

Full Automatic Classification of Suspicious Areas in Breast Thermo Images for Early Cancer Detection

A.A. Lashkari^{1*}, F. Pak², and M. Firouzmand³

¹Ph.D, Department of Bio-Medical Engineering, Institute of Electrical Engineering & Information Technology, Iranian Research Organization for Science and Technology (IROST), Tehran, Iran

²Colleague, Department of Bio-Medical Engineering, Institute of Electrical Engineering & Information Technology, Iranian Research Organization for Science and Technology (IROST), Tehran, Iran

³Assistant Professor, Department of Bio-Medical Engineering, Institute of Electrical Engineering & Information Technology, Iranian Research Organization for Science and Technology (IROST), Tehran, Iran

Receipt in the online submission system 15 September 2015, received in revised form 16 November 2015, accepted 29 December 2015

Abstract

Breast cancer is the most common type of cancer among women. The important key to treat the breast cancer is early detection of it because according to many pathological studies more 80% of all abnormalities are still benign at primary stages; so in recent years, many studies and extensive research done to early detection of breast cancer with higher precision and accuracy.

Infra-red breast thermography is an imaging technique based on recording temperature distribution patterns of breast tissue. Compared with breast mammography technique, thermography is more suitable technique because it is noninvasive, non-contact, passive and free ionizing radiation. In this paper, a full automatic high accuracy technique for classification of suspicious areas in thermogram images with the aim of assisting physicians in early detection of breast cancer has been presented. Proposed algorithm consists of four main steps: pre-processing & segmentation, feature extraction, feature selection and classification. At the first step, using full automatic operation, region of interest (ROI) determined and the quality of image improved. Using thresholding and edge detection techniques, both right and left breasts separated from each other. Then relative suspected areas become segmented and image matrix normalized due to the uniqueness of each person's body temperature. At feature extraction stage, 23 features, including statistical, morphological, frequency domain, histogram and Gray Level Co-occurrence Matrix (GLCM) based features are extracted from segmented right and left breast obtained from step 1. To achieve the best features, feature selection methods such as mRMR, SFS, SBS, SFFS, SFBS and GA have been used at step 3. Finally to classify and TH labeling procedures, different classifiers such as AdaBoost, SVM, kNN, NB and PNN are assessed to find the best suitable one. The results obtained on native database showed the best and significant performance of the proposed algorithm in comprise to the similar studies. According to experimental results, mRMR combined with AdaBoost with the maximum accuracy of 92%, and SFFS combined with AdaBoost with a maximum accuracy of 88%, are the best combination of feature selection and classifier for evaluation of the right and left breast images respectively.

Key words: *Breast Cancer, Breast Thermography, Thermogram, Feature Selection, Classification, TH*

* Corresponding author

Address: Department of Bio-Medical Engineering, Institute of Electrical Engineering & Information Technology, Iranian Research Organization for Science and Technology (IROST), Postal Code: 3353136846, Tehran, Iran.

Tel: +98-21-56276031

Fax: +98-21-56276032

E-mail: Lashgari.a@irost.ir

آشکارسازی و دسته‌بندی تمام خودکار نواحی مشکوک در تصاویر ترموگرام پستان برای تشخیص زودهنگام سرطان

امیراحسان لشکری^{۱*}، فاطمه پاک^۲، محمد فیروزمند^۳

^۱دکتری مهندسی پزشکی، پژوهشکده مهندسی برق و فناوری اطلاعات، سازمان پژوهش‌های علمی و صنعتی ایران، تهران
^۲همکار طرح، گروه بیوالکتریک، پژوهشکده مهندسی برق و فناوری اطلاعات، سازمان پژوهش‌های علمی و صنعتی ایران، تهران
^۳استادیار، گروه بیوالکتریک، پژوهشکده مهندسی برق و فناوری اطلاعات، سازمان پژوهش‌های علمی و صنعتی ایران، تهران

تاریخ ثبت در سامانه: ۱۳۹۴/۶/۲۴، بازنگری: ۱۳۹۴/۸/۲۵، پذیرش قطعی: ۱۳۹۴/۱۰/۸

چکیده

سرطان پستان رایج‌ترین نوع سرطان در بین زنان است. مطالعات پاتولوژیک نشان دادند که بیش از ۸۰٪ ناهنجاری‌های پستان در مراحل اولیه خوش‌خیم هستند، بنابراین مهم‌ترین مسأله در درمان آن تشخیص زودهنگام است. ترموگرافی مادون قرمز پستان یک روش تصویرگیری مبتنی بر ثبت الگوهای توزیع دمایی بافت پستان است و در مقایسه با ماموگرافی پستان به دلیل غیرتهاجمی، غیرتماسی، غیرفعال بودن و عدم استفاده از تابش یونیزان روشی بسیار مناسب در تشخیص زودهنگام سلول‌های سرطانی است. در این مقاله روشی به منظور آشکارسازی خودکار نواحی مشکوک در تصاویر ترموگرام پستان با هدف کمک رساندن به پزشکان در تشخیص زودهنگام این سرطان ارائه شده است، به نحوی که دقت و صحت را افزایش داده و درصد پذیرش اشتباه را کاهش می‌دهد. این الگوریتم شامل ۴ بخش اصلی پردازش تصویر، استخراج ویژگی، انتخاب ویژگی و طبقه‌بندی است. در مرحله پردازش، ابتدا براساس عملیاتی تمام خودکار، ناحیه‌ی مطلوب تعیین شده، کیفیت تصاویر ارتقاء یافته و سپس پستان راست و چپ از یکدیگر جدا می‌شوند. سپس درایه‌های ماتریس تصویر، نرمال‌سازی شده و نواحی نسبی مشکوک تعیین می‌شوند. پس از آن و در مرحله استخراج ویژگی، ویژگی‌های آماری، ویژگی‌هایی مبتنی بر هیستوگرام، ویژگی‌هایی مبتنی بر ماتریس هم‌وقوعی (GLCM)، ویژگی‌هایی بر اساس مورفولوژی نواحی مشکوک و ویژگی‌هایی در حوزه‌ی فرکانس از هر یک از نواحی بخش‌بندی شده‌ی پستان راست و چپ استخراج می‌شوند. در ادامه برای دستیابی به بهترین ویژگی‌ها، روش‌های انتخاب ویژگی نظیر کم‌ترین افزونگی و بیش‌ترین ارتباط (mRMR)، انتخاب متوالی روبه جلو (SFS)، انتخاب متوالی روبه عقب (SBS)، انتخاب متوالی سیال روبه جلو (SFBS)، انتخاب متوالی سیال روبه عقب (SFBS) و الگوریتم ژنتیک (GA) به کار گرفته می‌شود. در پایان برای طبقه‌بندی و تعیین معیار استاندارد برای تحلیل دمای عروقی پستان‌ها (TH)، روش‌های مختلف طبقه‌بندی مانند AdaBoost، ماشین‌های بردار پشتیبان (SVM)، نزدیک‌ترین همسایه (KNN)، بیزین ساده (NB) و شبکه‌ی عصبی احتمالی (PNN) مورد ارزیابی قرار گرفتند، تا از مناسب‌ترین آن‌ها به منظور طبقه‌بندی ویژگی‌ها استفاده شود. نتایج به‌دست آمده روی پایگاه داده‌ی بومی، بیانگر کارایی قابل‌توجه روش پیشنهادی است. با توجه به نتایج، ترکیب mRMR با AdaBoost با بیشینه صحت ۹۲٪ و ترکیب SFBS با AdaBoost با بیشینه صحت ۸۸٪، به ترتیب بهترین ترکیبات به‌دست آمده روی تصاویر پستان راست و چپ ارزیابی شدند.

کلیدواژه‌ها: سرطان پستان، ترموگرافی پستان، ترموگرام، انتخاب ویژگی، طبقه‌بندی، TH

*نویسنده مسئول

نشانی: گروه بیوالکتریک، پژوهشکده مهندسی برق و فناوری اطلاعات، سازمان پژوهش‌های علمی و صنعتی ایران، تهران، کد پستی: ۳۳۵۳۱۳۶۸۶

تلفن: ۰۲۱۶۲۲۷۳۱۰ (۹۸+)

دورنگار: ۰۲۱۶۲۲۷۳۱۰ (۹۸+)

پست الکترونیکی: lashgari.a@irost.ir

۱- مقدمه

سرطان پستان رایج‌ترین نوع سرطان و یکی از مهم‌ترین دلایل مرگ و میر در بین زنان است. مهم‌ترین مسئله در درمان این نوع سرطان، تشخیص زود هنگام آن است. در سال‌های اخیر، مطالعات و تحقیقات گسترده‌ای برای آشکارسازی هرچه دقیق‌تر و صحیح‌تر آن صورت گرفته است [۱].

