

Research Paper

Mammographic Image Processing for Classification of Breast Cancer Masses by Using Support Vector Machine Method and Grasshopper Optimization Algorithm



Naser Safdarian¹ , Shadi Yousefian Dezfoulinejad² 

1. Young Researchers and Elite Club, Tabriz Branch, Islamic Azad University, Tabriz, Iran.

2. Department of Biomedical Engineering, Faculty of Engineering, Dezfoul Branch, Islamic Azad University, Dezfoul, Iran.



Citation: Safdarian N & Yousefian Dezfoulinejad Sh. [Mammographic Image Processing for Classification of Breast Cancer Masses by Using Support Vector Machine Method and Grasshopper Optimization Algorithm (Persian)]. Journal of Arak University of Medical Sciences (JAMS). 2020; 23(2):246-263. <https://doi.org/10.32598/JAMS.23.2.4672.1>

 <https://doi.org/10.32598/JAMS.23.2.4672.1>



Article Info:

Received: 13 Nov 2019

Accepted: 21 Jan 2020

Available Online: 01 Jun 2020

Key words:

Mammographic image processing, Breast cancer, Geometric features extraction, Support Maachine, Grasshopper optimization algorithm

ABSTRACT

Background and Aim Breast cancer is the abnormal cell growth in the breast. In both benign and malignant masses, there is rapid and high cell growth. Nowadays, due to the development of technologies, the diagnosis of diseases has become non-invasive and physicians attempts to diagnose the disease without surgery and based on internal organ images.

Methods & Materials In this study, by using images prepared from the Digital Database for Screening Mammography (DDSM), a new method is proposed for detecting cancerous masses in the mammographic images using geometric features extraction and optimization of Support Vector Machine (SVM) parameters to classify breast cancer masses automatically. First, images were pre-processed and then boundaries were determined using threshold method. Next, morphological operators were used to improve these boundaries and the segmentation of images was carried out to classify cancerous masses. Finally, by using the SVM parameter optimization method, Grasshopper Optimization Algorithm (GOA), and 4-fold crossvalidation method, data were classified into two groups of benign and malignant (cancer) masses.

Ethical Considerations Images from DDSM database were used in this research, all images are open access in this database.

Results The accuracy, sensitivity and specificity values for applying the Radial Basis Function (RBF) kernel in SVM classifier (before optimization process) were obtained 97%, 100% and 96, respectively. After optimization of SVM parameters by the GOA, it was reported 100% for all accuracy, sensitivity and specificity indices for applying linear kernel function, indicating the high accuracy of the proposed method. The average values of accuracy, sensitivity and specificity indices for applying all three SVM kernel functions after optimization were 95.83, 100 and 94.81%, respectively.

Conclusion The extracted geometrical features from breast cancer masses are highly efficient for model training and the diagnosis of breast cancer. The GOA could improve the overall accuracy of the proposed method by optimizing the SVM parameters. The results showed the higher performance of the proposed method compared to other methods.

Extended Abstract

Introduction

B

reast cancer is a type of cancer that, due to the abnormal growth of cells, causes a lump

in the breast tissue. According to the World Health Organization (WHO), this cancer affects 2.1 million women each year, and also causes the highest number of cancer deaths among women [1]. Mammography X-ray imaging is one of the most common methods used by radiologists to diagnose and screen for breast cancer and to determine the pres-

* Corresponding Author:

Naser Safdarian, PhD.

Address: Department of Biomedical Engineering, Faculty of Engineering, Dezfoul Branch, Islamic Azad University, Dezfoul, Iran.

Tel: +98 (936) 8173934

E-mail: naser.safdarian@yahoo.com

ence of lumps and cysts. In mammography images, very fine calcareous particles are usually seen as noisy particles, and the masses have very low light intensity, making it difficult for radiologists and physicians to detect. Given that accurate and timely diagnosis of cancerous masses, as well as its various types, is of particular importance in the health of individuals in society, the difficulty of diagnosing breast cancer masses, which is mainly associated with diagnostic errors, should be addressed by researchers. In this study, we introduce a new and automatic method to diagnose and detect breast cancer masses with high accuracy. For this purpose, after pre-processing and detecting the borders of the cancerous masses from mammography images, a number of features were extracted from the detected masses and in the end, the feature matrix was applied to the Support Vector Machine (SVM) classifier input.

