

کاربرد منطق فازی در تشخیص کامپیوتری سرطان سینه

مژگان قنبرپور

گروه کامپیوتر دانشگاه آزاد اسلامی سروستان

M_ghanbarpoor@yahoo.co.

چکیده

دستورالعمل مقاله حاضر در مورد کاربرد شیوه منطق فازی برای رسمی جلوه دادن اصطلاحات بکار رفته در دانشگا

(ACR) توضیح می دهد. در عمل، رادیولوژیست ها بسیاری از اصطلاحات وابسته به تشخیص سرطان سینه را به طور نسبتاً ذهنی تعیین می نمایند. دو ویژگی مهم ذکر شده در واژگان ACR، لوبولاسیون و میکرو لوبولاسیون گرک ها نام دارد. در اینجا شیوه ای برای رسمی جلوه دادن تمایز بین این دو ویژگی و همچنین شرح موارد واسطه بین توده های لخته و ریزلخته پیشنهاد می کنیم. در این مقاله نشان داده شده است که برای حل این مشکل، منطق فازی یک ابزار موثر به شمار می رود. روش رسمی سازی پیشنهاد شده مبنایی ب

نماید:

الف - بازبینی مبسوط مطالعات مقایسه ای مبهم و نامعلوم

ب - استخراج اصول اولیه وابسته از تصویر به صورت خودکار

ج - تشخیص و آشکارسازی توده های لوبولاسیون و ریزلوبولاسیون بر اساس اصول اولیه

کلمات کلیدی

منطق فازی

1- مقدمه

گونه نشانه ای دال بر اندازه یا تعداد حرکات نوسانی و بدون تعریف مفهوم حرکت نوسانی تعریف می کند. کلمات توصیفی در هر طبقه پیوستار وضعیت خوش خیم تا بدخیم را توصیف می کنند. توده لوبولی بیشتر اوقات خوش خیم می باشد، هرچند تعداد معدودی از موارد بدخیم نیز لوبولی خواهند بود. این موارد معمولاً از لحاظ پاتولوژیکی از هم تفکیک شده اند. به علاوه، توده ای با حاشیه ریزلخته در مقایسه با حاشیه نامشخص شانس کمتری بر بدخیم شدن دارد. توده ریزلخته در طبقه شک کم رو به متوسط طبقه بندی خواهد شد و شانس بدخیم شدن آن 10-20 درصد خواهد بود. در این مقاله، مفهوم حرکت نوسانی به تقابل بین حداقل فرورفتگی های مجاور اتلاق شده است. عمق این فرورفتگی ها بین کوچک تا بسیار بزرگ متغیر می باشد. بنابراین، در تحلیل الگوریتمی کامپیوتری رسمی، بدان معناست که اگر توده دارای حرکت نوسانی کوچک، متوسط / بزرگی باشد، آنگاه الگوریتم در گروه لوبولی باید طبقه بندی شود. اما در موقعیت زندگی

روشهای تشخیصی فعلی در ماموگرافی دیجیتالی عمدتاً بر اساس شبکه های عصبی و بدون بهره گیری از منطق فازی عمل می کنند. با این حال، شایان توجه است که این روشها از درجات بی نظمی ودایره وار بودن استفاده می کنند که شبیه به مفاهیم کلیدی در منطق فازی می باشد. از این درجات در شبکه های عصبی به عنوان متغیرهای ورودی استفاده شده است. در این مقاله، از شیوه منطق فازی برای طبقه بندی توده یافت شده در ماموگرام به صورت لخته یا ریزلخته استفاده می کنیم. در تشخیص سرطان سینه، ویژگیهای توده لخته و ریزلخته از اهمیت بسزایی برخوردار می باشد. روش تحلیل پیشنهاد شده بر اساس تعاریف پزشکی دو اصطلاح قبلی عمل می کند که توسط ACR ارائه شده است. بر طبق این لغت نامه، در صورت تناقض با حرکت نوسانی گفته می شود که توده دارای شکل لوبولی می باشد. شایان توجه است که این لغت نامه مفهوم لوبولی را بدون ارائه هیچ