چندین سال است که از تصاویر سونوگرافی و ماموگرافی به منظور تشخیص زودهنگام سرطان پستان استفاده می‌شود، که بیش‌ترین روش مورد استفاده ماموگرافی است. از مضرات ماموگرافی می‌توان به تهاجمی بودن آن اشاره کرد، که متخصصان معتقدند تابش الکترومغناطیس آن می‌تواند یک عامل محرک برای رشد سرطان باشد. بیماران از اعمال فشار زیاد به پستان در روش استاندارد ماموگرافی شکایت دارند. همچنین خطر ناشی از تابش‌های یونیزان برای آشکارسازی تومور بسیار نگران‌کننده است. سونوگرافی نیز اگرچه فرآیندی بسیار کارآمد برای آشکارسازی توده‌ها در پستان است، اما در تشخیص ductal درجه ۱ شکست زیادی دارد [۲]. به منظور تکمیل این روش‌ها، می‌توان از ترموگرافی مادون قرمز برای تشخیص ناهنجاری‌های پستان در تصاویر دمایی استفاده کرد. ترموگرافی مادون قرمز با استفاده از یک دوربین بسیار حساس به مادون قرمز، تصویری از توزیع دما در بدن انسان به دست می‌آورد. این روش تصویرگیری به دلیل عدم استفاده از تابش‌های یونیزان و یا تزریق مواد کنتراست‌زا و همچنین غیرتماسی (عدم فشار پستان)، غیرتهاجمی و غیرفعال (پسیو) بودن، از نظر کاربردهای پزشکی و زیست‌شناسی در مقایسه و یا ترکیب با سایر روش‌های تصویرگیری بسیار مهم و ارزشمند است. از طرفی به دلیل ثبت الگوهای توزیع دمایی در بافت، تفسیر فیزیولوژیکی (تغییر دما بر اساس تغییرات الگوهای فعل و انفعال) از بافت ارائه می‌شود، بنابراین حتی زمانی که هیچ ناهنجاری آناتومیکی مبنی بر وجود سرطان وجود نداشته باشد و بافت از لحاظ ظاهری کاملاً سالم به نظر برسد، موارد مشکوک به سرطان را آشکار می‌سازد. درحالی که سایر روش‌های تشخیصی موجود بیانگر شکل‌های ساختاری بوده و تا زمانی که آناتومی بافت تغییر نکند، قادر به

آشکارسازی عارضه نیستند. نتایج ترموگرافی می‌تواند ۸ تا ۱۰ سال قبل از ماموگرافی مورد ارزیابی قرار گیرد و تشکیل هرگونه جرم یا بافتی را در بدن انسان نشان دهد [۳].

تغییرات در دمای پستان با تغییرات در بافت، متناسب و مرتبط است. قبل از آن که یک سلول سرطانی شود، بافت‌های اطراف آن برای حمایت از رشد سریع این سلول به یک منبع ثابت مغذی، شروع به ایجاد رگ‌های خونی جدید نموده تا مواد مغذی مورد نیاز برای رشد آن را فراهم نمایند. برای حفظ ثبات این منبع، سلول‌های به اصطلاح "بد"، مواد شیمیایی را به نواحی اطراف رها کرده که منجر به بازماندن رگ‌های خونی موجود شده و به طور مداوم کمک می‌کنند تا رگ‌های خونی، یکی پس از دیگری ساخته شوند. این عمل به عنوان رشد رگ‌های خونی جدید^۲ شناخته می‌شود، بنابراین یک بستر رگی قوی در پستان برای زمینه‌سازی شرایط رشد تومورهای مورد نیاز فراهم می‌آید. این رگ‌های خونی به صورت سخت، کاملاً فعال و سریع، مواد مغذی را به سلول‌های سرطانی گرسنه و تازه تشکیل شده می‌رسانند، که افزایش انتشار و تزریق مایعات (لنف و خون) در آن ناحیه به تولید گرمای اضافی در نواحی سطحی پوست اطراف آن و در نتیجه افزایش دمای محلی منجر می‌شود، که تشکیل‌دهنده‌ی "نقاط داغ" هستند. این نقاط داغ قبل از آن که تومورهای سرطانی خوش‌خیم رشد کنند، به وجود می‌آیند [۴]. معمولاً این تغییرات حرارتی و دمایی پستان در تومورهای بدخیم بیشتر هم می‌شود. بنابراین با آشکارسازی این الگوهای توزیع دمایی در بافت، می‌توان نواحی مشکوک را آشکارسازی و دسته‌بندی نمود.

ویلیامز و هاندلی در سال ۱۹۶۱، [۵]، امکان آشکارسازی سرطان پستان از افزایش دما در یک تصویر پستان را نشان دادند (در آن زمان حساسیت دمایی حدود ۲ درجه سانتیگراد بود، که این تصویر در چند دقیقه بدست می‌آمد). مطالعات بعدی نشان دادند که تصویر گرمایی می‌تواند برای آشکارسازی سرطان پستان مفید باشد. پاریسکی و همکارانش [۶] بعد از انجام یک آزمایش بالینی بر روی ۷۶۹ بیمار به مدت چهار سال، گزارش کردند که تصویرگیری دمایی مادون

^۲ Neogenesis^۱ Ductal Carcinoma insitu

بیش‌ترین ارتباط ($mrMR^4$)، انتخاب متوالی روبه جلو (SFS^5)، انتخاب متوالی روبه عقب (SBS^6)، انتخاب متوالی سیال روبه جلو ($SFFS^7$)، انتخاب متوالی سیال روبه عقب ($SFBS^8$) و الگوریتم ژنتیک (GA^9)، تعدادی ویژگی از ۲۳ ویژگی انتخاب می‌شوند. در نهایت ماتریس ویژگی حاصل برای تعیین THها به الگوریتم‌های طبقه‌بندی AdaBoost، ماشین‌های بردار پشتیبان (SVM^{10})، نزدیک‌ترین همسایه (KNN^{11})، بیزین ساده (NB^{12}) و شبکه‌ی عصبی احتمالی (PNN^{13}) داده می‌شود تا مناسب‌ترین روش طبقه‌بندی برای تعیین بهینه‌ی THها مورد استفاده قرار گیرد. «TH» معیاری استاندارد برای تحلیل دمای عروقی در پستان‌ها است، که تصاویر حرارتی را بر اساس نوع توده در ۵ دسته‌ی کلی دسته‌بندی می‌کند.

۲- مواد و روش‌ها

۲-۱- اخذ داده

برای ارزیابی کارایی سیستم پیشنهادی از یک پایگاه داده بومی استفاده شد. برای جمع‌آوری تصاویر این پایگاه داده، از افراد مراجعه‌کننده به بیمارستان امام خمینی که به تشخیص پزشک معالج، مشکوک به سرطان بودند، تصویر ترموگرافی گرفته شد. این تصاویر با همکاری شرکت فناوران مادون قرمز [۱۱] و با رعایت مسائل اخلاقی، توسط دوربین مدل (Thermoteknix VisIR 640, Resolution: 480 × 640) تهیه و جمع‌آوری شده و با نظر پزشک متخصص برچسب‌گذاری شدند. این پایگاه داده شامل تصاویر با زوایای مختلف ۰، $\pm 45^\circ$ و $\pm 90^\circ$ درجه قبل و بعد از تست یخ است که در مجموع ۱۰ تصویر برای هر فرد وجود دارد. تعداد کل افراد شرکت‌کننده در تصویرگیری ۵۰ نفر است (تعداد کل تصاویر ۵۰۰ عدد). در این مقاله فقط از تصاویر با نمای رو به‌رو و بعد از تست یخ استفاده شده‌است.

قرمز یک روش غیرتهاجمی و ایمن است، که می‌تواند در کنار ماموگرافی برای تعیین بدخیم یا خوش‌خیم بودن عارضه بسیار ارزشمند باشد. آرورا و همکارانش [۷] با انجام یک آزمایش بالینی روی ۹۲ بیمار، به تصویرگیری دمایی مادون قرمز دیجیتال در کنار ماموگرافی و اولتراسوند دست یافتند، که این تصویرگیری در زنان با بافت پستان چگال بسیار ارزشمند بود. کندی [۸] نشان داد که ترکیب ترموگرافی و سایر روش‌های تشخیصی می‌تواند حساسیت و تشخیص را بهبود بخشد.

مطالعات زیادی پیرامون کاربرد ترموگرافی در آشکارسازی سرطان پستان در دهه‌ی ۱۹۶۰ و ۱۹۷۰ انجام شده است [۹]. چون دسته‌بندی در آن زمان به طور حسی و مستقیم انجام می‌شد، درصد پذیرش اشتباه افزایش می‌یافت. اما بر اساس تکنولوژی‌های ارائه‌شده در سال‌های اخیر، اثبات شده‌است که افزایش فعالیت متابولیکی و رشد رگ‌های خونی در نقاط همسایگی محل‌های سرطانی، تغییراتی را در دمای سطح پستان ایجاد می‌کنند [۱۰].