Materials and Methods

The images used in this study were collected from the Digital Database for Screening Mammography (DDSM) database [17]. First, for pre-processing of raw images, a 3×3 median filter was applied on digital mammography images to remove noise using MATLAB software. Then, the threshold method was used to extract the cancerous masses. Since the mass edge extracted by the threshold method had inward direction, the brightness of the pixels around the edge was expanded towards the center by using the Dilation operator. After detecting the area of breast cancer masses, we extracted 19 structural features from this area by MATLAB software. Finally, using SVM parameter optimization method by Grasshopper Optimization Algorithm (GOA), as well as using 4-fold cross validation method, data were divided into two categories of benign and cancer.

Results

The values of accuracy, sensitivity, and specificity (along with their variances) resulting from the use of data (benign and cancer) classified by the SVM method using three kernel functions of Linear, Radial Basis Function (RBF) and Polynomial were presented in tables. The final results after applying the GOA were also shown in a separate table. The used training data was 85% and 15% of the data were considered as test data. In 4-fold crossvalidation method, the number of programs executed per kernel function was 100 times. The best results of accuracy, sensitivity and specificity indicators (features) for using RBF kernel function in SVM classifier (before process) were obtained 97%, 100% and 96%, respectively. For linear function after optimization of SVM parameters by GOA, it was obtained 100% for all accuracy, sensitivity, and specificity indicators, which shows the high accuracy of the proposed method.

The average values of accuracy, sensitivity and specificity indices for all three SVM kernel functions after applying the optimization algorithm were 95.83, 100 and 94.81%, respectively.

Discussion

Some advantages of this study include a large number of features extracted from masses detected from mammography images, the use of GOA to more accurately determine the type of breast tissue cancer mass, and high speed and accuracy of the proposed algorithm. The boundaries of cancerous tumors were extracted with high accuracy, and finally the classification was performed using simple morphological features. To our knowledge, no study has previously used the optimization methods in the final classification stage. So, it can be said that this the first study that use the GOA to optimize the kernel parameters of different SVM classifiers. This can be the advent of new methods in improving various classification processes in a variety of medical diagnoses. After the detection and diagnosis of breast cancer masses that was performed with high accuracy in this study, according to the morphological and simple features of cancer masses, classification operation was performed well and with high accuracy. The results of this study show the higher performance of the proposed method compared to other methods used in previous related studies [6-16].

Ethical Considerations

Compliance with ethical guidelines

Images from DDSM database were used in this research, all images are open access in this database.

Funding

This research did not receive any grant from funding agencies in the public, commercial, or non-profit sectors.

Authors' contributions

Conceptualization, research, methodology: Naser Safdarian; Data collection, resources, writing-original draft: Shadi Yousefian Dezfoulinejad; writing - review & editing: Naser Safdarian.

Conflicts of interest

The authors declared no conflict of interests.

Acknowledgements

The authors would like to thank the Young Researchers and Elite Club of Islamic Azad University, Tabriz branch.

پردازش تصاویر ماموگرافی به منظور طبقه‌بندی توده‌های سرطان سینه با استفاده از روش ماشینی بردار پشتیبان با بکارگیری الگوریتم بهینه‌سازی ملخ

* ناصر صفدریان^۱، شادی یوسفیان دزفولی نژاد^۲ 

۱. باشگاه پژوهشگران جوان و نخبگان، واحد تبریز، دانشگاه آزاد اسلامی، تبریز، ایران

۲. گروه مهندسی پزشکی، دانشکده فنی و مهندسی، دانشگاه آزاد اسلامی واحد دزفول، دزفول، ایران

چکیده

زمینه و هدف: سرطان سینه به دلیل رشد غیرقابل کنترل سلولهای غیرطبیعی در سینه ایجاد می‌شود. در هر دو نوع تومورهای سرطانی خوشخیم و بدخیم، رشد سریع و زیاد سلولها وجود دارد. امروزه، با توجه به گسترش تکنولوژی، تشخیص بیماریها از حالت تهاجمی خارج شده و تلاش پزشکان و متخصصان بر این است که تشخیص بیماری را بدون جراحی و از روی تصاویر ارگانه‌های داخلی انجام دهند.

مواد و روشها: پژوهش حاضر مطالعه‌ای بر اساس بررسی اطلاعات پایگاه داده‌ی تصاویر دیجیتال ماموگرافی (DDSM می‌باشد، که روشی جدید برای تشخیص و آشکارسازی توده‌های سرطانی تصاویر ماموگرافی با استفاده از استخراج ویژگیهای هندسی و بهینه‌سازی پارامترهای الگوریتم ماشینی بردار پشتیبان (SVM)، به منظور طبقه‌بندی خودکار توده‌های سرطان سینه ارائه شده است. در ابتدا پیش‌پردازش تصاویر ماموگرافی انجام می‌شود، سپس با استفاده از روش آستانه‌گذاری، مرزهای توده‌ها مشخص شد و سپس با عملگرهای مورفولوژی این مرزها بهبود یافته و در نهایت قطعه‌بندی تصاویر برای طبقه‌بندی نوع توده‌های سرطانی بافت سینه انجام شد. در مرحله نهایی با استفاده از روش بهینه‌سازی پارامترهای طبقه‌بندی کننده SVM با کمک الگوریتم بهینه‌سازی ملخ، و استفاده از روش اعتبارسنجی $fold-4$ طبقه‌بندی داده‌ها به دو دسته‌ی خوشخیم و بدخیم انجام شد.