ویژگیهای لوبولی را به درستی شناسایی کنند. این مقاله شیوه شناسایی توده لوبولاسیون / ریزلوبولاسیون پیشنهاد می کند که مسئله عملی و متدولوژیکی را بررسی می کند. این شیوه به مبنایی برای تحلیل و رسمی سازی سایر اصطلاحات لغت نامه ACR نیز تبدیل می شود. شیوه پیشنهاد شده به شیوه ای طراحی شده است که شیوه تصمیم گیری متخصصین را کپی می کند (تقلید می کند). به همین خاطر این مقاله فقط بر ارائه شیوه ای جهت رسمی سازی حالات لوبولاریته و میکرولوبولاریته در توده های یافت شده در ماموگرام تمرکز می کند. این مقاله راجع به توسعه ویژگیها بحث می کند که لوبولاریته و میکرو لوبولاریته و رسمی سازی این ویژگیها از نظر شیوه منطق فازی را توصیف می نماید .

2- رسمی سازی با منطق فازی

در این بخش دو تعریف قبلی را کمی تغییر می دهیم. در این جا می گوییم اگر توده با حرکات نوسانی بزرگ و عمیق طراحی شده باشد (حد فاصل آن مشخص باشد)، آنگاه لوبولی قلمداد می گردد. حاشیه توده در صورتی میکرولوبولاریته می باشد که دارای فرورفتگی های متعدد کوچکی (چرخه) باشد که حرکات نوسانی کوچک و سطحی تولید کنند. در نگاه اول، به نظر می رسد دقت تعاریف را ارتقاء نداد. اما، این قبیل روشهای رسمی سازی از اهمیت بسزایی برخوردار هستند. آنها به ما این امکان را می دهند تا از منطق فازی استفاده کرده و دو تعریف اصلی ACR را به عنوان توابع ثنائیه و اصطلاحات فازی بیان کنیم. ملاحظات فوق دربرگیرنده دو اصطلاح فازی مهم به نام های تعدادی و چندین می باشد. در صورت استفاده از آنها در محیطی با سایر اصطلاحات زبان طبیعی، معنای مشخص و روشنی پیدا می کنند. آنگاه می توانیم مجموعه فازی را با اصطلاحات

فازی $\{few, some, several, many\}$

به ازای تعداد حرکات نوسانی تعریف کنیم. شایان توجه است که تعداد حرکات نوسانی برابر با $0,1,2,3,...$ می باشد. به طور مثال، برای اصطلاح فازی "معدود یا اندک"، تعداد حرکات نوسانی برابر با صفر می باشد. یعنی خانواده نظیر چهار تابع عضویت عبارتند از:

$$\mu_{few}(x), \mu_{some}(x), \mu_{several}(x), \mu_{many}(x)$$

پاره ای از مقادیر نمونه برداری شده توابع عضویت عبارتند از:

واقعی این گونه نمی باشد، زیرا رادیولوژیست اندازه، تعداد حرکات، و عمق آنها را لحاظ می نماید. اما، لغت نامه ACR در تعریف رسمی لختگی، این ویژگیها را عنوان نمی کند. بنابراین احتمال آن می رود که رادیولوژیست های مختلف ادراکات گوناگونی راجع به اندازه و تعداد حرکات کافی برای طبقه بندی شکل توده در گروه لوبولی داشته باشند.

واژه حاشیه ریزلخته (بر طبق لغت نامه ACR) بدان معناست که حاشیه با چرخه کوتاه حرکت نوسانی کرده و بدین طریق حرکات نوسانی کوچکی ایجاد می کند. مجدداً، رادیولوژیست های مختلف ممکن است ادراکات متفاوتی از معنای چرخه کوتاه و حرکات نوسانی کوچک داشته باشند. لغت نامه ACR چارچوب واحدی برای تعریف این اصطلاحات به صورت همسان و هدفمند ارائه نکرده و به همین خاطر رادیولوژیست ها تصمیمات ذهنی و فردی در مورد این ویژگیها اتخاذ می نمایند.