در مطالعه‌ی حاضر با توجه به جذابیت‌های فراوان اشاره‌شده و نیز تشخیص زودهنگام سرطان پستان با استفاده از ترموگرافی پستان، روشی پیشنهاد شده‌است که به لحاظ دقت، حساسیت و صحت بسیار قابل‌قبول است. در این تحقیق بعد از تعیین ناحیه‌ی مطلوب (ROI^2)، کیفیت تصاویر حرارتی ارتقاء یافته و سپس به کمک تکنیک آستانه‌گیری و عملگر-های لبه‌یابی، دو پستان راست و چپ از یکدیگر جدا شده و نواحی نسبی مشکوک به‌دست می‌آیند. همچنین به دلیل منحصربه‌فرد بودن دمای بدن هر شخص، داده‌ها نرمال‌سازی شدند و ۲۳ ویژگی شامل ویژگی‌های آماری، ویژگی‌هایی مبتنی بر هیستوگرام، مبتنی بر ماتریس هم‌وقوعی ($GLCM^3$)، ویژگی‌هایی بر اساس مورفولوژی نواحی مشکوک و ویژگی‌هایی در حوزه‌ی فرکانس از تصاویر حاصل استخراج می‌شوند. سپس برای کاهش بار محاسبات و افزایش دقت و صحت الگوریتم پیشنهادی، به کمک الگوریتم‌های مرسوم انتخاب ویژگی نظیر کم‌ترین افزونگی و

⁴ Minimal Redundancy Maximal Relevance

⁵ Sequential Forward Selection

⁶ Sequential Backward Selection

⁷ Sequential Floating Forward Selection

⁸ Sequential Floating Backward Selection

⁹ Genetic Algorithm

¹⁰ Support Vector Machine

¹¹ K-nearest neighbor

¹² Naïve Bayes

¹³ Probabilistic Neural Network (PNN)

¹ Specificity

² Region Of Interest (ROI)

³ Gray Level Co-occurrence Matrix

برچسب‌گذاری بر اساس معیار استاندارد TH، که برای تحلیل دمای عروقی در پستان‌ها به کار می‌رود، انجام می‌شود. این معیار استاندارد در دهه‌ی ۱۹۸۰ ارائه گردید. بنابراین پزشکان با استفاده از ترکیبی از الگوی عروقی و حرارتی سراسر پستان، تصاویر حرارتی را در ۵ دسته‌ی کلی طبقه‌بندی و گزارش می‌کنند [۱۳].

- TH₁: بافت غیرعروقی نرمال
- TH₂: بافت عروقی نرمال
- TH₃: بافت دارای ابهام
- TH₄: بافت غیرنرمال
- TH₅: بافت به شدت غیرنرمال

جدول (۱)، پراکندگی درجه‌ی TH را در هر دو پستان چپ و راست که به کمک پزشک متخصص و نیز بر اساس مقاله‌ی گاترین و همکارانش [۱۴]، برچسب‌گذاری شده‌است، نشان می‌دهد. همان‌طور که مشاهده می‌شود، تعداد تصاویر پستان چپ و راست در سطوح ۲ و ۳ دارای TH برابر هستند، در حالی که در سطوح ۱، ۴ و ۵ این برابری داده وجود ندارد. روش‌های استفاده‌شده مبتنی بر داده بودند و میزان و نوع داده در نتایج نهایی تأثیرگذار است، که می‌تواند وجود تفاوت در نتایج مربوط به دقت و صحت گزارش شده برای آشکارسازی و دسته‌بندی بافت‌ها در پستان چپ و راست را توجیه نماید.

جدول (۱) - تعداد تصاویر پستان راست و چپ در هر TH

TH	۱	۲	۳	۴	۵
پستان راست	۱۶	۲۰	۸	۳	۳
پستان چپ	۵	۲۰	۸	۱۱	۶

* برای مثال: ۱۶ تصویر از ۵۰ تصویر پستان راست دارای برچسب TH₁ است

۲-۲- روش پیشنهادی

تکنیک ارائه‌شده شامل ۴ مرحله‌ی اصلی است:

- (۱) پیش‌پردازش و پردازش تصویر (۲) استخراج ویژگی،
- (۳) انتخاب ویژگی و (۴) طبقه‌بندی و تعیین TH. شکل (۱) ساختار روش ارائه‌شده را نشان می‌دهد. در روش پیشنهادی از تصویر با نمای روبه‌رو و بعد از تست یخ استفاده شد.

نواحی اسکن‌شده‌ی بدن که تابش از آن‌ها ساطع شده، عاری از لباس و یا جواهرات بودند. دوربین به‌گونه‌ای طراحی شده‌بود که همواره در جهت کاهش و حذف خطاهای تکنیکی در فرآیند تصویرگیری عمل کند. برای کاربردهای پزشکی وجود یک محیط کنترل‌شده برای افزایش دقت و صحت بسیار مفید است. دوربین با فاصله‌ای معین از فرد قرار می‌گیرد. تعدادی از پیکسل‌های آن برای پیش‌زمینه در نظر گرفته می‌شود. بیمار نباید پشت به زمینه یا صفحه قرار گیرد. هرکدام از این پارامترها می‌تواند در نتیجه‌ی منطقی آزمایش یا روند تشخیص یا درمان کامل، موثر باشد. علاوه بر موارد اشاره‌شده، سطحی از بدن که قرار است تصویرگیری شود باید فاقد هرگونه ماده‌ی پوششی مانند کرم پودر باشد. همچنین دو ساعت قبل از تصویرگیری نباید سیگار کشیده شود. ۷ تا ۱۰ دقیقه زمان برای انطباق گرمایی لازم است. ناحیه‌ی پستان باید توسط یک خنک‌کننده سرد شود. بیماران باید در حدود ۱/۵ متر از دوربین فاصله داشته باشند و دمای اتاق باید در حدود ۲۲ درجه‌ی سانتیگراد باشد. همه‌ی پنجره‌ها باید برای جلوگیری از هرگونه تابش گرمایی به داخل اتاق دارای پوششی^۱ باشند. استانداردسازی سیستم تصویرگر گرمایی مدنظر قرار گیرد. یک منبع جسم سیاه خارجی برای چک نمودن حساسیت دما لازم است. دمای آستانه باید کالیبره شود. دوربین نیز باید حدود ۱۵ دقیقه قبل از تصویرگیری روشن شود. باید به اسکرها فرصت پایدار شدن در مدت ۳ اسکن اولیه را داد. نور فلوروسنت از محدوده‌ی در حال اسکن دور شود. از بیمار خواسته شود تا دستانش را روی سرش قرار دهد و در حالت استراحت بماند. سپس ۵ تصویر ترموگرام با زوایای مختلف ۰، ±۴۵ و ±۹۰ درجه (قبل از تست یخ) گرفته شود. پس از آن دستانش را به مدت ۲۰ دقیقه درون ظرف حاوی آب و یخ قرار دهد تا دمای کلی بدنش کاهش یابد و سپس ۵ تصویر ترموگرام دیگر با زوایای مختلف ۰، ±۴۵ و ±۹۰ درجه گرفته شود (بعد از تست یخ) [۱۲].

بعد از گرفتن تصویر در درجات مختلف و در دو حالت قبل و بعد از تست یخ، باید تصاویر هر دو پستان چپ و راست را زیر نظر پزشک متخصص برچسب‌گذاری کرد.

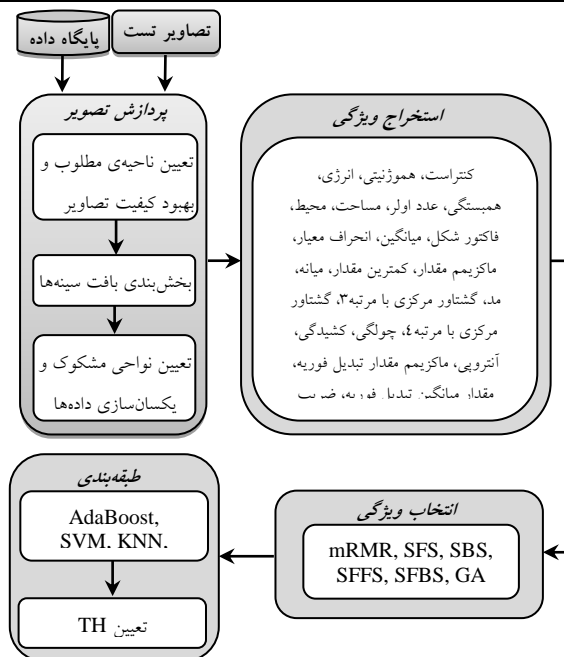
^۱ Shield

که پستان‌ها به‌طور کامل و به‌صورت قطعی در آن وجود دارند. سپس با توجه به اینکه پالت رنگ در تصاویر داده‌شده به صورت معکوس است و نواحی گرم‌تر، تیره‌تر و با شدت کمتر هستند، تصاویر مکمل شدند. به عبارتی دیگر با مکمل نمودن تصویر، الگوی رنگ تصاویر برای سازگاری با مفهوم دمایی، از پالت رنگ False به پالت رنگ True تبدیل شدند.

