مقاله پژوهشی

فصلنامهٔ بیماری‌های پستان ایران، سال پنجم، شماره چهارم، زمستان ۱۳۹۱

خوشه بندی و شناسایی کردن سرطان پستان توسط تصاویر حرارتی به کمک ترکیب شبکه عصبی SOM و SVM

حسین قیومی زاده[1]: دانشجو دکتری مهندسی پزشکی دانشگاه حکیم سبزواری

علی دروگرمقدم: دانشجو کارشناسی ارشد مهندسی پزشکی دانشگاه حکیم سبزواری، عضو مرکز تحقیقات فناوری‌های نوین پزشکی دانشگاه علوم پزشکی سبزوار

جواد حدادنیا: دانشیار گروه مهندسی پزشکی دانشگاه حکیم سبزواری، عضو مرکز تحقیقات فناوری‌های نوین پزشکی دانشگاه علوم پزشکی سبزوار

محمد محمد زاده: استادیار فیزیولوژی مرکز تحقیقات سلولی- مولکولی، گروه فیزیولوژی دانشگاه علوم پزشکی سبزوار

امید رحمانی سریاست: دانشجوی دکتری مهندسی برق- الکترونیک دانشگاه حکیم سبزواری

چکیده

مقدمه: تصویربرداری مادون قرمز یک روش تصویربرداری غیرتهاجمی است که به عنوان یک ابزار تشخیصی به کار می‌رود. این تحقیق تلاشی برای طبقه بندی کردن سرطان پستان به وسیله استخراج اطلاعات کمی و کیفی از تصاویر حرارتی و استخراج بهترین پارامترهای تشخیصی از میان کل پارامترهای موجود به کمک شبکه‌های عصبی خود سازمانده در جهت شناسایی سرطان پستان است.

موارد و روش‌ها: در روش پیشنهادی ابتدا از طریق شبکه‌های عصبی خود سازمانده به طبقه بندی کردن یک تصویر حرارتی می‌پردازیم و نواحی مشکوک را استخراج نموده سپس دوباره نواحی جداسازی شده را با الگوریتم دوم از شبکه SOM که دارای ساختار متفاوت با الگوریتم اول است به طبقه بندی نمودن نواحی جداسازی شده قبلی می‌پردازیم تا نهایتاً به دو کلاس طبقه بندی نماییم سپس وزن‌ها یا به عبارتی مراکز ایجاد شده این دو کلاس به عنوان ویژگی‌های تشخیص به ماشین بردار پشتیبان SVM داده خواهد شد تا نمونه‌های سرطانی از غیر سرطانی تفکیک گردد.

نتایج: به وسیله الگوریتم پیشنهادی توانستیم به شناسایی سرطان پستان از روی تصاویر حرارتی دست یابیم مقدار کارایی (Performance) خروجی شبکه‌های ماشین بردار پشتیبان با نتایج ماموگرافی برابر ۱/۷۳ است که نتیجه قابل ملاحظه‌ای صورت پذیرفته است.

بحث و نتیجه‌گیری: نتایج این تحقیق بیانگر آن است که مدل ترکیبی قادر است بهترین پارامتر تشخیصی را از روی تصاویر حرارتی استخراج نموده سپس توسط این پارامترها به کلاس بندی و جداسازی تصاویر سرطانی از غیر سرطانی با دقت بالا پرداخت.

واژه‌های کلیدی: تصاویر حرارتی، سرطان پستان، شبکه‌های عصبی خود سازمانده SOM، ماشین بردار پشتیبان SVM.

[1] نشانی نویسنده پاسخگو: خراسان رضوی، سبزوار، توحید شهر، پردیس دانشگاه حکیم سبزواری، دانشکده برق و کامپیوتر، گروه مهندسی پزشکی، کد پستی: ۹۶۱۷۹۷۶۴۸۷، حسین قیومی زاده.
نشانی الکترونیک: h.ghayoumizadeh@gmail.com

[Downloaded from ijbd.ir on 2026-11-19]

[DOR: 20.1001.1.17359406.1391.5.4.2.2]

مقدمه

تصویربرداری مادون قرمز یک روش تصویربرداری غیرتهاجمی است که به عنوان یک ابزار تشخیصی به کار می‌رود. منشأ اصلی اشعه‌های مادون قرمز گرمای برخاسته از بدن‌های مختلف است که دمای آنها از صفر مطلق بیشتر است. بنابراین ترموگرام از یک بیمار توزیع گرما در بدن را فراهم می‌کند. با توجه به نسبت‌های متابولیسمی بالا و پیشرفت عروق angiogenesis، سلول‌های سرطانی نسبت به سلول‌های طبیعی اطرافشان، دمای بیشتری دارند(۶ و ۷). بنابراین می‌توان در تصاویر مادون قرمز، سلول‌های سرطانی را به صورت کانون‌های بحرانی نشان داد.

سرطان پستان در کشورمان رتبه دوم را پس از سرطان ریه و رتبه نخست را در بین سرطان زنان دارد و یک پنجم علت مرگ زنان در اثر سرطان، مربوط به این نوع از سرطان است(۱). سرطان پستان شایع‌ترین سرطان در بین زنان است. طبق برآورد سال ۲۰۰۰ سالانه یک میلیون بیمار جدید به شمار بیماران سرطان پستان در جهان افزوده می‌شود. مقدار این افزایش در حدود ۲٪ و مرگ و میر ناشی از آن در سرتاسر دنیا سالانه ۴۲۰۰۰۰ مورد است(۲ و ۳). تشخیص به موقع سرطان پستان یکی از مهم‌ترین مسائلی است که محققان همیشه در پی آن بوده‌اند. با توجه به هزینه‌های بالای درمان و گسترش بیماری بین بانوان، شناسایی به موقع این عارضه، گام مهمی در کاهش عوارض اجتماعی و سلامتی جامعه خواهد بود(۴ و۵).