ملاحظات اخلاقی: در این پژوهش از تصاویر پایگاه داده‌ی DDSM استفاده شده است. لازم به ذکر است که تصاویر موجود در این پایگاه داده بصورت رایگان جهت دسترسی وجود دارند.

یافته‌ها: بهترین نتایج شاخص‌های صحت، حساسیت و اختصاصیت (ویژگی) به ازای به‌کارگیری تابع کرنل (RBF) پیش از انجام عملیات بهینه‌سازی پارامترهای طبقه‌بندی کننده SVM به ترتیب برابر با 0.97% ، 1.00% و 96% بدست آمد. همچنین، بهترین نتایج شاخص‌های صحت، حساسیت و اختصاصیت به ازای به‌کارگیری تابع کرنل خطی، پس از انجام عملیات بهینه‌سازی پارامترهای طبقه‌بندی کننده SVM با الگوریتم بهینه‌سازی ملخ، به ترتیب برابر با 1.00% ، 1.00% و 1.00% بدست آمد که نشان‌دهنده دقت بالای روش پیشنهاد شده در این مقاله است. متوسط مقادیر شاخص‌های صحت، حساسیت و اختصاصیت به ازای بکارگیری هر سه تابع کرنل SVM و پس از اعمال الگوریتم بهینه‌سازی، به ترتیب برابر با $0.95/83$ ، 1.00 و $94/81$ درصد به دست آمد.

نتیجه‌گیری: بر اساس روش پیشنهاد شده در این مقاله، ویژگیهای هندسی به‌دست آمده از بافت توده‌های سرطانی سینه جهت آموزش مدل و تشخیص نوع توده سرطان سینه دارای کارایی بالایی است و روش بهینه‌سازی ملخ با انجام بهینه‌سازی پارامترهای طبقه‌بندی کننده، دقت کلی تشخیص روش ارائه شده را بهبود بخشیده است. بنابراین نتایج حاصل از این پژوهش نشان‌دهنده عملکرد بالای روش پیشنهاد شده در مقایسه با سایر پژوهشهای پیشین انجام شده در این زمینه است.

اطلاعات مقاله:

تاریخ دریافت: ۲۲ آبان ۱۳۹۸

تاریخ پذیرش: ۰۱ بهمن ۱۳۹۸

تاریخ انتشار: ۱۲ خرداد ۱۳۹۹

کلیدواژه‌ها:

پردازش تصاویر ماموگرافی، سرطان سینه، استخراج ویژگی‌های هندسی، ماشینی بردار پشتیبان، الگوریتم بهینه‌سازی ملخ

مقدمه

می‌کند [۱]. بر اساس تحقیقات انجام شده در سال ۲۰۱۸ تعداد ۶۲۷۰۰۰ زن به علت سرطان سینه جان خود را از دست داده‌اند [۱]. در واقع، این سرطان تقریباً 15% از کل مرگومیرهای ناشی از انواع سرطان را در میان زنان تشکیل می‌دهد. این بیماری در ایران نیز شایع‌ترین سرطان و سومین علل مرگومیر در میان زنان است [۲]. لذا در صورتی که بتوان توسط روشی، این بیماری را تشخیص داد می‌توان شاهد کاهش مرگومیر انسان‌ها بود. روش

سرطان سینه نوعی سرطان است که به دلیل رشد غیرقابل کنترل سلول‌های غیرطبیعی سبب ایجاد توده در بافت پستان میشود. طبق گزارش سازمان بهداشت جهانی (WHO) هر ساله این سرطان ۲/۱ میلیون زن را تحت تأثیر قرار میدهد، همچنین بیشترین مرگومیر ناشی از سرطان را در میان زنان ایجاد

* نویسنده مسئول:

ناصر صفدریان

نشانی: باشگاه پژوهشگران جوان و نخبگان، واحد تبریز، دانشگاه آزاد اسلامی، تبریز، ایران.

