



Improve the Quality of Mammogram Images by Image Processing Techniques

Mahdi Hariri^{1*}, Hasan Najafy²

1. Assistant Professor, Electrical Engineering Department, University of Zanjan, Zanjan, Iran. (*Corresponding Author*) ma.hariri@znu.ac.ir
2. MSc Student, Electrical and Computer engineering department, Zanjan Branch, Islamic Azad University, Zanjan, Iran. najafe@gmail.com

Abstract.

Introduction: Unfortunately, breast cancer has become the most common disease among women, leading to death. Breast cancer can be diagnosed using the classification of breast masses. Among the various methods used to diagnose breast cancer, mammography is considered the most promising method and has been used many times by doctors. Mammogram images are usually noisy and have low contrast. Therefore, the creation of an automatic or semi-automatic diagnostic aid system can be of great help to the doctor in the disease diagnosis process. In this field, machine learning and image processing algorithms are used. Therefore, in this research, we have tried to improve the quality of mammogram images and increase segmentation accuracy to increase the accuracy of diagnosis and segmentation of breast cancer tumor masses. To achieve this goal, we used image processing and data mining techniques.

Method: This research is looking for a method that can improve the accuracy of the final diagnosis. Therefore, after the pre-processing stage, which includes rescaling and adjusting the texture of the image, highlighter transforms in the frequency domain such as Curvelet and Contourlet are used to highlight and increase the differentiation of areas with masses in the image for decision-making. Traditional enhancement methods are fixed neighborhood methods used to enhance images based on overall features. However, local features based on image zoning are used for image segmentation, and these methods are also used to increase the contrast of mammogram images concerning their surroundings. The improvement methods used in this research use features based on the wavelet domain.

Results: The input image to the system is subjected to the feature extraction process and three main categories of frequency, morphology, and histology features are extracted from it. This process is done through the cycle, size equalization, histogram adjustment, contourlet transform, and curvelet transform. Due to the sensitivity of the systems, it has been tried to extract the features with various levels and known matrices such as the matrix of events and gray groups. The classification results evaluated the said method. The best results on the data set were the proposed method, which reached an accuracy rate of 86.3 and showed a good improvement on the displayed data set.

Keywords: Mammogram images, Cancer diagnosis, Quality improvement, Histogram equalization, Feature extraction, Contourlet transform, Curvelet transform.

بهبود کیفیت تصاویر ماموگرافی با استفاده از تکنیک‌های پردازش تصویر

سال سوم، بهار ۱۴۰۱
شماره اول، صص: ۵۷-۶۹

تاریخ دریافت: ۱۴۰۱/۰۳/۲۰
تاریخ پذیرش: ۱۴۰۱/۰۶/۱۰

مهدی حریری^{۱*}، حسن نجفی^۲

۱. استادیار، گروه مهندسی برق، دانشگاه زنجان، زنجان، ایران. ma.hariri@znu.ac.ir

۲. کارشناس ارشد الکترونیک دیجیتال، واحد زنجان، دانشگاه آزاد اسلامی، زنجان، ایران. najafe@gmail.com

چکیده: با توجه به گسترش ابتلا به سرطان سینه، تشخیص زودهنگام آن از روی تصاویر ماموگرام با استفاده از روش‌های کامپیوتری، به‌عنوان روشی مؤثر در کاهش میزان مرگ و میر بیماران مورد توجه قرار گرفته است. از این رو در این پژوهش روشی مبتنی بر تکنیک‌های پردازش تصویر جهت بهبود کیفیت تصاویر ماموگرافی ارائه شده است. از این رو در این پژوهش در جهت ایجاد یک سیستم پزشکیار، سعی در بهبود کیفیت تصاویر ماموگرافی با تکنیک‌های پردازش تصویر داریم. این پژوهش دارای دو مرحله پیش پردازش شامل یکسان‌سازی ابعاد و تعدیل هیستوگرام تصاویر و مرحله استخراج ویژگی با استفاده از تبدیل‌های کانتورلت و کرولت از تصاویر ماموگرافی است که سه دسته ویژگی ریخت‌شناسی و بافت‌شناسی، آماری و فرکانسی را جهت بهبود تشخیص ارائه می‌نماید و سبب افزایش صحت تشخیص می‌شود. روش بهبود پیشنهادی روی مجموعه داده MIAS اجرا شد و زیرمجموعه‌ای از ویژگی‌های استخراج شده برای ورودی طبقه‌بند انتخاب گردید. با مقایسه عملکرد روش پیشنهادی روی طبقه‌بندهای مختلف، این روش میزان صحت ۸۶٫۳ را نشان می‌دهد که نسبت به بقیه روش‌ها نتیجه بهتری می‌باشد.

واژه‌های کلیدی: ماموگرافی، تشخیص سرطان، بهبود کیفیت تصویر، تعدیل هیستوگرام، استخراج ویژگی، تبدیل کانتورلت، تبدیل کرولت.

۱. مقدمه و بیان مسئله

سرطان سینه یکی از مهم‌ترین دلایل مرگ و میر در بین زنان در سراسر جهان است. مشاهده شده‌است که تشخیص زودهنگام بدخیم‌بودن تومور می‌تواند به فرآیند درمان بیماری کمک کرده و امید به بقا و نجات از بیماری را به‌طور چشمگیری افزایش دهد [۱]. متأسفانه سرطان سینه تبدیل به رایج‌ترین بیماری منجر به مرگ در بین زنان شده‌است. سرطان سینه می‌تواند با استفاده از کلاس‌بندی توده‌های سینه تشخیص داده‌شود. دو نوع مختلف از تومورها، تومورهای خوش‌خیم و تومورهای بدخیم وجود دارند. پزشکان نیازمند یک پروسه تشخیص قابل‌اطمینان برای تمایز بین این تومورها هستند. اما به‌طور کلی، تشخیص و جداسازی تومورها حتی توسط افراد مجرب نیز کار دشواری است. از این رو ایجاد یک سیستم کمک تشخیصی خودکار یا نیمه‌خودکار می‌تواند کمک شایانی به فرآیند تشخیص بیماری نماید. در این زمینه از الگوریتم‌های یادگیری ماشینی و پردازش تصویر استفاده می‌شود. پژوهشگران ثابت کرده‌اند که این الگوریتم‌ها، کارایی و عملکرد بهتری را در آشکارسازی و تشخیص از خود نشان داده‌اند [۲].

از بین روش‌های مختلفی که برای تشخیص سرطان سینه تاکنون به‌کار گرفته شده‌اند، ماموگرافی به‌عنوان امیدبخش‌ترین روش محسوب شده و به دفعات توسط پزشکان مورد استفاده قرار گرفته‌است. تصاویر ماموگرام معمولاً نویزدار و دارای کنتراست پایین هستند. در ماموگرافی سینه، معمولاً نواحی روشن نشان‌دهنده سرطان هستند. در برخی از تصاویر ماموگرام، ممکن است هم بافت‌های بدخیم و هم بافت‌های توده‌ای نرمال مشاهده شوند. تمایز بین بافت‌های توده‌ای بدخیم و نرمال، تنها از طریق آستانه‌یابی امکان‌پذیر نیست. به‌دست‌آوردن اطلاعات نواحی توده‌ای سرطانی در یک تصویر ماموگرام، امری ضروری به‌شمار آمده و به شناسایی تومور و قطعه‌بندی آن کمک می‌کند. از این رو، تشخیص ضایعات سرطانی در تصاویر ماموگرام تبدیل به یکی از حوزه‌های پژوهشی فعال شده‌است. تاکنون روش‌های متعددی از جمله سیستم‌های تشخیص به کمک کامپیوتر و روش‌های مبتنی بر شدت نور، برای قطعه‌بندی سرطان سینه در تصاویر ماموگرام ایجاد و معرفی شده‌اند. با این حال، هنوز راه‌حلی که بتواند معیارهای مربوط به ضایعات سرطانی را به صورت کامل تشخیص دهد، به‌وجود نیامده‌است [۱]. از این رو در این پژوهش برای افزایش دقت تشخیص و قطعه‌بندی توده‌های توموری سرطان سینه، سعی در بهبود کیفیت تصاویر ماموگرام و افزایش دقت قطعه‌بندی نموده‌ایم. برای نیل به این هدف از تکنیک‌های پردازش تصویر و داده‌کاوی استفاده کردیم.

۲. پیشینه پژوهش

روش‌هایی که اخیراً برای تشخیص سرطان سینه به کمک کامپیوتر اجرا می‌شوند، سریعتر و دقیق‌ترند. عملکرد این روش‌ها با عملکرد پاتولوژیست‌ها در تشخیص، مقایسه و ارزیابی می‌گردد.