تصاویر به‌دست آمده از مرحله‌ی پیش‌پردازش اولیه، دارای پس‌زمینه‌ی تقریباً یکنواخت بوده و تاحدودی تار بودند و کنتراست پایینی داشتند. برای بهبود کنتراست تصاویر، ابتدا درصدی از داده‌های پایین و بالا اشباع شدند و مقادیر شدت میانی به‌گونه‌ای تغییر یافتند که تمام تغییرات ممکن دیده شوند. سپس به منظور دستیابی هر چه بهتر به جزئیات تصویر، هیستوگرام تصاویر متعادل‌سازی شدند و موارد غیر-مهم در تصاویر حذف شدند. همچنین برای حذف نویزهایی مانند نویزهای نمک و فلفلی، فیلتر میانه بر روی تصاویر اعمال شد. در نتیجه، کنتراست بهتری بین مقادیر باقی‌مانده برقرار شد و جزئیات نیز با دقت و صحت بیشتری ارائه شدند. پس از آن با به‌کارگیری لبه‌یاب سوبل و پیدا کردن لبه‌های افقی وعمودی تصویر، محدوده‌ی بافت پستان مشخص شده و دوباره فضای اطراف بافت، برای کمینه نمودن خطای تخمین آماری، برش داده شدند تا ناحیه‌ی مناسب و مطلوبی که کیفیت آن ارتقاء یافت، در تصاویر حرارتی به‌دست آمد (شکل ۲ (الف-ح)).

۲-۲-۱-۲- بخش‌بندی بافت پستان‌ها

آخرین مرحله از بخش پردازش، بخش‌بندی بافت پستان‌های چپ و راست است. ابتدا با استفاده از عملگر سوبل با آستانه ۰/۰۵، لبه‌های کناری، بالایی و پایینی بافت پستان مشخص شدند که در آشکارسازی لبه‌ها مناطقی که پیوسته بوده و تعداد پیکسل‌های آن‌ها بالاتر از ۲۰ بودند، به عنوان لبه‌های واقعی تصویر شناخته شدند. بدین منظور در روش پیشنهادی از تکنیک آستانه‌گیری بر اساس آزمون و خطا استفاده شد. سپس مختصات پیکسل‌های لبه‌های کناری، بالایی و پایینی به‌دست آمد و خطوط مار بر آن‌ها رسم گردید. میانگین مختصات نقاط لبه‌های کناری نیز محاسبه شد و خط مار بر



شکل (۱) - ساختار الگوریتم پیشنهادی

۲-۲-۱-۱- پیش‌پردازش و پردازش تصاویر

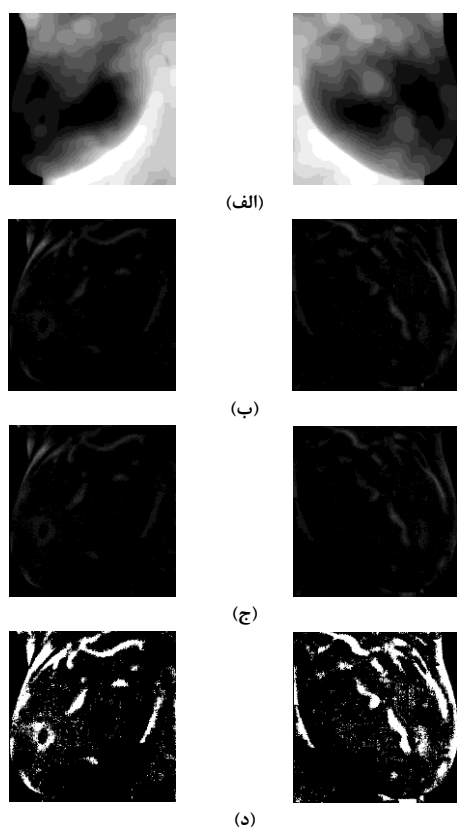
مرحله‌ی پیش‌پردازش از مراحل بسیار مهم است، به خصوص در مطالعاتی که مبتنی بر داده هستند و نوع داده در نتیجه‌ی نهایی تأثیرگذار است. بنابراین انجام عمل پیش‌پردازش و پردازش تصویر به‌صورت درست و بهینه، تأثیر زیادی در نتیجه دارد و در نهایت یک دسته‌بندی بهینه را به همراه خواهد داشت. البته منظور از بهینه بودن، علاوه بر دقت و صحت، سرعت انجام کار نیز است. بنابراین برای رسیدن به نتایج مطلوب و تمیز دادن بافت نرمال از غیرنرمال، در روش پیشنهادی عمل پیش‌پردازش و پردازش تصاویر در سه مرحله‌ی (۱) تعیین ناحیه‌ی مطلوب و بهبود کیفیت تصاویر حرارتی، (۲) بخش‌بندی بافت پستان‌ها و (۳) تعیین نواحی مشکوک و نرمال‌سازی داده‌ها انجام شد.

۲-۲-۱-۱-۱- تعیین ناحیه‌ی مطلوب و بهبود کیفیت تصاویر

اولین مرحله از روش پیشنهادی، حذف نواحی اضافی است، به نحوی که فضای باقی‌مانده فقط شامل پستان‌ها باشد. برای این منظور، ابتدا بافت‌های گردن و زیر پستان (شکم) از طریق برش اتوماتیک یا خودکار و دور ریختن آرایه‌های کناری ماتریس تصویر حذف شدند و آنچه باقی می‌ماند فضایی است

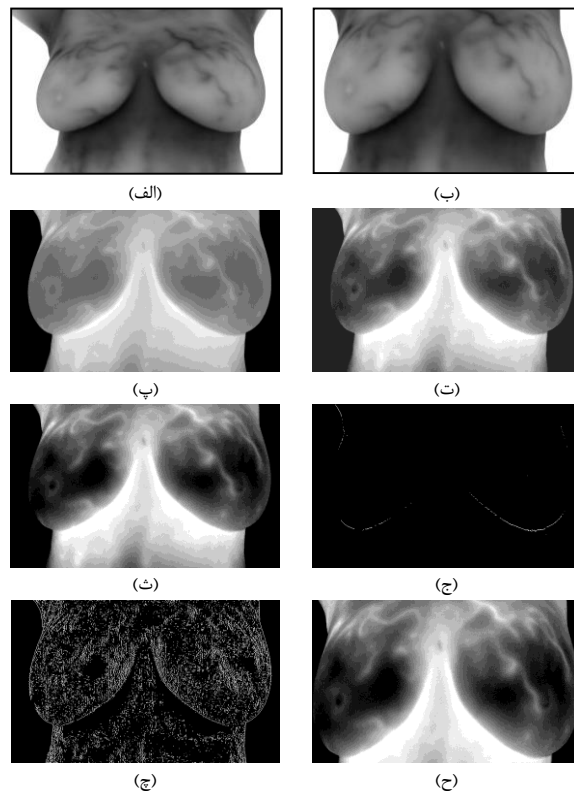
۳-۱-۲- تعیین نواحی مشکوک و نرمال سازی داده‌ها

برای استخراج دقیق‌تر نواحی مشکوک، عمل فرسایشی^۱ تصاویر به منظور یافتن تصویر زمینه‌ی تخمینی، به‌کار گرفته شد (شکل ۴(الف)) و تفاضل زمینه‌ی تخمینی و تصویر اصلی به دست آمد (شکل ۴(ب)). سپس به دلیل منحصر به فرد بودن دمای بدن هر شخص باید عمل یکسان‌سازی یا نرمال کردن داده صورت گیرد. بدین منظور درایه‌های ماتریس تصویر به نرم آن تقسیم شدند، تا کلیه‌ی نواحی و داده‌های مربوط به هر فرد یکسان‌سازی شوند (شکل ۴(ج)). در نهایت با به‌کارگیری تکنیک آستانه‌گیری تطبیقی، نواحی نسبی مشکوک به دست آمدند (شکل ۴(د)).

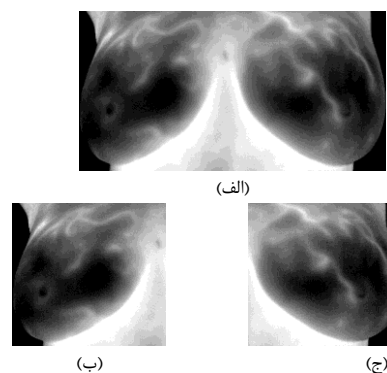


شکل ۴. تعیین نواحی مشکوک. الف) تصویر زمینه‌ی تخمینی بعد از اعمال erosion. ب) تفاضل زمینه‌ی تخمینی و تصویر اصلی، ج) نرمال‌سازی داده‌ها، د) تعیین نواحی نسبی مشکوک با به‌کارگیری تکنیک آستانه‌گیری تطبیقی. (ستون اول: پستان راست، ستون دوم: پستان چپ)

این نقطه به عنوان خط جدا کننده‌ی پستان چپ و راست مورد استفاده قرار گرفت. سپس تصویر از این نقطه به طور خودکار برش داده شد و دو تصویر مجزا از پستان چپ و راست به دست آمد (شکل ۳).



شکل (۲) - تعیین ناحیه‌ی مطلوب و بهبود کیفیت تصاویر. الف) تصویر اصلی، ب) حذف نواحی اضافی شامل بافت پستان و زیر شکم از تصویر، پ) مکمل کردن تصویر (ب)، ت) هیستوگرام تصویر (پ)، ث) اصلاح هیستوگرام تصویر، ج) اعمال لبه‌یاب سوبل به منظور دست‌یابی به لبه‌های افقی تصویر، چ) اعمال لبه‌یاب سوبل برای دست‌یابی به لبه‌های عمودی تصویر، ح) ناحیه‌ی مطلوب تعیین شده.