تلفن: ۸۱۷۳۳۹۳۴ (۹۳۶) +۹۸

پست الکترونیکی: naser.Safdarian@yahoo.com

سالم، و بافت دارای توده‌های خوش خیم و بدخیم، پیش‌پردازش تصویر، شناسایی و استخراج بخش‌هایی از بافت پستان که دارای توده‌های سرطانی است، گسترش ناحیه‌ی مشکوک به توده تا مرزهای توده‌ی آشکارسازی شده (و یا بالعکس؛ یعنی ابتدا آشکارسازی مرزهای توده و سپس تشخیص نهایی پیکسل‌های تصویر که حاوی توده سرطانی در بافت مورد نظر هستند)، استخراج ویژگی‌های مختلف (مانند ویژگی‌های مورفولوژیکی، هندسی، مبتنی بر بافت، ویولت و ...) از توده، و در نهایت انجام طبقه‌بندی نهایی ویژگی‌های استخراج شده برای تعیین نوع توده‌های سرطانی (خوش خیم یا بدخیم) توسط الگوریتم‌های هوشمند می‌باشند. در ادامه‌ی این بخش به معرفی برخی از مهمترین تحقیقات انجام شده در این حوزه اشاره شده است.

نیک روان و همکاران در سال ۱۳۹۵ روشی نوین جهت تشخیص توده‌های سرطان سینه ارائه کردند [۶]. این پژوهش شامل سه مرحله‌ی پیش‌پردازش، استخراج ویژگی و طبقه‌بندی بود. مرحله‌ی پیش‌پردازش شامل حذف نویز و پس‌زمینه از تصویر است. ویژگی‌های استخراج شده شامل ویژگی‌های بافتی بدست آمده از ماتریس هم‌رخداد و هیستوگرام توده و ویژگی‌های هندسی است. تعداد ۲۴ ویژگی استخراج شده از هر توده با استفاده از شبکه عصبی پرسپترون به دو دسته خوش خیم و بدخیم طبقه‌بندی شد. صحت بدست آمده از سیستم طراحی شده ۹۳/۳٪ گزارش شده است [۶].

ایمان عباس پور کازرونی و همکاران در سال ۱۳۹۲ روشی هوشمند برای جداسازی تصاویر ماموگرافی بر اساس میزان چگالی بافت و توده‌ها پیشنهاد دادند [۷]. مدلی برای کاهش اطلاعات موجود در تصاویر براساس اجزای اولیه به‌صورت دو بعدی ارائه شده است که با کاهش داده‌های اضافی موجود در تصاویر ماموگرافی به دقت و سرعت دسته‌بندی تصاویر کمک می‌کند. این مدل بر روی پایگاه داده جامع‌ه تحصیل ماموگرافی MIAS اعمال شده و دقت نهایی بدست آمده در حدود ۹۰٪ گزارش شده است [۷].

سعیده ناصری نوروزانی و همکاران در سال ۱۳۹۷ روشی برای تعیین مرحله بالینی بیماران مبتلا به سرطان پستان را ارائه کردند [۸]. در این پژوهش از دو پایگاه داده‌ی SEER و پایگاه داده‌ی محلی استفاده شده است. روش اصلی این پژوهش از دو بخش عمده تشکیل شده است. ابتدا مدلی مکانیزه با استفاده از سیستم TNM و الگوریتم کلونی مورچگان جهت تشخیص مرحله سرطان پستان بر اساس پایگاه داده‌ی SEER ایجاد شده است، و سپس هدف بومی‌سازی و ایجاد مدل تشخیص سرطان پستان بر اساس یک دیتاست محلی انجام گردیده است. دقت بدست آمده برای پایگاه داده‌ی SEER و برای پایگاه داده‌ی محلی بعد از اعمال الگوریتم کلونی مورچگان به ترتیب ۹۹/۴۳٪ و ۹۸/۹۵٪ بدست آمده و گزارش شده است [۸].

تصویربرداری ماموگرافی توسط اشعه X یکی از رایجترین روشهای مورد استفاده رادیولوژیست‌ها برای تشخیص و غربالگری سرطان سینه و تعیین وجود توده‌های سرطانی و کیست‌ها است. با این حال تفسیر تصاویر ماموگرافی دشوار است و طبق آمار مرکز ملی سرطان در آمریکا، ۱۰ تا ۳۰ درصد توده‌های موجود در سینه توسط رادیولوژیست قابل مشاهده نیستند [۳-۴]. توده‌ها و ذرات ریز آهکی که ذرات بسیار کوچک کلسیم هستند، از نشانه‌ها و علائم سرطان در تصاویر ماموگرافی هستند و تشخیص درست این علائم بسیار سخت و دشوار است. بطور کلی توده‌ها به دو دسته‌ی خوش خیم و بدخیم طبقه‌بندی می‌شوند، که هر کدام از لحاظ شکل ظاهری در تصویر دارای ویژگی‌های خاصی هستند. توده‌های خوشخیم بیضوی شکل، دارای لبه‌های مشخص و بدون زاویه هستند؛ در حالی که توده‌های بدخیم دارای شکل غیریکنواخت و غیرمنظم، لبه‌های نامشخص، زاویه‌دار و غالباً شبیه چکمه یا لوبورال^۱ هستند [۵].