Singh و همکارانش در سال ۲۰۱۵ یک رویکرد ساده و

آسان برای تشخیص بافت‌های سرطانی در تصاویر ماموگرام معرفی کردند. رویکرد آن‌ها برای قطعه‌بندی ناحیه تومور در تصاویر ماموگرام استفاده از روش‌های پردازش تصویر پایه‌ای و ساده مانند میانگین‌گیری و آستانه‌یابی می‌باشد. آن‌ها برای تشخیص تومور از روش حداکثر میانگین و حداقل واریانس استفاده نمودند. نتایج اعمال روش پیشنهادی آن‌ها روی تصاویر ماموگرام و مقایسه آن با تحلیل عینی تصاویر توانایی قطعه‌بندی موفق ناحیه سرطانی را در این تصاویر نشان داد. در حین قطعه‌بندی، پیکسل‌های ناحیه سرطانی نیز شناسایی شدند.

به دلیل استفاده آن‌ها از روش‌های پایه پردازش تصویر، عملکرد این روش ساده و سریع می‌باشد. روش مذکور می‌تواند در سایر برنامه‌های کاربردی تصویربرداری پزشکی، نظیر استخراج ویژگی و انطباق الگو نیز کاربرد مؤثری داشته باشد. انتخاب دستی پارامترهای آستانه و اندازه فیلتر میانگین از نقاط ضعف روش پیشنهادی آن‌ها می‌باشد [۱]. Kaymak و همکارانش در سال ۲۰۱۷ روشی جهت تشخیص سرطان سینه با شبکه عصبی مصنوعی پیشنهاد دادند. آن‌ها طبقه‌بندی تصاویر را با استفاده از شبکه عصبی پس‌انتشار خطا یا انتشار رو به عقب (BPNN) و شبکه‌های عصبی پایه‌ای شعاعی (RBNN) انجام دادند و به ترتیب دقت تشخیص برابر با ۵۹٪ و ۷۰٫۴٪ را گزارش نمودند [۳].

Ramadan و همکارانش در سال ۲۰۲۰ مقاله‌ای مروری تحت

عنوان "روش‌های مورد استفاده برای تشخیص رایانه‌ای، برای تشخیص سرطان سینه با استفاده از ماموگرافی" ارائه دادند. داده‌های آماری آن‌ها در این مقاله نشان می‌داد که طبق پیش‌بینی انجمن سرطان آمریکا در سال ۲۰۱۹، حدود ۲۶۸۶۰۰ مورد جدید سرطان تهاجمی سینه در زنان، حدود ۶۲۹۳۰ مورد جدید غیرتهاجمی و حدود ۴۱۷۶۰ مورد مرگ ناشی از سرطان سینه در ایالات متحده وجود دارد. این مسئله رویکرد زیاد متخصصان برای استفاده از تصاویر سینه برای تشخیص تومور را در پی داشت که در گزارش شورای تحقیقات ملی آن کشور نیز نشان داده شد. یکی از راه‌های پاسخگویی به این تقاضا، توسعه سیستم‌های تشخیص به کمک رایانه CAD^۲ برای تشخیص سرطان سینه با استفاده از ماموگرافی است. این مقاله با بررسی پیشرفت‌ها اخیر در سیستم‌های CAD برای تشخیص سرطان سینه با استفاده از ماموگرافی پرداخته و یک نمای کلی از روش‌های استفاده‌شده در مراحل مختلف آن شامل مرحله پیش‌پردازش، مرحله بهینه‌سازی و مرحله پایانی، طبقه‌بندی تومور، ارائه می‌نماید. سطح فعلی عملکرد سیستم‌های CAD دلگرم‌کننده است اما برای تشخیص مستقل بدون استفاده از تشخیص بالینی کافی نیست. تا زمان ارتقاء عملکرد سیستم‌های CAD از سطح موجود فعلی با استفاده از روش‌های تشخیص الگو نظیر افزایش داده‌ها در فرآیند یادگیری عمیق و بهره‌برداری از پیشرفت

قدرت محاسباتی رایانه‌ها، سیستم‌های CAD نقش خود را به صورت یک سیستم کمکی تشخیص بالینی بیماری ادامه می‌دهند [۴].

Kumar و همکارانش در سال ۲۰۲۰ در تحقیقی به تشخیص خودکار بر اساس الگوریتم‌های یادگیری ماشین پرداختند. آن‌ها با بررسی مرگ و میر صدها هزار نفر از زنان با سرطان سینه به عنوان شایعترین نوع سرطان در میان زنان به بررسی اهمیت داده‌ها در این سرطان پرداخته‌اند. برای جلوگیری از این میزان مرگ و میر، تشخیص به موقع برای پیشگیری و درمان آن بسیار مهم است که داده‌کاوی می‌تواند نقش تعیین‌کننده‌ای برای آن ایفا نماید. آن‌ها از تکنیک‌های مختلف داده‌کاوی برای طبقه‌بندی، در پیش‌بینی نوع بدخیم و خوش‌خیم استفاده نمودند [۵].

اخیرا نیز از شبکه‌های عصبی عمیق و انواع آن برای تشخیص دقیق سرطان سینه استفاده شده است. Ragab و همکاران با ارائه چارچوبی مبتنی بر شبکه‌های عصبی پیچشی عمیق چندگانه Multi-DCNN، استخراج و طبقه‌بندی ویژگی‌های تومورها را با استفاده از تکنیک‌های یادگیری عمیق DL برای کمک به رادیولوژیست‌ها در طبقه‌بندی ضایعات سرطانی در ماموگرام ارائه نموده‌اند. در این تجزیه و تحلیل مؤلفه اصلی - PCA برای کاهش بردار ویژگی بزرگ تولید شده، در همجوشی ویژگی‌ها و کاهش هزینه محاسباتی معرفی شده است. این آزمایش‌ها بر روی دو مجموعه داده شامل: (۱) زیرمجموعه تصویربرداری سینه از پایگاه داده دیجیتال برای غربالگری ماموگرام‌ها (CBIS-DDSM) و (۲) پایگاه داده دیجیتال ماموگرافی جامعه تحلیل تصویر ماموگرافی (MIAS) انجام شده است. دقت حاصل با استفاده از همجوشی ویژگی‌های عمیق برای هر دو مجموعه داده در مقایسه با پیشرفته‌ترین سیستم‌های CAD بیشترین میزان را نشان داد. در مقابل هنگام استفاده از PCA در مجموعه‌های همجوشی ویژگی، دقت بهبود نیافت، با این حال، با کاهش زمان اجرا، هزینه محاسباتی کاهش یافت [۶].

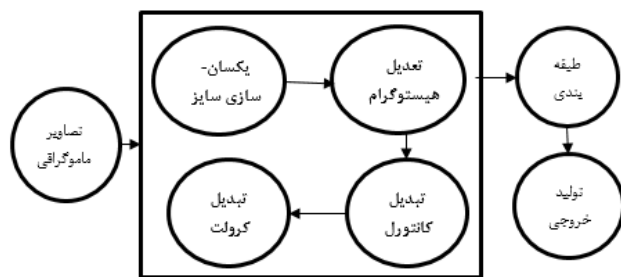
یادگیری عمیق با توجه به ساختار شبکه عمیق آن، علی‌رغم دقت بالا، نیاز به پردازنده‌های قوی و تنظیم پارامترهای بسیاری دارد که برای سیستمی با منابع محاسباتی محدود مانند سیستم‌های تعبیه‌شده مستقل چندان مناسب نمی‌باشد.

۳. مشکلات و چالش‌های تشخیص

به علت پیچیدگی بافت‌های سینه، شباهت بین توده‌های سرطانی با بافت‌های طبیعی پیرامون آن و استفاده از تابش اشعه ایکس با تراکم پایین در تهیه ماموگرام‌ها، تصاویر از کیفیت پایینی برخوردار هستند. در ضمن تومورها دارای اندازه‌ها و شکل‌های متفاوتی می‌باشند از این رو تشخیص ضایعات به خصوص در مراحل اولیه کار بسیار دشواری می‌باشد که ممکن است باعث عدم تشخیص صحیح و به موقع بافت‌های آلوده به سرطان شود به طوری که حتی گاهی اوقات یک رادیولوژیست یا پزشک

ماهر تفسیرهای گوناگونی را از توده‌های سرطانی و مکان دقیق آن ارائه می‌نماید. یکی از چالش‌های موجود در روش‌های تشخیص تصاویر ماموگرافی، مشکل تحلیل بافت‌های چگال است. به این ترتیب که ناحیه‌ای از سینه که اطراف توده است به رنگ سفید به نظر می‌رسد و بنابراین توده سرطانی و بافت پس‌زمینه آن به سختی از یکدیگر قابل-تفکیک‌اند. استفاده پزشک برای تشخیص توده سرطانی با استفاده از مشاهدات بصری منجر به بروز دو خطا می‌گردد. خطای اول زمانی رخ می‌دهد که یک تصویر ماموگرام در دو نوبت به یک پزشک نشان داده شود، ممکن است که تشخیص وی متفاوت باشد. خطای دوم ناشی از نشان دادن تصویر به دو پزشک متفاوت می‌باشد که تشخیص هر یک با دیگری مغایرت دارد. این عدم تشخیص عوامل متعددی دارد از جمله: کیفیت ضعیف تصویر، خستگی چشم و یا اشتباه پزشک در موقع تشخیص. بهبود تصاویر ماموگرافی فرآیند تغییر تصاویر به گونه‌ای است که مقدار کنتراست تصویر افزایش و مقدار نویز آن کاهش یابد تا رادیولوژیست بتواند به سهولت ناهنجاری‌ها را تشخیص دهد. بنابراین این تحقیق به دنبال روشی است که بتواند به وسیله آن دقت تشخیص نهایی را بهبود دهد. از این رو بعد از مرحله پیش‌پردازش که شامل تغییر مقیاس و تعدیل بافت‌نگار تصویر است از تبدیلات برجسته‌ساز در حوزه فرکانس نظیر کرولت و کانتورلت برای برجسته‌سازی و افزایش تمایز نواحی دارای توده برای تصمیم‌گیری استفاده می‌نماید.