در(۲۴) نویسندگان به بررسی آسیب‌های مانند کیست‌ها و نئوپلاسم خوش خیم پرداخته‌اند. که در نتیجه‌گیری، دمانگاری مادون قرمز را به عنوان یک روش امیدوار کننده به تشخیص زود هنگام آسیب‌های پستان می‌پردازند که می‌تواند به صورت غربالگری پستان در ادبیات پزشکی یافت شود.

در(۹) نویسنده با استخراج ویژگی حرارتی از تصاویر اگرچه توانسته است به شناسایی سرطان پستان اقدام نماید اما مشکلی که در مقاله وجود دارد نواحی مشکوک و پرحرارتی بود که در مکان‌هایی به جز پستان پدیدار می‌شد، در صورتی که سیستم‌های کامپیوتری و هوشمند می‌خواستند آنها را آنالیز کنند با مشکلاتی در جداسازی نواحی سرطانی و همچنین در تشخیص مواجه می‌شدند.

در(۲۲) نویسندگان تمرکز خودشان را بر روی تصویربرداری حرارتی از بافت خوش خیم پستان معطوف کردند و نتایج خود را با تصاویر میکروسکوپی مقایسه و آنالیز نموده‌اند. در این کار در واقع تفاوت حرارتی بافت‌های خوش خیم و بد خیم با یکدیگر مقایسه شده و در نتیجه تحقیق بر امکان تصویربرداری حرارتی به عنوان یک عامل تشخیصی مهم تاکید می‌نماید.

در(۲۳) به تشخیص سرطان سینه در مراحل اولیه بر اساس تصاویر مادون قرمز پرداخته شده است، سپس برخی از روش‌ها، پروتکل‌ها و شبیه سازی‌های عددی و یا توسعه یافته انجام بر روی تصاویر اعمال شده است. دو مسئله مهم در این مقاله توسعه یک پروتکل استاندارد برای به دست آوردن تصاویر سینه‌های حرارتی از جمله طراحی، ساخت و نصب و راه اندازی دستگاه‌های مکانیکی و دوم محاسبات عددی پروفایل درجه حرارت پستان است که با عدم قطعیت از ارزش‌های واقعی از پارامترهای ترموگرافی برخی از بافت‌ها ایجاد می‌شود. سپس یک روش برای تخمین خواص حرارتی بر اساس این تصاویر ارائه شده است.

اطلاعات لازم برای این پژوهش از یک تحقیق عملی در شهر سبزوار در دانشگاه حکیم سبزواری و با همکاری دانشگاه علوم پزشکی سبزوار به دست آمده است. این تحقیق روی اطلاعات جمع آوری شده از تصاویر حرارتی ۲۰۰ زن که شامل ۱۵ مورد ضایعه غیر طبیعی در ناحیه پستان در محدوده سنی ۱۸ تا ۳۵ سال، انجام شده است. پس هدف در این کار طبقه بندی نمودن تصاویر بر مبنای حرارت نواحی مشکوک است، که هم ارز با نواحی پر حرارت خواهد بود.

در کارهای قبلی(۸ و ۹) توسط نویسندگان همین مقاله یکی از مشکلاتی که ارایه شد نواحی مشکوک و پرحرارتی بود که در مکان‌هایی به جز پستان پدیدار می‌شد، در صورتی که سیستم‌های کامپیوتری و هوشمند می‌خواستند آنها را آنالیز کنند با مشکلاتی در جداسازی نواحی سرطانی و همچنین در تشخیص مواجه خواهند می‌شدند. پس در این کار با ارایه سیستمی هوشمندانه‌تر و دقت بالا سعی در رفع عیوب خواهد شد. در روش پیشنهادی ابتدا از طریق شبکه‌های عصبی خود سازمانده به طبقه بندی نمودن یک تصویر حرارتی می‌پردازیم و نواحی مشکوک را

[Downloaded from ijbd.ir on 2026-11-19]

[DOR: 20.1001.1.17359406.1391.5.4.2.2]

استخراج نموده سپس دوباره نواحی جداسازی شده را با الگوریتم دوم از شبکه SOM که دارای ساختار متفاوت با الگوریتم اول است به طبقه‌بندی نمودن نواحی جداسازی شده قبلی می‌پردازیم وزن‌ها یا به عبارتی مراکز ایجاد شده به عنوان ویژگی‌های تشخیص به شبکه‌های ماشین بردار SVM داده خواهد شد.

پردازش تصویر:

در فرایند تشخیصی روش ارایه شده لازم است مراحلی از جمله خوشه‌بندی، پیش پردازش تصاویر و جداسازی نواحی سرطانی صورت پذیرد که در ذیل به ترتیب به آنها اشاره می‌شود.

خوشه‌بندی را می‌توان به عنوان مهم‌ترین مسئله در یادگیری بدون نظارت در نظرگرفت. خوشه‌بندی با یافتن یک ساختار درون یک مجموعه از داده‌های بدون برچسب درگیر است. خوشه به مجموعه‌ای از داده‌ها گفته می‌شود که به هم شباهت داشته باشند. در خوشه‌بندی سعی می‌شود تا داده‌ها به خوشه‌هایی تقسیم شوند که شباهت بین داده‌های درون هر خوشه حداکثر و شباهت بین داده‌ای درون خوشه‌های متفاوت حداقل شود. از آنجا که خوشه‌بندی یک روش یادگیری بدون نظارت محسوب می‌شود، در موارد بسیاری از جمله داده کاوی، در تقسیم بندی تصاویر و غیره می‌تواند کاربرد داشته باشد(۱۲-۱۰). در طبقه‌بندی الگو به شرایطی بر می‌خوریم که در آن اطلاعاتی که شبکه درباره پاسخ صحیح در اختیار ما قرار می‌دهد، نمی‌توانست ما را به یکی از پاسخ‌ها راهنمایی کند. در این شبکه‌ها در بعضی موارد، پاسخ شبکه شامل اطلاعاتی بیشتر از حد نیاز بود، به گونه‌ای که نمی‌توانستیم این اطلاعات اضافی را برای رسیدن به پاسخ مورد نظر استفاده کنیم. با این پیش زمینه و با توجه به روش پیشنهادی شبکه‌های عصبی خود سازمانده می‌تواند در جهت خوشه بندی تصاویر استفاده کنیم.

