

## کلاس‌بندی تومورهای خوش‌خیم و بدخیم در تصاویر اولتراسوند پستان با کمک ویژگی‌های ریخت‌شناسی

هدا نعمت<sup>۱</sup>، علی محلوچی‌فر<sup>۲</sup>، علی گویا<sup>۳</sup> و نسرين احمدی‌نژاد<sup>۴</sup>

### چکیده

سرطان پستان، دومین عامل مرگ‌ومیر زنان در جهان محسوب می‌شود و به دلیل ناشناخته بودن علت این بیماری، تنها روش کنترل آن شناسایی و تشخیص زودهنگام است. مهمترین روش تشخیص سرطان پستان، نمونه‌برداری از بافت مشکوک و انجام آزمایش‌های آسیب‌شناسی است. از آنجا که انجام این روش تهاجمی بوده و در اکثر موارد غیرضروری می‌باشد، به همین جهت محققان در تلاش‌اند تا با ارائه سیستم‌های تشخیصی کمک-رایانه‌ای با قابلیت اطمینان بالا، تعداد نمونه‌برداری‌های غیرضروری را کاهش دهند. این سیستم‌ها از چهار بخش پیش‌پردازش، ناحیه‌بندی، استخراج و انتخاب ویژگی، و کلاس‌بندی تشکیل می‌شوند و ابزاری سودمند برای تشخیص سرطان پستان هستند. در این پژوهش به منظور طبقه‌بندی توده‌های پستان به دو گروه خوش‌خیم و بدخیم، پس از پیش‌پردازش تصاویر، به ناحیه‌بندی آنها و تعیین مرز توده، با ترکیب دو رویکرد دستی و کامپیوتری، پرداخته شده است. در مرحله بعد ۸۲۷ ویژگی شامل ۲۴ ویژگی ریخت‌شناسی مبتنی بر شکل و ۸۰۳ ویژگی ریخت‌شناسی مبتنی بر مرز از هر تصویر استخراج شده که ۶۰۴ ویژگی از آنها به تازگی در این پژوهش ارائه شده‌اند. پس از آن با استفاده از کلاسبند رگرسیون لجستیک تنک به حذف ویژگی‌های نامرتب و کلاس‌بندی تصاویر پرداخته شده است. پایگاه داده مورد استفاده در این پژوهش شامل ۱۰۴ تصویر سونوگرافی از توده‌های پستان (۷۲ تصویر مربوط به توده‌های خوش‌خیم و ۳۲ تصویر مربوط به توده‌های بدخیم) است که با اعمال الگوریتم پیشنهادی به این تصاویر، نوع توده با صحت ۸۹/۴۲٪، حساسیت ۷۸/۱۳٪ و دقت ۹۴/۴۴٪ تشخیص داده شده است.

### کلیدواژه‌ها

تصاویر اولتراسوند، رگرسیون لجستیک، سیستم تشخیصی کمک-رایانه‌ای، کلاس‌بندی

است. روش‌های تشخیصی سرطان پستان شامل معاینات بالینی، تصویربرداری پزشکی و نمونه‌برداری است [۱]، که در میان روش‌های تصویربرداری، اولتراسوند به دلیل داشتن مزایایی مثل زمان حقیقی بودن، کم هزینه بودن، عدم استفاده از تشعشعات یونیزان و همچنین داشتن حساسیت بالا در بافت‌های چگال به عنوان یک ابزار مناسب در تشخیص این نوع سرطان استفاده می‌شود [۲-۸]. اما از طرف دیگر ماهیت نویزی تصاویر اولتراسوند، هم‌پوشانی برخی از ویژگی‌های توده‌های خوش‌خیم و بدخیم با یکدیگر و درک متفاوت از تصویر باعث متفاوت بودن تفسیر رادیولوژیست‌ها و در نتیجه تشخیص متفاوت می‌شود، از این رو در اکثر موارد روش نمونه‌برداری، که دقیق‌ترین روش برای

### ۱ مقدمه

با توجه به شیوع بالای سرطان پستان و ناشناخته بودن علت آن، تشخیص زودهنگام این بیماری از اهمیت بالایی برخوردار

این مقاله در مردادماه سال ۱۳۹۵ دریافت، در آبان‌ماه بازنگری و در اسفندماه همان سال پذیرفته شد.

<sup>۱</sup> دانش‌آموخته کارشناسی ارشد دانشگاه تربیت مدرس، دانشکده برق و کامپیوتر، گروه مهندسی پزشکی.

رایانامه: [h.nemat@modares.ac.ir](mailto:h.nemat@modares.ac.ir)

<sup>۲</sup> دانشگاه تربیت مدرس، دانشکده برق و کامپیوتر، گروه مهندسی پزشکی.

رایانامه: [mahlooji@modares.ac.ir](mailto:mahlooji@modares.ac.ir)

رایانامه: [a.gooya@modares.ac.ir](mailto:a.gooya@modares.ac.ir)

<sup>۴</sup> دانشگاه علوم پزشکی تهران، دانشکده پزشکی، گروه رادیولوژی.

رایانامه: [n\\_ahmadinejad@yahoo.com](mailto:n_ahmadinejad@yahoo.com)

نویسنده مسئول: هدا نعمت

وجود اینکه تحلیل بافت برای رادیولوژیست ها نقش مهمی در تشخیص سرطان پستان دارد ولی تشخیص کامپیوتری را به صورت معناداری بهبود نمی بخشد و ترکیب ویژگی های بافت با ویژگی های ریخت شناسی، قدرت تمایز دهی توده های خوش خیم و بدخیم را افزایش نمی دهد. تنها با استفاده از ویژگی های ریخت شناسی هم می توان توده ها را به دو نوع خوش خیم و بدخیم کلاس بندی کرد و نیازی به ترکیب ویژگی های بافت وجود ندارد."

در پژوهش حاضر پس از پیش پردازش تصاویر، به منظور ناحیه بندی آنها و جدا کردن توده از بافت های اطرافش، از رویکردی مبتنی بر ترکیب مرز بندی دستی و ناحیه بندی کامپیوتری استفاده شده است که دقت تعیین مرز توده را تا حد زیادی افزایش می دهد. پس از آن، با ارائه روشی خودکار، به استخراج ناحیه مطلوب پرداخته شده است. در مرحله استخراج ویژگی، مطالعه جامعی بر ویژگی های مختلف ریخت شناسی صورت گرفته و ضمن پیاده سازی مجموعه ای شامل ۲۲۳ ویژگی ارائه شده در مقالات مختلف، ۶۰۴ ویژگی جدید در جهت تکمیل ویژگی های موجود ارائه گشته است. همچنین به جای استفاده از الگوریتم مستقل انتخاب ویژگی، مجموعه ۸۲۷ ویژگی استخراجی به عنوان ورودی کلاس بند اعمال شده، با این تفاوت که با استفاده از کلاس بند رگرسیون لجستیک تنک، خود کلاس بند با کمک الگوریتم تعیین ارتباط خودکار، قادر به حذف ویژگی های نامرتب می باشد. به این ترتیب ضمن حذف خودکار ویژگی های نامرتب و انجام کلاس بندی، مجموعه ای شامل مرتب ترین و متمایز کننده ترین ویژگی ها به منظور تشخیص معرفی گشته است.

## ۲ الگوریتم پیشنهادی

### ۲-۱ پیش پردازش

در این پژوهش در ابتدا و به منظور بهبود کیفیت فرایند ناحیه بندی، به پیش پردازش تصاویر، که شامل بهبود کنتراست و کاهش نویز نقطه ای است، پرداخته شده است. شایان ذکر است تصاویر اصلی از نوع عدد صحیح بی علامت هشت بیتی<sup>۴</sup> (uint8) بوده اند. **بهبود کنتراست:** به منظور بهبود کنتراست تصاویر، در ابتدا شدت سطوح خاکستری تصویر اصلی در محدوده بین صفر تا ۲۵۵ نرمالیزه شده و سپس برای برجسته نمودن ضایعه نسبت به بافت های اطراف، الگوریتم contrast-limited adaptive histogram equalization (CLAHE) به تصاویر اعمال شده است [۱۴].

تشخیص نوع توده می باشد، مورد استفاده قرار می گیرد، اما این روش تهاجمی بوده و باعث تحمیل هزینه های سنگین به بیمار و سیستم بهداشتی و درمان جامعه می شود. از طرفی طبق بررسی های آماری، تنها ۱۰-۳۰ درصد بایوپسی های انجام شده ضروری می باشند؛ لذا محققان به منظور سهولت در امر تشخیص، درصدد ارائه سیستم های تشخیصی کمک-رایانه ای با قابلیت افتراقی بالا و قابل اطمینان هستند.

اساس یک سیستم کمک-رایانه ای تشخیص سرطان پستان، بر پایه مدلسازی دو دسته ویژگی مبتنی بر چگونگی تولید اکو و مبتنی بر شکل و مرز توده می باشد که ویژگی های دسته اول، ویژگی های بافت<sup>۱</sup> و دسته دوم، ویژگی های ریخت شناسی<sup>۲</sup> نامیده می شوند [۹، ۳-۱]. ویژگی های بافت وابسته به دستگاه سونوگرافی و تنظیمات دستگاه می باشند ولی ویژگی های ریخت شناسی براساس طبیعت خود توده بوده و لذا به نوع دستگاه تصویربرداری و تنظیمات آن وابسته نیستند [۱]. تفاوت سیستم های ارائه شده در واقع تفاوت در بخش های مختلف یک سیستم تشخیصی، شامل پیش پردازش، ناحیه بندی، استخراج و انتخاب ویژگی و کلاس بندی، می باشد [۱۰، ۹، ۷، ۶، ۲، ۱].

در مقاله [۱۰] پس از استخراج دستی ناحیه مطلوب<sup>۳</sup> (ROI)، با کمک الگوریتم نیمه خودکار ناحیه بندی مبتنی بر تبدیل حوضچه به تعیین مرز توده پرداخته شده و با استخراج سه ویژگی ریخت شناسی و استفاده از کلاس بند تفکیک کننده خطی (LDA) به طبقه بندی تصاویر پرداخته است.

در مقاله [۱۱] نیز پس از استخراج دستی ROI، با استفاده از فیلتر انتشار ناهمسانگرد به کاهش نویز نقطه ای تصاویر پرداخته شده و با کمک الگوریتم ناحیه بندی مبتنی بر سطوح همتراز، تصاویر ناحیه بندی شده اند و سپس با استخراج شش ویژگی ریخت شناسی و استفاده از کلاس بند ماشین بردار پشتیبان به کمک هسته گوسی به کلاس بندی تصاویر پرداخته است.

در مقاله [۱۲] با استفاده از فیلتر انتشار ناهمسانگرد مبتنی بر ضرایب موجک به کاهش نویز نقطه ای پرداخته شده و سپس با استفاده از الگوریتم ناحیه بندی مبتنی بر سطوح همتراز، توده را از بافت اطرافش جدا کرده و با استخراج شش ویژگی ریخت شناسی و استفاده از الگوریتم پرسپترون چند لایه شبکه عصبی به طبقه بندی توده ها به دو گروه خوش خیم و بدخیم پرداخته است.

در مقاله [۱۳] نیز پس از استخراج دستی ROI، با استفاده از الگوریتم تعمیم یافته کانتور فعال به ناحیه بندی تصاویر پرداخته شده و پس از استخراج شش ویژگی ریخت شناسی و بافتی و استفاده از ماشین بردار پشتیبان به طبقه بندی توده ها پرداخته است.

