

ارائه روشی جهت بهبود کیفیت تصاویر ماموگرافی مبتنی بر ترکیب تکنیک‌های کانتورلت و کرولت

حسن نجفی^{۱*}

^۱ دانشجوی کارشناسی ارشد، رشته برق، گرایش الکترونیک دیجیتال، دانشگاه آزاد اسلامی واحد زنجان، ایران.

چکیده

با توجه به اینکه سرطان سینه تبدیل به امری معمول و رایج شده، تشخیص زود هنگام کامپیوتری جهت کاهش میزان مرگ و میر به امری ضروری و چالش بر انگیز تبدیل شده است. از این رو در این پژوهش روشی مبتنی بر تکنیک‌های پردازش تصویر جهت بهبود کیفیت تصاویر ماموگرافی ارائه شده است. این پژوهش یک سیستم پزشکیار است که دارای دو مرحله پیش‌پردازش (استفاده از یکسان‌سازی سایز تصاویر و تعدیل هیستوگرام) و استخراج ویژگی (استفاده از تبدیل‌های کانتورلت و کرولت در تصاویر ماموگرافی دریافتی بیمار که شامل سه دسته ویژگی اصلی ریخت‌شناسی و بافت‌شناسی، آماری و دسته آخر فرکانسی) جهت بهبود ارائه می‌دهد و سبب افزایش صحت تشخیصی می‌شود. برای شبیه‌سازی روش پیشنهادی مجموعه داده تصاویر دیجیتال ماموگرافی غربالگری MIAS استفاده شده و زیر مجموعه ویژگی‌های استخراج شده برای ورودی طبقه‌بند انتخاب و در نهایت جهت ارزیابی روش پیشنهادی از طبقه‌بندها و معیارهای مناسب استفاده می‌شود. در بخش آخر شبیه‌سازی روش پیشنهادی مبتنی بر طبقه‌بندهای مختلف مورد ارزیابی قرار گرفت. بهترین نتیجه روی مجموعه داده، مربوط به روش پیشنهادی بود که میزان صحت ۸۶.۳٪ از خود نشان داد و به نسبت بقیه روش‌ها نتیجه مطلوب تری داشته است.

کلمات کلیدی: تصاویر ماموگرافی، تشخیص سرطان، پردازش تصویر، بهبود کیفیت، تعدیل هیستوگرام، استخراج ویژگی، تبدیل‌های کانتورلت و کرولت.

تاریخچه مقاله:

تاریخ ارسال: ۱۳۹۹/۰۸/۱۵

تاریخ اصلاحات: ۱۳۹۹/۰۹/۲۵

تاریخ پذیرش: ۱۳۹۹/۱۲/۲۰

تاریخ انتشار: ۱۳۹۹/۱۲/۲۹

Keywords:

Mammography Images,
Cancer Diagnosis,
Image Processing,
Quality Improvement,
Histogram Equalization,
Feature Extraction,
Contourlet and Curvelet
Transformations

*ایمیل نویسنده مسئول:

najafe@gmail.com

Provide an approach the Quality of Mammography Images Based on a Combination of Contourlet and Curvelet Techniques

Hassan Najafy^{*1}

¹Graduated Masters Department Electrical engineering, Islamic Azad university of Zanjan Branch, Iran.

Abstract

Given that breast cancer has become commonplace, Early computer detection to reduce mortality has become a necessity and a challenge. In this study, a method based on image processing techniques to improve the quality of mammographic images is presented. This research is a medical system that has two stages of preprocessing (Histogram equalization and Size equalization image) and feature extraction (The use of contourlet and Curvelet transformations in mammographic images received from patients include three main categories of morphological and histological features, statistical and frequency.) for improvement and increases diagnostic accuracy. To simulate the proposed method, the MIAS digital mammography screening digital image dataset was used and the extracted feature subset is selected for the classifier input, Finally, appropriate classifications and criteria are used to evaluate the proposed method. In the last part of the simulation, the proposed method based on different classifications was evaluated. The best result on the data set was related to the proposed method. The accuracy of the proposed method was 86.3 and it had better results than other methods.

است و به این سوال پاسخ می‌دهد که آیا ارائه مدلی جهت بهبود کیفیت تصاویر ماموگرافی مبتنی بر تکنیک‌های پردازش تصویر و داده‌کاوی امکان‌پذیر است؟

۲ - کارهای مرتبط

کایماکا و همکارانش در سال ۲۰۱۷ روشی جهت تشخیص سرطان سینه با شبکه عصبی مصنوعی پیشنهاد دادند. آنها اظهار داشتند که سرطان سینه بیشتر در ناحیه‌ی سینه‌ی زنان ایجاد می‌شود. معاینه‌های منظم برای تشخیص و درمان زود هنگام این نوع از سرطان بسیار حیاتی است. پاتولوژیست تشخیص سرطان را انجام می‌دهد. روش‌های اخیر که برای تشخیص سرطان سینه با کمک کامپیوتر ارائه شده‌اند، نوع دیگر و سریع‌تری از تشخیص سرطان سینه را ممکن می‌سازند. بنابراین، روش‌های کامپیوتری برای تشخیص سرطان سینه در حال توسعه و بهبود هستند. در پژوهش آنها، روشی برای دسته‌بندی خودکار تصاویر برای تشخیص سرطان سینه ارائه می‌شود. دسته‌بندی تصاویر با استفاده از شبکه عصبی پس انتشار خطا^۱ یا انتشار رو به عقب (BPNN) انجام می‌شود. عملکرد دسته‌بندی خودکار تصاویر سرطان سینه با استفاده از شبکه‌های عصبی شعاع مینا^۲ (RBNN) نیز بهبود بیشتری یافته است. دقت روش‌های BPNN و RPNN به ترتیب برابر با ۵۹ درصد و ۷۰/۴ درصد گزارش شده است [۳].

سینگ و همکارانش در سال ۲۰۱۵ در مقاله خود یک روش جدید برای تشخیص سرطان سینه و قطعه بندی در تصویر ماموگرام پیشنهاد دادند آنها یک رویکرد ساده و آسان برای تشخیص بافت‌های سرطانی در تصاویر ماموگرام معرفی کردند. پس از فاز تشخیص، قطعه‌بندی ناحیه تومور در تصویر ماموگرام انجام می‌شود. رویکرد آنها از روش‌های پردازش تصویر ساده مانند میانگین‌گیری و آستانه‌یابی استفاده می‌کند. برای تشخیص تومور، روش حداکثر میانگین و حداقل واریانس معرفی شد و نتایج آزمایشی، اثربخشی و کارایی روش آنها را نشان داد. نتیجه کار آنها به این ترتیب بود که روش پیشنهادی آنها روی دو تصویر ماموگرام اجرا شد. با استفاده از تحلیل عینی، به خوبی نتیجه شد که روش آنها می‌تواند به صورت کامل و موفقیت آمیز، ناحیه سرطانی را در تصاویر ماموگرام قطعه‌بندی کند. در حین قطعه‌بندی، پیکسل‌های ناحیه سرطانی نیز شناسایی شدند. روش آنها بسیار ساده و سریع است، زیرا از روش‌های پایه پردازش تصویر در آن استفاده شده است. همچنین روش آنها می‌تواند در سایر برنامه‌های کاربردی تصویربرداری پزشکی، انطباق الگو استخراج ویژگی نیز کاربرد مؤثری داشته باشد.

سرطان سینه یکی از مهم‌ترین دلایل مرگ و میر در بین زنان در سراسر جهان است. مشاهده شده است که تشخیص زود هنگام بدخیم بودن می‌تواند به آشکارسازی بیماری کمک کرده و امید به بقا و نجات از بیماری را به طور چشمگیری افزایش دهد [۱]. از جاییکه سرطان سینه تبدیل به رایج‌ترین بیماری در بین زنان شده است که منجر به مرگ می‌شود. سرطان سینه می‌تواند با استفاده از کلاس‌بندی تومورها تشخیص داده شود. دو نوع مختلف از تومورها، تومورهای خوش خیم و تومورهای بدخیم وجود دارند. پزشکان نیازمند یک پروسه تشخیص قابل اطمینان برای تمایز قائل شدن بین این تومورها هستند. اما به طور کلی، تشخیص و جداسازی تومورها حتی توسط افراد مجرب نیز کار بسیار دشواری به حساب می‌آید. از این رو، جهت تشخیص تومورها، نیاز به یک سیستم تشخیص به صورت خودکار است. بسیاری از پژوهشگران تلاش داشته‌اند تا از طریق الگوریتم‌های یادگیری ماشین، راهی را برای تشخیص نجات انسان از سرطان‌های مختلف پیدا کنند. همچنین پژوهشگران ثابت کرده‌اند که این الگوریتم‌ها، کارایی و عملکرد بهتری را در تشخیص و آشکارسازی سرطان از خود نشان داده‌اند [۲].