در تصاویر ماموگرافی، ذرات بسیار ریز آهکی معمولاً بصورت ذرات نویزی مشاهده شده و توده‌ها دارای شدت روشنایی بسیار کمی هستند، که تشخیص آنها را توسط رادیولوژیست‌ها و پزشکان دشوار می‌کند. با توجه به اینکه تشخیص دقیق و بهموقع توده سرطانی و انواع مختلف آن از اهمیت ویژه‌ای در سلامتی افراد جامعه برخوردار است، لذا باید به‌نحوی دشواری تشخیص توده‌های سرطان سینه که به‌طور عمده با خطاهای انسانی در دقت تشخیص نیز همراه است توسط محققین مورد توجه قرار گیرد. به همین دلیل ارائه و گسترش یک روش خودکار به کمک روشهای پردازش تصویر و الگوریتم‌های تشخیص و همچنین بهینه‌سازی و ارتقای سیستم‌های تشخیصی موجود برای کاهش خطای انسانی و کمک به تشخیص زود هنگام و بهبود روند درمان این بیماری حائز اهمیت است. در این مقاله کوشیده ایم با معرفی روشی خودکار و جدید، در تشخیص و استخراج توده‌های سرطان سینه با دقت بالا دست پیدا کنیم. برای این منظور، پس از پیش‌پردازش و تشخیص مرزهای توده‌های سرطانی از تصاویر ماموگرافی، تعدادی ویژگی از توده‌های تشخیص داده شده استخراج شده و در نهایت ماتریس ویژگی‌ها به ورودی طبقه‌بندی‌کننده‌ی SVM اعمال شده است. در این مقاله با به‌کارگیری این روش، تشخیص و طبقه‌بندی توده‌های سرطان سینه با سرعت و دقت بالا انجام شده است.

مروری بر پژوهش‌های پیشین

در سال‌های گذشته تحقیقات گسترده‌ای برای کاهش خطا در شناسایی و تشخیص توده‌های سرطان پستان و افزایش سرعت و دقت برای کمک به رادیولوژیست‌ها انجام گرفته است. بطور کلی، تحقیقات انجام شده در این حوزه؛ شامل بخش‌های انتخاب پایگاه داده‌ی مناسب که شامل تصاویر دیجیتال ماموگرافی (تصاویر بافت

1. Lobular

نهایی بدست آمده حاصل از این پژوهش ۹۱.۳۷٪ گزارش شده است [۱۲].

بوشرا موگال^۳ و همکاران در سال ۲۰۱۷ یک سیستم تشخیص تومور را برای طبقه‌بندی صحیح توده‌های سینه به دسته‌های طبیعی، غیرطبیعی، خوش‌خیم و یا بدخیم ارائه کردند [۱۳]. روش پیشنهادی دو مجموعه داده DDSM و MIAS را مورد بررسی قرار داده است. یک الگوریتم جدید برای طبقه‌بندی براساس ترکیب تبدیل top-hat و ماتریس هم‌رخداد سطح خاکستری با شبکه عصبی انتشار برگشتی مورد استفاده قرار گرفته است. دقت نهایی به‌دست آمده به طور متوسط ۹۷/۶۲٪ گزارش شده است [۱۳].

ویلیام تورس^۴ و همکاران در سال ۲۰۱۸ روشی برای کاهش تشخیص مثبت کاذب در طبقه‌بندی تصاویر ماموگرافی ارائه کردند [۱۴]. در این پژوهش برای ارزیابی و تایید اعتبار، پایگاه داده‌ی DDSM با استفاده از سیستم CAD مورد بررسی قرار گرفت. از توصیف‌گرهای بافتی مبتنی بر شاخص‌های تفاوت کارکردی برای استخراج ویژگی‌ها و پس از آن طبقه‌بندی مناطق مورد نظر بر اساس وجود یا عدم وجود توده استفاده شده است. صحت نهایی بدست آمده در این کار ۹۲/۲۹٪ گزارش شده است [۱۴].