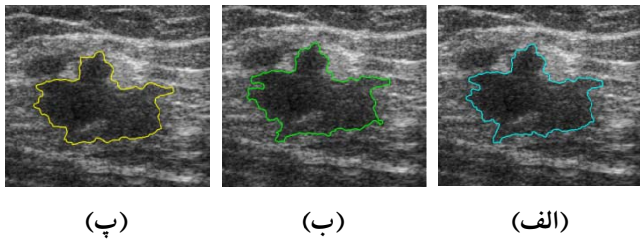
در مقاله [۹] با بررسی ۱۴۶۵ ویژگی بافت و ۲۶ ویژگی ریخت شناسی و بررسی ترکیب این ویژگی ها، بیان شده که: "با

<sup>۱</sup> Texture features

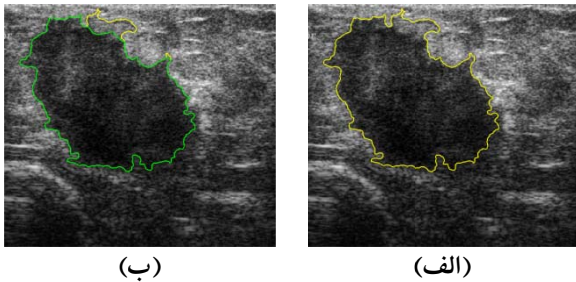
<sup>۲</sup> Morphological features

<sup>۳</sup> Region of interest

<sup>۴</sup> 8-bit unsigned integer



شکل ۲ نمایش سه مرز کامپیوتری یک تومور بدخیم به ازای سه مقدار (الف) 2، (ب) 2.2 و (ج)  $2\pi$  برای پارامتر مرتبط با عرض گوسی تابع محدودیت گوسی الگوریتم ناحیه بندی.



شکل ۳ (الف) تصویر سونوگرافی یک تومور بدخیم پستان و مناسب ترین مرز کامپیوتری آن (زرد رنگ)، (ب) مرز اصلاح شده توده با کمک نرم افزار Didger5 (مرز سبز رنگ) و حذف قسمت های اضافه ای که در مرز کامپیوتری به اشتباه جزء تومور محسوب شده است (قسمت های زرد رنگ)

### ۲-۳ استخراج خودکار ناحیه مطلوب

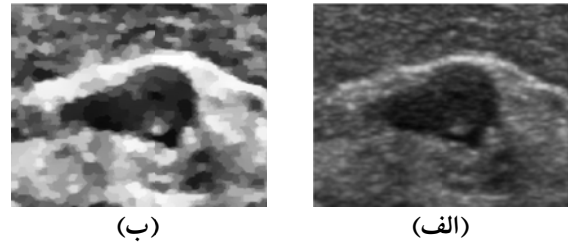
پس از تعیین مرز توده و به منظور حذف نواحی اضافی در تصویر، در این پژوهش الگوریتمی جهت استخراج خودکار ناحیه مطلوب پیشنهاد شده است، به این صورت که ابتدا با استفاده از عملگر اتساع ریخت شناسی<sup>۲</sup> و با بهره گیری از یک المان ساختاری<sup>۳</sup> به شکل دیسک و شعاع ۳۰ پیکسل، یک مرز خارجی برای توده مشخص شده (شکل ۴-ب)، سپس مستطیل محیطی به این مرز خارجی تعیین (شکل ۴-ج) و تصویر در مختصات مربوط به این مستطیل برش داده شده است. شکل ۴-الف تصویر اصلی و شکل ۴-د ناحیه مطلوب استخراج شده از این تصویر را نشان می دهد.

### ۲-۴ استخراج ویژگی

توده های پستان در تصاویر اولتراسوند دارای شکل و الگوی بافتی متفاوتی هستند. بررسی ها نشان داده که غالب توده های خوش خیم دارای شکلی بیضوی با لبه های مشخص و بدون زاویه اند و در جهت موازی با پوست قرار دارند، در حالی که توده های بدخیم اکوهای غیریکنواخت داشته و دارای شکل های نامنظم با مرزهای نامشخص و زاویه دار هستند [۱]. ویژگی های ریخت شناسی، که بر مبنای ویژگی های ناحیه ای توده و مرزهای آن تعریف می گردند، از اطلاعات شکل و یا حاشیه ضایعه استفاده

کاهش نویز نقطه ای: به منظور نویز زدایی تصاویر، الگوریتمی مبتنی بر فیلتر انتشار ناهمسانگرد<sup>۱</sup> که در مقاله [۱۴] ارائه شده، مورد استفاده قرار گرفته است.

شکل ۱ تصویر اصلی و تصویر پیش پردازش شده مربوط به یکی از تصاویر پایگاه داده را نمایش می دهد.



شکل ۱ (الف) تصویر اصلی سونوگرافی یک توده خوش خیم پستان، (ب) تصویر پیش پردازش شده

### ۲-۲ ناحیه بندی

در مقالات مرتبط به منظور جدا کردن تومور از بافت اطرافش، یا از روش مرز بندی دستی تومور توسط رادیولوژیست و یا از یک روش خودکار و یا نیمه خودکار ناحیه بندی کامپیوتری استفاده شده است. اما استفاده تنها از هریک از این رویکردها مشکلاتی به همراه دارد، از جمله اینکه در مرز بندی دستی تومور ممکن است جزئیات مربوط به مرز به خصوص در تومورهای بدخیم، که برای مرحله استخراج ویژگی اهمیت زیادی دارند، از بین برود، همچنین هیچ یک از الگوریتم های ناحیه بندی کامپیوتری قادر به تعیین دقیق مرز در همه تصاویر نبوده و مرز مشخص شده کاملاً منطبق بر نظر متخصص رادیولوژیست نیست. به این ترتیب استنباط می شود که ترکیب این روش ها بتواند بهترین نتیجه را برای ناحیه بندی تصاویر به دست دهد. از این رو در این پژوهش با استفاده از ترکیب هر دو روش دستی و کامپیوتری به تعیین مرز توده پرداخته شده است، به این صورت که ابتدا توسط الگوریتم نیمه خودکار کامپیوتری مبتنی بر تبدیل حوضچه و با تغییر مقدار پارامتر مربوط به عرض گوسی تابع محدودیت گوسی هم مرکز با ضایعه، که در اولین مرحله در روند ناحیه بندی در تصویر فیلتر شده ضرب شده است [۱۴]، برای هر تومور سه مرز مشخص گردیده است. (مقادیر مربوط به این پارامتر به پیشنهاد آقای Gómez Flores، نویسنده اول مقاله [۱۴] و طی مکاتبه با ایشان انتخاب شده اند.) شکل ۲ سه مرز کامپیوتری حاصل از تغییر این پارامتر را برای یک تومور بدخیم پستان نشان می دهد. پس از تعیین سه مرز کامپیوتری، هر یک از این مرزها توسط متخصص رادیولوژیست با تجربه به دقت مورد بررسی قرار گرفته و در صورت نیاز قسمت هایی از مرز با کمک نرم افزار Didger 5 ویرایش شده و مرز اصلاح شده به عنوان مرز تومور مشخص گردیده است (شکل ۳).

<sup>2</sup>Morphological dilation operator

<sup>3</sup>Structure element

<sup>1</sup>Anisotropic diffusion filtering

(۷) جمود<sup>۷</sup> [۱، ۹-۱۱، ۱۳، ۱۶، ۱۷، ۱۹، ۲۰]. (رابطه مربوط به این ویژگی ویژگی در مراجع [۹، ۱۰، ۱۹]. با عنوان نسبت هم پوشانی<sup>۸</sup> بیان شده است.)

(۸) جهت<sup>۹</sup> [۱، ۹].

(۹) گریز از مرکز<sup>۱۰</sup> [۱، ۱۳].

(۱۰) کشیدگی<sup>۱۱</sup> [۱، ۱۷، ۲۱].

(۱۱) اختلاف مساحت<sup>۱۲</sup> [۱، ۱۲].

(۱۲) مقدار باقیمانده نرمال شده<sup>۱۳</sup> [۹، ۱۰].

(۱۳) فشردگی<sup>۱۴</sup> [۲، ۱۰، ۲۱، ۲۲].

(این ویژگی در مراجع [۲، ۱۰، ۲۱] با عنوان مدوری<sup>۱۵</sup> مطرح شده است.)

(۱۴) شاخص تحدب<sup>۱۶</sup> [۲۳].

(۱۵) نسبت محدوده مدوری<sup>۱۷</sup> [۲۳].

(۱۶) نسبت بی نظمی<sup>۱۸</sup> [۲۳].

(ویژگی های ۱۶-۱۴ برای تشخیص نوع تومور در تصاویر ماموگرافی ارائه شده اند.)

(۱۷) نسبت بسته ریخت شناسی<sup>۱۹</sup> [۱۰].

(۱۸) اسکلت نرمال شده بیضوی<sup>۲۰</sup> [۲، ۹، ۱۸].

(۱۹) تعداد نقاط انتهایی اسکلت<sup>۲۱</sup> [۹].

(۲۰) مقدار میانگین تابع تغییرات<sup>۲۲</sup>

(۲۱) مقدار واریانس تابع تغییرات<sup>۲۳</sup>

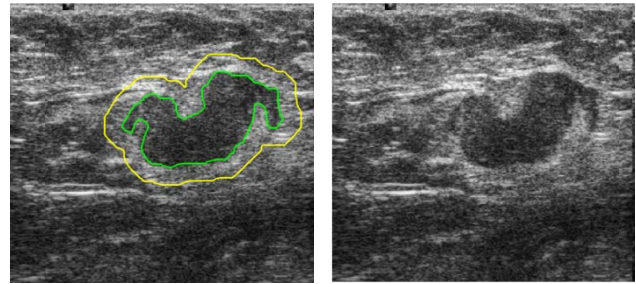
(۲۲) مقدار چولگی (عدم تقارن) تابع تغییرات<sup>۲۴</sup>

(۲۳) مقدار کشیدگی تابع تغییرات<sup>۲۵</sup>

(۲۴) مقدار بی نظمی تابع تغییرات<sup>۲۶</sup> [۱، ۱۲].

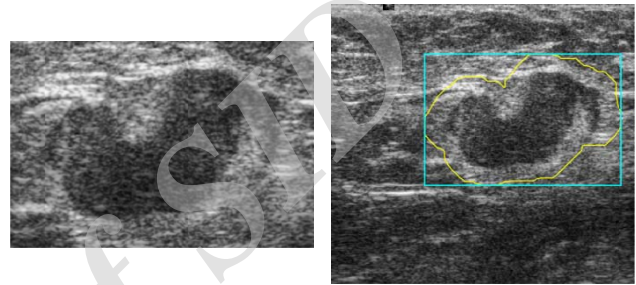
(برای به دست آوردن تابع تغییرات کفایت قسمتی از تصویر، که شامل توده است، در زوایای مختلف با پله های ۵-درجه ای چرخانده شود سپس در هر زاویه پیکسل هایی با کمترین و بیشترین مقدار  $y$  یافت شده و تصویر خط واصل این پیکسل ها روی محور  $y$  محاسبه شود. این تابع به نوعی بیانگر میزان گسترش توده در

می کنند و روی شاخص های محلی تومور تمرکز دارند. این ویژگی ها را می توان در دو دسته کلی ویژگی های مبتنی بر شکل و ویژگی های مبتنی بر مرز توده قرار داد [۳، ۵، ۹، ۱۵]. در ادامه به بررسی ویژگی های هر یک از این گروه ها پرداخته می شود.



(ب)

(الف)



(ت)

(پ)

شکل ۴ (الف) تصویر اصلی شامل توده و نواحی اطراف آن، (ب) مرز توده (سبز رنگ) و مرز خارجی در فاصله ۳۰-پیکسلی از آن (زرد رنگ)، (پ) مرز خارجی توده (زرد رنگ) و مستطیل محیط به آن، (ت) ناحیه مطلوب استخراج شده از تصویر اولیه.