روش‌های مختلفی برای تشخیص سرطان سینه تاکنون بکار گرفته شده‌اند که در بین آنها، ماموگرافی به‌عنوان امیدبخش‌ترین روش محسوب شده و به کرات توسط پزشکان مورد استفاده قرار می‌گیرد. تصاویر ماموگرام معمولاً نویزدار و دارای کنتراست پایین هستند. در ماموگرافی سینه، نواحی روشن نشان دهنده سرطان هستند. در برخی از تصاویر ماموگرام، ممکن است هم بافت‌های بدخیم و هم بافت‌های توده‌ای نرمال مشاهده شوند. تمایز قائل شدن بین بافت‌های توده‌ای بدخیم و نرمال، تنها از طریق آستانه‌یابی امکان‌پذیر نیست. به‌دست آوردن اطلاعات نواحی توده‌ای ضایعات سرطانی در یک تصویر ماموگرام، امری ضروری به شمار 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 شده است. تا کنون 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 هیچ راه حلی که به صورت کامل امیدوار کننده باشد و یا بتواند فقط معیارهای مربوط به ضایعات سرطانی را به صورت کامل تشخیص دهد، به وجود نیامده است [۱]. از این رو در این پژوهش یک الگوریتم نوین پزشکیار جهت تشخیص سرطان سینه مبتنی بر بهبود کیفیت تصاویر ماموگرافی ارائه شده

²Radial Basis Neural Networks

¹Back Propagation Neural Network

خیمپاریا و همکارانش در سال (۲۰۲۱) مقاله ای تحت عنوان "تشخیص سرطان سینه بر اساس ماموگرافی مدرن با استفاده از یادگیری انتقال ترکیبی" ارائه دادند. آنها معتقدند که سرطان سینه یک سرطان شایع در زنان است. تشخیص به موقع سرطان سینه به طور خاص و به طور کلی، می تواند میزان بقای زنان را به میزان قابل توجهی افزایش دهد و می تواند بسیار موثرتر باشد. این مقاله عمدتاً بر روی فرآیند یادگیری انتقال برای تشخیص سرطان سینه تمرکز دارد. ¹ VGG¹ اصلاح شده (MVGG²) در مجموعه داده های تصاویر دو بعدی و سه بعدی ماموگرافی پیشنهاد و اجرا می شود. نتایج تجربی نشان داد که مدل پیشنهادی یادگیری انتقال ترکیبی (تلفیقی از MVGG و ImageNet) صحت ۹۴/۳ درصد را فراهم می کند همچنین از طرف دیگر، فقط معماری پیشنهادی دقت MVGG ۸۹/۸ درصد را ارائه می دهد. بنابراین، به طور دقیق بیان شده است که شبکه ترکیبی پیشنهادی ترکیبی بهتر از سایر شبکه های عصبی کانولوشن در مقایسه عمل می کند. معماری پیشنهادی را می توان ابزاری موثر برای رادیولوژیست ها جهت کاهش میزان منفی کاذب و مثبت کاذب دانست که کارایی تجزیه و تحلیل ماموگرافی بهبود می یابد [۵].

رمدان و همکارانش در سال (۲۰۲۰) مقاله ای مروری تحت عنوان "روش های مورد استفاده رایانه ای برای تشخیص سرطان سینه با استفاده از ماموگرافی" ارائه دادند. آنها معتقدند که طبق پیش بینی های انجمن سرطان آمریکا برای سال ۲۰۱۹، حدود ۲۶۸،۶۰۰ مورد جدید در ایالات متحده با سرطان مهاجم سینه در زنان، حدود ۶۲۹۳۰ مورد غیر تهاجمی جدید و حدود ۴۱،۷۶۰ مورد مرگ ناشی از سرطان سینه وجود دارد. در نتیجه، تقاضای زیادی برای متخصصان تصویربرداری از سینه وجود دارد که در گزارش اخیر موسسه پزشکی و شورای تحقیقات ملی نشان داده شده است. یکی از راه های پاسخگویی به این تقاضا، توسعه سیستم های تشخیص به کمک رایانه^۳ برای تشخیص سرطان سینه با استفاده از ماموگرافی است. این مطالعه با هدف بررسی پیشرفت ها اخیر در سیستم های CAD برای تشخیص سرطان سینه با استفاده از ماموگرافی و ارائه یک نمای کلی از روش های استفاده شده که در مراحل آن مرحله پیش پردازش، تقویت و پایان دادن به مرحله طبقه بندی ارائه می شود. سطح فعلی عملکرد برای سیستم های CAD⁴ دلگرم کننده است، اما برای تشخیص مستقل سیستم های CAD و تشخیص سیستم های بالینی کافی نیست. تا زمانی عملکرد سیستم های CAD با افزایش روش های موجود به طور چشمگیری از سطح فعلی افزایش یابد، بهره گیری از روش های نویدبخش در

اصلی ترین عیب روش آنها، انتخاب دستی پارامتر آستانه یابی و اندازه فیلتر میانگین گیری است [۱].
کار و همکارانش در سال (۲۰۱۵) تحقیق خود را با موضوع، انتخاب ژن از داده های به روز ژن ریز آرایه، برای طبقه بندی زیر گروه های سرطان با بکارگیری تکنیک PSO و مجاورت نزدیک ترین K تطبیقی را ارائه کرده اند. آنها در این مقاله روش انتخاب ژن KNN تطبیقی PSO برای طبقه بندی صحیح داده های ریز آرایه پیشنهاد کرده اند. یک بررسی برای انتخاب مقادیر بهینه ی K نیز پیشنهاد شده است. سپس آن ها بعضی از روش های اکتشافی را برای انتخاب زیرمجموعه ی ژن های حاوی اطلاعات مفید به کار بستند. به هر حال، بسته به کاهش ابعاد، بعضی از ژن های دارای اطلاعات مفید، ممکن است در روش اکتشافی شرکت نکنند. بدین ترتیب، مشارکت این پژوهش می تواند در غالب روش انتخاب ژن پیشنهاد شده ای لحاظ شود که قادر به کنار هم قرار دادن تمامی ژن های مورد نیاز است. به علاوه، بهره وری و ویژگی کلی زیرمجموعه های انتخاب شده با استفاده از هسته های مختلف SVM مجدداً تایید شده اند. نتایج این مطالعه نشان می دهند که روش پیشنهادی می تواند ابزاری مناسب برای انتخاب ژن از داده های به روز ژن فراهم سازد. روش پیشنهادی با دیگر روش ها نیز مورد مقایسه قرار گرفته است. مشاهده شده است که این روش ۶، ۳ و ۴ ژن را برای به دست آوردن دقت های ۱۰۰ درصدی، ۹۷/۰۵۸۸ درصدی و ۱۰۰ درصدی برای مجموعه داده های SRBCT، ALL-AML و MLL می یابد. یکی از نواحی بالقوه در بکارگیری فناوری ریز آرایه، میکروبیولوژی بالینی می باشد. این فناوری از ریز آرایه های تراکم میانه برای ارزیابی فی البداهه ی تعداد بالایی از اهداف ژنتیکی میکروبی استفاده می کند. روش پیشنهاد شده حاکی از احتمال توسعه ی روش های ساده شده برای تشخیص آسان زیرگروه های سرطان می باشد. این احتمال می تواند برای طراحی تراشه های ریز آرایه برای این اهداف باشد. سپس، آسیب شناسان می توانند ژن های مرتبط را به طور کارآمد شناسایی کنند و نمونه ها را بطور صحیح طبقه بندی نمایند. به علاوه، برای شناخت مزایای فناوری ریز آرایه در تشخیص سرطان، بهبود فناوری های پیش بینی کننده باید توسعه یابد. سکوه های ریز آرایه ی جدید باید با آمار و نرم افزار برای تحلیل و فناوری های داده کاوی به میان آیند. سپس، تنها رویه های دقیق تر در زمینه ی گسترده موثر خواهد بود. علاوه بر این مسئله، تکنیک های تشخیصی باید در کنار آزمایشگاه های مختلف برای ابر پروفایل ها در مسیر پژوهش سرطان شکل گیرند [۴].

³Computer-Aided Diagnosis

⁴Computer-Aided Diagnosis

¹Visual Geometry Group

²Modified Visual Geometry Group

۳ - تحلیل روش پیشنهادی

به علت پیچیدگی بافت‌های سینه، شباهت بین توده‌های سرطانی و بافت‌های طبیعی 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 ذرات 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 دهند. یکی از چالش‌های موجود در روش‌های تشخیص تصاویر ماموگرافی، مشکل بودن تحلیل بافت‌های چگال است. به این ترتیب که ناحیه‌ای از سینه که اطراف توده است به رنگ سفید به نظر می‌رسد و بنابراین توده و بافت پس‌زمینه از یکدیگر قابل تفکیک نیستند. 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 زیادی دارد که برخی از آنها عبارتند از: 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. از این رو این مقاله دنبال روشی است که بتواند به وسیله آن میزان تشخیص نهایی را بهبود داده و از این رو بعد از مرحله پیش پردازش که خود شامل یکسان‌سازی و تعدیل هیستوگرام است به جهت استخراج بهینه‌تر ویژگی‌های تصمیم‌گیری از تبدیل‌های کانتورلت و کرولت استفاده می‌نماید.