باسما موهام^۵ و همکاران در سال ۲۰۱۸ روشی خودکار جهت طبقه‌بندی تصاویر ماموگرام ارائه کردند [۱۵]. این مطالعه پایگاه اطلاعاتی دیجیتال غربالگری ماموگرافی (DDSM) را مورد بررسی قرار داده است. الگوریتم پیشنهادی شامل سه مرحله اصلی است. اول، سه نوع مختلف از ویژگی‌ها از توده جدا شده است. سپس مرتبط‌ترین ویژگی‌ها با استفاده از الگوریتم آزمایش انتخاب شده و در نهایت، طبقه‌بندی برای تمایز میان توده‌های خوش‌خیم و بدخیم با استفاده از سه طبقه‌بندی کننده، شبکه عصبی مصنوعی، ماشین‌های بردار پشتیبان، k-نزدیک‌ترین همسایه انجام شده است. شبکه عصبی مصنوعی بهترین نتایج را با دقت ۹۸/۹٪ بدست آورده است [۱۵].

ماریا پرز^۶ و همکاران در سال ۲۰۱۷ روشی برای طبقه‌بندی تصاویر ماموگرام ارائه کردند [۱۶]. در این پژوهش از پایگاه داده دیجیتال غربالگری ماموگرافی (DDSM) استفاده شده است. از شبکه عصبی مصنوعی (ANN) با استفاده از یک روش برون سپاری (Back-Propagation) و همچنین ویژگی‌های بافتی برای طبقه‌بندی تصاویر به سه دسته نرمال، خوش‌خیم و سرطان استفاده شده است. صحت بدست آمده حاصل از این روش به طور متوسط ۸۴.۷۲٪ گزارش شده است [۱۶].

با توجه به روش‌های گزارش شده در پژوهش‌های پیشین در

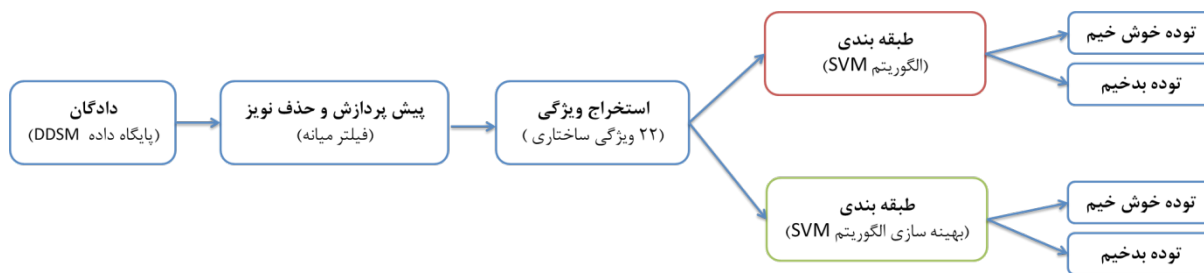
حامد جباری و همکاران در سال ۱۳۹۵ روشی ترکیبی جهت قطع‌بندی و تشخیص تومورهای تصاویر ماموگرافی بافت پستان ارائه کردند [۹]. در این مطالعه از داده‌های پایگاه داده DDSM شامل ۱۵۰ تصویر تومور خوش‌خیم و ۱۵۰ تصویر تومور بدخیم استفاده شده است. پس از پیش‌پردازش و حذف نواحی اضافی، قطع‌بندی بر مبنای هیستوگرام تصویر و ترکیب تبدیل موجک و الگوریتم ژنتیک و مورفولوژی انجام شده است. به علاوه الگوریتم‌های بهینه‌سازی کلونی مورچگان و ازدحام ذرات نیز در آن مطالعه مورد استفاده و مقایسه قرار گرفته‌اند. نتایج نشان می‌دهند که قطع‌بندی با استفاده از الگوریتم ژنتیک نسبت به الگوریتم کلونی مورچگان و ازدحام ذرات دارای صحت بالاتری است. در نهایت مدل پیشنهادی به طبقه‌بندی داده‌ها به دو دسته‌ی خوش‌خیم و بدخیم پرداخته است و صحت نهایی بدست آمده ۹۱/۴٪ گزارش شده است [۹].

پگاه توکل خواه و همکاران در سال ۱۳۸۶ روشی برای دسته‌بندی و تقطیع تصاویر سرطان سینه ارائه کردند [۱۰]. در این پژوهش قطع‌بندی تصویر با استفاده از الگوریتم رشد ناحیه انجام شده است. در آن مطالعه، ویژگی‌های آماری و وابستگی مکانی از تصاویر استخراج شده است، و در ادامه از روش آستانه‌سازی فازی تطبیقی مبتنی بر انتروپی به منظور استخراج دقیق‌تر توده‌های سرطانی استفاده شده است. دقت نهایی بدست آورده در این پژوهش با استفاده از ۳۸ ویژگی، جهت طبقه‌بندی داده‌ها به دو گروه سالم و بیمار به ترتیب برابر ۸۶/۶۶٪ و ۸۲/۵٪ است [۱۰].