دو مثال زیر نیاز به چارچوب واحد برای تعریف اصطلاحات وابسته به شکل توده ها در ماموگرامها را روشن می کنند:

1. فرض کنید رادیولوژیست در توده یک حرکت نوسانی بزرگ و دو حرکت نوسانی کوچک یافته است. به نظر شما آیا این موضوع به معنای لوبولی یا ریزلوبولی بودن است یا اینکه هر دو ویژگی باهم وجود دارد؟ همچنین فرض کنید در مورد این توده رادیولوژیست دوم به این نتیجه رسیده است که دو حرکت نوسانی بزرگ و یک حرکت نوسانی کوچک وجود دارد. در این مورد نیز این سؤال مطرح می شود: آیا توده از نوع لوبولی یا ریزلوبولی بوده یا اینکه هر دو ویژگی باهم در توده مشاهده می شوند؟

2. فرض کنید که در بعضی از مطالعات، پنج نفر از 10 رادیولوژیست به این نتیجه رسیده اند که توده خاصی از نوع لوبولی است، اما پنج نفر دیگر به نتیجه عکس آن رسیده اند. با استفاده از این تجربه متناقض، چگونه می توان سیستم کامپیوتری را آموزش داد به طوری که بتواند توده لوبولی را آشکار نماید (تشخیص دهد)؟ آیا لازم است این موارد از مجموعه آموزشی حذف شوند؟ در زندگی واقعی موارد مشابهی ظاهر می شوند. در صورت حذف این موارد، سیستم های تشخیص آموزش دیده آنها را به صورت اختیاری تشخیص خواهند داد، هرچند بیشتر آنها نمی توانند

$$\mu_{few}(2)=1/3, \mu_{some}(2)=2/3, \mu_{several}(3)=1, \mu_{many}(2)=0$$

رادیولوژیست های مصاحبه شده از این روش رسمی سازی راضی بودند. اگرچه باکمک مقادیر عددی توابع عضویت فوق می توان استدلال نمود، اما در این جا مسئله اصلی آن است که تعیین کمی مفاهیم فازی مهم برای طبقه بندی همسان و هدفمند توده ها در گروه های لوبولی یا ریزلوبولی میسر می باشد. سپس، معنای اصطلاحات مجموعه فازی $\{small, big\}$ را تعریف می کنیم.

این مجموعه در تعریف اندازه حرکات نوسانی نقش مهمی ایفا می کند. اولاً به مقیاس کافی برای اندازه گیری طول حرکت نوسانی نیاز داریم. به همین خاطر طول یک حرکت نوسانی در اصطلاحات نسبی را در نظر می گیریم، زیرا توده های مختلف هم اندازه نمی باشند. به طور مثال، حرکت نوسانی به طول 3 میلی متر در توده بزرگ، ریزلوبولی خوانده می شود، در حالیکه توده کوچکی با همین اندازه حرکت نوسانی را می توان در گروه لوبولی طبقه بندی نمود. بنابراین، ابتدا L ، ماکزیمم طول توده را محاسبه می کنیم. این شیوه به ما امکان برآورد طول حرکت نوسانی به صورت تابع L را می دهد. با توصیف حرکات نوسانی در مقیاس تعیین شده بر حسب طولی نسبی حرکت نوسانی می توانیم اصطلاحات فازی "کوچک" و "بزرگ" را به صورت رسمی مطرح کنیم. از آنجایی که توده ها دارای عمق لوبولاریته متغیری هستند، در نتیجه توابع عضویت فازی را می توان با توجه به عمق (سطحی یا عمیق) حرکات نوسانی نیز تعریف نمود. بنابراین، معیار نسبی عمق لوبولاریته را معرفی می کنیم که به صورت کسر ماکزیمم طول (L نشان داده شده است) توده تعریف شده است. این مرحله شبیه به مرحله شرح داده شده در مجموعه های فازی پیشین می باشد. مفهوم توده لوبولاریته را به شرح ذیل می توان رسمی جلوه داد: توده در صورتی لوبولاریته خوانده می شود که با طول و عمق بیشتر از $L/4$ حداقل 3 حرکت نوسانی داشته باشد. مفهوم حاشیه توده میکرولوبولاریته را نیز می توانیم به صورت رسمی ارائه کنیم. حاشیه توده در صورتی میکرولوبولاریته خوانده می شود که حداقل 6 حرکت نوسانی با طول و عمق کمتر از $L/12$ داشته باشد. بر اساس وابستگی متقابل مفاهیم اندازه، عمق و تعداد حرکات نوسانی تعاریف مذکور مطرح شده و از آنها می توان برای