## ۲-۴-۱ ویژگی های ریخت شناسی مبتنی بر شکل توده

به منظور بررسی ویژگی های ریخت شناسی مبتنی بر شکل توده، ۲۴ ویژگی از این گروه، که ۲۱ ویژگی آن مربوط به ویژگی های ارائه شده برای تشخیص نوع تومور در تصاویر اولتراسوند و سه ویژگی آن مربوط به این تشخیص در تصاویر ماموگرافی بوده، پیاده سازی شده اند، که این ویژگی ها به قرار زیر می باشند:

(۱) ضریب شکل<sup>۱</sup> [۱، ۹، ۱۱، ۱۶].

(۲) نسبت قطر بزرگ به قطر کوچک<sup>۲</sup> [۱، ۲، ۹، ۱۱، ۱۶، ۱۷].

(۳) نسبت محور بزرگ به محور کوچک<sup>۳</sup> [۱، ۲، ۱۶-۱۸].

(۴) گردی<sup>۴</sup> [۱، ۹، ۱۱، ۱۶، ۱۷].

(۵) تحدب<sup>۵</sup> [۱، ۹، ۱۱، ۱۶، ۱۷، ۱۹، ۲۰].

(۶) وسعت<sup>۱</sup> [۱، ۹، ۱۱، ۱۶، ۱۷، ۲۰].

<sup>7</sup>Solidity

<sup>8</sup>Overlap-ratio

<sup>9</sup>Orientation

<sup>10</sup>Eccentricity

<sup>11</sup>Elongation

<sup>12</sup>Difference area

<sup>13</sup>Normalized residual value

<sup>14</sup>Compactness

<sup>15</sup>Circularity

<sup>16</sup>Convexity index

<sup>17</sup> Circularity range ratio

<sup>18</sup> Irregularity ratio

<sup>19</sup> Morphological-closing ratio

<sup>20</sup> Elliptic-normalized skeleton

<sup>21</sup> Number of skeleton end-points

<sup>22</sup>Mean-variation

<sup>23</sup>Variance-variation

<sup>24</sup> Skewness-variation

<sup>25</sup>Kurtosis-variation

<sup>26</sup>Entropy-variation

<sup>1</sup>Form-factor

<sup>2</sup>Aspect-ratio

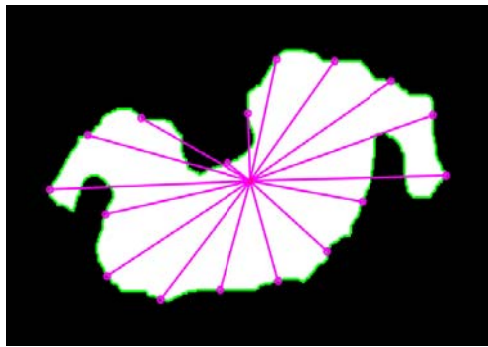
<sup>3</sup> Long axis to short axis ratio

<sup>4</sup> Roundness

<sup>5</sup>Convexity

<sup>6</sup>Extent



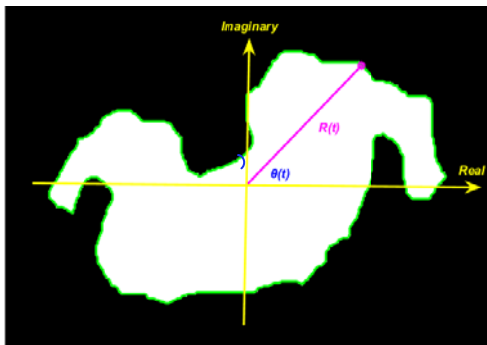


شکل ۵ مرکز توده (سبز رنگ)، مرکز توده (\*) و فاصله شعاعی (ارغوانی رنگ) که معادل فاصله نقاط مرکز از مرکز توده می باشد.

### ۳) مختصات قطبی (PC)

مشخصه مختصات قطبی، از ترکیب مختلط مشخصه فاصله شعاعی  $R(t)$  و زاویه قطبی  $\theta(t)$  مطابق رابطه ۳ به دست می آید (شکل ۶) [۲۷].

$$PC(t) = R(t) + j\theta(t) \quad (۳)$$



شکل ۶ نمایش مختصات قطبی نقاط مرکز (سبز رنگ) به صورت فاصله شعاعی  $R(t)$  (ارغوانی رنگ) و زاویه قطبی  $\theta(t)$  (آبی رنگ)

### ۴) فاصله از دورترین نقطه (FPD)

مقدار این مشخصه برای هر نقطه از مرکز به صورت مجموع فواصل شعاعی مربوط به آن نقطه و دورترین نقطه مرکز نسبت به آن نقطه به دست می آید. به عبارتی اگر بخواهیم این مقدار را برای نقطه  $(x(t), y(t))$  به دست بیآوریم، ابتدا باید نقطه  $(x_{fp}(t), y_{fp}(t))$  که دورترین نقطه نسبت به  $(x(t), y(t))$  است را بیابیم (شکل ۷) سپس مطابق رابطه ۴ فاصله شعاعی این دو نقطه را با هم جمع کنیم [۲۶، ۲۷، ۳۲، ۳۵].

$$FPD(t) = \sqrt{(x(t) - x_c)^2 + (y(t) - y_c)^2} + \sqrt{(x_{fp}(t) - x_c)^2 + (y_{fp}(t) - y_c)^2} \quad (۴)$$

بافت های اطراف بوده و نامنظمی و شاخه شاخه بودن توده را در جهات مختلف نشان می دهد).

### ۲-۴-۲ ویژگی های ریخت شناسی مبتنی بر مرکز توده

با بررسی ویژگی های ریخت شناسی مبتنی بر مرکز ارائه شده در مقالات، این گونه استنباط شد که همه این ویژگی ها بر اساس یک مشخصه<sup>۱</sup> از مرکز توده هستند، که این مشخصه در واقع بیان یک بعدی مرکز می باشد. مشخصه هایی که به این منظور در مقالات مربوطه استفاده شده اند، مشخصه های فاصله شعاعی و مختصات مختلط بوده اند. پس از آن با بررسی روش های یک بعدی سازی ارائه شده در مقالات مربوط به بازیابی شکل<sup>۲</sup>، هفت مشخصه دیگر نیز نیز پیاده سازی شده و در مجموع نه مشخصه مورد بررسی قرار گرفته و به بررسی ویژگی های تعمیم داده شده به این مشخصه ها پرداخته شده است.

#### ➤ مشخصه های مرکز توده

به طور کلی هر بازنمایی یک بعدی از مرکز یا ناحیه دوبعدی که بتواند یک شکل را به طور یکتا بیان کند، مشخصه شکل نامیده می شود [۲۸-۲۴، ۲۰]. در این قسمت به منظور بررسی جامع اطلاعات ریخت شناسی مبتنی بر مرکز توده، ۹ مشخصه یک بعدی از مرکز هر توده استخراج شده است که این مشخصه ها به قرار زیر می باشند: (در ادامه فرض می کنیم مختصات هر یک از نقاط مرکز توده به صورت  $(x(t), y(t))$ ،  $t = 1, 2, \dots, N$  است که  $N$  تعداد این نقاط می باشد و مختصات مرکز توده به صورت  $(x_c, y_c)$  بیان می شود.)

#### ۱) مختصات مختلط<sup>۳</sup> (Z)

تابع مختصات مختلط به صورت عدد مختلطی از مختصات نقاط مرکز بیان می گردد که به منظور حذف اثر بایاس، از مختصات شیفت یافته نسبت به مرکز توده استفاده می شود [۳۳-۲۴، ۲۰]. (رابطه ۱)

$$Z(t) = (x(t) - x_c) + j(y(t) - y_c) \quad (۱)$$

#### ۲) فاصله شعاعی<sup>۴</sup> (R)

مقدار این تابع مطابق شکل ۵ به صورت فاصله هر نقطه مرکز از مرکز توده به دست می آید [۳۵-۳۳، ۳۱، ۳۰، ۲۴-۲۸، ۲۰]. (رابطه ۲)

$$R(t) = \sqrt{(x(t) - x_c)^2 + (y(t) - y_c)^2} \quad (۲)$$

<sup>1</sup> Signature

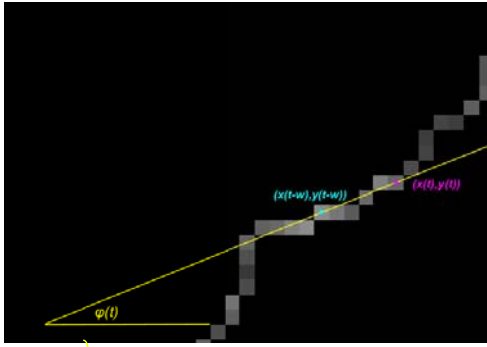
<sup>2</sup> Shape retrieval

<sup>3</sup> Complex coordinate

<sup>4</sup> Radial distance

<sup>5</sup> Polar coordinate

<sup>6</sup> Farthest point distance



شکل ۹ تابع زاویه‌ای  $\varphi(t)$  محاسبه شده در پنجره‌ای به طول  $w = 5$

(۷) مختصات زاویه شعاعی<sup>۳</sup>

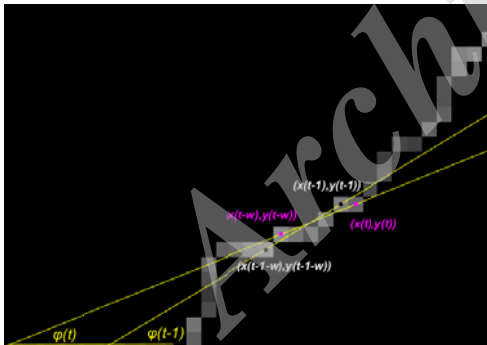
این مشخصه مانند مشخصه مختصات قطبی است با این تفاوت که مطابق رابطه ۷ به جای زاویه قطبی  $\theta(t)$ ، از تابع زاویه‌ای  $\varphi(t)$  استفاده می‌شود [۲۷].

$$ARC(t) = R(t) + j\varphi(t) \quad (۷)$$

(۸) خمیدگی مرز<sup>۴</sup>

خمیدگی، دومین مشتق مرز و اولین مشتق مماس بر مرز را بیان می‌کند. تابع خمیدگی، تابع تغییرات زاویه‌ای خطوط مماس بر مرز توده می‌باشد و به صورت تفاضل زوایای خطوط مماس محاسبه شده در یک پنجره به طول  $w$  برای دو نقطه متوالی از مرز (مطابق رابطه ۸) به دست می‌آید (شکل ۱۰) [۲۴-۲۶، ۲۸، ۳۲، ۳۵-۳۸].

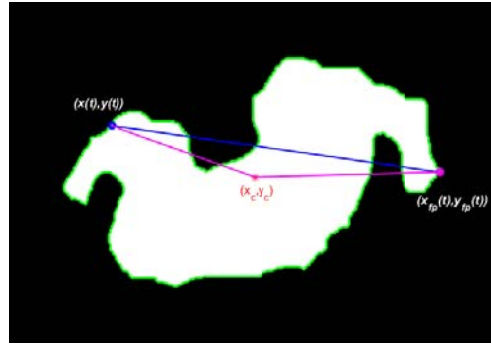
$$K(t) = \varphi(t) - \varphi(t-1) \quad (۸)$$



شکل ۱۰ زوایای مماسی  $\varphi(t)$  و  $\varphi(t-1)$  محاسبه شده در پنجره‌ای به طول  $w = 5$  به منظور محاسبه خمیدگی مرز  $(\varphi(t) - \varphi(t-1))$

(۹) تابع مساحت سه‌گوش<sup>۵</sup> (TAF)

مقدار این تابع مطابق شکل ۱۱ بر اساس مساحت مربوط به سه‌گوش متشکل از دو نقطه متوالی مرز و مرکز توده به دست می‌آید [۲۰، ۲۴، ۲۷، ۳۲].