مدلی که برای شبکه عصبی در مقاله استفاده شده شبکه خود سازمانده است. در شبکه خود سازمانده، از روش یادگیری رقابتی برای آموزش استفاده می‌شود و مبتنی بر مشخصه‌های خاصی از مغز انسان توسعه یافته است. سلول‌ها در مغز انسان در نواحی مختلف طوری سازمان دهی شده‌اند که در نواحی حسی مختلف، با نقشه‌های محاسباتی مرتب و معنی‌دار ارائه می‌شوند. برای نمونه، ورودی‌های حسی لامسه، شنوائی و ... با یک ترتیب هندسی معنی‌دار به نواحی مختلف مرتبط هستند(۱۳). نقشه خود سازماندهی(SOM) یک شبکه عصبی بدون نظارت است که از نورون‌های عصبی در یک ساختار گرید منظم با ابعاد پایین تشکیل شده است. هر نورون دارای یک بردار وزن n بعدی است که در آن n برابر با ابعاد بردارهای ورودی است. بردارهای وزن(سیناپس‌ها) لایهٔ ورودی را به لایهٔ خروجی(که نقشه یا لایهٔ رقابتی نامیده می‌شود) متصل می‌کنند. نورون‌ها توسط یک تابع همسایگی به یکدیگر متصل شده‌اند. هر بردار ورودی، بر اساس بیشترین شباهت، نورونی در لایهٔ خروجی را که سلول برنده خوانده می‌شود، فعال می‌کند. شباهت معمولاً بر اساس فاصله اقلیدسی بین دو بردار اندازه‌گیری می‌شود (۱۴).

$$D_j = \sum_{i=1}^{n} \left\| W_{i,j} - x_i \right\|^2 \quad (۱)$$

که در آن xi ، i امین بردار ورودی بوده، Wi, j بردار وزنی است که ورودی i را به نورون خروجی j متصل می‌کند و Dj حاصل جمع فاصلهٔ اقلیدسی بین نمونهٔ ورودی xi و بردار وزن ارتباطی آن به j امین سلول خروجی است که یک واحد نقشه نامیده می‌شود. مهم‌ترین تفاوت الگوریتم آموزشی SOM با سایر الگوریتم‌های اندازه‌گیری برداری[1] این است که علاوه بر وزن ارتباطی واحد دارای بیشترین انطباق(سلول عصبی برنده)، وزن‌های سلول‌های همسایهٔ سلول برنده نیز به روزرسانی می‌شوند. مشاهدات نزدیک به هم در فضای ورودی، دو واحد نزدیک به هم در نقشه را فعال می‌سازند. مرحلهٔ آموزش تا زمانی که بردارهای وزن به حالت پایداری برسند و دیگر تغییر نکنند، ادامه می‌یابد.

$$W_{i,j}^{new} = W_{i,j}^{old} + h_{i,j}(x_{.i} - W_{i,j}^{old}) \quad (۲)$$

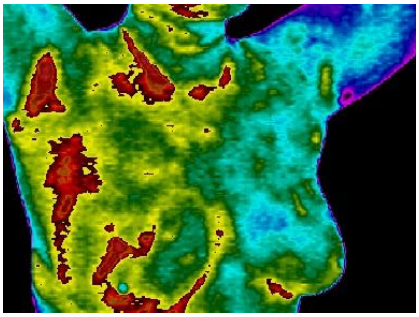
که x.i نمونهٔ ورودی، Wi, j old بردار وزن قبلی بین بردار ورودی xi و بردار وزن ارتباطی به سلول عصبی خروجی j است. hi, j تابع همسایگی و Wi, j new بردار وزن به روزرسانی شده بین سلول ورودی i و سلول خروجی j است. در این مرحله لازم است بر روی داده ها پیش پردازشی صورت پذیرد. داده‌هایی که برای طبقه بندی نمودن یک تصویر حرارتی وارد شبکه عصبی

[1] vector quantization

[Downloaded from ijbd.ir on 2026-11-19]

[DOR: 20.1.1.17359406.1391.5.4.2.2]

می‌نماییم شامل ۵ فاکتور است. اگر شدت روشنایی پیکسل های تصویر حرارتی مانند تصویر۱، f (x,y) باشد به نوعیx, y موقعیت مکانی پیکسل‌هارا تشکیل دهند g را می‌تواند طبق رابطه ۱ به عنوان ورودی وارد شبکه som نماییم.



شکل۱: تصویر حرارتی گرفته شده از بیمار که بنا بر اختلاف دمای سطحی بدن دارای رنگ‌های متفاوتی از R,G,B است.

$$G=\begin{bmatrix} \widetilde{W1 * \vec{X}_i} \\ \widetilde{W2 * \vec{Y}_J} \\ \widetilde{W3 * \vec{R}_{IJ}} \\ \widetilde{W4 * \vec{G}_{IJ}} \\ \widetilde{W5 * \vec{B}_{IJ}} \end{bmatrix}$$

(۳)