روش‌های بهبود مرسوم، روش‌های همسایگی ثابت هستند و به منظور بهبود تصویر مبتنی بر ویژگی‌های سراسری به کار می‌روند. روش‌های مبتنی بر ناحیه، در قطعه‌بندی تصویر به کار می‌روند، به علاوه این روش‌ها در بهبود کنتراست ویژگی‌های ماموگرام با توجه به پیرامون آن‌ها مورد استفاده قرار می‌گیرند. روش‌های بهبود مبتنی بر ویژگی در حوزه موجک عمل می‌کنند و در نهایت روش‌های بهبود فازی، روش‌هایی هستند که با اعمال عملگرهای فازی و ویژگی‌ها به بهبود ویژگی‌های ماموگرام

تشخیص الگو مانند افزایش داده‌ها در یادگیری عمیق و بهره‌برداری از پیشرفتهای قدرت محاسباتی رایانه‌ها، سیستم‌های CAD همچنان به عنوان یک روش بالینی ادامه خواهند داشت [۶].

راگاب و همکارانش در سال (۲۰۲۱) مقاله‌ای مروری تحت عنوان "چارچوبی برای طبقه‌بندی سرطان scki با استفاده از Multi-DCNN" ارائه دادند. در این مقاله یک سیستم جدید با کمک رایانه (CAD) مبتنی بر استخراج و طبقه‌بندی ویژگی‌ها با استفاده از تکنیک‌های DL برای کمک به رادیولوژیست‌ها در طبقه‌بندی ضایعات سرطان scki در ماموگرام ارائه می‌دهد. این کار توسط چهار آزمایش مختلف انجام می‌شود تا روش بهینه تعیین شود. مرحله اول شامل شبکه‌های از قبل آموزش دیده DCNN است. در مرحله دوم، ویژگی‌های عمیق DCNN استخراج شده و به یک طبقه‌بندی دستگاه بردار پشتیبانی (SVM) با عملکردهای مختلف هسته منتقل می‌شود. آزمایش سوم همجوشی ویژگی‌های عمیق را انجام می‌دهد تا نشان دهد که ترکیب ویژگی‌های عمیق دقت طبقه‌بندی کننده-های SVM را افزایش می‌دهد. سرانجام، در آزمایش چهارم، تجزیه و تحلیل مولفه اصلی PCA برای کاهش بردار ویژگی بزرگ تولید شده، در همجوشی ویژگی‌ها و کاهش هزینه محاسباتی معرفی شده است. این آزمایشات بر روی دو مجموعه داده شامل: (۱) زیر مجموعه تصویربرداری سینه از پایگاه داده دیجیتال برای غربالگری ماموگرافی (CBIS-DDSM¹) و (۲) پایگاه داده دیجیتال ماموگرافی جامعه تجزیه و تحلیل تصویر ماموگرافی (MIAS²) انجام شد. نتیجه دقت به دست آمده 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 [۷].

کومار و همکارانش در سال (۲۰۲۰) مقاله‌ای تحت عنوان "تشخیص خودکار سرطان سینه بر اساس الگوریتم‌های یادگیری ماشین" ارائه دادند. آنها معتقدند همانطور که علم داده در همه جا نقش محوری دارد، همچنین مراقبت‌های بهداشتی نیز در آن کاربرد برجسته‌ای دارد. سرطان سینه بالاترین نوع سرطان در میان زنان است، که ۶۲۷۰۰۰ نفر را به تنهایی از بین برده است. این میزان بالای مرگ و میر ناشی از سرطان سینه به تشخیص به موقع نیاز دارد تا بتواند به موقع پیشگیری انجام شود. به عنوان یک کمک کننده بالقوه در توسعه فن‌آوری پیشرفته، داده‌کاوی یک برنامه چند برابری در پیش‌بینی سرطان سینه می‌باشد. این کار بر روی اجرای تکنیک‌های مختلف طبقه‌بندی برای داده‌کاوی در پیش‌بینی سرطان بدخیم و خوش خیم سینه متمرکز است [۸].

²Mammographic Image Analysis Society

¹Curated Breast Imaging Subset of the Digital Database for Screening Mammography

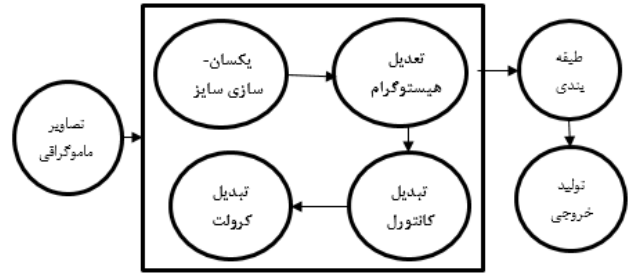
که در آن پارامترهای s و τ به ترتیب پارامترهای انتقال و مقیاس هستند انتقال یک ویولت به معنای تأخیر انداختن آن است و باعث می شود ویولت مورد نظر به سمت راست انتقال پیدا کرده و از اول تا انتهای سیگنال را طی کند. در تبدیل ویولت برخلاف تبدیل فوری به طور مستقیم پارامتر فرکانس وجود ندارد، در عوض پارامتر مقیاس وجود دارد که به طور معکوس با فرکانس در ارتباط است. مقیاس کردن همان طور که از معنی آن بر می آید به عنوان یک عملگر ریاضی سیگنال را منبسط و منقبض می کند، مانند مفهوم مقیاس در نقشه، مقیاس های بالا متناظر با دید کلی و فارغ از جزئیات به سیگنال است (متناظر با فرکانس های پایین) و مقیاس های کوچک متناظر با نگاه به جزئیات سیگنال است و لذا در تناظر با فرکانس های بالا خواهد بود.

در حقیقت تبدیل ویولت یک شباهت سنجی بین سیگنال و توابع پایه (ویولت ها) است و منظور از شباهت سنجی در این بحث شباهت بین محتوای فرکانسی است. نتیجه تبدیل ویولت یک ماتریس است که ستون های آن جابجایی در زمان را نشان می دهد و سطرها آن مقیاس مورد بررسی را نشان می دهد. به بیان دیگر ضرایب تبدیل ویولت بیانگر میزان نزدیکی سیگنال به ویولت در مقیاس مورد نظر می باشد. با انجام تبدیل موجک پیوسته ضرایب به دست می آید که در مقیاس های مختلف با بخش های مختلف سیگنال تولید می شوند. [۹].

۳-۲- مرحله اول پیش پردازش در روش پیشنهادی

موجک گابور: فاز بعدی مرحله پیش پردازش است که از موجک گابور برای این کار استفاده شده است. یک موجک گابور دو بعدی یک تابع گوسی مدوله شده با یک تابع سینوسی مختلط می باشد. این تابع را می توان به وسیله فرکانس تابع سینوسی و انحراف معیارهای پوش تابع گوسی بیان کرد. علیرغم این شکل کلی، هیچ تعریف استاندارد و دقیقی از تابع گابور دو بعدی وجود ندارد و اشکال مختلفی از این تابع در مقالات ظاهر شده است. بیشتر این اختلافات به اندازه های مختلف پوش تابع گوسی و فرکانس تابع سینوسی مربوط می شود. بر اساس تئوری موجک، تابع گابور که به نحو مناسبی نرمالیزه شده باشد، می تواند به عنوان یک موجک مادر برای تولید خانواده ای از موجک های گابور غیر متعامد به کار رود. با این وجود، گرچه تابع گابور با حذف پاسخ dc آن می تواند یک موجک قابل قبول قلمداد شود، لیکن این تابع به یک پایه متعامد منجر نخواهد شد و این بدان معنی است که تبدیل موجک بر پایه موجک گابور حاوی افزونگی است. برغم غیر متعامد بودن این خانواده موجک ها، تابع گابور تنها تابعی است که می تواند حد تئوری دقت تفکیک توأم اطلاعات در هر دو

می پردازند. بلوک دیاگرام روش پیشنهادی که مبتنی بر ویژگی است در (شکل-۱) نشان داده شده است.



(شکل-۱): بلوک دیاگرام روش پیشنهادی

همانگونه که در شکل ۱ نشان داده شده روش پیشنهادی شامل ۴ مرحله بوده که در پیش پردازش و استخراج ویژگی به عنوان کمک به بهبود مراحل طبقه بندی و تشخیص مورد استفاده قرار می گیرد. در ادامه هر یک از روش های عنوان شده در روش پیشنهادی مورد بررسی قرار می گیرد.

۳-۱- استخراج ویژگی در روش پیشنهادی: در گام اول روش پیشنهادی به استخراج ویژگی پرداخته می شود. برای استخراج ویژگی بعد از یکسان سازی سائز و تعدیل هیستوگرام، از تبدیل کانتورول، تبدیل کرولت و تبدیل موج ویولت استفاده می شود.