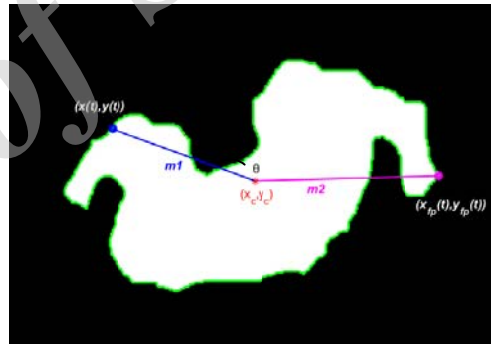


شکل ۷ خط واصل نقطه  $(x(t), y(t))$  و دورترین نقطه نسبت به آن  $(x_{fp}(t), y_{fp}(t))$  (آبی‌رنگ) و فواصل شعاعی این دو نقطه (ارغوانی‌رنگ)

(۵) زاویه با دورترین نقطه<sup>۱</sup> (FPA)

زاویه میان طول شعاعی یک نقطه از مرز  $(m_1)$  و طول شعاعی دورترین نقطه از نقاط مرز نسبت به آن نقطه  $(m_2)$  مقدار این تابع را به دست می‌دهد (شکل ۸). مقدار این زاویه مطابق رابطه ۵ به دست می‌آید [۳۵].

$$FPA(t) = \tan^{-1} \left( \frac{m_2 - m_1}{1 + m_1 \times m_2} \right) \quad (۵)$$



شکل ۸ زاویه میان خطوط شعاعی مربوط به نقطه  $(x(t), y(t))$  (آبی‌رنگ) و دورترین نقطه نسبت به آن  $(x_{fp}(t), y_{fp}(t))$  (ارغوانی‌رنگ)

(۶) تابع زاویه‌ای<sup>۲</sup>

تابع زاویه‌ای  $\varphi(t)$  تغییرات جهت‌های مرز را بیان می‌کند. در واقع این تابع مطابق شکل ۹ زاویه مماسی را برای پنجره‌ای به طول  $w$  و به صورت رابطه ۶ به دست می‌دهد. البته تابع زاویه  $\varphi(t)$  تنها می‌تواند مقادیر را در محدوده‌ای به طول  $2\pi$  فرض کند که این محدوده معمولاً در بازه  $[0, 2\pi]$  یا  $[-\pi, \pi]$  است، بنابراین  $\varphi(t)$  در حالت کلی شامل ناپیوستگی‌هایی به اندازه  $2\pi$  می‌باشد [۲۰، ۲۴، ۲۵، ۲۷، ۳۶].

$$\varphi(t) = \tan^{-1} \left( \frac{y(t) - y(t-w)}{x(t) - x(t-w)} \right) \quad (۶)$$

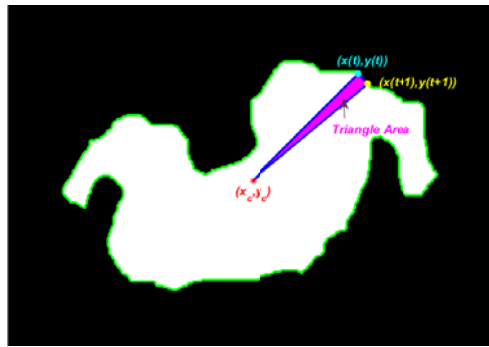
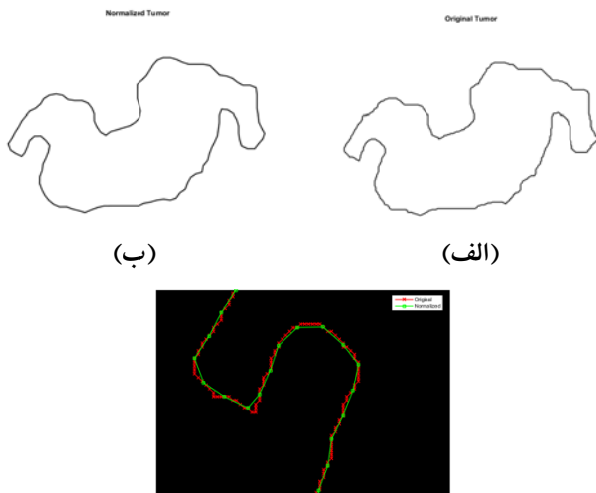
<sup>3</sup> Angular radial coordinate

<sup>4</sup> Contour curvature

<sup>5</sup> Triangle area function

<sup>1</sup> Farthest point angel

<sup>2</sup> Angular function



شکل ۱۱ مرز توده (سبز رنگ) و مقدار تابع مساحت سه‌گوش (ناحیه ارغوانی رنگ) حاصل از اتصال مرکز توده (\*) و دو نقطه متوالی مرز (نقاط آبی رنگ و زرد رنگ)

### ➤ ویژگی‌های مبتنی بر اطلاعات مکانی مشخصه‌های مرز

به منظور استفاده از اطلاعات مکانی مشخصه‌ها، مقادیر ویژگی‌های میانگین، انحراف معیار، شاخص همواری، شاخص ناهمواری و عبور از صفر [۱، ۱۰، ۲۱] برای چهار مشخصه حقیقی فاصله شعاعی، فاصله از دورترین نقطه، زاویه با دورترین نقطه و تابع مساحت سه‌گوش محاسبه شده است. به این ترتیب در مجموع ۲۰ ویژگی ریخت‌شناسی مبتنی بر اطلاعات مکانی مشخصه‌های مرز ارائه شده است.

### ➤ ویژگی‌های مبتنی بر اطلاعات فرکانسی مشخصه‌های

مرز

به منظور بررسی اطلاعات فرکانسی مرز و استخراج ویژگی‌های توصیف‌گر فوری، ابتدا تعداد نقاط توصیف‌کننده مرز توده نرمال و پس از آن با اعمال تبدیل فوری گسسته<sup>۱</sup> (DFT)، از ضرایب فوری و همچنین ویژگی‌های آماری این ضرایب به عنوان ویژگی برای تمایز میان تومورهای خوش‌خیم و بدخیم استفاده شده است.

### ✓ نرمال‌سازی نقاط مرز

از آنجا که تعداد نقاط بیان‌کننده مرز توده در عملکرد بهینه تبدیل فوری اهمیت زیادی دارد، از این‌رو نمونه‌برداری از نقاط مرز، یک گام ضروری قبل از استخراج مشخصه‌های مرز است [۲۴، ۲۵]. اگر  $N$  تعداد کل نقاط بیان‌کننده مرز توده باشد، که این نقاط به روش ردیابی ۸-اتصال مرز<sup>۲</sup> استخراج شده‌اند، و  $L = 2^7 = 128$  تعداد نقاط مطلوب ما پس از نمونه‌برداری باشد، نقاط انتخابی در فواصل طول کمان برابر با  $P/L$  قرار می‌گیرند که  $P$  محیط مرز توده می‌باشد [۲۵، ۲۹، ۳۰، ۳۵]. نقاط انتخابی باید بتوانند در حین حذف نویز، جزئیات شکل توده را به خوبی حفظ کنند. شکل ۱۲ مرز توده را در دو حالت اصلی و نرمال‌شده و نمای نزدیکی از نقاط این مرزها را نشان می‌دهد.

(ب)

شکل ۱۲ (الف) مرز اصلی توده با تعداد  $(N = 838)$  نقطه که این نقاط به روش ردیابی ۸-اتصال مرز به دست آمده‌اند، (ب) مرز نرمال‌شده توده با تعداد  $(L = 128)$  نقطه، (پ) نمای نزدیک نقاط مرز اصلی توده (قرمز رنگ) و مرز نرمال‌شده (سبز رنگ)

### ✓ توصیف‌گرهای فوری

ضرایب تبدیل فوری گسسته، توصیف‌گرهای فوری نامیده می‌شوند و در واقع این توصیف‌گرها بیان مجددی از شکل در حوزه فرکانس هستند. توصیف‌گرهای فرکانس پایین‌تر، شامل اطلاعاتی درباره ویژگی‌های کلی شکل و توصیف‌گرهای فرکانس بالاتر، شامل اطلاعاتی در مورد جزئیات شکل هستند.

فرض کنیم  $t = 0, 1, \dots, L-1, s(t)$  یکی از مشخصه‌های بیان شده است که مرز توده را پس از نرمال‌سازی  $L$ -نقطه‌ای به صورت یک‌بعدی بیان می‌کند، تبدیل فوری گسسته  $s(t)$  توسط رابطه ۹ به دست می‌آید:

$$a_n = \frac{1}{L} \sum_{t=0}^{L-1} s(t) \exp\left(\frac{-j2\pi nt}{L}\right), \quad n=0, 1, \dots, L-1 \quad (9)$$

ضرایب  $n = 0, 1, \dots, L-1, a_n$  توصیف‌گرهای فوری نامگذاری می‌شوند [۳۳-۳۵، ۳۰-۳۲، ۲۶، ۲۵، ۲۰].

#### • نامتغیر بودن نسبت به جابجایی<sup>۳</sup>

از آنجا که تمام مشخصه‌های بیان شده نسبت به جابجایی نامتغیرند، بنابراین توصیف‌گرهای فوری مربوط به آنها نیز نسبت به جابجایی نامتغیر هستند [۳۰، ۲۸، ۲۵، ۲۰].

#### • نامتغیر بودن نسبت به چرخش<sup>۴</sup>

در حالت عادی نامتغیر بودن نسبت به چرخش توسط نرمال‌سازی فاز به دست می‌آید، اما از آنجا که در این کاربرد از اطلاعات فاز توصیف‌گرها استفاده نمی‌شود و تنها مقادیر اندازه آنها مورد استفاده قرار می‌گیرد، نامتغیر بودن نسبت به چرخش برای این ویژگی‌ها برقرار است [۳۰، ۲۸، ۲۵].

<sup>3</sup> Translation invariant

<sup>4</sup> Rotation invariant

<sup>1</sup> Discrete fourier transform

<sup>2</sup> 8-connectivity contour tracing procedure

روش تعیین ارتباط خودکار (ARD)<sup>۳</sup> در مرجع [۳۹] تعمیم داده شده است. در ادامه به بررسی این کلاس بند تک پرداخته می شود. در این بخش اعداد به صورت حروف کج (مانند:  $x$ ) و بردارها به صورت درشت و سیاه (مانند:  $\mathbf{x}$ ) نمایش داده شده اند.