$\overrightarrow{R_{I,J}}$و$\overrightarrow{G_{I,J}}$و$\overrightarrow{B_{I,J}}$ نیز شدت روشنایی پیکسل مورد نظر در سه بعدی رنگی است. $\vec{X}_i$ موقعیت پیکسل در جهت عمودی و $\vec{Y}_J$ موقعیت در جهت افقی می‌باشد اما این نکته قابل ذکر است که این دو موقعیت با توجه به سایز تصاویر حرارتی متغیر هستند پس بهتر است نرمالیزه گردند. برای نرمال کردن داده‌ها با توجه به ماتریس کواریانس حاصل متوجه خواهیم شد که داده‌ها ارتباط مستقیمی با یکدیگر نخواهند داشت پس برای نرمالیزه کردن داده‌ها می‌توان از رابطه ۲ استفاده نمود:

$$\hat{x} = \frac{x_{i-\mu}}{\sigma}$$

(۴)

μو σ میانگین و انحراف معیار هر بعد از داده‌ها است، Wها مقادیر وزن‌هایی است که به هر بعد از داده‌ها می‌دهیم نکته قابل توجه این است که چون در تشخیص سرطان به نواحی قرمز تمرکز ما معطوف خواهد شد پس باید در شبکه عصبی به داده‌های $R_{I,J}$ وزن بیشتری دهیم، پس رنگ‌های قرمز باید وزن بیشتری داشته باشند(۱۵). در این شبکه‌ها فرض می‌شود که سیگنال‌های ورودی n تایی در m خوشه قرار می‌گیرند و خوشه‌ها دارای آرایش یک بعدی یا دو بعدی منظم هستند. بردار وزن برای هر خوشه، بردار نمونه الگوهای ورودی پیوند یافته با آن خوشه است. در کاربرد مورد نظر ما، آموزش از طریق کمینه کردن فاصله اقلیدسی بین نمونه‌های ورودی و نمونه‌های اولیه نقشه که توسط تابع همسایگی hi, j وزن‌دهی می‌شوند، انجام می‌گیرد. بنابراین، معیاری که باید کمینه شود به صورت زیر تعریف می‌شود(۱۶):

$$E_{SOM} = \frac{1}{N}\sum_{k=1}^{N}\sum_{j=1}^{M} h_{j,N(x(k))} \left\| w_j - x^{(k)} \right\|^2$$

(۵)

که در رابطه فوق، N تعداد نمونه‌های داده، M تعداد واحدهای نقشه، N(xk) سلول عصبی با نزدیک‌ترین فاصله نسبت به نمونهٔ داده N(xk) است و h تابع همسایگی گوسی است که به صورت رابطه(۴-۵) تعریف می‌شود:

$$h_{i,j}(t) = \exp\left(-\frac{\|r_j - r_i\|^2}{2\sigma_t^2}\right)$$

(۶)

که در آن $\|r_j - r_i\|^2$ فاصله بین واحد نقشه j و نمونه ورودی i بوده و σ_t شعاع همسایگی در زمان t است که به صورت زیر تعریف می‌شود:

$$\sigma(t) = \sigma_0 \exp(-$$

(۷)

که t شماره تکرار آموزش، T حداکثر تعداد تکرار یا طول آموزش است. فاصله بین Xk و بردارهای وزن همه سلول‌های عصبی نقشه محاسبه می‌شوند. سلولN(Xk) که کمترین فاصله را با نمونه ورودی Xk داشته باشد، در مرحله رقابت، برنده خواهد شد:

$$(1\text{-}1)N(X_k) = \arg_{1<j<m} \min\|W_j -$$

(۸)

[Downloaded from ijbd.ir on 2026-11-19]

[DOR: 20.1001.1.17359406.1391.5.4.2.2]

شعاع همسایگی در شروع، مقدار بزرگی است و با افزایش زمان الگوریتم آموزش، در هر بار تکرار، کاهش می‌یابد. بعد از مرحله رقابت، SOM باید بردار وزن سلول برنده N(Xk) و همه سلول‌های همسایه آن را که در شعاع همسایگی RN(Xk) قرار گرفته‌اند را به روزرسانی کند.

اگر $W_{.j} \in R^{N(X_k)}$ آنگاه:

$$w_j(t+1) = w_j(t) + \alpha(t) h_{j,N(xk)}(t)_k \left(x(t) - w_j(t)\right)$$

(۹)

در غیر این صورت:

$$w_j(t+1) = w_j(t)$$

(۱۰)

که در روابط بالا، $h_{j,N(xk)}(t)_k$ تابع همسایگی در زمان t و عامل خطی آموزش در زمان t است که به صورت زیر تعریف می‌شود:

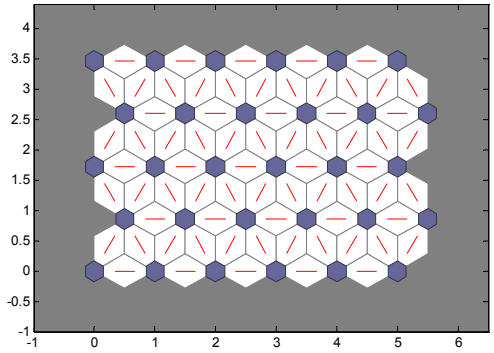
$$\alpha(t) = \alpha_0 \exp\left(-\frac{t}{T}\right)$$

(۱۱)

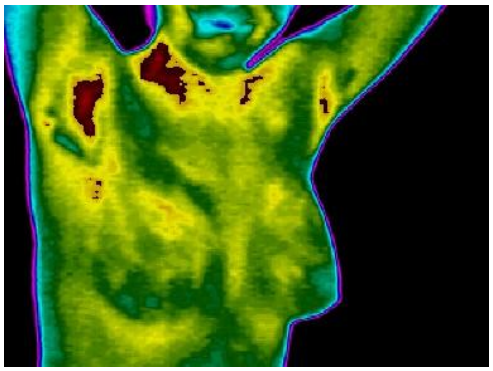
در رابطهٔ فوق α_0 نرخ اولیه آموزش، t شماره تکرار الگوریتم آموزش و T حداکثر تکرار یا طول آموزش است. مرحله آموزش تا زمان پایداری بردارهای وزن تکرار می‌شود یعنی تا زمانی که دیگر تغییری در وزن ورودی‌ها دیده نشود. بعد از مرحله آموزش، شبکه عصبی نقشه خودسازماندهی، قادر خواهد بود n نمونه داده را در قالب m واحد نقشه(خوشه)، خوشه‌بندی کند.