تبدیل ویولت^۱ یا موجک: تبدیل ویولت اطلاعات را به صورت همزمان در دو حوزه زمان و فرکانس ارائه می کند. در تبدیل ویولت برخلاف تبدیل فوری، که سیگنال یا سری اطلاعاتی را بر روی توابع سینوسی و کسینوسی و هارمونیک های آنها تجزیه می کرد، سیگنال بر روی یک دسته از توابع (موجک)، تصویر می گردد. بر خلاف توابع سینوسی و کسینوسی در تبدیل فوری تابع موجک در فضای زمان محدود بوده و بعد از چند ارتعاش به سرعت به سمت صفر میل می کند.

تئوری موجک برای غلبه بر مشکلات تبدیل فوری ارائه گردیده است. در این روش مسئله تقسیم سیگنال به بخش های مختلف با استفاده از مقیاس گذاری و انتقال دادن یک تابع حل می شود. این تابع در طول سری اطلاعاتی انتقال پیدا می کند و برای هر موقعیت آن، طیف سری اطلاعاتی محاسبه می شود. این مراحل برای توابعی با مقیاس های مختلف تکرار می شود و در نهایت نتیجه حاصل به صورت مجموع های از اطلاعات آرگومان-فرکانس به دست می آید. 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 می آیند. شکل کلی تبدیل پیوسته ویولت به صورت (رابطه-۱) است:

$$CWT_x^\psi = \frac{1}{\sqrt{s}} \int_{-\infty}^{+\infty} x(t) \psi\left(\frac{t-\tau}{s}\right) dt \quad (\text{رابطه-۱})$$

^۱Wavelet Transform

و تکنیکی که برای ایجاد یک هیستوگرام متعادل بکار می‌رود، متعادل‌سازی هیستوگرام نامیده می‌شود [۱۵]. در این تکنیک هیستوگرام تجمعی H از سطوح خاکستری G به عنوان قسمت اصلی تابع $F_{HE}(G)$ بکار می‌رود که سطوح خاکستری اصلی را به سطوح خاکستری تبدیل شده می‌نگارد:

$$F_{HE}(G) = G'_{min} + \Delta G' \frac{H(G)}{N} \quad \text{with } \Delta G' = G'_{max} - G'_{min} \quad (\text{رابطه-۴})$$

که در آن G'_{min} و G'_{max} بیانگر حدود بالا و پایین سطوح خاکستری تبدیل یافته‌اند و N بیانگر تعداد پیکسل‌هایی است که هیستوگرام بر روی آن اعمال می‌شود [۱۶].

یک روش عمومی HE تکنیک متعادل‌سازی هیستوگرام که بر مبنای توزیع سطوح خاکستری در کل تصویر اعمال می‌شود و برای بهبود جزئیات تصویر مناسب نمی‌باشد. با اعمال متعادل‌سازی هیستوگرام به نواحی محلی تصویر، متعادل‌سازی هیستوگرام تطبیقی به‌دست می‌آید:

$$F_{AHE}(G) = G'_{min} + \Delta G' \frac{H_{AHE}(G)}{N_{AHE}} \quad (\text{رابطه-۵})$$

متعادل‌سازی هیستوگرام تطبیقی به عنوان ابزاری مؤثر در جهت بهبود کنتراست در پردازش تصاویر بکار برده می‌شود. از خواص منحصر به فرد این روش این است که به شیوه محلی عمل می‌کند در حالیکه در روش‌های دیگر به‌صورت سراسری اعمال می‌شود در واقع در روش‌های قبلی یک آستانه‌گیری یا در کل تصویر انجام می‌شد و نسبت به آن بهبود یا متعادل‌سازی اعمال می‌شد. اما در روش AHE تصویر را به خانه‌های مانند کاشی تقسیم می‌کند و هر کاشی یا قسمت خود به ۹ خانه تقسیم می‌شود و برای متعادل‌سازی هر قسمت با مرکز خودش نقطه‌یابی شده و سنجش می‌شود 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 می‌شود و همین دلیل علاوه بر متعادل‌سازی مقداری نویز تصویر را زیاد می‌کند چون نویزها نیز خود باعث ایجاد لبه می‌شوند، جهت تکمیل آن و رفع عیب آن از روش CLAHE² استفاده می‌شود.

در این روش تابع انتقال از طریق محور γ ها و حول انتقال CDF کنتراست محاسبه می‌شود. یکی دیگر از تفاوت‌های مهم این روش با روش‌های قبلی این است که مقدار آستانه یا آستانه قطع در محور γ ها انتخاب می‌شود. پس از تعیین نقطه آستانه در محور γ ها تصویر قطع می‌شود. نقاطی که دارای فراوانی زیادی هستند و در روش پیشنهادی به این فراوانی‌ها نیازی نیست بنابراین آنها حذف

حوزه مکانی و فرکانسی را حاصل نماید. فیلتر گابور از عملگرهای لگاریتمی برای تصویر استفاده می‌کند. برای این عمل از روابط ۲ استفاده می‌شود.

$$g_{\theta}(x,y) = \exp\left(-\pi\left[\frac{x'^2}{\sigma_x^2} + \frac{y'^2}{\sigma_y^2}\right]\right) \cos(2\pi f x')$$

$$x' = x \cos\theta + y \sin\theta$$

$$y' = x \sin\theta + y \cos\theta \quad (\text{رابطه-۲})$$

در این روابط θ جهت فیلتر و σ انحراف معیار و f فرکانس کسینوسی هستند. فیلتر گابور می‌تواند در جهت‌های مختلف با تغییر مقیاس s کار کند. این الگوها سبب استخراج مقادیر به ازای هر بلوک از تصویر می‌گردد [۱۰].

۳-۳- مرحله دوم پیش‌پردازش روش پیشنهادی

مقارن نمودن و مربع نمودن ابعاد تصویر: در ابتدا برای حذف آرتیفکت‌های حاشیه و لبه‌ها که در اثر استفاده از تبدیل موجک در مرحله پردازش اصلی به‌وجود می‌آیند، تصویر ورودی به صورت مقارن گسترش داده می‌شود. گسترش مقارن، پیوستگی را در قطعات کپی شده ورودی حفظ نموده و ضرایب بزرگ موجک ناشی از ناپیوستگی حاشیه را حذف می‌نماید. در گسترش مقارن فرض می‌شود که سیگنال مورد پردازش که در این‌جا سیگنال تصویر است در مرزها یا حاشیه تصویر مقارن است.

بهبود کنتراست با تعدیل هیستوگرام: برای این

عمل از متعادل‌سازی هیستوگرام تطبیقی AHE¹ استفاده شده است. در این روش به وسیله‌ی بخش‌بندی تصویر به بلوک‌های کوچک، بهترین سطح آستانه برای هر بلوک به وسیله‌ی روش‌هایی نظیر HE و BBHE تعیین می‌گردد سپس می‌توان سطوح آستانه برای هر پیکسل را به وسیله‌ی درون‌یابی بین مراکز بلوک‌ها تعیین نمود. تکنیک‌های مختلفی از متعادل‌سازی هیستوگرام تصویر، برای بهبود کنتراست تصاویر گزارش شده است [۱۱-۱۴]. در این تکنیک با فرض مقادیر گسسته برای سطوح خاکستری تابع احتمال به‌صورت رابطه ۳ قابل تعریف است:

$$P_r(r_k) = \frac{n_k}{n} \quad 0 \leq r_k \leq 1 \quad . \quad k = 0.1 \dots L-1 \quad (\text{رابطه-۳})$$

که در این رابطه r_k سطح خاکستری پیکسل‌هایی 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 می‌شود، L تعداد سطوح خاکستری، $P_r(r_k)$ احتمال رخدادن سطح خاکستری Kام و n_k تعداد دفعاتی است که این سطح خاکستری در تصویر رخ می‌دهد. نمودار $P_r(r_k)$ بر حسب r_k معمولاً هیستوگرام نامیده می‌شود

²Contrast Limited Adaptive Histogram Equalization (CLAHE)

¹Adaptive Histogram Equalization (AHE)

که D_a ماتریس مقیاس گذاری سهمی و R_θ چرخش به اندازه θ رادیان است اگر تابع ψ یک مربع واحد باشد، دیده می شود که a, b, θ, ψ در محدوده قرار می گیرند که پوش آن یک مستطیل با ابعاد a در \sqrt{a} است که محور کوچک آن در امتداد θ جهت می گیرد. یکی از ویژگی های مهم کرولت ها این است که از قانون هارمونیک تبعیت می کنند که بیان می کند تحلیل و بازسازی یک تابع دلخواه $f(x_1, x_2)$ توسط برهم نهی امکان پذیر است. هر کدام از این کمیت های مقیاس، موقعیت و زاویه را می توان به طور مجزا به شکل زیر بیان کرد:

$$\begin{aligned} \alpha_j &= 2^{-j}, j = 0, 1, 2, \dots \\ \theta_{j,l} &= 2\pi l \cdot 2^{-|j/2|}, l = 0, 1, \dots, 2^{|j/2|} \\ b_k^{(j,l)} &= R_{\theta_{j,l}} \left(k_1 2^{-j}, k_2 2^{-\frac{j}{2}} \right), k_1, k_2 \in \mathbb{Z}, \\ \Psi_{j,l,k} &= \Psi_{\alpha_j b_k^{(j,l)}, \theta_{j,l}} \end{aligned} \quad (\text{رابطه-۸})$$

رابطه k, a, z, ψ به صورت زیر بیان می شود:

$$\begin{aligned} f &= \sum_{j,l,k} \langle f, \Psi_{j,l,k} \rangle \Psi_{j,l,k} \\ \|f\|_{L^2}^2 &= \sum_{j,l,k} |\langle f, \Psi_{j,l,k} \rangle|^2 \end{aligned} \quad (\text{رابطه-۹})$$

خروجی تبدیل، مجموعه ای از ضرایب کرولت گسسته $(j, l, k) \in \mathbb{C}^D$ است.