تعیین مفاهیم توده های لوبولاسیون و میکرولوبولاسیون به صورت عینی و همسان استفاده نمود. هر حرکت نوسانی با طول و عمق مختص خود نشان داده شده است. کلیه حرکات نوسانی نشان داده شده بزرگ و عمیق می باشند. بنابراین، کلیه توابع عضویت برابر با 1.00 بوده و بر طبق روش رسمی سازی بکار رفته می توان به این نتیجه رسید که توده لوبولاریته می باشد. تعاریف قبل شرایطی برای طبقه بندی بعضی از توده ها در گروه لوبولاسیون و میکرولوبولاسیون فراهم می آورند در صورتی که توده حداقل نه حرکت نوسانی داشته باشد (از این تعداد سه مورد لوبولاریته و شش مورد میکرولوبولاریته می باشد). در اینجا سه حرکت نوسانی بزرگ و عمیق را انتخاب کرده و مینیموم مقادیر تابع عضویت آنها برای جملات "بزرگ" و "عمیق" را محاسبه می کنیم. این مقدار را درجه لوبولاریته (یا DL) می نامیم. اگر تعداد حرکات نوسانی کمتر از 3 باشد، آنگاه توابع عضویت طول و عمق را با تابع عضویت تعداد حرکات نوسانی ترکیب می کنیم. در این ترکیب، مینیمم سه مقدار را بر طبق شیوه منطق فازی استاندارد محاسبه می کنیم. سپس استدلال های مطرح شده در مورد مینیمم را در بخش بعدی مورد تجزیه و تحلیل قرار می دهیم. شایان توجه است که توابع عضویت فازی پیشنهاد شده نشانه می باشند. شکل درست آنها را می توان به کمک داده های تاریخی یا شیوه ای تعیین نمود که رادیولوژیست ها با آن موافقت داشته باشند. در بخش بعدی، ایده های DL و DM به طور رسمی مطرح می گردد.

3- درجات لوبولاریته و ریزلوبولاریته

رادیولوژیست ها از یک شیوه غیر رسمی برای تعیین لوبولاریته و ریزلوبولاریته (لوبولاریته میکرو) توده استفاده می کنند. به منظور حفظ ثبات و همسانی در این ارزیابی ها و افزایش عینیت، نیاز داریم این مفاهیم را به صورت رسمی ارائه دهیم. ابتدا دو توده را در نظر می گیریم. توده اول دارای حرکات نوسانی عمیق و توده دوم دارای حرکات نوسانی سطحی می باشد. برای رسمی نمودن این وجه تمایز، از معیارهای مختلفی می توان استفاده نمود. مقادیر $\mu_{deep}(d1), \mu_{deep}(d2)$ با

استفاده از تابع عضویت محاسبه شده اند، به گونه ای که در مورد یک حرکت نوسانی مشخص، دو معیار پیشین به یک درجه لوبولاریته تبدیل شده اند. به خاطر بیاورید که در مورد طول و عمق حرکات نوسانی نیز از این توابع عضویت استفاده می کنیم. این کار با جانشینی جملات بزرگ برای عمیق و کوچک برای سطحی انجام می شود.