شکل (۳) - بخش‌بندی بافت پستان‌ها. الف) ناحیه‌ی مطلوب تعیین شده ب) تصویر پستان راست ج) تصویر پستان چپ

^۱ Erosion

۲-۲-۲- استخراج ویژگی

ویژگی استفاده می‌شود تا ویژگی‌های مؤثر انتخاب شوند و با کاهش پیچیدگی طبقه‌بندی، دقت و صحت سیستم افزایش یابد. در این مقاله از ۶ روش انتخاب ویژگی [15mRMR]، [16SFS]، [16SBS]، [17SFBS] و [18GA] استفاده شد و نتایج آن‌ها با یکدیگر مقایسه شدند. از میان این روش‌ها، از mRMR برای انتخاب ویژگی از تصاویر پستان راست و از SFBS برای انتخاب ویژگی از پستان چپ استفاده شد.

۲-۲-۴- طبقه‌بندی و تعیین TH

آخرین گام در روش پیشنهادی، مرحله طبقه‌بندی است که در آن ماتریس ویژگی شامل ویژگی‌های انتخاب‌شده، برای تعیین THها به یک الگوریتم طبقه‌بندی داده می‌شود و سپس TH-های تعیین‌شده توسط الگوریتم با THهای تعیین‌شده توسط پزشک مقایسه می‌شوند. در این تحقیق، برای نشان دادن کارایی سیستم، از روش‌های مختلف طبقه‌بندی استفاده شد و نتایج حاصل از الگوریتم‌های طبقه‌بندی، مورد ارزیابی قرار گرفته و با یکدیگر مقایسه شدند تا بهترین نتیجه به دست آید. از الگوریتم‌های مورد آزمایش، می‌توان به [19AdaBoost] و [20]، ماشین‌های بردار پشتیبان (SVM) [21]، نزدیک‌ترین همسایه (KNN) [22]، بی‌زین ساده (NB) [22] و شبکه‌ی عصبی احتمالی (PNN) [23] اشاره کرد.

۳- یافته‌ها و بحث

در این مقاله برای ارزیابی کارایی سیستم پیشنهادی از یک پایگاه داده بومی استفاده شد. این تصاویر از بیماران بیمارستان امام‌خمینی و با کمک شرکت فناوران مادون‌قرمز [11] جمع‌آوری شد. برچسب‌گذاری تصاویر نیز با کمک و مشاوره این شرکت و تحت نظر پزشک صورت گرفت. عمل تصویرگیری نیز از ۵۰ نفر در زوایای مختلف ۰، ±۴۵ و ±۹۰ درجه قبل و بعد از تست یخ انجام گرفت. در این مقاله فقط از تصاویر با نمای رو به رو و بعد از تست یخ استفاده شد.

در روش پیشنهادی، بعد از تعیین نواحی مطلوب و بهبود کیفیت تصاویر حرارتی، بافت پستان‌های راست و چپ به کمک عملگرهای لبه‌یابی از یکدیگر جدا شده و برای تعیین

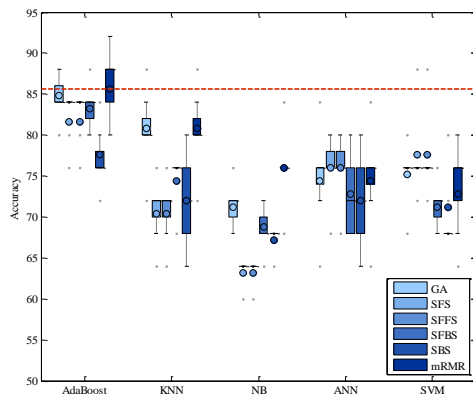
در مرحله‌ی استخراج ویژگی ضروری است اطلاعاتی را از تصاویر استخراج نموده تا سیستم بتواند به درستی بافت نرمال را از غیرنرمال تشخیص دهد. هر یک از ویژگی‌ها به‌طور مجزا بر روی پستان‌های راست و چپ محاسبه شدند. ویژگی‌های مورد استفاده عبارتند از: کنتراست^۱، هموژنیتی^۲، انرژی، همبستگی^۳، عدد اویلر^۴، مساحت، محیط، فاکتور شکل، میانگین، انحراف‌معیار، ماکزیمم مقدار، کم‌ترین مقدار، میانه، مد، گشتاور مرکزی با مرتبه‌ی ۳، گشتاور مرکزی با مرتبه‌ی ۴، چولگی^۵، کشیدگی^۶، آنتروپی، بیشینه مقدار تبدیل فوریه^۷، مقدار میانگین تبدیل فوریه^۸، ضریب Hurst و نرم^۹. موارد ذکر شده شامل ویژگی‌های آماری، ویژگی‌های مبتنی بر هیستوگرام، ویژگی‌های مبتنی بر ماتریس هم‌وقوعی، ویژگی‌هایی بر اساس مورفولوژی نواحی مشکوک و ویژگی‌هایی در حوزه‌ی فرکانس هستند. ویژگی‌های کنتراست، هموژنیتی، انرژی و همبستگی، مبتنی بر ماتریس هم‌وقوعی؛ ویژگی‌های عدد اویلر، مساحت، محیط و فاکتور شکل بر اساس ویژگی‌های مورفولوژیک نواحی مشکوک؛ ویژگی‌های میانگین، انحراف‌معیار، ماکزیمم مقدار و کم‌ترین مقدار، ویژگی‌های مبتنی بر هیستوگرام؛ ویژگی‌های میانه، مد، چولگی و کشیدگی، ویژگی‌های آماری و ویژگی‌های ماکزیمم مقدار تبدیل فوریه و مقدار میانگین تبدیل فوریه ویژگی‌هایی در حوزه‌ی فرکانس هستند.

۲-۲-۳- انتخاب ویژگی

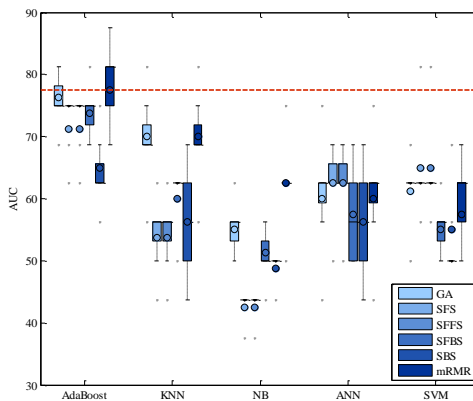
اطلاعات متنوع منجر به به‌دست‌آوردن بردارهای ویژگی با ابعاد بالا می‌شوند که به سبب پیچیدگی، دقت و صحت را کاهش و بارمحاسباتی را افزایش می‌دهند. بنابراین برای کارایی بالاتر طبقه‌بندی‌کننده، از روش‌های مختلف انتخاب

^۱ Contrast^۲ Homogeneity^۳ Correlation^۴ Euler^۵ Skewness^۶ Kurtosis^۷ Max value of FFT^۸ Mean value of FFT^۹ Norm

پیشنهادی روی تصاویر پستان راست و چپ است. در نمودار box-whisker، دایره‌ی موجود در هر box ای که برای بررسی روش‌های مختلف طبقه‌بندی با روش‌های مختلف انتخاب ویژگی رسم شدند، بیانگر میانگین و خط افقی رسم شده میان این boxها بیانگر میانه است. نقاط و خطوط افقی خارج از boxها نیز بیانگر داده‌های پرت هستند. خط‌چین رسم شده بیانگر بالاترین صحت میانگین به دست آمده از ترکیب یکی از روش‌های انتخاب ویژگی با یکی از روش‌های طبقه‌بندی در تصاویر پستان چپ و راست است و همان‌طور که گفته شد از ترکیب AdaBoost با SFBS بهترین صحت برای تصاویر پستان چپ و از ترکیب AdaBoost با mRMR بهترین صحت برای تصاویر پستان راست حاصل می‌شود. مقایسه‌ی صحت حاصل از این روش با سایر روش‌ها در تصاویر پستان راست و چپ، در شکل (۱۱) و بهترین صحت‌های حاصل از الگوریتم‌های mRMR و SFBS در جدول (۲) قابل مشاهده است.



شکل (۵) - بررسی صحت سیستم براساس روش‌های مختلف انتخاب ویژگی و طبقه‌بندی در تصاویر پستان راست



شکل (۶) - بررسی AUC سیستم براساس روش‌های مختلف انتخاب ویژگی و طبقه‌بندی در تصاویر پستان راست

نواحی نسبی مشکوک در هر یک از پستان‌ها از عملگر فرسایشی با دیسک دایره‌ای به شعاع ۱۰ استفاده شد. سپس ویژگی‌های ذکر شده از تصاویر پستان راست و چپ استخراج شدند و برای دست‌یابی به ویژگی‌های بهینه، روش‌های مختلف انتخاب ویژگی به‌کار گرفته شدند. در نهایت ماتریس ویژگی تشکیل شده برای تعیین TH با الگوریتم‌های مختلف طبقه‌بندی مورد بررسی قرار گرفت که از میان الگوریتم‌های مورد استفاده، بهترین نتیجه از الگوریتم AdaBoost حاصل شد. نتایج آزمایشات نشان می‌دهند، زمانی که تعداد نقاط همسایگی در روش طبقه‌بندی KNN برابر با ۳، مقدار پارامتر هموارساز در شبکه‌ی عصبی احتمالی برابر با ۰/۳۵ و برای روش طبقه‌بندی SVM نیز کرنل Gaussian با درجه‌ی ۶ و پارامتر C برابر با ۱ در نظر گرفته شود، نتایج خوبی به دست می‌آید. همچنین برای ارزیابی روش پیشنهادی، از 5-fold cross validation استفاده شد. به علاوه، معیارهایی شامل حساسیت^۱، تشخیص، صحت^۲، سطح زیر منحنی ROC^۳ (AUC^۴)، پذیرش اشتباه (FPR^۵) نیز برای نشان دادن کارایی سیستم، به‌کار گرفته شدند.

