7(4):238-44.
30. Li, Y., Chen, H., Yang, Y., Cheng, L., and Cao, L., A bilateral analysis scheme for false positive reduction in mammogram mass detection. *Comput. Biol. Med.* 2015; 57:84-95.
 31. Kozegar, E., Soryani, M., Minaei, B., Domingues, I., et al., Assessment of a novel mass detection algorithm in mammograms. *J. Cancer Res.* 2013; 9 (4): 592-600.
 32. Sampat, M.P, Bovik, A.C, Whitman, G.J, Markey, M.K., 2008. A model-based framework for the detection of spiculated masses on mammography. *Med. Phys.* 35 (5), 2110-23.
 33. Malek, A. A, Rahman, W.E.Z.W.A, Ibrahim, A, Mahmud, R, Yasiran, S.S, and Jumaat, A.K, Region and boundary segmentation of microcalcifications using seed-based region growing and mathematical morphology. *Procedia-Social Behavioral Scie*, 2010; 8: 634-9.
 34. Deepthi Sehwat, Abhishek Sehwat, Dhirendra Jaiswal, Dr. Anima Sen, Detection and Classification of Tumor in Mammograms using Discrete Wavelet Transform and Support Vector Machine, *International Research Journal of Engineering and Technology (IRJET)*, 2017; 4(5): 1328-34.
 35. N. Dhungel, G. Carneiro, A.P. Bradley, Deep structured learning for mass segmentation from mammograms, *IEEE International Conference*, 2015: 2950-4.
 36. N. Dhungel, G. Carneiro, A.P. Bradley, A deep learning approach for the analysis of masses in mammograms with minimal user intervention, *Med. Image Anal.* 2017; 37 (1): 114-28.
 37. Kim, S.Y, et al, Breast Cancer Detected at Screening US: Survival Rates and Clinical-Pathologic and Imaging Factors Associated with Recurrence. *Radiology*, 2017; 284(2): 354-64.
 38. Tosteson, A.N, et al, Consequences of false-positive screening mammograms. *JAMA internal medicine.* 2014;174(6): 954-61.
 39. Poorolajal, J, et al, Breast cancer screening (BCS) chart: a basic and preliminary model for making screening mammography more productive and efficient. *J Public Health (Oxf)*, 2017; 40(2):1-8.
 40. HongyuWang, Jun Feng, Qirong Bu, Feihong Liu, Min Zhang, Yu Ren, et al. Breast Mass Detection in Digital Mammogram Based on Gestalt Psychology, *Journal of Healthcare Engineering*, 2018; 12(2): 1-13.