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

که k, a, z, ψ شکل موج گسسته می باشد [۱۷]. ساختار

پایه سازی FDCT به شیوه پیچشی شامل مراحل زیر می باشد:

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

های فوریه $t / > -t_1$ و $t_2 < 2t$ به دست می آید.

مرحله ۲: برای هر مقیاس z و هر زاویه a حاصل ضرب

$$f^*[t_1, t_2] \cdot \psi[z, a, t_1, t_2]$$

مرحله ۳: این حاصل ضرب دور مبدا پیچشی (متناوب)

می شود. رابطه ۱۱ برقرار می شود:

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

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, z, a اعمال

شده، بنابراین ضرایب گسسته $CD(j, l, k)$ به دست می آیند [۱۷].

ورودی مرحله استخراج ویژگی، تصویر نرمال SHC شده

است. تبدیل کرولت بر روی تصویر ورودی اعمال می شود، که نتیجه

تبدیل کرولت ضرایب کرولت می باشند. برای هر دسته از ضرایب

کرولت، مقادیر میانگین و انحراف معیار یا ضرایب به تنهایی را به

عنوان ویژگی به دست آورده می شود. بردار ویژگی می تواند هر یک

شده و به سطح دامنه اضافه می شود. این امر سبب می شود مقداری از نمودار از سطح قطع شود که مجدداً نیز به دامنه اضافه می شود تا تقریباً به سطح مطلوبی برسد و آنچه که باقی می ماند و در سطح متعادل سازی بیشتر تقویت می شود همان نقاط مطلوب و مورد نیاز روش پیشنهادی در تصویر است.

۳-۴- مرحله استخراج بهینه تر ویژگی ها مبتنی بر کانتورلت و کرولت

کانتورلت: تبدیل کانتورلت، تبدیل دو بعدی جهت دار

جدایی ناپذیری است که برای توصیف منحنی ها و جزئیات ظریف در تصاویر به کار می رود. برخلاف دیگر تبدیل ها که ابتدا در یک حوزه پیوسته ساخته می شوند و سپس برای نمونه برداری داده گسسته می شوند، تبدیل کانتورلت با کمک فیلتر بانک ها از یک حوزه گسسته شروع شده و سپس از طریق یک چهارچوب تحلیلی مالتی رزولوشن به یک حوزه پیوسته همگرا می شود. برای تبدیل گسسته کانتورلت یک الگوریتم سریع نیز وجود دارد. تبدیل کانتورلت از دو بخش عمده تشکیل شده است: تجزیه گر هرم لاپلاس و بانک های فیلتر جهت دار^۱ تصویر اصلی توسط هرم لاپلاس به دو تصویر تبدیل می شود: تصویر پایین گذر و تصویر میان گذر. هر تصویر میان گذر در مرحله ی بعد توسط بانک فیلتر جهت دار تجزیه می گردد و با تکرار مراحل ذکر شده روی تصویر پایین گذر، تجزیه ی چند مقیاسه و چند جهته از تصویر به دست می آید [۱۷].

کرولت: تبدیل کرولت یک تبدیل چند مقیاسی و چند

جهتی با توابع پایه ی سوزنی شکل و بسیار حساس به زاویه می باشد. توابع پایه تبدیل ویولت، ایزوتروپیک هستند و بنابراین برای بیان بخش منفرد منحنی ها به تعداد ضرایب زیادی نیاز دارد. کرولت ها مقیاس بندی سهمی شکل را نیز می پذیرند. به خاطر این خواص، تبدیل کرولت، نمایش بهینه و کم حجمی از منحنی های منفرد ارائه می دهد. تبدیل کرولت در مقیاس ها و زوایای مختلف، برخلاف دیگر تبدیل های زاویه ای مثل ویولت و گابور، محدوده ی فرکانسی را به طور کامل پوشش می دهد [۱۸].

کرولت ها به مانند امواج تختی هستند که در فضای دو

بعدی توسط دو متغیر جهت و فرکانس مرکزی به دست مشخص می شوند. کرولت از اعمال اتساع^۲، چرخش و انتقال روی تابع ψ به دست می آیند.

کرولت ها توسط متغیر مقیاس a که $0 < a(1)$ موقعیت b

و چرخش θ ، اندیس گذاری می شوند و دارای رابطه زیر می باشد.

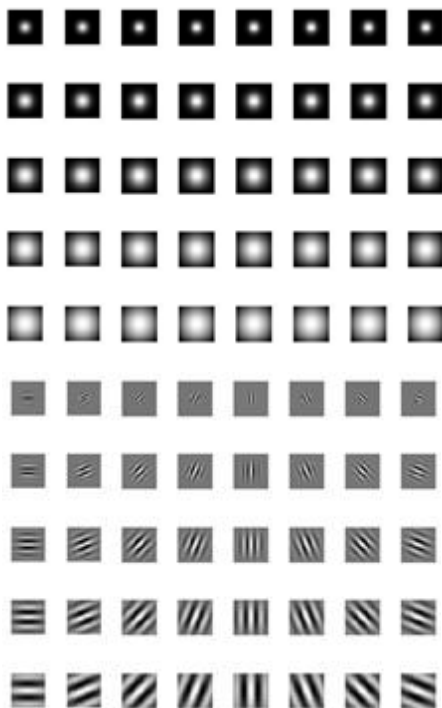
$$\Psi_{a,b,\theta}(x) = \alpha^{\frac{3}{4}} \Psi(D_a R_\theta(x - b)) \quad (\text{رابطه-۶})$$

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

²Dilation

¹Directional Filter Bank

تبدیل ویولت اطلاعات را به صورت همزمان در دو حوزه زمان و فرکانس ارائه می‌کند. در تبدیل ویولت بر خلاف تبدیل فوریه، که سیگنال یا سری اطلاعاتی را بر روی توابع سینوسی و کسینوسی و هارمونیک‌های آنها تجزیه می‌کند (استفاده می‌شود. در فاز پیش‌پردازش از موجک گابور استفاده شده است. فیلتر گابور بر پایه فیلترهای تکی بوده که با توجه به انتخاب اولیه به بلوک‌های مجاور و هم‌اندازه اعمال شده تا نویز مربوطه را به نحوی تغییر دهد که در ویژگی‌های نهایی استخراج شده از تصاویر تاثیرگذار نباشد. نمونه‌ای از فیلترهای اعمالی در بلوک‌های ۵*۸ در (شکل ۲-۲) نشان داده شده است.



(شکل-۲): نمونه‌ای از فیلتر گابور استفاده شده در تصاویر

برای استخراج ویژگی با توجه به مراحل بیان شده در بخش قبل ابتدا فیلترهایی بر روی تصاویر اجرا شده است و این فیلترها سبب بهبود خروجی تصاویر می‌گردد. قبل از اعمال این فیلترها بسیاری از استخراج ویژگی‌ها می‌تواند دچار خطا شده و به این ترتیب داده‌های استخراجی مناسبی شکل نخواهد گرفت. نمودار اولیه و نمودار بعد از فیلتر اعمال شده در (شکل-۳) و (شکل-۴) نشان داده شده است.

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 مقیاس-بندی 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 زیر باند 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 است. پس برای تبدیل کرولت ۴ سطحی، ۲۶ زیرباند از ضرایب کرولت تولید می‌شود که هر زیرباند مجموعه‌ای از ضرایب است.

۴- نتایج تجربی

در این بخش روش گفته شده مذکور مبتنی بر نرم‌افزار Matlab و Weka پیاده سازی و نتایج مراحل مختلف ارزیابی نمایش داده می‌شود. قابل به ذکر است که سیستم استفاده شده در ارزیابی دارای پردازنده Intel Pentium T4400 @2.2 GHz، حافظه Processor، Ram 3.0 GB، Hard Disk 320 GB و سیستم عامل Windows 7 است.