$$\text{سپس } \min\{\mu_{deep}(d1), \mu_{deep}(d2)\}$$

را محاسبه می کنیم که می توان آن را درجه عمق حرکت نوسانی در نظر گرفت. یعنی:

$$\mu_{deep}(undulation) = \min\{\mu_{deep}(d1), \mu_{deep}(d2)\}$$

همچنین، درجه سطحی بودن یک حرکت نوسانی از رابطه زیر بدست می آید:

$$\mu_{shallow}(undulation) = \min\{\mu_{shallow}(d1), \mu_{shallow}(d2)\}$$

طول حرکت نوسانی (به عبارتی U_1) به صورت طول حاشیه توده بین نقاط A و B اندازه گیری شده است. حال می توان درجه لوبولاریته (DL) توده را از رابطه زیر تعریف کرد:

$$DL(mas) = \min\{\mu_{some}(k), \min_{k \geq 1} \{\mu_{big}(U_i), \mu_{deep}(U_i)\}\} \quad (1)$$

در این رابطه U_1, U_2, \dots, U_k حرکات نوسانی را نشان می دهند، به نحوی که:

$$\min_{k \geq 1} \{\mu_{big}(U_i), \mu_{deep}(U_i)\} \geq 0.50$$

همچنین، درجه ریزلوبولاریته (DM) توده ای با k حرکت نوسانی از رابطه زیر بدست می آید:

$$DM(mas) = \min\{\mu_{several}(k), \min_{k \geq 1} \{\mu_{small}(U_i), \mu_{shallow}(U_i)\}\} \quad (2)$$

در این رابطه U_1, U_2, \dots, U_k حرکات نوسانی را نشان می دهند، به نحوی که:

$$\min_{k \geq 1} \{\mu_{small}(U_i), \mu_{shallow}(U_i)\} \geq 0.5$$

در موارد $k=0$ ، داریم

$$\mu_{some}(k)=0, \mu_{several}(k)=0$$

بنابراین، درجه لوبولاریته و ریزلوبولاریته برابر با 0 می باشد، به عبارتی نتیجه بدست آمده همان چیزی است که انتظار آن را داشتیم. عموماً با تنظیم معادلات 1 و 2 شماری استدلال نظری و آزمایشی مطرح شده است. اما می توان از استدلال های دیگری نیز در

مورد مسئله ماموگرافی استفاده نمود. سیستم تشخیص کامپیوتری سرطان سینه ماموگرافی با ویژگیهای مشکوک را تشخیص نمی دهد. مشکوک بودن ویژگی با درجه بین 0 و 1 بیان می گردد، بالاترین درجه شک 0.50 می باشد. مقادیر DL و DM از جمله نمونه های این قبیل درجات محسوب می شوند. در مورد این قبیل ویژگیهای نامعلوم (مشکوک)، سیستم CAD حضور یک ویژگی خاص اما با پایایی نسبی را پیشنهاد می کند. پایایی ممکن است بسیار پائین بود. همچنین، درجه پایایی به مقادیر خاص معیارهای DL و DM بستگی دارد. به همین خاطر معادلات بکار رفته برای تعریف DL و DM اهمیت بیشتری پیدا می کنند. فرض کنید حرکت نوسانی پنج تابع عضویت اول برابر با 1.00 و تابع ششم برابر با 0.60 است (به عبارتی

$$(\mu_{deep}(undulation3)=0.60)$$

آنگاه معادله 1 برآورد بدبینانه ای حاصل می کند، به عبارتی درجه قطعیت پائین برای حضور لوبولاریته،

$DL=0.6$. از طریق جانشینی در معادله 1، مینیمم عملیات برای نیل به ماکزیمم سنجش خوشبینانه ای ارائه می دهد، به عبارتی درجه بالای لوبولاریته، $DL=1.00$. در آخرین برآورد خوشبینانه، اطلاعات اخطار دهنده را نادیده گرفته و آن را از دست می دهیم (به عبارتی اینکه $\mu_{deep}(undulation1)=0.60$). مقدار