روش‌های بهبود مرسوم، روش‌های با همسایگی ثابت هستند که به منظور بهبود تصاویر مبتنی بر ویژگی‌های سراسری به کار می‌روند. اما برای قطعه‌بندی تصویر از ویژگی‌های محلی مبتنی بر ناحیه بندی تصاویر استفاده می‌شود، همچنین این روش‌ها در افزایش تباین تصاویر ماموگرام با توجه به پیرامون آن‌ها نیز مورد استفاده قرار می‌گیرند. روش‌های بهبود مورد استفاده در این تحقیق از ویژگی‌های محلی مبتنی بر حوزه موجک استفاده- می‌نمایند. بلوک دیاگرام روش پیشنهادی مبتنی بر این ویژگی‌ها در شکل (۱) نشان داده شده است.



شکل ۱: بلوک دیاگرام روش پیشنهادی

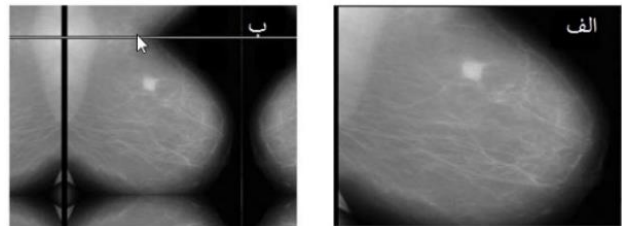
همانگونه که در شکل ۱ نشان داده شده، روش پیشنهادی شامل مراحل پیش‌پردازش و استخراج ویژگی از تصویر است که برای افزایش دقت طبقه‌بندی و تشخیص مورد استفاده قرار می‌گیرد.

۴. پیش پردازش و استخراج ویژگی های تصویر

پیش پردازش تصویر برای کمک به استخراج ویژگی برای بهبود مراحل طبقه بندی و تشخیص مورد استفاده قرار می گیرد.

۱.۴. تقارن و تغییر مقیاس تصویر

در ابتدا برای حذف آرتیفکت های حاشیه و لبه ها که در اثر استفاده از تبدیل موجک در مرحله پردازش اصلی به وجود می آیند، تصویر ورودی به صورت متقارن گسترش داده می شود. گسترش متقارن، پیوستگی را در قطعات کپی شده ورودی حفظ نموده و ضرایب بزرگ موجک ناشی از ناپیوستگی حاشیه را حذف می نماید. در گسترش متقارن فرض می شود که سیگنال مورد پردازش که در این جا سیگنال تصویر است در مرزها یا حاشیه تصویر متقارن است.



شکل ۲: (الف) تصویر اصلی از پایگاه داده MIAE (ب) تصویر

گسترش یافته در مرزها

۲.۴. بهبود تباین تصویر با تعدیل بافت نگار آن

برای این عمل از متعادل سازی بافت نگار تطبیقی AHE^v استفاده شده است. در این روش به وسیله بخش بندی تصویر به بلوک های کوچک، بهترین سطح آستانه برای هر بلوک به وسیله روش هایی نظیر HE تعیین می گردد، سپس می توان سطوح آستانه برای هر پیکسل را به وسیله درون یابی بین مراکز بلوک ها تعیین نمود. تکنیک های مختلفی از متعادل سازی هیستوگرام تصویر، برای بهبود کنتراست تصاویر گزارش شده است [۷-۱۰]. در این تکنیک با فرض مقادیر گسسته برای سطوح خاکستری تابع احتمال به صورت رابطه ۱ قابل تعریف است:

$$P_r(r_k) = \frac{n_k}{n} \quad 0 \leq r_k \leq 1 \quad k = 0.1 \dots L-1 \quad (1) \text{ رابطه ۱}$$

که در این رابطه r_k سطح خاکستری پیکسل هایی است که قرار است، بهبود یابند، و به ازای $r_k = 0$ سطح خاکستری سیاه و به ازای $r_k = 1$ سطح خاکستری سفید در نظر می شود، L تعداد سطوح خاکستری، $P_r(r_k)$ احتمال رخ دادن سطح خاکستری k ام و n_k تعداد دفعاتی است که این سطح خاکستری در تصویر رخ می دهد. نمودار $P_r(r_k)$ بر حسب r_k معمولاً هیستوگرام نامیده می شود و تکنیکی که برای ایجاد یک هیستوگرام متعادل به کار می رود، متعادل سازی هیستوگرام نامیده می شود [۱۱]. در این تکنیک هیستوگرام تجمعی H از سطوح خاکستری

G به عنوان قسمت اصلی تابع $F_{HE}(G)$ به کار می رود که سطوح خاکستری اصلی را به سطوح خاکستری تبدیل شده می نگارد:

$$F_{HE}(G) = G'_{min} + \Delta G' \frac{H(G)}{N} \quad \text{with } \Delta G' = G'_{max} - G'_{min}$$

رابطه (۲)

که در آن G'_{min} و G'_{max} بیانگر حدود بالا و پایین سطوح خاکستری تبدیل یافته اند و N بیانگر تعداد پیکسل هایی است که هیستوگرام بر روی آن اعمال می شود [۱۲].

تکنیک متعادل سازی هیستوگرام HE یک روش عمومی است که بر مبنای توزیع سطوح خاکستری در کل تصویر اعمال می شود و برای بهبود جزئیات تصویر مناسب نمی باشد. با اعمال متعادل سازی هیستوگرام به نواحی محلی تصویر، متعادل سازی هیستوگرام تطبیقی به دست می آید:

$$F_{AHE}(G) = G'_{min} + \Delta G' \frac{H_{AHE}(G)}{N_{AHE}} \quad (3) \text{ رابطه (۳)}$$

متعادل سازی هیستوگرام تطبیقی به عنوان ابزاری مؤثر در جهت بهبود کنتراست در پردازش تصاویر به کار برده می شود. از خواص منحصر به فرد این روش این است که به شیوه محلی عمل می کند در حالی که در روش های دیگر به صورت سراسری اعمال می شود در واقع در روش های قبلی یک آستانه گیری یا در کل تصویر انجام می شد و نسبت به آن بهبود یا متعادل سازی اعمال می شد. اما در روش AHE تصویر را به خانه های مانند کاشی تقسیم می کند و هر کاشی یا قسمت خود به ۹ خانه تقسیم می شود و برای متعادل سازی هر قسمت با مرکز خودش نقطه یابی شده و سنجش می شود و این مرکزیت لزوماً مرکز مربع نیست بلکه این نقطه مرکز حتی ممکن است در گوشه یک خانه تصویر باشد. تعیین نقطه مرکز به تفاوت کنتراست یا لبه تصویر بستگی دارد، هر کجا این تفاوت سایه روشن بیشتر باشد نقطه مرکز محسوب می شود و همین دلیل علاوه بر متعادل سازی مقداری نویز تصویر را زیاد می کند چون نویزها نیز خود باعث ایجاد لبه می شوند، جهت تکمیل مسئله و رفع عیب آن از روش CLAHE^a استفاده می شود.

در این روش تابع انتقال از طریق محور Y ها و حول مقدار تابع توزیع تجمعی CDF تباین محاسبه می شود. یکی دیگر از تفاوت های مهم این روش با روش های قبلی این است که مقدار آستانه قطع در محور Y ها انتخاب می شود. پس از تعیین نقطه آستانه در محور Y ها مقادیر پیکسل های تصویر تغییر داده می شوند. نقاطی که دارای فراوانی زیادی هستند که برای استخراج ویژگی در روش پیشنهادی به آن ها نیازی نیست به پس زمینه اضافه می شوند این مسئله باعث تغییر و جابجایی سطوح آستانه به مقادیر بیشتر می شود تا تقریباً به سطح مطلوب برسد. در این حالت آن چیز که در تصویر متعادل سازی شده باقی می ماند همان نقاط مطلوب و مورد نیاز برای قطعه بندی و استخراج ویژگی های مورد نیاز در روش پیشنهادی است؛ که باید بیشتر تقویت شوند.