## ۲-۵-۱ رگرسیون لجستیک (LR)

LR یک مدل احتمالی<sup>۴</sup> برای حل مسائل کلاس بندی باینری است که پارامترهای آن توسط روش درست نمایی بیشینه<sup>۵</sup> تخمین زده می شوند. تابع خطی تفکیک کننده<sup>۶</sup>، که دو کلاس S1 و S2 را جدا می کند، توسط رابطه ۱۲ به دست می آید.

$$f(\mathbf{x}; \mathbf{w}) = \sum_{i=1}^D w_i x_i + w_0 = \mathbf{w}^t \mathbf{x} \quad (12)$$

به طوری که  $\mathbf{x} = (x_1, \dots, x_D) \in R^D$  بردار ویژگی ورودی در فضای D-بعدی و  $\mathbf{w} = (w_1, \dots, w_D)$  بردار وزن در همین فضا می باشد. در این کلاس بند، ابرصفحه<sup>۷</sup>  $f(\mathbf{x}; \mathbf{w}) = 0$  مرز میان دو کلاس را مشخص می کند. LR با استفاده از یک تابع لجستیک<sup>۸</sup> (رابطه ۱۳) (رابطه ۱۳) احتمال تعلق بردار ویژگی ورودی را به هر یک از کلاس های S1 و S2 مطابق رابطه ۱۴ محاسبه می کند.

$$\sigma(x) = \frac{1}{1 + \exp(-x)} \quad (13)$$

$$p = \frac{1}{1 + \exp(-\mathbf{w}^t \mathbf{x})} \equiv P(S2|\mathbf{x}), P(S1|\mathbf{x}) = 1 - p \quad (14)$$

برای فرمول بندی ریاضی، متغیر باینری  $y$  را به گونه ای تعریف می کنیم که  $y = 0$  برای کلاس S1 و  $y = 1$  برای کلاس S2 باشد. اگر ورودی-خروجی  $N$  داده نمونه را به صورت  $\{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_N, y_N)\}$  داشته باشیم، تابع درست نمایی بر اساس رابطه ۱۵ به دست می آید.

$$P(\mathbf{y}|\mathbf{X}; \mathbf{w}) = P(y_1, \dots, y_N | \mathbf{x}_1, \dots, \mathbf{x}_N; \mathbf{w}) \quad (15)$$

$$= \prod_{n=1}^N P(y_n | \mathbf{x}_n; \mathbf{w})$$

$$= \prod_{n=1}^N p_n^{y_n} (1 - p_n)^{1 - y_n}$$

$$; p_n \equiv P(y_n = 1 | \mathbf{x}_n; \mathbf{w}) = \frac{1}{1 + \exp(-\mathbf{w}^t \mathbf{x}_n)}$$

در نهایت بردار وزن  $\mathbf{w}$  مطابق رابطه ۱۶ با بیشینه کردن لگاریتم این تابع درست نمایی حاصل می شود.

$$l(\mathbf{w}) = \sum_{i=1}^N [y_n \log p_n + (1 - y_n) \log(1 - p_n)] \quad (16)$$

شایان ذکر است که این بهینه سازی به دلیل مثبت معین بودن ماتریس هشین<sup>۹</sup> در همه جا، به یک مقدار بیشینه یکتا همگرا می شود. پس به این ترتیب اگر برای نمونه آزمایشی  $\mathbf{x}_{test}$  با کلاس نامشخص،  $f(\mathbf{x}; \mathbf{w}) > 0$  باشد؛ این داده به کلاس S2 و اگر  $f(\mathbf{x}; \mathbf{w}) < 0$  باشد، به کلاس S1 تعلق می گیرد.

## • نامتغیر بودن نسبت به مقیاس<sup>۱</sup>

برای مشخصه های مختصات مختلط، مختصات قطبی و مختصات زاویه شعاعی به دلیل مختلط بودن مقادیر آنها، به منظور بیان توده از تمام توصیف گرهای به جز اولین توصیف گر (مؤلفه DC) استفاده می شود. با تقسیم مقادیر اندازه سایر توصیف گرهای به اندازه دومین توصیف گر، نامتغیر بودن نسبت به مقیاس ایجاد می شود. بردار ویژگی حاصل از مجموعه این توصیف گرهای توسط رابطه ۱۰ به دست می آید [۲۵، ۲۷-۲۹، ۳۳].

$$\mathbf{f} = \left[ \frac{|FD_2|}{|FD_1|}, \frac{|FD_3|}{|FD_1|}, \dots, \frac{|FD_{L-1}|}{|FD_1|} \right] \quad (10)$$

برای مشخصه های فاصله شعاعی، فاصله از دورترین نقطه، زاویه با دورترین نقطه، تابع زاویه ای، خمیدگی مرز و تابع مساحت سه گوش به دلیل حقیقی بودن توابع آنها و این که تنها  $L/2$  فرکانس ها در تبدیل فوریه متفاوت هستند، تنها نیمی از توصیف گرهای مورد نیاز است و نامتغیر بودن نسبت به مقیاس با تقسیم مقادیر اندازه نیمه اول این توصیف گرهای به اندازه دومین توصیف گر به دست می آید (مطابق رابطه ۱۱) [۲۵-۲۸، ۳۰، ۳۲].

$$\mathbf{f} = \left[ \frac{|FD_2|}{|FD_1|}, \frac{|FD_3|}{|FD_1|}, \dots, \frac{|FD_{L/2}|}{|FD_1|} \right] \quad (11)$$

## ✓ ویژگی های آماری توصیف گرهای فوریه

پس از محاسبه ضرایب فوریه مربوط به هر مشخصه، ویژگی های آماری میانگین، واریانس و آنتروپی مربوط به این ضرایب محاسبه شده اند که در مجموع برای ۹ مشخصه مورد استفاده، ۲۷ ویژگی حاصل شده است.

در مجموع در این بخش، از هر تصویر ۸۲۷ ویژگی ریخت شناسی استخراج گردیده که ۲۴ ویژگی مبتنی بر شکل توده و ۸۰۳ ویژگی مبتنی بر مرز توده (۷۸۳ ویژگی مبتنی بر اطلاعات فرکانسی و ۲۰ ویژگی مبتنی بر اطلاعات مکانی) بوده است.

## ۲-۵ انتخاب ویژگی و کلاس بندی

پس از استخراج مجموعه ای شامل ۸۲۷ ویژگی، حال باید تعدادی از مناسب ترین این ویژگی ها انتخاب شود، در مقالات مشابهی که با هدف کلاس بندی نوع توده، تعداد زیادی ویژگی را مورد بررسی قرار داده اند، این کار توسط یک روش انتخاب ویژگی انجام شده است. به این ترتیب که تعدادی از ویژگی های استخراج شده توسط یک الگوریتم انتخاب ویژگی و به صورت مستقل از سایر مراحل، انتخاب و به کلاس بند داده شده اند. در پژوهش حاضر مرحله مجزای انتخاب ویژگی حذف شده و تمام ویژگی های استخراج شده به کلاس بند داده شده اند با این تفاوت که خود کلاس بند قادر به حذف خودکار ویژگی های نامرتب بوده و مرحله انتخاب ویژگی به طور هم زمان با کلاس بندی انجام می شود. کلاس بندی که در این پژوهش مورد استفاده قرار گرفته است، مدلی از رگرسیون لجستیک (LR) است که بر اساس چارچوب بیسی<sup>۲</sup> با استفاده از

<sup>3</sup>Automatic relevance determination

<sup>4</sup>Probabilistic model

<sup>5</sup>Maximum likelihood

<sup>6</sup>Linear discriminant function

<sup>7</sup>Hyper-plane

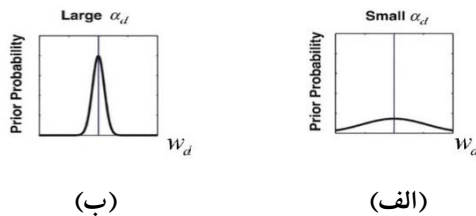
<sup>8</sup>Logistic function

<sup>9</sup>Hessian

<sup>1</sup> Scale invariant

<sup>2</sup>Bayesian framework





(ب)

(الف)

شکل ۱۳ توزیع‌های پیشین پارامتر وزن  $w_d$  و چگونگی کنترل این پارامتر توسط پارامتر ارتباط  $\alpha_d$ ، (الف) مقدار کوچک  $\alpha_d$  که منجر به  $w_d$  بزرگ می‌شود، (ب) مقدار بزرگ  $\alpha_d$  که منجر به  $w_d$  صفر می‌شود [۳۹]

### ۳ ارزیابی

#### ۱-۳ پایگاه داده تصاویر

پایگاه داده مورد استفاده در این مطالعه، شامل ۱۰۴ تصویر سونوگرافی توده‌های پستان (۷۲ تصویر مربوط به توده‌های خوش‌خیم و ۳۲ تصویر مربوط به توده‌های بدخیم) است. این تصاویر از بیماران مراجعه‌کننده به سه مرکز تصویربرداری جمع‌آوری شده‌اند:

(۱) مرکز تصویربرداری بیمارستان امام خمینی: ۵۶ تصویر سونوگرافی توده‌های پستان (۳۴ تصویر مربوط به توده‌های خوش‌خیم و ۲۲ تصویر مربوط به توده‌های بدخیم) با فرمت DICOM که با استفاده از دستگاه سونوگرافی کمپانی زیمنس آلمان، مدل Antares با فرستنده خطی VFX 5-13M به دست آمده‌اند.

(۲) مرکز تصویربرداری دزاشیب: ۳۵ تصویر سونوگرافی توده‌های پستان (۲۷ تصویر مربوط به توده‌های خوش‌خیم و ۸ تصویر مربوط به توده‌های بدخیم) با فرمت jpg که با استفاده از دستگاه سونوگرافی کمپانی اولتراسونیکس آلمان، مدل Sonix OP با فرستنده خطی 5-14 Mhz به دست آمده‌اند.

(۳) مرکز تصویربرداری دکتر گیتی: ۱۳ تصویر سونوگرافی توده‌های پستان (۱۱ تصویر مربوط به توده‌های خوش‌خیم و ۲ تصویر مربوط به توده‌های بدخیم) با فرمت tiff که با استفاده از دستگاه سونوگرافی کمپانی جنرال الکتریک آمریکا با فرستنده خطی SP6-12 به دست آمده‌اند.

در پژوهش حاضر کلاس واقعی (تشخیص بالینی) توده‌ها در تصاویر سونوگرافی بر اساس انجام بایوپسی و نتایج آزمایش‌های پاتولوژی و یا دنبال کردن برنامه زمانی معاینه دو ساله مشخص گردیده است. به این صورت که پس از اخذ تصاویر سونوگرافی توسط متخصص رادیولوژیست با تجربه، توده‌های مشکوک به بدخیمی، بایوپسی شده‌اند و بافت‌های حاصل از نمونه‌برداری این توده‌ها، به منظور انجام مطالعات میکروسکوپی و تشخیص خوش‌خیمی یا بدخیمی آنها، به آزمایشگاه پاتولوژی فرستاده شده‌اند؛ در آزمایشگاه پاتولوژی پس از انجام مراحل آسیب‌شناسی، نوع توده توسط متخصص پاتولوژی تشخیص داده شده است. در مورد تعدادی از توده‌های خوش‌خیم نیز که نیاز به انجام بایوپسی نبوده، بیمار هر شش ماه یکبار مورد بررسی و تصویربرداری قرار

#### ۲-۵-۲ تعیین ارتباط خودکار (ARD)

ARD یک روش مؤثر برای حذف ویژگی‌های نامرتبب ایجاد می‌کند، به طوری که وزن مربوط به این ویژگی‌ها را به طور خودکار صفر کرده و منجر به تنک شدن بردار وزن برای کلاس بندی می‌شود. با فرض یک تابع پیشین<sup>۱</sup> گوسی با میانگین صفر و ماتریس کواریانس قطری، که المان‌های قطری آن ابرپارامترهای<sup>۲</sup> قابل تنظیمی هستند، می‌تواند محدوده ممکن برای پارامترهای وزن مربوطه را به دست دهد.