توپولوژی شبکه SOM که استفاده شده است از نوع exagonal است تعداد نورون‌های که برای این شبکه در نظر گرفته شده به صورت ماتریسی ۶ در ۵ است، توپولوژی که حاصل خواهد شد مانند شکل ۲ است.

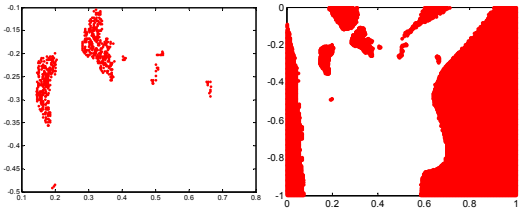
اگر یک نمونه تصویر را وارد شبکه som نماییم و خروجی شبکه را نمایش دهیم به صورت شکل ۳ خواهد بود.



شکل۲: توپولوژی شبکه استفاده شده نورون‌ها در som که به صورت ماتریسی شش ضلعی است.



(الف)



(ب) (ج)

شکل۳: الف) تصویر اولیه حرارتی از فرد بیمار که دارای نواحی قرمز و مشکوک است اما فرد دارای سرطان نیست.
ب) جداسازی اولیه نواحی مشکوک توسط som.
ج) تطابق و هم ارز نمودن تصویر اولیه حرارتی با تصویر جداسازی شده توسط som همراه با ایجاد بزرگنمایی تصویری.

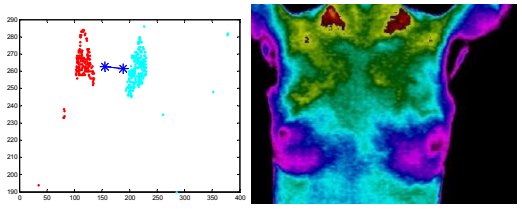
جداسازی نواحی سرطانی:

با دقت و بررسی در نواحی جداسازی شده توسط som متوجه خواهیم شد برخی از نواحی دارای حرارت بالایی هستند مانند نواحی شاهرگ گردن، اما طبیعتاً این امر عادی است پس روشی که در این مرحله به نظر می‌رسد جداسازی دوباره این نواحی است. روش و ایده‌ای که برای حل این مسله مطرح می‌شود این است که از شبکه خود سازمانده استفاده نماییم اما این بار به این صورت که مرکز

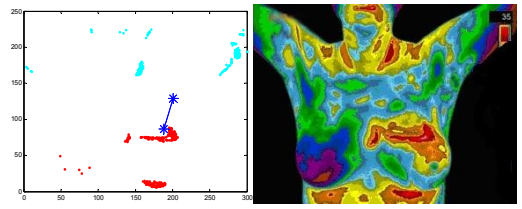
[Downloaded from ijbd.ir on 2026-11-19]

[DOR: 20.1.1.17359406.1391.5.4.2.2]

هر ناحیه جداسازی شده را استخراج نماییم تا نواحی از یکدیگر تفکیک شوند تا در مرحله تشخیص بتوانیم با توجه به مراکز جداسازی شده به تشخیص بپردازیم(۱۷). توپولوژی شبکه SOM که این بار برای استفاده شده است از نوع grid است تعداد نورون‌های که برای این شبکه در نظر گرفته شده به صورت ماتریسی ۲ در ۱ است. اگر مراکز کلاس‌های تصاویر حرارتی مختلف را به دست آوریم مانند شکل ۳ به نکاتی پی خواهیم برد در نمونه شکل ۳: الف که فرد غیر سرطانی است مراکز کلاس‌های که توسط شبکه عصبی خود سازمانده ایجاد شده در نواحی شاهرگ است اما در نمونه سرطانی که مراکز کلاس قرار گرفته در ناحیه پستان صورت پذیرفته پس این مراکز کلاس می‌تواند به عنوان یک ویژگی مهم در تفکیک فرد سرطانی از فرد غیر سرطانی انتخاب گردد.



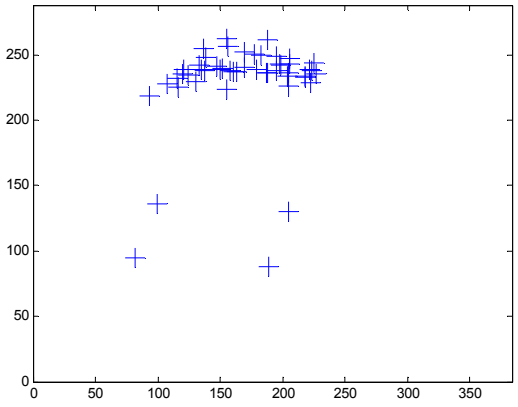
(الف)



(ب)

شکل۳: به دست آوردن مراکزکلاس‌های جداسازی شده که در مرحله خوشه بندی تصاویر حاصل شده است. در نمونه الف فرد غیرسرطانی است که مراکز کلاس‌های جداسازی شده در ناحیه شاهرگ ایجاد شده اما در نمونه ب که فرد سرطانی است مراکز کلاس در ناحیه پستان ایجاد شده است.

اگر تمام تصاویر را در این فرایند قرار دهیم و تمام مراکز کلاس‌هایی که از تصاویر به دست می‌آید رسم نماییم تصویری همانند شکل۴ به دست خواهیم آورد. که با دقت در شکل متوجه می‌شویم چهار مرکز کلاس در ناحیه جدا از مراکزی که غیرسرطانی است ایجاد شده‌اند.