از مهمترین گام‌های روش پیشنهادی، استخراج ویژگی‌های مناسب تشخیص است که بعد از انجام مراحل پیش‌پردازش تصویر، تغییر و همسان‌سازی سایز و تعدیل هیستوگرام تصویر، از تبدیلات حوزه فرکانس برای آن استفاده می‌شود. برای استخراج ویژگی بعد از یکسان‌سازی سایز و تعدیل بافت‌نگار، از تبدیل کانتورلت، تبدیل کرولت و تبدیل موج ویولت استفاده می‌شود.

۳،۴. تبدیل ویولت^۱ یا موجک

نا توانی تبدیل فوری در آنالیز سیگنال‌های با خصوصیات غیرایستایی منجر به ارائه تحلیل فرکانس‌زمان توسط گابور گردید که با پنجره‌های با اندازه ثابت محدود می‌شد. تبدیل موجک در واقع یک آنالیز پنجره‌ای در مقیاس متغیر است. در تبدیل موجک، موج کوچک، تابع پنجره دارای طول محدود است. در این تبدیل سیگنال بر روی یک دسته از توابع، موجک‌ها، تصویر می‌گردد. تابع مادر موجک در حوزه زمان محدود بوده که پس از چند ارتعاش به سرعت به سمت صفر میل می‌کند.

تئوری موجک برای غلبه بر مشکلات تبدیل فوری ارائه شده است. در این روش مسئله تقسیم سیگنال به بخش‌های مختلف با استفاده از مقیاس‌گذاری و انتقال یک تابع حل می‌شود. این تابع در طول سری اطلاعاتی انتقال می‌یابد و برای هر موقعیت آن، طیف سری اطلاعاتی محاسبه می‌شود. این مراحل برای توابعی با مقیاس‌های مختلف تکرار می‌شود و در نهایت نتیجه حاصل به صورت مجموع‌های از اطلاعات آرگومان -فرکانس به دست می‌آید. ویژگی اصلی تبدیل موجک در مقابل تبدیل فوری زمان کوتاه این است که تمامی توابع پایه از انتقال و مقیاس یک تابع (موجک مادر) به دست می‌آیند. شکل کلی تبدیل پیوسته ویولت به صورت رابطه ۱ است:

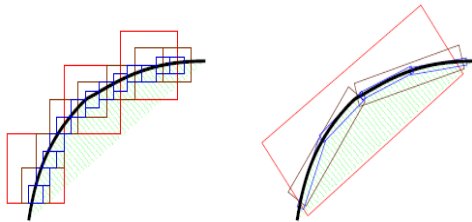
$$CWT_x^\psi = \frac{1}{\sqrt{s}} \int_{-\infty}^{+\infty} x(t) \psi\left(\frac{t-\tau}{s}\right) dt \quad \text{رابطه (۴)}$$

که در آن پارامترهای s و τ به ترتیب پارامترهای انتقال و مقیاس هستند انتقال یک ویولت به معنای تأخیر انداختن آن است و باعث می‌شود ویولت مورد نظر به سمت راست انتقال پیدا کرده و از اول تا انتهای سیگنال را طی کند. در تبدیل ویولت برخلاف تبدیل فوری به طور مستقیم پارامتر فرکانس وجود ندارد، در عوض پارامتر مقیاس وجود دارد که به طور معکوس با فرکانس در ارتباط است. مقیاس همان طور که از معنی آن برمی‌آید به عنوان یک عملگر ریاضی سیگنال را منبسط و منقبض می‌کند، مانند مفهوم مقیاس در نقشه، مقیاس‌های بالا متناظر با دید کلی و فارغ از جزئیات به سیگنال است (متناظر با فرکانس‌های پایین) و مقیاس‌های کوچک متناظر با نگاه به جزئیات سیگنال است و لذا در تناظر با فرکانس‌های بالا خواهد بود. تبدیل ویولت یک شباهت‌سنجی بین سیگنال و توابع پایه (ویولت‌ها) است و منظور از شباهت‌سنجی نیز شباهت بین محتوای فرکانسی است. نتیجه تبدیل ویولت یک ماتریس است که ستون‌های آن جایجایی در زمان را نشان می‌دهد و سطرهای آن مقیاس مورد بررسی را

نشان می‌دهد. به بیان دیگر ضرایب تبدیل ویولت بیانگر میزان نزدیکی سیگنال به ویولت در مقیاس مورد نظر می‌باشد. با انجام تبدیل موجک پیوسته ضرایبی به دست می‌آید که در مقیاس‌های مختلف با بخش‌های مختلف سیگنال تولید می‌شوند. [۱۳].

۴،۴. تبدیل کانتورلت

تبدیل‌های جدایی‌پذیر یک‌بعدی نظیر تبدیل تبدیل موجک در نمایش هندسی لبه‌های تصویر محدودیت دارند. برای برجسته‌کردن بهتر ویژگی‌های هندسی، به صورت کانتور، می‌توان از تبدیل کانتورلت استفاده نمود. شکل (۳) تفسیر مرزها با تبدیل موجک و تبدیل کانتورلت را با هم مقایسه می‌نماید.



شکل ۳: مقایسه تفسیر منحنی مرزی با تبدیل کانتورلت (راست) و تبدیل موجک (چپ)

تبدیل کانتورلت، تبدیل دو بعدی جهت‌داری است که برای توصیف منحنی‌ها و جزئیات ظریف در تصاویر به کار می‌رود. تبدیل کانتورلت از توابع پایه‌ای که در جهات مختلف به اشکال و مقیاس‌های مختلف گرایش دارند تشکیل شده است. با استفاده از مجموعه توابع پایه غنی تبدیل کانتورلت به طور کارآمد قابلیت توصیف کانتورهای هموار را که اجزای اصلی تصاویر طبیعی می‌باشند، فراهم می‌کند. تبدیل موجک همانند یک هرم لاپلاسین یک تبدیل چندمقیاسی است که برا تفسیر نقاط ناپیوسته تصویر و لبه‌های تصویر به کار می‌رود. اگر این تبدیل با یک فیلتربانک جهت‌دار که برای ارتباط نقاط ناپیوستگی به یک ساختار خطی به کار می‌رود، ترکیب شود یک بانک فیلتر جهت‌دار هر می به وجود می‌آید که به بسط حاصل از آن تبدیل کانتورلت می‌گویند. تبدیل کانتورلت از دو بخش عمده تشکیل شده است: تجزیه گر هرم لاپلاس و بانک‌های فیلتر جهت‌دار^۱ تصویر اصلی توسط هرم لاپلاس به دو تصویر تبدیل می‌شود: تصویر پایین‌گذر و تصویر میان‌گذر. هر تصویر میان‌گذر در مرحله بعد توسط یک فیلتر جهت‌دار تجزیه می‌گردد. با تکرار مراحل روی تصویر پایین‌گذر، تجزیه چندمقیاسی و چندجهتی از تصویر به دست می‌آید [۱۴].

۵،۴. تبدیل کرولت

تبدیل کرولت یک تبدیل چندمقیاسی و چندجهتی با توابع پایه سوزنی شکل و بسیار حساس به زاویه می‌باشد. توابع پایه تبدیل ویولت، ایزوتروپیک هستند و بنابراین برای بیان بخش منفرد منحنی‌ها به تعداد ضرایب زیادی نیاز دارند. کرولت‌ها مقیاس‌بندی سهمی شکل را نیز می‌پذیرند. به خاطر

این خواص، تبدیل کرولت، نمایش بهینه و کم حجمی از منحنی‌های منفرد ارائه می‌دهد. تبدیل کرولت در مقیاس‌ها و زوایای مختلف، برخلاف دیگر تبدیل‌های زاویه‌ای مثل ویولت و گابور، محدوده فرکانسی را به طور کامل پوشش می‌دهد [۱۵].

کرولت‌ها به مانند امواج تختی هستند که در فضای دو بعدی توسط دو متغیر جهت و فرکانس مرکزی مشخص می‌شوند. کرولت از اعمال اتساع^{۱۱}، چرخش و انتقال روی تابع Ψ به دست می‌آیند.