۴-۱ مجموعه داده:

مجموعه داده تصاویر دیجیتال ماموگرافی غربالگری MIAS¹ یک منبع^۲ برای استفاده جامعه پژوهشی در تحلیل تصویر ماموگرافی است. این مجموعه داده تقریباً در ۵۰۰ مطالعه 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 ACR^۳، امتیاز ظریف برای ناهنجاری‌ها، شرح کلمات کلیدی ACR از اختلالات) و اطلاعات تصویر (اسکنر، رزومه مکانی، ...) است. تصاویر حاوی مناطق مشکوک مربوط به اطلاعات مکان و نوع مناطق در سطح پیکسل است که پیکسل کوچکترین جز عکس در آن سطح است. در این مقاله تنها از ۱۵۰ نمونه تصادفی از ۵ دسته عادی، سرطان خوش‌خیم، سرطان بدخیم، میکروکلسیفیکیشن^۴ خوش‌خیم و میکروکلسیفیکیشن بدخیم استفاده شده است که در هر دسته به طور یکسان ۳۰ نمونه قرار دارد.

۴-۲ ارزیابی شبیه‌سازی

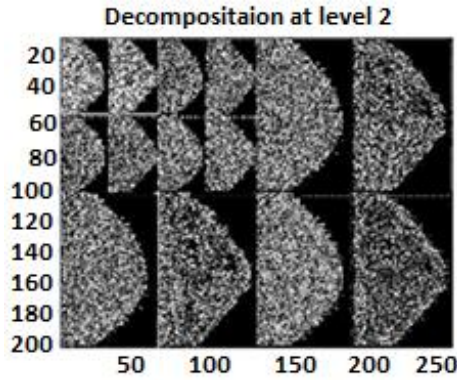
در گام اول روش پیشنهادی به استخراج ویژگی پرداخته می‌شود. برای استخراج ویژگی بعد از یکسان‌سازی سایز و تعدیل هیستوگرام، از تبدیل کانورت، تبدیل کرولت و تبدیل موج ویولت

³American College of Radiology

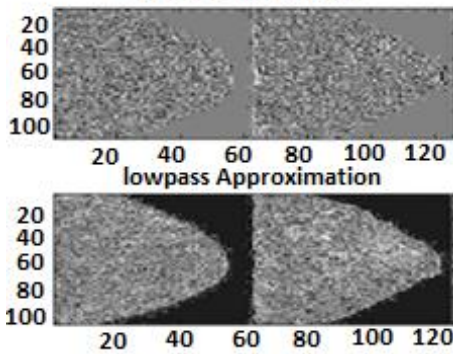
⁴Microcalcifications

¹Mammographic Image Analysis Society

²<http://peipa.essex.ac.uk/info/mias.html>



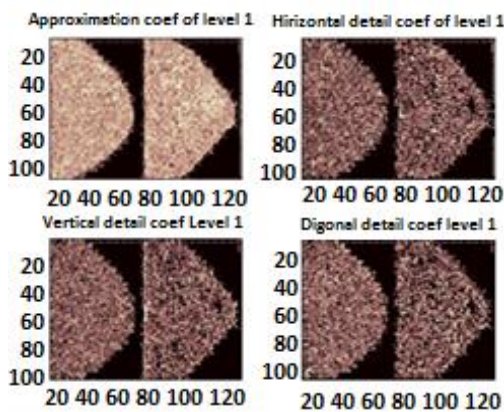
الف) استخراج با سطح ۴
Verticla Detail Image



ب) استخراج با سطح ۲

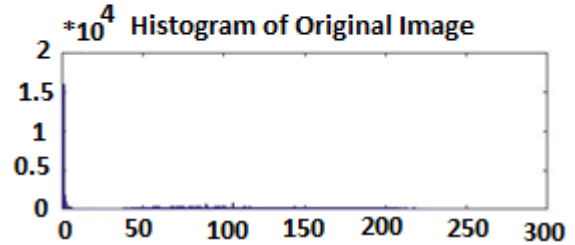
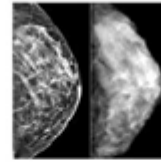
(شکل-۵): استخراج ویژگی سطوح خاکستری

در تصاویر ۵ و ۶ نمونه ای از ویژگی های برداری سطح-بندی های مختلف تبدیل عنوان ویژگی نهایی با فیلتر گابور استخراج شده و برای یک تصویر مجموعه داده نشان داده شده است. همانگونه که در (شکل-۵) الف نشان داده شده است برای هر تصویر سطوح مختلف آن استخراج می گردد همچنین تغییر سطوح هر تصویر از نمونه عادی و تغییر شکل سطوح آن همانند (شکل-۵) ب استخراج شده و بردارهای مربوطه به عنوان ویژگی های نهایی تصاویر در روش پیشنهادی محسوب می گردد.



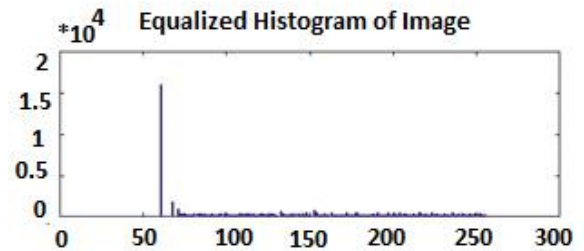
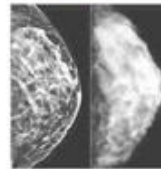
الف) لبه یابی

Original Image



(شکل-۳): نمودار هیستوگرام یک تصویر از مجموعه داده

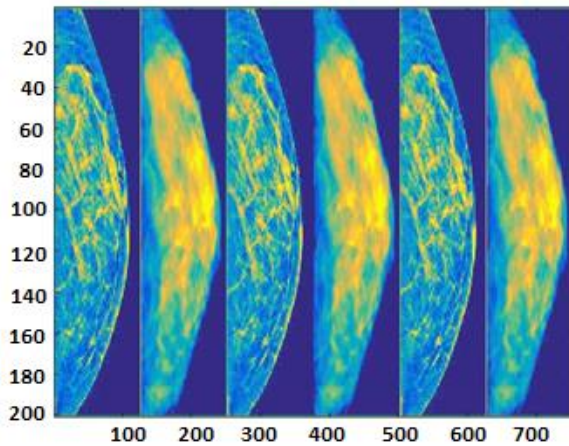
Histogram Equalized Image



(شکل-۴): نمودار هیستوگرام تصویر استفاده شده در شکل ۳

بعد از اعمال AHE

در مرحله ساخت داده ها بعد از اعمال تحلیل های گفته شده به استخراج داده پرداخته می شود. با توجه به سطوح استفاده شده و فیلترهای ترکیبی این تبدیل برای تصاویر رنگی و خاکستری و انواع مختلف تبدیلات و فیلترها (رنگی تبدیل شده به خاکستری، این فیلتر عکس رنگی را به خاکستری تبدیل می کند) و همچنین آزمون های در برگزیده آن به تعداد ۹۵ ویژگی مستخرج این فیلترها از این نوع تصاویر پرداخته شده است. همچنین یک ویژگی بیان کننده نوع تصویر نیز قرار دارد که مجموعاً ۹۶ ویژگی است.



(شکل-۷): طبقه بندی و دسته بندی تصاویر

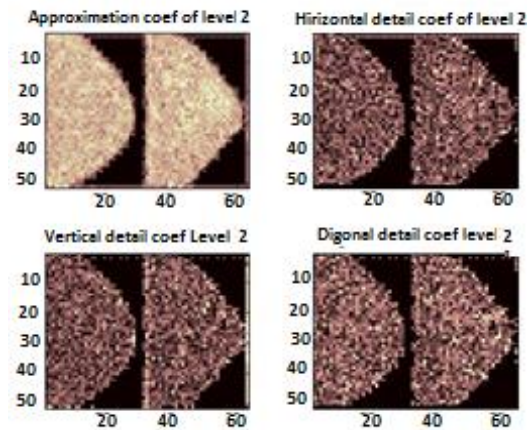
همانگونه که در (شکل-۷) نشان داده شده است هر یک از تصاویر در مجموعه با دسته‌های مختلف در جایگاه‌های خود برای طبقه‌بندی نهایی قرار می‌گیرند و این ایجاد تفاوت‌ها در نهایت در تشخیص نهایی تصاویر و دسته‌هایشان کمک کننده خواهند بود.

۴-۳ ارزیابی شبیه‌سازی

بعد از شبیه‌سازی روش پیشنهادی نوبت به ارزیابی روش پیشنهادی می‌رسد. برای این امر روش پیشنهادی را با طبقه‌بندی‌های پرکاربرد و اساسی که در بیشتر مقالات مطالعه شده مورد استفاده قرار گرفته مورد ارزیابی قرار گرفت. معیارهای مورد نیاز در بخش زیر دسته‌بندی شده است:

صحت^۱: صحت، به سنج‌ای از میزان تطابق پیش‌بینی‌های یک مدل اشاره دارد که با واقعیت مدل‌سازی شده هم‌خوانی دارد. صحت یک طبقه‌بندی بر روی داده‌ی آزمایشی، می‌تواند به صورت تقسیم تعداد اشیای درست طبقه‌بندی شده به تعداد کل اشیا محاسبه شود.