رباب شیخ‌پور و همکاران در سال ۱۳۹۵ روشی برای تشخیص سرطان پستان با استفاده از برآورد ناپارامتری چگالی احتمال ارائه کردند [۱۱]. در این پژوهش از پایگاه داده WBCD شامل ۶۹۹ نمونه‌ی خوش‌خیم و بدخیم با ۹ ویژگی و WDCB شامل ۵۶۹ نمونه‌ی خوش‌خیم و بدخیم با ۳۰ ویژگی استفاده شده است و سپس با استفاده از روش‌های تخمین چگالی مبتنی بر هسته، مدلی برای طبقه‌بندی مجموعه داده‌ها ارائه شده است. نتایج پژوهش نشان می‌دهند که روش برآورد چگالی هسته‌ای گوسین مبتنی بر فاصله اقلیدس با دقت ۹۷/۹۳٪ بالاترین دقت را در میان سایر روش‌ها داشته است [۱۱]. پزشکی^۲ و همکاران در سال ۲۰۱۹ روشی نوین برای طبقه‌بندی سرطان سینه ارائه کردند [۱۲]. در این پژوهش در ابتدا، مجموع تفاوت‌های بین پیکسل مرکزی و پیکسل‌های همسایه در جهت‌های مختلف محاسبه شده است. کمترین تفاوت بین دو جهت عمود متعام شباهت پیکسل‌ها در یک جهت را به عنوان قسمت سوزنی شکل توده مشخص می‌کند. این بخش‌ها به تصویر قطع‌بندی شده اضافه شده است تا شکل تومور را بزرگتر کنند. در نهایت ویژگی‌های تومور از تصویر قطع‌بندی شده‌ی نهایی استخراج شده و تومورها به دو دسته‌ی خوش‌خیم و بدخیم طبقه‌بندی شده است. دقت

3. Bushra Mughal
4. William Torres
5. Basma A. Mohame
6. Maria Perez

2. Pezeshki



تصویر ۱. دیاگرام بلوکی روش پیشنهادی

استفاده از استخراج ویژگیهای ساختاری از توده‌های سرطانی تشخیص داده شده، به منظور طبقه‌بندی خودکار توده‌های سرطان سینه ارائه شده است. این پژوهش بر اساس بررسی اطلاعات پایگاه داده‌ی تصاویر ماموگرافی دیجیتال DDSM انجام شده است. در این بخش ابتدا به معرفی پایگاه داده‌ی به کار گرفته شده در این مقاله پرداخته شده است. سپس به معرفی انواع روش‌های پیش‌پردازش تصاویر ماموگرافی بافت سینه، استخراج توده‌ها و جداسازی آنها از بافت سالم سینه، حذف نواحی اضافه تشخیص داده شده و همچنین بهبود مرزهای توده‌های آشکارسازی شده پرداخته شده است. در نهایت روش‌های استخراج ویژگی از توده استخراج شده است و طبقه‌بندی داده‌ها به دو دسته خوشخیم و بدخیم با استفاده از شبکه عصبی ماشین بردار پشتیبان (Support Vector Machine) و بهینه‌سازی صورت گرفته بر روی آن انجام شده است. تصویر شماره ۱ دیاگرام بلوکی روش پیشنهاد شده در این مقاله را نشان میدهد.

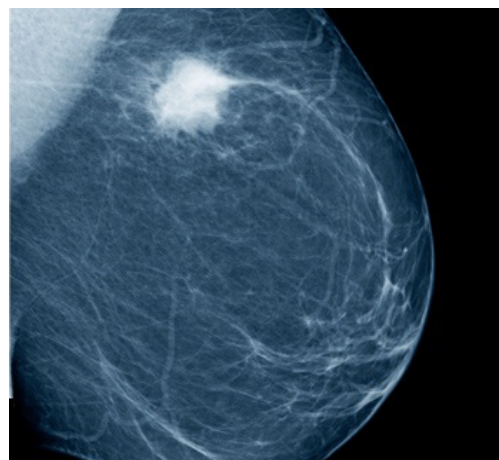
پایگاه داده

تصاویر مورد استفاده در این پژوهش از پایگاه داده (DDSM) Digital Database for Screening Mammography جمع‌آوری شده است [۱۷]. این پایگاه داده شامل ۳۲۷ تصویر