کرولت‌ها توسط متغیر مقیاس a که $0 < a(1)$ موقعیت b و چرخش θ ، اندیس‌گذاری می‌شوند و دارای رابطه زیر می‌باشند.

$$\Psi_{a,b,\theta}(x) = \alpha^3 \Psi(D_a R_\theta(x - b)) \quad \text{رابطه (۶)}$$

$$D_a = \begin{pmatrix} 1/\alpha & 0 \\ 0 & 1/\sqrt{\alpha} \end{pmatrix} \quad \text{رابطه (۷)}$$

که D_a ماتریس مقیاس‌گذاری سهمی و R_θ چرخش به اندازه θ رادیان است اگر تابع Ψ یک مربع واحد باشد، دیده می‌شود که $\Psi_{a,b,\theta}$ در محدوده‌ای قرار می‌گیرد که پوش آن یک مستطیل با ابعاد a در \sqrt{a} است که محور کوچک آن در امتداد θ جهت می‌گیرد. یکی از ویژگی‌های مهم کرولت‌ها این است که از قانون هارمونیک تبعیت می‌کنند که بیان می‌کند تحلیل و بازسازی یک تابع دلخواه $f(x_1, x_2)$ توسط برهم‌نهی امکانپذیر است. هر کدام از این کمیت‌های مقیاس، موقعیت و زاویه را می‌توان به طور مجزا به شکل زیر بیان کرد:

$$\begin{aligned} \alpha_j &= 2^{-j}, j = 0.1.2. \dots \\ \theta_{j,l} &= 2\pi l \cdot 2^{-|j/2|}, l = 0.1. \dots 2^{|j/2|} \\ b_k^{(j,l)} &= R_{\theta_{j,l}} \left(k_1 2^{-j}, k_2 2^{-\frac{j}{2}} \right), k_1, k_2 \in Z. \\ \Psi_{j,l,k} &= \Psi_{a_j, b_k^{(j,l)}, \theta_{j,l}} \end{aligned}$$

رابطه (۸)

رابطه $\Psi_{j,l,k}$ به صورت زیر بیان می‌شود:

$$\begin{aligned} f &= \sum_{j,l,k} \langle f, \Psi_{j,l,k} \rangle \Psi_{j,l,k} \\ \|f\|_{L^2}^2 &= \sum_{j,l,k} |\langle f, \Psi_{j,l,k} \rangle|^2 \end{aligned}$$

رابطه (۹)

خروجی تبدیل، مجموعه‌ای از ضرایب کرولت گسسته $C^D(j,l,k)$ است.

$$C^D(j,l,k) := \sum_{0 \leq t_1, t_2 < n} f[t_1, t_2] \overline{\varphi_{j,l,k}^D[t_1, t_2]} \quad (۱۰)$$

که $\Psi_{j,l,k}$ شکل موج گسسته می‌باشد [۱۷]. ساختار پیاده‌سازی

FDCT به شیوه پیچشی شامل مراحل زیر می‌باشد:

مرحله ۱- FFT دو بعدی بر روی تابع f اعمال شده و نمونه‌های

فوریه $-t/2 < t < t/2$ به دست می‌آید.

مرحله ۲- برای هر مقیاس j و هر زاویه l حاصل ضرب $\Psi_{j,l}$

$$f[t_1, t_2] \cdot \Psi_{j,l}[t_1, t_2]$$

مرحله ۳- این حاصل ضرب دور مبدا پیچشی (متناوب) می‌شود.

رابطه ۱۱ برقرار می‌شود:

$$\hat{f}_{j,l}[n_1, n_2] = \Psi(\hat{\Psi}_{j,l}, \hat{f})[t_1, t_2] \quad \text{رابطه (۱۱)}$$

که در بازه تغییرات t_1 و t_2 برابر است با $0 < t_2 < L_2$ و $0 < t_1 < L_1$ ، برای زمانی که در بازه $(-\pi/4, \pi/4)$ قرار داشته باشد.

مرحله ۴- معکوس FFT دو بعدی بر روی هر $f_{j,l}$ اعمال شده، بنابراین ضرایب گسسته $CD(j,l,k)$ به دست می‌آیند [۱۴].

ورودی مرحله استخراج ویژگی، تصویر نرمال SHC شده است. تبدیل کرولت بر روی تصویر ورودی اعمال می‌شود، که نتیجه تبدیل کرولت ضرایب کرولت می‌باشند. برای هر دسته از ضرایب کرولت، مقادیر میانگین و انحراف معیار یا خود ضرایب به تنهایی به عنوان ویژگی استفاده می‌شود. بردار ویژگی می‌تواند هر یک از این مقادیر یا ترکیب این دو کمیت یا ضرایب خود کرولت باشد. در تبدیل کرولت، هر تصویر را می‌توان به ۳، ۴ یا ۵ سطح مقیاس‌بندی کرد.

تعداد زیرباندها در هر یک از سطوح متفاوت است و برای هر ۳ سطح، تعداد زیر باندها ۱۸ یعنی $1+16+1$ (به ترتیب مربوط به سطوح ۱ و ۲ و ۳) می‌باشد. و برای ۴ سطح، تعداد زیرباند ۵۰ تا یعنی $1+16+32+1$ (به ترتیب مربوط به سطوح ۱، ۲، ۳ و ۴) است. ضرایبی که هر تبدیل کرولت در ۱۸۰ درجه تولید می‌کند، تکراری است. لذا برای سطوح ۲ و ۳ نصف زیرباندها کافی است. پس برای تبدیل کرولت ۴ سطحی، ۲۶ زیرباند از ضرایب کرولت تولید می‌شود که هر زیرباند مجموعه‌ای از ضرایب است.

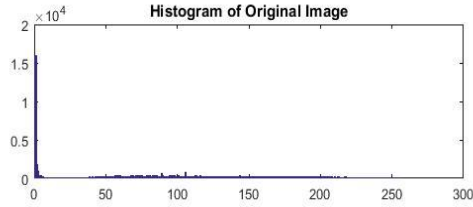
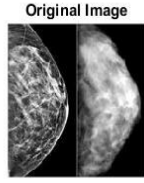
۵. نتایج

در این تحقیق روش‌ها و تبدیل‌های مذکور با نرم‌افزار Matlab و Weka پیاده‌سازی شدند و نتایج مراحل مختلف برای ارزیابی نمایش داده می‌شوند. قابل ذکر است که سیستم استفاده شده در ارزیابی دارای پردازنده Intel Pentium T4400 @ 2.2 GHz Processor، حافظه Ram 3.0 GB و سیستم عامل Windows 10 است.

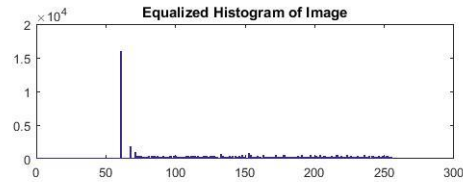
۵.۱. مجموعه داده

مجموعه تصاویر دیجیتال مورد استفاده، تصاویر ماموگرافی انجمن تحلیل تصاویر ماموگرافی MIAS است که یک منبع استاندارد برای استفاده جامعه پژوهشی در تحلیل تصاویر ماموگرافی است. این مجموعه داده در بیش از ۵۰۰ تحقیق استفاده شده است و شامل ۲۵۰۰ نمونه است. هر تصویر شامل دو تصویر از هر سینه، همراه با برخی از اطلاعات مربوط به بیمار (سن در زمان مطالعه، رتبه‌بندی تراکم سینه^{۱۲} ACR، براساس رادیوگرافی کالج آمریکایی، ناهنجاری‌ها، شرح کلمات کلیدی ACR از اختلالات) و اطلاعات تصویر (اسکنر، دقت تفکیک مکانی، ...) است. تصاویر حاوی مناطق مشکوک مربوط به اطلاعات مکانی مناطق و نوع مناطق در سطح پیکسل است که پیکسل کوچکترین جز تصویر در آن سطح است. در این تحقیق از ۱۵۰ نمونه تصادفی از ۵ دسته عادی، سرطان خوش خیم، سرطان بدخیم، میکروکلسیفیکیشن^{۱۳} خوش خیم و

میکروکلسیفیکیشن بدخیم استفاده شده است که در هر دسته به طور یکسان ۳۰ نمونه قرار دارد.