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



(ب) تراکم بافت

(شکل-۶): استخراج ویژگی رنگی تصاویر

همانگونه که در (شکل-۶) نشان داده شده است برای هر تصویر رنگی نیز چهار ویژگی عمده با روش‌های توضیح داده شده در بخش قبل استخراج و مورد بررسی قرار می‌گیرد. اولین ویژگی ریخت شناسی لبه‌یابی و شکل ظاهری است (شکل-۶) الف. در ویژگی بعدی ریخت‌شناسی تراکم بافت‌ها و میزان نزدیکی و چسبندگی و به اصطلاح حجم بافت استخراج می‌شود. ویژگی بعدی ریخت‌شناسی مورد ارزیابی قرار گرفته میزان شکل بافت و توده قسمت سمت چپ و راست موجود در عکس بوده که بررسی و استخراج می‌شود (شکل-۶) ب. کلیه این موارد با استفاده از فیلترهای تبدیل کانتورلت و تبدیل کرولت مورد استخراج قرار می‌گیرد.

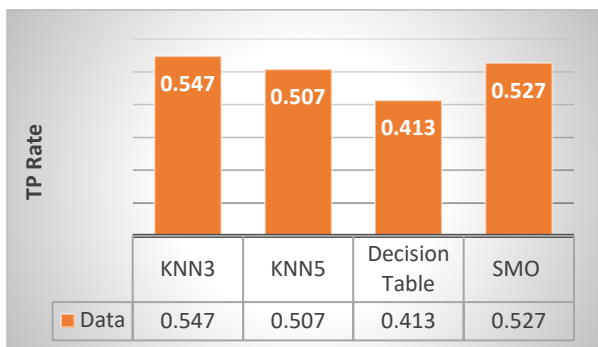
فاز نهایی و انتهایی، انتخاب استخراج ویژگی و در نظر گرفتن شکل‌های گوناگون ویژگی می‌باشد که به عنوان ورودی دسته‌بندی و انتخاب نهایی استفاده خواهد شد. نمونه‌ای از این ایجاد تمایز برای تشخیص مناسب در (شکل-۷) نشان داده شده است.

^۱Accuracy

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

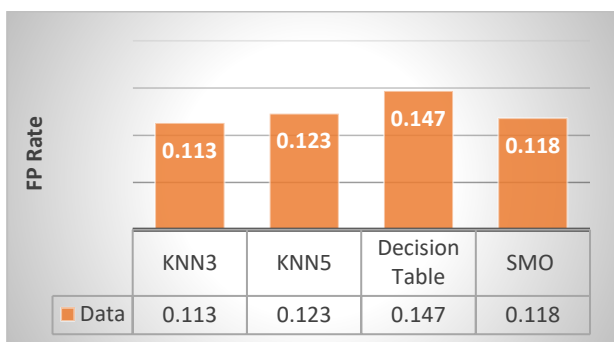
بر اساس معیارهای مذکور ارزیابی کارایی الگوریتم پیشنهادی را مبتنی بر طبقه‌های K نزدیک‌ترین همسایه، Decision Table و SMO انجام شده است.

در شکل ۸ ارزیابی روش‌ها با معیار TP در مجموعه‌های داده مختلف نشان داده شده است. همانگونه که در نشان داده شده است با توجه به نحوه استخراج داده ارایه شده روش K نزدیکترین همسایه با مقدار ۳ در مجموعه Data و روش SMO بهترین نتیجه را داشته است. با توجه به نحوه استخراج ویژگی و انتخاب ویژگی معرفی شده روش درخت تصمیم پایین‌ترین نتایج را از خود نشان داده است.



(شکل ۸): ارزیابی روش‌ها مختلف با استخراج داده پیشنهادی با معیار TP

در (شکل ۹) ارزیابی روش‌ها با معیار FP نشان داده شده است. بیشترین تشخیص مقدار FP همانگونه که در (شکل ۹) نشان داده شده است روش درخت تصمیم بالاترین نرخ تشخیص این مقدار را دارا است. روش‌های دیگر نیز در رتبه‌های بعدی قرار دارند.



(شکل ۹): ارزیابی روش‌ها مختلف با استخراج داده پیشنهادی با معیار FP

معیار FP

همانگونه که در (شکل ۱۰) نشان داده شده است روش K نزدیکترین همسایه با مقدار K برابر ۷ بهترین نتیجه را در معیار

(جدول ۱): کلاس تخصیص یافته توسط مدل

منفی	مثبت		
<p>FN² Rate: بیانگر تعداد رکوردهایی است که دسته واقعی آنها مثبت بوده و الگوریتم طبقه‌بندی نیز دسته آنها را به اشتباه منفی تشخیص داده است.</p>	<p>TP¹ Rate: بیانگر تعداد رکوردهایی است که دسته واقعی آنها مثبت بوده و الگوریتم طبقه‌بندی نیز دسته آنها را بدرستی مثبت تشخیص داده است.</p>	مثبت	کلاس واقعی
<p>TN⁴ Rate: بیانگر تعداد رکوردهایی است که دسته واقعی آنها منفی بوده و الگوریتم طبقه‌بندی نیز دسته آنها را به درستی منفی تشخیص داده است.</p>	<p>FP³ Rate: بیانگر تعداد رکوردهایی است که دسته واقعی آنها مثبت بوده و الگوریتم طبقه‌بندی نیز دسته آنها را به اشتباه مثبت تشخیص داده است.</p>	منفی	

دقت و بازخوانی^۵: دقت و بازخوانی معیارهای کاربردی در

حوزه بازیابی اطلاعات هستند که میزان تناسب اسناد بازیابی شده توسط سیستم را با نیاز کاربر تعیین می‌کنند. این معیارها به صورت زیر تعریف می‌شوند:

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

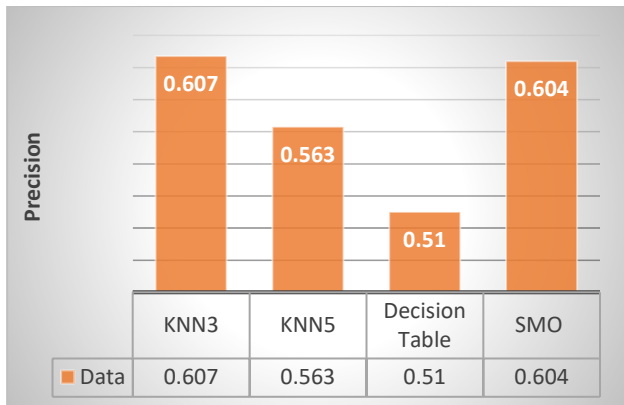
$$Precision = \frac{TP}{FP+TP} \quad (\text{رابطه-۱۴})$$

به جای این دو معیار، می‌توان از یک معیار ترکیبی برای ارزیابی کارایی بازیابی به نام F-measure استفاده نمود. پارامتر F-Measure نیز ترکیب Recall و Precision است [۱۹،۲۰].

⁴True Negative
⁵Precision and Recall

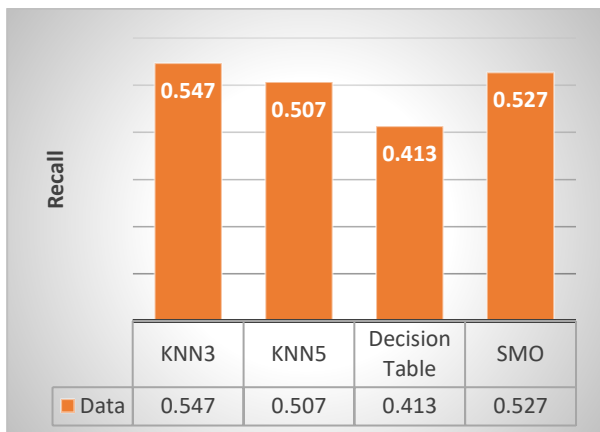
¹True Positive
²False Negative
³False Positive

در (شکل-۱۱) ارزیابی روش براساس معیار Precision نشان داده شده است.



(شکل-۱۱): ارزیابی روش‌ها مختلف با استخراج داده پیشنهادی با معیار Precision

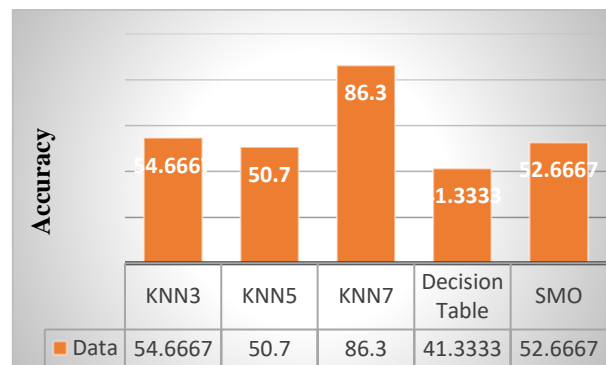
همانگونه که در (شکل-۱۱) نشان داده شده است روش K نزدیکترین همسایه با مقدار K برابر ۳ بهترین نتیجه را داشته و بعد از آن روش SMO دقت مناسبی داشته است. با توجه به نحوه استخراج ویژگی و انتخاب ویژگی معرفتی شده روش درخت تصمیم پایین‌ترین نتایج را از خود نشان داده است. البته روش‌های نزدیکترین همسایه با K برابر ۵ نیز دقت قابل قبولی دارد.