۳-۱- ارزیابی کارایی روش پیشنهادی

در این قسمت کارایی حاصل از روش پیشنهادی با ترکیب-های مختلف روش‌های انتخاب ویژگی و طبقه‌بندی، ارزیابی شد. در این آزمایش از ۵ روش طبقه‌بندی AdaBoost، SVM، KNN، NB و PNN و ۶ روش انتخاب ویژگی mRMR، SFBS، SFSS، SFBS، SBS و GA استفاده شد. نتایج به کمک نمودار box-whisker در شکل‌های ۵ تا ۷ برای پستان راست و در شکل‌های ۸ تا ۱۰ برای پستان چپ ارائه شدند. در نمودارهای رسم شده محور X بیانگر روش‌های مختلف طبقه‌بندی و محور Y در اشکال مختلف بیانگر صحت، AUC و F-measure به دست آمده در سیستم

^۱ Sensitivity

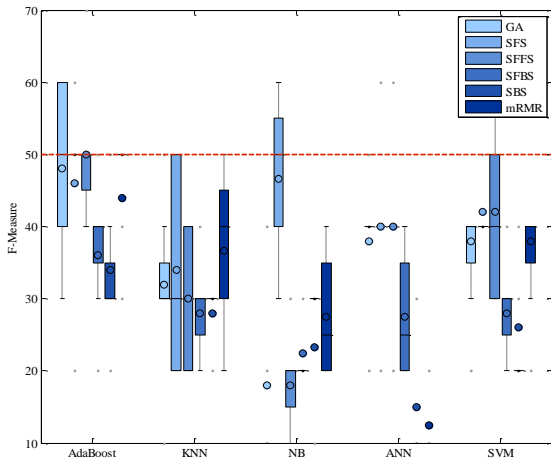
^۲ Accuracy

^۳ Receiver Operating Characteristic

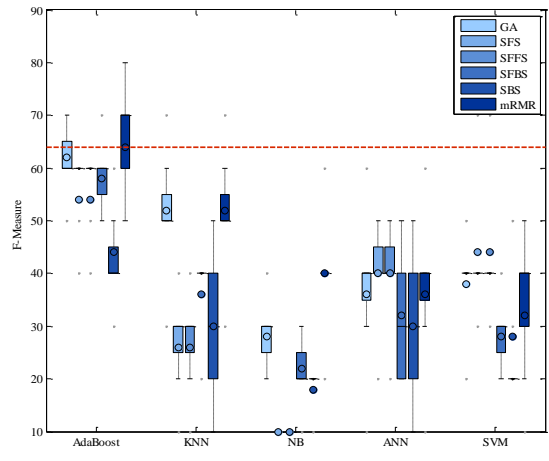
^۴ Area Under Curve

^۵ Equal Error Rate

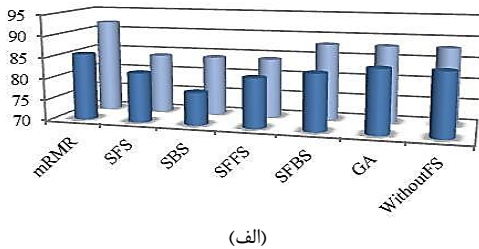
^۶ False Positive Rate



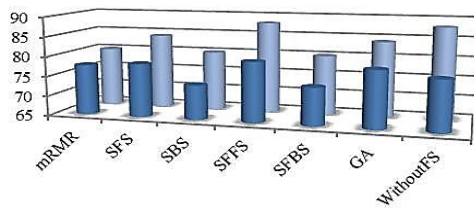
شکل (۱۰)- بررسی F-Measure سیستم براساس روش‌های مختلف انتخاب ویژگی و طبقه‌بندی در تصاویر پستان چپ



شکل (۷)- بررسی F-Measure سیستم براساس روش‌های مختلف انتخاب ویژگی و طبقه‌بندی در تصاویر پستان راست



(الف)

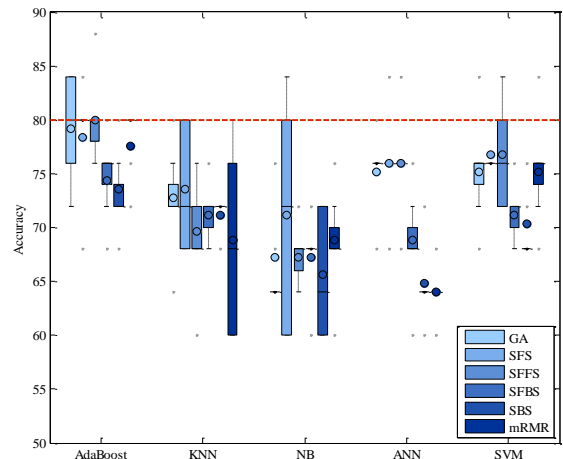


(ب)

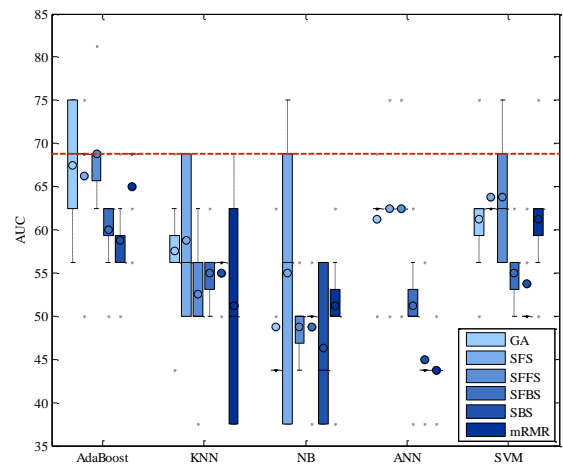
شکل (۱۱)- مقایسه‌ی صحت‌های حاصل از الگوریتم‌ها (الف) روی تصاویر پستان راست (ب) روی تصاویر پستان چپ

جدول (۲)- ارزیابی بهترین ترکیب حاصل روی پستان راست و چپ

ترکیب AdaBoost با mRMR روی پستان راست	ترکیب AdaBoost با SFSS روی پستان چپ	بهترین ترکیب‌های حاصل شده
۸۵/۶۰٪	۸۰٪	Maen Accuracy
۰/۶۴	۰/۵	Mean Sensitivity
۰/۹۱	۰/۸۷۵	Mean specificity
۷۷/۵٪	۶۷/۷۵٪	Mean AUC
۰/۲۲۵۰	۰/۳۱۲۵	Mean EER
۶۴٪	۵۰٪	Mean F-Measure
۹۲٪	۸۸٪	Max Accuracy
۰/۰۹	۰/۱۲۵	Mean FPR



شکل (۸)- بررسی صحت سیستم براساس روش‌های مختلف انتخاب ویژگی و طبقه‌بندی در تصاویر پستان چپ



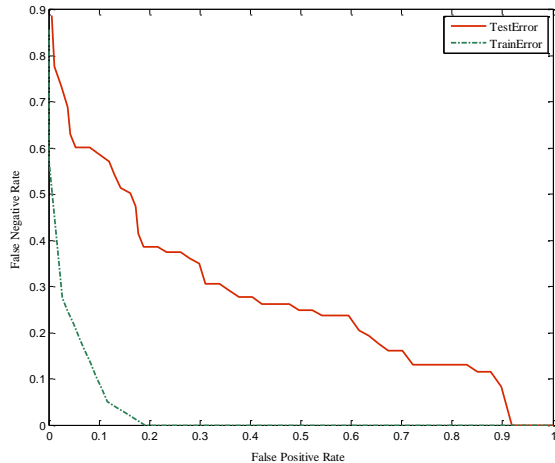
شکل (۹)- بررسی AUC سیستم براساس روش‌های مختلف انتخاب ویژگی و طبقه‌بندی در تصاویر پستان چپ