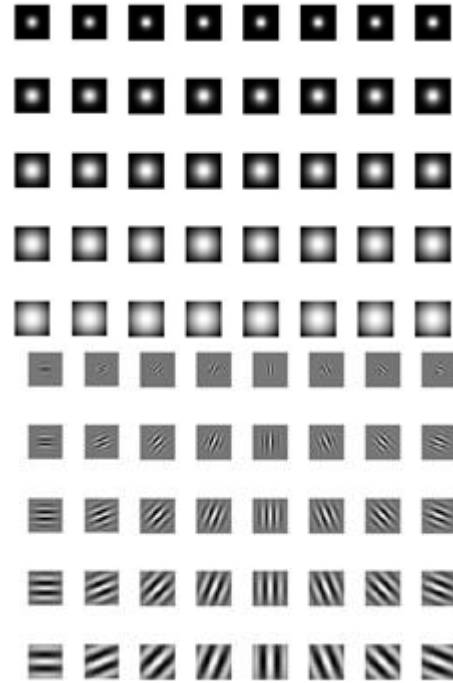
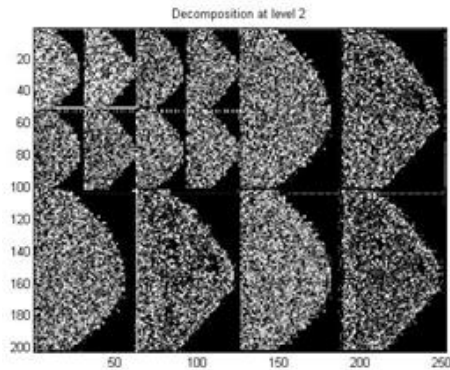
شکل ۵: نمودار هیستوگرام یک تصویر از مجموعه داده‌ها



شکل ۶: نمودار هیستوگرام تصویر شکل (۵) بعد از اعمال هم‌سازی

هیستوگرام تطبیق AHE

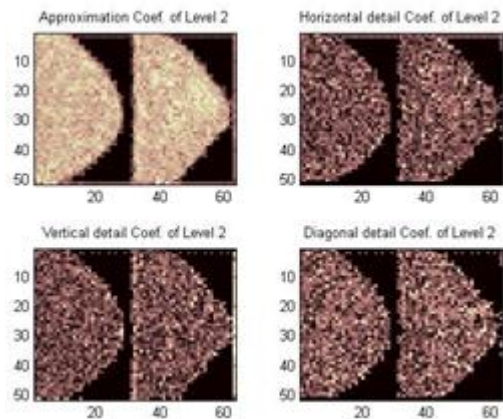
در مرحله آماده‌سازی داده‌ها بعد از اعمال تحلیل‌های مذکور بر روی تصاویر به استخراج ویژگی‌ها از تصاویر می‌پردازیم. با توجه به سطوح مختلف تبدیلات کرولت و کانتورلت و فیلترهای گابور مختلف و اعمال آن‌ها روی تصاویر رنگی و خاکستری به استخراج ویژگی‌ها جهت دسته‌بندی و مقایسه روی تصاویر ماموگرام مطابق تصاویر شکل (۷) الف و ب، می‌پردازیم. بطور کل ۹۵ ویژگی با استفاده از فیلترهای مذکور استخراج شده که به همراه نوع تصویر سرجمع ۹۶ ویژگی می‌شوند.



شکل ۴: نمونه‌ای از فیلترهای گابور استفاده شده در تصاویر

در گام اول روش پیشنهادی به استخراج ویژگی پرداخته می‌شود. برای استخراج ویژگی بعد از یکسان‌سازی سایز و تعدیل هیستوگرام، از تبدیل کانتورلت، تبدیل کرولت استفاده می‌شود. در فاز پیش‌پردازش از بانک فیلتر جهتدار گابور استفاده شده است. فیلتر گابور بر پایه فیلترهای تکی بوده که با توجه به انتخاب اولیه به بلوک‌های مجاور و هم‌اندازه هم اعمال می‌شود. نمونه‌ای از فیلترهای گابور اعمالی بصورت بلوک‌های ۵*۸ در شکل ۴ نشان داده شده است.

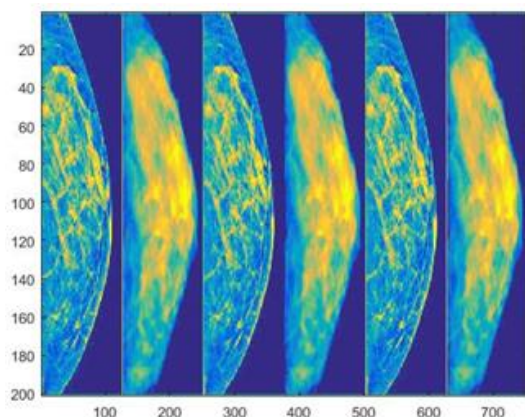
در ابتدا برای استخراج بهتر ویژگی‌ها فیلترهایی بر روی تصاویر اجرا شده است و این فیلترها سبب بهبود خروجی تصاویر می‌گردند. قبل از اعمال این فیلترها در بسیاری از موارد استخراج ویژگی‌ها می‌تواند دچار خطا شده و به این ترتیب داده‌های استخراجی مناسبی شکل‌نگیرد. تصاویر اولیه همراه هیستوگرام آن‌ها و تصاویر بعد از اعمال فیلترها در شکل ۵ و ۶ نشان داده شده‌اند.



ب) تراکم بافت

شکل ۸: استخراج ویژگی رنگی تصاویر

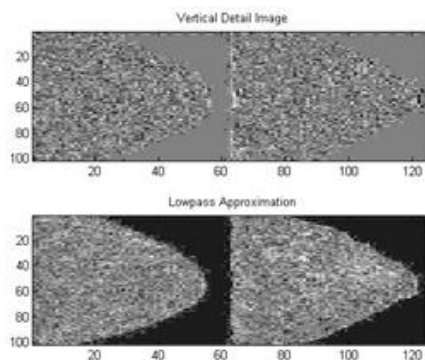
مصابق شکل ۸، برای هر تصویر رنگی نیز چهار ویژگی عمده با تبدیلات کرولت و کانتورلت استخراج و مورد استفاده قرار می‌گیرد. مطابق شکل ۸ الف اولین ویژگی ریخت‌شناسی، لبه‌یابی و شکل ظاهری است. برای ویژگی ریخت‌شناسی بعدی نیز تراکم بافت‌ها و میزان نزدیکی و چسبندگی اجزا آن‌ها به یکدیگر و به اصطلاح حجم بافت استخراج می‌شود. به عنوان ویژگی‌های بعدی ریخت‌شناسی نیز شکل بافت و میزان توده موجود در سمت چپ و راست عکس بررسی و استخراج می‌شود. فاز نهایی تشخیص، انتخاب ویژگی و در نظر گرفتن شکل‌های گوناگون ویژگی می‌باشد که به عنوان ورودی دسته‌بند و تصمیم‌گیری نهایی استفاده می‌شود. نمونه‌ای از استخراج ویژگی‌های متمایز برای تشخیص مناسب در شکل ۹ نشان داده شده است.



شکل ۹: طبقه‌بندی و دسته‌بندی تصاویر

مطابق شکل ۹، ویژگی‌های متمایز موجب ایجاد تصاویر متمایز برای طبقه‌بندی و در نتیجه کمک به دسته‌بندی درست و تشخیص نهایی تصاویر می‌شود.

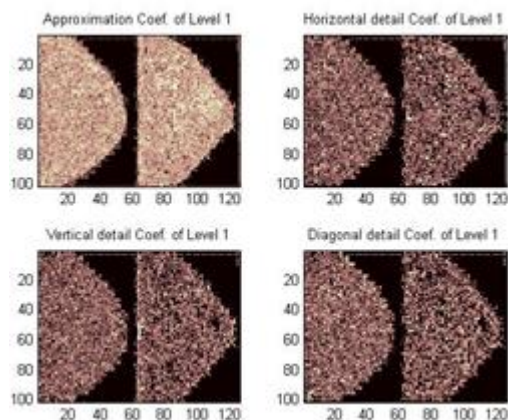
الف) استخراج با سطح ۴



ب) استخراج با سطح ۲

شکل ۷: استخراج ویژگی از روی سطوح خاکستری تصویر

در تصاویر ۷ و ۸ نمونه‌ای از ویژگی‌های برداری سطح‌بندی‌های مختلف تبدیلات کرولت و کانتورلت با اعمال فیلتر گابور استخراج شده و برای یکی از تصاویر مجموعه داده نشان داده شده است. همانگونه که در شکل ۷ الف می‌بینیم برای هر تصویر ویژگی‌های مختلف یک سطح تفکیک استخراج می‌گردند. همچنین تغییر سطوح هر تصویر از نمونه ابتدایی و تغییر شکل آن نسبت به تغییر سطح تفکیک همانند شکل ۷ ب استخراج شده و بردارهای مربوطه به عنوان ویژگی‌های نهایی تصاویر برای دسته‌بندی در روش پیشنهادی مورد استفاده قرار می‌گیرند.