#### ۳-۵-۲ رگرسیون لجستیک تنک (SLR)

با ترکیب LR و ARD، مدل SLR به دست می‌آید. در واقع SLR تعمیم بیزی LR است که به طور هم‌زمان انتخاب ویژگی و آموزش پارامترهای مدل را برای کلاس بندی انجام می‌دهد. SLR با استفاده از روش ARD اهمیت هر پارامتر را هم‌زمان با تخمین مقادیر پارامتر تعیین می‌کند. در واقع این روش سه گام انتخاب ویژگی، آموزش کلاس بند و ارزیابی عملکرد کلی را شامل می‌شود. SLR با کمک الگوریتم ARD برای بردار وزن مقدار پیشینی مطابق رابطه ۱۷ فرض می‌کند، به طوری که  $w_d$ ، المان  $d$ -ام بردار  $w$  است.

$$P(w_d|\alpha_d) = N(0, \alpha_d^{-1}) \quad ; \quad d=1, \dots, D \quad (17)$$

در واقع SLR با پارامترهای وزن به عنوان متغیرهایی تصادفی با توزیع‌های پیشین برخوردار می‌کند که توزیع پیشین هر پارامتر  $w_d$ ، توزیعی نرمال با میانگین صفر فرض شده است. همچنین توزیع پیشین برای ابرپارامترها مطابق رابطه ۱۸ در نظر گرفته می‌شود.

$$P_0(\alpha_d) = \alpha_d^{-1} \quad ; \quad d=1, \dots, D \quad (18)$$

ابریارامتر  $\alpha_d$  که پارامتر ارتباط<sup>۳</sup> نامیده می‌شود، میزان ارتباط پارامتر وزن مربوطه را بیان و محدوده ممکن برای این پارامتر را کنترل می‌کند. اگر پارامتر ارتباط کوچک باشد، توزیع احتمال  $w_d$  گسترده بوده و بنابراین پارامتر وزن تخمین زده شده، می‌تواند پس از مشاهده، مقدار زیادی بگیرد (شکل ۱۳-الف) و اگر پارامتر ارتباط بزرگ باشد، پیک تیز توزیع احتمال  $w_d$  در صفر، به عنوان دانش پیشین در نظر گرفته می‌شود و بنابراین پارامتر وزن تخمین زده شده، حتی پس از مشاهده به سمت صفر میل می‌کند. (شکل ۱۳-ب). الگوریتم به صورت تکراری<sup>۴</sup> توزیع‌های پسین<sup>۵</sup> مدل را محاسبه کرده و از آنجا که بیشتر پارامترهای ارتباط به بی‌نهایت واگرا می‌شوند، بنابراین پارامترهای وزن مربوطه، به طور مؤثر صفر و از مدل حذف می‌گردند؛ که این روند برای تعیین ارتباط پارامترها، ARD نامیده می‌شود [۳۹].

<sup>1</sup> Prior

<sup>2</sup> Hyper-parameters

<sup>3</sup> Relevance parameter

<sup>4</sup> Iterative

<sup>5</sup> Posterior distributions

(PPV) و مقدار پیش بینی منفی<sup>۱۱</sup> (NPV) به منظور ارزیابی استفاده شده اند [۱۰، ۱۲، ۱۳، ۱۶]:

$$Accuracy = \frac{TP+TN}{(TP+TN+FP+FN)} \quad (19)$$

$$Sensitivity = \frac{TP}{(FN+TP)} \quad (20)$$

$$Specificity = \frac{TN}{(FP+TN)} \quad (21)$$

$$Positive-predictive-value = \frac{TP}{(FP+TP)} \quad (22)$$

$$Negative-predictive-value = \frac{TN}{(FN+TN)} \quad (23)$$

### ۳-۳ نتایج

در جدول (۲) نتایج ارزیابی کلاس بندی با کمک رگرسیون لجستیک تنک و استفاده از دسته های مختلف ویژگی ارائه شده است. همچنین آن دسته از ویژگی هایی که توسط الگوریتم ARD به عنوان ویژگی های مرتبط انتخاب شده و در واقع کلاس بندی با کمک آنها انجام گشته، معرفی شده اند.

همان طور که از نتایج مشاهده می شود ویژگی های مبتنی بر مرز مشابه ویژگی های مبتنی بر شکل قدرت تمایزدهی مناسبی دارند و همچنین تعمیم این ویژگی ها به مشخصه های یک بعدی دیگر باعث بهبود عملکرد این مجموعه ویژگی شده است.

در جدول (۳) نتایج مربوط به کلاس بندی تصاویر با استفاده از SLR و زیرگروه های مجموعه ویژگی ریخت شناسی مبتنی بر مرز توده ارائه شده است. همان طور که مشاهده می شود توصیف گرهای فوریه در مقایسه با مشخصات آماری این توصیف گرها قدرت تمایزدهی بالاتری دارند.

گرفته و پس از دوره نظارت دو ساله، در صورت اطمینان از عدم رشد توده، تشخیص نهایی خوش خیمی به توده مربوطه داده شده است.

### ۲-۳ اعتبارسنجی و ارزیابی

در این پژوهش به منظور اعتبارسنجی و ارزیابی الگوریتم پیشنهادی، اعتبارسنجی متقابل<sup>۱</sup> انجام شده که به دلیل کوچک بودن پایگاه داده مورد استفاده، از پورتکل Leave-one-out استفاده شده است. در این پورتکل یک داده برای آزمایش و بقیه داده ها به منظور آموزش کلاسبند استفاده می شوند و همه داده ها یک بار به عنوان داده آزمایش مورد استفاده قرار می گیرند به این ترتیب به تعداد داده های موجود (در اینجا ۱۰۴ تصویر دارای برجسب)، الگوریتم مورد آزمایش قرار گرفته، به طوری که هر بار، با داشتن ۱۰۳ تصویر آموزشی، برجسب کلاس تصویر آزمایش توسط کلاسبند مشخص شده است. پس از این مرحله با مقایسه نتیجه حاصل از کلاس بندی الگوریتم و نتیجه تشخیصی بالینی، ماتریس درهم ریختگی<sup>۲</sup>، که شامل مقادیر TP<sup>۳</sup> (تعداد تومورهای بدخیمی که به درستی بدخیم شناسایی شده اند)، TN<sup>۴</sup> (تعداد تومورهای خوش خیمی که به درستی خوش خیم شناسایی شده اند)، FP<sup>۵</sup> (تعداد تومورهای خوش خیمی که به اشتباه بدخیم شناسایی شده اند)، و FN<sup>۶</sup> (تعداد تومورهای بدخیمی که به اشتباه خوش خیم شناسایی شده اند) است، مطابق جدول (۱) به دست آمده است.

جدول ۱ نمایشی از ماتریس درهم ریختگی و چگونگی محاسبه

مقادیر TP، TN، FP و FN

		تشخیص بالینی	
		بدخیم	خوش خیم
تشخیص الگوریتم	بدخیم	TP	FP
	خوش خیم	FN	TN

سپس مطابق روابط ۱۹-۲۳ معیارهای صحت<sup>۷</sup> (Acc)، حساسیت<sup>۸</sup> (Sen)، دقت<sup>۹</sup> (Spe)، مقدار پیش بینی مثبت<sup>۱۰</sup> (PPV)

<sup>1</sup> Cross validation

<sup>2</sup> Confusion matrix

<sup>3</sup> True Positive

<sup>4</sup> True Negative

<sup>5</sup> False Positive

<sup>6</sup> False Negative

<sup>7</sup> Accuracy

<sup>8</sup> Sensitivity

<sup>9</sup> Specificity

<sup>10</sup> Positive-predictive-value

<sup>11</sup> Negative-predictive-value

جدول ۲ نتایج ارزیابی کلاس بندی با کمک الگوریتم SLR و ویژگی های ریخت شناسی

ویژگی های مورد استفاده	ویژگی های انتخاب شده توسط ARD	Acc (%)	Sen (%)	Spe (%)	PPV (%)	NPV (%)
ویژگی های ریخت شناسی مبتنی بر شکل توده (۲۴ ویژگی)	مقدار باقیمانده نرمال شده، مقدار میانگین، واریانس و بی نظمی تابع تغییرات	۸۹/۴۲	۷۱/۸۸	۹۷/۲۲	۹۲/۰۰	۸۸/۶۱
ویژگی های ریخت شناسی مبتنی بر مرز توده (۸۰۳ ویژگی)	نوزدهمین، چهل و سومین، نود و پنجمین و صد و بیست و سومین مؤلفه فرکانسی مشخصه Z، شانزدهمین مؤلفه فرکانسی مشخصه PC، سی و هفتمین مؤلفه فرکانسی مشخصه R، هشتمین مؤلفه فرکانسی مشخصه FPD، سی و هفتمین و پنجاه و هفتمین مؤلفه فرکانسی مشخصه TAF	۸۹/۴۲	۷۵/۰۰	۹۵/۸۳	۸۸/۸۹	۸۹/۶۱
ویژگی های ریخت شناسی مبتنی بر شکل و مرز توده (۸۲۷ ویژگی)	مقدار باقیمانده نرمال شده، مقدار میانگین تابع تغییرات، چهل و چهارمین، نود و چهارمین و صد و چهاردهمین مؤلفه فرکانسی مشخصه Z، چهل و دومین مؤلفه فرکانسی مشخصه PC، شصت و سومین مؤلفه فرکانسی مشخصه R، هشتاد و سومین مؤلفه فرکانسی مشخصه TAF	۸۹/۴۲	۷۸/۱۳	۹۴/۴۴	۸۶/۲۱	۹۰/۶۷

جدول ۳ نتایج ارزیابی کلاس بندی با کمک الگوریتم SLR و ویژگی های ریخت شناسی مبتنی بر مرز توده

ویژگی های مورد استفاده		Acc (%)	Sen (%)	Spe (%)	PPV (%)	NPV (%)
ویژگی های ریخت شناسی مبتنی بر مرز توده (۸۰۳ ویژگی)	ویژگی های مبتنی بر اطلاعات مکانی مشخصه ها (۲۰ ویژگی)	۸۱/۷۳	۹۰/۲۸	۶۲/۵۰	۷۴/۰۷	۸۴/۴۲
	توصیف گرهای فوریه (۷۵۶ ویژگی)	۸۵/۵۸	۶۸/۷۵	۹۳/۰۶	۸۱/۴۸	۸۷/۰۱
	مشخصات آماری توصیف گرهای فوریه (۲۷ ویژگی)	۸۲/۶۹	۶۸/۷۵	۸۸/۸۹	۷۳/۳۳	۸۶/۴۹

جدول ۴ خلاصه ای از مشخصات و نتایج ارائه شده در مقالات پیاده سازی شده

عنوان مقاله	نویسنده اول	سال انتشار	ویژگی های پیشنهادی مقاله	کلاس بند	پایگاه داده تصاویر	Acc (%)	Sen (%)	Spe (%)	PPV (%)	NPV (%)
Breast mass classification on sonographic images on the basis of shape analysis [12]	H. Behnam	۲۰۱۰	اختلاف مساحت، مقدار میانگین، واریانس، چولگی، کشیدگی و بی نظمی تابع تغییرات	NN (MLP)	47 B 34 M	۹۳/۸۳	۹۱/۱۸	۹۵/۷۴	۹۳/۹۴	۹۳/۷۵
Assessing the performance of morphological parameters in distinguishing breast tumors on ultrasound images [10]	A. V. Alvarenga	۲۰۱۰	مقدار باقیمانده نرمال شده، فشردگی و شاخص ناهمواری	LDA	69 B 177 M	۸۳/۷۴	۸۳/۰۵	۸۵/۵۱	۹۳/۶۳	۶۶/۲۹
Automatic ultrasound segmentation and morphology based diagnosis of solid breast tumors [11]	R. F. Chang	۲۰۰۵	ضریب شکست، نسبت قطر بزرگ به قطر کوچک، گردی، وسعت، تحدب و جمود	SVM (rbf)	120 B 90 M	۹۰/۹۵	۸۸/۸۹	۹۲/۵۰	۸۹/۸۹	۹۱/۷۴

B: تصویر اولتراسوند مربوط به یک توده خوش خیم پستان، M: تصویر اولتراسوند مربوط به یک توده بدخیم پستان.

## ۳-۴ مقایسه عملکرد الگوریتم ارائه شده

به دلیل عدم دسترسی به پایگاه داده مورد استفاده در هر یک از مقالات مورد بررسی و همچنین به منظور انجام مقایسه بهتر الگوریتم های استخراج و انتخاب ویژگی و کلاس بندی میان روش پیشنهادی و روش های ارائه شده در سایر مقالات، الگوریتم های مذکور بر روی تصاویر مرز بندی شده مورد استفاده در پژوهش حاضر اعمال شده اند. همچنین از پورتکل Leave-one-out به

در این قسمت به بررسی عملکرد الگوریتم پیشنهادی در مقایسه با الگوریتم های ارائه شده در مقالات مشابه می پردازیم. به این منظور سه مقاله در زمینه موضوع پژوهش حاضر انتخاب و پیاده سازی شده اند که اطلاعات مربوط به هر یک از این مقالات به طور مختصر در جدول (۴) بیان شده است.