۲-۳- بررسی خطای آموزش و تست روش پیشنهادی

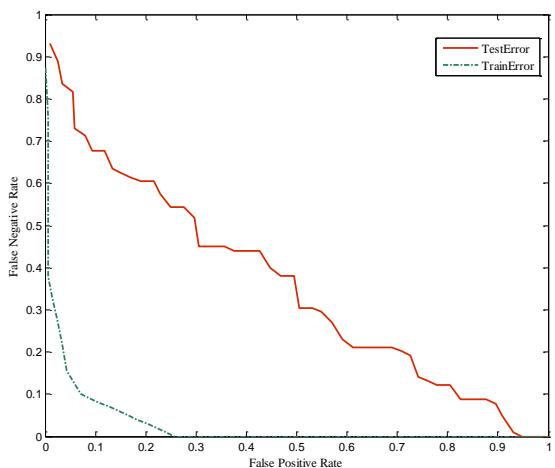
در این قسمت برای بررسی درستی الگوریتم پیشنهادی، خطای آموزش و تست مورد بررسی قرار گرفت. همانطور که در نمودار^۱ DET در شکل‌های (۱۲ و ۱۳) نیز مشاهده می‌شود، خطای آموزش سیستم پیشنهادی از خطای تست کمتر است که این امری بدیهی و منطقی است و در این نمودارها نیز درستی آن به اثبات رسیده است. در نمودار DET، محور X بیانگر FPR و محور Y بیانگر FNR است. هر دو خطای نرخ پذیرش اشتباه (FPR) و نرخ رد اشتباه (FNR) در سیستم تشخیص سرطان مهم هستند. خطای FNR موقعی رخ می‌دهد که تصویر نرمال باشد اما غیرنرمال تشخیص داده شود، یعنی به اشتباه گفته شده که غیرنرمال است. خطای FPR موقعی رخ می‌دهد که تصویر غیرنرمال است ولی نرمال تشخیص داده شده، یعنی به اشتباه گفته شده که نرمال است. بنابراین در صورت رخ دادن خطای FPR باید هزینه‌ی بیشتری را متحمل شد، زیرا تصویری که غیرنرمال بوده، نرمال تشخیص داده شده است که این خطا ممکن است غیر قابل جبران باشد. مسئله‌ای که در تشخیص سرطان پستان نیز از اهمیت ویژه‌ای برخوردار است، کاهش درصد پذیرش اشتباه (FPR) است. "FPR" در این سیستم برای تصاویر پستان راست، در حالت میانگین ۰/۰۹ و در کم‌ترین حالت (حالتی که سیستم دارای صحت بیشینه است) برابر با ۰/۰۵ و برای تصاویر پستان چپ، در حالت میانگین ۰/۱۲۵ و در کم‌ترین حالت برابر با ۰/۰۷۵ است، که این مقدار به دست آمده برای FPR نیز بیانگر کارایی سیستم پیشنهادی است.

۳-۳- مقایسه با روش‌های پیشین

در جدول (۳) مقایسه‌ای بین روش‌های مختلف انجام گرفته است، که نتایج حاصل بیانگر کارایی بالای الگوریتم پیشنهادی است. لازم به ذکر است که برای روش‌های مطرح شده بهترین دقت و صحت حاصل، نمایش داده شد و مورد مقایسه قرار گرفت. همچنین هر یک از مقالات اشاره شده از پایگاه داده‌های بومی برای انجام آزمایشات استفاده کرده‌اند.



شکل (۱۲)- نمودار DET برای بررسی خطای آموزش و تست در تصاویر پستان راست



شکل (۱۳)- نمودار DET برای بررسی خطای آموزش و تست در تصاویر پستان چپ

در [Acharya۲۴] (نام فارسی) و همکارانش از ۵۰ تصویر (۲۵ تصویر نرمال و ۲۵ تصویر با بافت سرطانی) که از بیماران بیمارستان سنگاپور گرفته شده است، استفاده کرده‌اند. آن‌ها از ویژگی‌های بافت بر اساس ماتریس هم‌وقوعی بهره گرفته‌اند و سپس ویژگی‌های استخراج شده را برای تعیین وضعیت نرمال و بدخیمی به روش طبقه‌بندی ماشین‌های بردار پشتیبان (SVM) داده‌اند. برای روش مورد اشاره صحت برابر با ۰/۸۸/۱۰، حساسیت برابر با ۰/۸۵/۷۱ و تشخیص برابر با ۰/۹۰/۴۸ گزارش شده است.

در مقاله‌ی [۲۵]، قیومی‌زاده و همکارانش از پایگاه داده‌ای شامل ۲۰۰ بیمار استفاده کرده‌اند. آن‌ها برای به دست آوردن

^۱ Detection Error Trade-off

^۲ False Negative Rate

جدول (۳) - مقایسه‌ی روش پیشنهادی با روش‌های پیشین

Specificity	Sensitivity	Accuracy	روش‌ها
٪۹۰/۴۸	٪۸۵/۷۱	٪۸۸/۱۰	Acharya و همکارانش [۲۴]
٪۹۷	٪۹۳	-	[۲۵] Ghayoumi Zadeh
-	-	٪۷۸/۵۶	Yaneli و همکارانش [۲۶]
٪۸۶/۵	٪۸۵/۷	-	Araujo و همکارانش [۲۷]
٪۹۵	٪۸۰	٪۸۸	روش پیشنهادی (پستان راست)
٪۹۲/۵	٪۷۰	٪۹۲	روش پیشنهادی (پستان چپ)

ترموگرافی مادون قرمز یک تکنولوژی امیدبخش برای آشکار-سازی سرطان پستان است که می‌تواند به عنوان یکی از روش‌های تصویرگیری مکمل به منظور تکمیل نتایج ماموگرافی به کار رود. مزایای تصویرگیری دمایی نسبت به ماموگرافی عبارتند از:

- ۱- پرهیز از اعمال پرتوهای یونیزان و مضر
- ۲- اولین نشانه‌های سرطان پستان توسط ماموگرافی‌های تجاری حدود ۸ تا ۱۰ سال دیرتر از ترموگرافی تشخیص داده می‌شوند که با توجه به اهمیت مسئله‌ی زمان در درمان سرطان، در حدود ۹۹٪ شانس درمان افزایش می‌یابد.
- ۳- ترموگرافی می‌تواند منجر به حذف بیوپسی یا نمونه‌برداری‌های غیرضروری گردد. مطالعات نشان داده‌اند که در ۷۰ تا ۸۰ درصد مواردی که ماموگرام‌ها آن‌ها را آشکار نموده‌اند، نمونه‌برداری بیهوده و غیرلازم انجام می‌گیرد که هیچ علامتی از سرطان پستان در آن‌ها وجود ندارد. ترموگرافی پستان صحتی در حدود ۹۰ درصد دارد که طبق تحقیقات صورت گرفته می‌تواند نرخ امید به زندگی را افزایش دهد. به علاوه یک روش انفعالی است، چون هیچ پرتوی خطرناک یونیزان به طرف بیمار ساطع نکرده و به دور از هرگونه خطری است. بنابراین، در این مقاله روش جدیدی برای تشخیص سرطان پستان بر اساس ترموگرافی ارائه شد تا بتوان نواحی مشکوک و TH مربوطه را تعیین نمود. مراحل انجام کار در الگوریتم پیشنهادی، برای دست‌یابی به ۵ دسته‌ی TH₁-TH₅، شامل ۴ بخش اصلی پیش‌پردازش و پردازش تصویر، استخراج ویژگی، انتخاب ویژگی و طبقه‌بندی

پارامترهای مهم تصویر، از الگوریتم ژنتیک استفاده نموده و سپس برای خوشه‌بندی مولفه‌های اصلی از شبکه‌های عصبی فازی استفاده کرده‌اند. حساسیت و تشخیص روش قیومی‌زاده و همکارانش به ترتیب ۹۳٪ و ۹۷٪ است. Yaneli و همکارانش [۲۶] نیز با استفاده از شبکه‌ی بیزین از روی ۹۸ بیمار که مشکوک به سرطان بودند، مقادیری را برای ویژگی‌های حرارتی تخمین زدند. آن‌ها ادعا کردند که این مدل می‌تواند تعاملات غیرمنتظره میان ویژگی‌های حرارتی را نشان دهد. صحت حاصل از این روش ۷۸/۵۶٪ گزارش شده‌است.

هدف از بررسی صورت‌گرفته توسط Araujo و همکارانش [۲۷]، ارزیابی امکان استفاده از داده‌های فاصله در تجزیه و تحلیل داده‌های نمادین (SDA^۱) است که بر این اساس توانستند مدلی را برای اختلالات پستان (بدخیم، خوش‌خیم و کیست) به منظور تشخیص سرطان به دست آورند. در این راستا، آن‌ها روش استخراج ویژگی‌ای را در ۳ مرحله ارائه کردند. در مرحله‌ی اول، ۴ مورد از متغیرهای فاصله را از طریق مقادیر حداقل و حداکثر درجه‌ی حرارت ماتریس مورفولوژی و حرارتی به دست آوردند. در مرحله‌ی دوم، اپراتورها یا کاربران بر اساس تفاوت در فواصل در نظر گرفته شدند و سپس ویژگی‌های پیوسته و متوالی را به دست آوردند. در آخرین مرحله، ویژگی‌های پیوسته که توسط معیار fisher تغییر شکل می‌دهند، برای انجام عمل طبقه‌بندی به روش‌های مختلف طبقه‌بندی داده شدند. آن‌ها آزمایشاتشان را بر روی پایگاه داده‌ای شامل تصاویر ترموگرافی پستان در برزیل انجام دادند. نتایج حاصل را با حساسیتی برابر با ۸۵/۷٪ و تشخیصی برابر با ۸۶/۵٪ گزارش کردند.