الف) لبه‌یابی

۲.۵. ارزیابی شبیه‌سازی

بعد از پیاده‌سازی روش پیشنهادی نوبت به ارزیابی این روش و مقایسه- عملکرد آن با روش‌های موجود می‌رسد. برای مقایسه، عملکرد روش پیشنهادی را با طبقه‌بندی‌های پرکاربرد که در بیشتر مقالات ذکر شده، مورد ارزیابی قرار می‌دهیم. معیارهای مورد استفاده برای ارزیابی عبارتند از:

صحت^{۱۴} تشخیص: صحت یا درستی تشخیص، عبارت از میزان تطابق پیش‌بینی‌های یک مدل می‌باشد که با واقعیت مدل‌سازی شده هم‌خوانی دارد. صحت طبقه‌بندی بر روی داده آزمایشی، از تقسیم تعداد اشیائی که درست طبقه‌بندی شده‌اند به تعداد کل اشیاء مورد طبقه‌بندی طبق رابطه (۱۲) حاصل می‌شود.

$$\text{Accuracy} = \frac{TN+TP}{TN+FN+TP+FP} \quad \text{رابطه (۱۲)}$$

دقت و بازخوانی^{۱۵}: دقت و بازخوانی معیارهای کاربردی در حوزه بازیابی اطلاعات هستند که میزان تناسب اسناد بازیابی شده توسط سیستم را با

جدول ۱: ماتریس پراکندگی و مؤلفه‌های آن

منفی	مثبت		
Rate FN^{۱۷}: بیانگر تعداد رکوردهایی است که دسته واقعی آن‌ها مثبت بوده و الگوریتم طبقه‌بندی دسته آن‌ها را به‌درستی مثبت تشخیص داده‌است.	Rate TP^{۱۶}: بیانگر تعداد رکوردهایی است که دسته واقعی آن‌ها مثبت بوده و الگوریتم طبقه‌بندی نیز دسته آن‌ها را به‌درستی مثبت تشخیص داده‌است.	مثبت	کلاس واقعی
Rate TN^{۱۹}: بیانگر تعداد رکوردهایی است که دسته واقعی آن‌ها منفی بوده و الگوریتم طبقه‌بندی نیز دسته آن‌ها را به‌درستی منفی تشخیص داده‌است.	Rate FP^{۱۸}: بیانگر تعداد رکوردهایی است که دسته واقعی آن‌ها منفی بوده و الگوریتم طبقه‌بندی دسته آن‌ها را به‌اشتباه مثبت تشخیص داده‌است.	منفی	

نیاز کاربر تعیین می‌کنند. این معیارها به صورت زیر تعریف می‌شوند:

$$\text{Recall} = \frac{TP}{FN+TP} \quad \text{(رابطه ۱۳)}$$

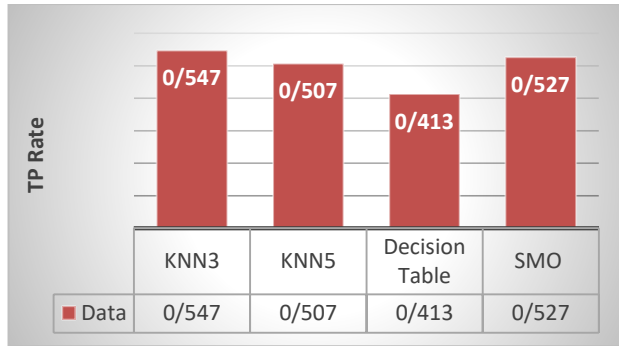
$$\text{Precision} = \frac{TP}{FP+TP} \quad \text{(رابطه ۱۴)}$$

به جای این دو معیار، می‌توان از یک معیار ترکیبی برای ارزیابی کارایی بازیابی به نام F-measure استفاده نمود. پارامتر F-Measure از ترکیب

Recall و Precision طبق رابطه ۱۵ بدست می‌آید. [۱۶، ۱۷]

$$F - \text{Measure} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad \text{(رابطه ۱۵)}$$

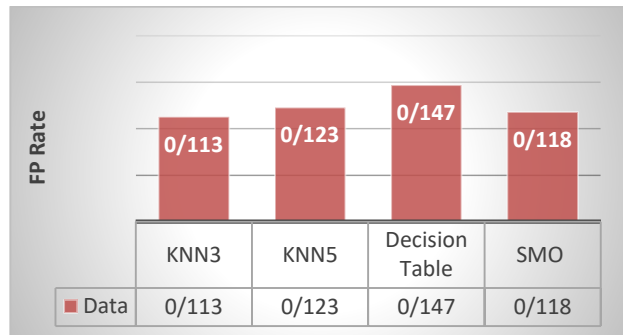
با استفاده از معیارهای مذکور عملکرد روش پیشنهادی را با روش‌های طبقه‌بند K نزدیک‌ترین همسایگی با تعداد همسایه‌های ۳ و ۵ جدول تصمیم و شبکه بردار پشتیبان SMO انجام می‌دهیم.



شکل ۹: ارزیابی روش‌های مختلف با استخراج داده پیشنهادی با معیار TP در شکل ۱۰ ارزیابی روش‌ها با معیار FP نشان داده شده‌است. در شکل ۹ بیشترین تشخیص مقدار FP و ارزیابی روش‌ها با معیار TP در مجموعه‌های داده مختلف آمده‌است. با توجه به نحوه استخراج داده ارائه شده روش K نزدیکترین همسایه با مقدار ۳ در مجموعه Data و روش SMO بهترین نتیجه را داشته‌است. با توجه به نحوه استخراج ویژگی و انتخاب ویژگی معرفی شده روش درخت تصمیم پایین‌ترین نتایج را از خود نشان داده‌است.

روش درخت تصمیم بالاترین نرخ تشخیص این مقدار را دارا است.

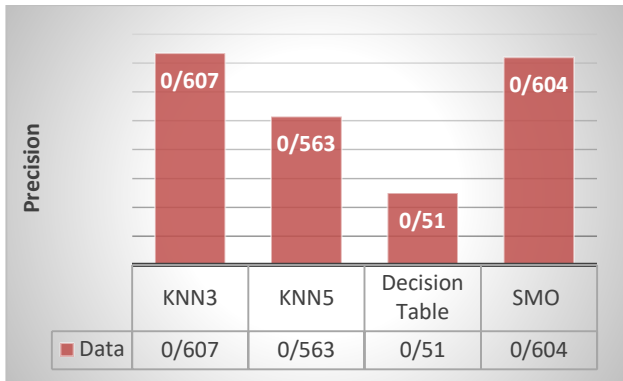
روش‌های دیگر نیز در رتبه‌های بعدی قرار دارند.



شکل ۱۰: ارزیابی روش‌ها مختلف با استخراج داده پیشنهادی با معیار FP

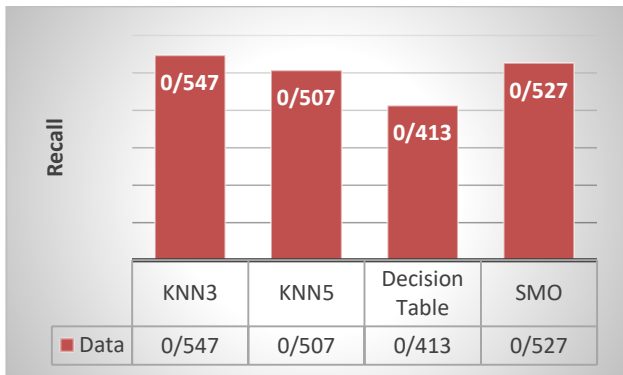
طبق شکل ۱۰، روش K نزدیکترین همسایه با مقدار K برابر ۷ بهترین نتیجه را در معیار صحت داشته و بعد از آن روش K نزدیکترین همسایه با مقدار K برابر ۳ و سپس روش SMO مقدار مناسبی را از خود نشان داده‌است. با توجه به نحوه استخراج ویژگی و انتخاب ویژگی معرفی شده روش درخت تصمیم پایین‌ترین نتایج را از خود نشان داده‌است. البته روش‌های نزدیکترین همسایه با K برابر ۵ نیز صحت قابل‌قبولی دارد.

در شکل ۱۲ ارزیابی روش براساس معیار Precision نشان داده شده است.



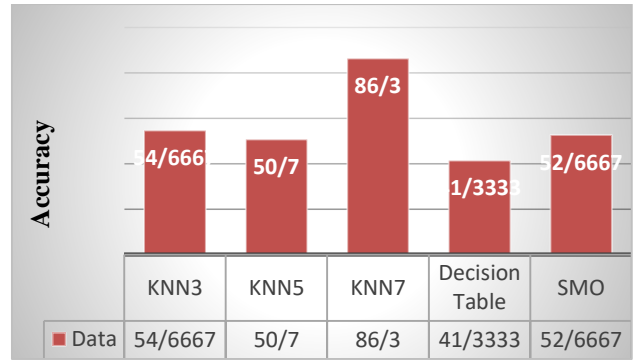
شکل ۱۲: ارزیابی روش‌های مختلف استخراج داده پیشنهادی با معیار Precision

طبق شکل ۱۲، روش K نزدیکترین همسایه با مقدار K برابر ۳ بهترین نتیجه و بعد از آن روش SMO صحت مناسبی داشته است. با توجه به نحوه استخراج ویژگی و انتخاب ویژگی معرفی شده روش درخت تصمیم پایین‌ترین نتایج را از خود نشان داده است. البته روش‌های نزدیکترین همسایه با K برابر ۵ نیز دقت قابل قبولی دارد.