(شکل-۱۲): 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 در رتبه‌های بعدی قرار دارند. در

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 داده است. با توجه به نحوه استخراج ویژگی و انتخاب ویژگی معرفتی شده روش درخت تصمیم پایین‌ترین نتایج را از خود نشان داده است. البته روش‌های نزدیکترین همسایه با K برابر ۵ نیز صحت قابل قبولی دارد.



(شکل-۱۰): ارزیابی روش‌ها مختلف با 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 همسان از نظر صحت

ردیف	مرجع	سال انتشار	روش	صحت به درصد
۲	[۲۱]	۲۰۱۵	بردار پشتیبان تصمیم	۵۲
۳	[۲۲]	۲۰۱۵	KNN	۸۲.۱
۴	[۲۳]	۲۰۱۵	روش‌های آماری و شبکه عصبی	۸۵
۵			KNN با K برابر ۷ داده‌های استخراج شده با روش پیشنهادی	۸۶.۳

مقایسه با معیارهای دیگر نیز سنجیده شد و نتایج نشان داد در آن معیارها نیز بهبود قابل توجهی داشته است.

معیارهای ارزیابی روی طبقه بندهای مشخصی باهم مقایسه شدند که در بین آنها روش نزدیکترین همسایگی دارای دقت بیشتری بوده و در مرحله بعد از آن طبقه SMO با بیشترین دقت و حساسیت قرار دارد. در کارهای آینده می توان با افزایش تعداد زاویه ها در فیلتر کانتورلت لبه ها را با دقت بیشتری از تصاویر استخراج کرد.

۶- 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] Kar, Subhajit, Kaushik Das Sharma, and Madhubanti Maitra. "Gene selection from microarray gene expression data for classification of cancer subgroups employing PSO and adaptive K-nearest neighborhood technique." *Expert Systems with Applications* 42, no. 1 (2015): 612-627.
- [5] Khamparia, Aditya, Subrato Bharati, Prajyo Podder, Deepak Gupta, Ashish Khanna, Thai Kim Phung, and Dang NH Thanh. "Diagnosis of breast cancer based on modern mammography using hybrid transfer learning." *Multidimensional systems and signal processing* 32, no. 2 (2021): 747-765.
- [6] Ramadan, Saleem Z. "Methods used in computer-aided diagnosis for breast cancer detection using mammograms: a review." *Journal of healthcare engineering* 2020 (2020).
- [7] Ragab, Dina A., Omneya Attallah, Maha Sharkas, Jinchang Ren, and Stephen Marshall. "A framework for breast cancer classification using multi-DCNNs." *Computers in Biology and Medicine* 131 (2021): 104245.
- [8] 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.

(شکل-۱۳) ارزیابی روشها با معیار F-Measure نشان داده شده است.



(شکل-۱۳): ارزیابی روشها مختلف با استخراج داده پیشنهادی

با معیار F-Measure

همانگونه که در (شکل-۱۳) نشان داده شده است 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.

۵- نتیجه گیری و پژوهش های آتی

در این مقاله یک روش استخراج ویژگی از تصاویر ماموگرافی دریافتی از بیمار ارایه شده است. تصویر ورودی به سیستم طی فرآیند استخراج ویژگی قرار گرفته و سه دسته ویژگی اصلی ریخت شناسی و بافت شناسی، از آن استخراج می شود. این فرآیند طی چرخه، یکسان سازی سایز، تعدیل هیستوگرام، تبدیل کانتورلت و تبدیل کرولت انجام می شود. به دلیل حساس بودن سیستمها سعی در استخراج مناسب ویژگیها با انواع سطوح و ماتریسهای شناخته شده مانند ماتریس رخدادها و سطوح خاکستری شده است. بعد از استخراج ویژگی به جهت بهبود صحت خروجی انتهایی از روشهای آماری برای انتخاب ویژگی استفاده شد و زیر مجموعه ای از ویژگیهای استخراج شده برای ورودی طبقه بند انتخاب شدند. بعد از این در گام بعد خروجی روش پیشنهادی در گام بعد برای مقایسه تاثیر آنها به طبقه بندهای مختلف اعمال شده و خروجی آنها با یکدیگر مقایسه گردید در گام آخر خروجی، روش پیشنهادی در طبقه بندیهای مختلف مورد ارزیابی قرار گرفت. بعد از اعمال ویژگیهای استخراج شده به ارزیابی روش گفته شده توسط نتایج طبقه بندها 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 پیشنهادی در مقایسه با روشهای مورد

- [22] E. Honda, R. Nakayama, H. Koyama, A. Yamashita, *Computer-aided diagnosis scheme for distinguishing between benign and malignant masses in breast DCEMRI*, *J. Digit. Imaging* (2015) 1–6.
- [23] S. Song, B.K. Seo, K.R. Cho, O.H. Woo, G.S. Son, C. Kim, S.B. Cho, S.-S. Kwon, *Computer-aided detection (CAD) system for breast MRI in assessment of local tumor extent, nodal status, and multifocality of invasive breast cancers: preliminary study*, *Cancer Imaging* 15 (1) (2015).



حسن نجفی مدرک کارشناسی خود را در رشته برق گرایش الکترونیک در سال ۱۳۹۰ از دانشگاه شیراز، کارشناسی ارشد خود را در رشته برق گرایش الکترونیک دیجیتال در سال ۱۴۰۰ از دانشگاه آزاد اسلامی زنجان اخذ کرده است. نشانه رایانامه ایشان عبارتند از najafe@gmail.com

روش ارجاع به مقاله : ح. نجفی، ارائه روشی جهت بهبود کیفیت تصاویر ماموگرافی مبتنی بر ترکیب تکنیک‌های کانتورلت و کرولت، دوفصلنامه محاسبات و سامانه های توزیع شده، سال سوم، شماره دوم، شماره پیاپی ۶، صفحه ۵۹ تا ۷۲، سال ۱۳۹۹.

How to cite: Hassan Najafy. A New Energy Efficient Clustering Method for Wireless Sensor Networks Based on Cuckoo Optimization Algorithms and Genetic Composition Operator, *Journal of Distributed Computing and Systems (JDCCS)*, Vol 3, Issue 2, Page 59-72, 2021.

- [9] Addison, Paul S. "Introduction to redundancy rules: the continuous wavelet transform comes of age." (2018): 20170258.
- [10] Lee, Tai Sing. "Image representation using 2D Gabor wavelets." *IEEE Transactions on pattern analysis and machine intelligence* 18, no. 10 (1996): 959-971.
- [11] 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.
- [12] Wang, X. and Wong, B.S. (2004). *Image Enhancement for Radiography Inspection. The Third international Conference on Experimental Mechanics, SPIE Proceedings*, 5852.
- [13] Wang, G. and Liao, T.W.(2002). *Automatic Identification of Different Types of Welding Defects in Radiographic Images*. *NDT & E International*, 35:519-528.
- [14] 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 & E International*, 37:291-299.
- [15] Gonzalez, R.C. and Wintz, P.(1987). *Digital Image Processing*. Addison-Wesley, 2nd Edition.
- [16] Castleman, K.R.(1979). *Digital Image Processing*. Englewood Cliffs N.J., Prentice Hall, 90-94.
- [17] [Candes, E. J., and Donoho, D. L., "Continuous curvelet transform", II. Discretization and frames: Applied and Computational Harmonic Analysis, 19, 198-222, 2003.
- [18] [Werner Ebeling, H. M., "Evolutionary Algorithms And Optimization", Berlin: Eingereicht An Der Mathematisch-Naturwissenschaftlichen Fakultät I Humboldt-Universität Zu Berlin , Dezember 2002.
- [19] 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 for Technology of Information and Communication*, pp. 411-416. IEEE, 2018.
- [20] Sa'di, Sadri, Amanj Maleki, Ramin Hashemi, Zahra Panbechi, and Kamal Chalabi. "Comparison of data mining algorithms in the diagnosis of type II diabetes." *International Journal on Computational Science & Applications (IJCSA)* 5, no. 5 (2015): 1-12.
- [21] M. Abdel-Nasser, H.A. Rashwan, D. Puig, A. Moreno, *Analysis of tissue abnormality and breast density in mammographic images using a uniform local directional pattern*, *Expert Syst. Appl.* 42 (24) (2015) 9499–9511.