0.60 حاکی از آن است که این مسئله باید به تفصیل مورد مطالعه قرار گیرد. اما، در صورت استفاده از عملیات مینیمم بدبینانه در معادلات 1 و 2، اطلاعات اخطار دهنده از دست نمی رود. به همین خاطر در مورد سئوالات حساس در مورد تشخیص سرطان، می بینیم که استراتژی بدبینانه مطمئن ترین راه می باشد. همچنین عباراتی با درجه پائین پایایی را نیز به عنوان پیشنهاد مقدماتی در نظر می گیریم که تغییر مجموعه ویژگیها به سطح بالاتر جزئیات جهت ارزیابی کامل پیچیده مسئله را منعکس می کنند. بعضی از آزمایشات دویی و و نشان داده است که مسائل نسبتاً ساده را می توان در فضای ویژگی کوچک تشخیص داد. در مورد مسائل پیچیده تر، به نمونه آموزشی تائید شده از لحاظ پاتولوژیکی با ویژگیهای بیشتر و مخصوصاً روش تشخیصی طراحی شده نیاز داریم. سیستم CAD، بر اساس

شیوه توصیف شده، به قابلیت های تغییر دهنده مجهز می باشد.

4- نتیجه گیری

رادیولوژیست ها اغلب بسیاری از ویژگیهای وابسته به تشخیص سرطان سینه را به صورت ذهنی مشخص می کنند. در اینجا شماری از ویژگیهای مهم را با کمک واژه نامه عکسبرداری از سینه ACR به عبارتی گرهک های لوبولاسیون و ریزلوبولاسیون به صورت رسمی ارائه کرده ایم. این کار مبنایی برای سه مرحله بعد به شمار می رود.

1. اعتباریابی وسیع رادیولوژیکی
2. تشخیص خودکار لوبولاسیون / ریزلوبولاسیون در عکس ماموگرافی
3. رسمی سازی مشابه سایر جملات با کمک واژه نامه عکسبرداری از سینه
مطالعه انجام شده نشان می دهد که منطق فازی ابزاری موثر برای رسیدگی به این نوع مسئله پزشکی به شمار می رود. شایان توجه است که واژه نامه عکسبرداری از سینه ACR محتوی مفاهیمی است که در شیوه منطق فازی شبیه به تحلیل لوبولاسیون پیشنهاد شده می توان آن را تعریف نمود. اما، با توجه به کارهای قبلی در خصوص آهکی شدن سینه، ویژگیهای مختلف مطرح شده در واژه نامه با مسائل مختلفی روبرو هستند که به راه حل های مناسب نیاز دارند. تحلیل کلیه مفاهیم مندرج در واژه نامه ACR (تقریباً 30) خارج از حیطه این مقاله بوده و به مقاله های مشابه زیادی نیاز دارد. اما در تشخیص سرطان سینه، هدف مذکور از اهمیت بسزایی برخوردار می باشد، زیرا همان گونه که شرح داده شده است، روشهای هوش مصنوعی سنتی و روشهای آماری تشخیص الگو معتبر نمی باشد. شیوه منطق فازی پیشنهاد شده امکان پذیر و موثر می باشد زیرا این تیپ شیوه منطق فازی در بسیاری از بخشهای دیگر موفق بوده است. کاربرد آن در زمینه تشخیص سرطان سینه یکی از این بخشها می باشد. شیوه منطق فازی پیشنهاد شده، وقتی با موفقیت کاربرد منطق فازی در بسیاری از حوزه های دیگر در نظر گرفته شده باشد، راهی برای تشخیص موثر و زود هنگام سرطان سینه باز می نماید.

مراجع

- [1] Kenneth H. Rosen. Discrete Mathematics and Its Applications, 4th Edition. WCB/McGraw-Hill, 1999
- [2] William A. Dembski, and Robert J. Marks II, "Bernoulli's Principle of Insufficient Reason and Conservation of Information in Computer Search," *Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics San Antonio, TX, USA*, 2647-2652 (October 2009).
- [3] Richard v. Sternberg, "DNA Codes and Information: Formal Structures and Relational Causes," *Acta Biotheoretica*, Vol. 56(3):205-232 (September, 2008).
- [4] Ian Stewart. "Mathematical Recreations", *Scientific American*. February 1993
- [5] Douglas D. Axe, "The Limits of Complex Adaptation: An Analysis Based on a Simple Model of Structured Bacterial Populations," *BIO-Complexity*, Vol. 2010(4):1-10