شکل ۱۳: ارزیابی روش‌ها مختلف با استخراج داده پیشنهادی با معیار Recall

در شکل ۱۲ ارزیابی روش‌ها با معیار Recall نشان داده شده است. بیشترین تشخیص مقدار Recall همانگونه که در شکل ۱۲ نشان داده شده است در مجموعه داده مربوط به روش K نزدیکترین همسایه با K برابر ۳ بالاترین نرخ تشخیص این مقدار را دارا است. روش‌های دیگر نیز



شکل ۱۱: ارزیابی روش‌های مختلف استخراج داده پیشنهادی با معیار

صحت

همچنین در جدول ۲ مقایسه‌ای بین روش پیشنهادی و صحت کارهای مرتبط انجام شده است و کاملاً مشخص است که روش پیشنهادی به نسبت سایر روش‌ها صحت بهتری را نتیجه داده است.

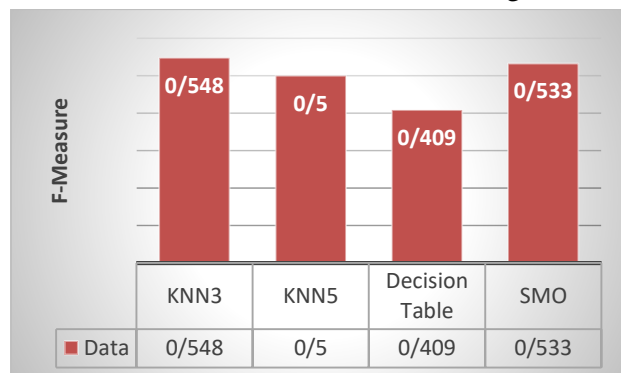
جدول ۲: مقایسه روش‌های طبقه‌بندی از نظر صحت

ردیف	روش	صحت به درصد
۱	بردار پشتیبان تصمیم	۵۲
۲	KNN	۸۲٫۱
۳	روش‌های آماری و شبکه عصبی	۸۵
۴	KNN با K برابر ۷ با داده‌های استخراج شده روش پیشنهادی	۸۶٫۳

مراجع

- [1] Singh, Anuj Kumar, and Bhupendra Gupta. "A novel approach for breast cancer detection and segmentation in a mammogram." *Procedia Computer Science* 54 (2015): 676-682
- [2] Gayathri, B. M., C. P. Sumathi, and T. Santhanam. "Breast cancer diagnosis using machine learning algorithms-a survey." *International Journal of Distributed and Parallel Systems* 4, no. 3 (2013): 105.
- [3] Kaymak, Sertan, Abdulkader Helwan, and Dilber Uzun. "Breast cancer image classification using artificial neural networks." *Procedia computer science* 120 (2017): 126-131.
- [4] Ramadan, Saleem Z. "Methods used in computer-aided diagnosis for breast cancer detection using mammograms: a review." *Journal of healthcare engineering* 2020 (2020).
- [5] Kumar, Vivek, Brojo Kishore Mishra, Manuel Mazzara, Dang NH Thanh, and Abhishek Verma. "Prediction of malignant and benign breast cancer: A data mining approach in healthcare applications." In *Advances in data science and management*, pp. 435-442. Springer, Singapore, 2020.
- [6] Ragab, Dina A., Omneya Attallah, Maha Sharkas, Jinchang cancer classification using multi-DCNNs." *Computers in Biology and Medicine* 131 (2021): 104245.
- [10] Lee, Tai Sing. "Image representation using 2D Gabor wavelets." *IEEE Transactions on pattern analysis and machine intelligence* 18, no. 10 (1996): 959-971.
- [7] da Silva, R.R., Caloba, L.P., Siqueira, M.H.S. and Rebello, J.M.A.(2004). *Pattern Recognition of Weld Defects Detected by Radiographic Test*. NDT & E International, 37:461-470.
- [8] Wanga, X. and Wong, B.S. (2004). *Image Enhancement for Radiography Inspection*. The Third international Conferenceon Experimental Mechanics, SPIE Proceedings, 5852.
- [9] Wang, G. and Liao, T.W.(2002). *Automatic Identification of Different Types of Welding Defects in Radiographic Images*. NDT & E International,35:519-528.
- [10] Shafeek, H.I., Gadelmawla, E.S., Abdel-Shafy, A.A. and Elewa, I.M. (2004). *Assessment of Welding Defects in Pipeline Radiographs Using Computer Vision*. NDT & EInternational,37:291-299.
- [11] Gonzalez, R.C. and Wintz, P.(1987). *Digital Image Processing*. Addison-Wesley, 2nd Edition.
- [12] Castleman, K.R.(1979). *Digital Image Processing*. Englewood Cliffs N.J., Prentice Hall, 90-94.
- [13] Addison, Paul S. "Introduction to redundancy rules: the continuous wavelet transform comes of age." (2018): 20170258.
- [14] Candes, E. J., and Donoho, D. L, "Continuous curvelet transform", II. Discretization and frames: Applied and Computational Harmonic Analysis, 19, 198-222, 2003.
- [15] Werner Ebeling, H. M., "Evolutionary Algorithms And Optimization", Berlin: Eingereicht An Der Mathematisch-Naturwissenschaftlichen Fakultät I Humboldt-Universität At Zu Berlin , Dezember 2002.
- [16] Rachmawanto, Eko Hari, Galang Rambu Anarqi, and Christy Atika Sari. "Handwriting Recognition Using Eccentricity and Metric Feature Extraction Based on K-Nearest Neighbors." In 2018 International Seminar on Application

در رتبه‌های بعدی قرار دارند. در شکل ۱۳ ارزیابی روش‌ها با معیار F-Measure نشان داده شده است.



شکل ۱۴: ارزیابی روش‌ها مختلف با استخراج داده پیشنهادی با معیار F-Measure

طبق شکل ۱۴، بهترین F-Measure که بالاترین نتیجه در آن به ثبت رسیده مربوط به روش K نزدیکترین همسایه با K برابر ۳ است و در ادامه روش SMO است.

۶. نتیجه‌گیری و کارهای آتی

در این مقاله یک روش استخراج ویژگی از تصاویر ماموگرافی دریافتی از بیمار ارائه شده است. تصویر ورودی به سیستم طی فرآیند استخراج ویژگی قرار گرفته و سه دسته ویژگی اصلی فرکانسی، ریخت‌شناسی و بافت‌شناسی، از آن استخراج می‌شود. این فرآیند طی چرخه، یکسان‌سازی سایز، تعدیل هیستوگرام، تبدیل کانتورلت و تبدیل کرولت انجام می‌شود. به دلیل حساس بودن سیستم‌ها سعی در استخراج مناسب ویژگی‌ها با انواع سطوح و ماتریس‌های شناخته شده مانند ماتریس رخداده‌ها و سطوح خاکستری شده است. بعد از استخراج ویژگی به جهت بهبود صحت خروجی انتهایی از روش‌های آماری برای انتخاب ویژگی استفاده شد و زی مجموعه‌ای از ویژگی‌های استخراج شده برای ورودی طبقه‌بند انتخاب شدند. بعد از این در گام بعد خروجی روش پیشنهادی در گام بعد برای مقایسه عملکرد روش پیشنهادی طبقه‌بندهای مختلف استفاده و خروجی آن‌ها با یکدیگر مقایسه شد. بعد از اعمال ویژگی‌های استخراج شده به ارزیابی روش گفته شده توسط نتایج طبقه‌بندها پرداخته شد. بهترین نتایج بر روی مجموعه داده روش پیشنهادی بوده است که به میزان صحت ۸۶٫۳ رسیده و بر روی مجموعه داده مذکور بهبود مناسبی از خود نشان داده است. علاوه بر معیار صحت، روش مذکور با روش‌های موجود با معیارهای دیگری نیز مورد مقایسه قرار گرفت که بین آن‌ها روش نزدیکترین همسایگی دارای دقت بیشتری است و در مرحله بعد طبقه SMO با بیشترین دقت و حساسیت قرارداد. در کارهای آینده می‌توان با افزایش تعداد زاویه‌ها در فیلتر کانتورلت لبه‌ها را با دقت بیشتری از تصاویر استخراج کرد.

پی‌نوشت

- ¹⁰ Directional Filter Bank
- ¹¹ Dilation
- ¹² American College of Radiology
- ¹³ Microcalcifications
- ¹⁴ Accuracy
- ¹⁵ Precision and Recall
- ¹⁶ True Positive
- ¹⁷ False Negative
- ¹⁸ False Positive
- ¹⁹ True Negative

- ¹ Back Propagation Neural Network
- ² Radial Basis neural network
- ³ Computer-Aided Diagnosis
- ⁴ Deep convolutional neural networks
- ⁵ Curated Breast Imaging Subset of the Digital Database for Screening Mammography
- ⁶ mammographic image analysis society
- ⁷ Adaptive Histogram Equalization (AHE)
- ⁸ Contrast Limited Adaptive Histogram Equalization (CLAHE)
- ⁹ Wavelet Transform