شکل۴: رسم تمام مراکز کلاس‌هایی که از طریق شبکه عصبی خود سازمانده خوشه‌بندی شده است و شامل موارد سرطانی و غیر سرطانی بوده که نمونه‌های سرطانی به وضوح قابل تفکیک و مشاهده هستند.

شناسایی سرطان:

در مرحله قبل ویژگی‌هایی که استخراج شد یعنی مراکز کلاس‌ها می‌تواند به عنوان ویژگی برای شناسایی استفاده گردد. حال می‌توانیم از این ویژگی‌ها برای شناسایی استفاده نماییم. برای این کار از الگوریتم ماشین بردار پشتیبان[۲] کمک می‌گیریم. ماشین‌های بردار پشتیبان را می‌توان از موثرترین روش‌های آموزشی آماری برای دسته بندی داده‌ها به حساب آورد. این روش بر پایه پیدا کردن بردارهای پشتیبان جهت شناسایی مرزهای تصمیم گیری بین کلاس‌های مختلف مبتنی است. از آنجایی که روش ماشین بردار پشتیبان بر اساس پیدا کردن ماشین‌های خطی در فضای ویژگی با ابعاد بالا ولی به صورت غیرخطی مرتبط با فضای داده‌های ورودی عمل می‌کند. به عنوان یکی از روش‌های سریع در دسته‌بندی داده‌ها کاربرد فراوان دارد. علاوه بر این استفاده از این روش برای تشخیص میکروکلسیفیکاسیون‌ها در تصاویر ماموگرافی کاربرد گسترده‌ای دارد(۱۸).

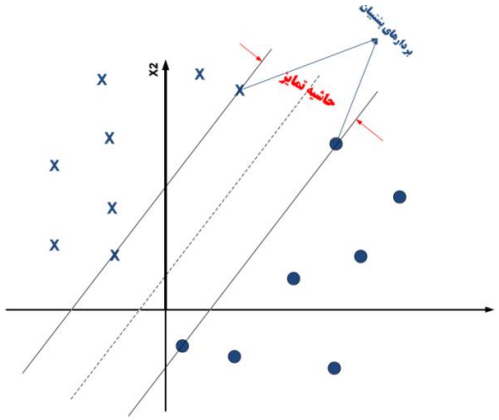
ایده اصلی یک طبقه بندی ماشین بردار پشتیبان در شکل (۵) مشخص شده است. این شکل ساده‌ترین حالت را نشان می‌دهد که در آن بردارهای داده که با علامت x ,o نشان داده شده‌اند را می‌توان توسط یک فوق صفحه از هم جدا کرد. در چنین حالتی می‌توان فوق صفحه‌های بسیاری برای جداسازی دو کلاس تعیین نمود. در بین تمام این فوق صفحه‌ها، طبقه بندی ماشین بردار پشتیبان،

[2] Support Vector Machines (SVM)

[Downloaded from ijbd.ir on 2026-11-19]

[DOR: 20.1001.1.17359406.1391.5.4.2.2]

فوق صفحه‌ای را تعیین می‌کند که بیشترین حاشیه را داشته باشد. چنین عملی با استفاده از روش مینیمم کردن ریسک عملیاتی ممکن است(۱۹).



شکل۵- مدل ساده شده از کلاسیفایر ماشین بردار پشتیبان خطی

اما در حالت کلی‌تر زمانی که داده‌ها به صورت خطی قابل تفکیک نباشند از یک تبدیل غیرخطی برای نگاشت کردن بردار داده x به یک فضای با ابعاد بالاتر(فضای ویژگی) استفاده می‌شود. برای اینکه انتقال به فضای با ابعاد بالاتر باعث خطای over –fitting نشود، ماشین بردار پشتیبان از تابع کرنل استفاده می‌کند که در این کرنل، نگاشت غیر خطی تلویحا جاسازی شده است. با توجه به تئوری cover، تابعی که شرایط کرنل بودن را مهیا کند، شرایط mercer را برقرار می‌سازد(۲۰).

کلاسیفایر ماشین بردار پشتیبان دارای فرم ذیل است:

$$g(x) = \sum_{i=1}^{L_s} \alpha_i d_i K(x_i, x) + \alpha_0$$

(۱۲)

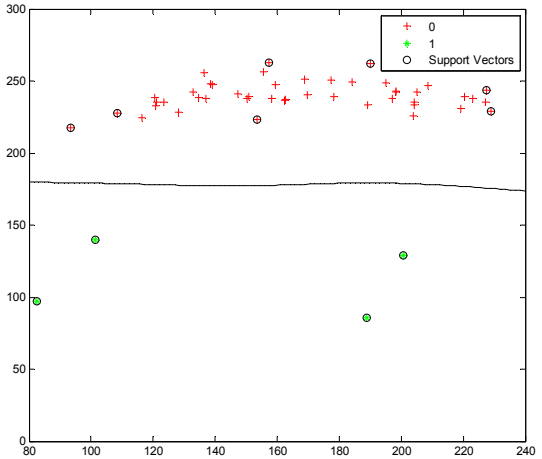
که در آن k تابع کرنل و xi معرف بردار پشتیبان است که داده‌های آموزشی به دست می‌آید. و Ls معرف تعداد بردارهای پشتیبان و di معرف شماره کلاس xi مربوطه است و در نهایت αi ها اعداد ثابتی هستند که در مرحله آموزشی به دست می‌آیند. با توجه به تعریف، بردارهای پشتیبان المان‌هایی از داده‌های آموزشی هستند که دقیقاً بر روی و یا داخل مرزهای تصمیم‌گیری طبقه بندی قرار می‌گیرند. به عبارتی دیگر این بردارها شامل نمونه‌هایی هستند که دسته‌بندی آنها از بقیه نمونه‌ها مشکل‌تر است. باید به این نکته توجه شود که در یک مسئله ماشین بردار پشتیبان تنها کسر کوچکی از داده‌های آموزشی به عنوان بردارهای پشتیبان تعیین می‌گردند. توابع کرنل نقش مرکزی و اصلی را در نگاشت کردن منحنی بردار ورودی به فضای ویژگی با ابعاد بالاتر، برای به دست آوردن تفکیک پذیری بیشتر بازی می‌کنند، در این مقاله از کرنل گوسی شعاعی RBF استفاده شده است که به فرم زیر است(۲۱):

$$K(x,y) = \exp\left(-\frac{\|x-y\|^2}{2\sigma^2}\right)$$

(۱۳)