۴- نتیجه‌گیری

در روش پیشنهادی از تکنیک تصویرگیری ترموگرافی استفاده شد تا بتوان به تشخیص زودهنگام تغییرات در حال وقوع در بافت پستان و سلول‌های سرطانی دست یافت. از ترموگرافی بر اساس این اصل که فعالیت متابولیکی و گردش خون در هر دو نوع بافت قبل سرطان و نواحی اطراف دچار سرطان پستان، اغلب بالاتر از بافت نرمال هستند، استفاده می‌شود.

^۱ Symbolic data analysis

- data system: positive predictive value of mammographic features and final assessment categories”, *AJR*, Vol. 171, pp. 35 – 40, 1998.
- [2] L. Liberman, “Clinical management issues in percutaneous core breast biopsy”, *Radiol Clin North Am*, Vol. 38, pp. 791 – 807, 2000.
- [3] S. B. Fox, K. C. Gatter, R. D. Leek, A. L. Harris, J. Bliss, J. L. Mansi and B. Gusterson, “Association of tumor angiogenesis with bone marrow micrometastases in breast cancer patients”, *Journal of the National Cancer Institute*, Vol. 89, No. 14, pp. 1044 – 1049, 1997.
- [4] INCA, Instituto Nacional do Câncer. Available from: URL: <http://www2.inca.gov.br>.
- [5] K. Lloyd-Williams, R. S. Handley, “Infrared thermometry in the diagnosis of breast disease”, *Lancet*, Vol. 2, pp. 1378 – 1381, 1961.
- [6] Y. R. Parisky, A. Sardi, R. Hamm, K. Hughes, L. Esserman, S. Rust and K. Callahan, “Efficacy of computerized infrared imaging analysis to evaluate mammographically suspicious lesions”, *AJR*, Vol. 180, pp. 263 – 269, 2003.
- [7] A. Nimmi, D. Martins, D. Ruggerio, E. Tousimis, A. J. Swistel, M. P. Osborne and R. M. Simmons, “Effectiveness of a noninvasive digital infrared thermal imaging system in the detection of breast cancer”, *The American Journal of Surgery*, Vol. 196, No. 4, pp. 523-526, 2008.
- [8] D. A. Kennedy, T. Lee, “A comparative review of thermography as a breast cancer screening technique”, *Integrative Cancer Therapies*, Vol. 8, No. 1, pp. 9-16, 2009.
- [9] A. Drastich, “Záření v lékařství III. Infradiometrické detekční a zobrazovací systémy”, *Skripta FEI VUT v Brně*, 1980.
- [10] R. C. Serrano, R. Lima, “Using Hurst coefficient and Lacunarity to diagnosis early breast diseases”, *Proc. 17th International Conference on Systems, Signals and Image Processing*, pp. 550-553, 2010.
- [11] Fanavaran Madoon Ghermez (FMG) Co., Ltd. Available from: www.fmg-med.ir.
- [12] Expanding the use of thermal imaging in the diagnosis of breast cancer. Doc 002–005.
- [13] W. C. Amalu, W. B. Hobbins, J. F. Head, R. L. Elliott, “Infrared imaging of the breast-An overview”, *Medical Devices & Systems*, The Biomedical Engineering Handbook, Edited by Joseph D. Bronzino, 3rd Eds, 2006.
- [14] M. Gautherie, P. Haehnel, J. P. Walter, L. G. Keit, “Thermovascular changes associated with in situ and minimal breast cancers, Results of an ongoing prospective study after four years. The Journal of reproductive medicine”, Vol. 32. No. 11, pp. 833 – 842, 1987.

ویژگی‌های انتخاب‌شده است. از مزایای این روش می‌توان به موارد ذیل اشاره کرد: اول، در روش پیشنهادی تمامی مراحل پیش‌پردازش شامل تعیین ناحیه‌ی مطلوب و بهبود کیفیت تصاویر حرارتی، بخش‌بندی بافت پستان‌ها و تعیین نواحی نسبی مشکوک به طور خودکار انجام گرفت. دوم، به دلیل منحصر به فرد بودن دمای بدن هر شخص، عمل نرمال‌سازی داده با تقسیم درایه‌های ماتریس تصویر به نرم آن صورت گرفت. سوم، ۲۳ ویژگی با انواع مختلف از هر یک از تصاویر پستان راست و چپ استخراج شد و سپس به کمک روش‌های مختلف انتخاب ویژگی، ویژگی‌های مؤثر شناسایی شدند و در نهایت ویژگی‌های انتخاب‌شده با روش‌های مختلف طبقه‌بندی مورد ارزیابی قرار گرفتند. با بررسی‌های صورت‌گرفته، ویژگی‌های کنتراست، انرژی، عدد اوپلر و کشیدگی به عنوان ویژگی‌های تأثیرگذار مشخص شدند. در نهایت برای ارزیابی کارایی روش پیشنهادی، از استراتژی 5-fold cross validation استفاده شد. از ترکیب AdaBoost با روش انتخاب ویژگی SFFS بهترین میانگین صحت ۸۰٪ و صحت بیشینه‌ی ۸۸٪ برای تصاویر پستان چپ به دست آمد و از ترکیب AdaBoost با روش انتخاب ویژگی mRMR بهترین میانگین صحت ۸۵/۶۰٪ و صحت بیشینه‌ی ۹۲٪ برای تصاویر پستان راست حاصل شد. همچنین درصد پذیرش اشتباه برای بهترین ترکیبات به دست آمده برای تصاویر پستان چپ و راست به ترتیب بیانگر ۰/۱۲۵ و ۰/۰۹ خواهند بود. با توجه به این‌که روش‌های دسته‌بندی خودکار مبتنی بر داده بودند، عدم تساوی محتوای تصویری در TH‌های ۱، ۴ و ۵، می‌تواند یکی از دلایل وجود تفاوت در صحت باشد.

۵- سپاسگزاری

از صندوق حمایت از پژوهشگران به خاطر حمایت مالی و از شرکت فناوران مادون قرمز، جناب آقای منصور علیدوستی و سرکار خانم مهندس نوید بابت همکاری صمیمانه در این پروژه سپاسگزاریم.

۶- مراجع

- [1] L. Liberman, A. F. Abramson, F. B. Squires, J. R. Glassman, E. A. Morris and D. D. Dershaw, “The breast imaging reporting and

- [22] Mitchell TM. *Machine Learning*, 1997.
- [23] D. F. Specht DF, "Probabilistic neural network for classification, map, or associative memory", In *Proceeding of the IEEE international conference on neural network*, pp. 525 – 532, 1988.
- [24] U. R. Acharya, E. Y. Ng, J. H. Tan, S. V. Sree, "Thermography based breast cancer detection using texture features and support vector machine", *Journal of Medical Systems*, Vol. 36, No. 3, pp.1503 – 1510, 2010.
- [25] H. Ghayoumi Zadeh, O. Pakdelazar, J. Haddadnia, G. RezaeiRad and M. Mohammad-Zadeh, "Diagnosing breast cancer with the aid of fuzzy logic based on data mining of a genetic algorithm in infrared images", *Middle East Journal of Cancer*, Vol. 3, No. 4, pp. 119 – 129, 2011.
- [26] A. A. M. Yaneli, C. R. Nicandro, M. M. Efren, P. C. Nancy, A. M. H. Gabriel, "Assessment of bayesian network classifiers as tools for discriminating breast cancer pre-diagnosis based on three diagnostic methods", In *Advances in Artificial Intelligence*, pp.419-431, 2013.
- [27] M. C. Araujo, R. C. Lima, R. M. De Souza, "Interval symbolic feature extraction for thermography breast cancer detection", *Expert System with Applications*, Vol. 41, No. 15, pp. 6728 – 6737, 2014.
- [15] H. C. Peng, F. H. Long, C. Ding, "Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy", *IEEE Transactions on Pattern Analysis Machine Intelligence*, Vol. 27, No. 8, pp.1226– 1238, 2005.
- [16] A. W. Whitney, "A direct method of nonparametric measurement selection", *IEEE Transaction on Computers*, Vol. 20, No. 9, pp. 1100 – 1103, 1971.
- [17] P. Pudil, J. Novovicova, J. Kittler, "Floating search methods in feature-selection", *Pattern Recognition Letters*, Vol. 15, No. 11, pp.1119 – 1125, 1994.
- [18] M. J. Martin-Bautista, M. A. Vila, "A Survey of Genetic Feature Selection in Mining Issues", *Proceedings of Congress on Evolutionary Computation*, pp. 1314–1321, 1999.
- [19] Y. Freund, R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting", *J Computer System Sciences*, Vol. 55, No. 1, pp.119 – 139, 1997.
- [20] R. E. Schapire, Y. Singer, "Improved boosting algorithms using confidence-rated predictions", *Mach Learn*, Vol. 37, No.3, pp. 297–336, 1999.
- [21] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition", *Data Mining and Knowledge Discovery*, pp. 121 – 167, 1998.