[۴۵] Fisher score و [۴۶] T<sub>test</sub> استفاده شده که پس از انتخاب بهترین ویژگی ها از میان ۸۲۷ ویژگی ورودی و با استفاده از کلاس بند رگرسیون لجستیک تصاویر طبقه بندی شده اند [۴۷، ۳۹]. جدول (۶) نتایج معیارهای ارزیابی مربوط به استفاده از این الگوریتم های مختلف انتخاب ویژگی به همراه کلاس بند LR و همچنین کلاس بندی بدون انجام انتخاب ویژگی را نمایش می دهد و همان طور که مشاهده می شود الگوریتم SLR در مقایسه با سایر الگوریتم های انتخاب ویژگی دارای عملکرد بهتری است. ناحیه بندی تصاویر پرداخته شده است. سپس با ارائه یک الگوریتم خودکار، ناحیه مطلوب استخراج و در مرحله بعد با بررسی ویژگی های ریخت شناسی مختلف به پیاده سازی و تعمیم و ارائه ویژگی های جدید پرداخته شده و در مجموع از هر تصویر ۸۲۷ ویژگی استخراج گردیده است. پس از آن به جای استفاده از الگوریتمی جهت انتخاب ویژگی با استفاده از الگوریتم تعمیم یافته رگرسیون لجستیک به انتخاب ویژگی هم زمان با کلاس بندی پرداخته شده است.

در این پژوهش سعی شده تا با استفاده از تصاویر اخذ شده توسط سیستم های مختلف تصویربرداری تا حد امکان وابستگی نتایج الگوریتم پیشنهادی به دستگاه و شرایط تصویربرداری کاهش داده شود.

ترکیب دو رویکرد مرز بندی دستی توده و ناحیه بندی کامپیوتری تصاویر به جای استفاده تنها از هر یک از این رویکردها، با پرداخت هزینه زمانی بیشتر، باعث افزایش دقت و صحت ناحیه بندی شده است، همچنین الگوریتم جدید ارائه شده برای تعیین خودکار ROI، موجب افزایش سرعت و کیفیت ناحیه مطلوب استخراج شده، گشته است.

استفاده از مشخصه های یک بعدی مرز و اطلاعات مکانی و فرکانسی آنها، که برای پزشک قابل مشاهده نیستند، در کنار ویژگی های ابتدایی مرز باعث تمایز بیشتر میان توده های خوش خیم و بدخیم شده است. همچنین با بررسی جامع ویژگی های ریخت شناسی مشاهده شده که ترکیب دو مجموعه ویژگی ریخت شناسی مبتنی بر شکل و مرز توده نسبت به استفاده تنها از هر یک از این مجموعه ویژگی ها، تأثیر قابل ملاحظه ای در نتایج ایجاد نکرده است.

در میان ویژگی های مورد بررسی، مقدار باقیمانده نرمال شده، مقدار میانگین، واریانس و بی نظمی تابع تغییرات در بین ویژگی های ریخت شناسی مبتنی بر شکل توده و مؤلفه های فرکانسی مشخصه های Z، PC، R، FPD و TAF در بین ویژگی های ریخت شناسی مبتنی بر مرز توده بهترین عملکرد را داشته اند؛ پس مشاهده می شود که ویژگی های جدید ارائه شده در این پژوهش دارای عملکرد مناسبی بوده اند.

استفاده از الگوریتم رگرسیون لجستیک تنک که به طور هم زمان با کلاس بندی، انتخاب ویژگی را هم انجام می دهد، در مقایسه با سایر الگوریتم های انتخاب ویژگی، نتایج بهتری را ارائه

منظور اعتبار سنجی، استفاده شده است. در جدول (۵) نتایج مربوط به پیاده سازی این مقالات ارائه شده و همان طور که مشاهده می شود الگوریتم پیشنهادی در مقایسه با سایر مقالات، مقادیر پارامترهای ارزیابی را افزایش داده و دارای عملکرد بهتری است. در این پژوهش به منظور ارزیابی الگوریتم انتخاب ویژگی و کلاس بندی علاوه بر استفاده از الگوریتم رگرسیون لجستیک تنک از هفت الگوریتم مختلف انتخاب ویژگی شامل: [۴۰] Chisquare، [۴۱] InfoGain، [۴۲] CFS، [۴۳] KruskalWallis، [۴۴] FCBF،

جدول ۵ مقایسه نتایج الگوریتم ارائه شده با نتایج پیاده سازی مقالات بر روی پایگاه داده موجود

	Acc (%)	Sen (%)	Spe (%)	PPV (%)	NPV (%)
نتایج الگوریتم پیشنهادی پژوهش حاضر	۸۹/۴۲	۷۸/۱۳	۹۴/۴۴	۸۶/۲۱	۹۰/۶۷
نتایج پیاده سازی مقاله [12] بر روی پایگاه داده موجود	۸۷/۵۰	۷۵/۰۰	۹۳/۰۶	۸۲/۷۶	۸۹/۳۲
نتایج پیاده سازی مقاله [10] بر روی پایگاه داده موجود	۸۴/۶۲	۶۸/۷۵	۹۱/۶۷	۷۸/۵۷	۸۶/۸۴
نتایج پیاده سازی مقاله [11] بر روی پایگاه داده موجود	۸۱/۷۳	۶۸/۷۵	۸۷/۵۰	۷۰/۹۷	۸۶/۳۰

جدول ۶ مقایسه نتایج استفاده از الگوریتم رگرسیون لجستیک تنک و چندین الگوریتم انتخاب ویژگی

	Acc (%)	Sen (%)	Spe (%)	PPV (%)	NPV (%)
SLR	۸۹/۴۲	۷۸/۱۳	۹۴/۴۴	۸۶/۲۱	۹۰/۶۷
Chisquare & LR	۸۴/۶۲	۶۵/۶۳	۹۳/۰۶	۸۰/۷۷	۸۵/۹۰
InfoGain & LR	۸۳/۶۵	۶۲/۵۰	۹۳/۰۶	۸۰/۰۰	۸۴/۸۱
CFS & LR	۸۱/۳۷	۶۵/۶۳	۸۸/۵۷	۷۲/۴۱	۸۴/۹۳
KruskalWallis & LR	۸۰/۳۹	۶۸/۷۵	۸۵/۷۱	۶۸/۷۵	۸۵/۷۱
FCBF & LR	۸۰/۳۹	۶۲/۵۰	۸۸/۵۷	۷۱/۴۳	۸۳/۷۸
Fisher score & LR	۶۹/۶۱	۳۱/۲۵	۸۷/۱۴	۵۲/۶۳	۷۳/۴۹
T <sub>test</sub> & LR	۶۷/۶۵	۴۰/۶۳	۸۰/۰۰	۴۸/۱۵	۷۴/۶۷
LR	۶۴/۷۱	۶۸/۷۵	۶۲/۸۶	۴۵/۸۳	۸۱/۴۸

#### ۴ جمع بندی و نتیجه گیری

این پژوهش با هدف تشخیص نوع توده در تصاویر سونوگرافی پستان و مقایسه عملکرد دسته های مختلف ویژگی های ریخت شناسی به دنبال ارائه یک سیستم تشخیصی کمک-رایانه ای بوده است. برای این منظور پس از جمع آوری پایگاه داده ای شامل ۱۰۴ تصویر سونوگرافی از توده های پستان به پیش پردازش و



- Liu, "Fully automatic and segmentation-robust classification of breast tumors based on local texture analysis of ultrasound images," *Pattern Recognit.*, vol. 43, no. 1, pp. 280–298, 2010.
- [6] Y. Su, Y. Wang, J. Jiao, and Y. Guo, "Automatic detection and classification of breast tumors in ultrasonic images using texture and morphological features," *Open Med. Inform. J.*, vol. 5, no. Suppl 1, pp. 26–37, 2011.
- [7] Y. Y. Liao, P. H. Tsui, and C. K. Yeh, "Classification of benign and malignant breast tumors by ultrasound B-scan and Nakagami-based images," *J. Med. Biol. Eng.*, vol. 30, no. 5, pp. 307–312, 2010.
- [8] S. M. El-Khateeb, A. E. Abou-Khalaf, and M. M. Farid, "A prospective study of three diagnostic sonographic methods in differentiation between benign and malignant salivary gland tumours," *Dentomaxillofacial Radiol.*, 2014.
- [9] W. Gómez Flores, W. C. D. A. Pereira, and A. F. C. Infantosi, "Improving classification performance of breast lesions on ultrasonography," *Pattern Recognit.*, vol. 48, no. 4, pp. 1125–1136, 2014.
- [10] A. V. Alvarenga, a. F. C. Infantosi, W. C. a Pereira, and C. M. Azevedo, "Assessing the performance of morphological parameters in distinguishing breast tumors on ultrasound images," *Med. Eng. Phys.*, vol. 32, no. 1, pp. 49–56, 2010.
- [11] R. F. Chang, W. J. Wu, W. K. Moon, and D. R. Chen, "Automatic ultrasound segmentation and morphology based diagnosis of solid breast tumors," *Breast Cancer Res. Treat.*, vol. 89, no. 2, pp. 179–185, 2005.
- [12] H. Behnam, F. S. Zakeri, and N. Ahmadinejad, "Breast mass classification on sonographic images on the basis of shape analysis," *J. Med. Ultrason.*, vol. 37, no. 4, pp. 181–186, 2010.
- [13] F. S. Zakeri, H. Behnam, and N. Ahmadinejad, "Classification of benign and malignant breast masses based on shape and texture features in sonography images," *J. Med. Syst.*, vol. 36, no. 3, pp. 1621–1627, 2012.
- [14] W. Gómez, L. Leija, A. V. Alvarenga, A. F. C. Infantosi, and W. C. A. Pereira, "Computerized lesion segmentation of breast ultrasound based on marker-controlled watershed transformation," *Med. Phys.*, vol. 37, no. 1, pp. 82–95, 2010.
- [15] K. M. Prabusankarlal, P. Thirumoorthy, and R. Manavalan, "Computer Aided Breast Cancer Diagnosis Techniques in Ultrasound: A Survey," *J. Med. Imaging Heal. Informatics*, vol. 4, no. 3, pp. 331–349, 2014.
- [۱۶] جهان‌دیده رستم علی، بهنام حمید و احمدی نژاد نسرين. "طبقه‌بندی توده‌های سرطانی سینه با استفاده از ویژگی‌های ریخت‌شناسی توده و ویژگی‌های بافتی تصاویر سونوگرافی در ناحیه دارای توده و نواحی اطراف آن." ۲۵۳–۲۴۷.
- [17] M. H. Yap, E. A. Edirisinghe, and H. E. Bez, "A Comparative Study in Ultrasound Breast Imaging

داده است. پس می‌توان نتیجه گرفت، در کاربردهای تشخیصی مشابهی که قصد انتخاب مهم‌ترین ویژگی‌ها را داریم، استفاده از الگوریتم کلاس‌بندی تنکی مانند SLR که به طور خودکار ویژگی‌های غیرمرتبط را حذف می‌کند، می‌تواند مفید باشد. همچنین صحت بالاتر نتایج کلاس‌بندی به همراه انتخاب ویژگی، نسبت به کلاس‌بندی بدون انتخاب ویژگی، مؤید این موضوع می‌باشد که در مواقعی که می‌خواهیم با تعداد زیادی ویژگی، به کلاس‌بندی داده‌ها بپردازیم، کاهش تعداد ویژگی‌ها و حذف ویژگی‌های نامرتبط نتایج را بهبود می‌دهد. الگوریتم SLR قادر است در زمان بسیار کوتاه (زمانی حدود ۲/۵ ثانیه برای تعیین نوع توده یک تصویر، با داشتن ۱۰۳ تصویر آموزشی با استفاده از ۸۲۷ ویژگی و با کمک نرم‌افزار Matlab 2014b در یک رایانه با پردازنده پنج هسته‌ای) با تعیین مهم‌ترین ویژگی‌ها به کلاس‌بندی تصاویر بپردازد.