که $\sigma \geq 0$، ثابتی است که عرض کرنل را تعریف می‌کند، در این مقاله عرض کرنل را ۰/۶ در نظر گرفتیم. این کرنل شرایط mercer را برآورده می‌سازد و یکی از پرکابردترین کرنل‌ها برای ماشین بردار پشتیبان است.

در این مرحله مرکز کلاس‌هایی که به وسیله شبکه عصبی som بدست آورده‌ایم را توسط ماشین بردار پشتیبان (svm) به دو گروه سرطانی و غیر سرطانی جداسازی می‌کنیم.



شکل۶: جداسازی و تفکیک نمونه‌ها توسط ماشین بردار پشتیبان که در آن نمونه‌های سرطانی با رنگ سبز و نمونه‌های غیر سرطانی به رنگ قرمز مشخص شده‌اند.

همان‌گونه که در شکل ۶ مشخص است، ماشین بردار پشتیبان توانسته نمونه‌ها را از یکدیگر تفکیک نماید.

آزمایشات:

برای ارزیابی مدل از پایگاه ایجاد شده کمک گرفته شده که شامل ۲۰۰ مورد است. که در این پایگاه سعی شده تصاویر با رعایت تمام استانداردهای لازم برای ایجاد یک تصویر مناسب حرارتی اعمال شده است. معیارهای ما

[Downloaded from ijbd.ir on 2026-11-19]
[DOR: 20.1.1.17359406.1391.5.4.2.2]

برای ارزیابی مدل میزان دقت، حساسیت و ویژگی تشخیصی برای این پایگاه است. نتایج حاصل از پایگاه در جدول ۱ قابل مشاهده است.

جدول۱: نتایج به دست آمده از حساسیت و ویژگی در دو پایگاه داده تصاویر حرارتی

	پایگاه داده
TN	۱۸۱
TP	۱۴
FN	۴
FP	۱
SEN	٪۹۳
SPE	٪۹۷

بحث

سرطان سینه، رشد مهار نشده سلول‌های غیر طبیعی است که در نواحی مختلف سینه ایجاد می‌شود. این اتفاق ممکن است در بافت‌های مختلف مانند مجاریی که شیر را انتقال می‌دهند، در بافت تولید کننده شیر و در بافت غیر غددی رخ دهد. سرطان سینه پس از سرطان پوست دومین سرطان شایع در زنان است. هر ساله تعداد زیادی از مبتلایان به سرطان سینه تشخیص داده می‌شوند و تعدادی نیز جان خود را از دست می‌دهند. خطر ابتلا به سرطان سینه با افزایش سن بیشتر می‌شود. سرطان پستان در صورتی که به موقع تشخیص داده شود به راحتی قابل درمان است. دلایل قطعی و مسلم سرطان سینه ناشناخته است اما تحقیقات نشان می‌دهد که با بالا رفتن سن در زنان خطر ابتلا هم بیشتر می‌شود. این بیماری در زنان کمتر از سی و پنج سال بسیار کم دیده می‌شود و بیشتر در زنان بالای پنجاه سال دیده شده است. با توجه به موارد ذکر شده غربالگری سرطان پستان از اهمیت ویژه‌ای برخوردار است، غربالگری به این معنی است که در جمعیت زنان سالم حتی قبل از آنکه هر نوع علائمی در پستان تظاهر کند بتوان با انجام آزمایشات و معاینه تغییرات غیر طبیعی را کشف نمود و تصمیمات لازم جهت درمان به کار گرفت. یکی از روش‌های معاینه که می‌تواند بدون ضرر برای خود شخص بیمار باشد استفاده از فناوری دوربین‌های حرارتی است، که در صورتی که بخواهیم این تکنیک را در جهت غربالگری استفاده نماییم باید به موازات آن بر روی هوش مصنوعی و تشخیص‌های هوشمندانه که دارای نقص کمتر هستند بیشتر تحقیق نمود. با توجه به این موارد این تحقیق می‌کوشد که با استخراج بهترین ویژگی‌های تشخیصی از تصویر حرارتی به کمک شبکه‌های عصبی som و طبقه بندی نمودن این ویژگی‌ها به وسیله ماشین بردار پشتیبان به این مهم دست یابد. به طوری‌که مقدار کارایی خروجی ماشین بردار پشتیبان با نتایج ماموگرافی برابر ۱.۷۳ است که نتیجه قابل ملاحظه‌ای صورت پذیرفته است.

درصورتی‌که قرار باشد محدودیت‌های طرح مورد ارزیابی قرار گیرد می‌توان به مواردی از جمله: امکان دسترسی به دوربین‌های حرارتی پیشرفته که به ویژه برای موارد پزشکی طراحی شده‌اند، ایجاد آرامش و دور نمودن استرس از بیمار یا هر گونه عاملی که سبب افزایش غیر عادی دمای بدن گردد، ایجاد نمودن یک پایگاه داده کلی و جامع که شامل تمام موارد خاص مربوط به سرطان پستان که در فرایند آموزش در قسمت هوش مصنوعی، سیستم دچار مشکل نگردد، اشاره نمود. برای پیشنهادات آینده که محقیقن بتوانند این تحقیق را ادامه دهند مقایسه همزمان فرآیند تشخیصی با استفاده از روش‌های از جمله PET–CT یا سونوگرافی و همچنین تلفیق و ترکیب این روش‌ها با تکنیک حرارتی که در مقاله مورد ارزیابی قرار گرفت، می‌باشد.

References

1. Anderson WF, Pfeiffer RM, Dores GM, Sherman ME. Comparison of age distribution patterns for different histopathologic types of breast carcinoma. Cancer Epidemiol Biomarkers Prev. 2006; 15(10):1899-905.
2. Nogaret JM. Breast cancer screening Rev Med Brux 2006; 27(4): S215-7.
3. Piasini P, Parkin DM, Ferlay J. Estimates of the world wide mortality from eighteen major cancers in 1985 , implications for prevention

[Downloaded from ijbd.ir on 2026-11-19]

[DOR: 20.1001.1.17359406.1391.5.4.2.2]

and projection of future burden. Int J Cancer 1993; 55:891-903.

4. Ge J, Sahiner B, Hadjiiski LM, et al. Computer aided detection of clusters of micro calcifications on full field digital mammograms. Med Phys 2006; 33(8): 2975-88.

5. Markey MK, Lo JY, Floyd CE. Differences between Computer-aided Diagnosis of Breast Masses and that of Calcifications. Radiology 2002; 223: 489-93.

6. Hay GA. Medical Image: Formation, Perception and Measurement. The Institute of Physics and John Wiley & Sons, 1976.

7. Watmough DJ. The role of thermographic imaging in breast screening, discussion by c r hill. In Medical Images: formation, perception and measurement 7th L H Gray Conference: Medical Images 1976, 142-58.

8. H.Ghayoumi Zadeh, I.Abaspur Kazerouni, J.Haddadnia. Distinguish Breast Cancer Based On Thermal Features in Infrared Images. Canadian Journal on Image Processing and Computer Vision 2011; 2(6).

۹.حسین قیومی زاده، ایمان عباسپور کازرونی، جواد حدادنیا، مریم هاشمیان. شناسایی سرطان پستان براساس الگوی حرارتی در تصاویر مادون قرمز. فصلنامه بیماری‌های پستان ایران ۱۳۹۰؛ ۴ (۱و۲).

10. Tubiana-Hulin M, Hacène1 K, Martin PM, Spyratos F. Prognostic factor clustering in breast cancer: Biology or chronology. European Journal of Cancer 1995; 31(2): 282-3.

11. Soria D, Garibaldi JM, Ambrogi F, Green AR, Powe D, Rakha E, Macmillan RD, Blamey RW, Ball G, Lisboa PJ, Etchells TA, Boracchi P, Biganzoli E, Ellis IO. A methodology to identify consensus classes from clustering algorithms applied to immunohistochemical data from breast cancer patients. Computers in Biology and Medicine 2010; 40(3): 318-30.

12. Li C, Liu L, Jiang W. Objective function of semi-supervised fuzzy c-means clustering algorithm. In IEEE International Conference on Industrial Informatics 2008; 737-42.

13. Riber S, Dramicanin M, Dramicanin T, Matija L. Classification of breast cancer luminescence data using self organizing mapping nerual network. FME Transactions 2006; 34: 87-91.

14. Yang L, Ouyang Z, Shi Y. A Modified Clustering Method Based on Self-Organizing Maps and Its Applications. Proceedings of the International Conference on Computational Science, ICCS 2012; 9: 1371-9.

15. Scheidta S, Ramseya M, Lancaster N. Radiometric normalization and image mosaic generation of ASTER thermal infrared data: An application to extensive sand sheets and dune fields. Remote Sensing of Environment 2008; 112(3): 920-33.

16. Yeo NC, Lee KH, Venkatesh YV, Ong SH. Colour image segmentation using the self-organizing map and adaptive resonance theory. Image and Vision Computing 2005; 23(12): 1060-79.

17. Tan TZ, Quek C, Ng GS, Ng EYK. A novel cognitive interpretation of breast cancer thermography with complementary learning fuzzy neural memory structure. Expert Systems with Applications 2007; 33 (3): 652-66.

18. Sattlecker M, Baker R, Stone N, Bessant C. Support vector machine ensembles for breast cancer type prediction from mid-FTIR micro-calcification spectra. Chemometrics and Intelligent Laboratory Systems 2011; 107 (2): 363-70.

19. Haykin S. Neural Networks: A Comprehensive Foundation. Second edition, Prentice-Hall 1999.

20. Vapnik V. Statistical Learning Theory. Wiley 1998.

21. Fatih Akay M. Support vector machines combined with feature selection for breast cancer diagnosis. Expert Systems with Applications 2009; 36 (2): 3240-7.

22. Soria D, Garibaldi JM, Ambrogi F, Green AR, Des Powe, Rakha E, Macmillanf RD, Blameyf RW, Balld G, Lisboae PJG, Etchellse TA, Boracchic P, Biganzolic E, Ellisb IO. Diagnosing benign and malignant lesions in breast tissue sections by using IR-microspectroscopy. Computers in Biology and Medicine 2010; 40(3): 318-30.

23. Bezerra LA, Oliveira MM, Rolim TL, Conci A, Santos FGS, Lyra PRM, Lima

[Downloaded from ijbd.ir on 2026-11-19]

[DOR: 20.1.1.17359406.1391.5.4.2.2]

RCF. Estimation of breast tumor thermal properties using infrared images. Signal Processing. Available online 12 June 2012.

24. Borchartt TB, Conci A, Lim RCF, Resmini R, Sanchez A. Breast thermography from an image processing viewpoint: A survey. Signal processing. Available online 29 August 2012.

[Downloaded from ijbd.ir on 2026-11-19]

[DOR: 20.1001.1.17359406.1391.5.4.2.2]