در نهایت به منظور تکمیل و بهبود پژوهش انجام شده، پیشنهاد می‌گردد تا تعداد تصاویر پایگاه داده، افزایش یافته و در صورت امکان از داده اعتبارسنجی استفاده شود. همچنین می‌توان اطلاعات تبدیل موجک مشخصه‌های یک‌بعدی مرز توده را به جای تبدیل فوریه گسسته این مشخصه‌ها مورد استفاده قرار داد. از سوی دیگر از آنجا که تعامل میان پزشک و کامپیوتر موجب اعتبار بیشتر تشخیص می‌گردد، پیشنهاد می‌شود از اطلاعات کیفی تشخیصی و نیز ویژگی‌های بالینی بیمار در کنار ویژگی‌های کامپیوتری، استفاده گردد.

## سپاسگزاری

از جناب آقای دکتر بهنام، عضو هیئت علمی دانشگاه علم و صنعت ایران، و سرکار خانم دکتر گیتی، عضو هیئت علمی دانشگاه علوم پزشکی تهران، که در جمع‌آوری پایگاه داده مورد استفاده، صمیمانه همکاری داشته‌اند، کمال تشکر را داریم.

## مراجع

- [۱] بهنام حمید، ذاکری فهیمه سادات، گیفانی پریسا و همکاران. "پردازش تصاویر اولتراسوند." نشر اشراقیه، ۱۳۹۰.
- [2] H. D. Cheng, J. Shan, W. Ju, Y. Guo, and L. Zhang, "Automated breast cancer detection and classification using ultrasound images: A survey," *Pattern Recognit.*, vol. 43, no. 1, pp. 299–317, 2010.
- [3] A. Jalalian, S. B. T. Mashohor, H. R. Mahmud, M. I. B. Saripan, A. R. B. Ramli, and B. Karasfi, "Computer-aided detection/diagnosis of breast cancer in mammography and ultrasound: A review," *Clin. Imaging*, vol. 37, no. 3, pp. 420–426, 2013.
- [4] P. Telagarapu and S. Poonguzhali, "Analysis of Contourlet Texture Feature Extraction to Classify the Benign and Malignant Tumors from Breast Ultrasound Images," vol. 6, no. 1, pp. 293–305, 2014.
- [5] B. Liu, H. D. Cheng, J. Huang, J. Tian, X. Tang, and J.

- Commun. Image Represent.*, vol. 26, pp. 161–167, 2015.
- [33] D. Varga, “American Hand Sign Recognition in Video Streams,” 2014.
- [34] J.-H. Lee, Y. K. Seong, C.-H. Chang, J. Park, M. Park, K.-G. Woo, and E. Y. Ko, “Fourier-based shape feature extraction technique for computer-aided B-Mode ultrasound diagnosis of breast tumor.,” *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, vol. 2012, pp. 6551–4, 2012.
- [35] M. R. Mani, G. P. S. Varma, P. D.M., and C. Satyanarayana, “Design of a Novel Shape Signature by Farthest Point Angle for Object Recognition,” *Int. J. Image, Graph. Signal Process.*, vol. 7, no. 1, pp. 35–46, 2014.
- [36] C. T. Zahn and R. Z. Roskies, “Fourier Descriptors for Plane Closed Curves,” vol. c, no. 3, 1972.
- [37] A. C. Jalba, M. H. F. Wilkinson, and J. B. T. M. Roerdink, “Shape Representation and Recognition Through Morphological Curvature Scale Spaces,” vol. 15, no. 2, pp. 331–341, 2006.
- [38] P. D. Silva and P. Bhuvaneshwari, “Various Shape Descriptors in Image Processing – A Review,” vol. 4, no. 3, pp. 2338–2342, 2015.
- [39] O. Yamashita, “Sparse estimation automatically selects voxels relevant for the decoding of fMRI activity patterns,” *Neuroimage*, vol. 42, no. 4, pp. 1414–1429, 2008.
- [40] H. L. H. Liu and R. Setiono, “Chi2: Feature selection and discretization of numeric attributes,” in *tai*, 1995, no. DECEMBER, p. 388.
- [41] G. Cawley, “Sparse multinomial logistic regression via bayesian l1 regularisation,” *Adv. Neural Inf. Process. Syst.*, vol. 19, p. 209, 2007.
- [42] M. A. Hall and L. A. Smith, “Feature Selection for Machine Learning: Comparing a Correlation-Based Filter Approach to the Wrapper.,” in *FLAIRS conference*, 1999, vol. 1999, pp. 235–239.
- [43] D. Cai, C. Zhang, and X. He, “Unsupervised Feature Selection for Multi-Cluster Data,” *Conf. Knowl. Discov. data*, pp. 333–342, 2010.
- [44] L. Yu and H. Liu, “Feature selection for high-dimensional data: A fast correlation-based filter solution,” in *ICML*, 2003, vol. 3, pp. 856–863.
- [45] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*. John Wiley & Sons, 2012.
- [46] “Student’s t-test.” [Online]. Available: [https://en.wikipedia.org/wiki/Student%2527s\\_t-test](https://en.wikipedia.org/wiki/Student%2527s_t-test).
- [47] C. M. Bishop, *Pattern recognition and machine learning*. springer, 2006.
- Classification,” vol. 7259, pp. 1–11, 2009.
- [18] C.-M. Chen, Y.-H. Chou, K.-C. Han, G.-S. Hung, C.-M. Tiu, H.-J. Chiou, and S.-Y. Chiou, “Breast lesions on sonograms: computer-aided diagnosis with nearly setting-independent features and artificial neural networks.,” *Radiology*, vol. 226, no. 2, pp. 504–514, 2003.
- [19] W. Gómez and W. Pereira, “Computerized diagnosis of breast lesions on ultrasonography,” *XXII Congr. ...*, pp. 399–402, 2010.
- [20] M. Yang, K. Kpalma, and J. Ronsin, “A survey of shape feature extraction techniques,” *Pattern Recognit.*, vol. 2008, no. November, pp. 43–90, 2008.
- [21] X. Diao, T. Wang, and S. Chen, “Computer-aided Diagnosis of Breast Tumor Based on B-mode Ultrasound and Color Doppler Flow Imaging,” 2009.
- [22] K. G. Kim, J. H. Kim, and B. G. Min, “Classification of malignant and benign tumors using boundary characteristics in breast ultrasonograms.,” *J. Digit. imaging Off. J. Soc. Comput. Appl. Radiol.*, vol. 15 Suppl 1, no. March, pp. 224–227, 2002.
- [23] S. Ge and H. D. Choi, “Breast Cancer Classification of Mammographic Masses using Improved Shape Features,” pp. 188–194, 2015.
- [24] D. Zhang and G. Lu, “Study and evaluation of different Fourier methods for image retrieval,” *Image Vis. Comput.*, vol. 23, no. 1, pp. 33–49, 2005.
- [25] D. Zhang and G. Lu, “A Comparative Study on Shape Retrieval Using Fourier Descriptors with Different Shape Signatures,” *Victoria*, vol. 1, no. January, pp. 1–9, 2001.
- [26] “A new shape signature for fourier descriptors,” *Computer (Long Beach Calif.)*, no. 1, pp. 161–164, 2007.
- [27] A. El-ghazal, O. Basir, and S. Belkasim, “Farthest point distance: A new shape signature for Fourier descriptors,” *Signal Process. Image Commun.*, vol. 24, no. 7, pp. 572–586, 2009.
- [28] H. Kauppinen, T. Seppanen, and M. Pietikainen, “An experimental comparison of autoregressive and Fourier-based descriptors in 2D shape classification,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 17, no. 2, pp. 201–207, 1995.
- [29] S. Conseil, S. Bourenane, and L. Martin, “Comparison of fourier descriptors and HU moments for hand posture recognition,” *Eur. Signal Process. Conf.*, no. Eusipco, pp. 1960–1964, 2007.
- [30] R. Kasaudhan, S. I. Jeon, and S. H. Son, “Similarity measurement with Mesh Distance Fourier Transform in 2D binary image,” pp. 183–187, 2015.
- [31] C. Mattrand, A. Béakou, and K. Charlet, “Numerical modeling of the flax fiber morphology variability,” *Compos. Part A Appl. Sci. Manuf.*, vol. 63, pp. 10–20, 2014.
- [32] X. Shu, L. Pan, and X. J. Wu, “Multi-scale contour flexibility shape signature for Fourier descriptor,” *J. Vis.*



هدا نعمت در سال ۱۳۸۷ موفق به اخذ مدرک دیپلم گردید و تحصیلات کارشناسی خود را در سال ۱۳۸۸ در رشته مهندسی پزشکی، گرایش بیوالکترونیک در دانشگاه اصفهان آغاز کرد و در سال ۱۳۹۲ فارغ التحصیل شد. پس از آن در همان سال تحصیلات کارشناسی ارشد خود را در دانشگاه تربیت مدرس آغاز و در سال ۱۳۹۴ فارغ التحصیل شد. ایشان در حاضر مشغول به تدریس در دانشگاه و انجام پژوهش می باشند.



علی محلوچی فر مدرک لیسانس الکترونیک خود را در سال ۱۹۸۸ از دانشگاه تهران و مدرک فوق لیسانس الکترونیک را در سال ۱۹۹۱ از دانشگاه صنعتی شریف دریافت کردند و در سال ۱۹۹۵ موفق به اخذ مدرک دکتری در رشته ابزار دقیق پزشکی از دانشگاه منچستر شدند. ایشان از سال ۱۹۹۶ عضو هیئت علمی دانشگاه تربیت مدرس شدند و در حال حاضر به عنوان دانشیار گروه مهندسی پزشکی در دانشکده برق و کامپیوتر این دانشگاه، مشغول به تدریس و پژوهش می باشند.



علی گویا دارای فوق لیسانس مهندسی پزشکی شاخه بیوالکترونیک از دانشگاه تهران و دکترای تکنولوژی اطلاعات از دانشگاه توکیو (سال ۲۰۰۷) می باشد. وی در گذشته استادیار و عضو هیأت علمی دانشگاه تربیت مدرس بوده و در حال حاضر عضو آکادمیک دانشگاه شفیلد انگلستان می باشد. زمینه تخصصی وی پردازش تصاویر پزشکی و یادگیری ماشین می باشد. در گذشته، ایشان موفق به اخذ فلوشیپ های پست دکترا از ژاپن و اروپا (ماری کوری) شده اند.

نسرين احمدی نژاد تحصیلات عالی خود را در سال ۱۳۶۱ در رشته پزشکی دانشکده پزشکی دانشگاه تهران آغاز کرد و به سال ۱۳۶۸ فارغ التحصیل شد. وی در سال ۱۳۷۲ موفق به اخذ درجه تخصصی در رشته رادیولوژی از دانشکده پزشکی دانشگاه علوم پزشکی تهران گردید. وی دوره های تکمیلی سی تی اسکن را در بیمارستان شریعتی دانشگاه علوم پزشکی تهران و همچنین دوره MRI را در کشور کانادا دانشگاه مگ گیل گذرانده اند. ایشان اکنون به عنوان عضو هیئت علمی و استادیار گروه رادیولوژی در بخش مرکز تصویر برداری بیمارستان امام خمینی (ره) مشغول به کار می باشند.