خصوصاً انجام عملیات شناسایی، تشخیص و طبقه‌بندی توده‌های سرطان سینه، لزوم انجام یک پژوهش که بتواند روشی خودکار با دقت و سرعت بالا جهت انجام عملیات ذکر شده را ارائه دهد وجود دارد. لذا در این مقاله، روشی ارائه شده است که در آن پس از تشخیص و آشکارسازی توده‌های سرطان سینه از تصاویر ماموگرافی و جداسازی آنها از بافت سالم پستان با بکارگیری روش‌های پردازش تصاویر دیجیتال، و استخراج ویژگیهای ساختاری مختلف از توده‌ها، به بهینه‌سازی پارامترهای طبقه‌بندی کننده‌ی ماشین بردار پشتیبان (SVM) جهت تشخیص توده‌های سرطان سینه به دو کلاس خوشخیم و بدخیم پرداخته شده است. در ادامه پس از معرفی پایگاه داده و پیش‌پردازش تصاویر، به بررسی الگوریتم قطعه‌بندی تصویر پرداخته شده است. در مرحله‌ی بعد ویژگیهای هندسی و ساختاری مورد استفاده در این پژوهش و طبقه‌بندی داده‌ها با استفاده از ماشین بردار پشتیبان و بهینه‌سازی آن شرح داده شده است. در نهایت، نتایج به دست آمده مورد تجزیه و تحلیل قرار گرفته است.

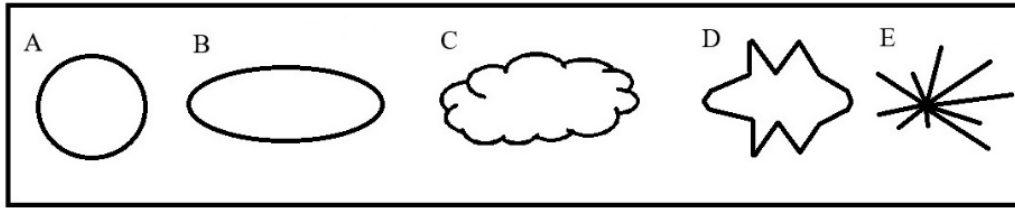
مواد و روش‌ها

پژوهش حاضر به ارائه‌ی روشی کاربردی و جدید جهت آشکارسازی و تشخیص توده‌های سرطانی تصاویر ماموگرافی با



تصویر ۲. تصویر ماموگرافی نمونه دارای توده سرطانی اخذ شده از پایگاه داده DDSM





تصویر ۳. شکل انواع توده‌های سرطانی سینه



در تحلیل تصاویر پزشکی و به‌خصوص تصاویر ماموگرافی، بسیار مهم و ضروری است، زیرا تصاویر پزشکی کیفیت پایینی دارند و تفکیک اجزای مختلف تصویر بسیار دشوار است و خطا در تشخیص بسیار زیاد است. بنابراین در این مرحله توده‌ها و لبه‌های تصویر به‌طور کامل مجزا شده و به صورت یک تصویر تفکیک شده نمایش داده می‌شوند.

روش قطعه‌بندی آستانه‌گذاری یکی از ساده‌ترین و مؤثرترین روش‌های قطعه‌بندی تصویر است که با استفاده از یک مقدار آستانه، تصویر را به دو بخش قطعه‌بندی می‌کند. در این روش ابتدا نقطه‌ی میانی هیستوگرام تصویر به عنوان مقدار آستانه T_0 در نظر گرفته شده، سپس تبدیل شدت روشنایی‌های تصویر بر اساس این مقدار قطعه‌بندی تصویر مطابق فرمول ۱ انجام میشود.

$$\begin{cases} \text{If } f(i, j) \geq T_0 & \text{then } f(i, j) = 255 \\ \text{Else} & f(i, j) = 0 \end{cases} \quad (1)$$

که در این رابطه نشان‌دهنده شدت روشنایی در سطح ام و ستون ام است. این الگوریتم برای شدت روشنایی‌های و میانگین‌گیری انجام داده و بعد از آن میانگینی که از دیگری بیشتر است را یافته و در نهایت مقادیر شدت روشنایی به دو دسته A و B تقسیم میشوند، که در آن A دارای حداقل مقادیر شدت روشنایی و B دارای حداکثر مقادیر شدت روشنایی در تصویر است [۲۰].

بهبود مرز توده‌های سرطانی

الگوریتم ریخت‌شناسی ریاضی (Mathematical Morphology) یا مورفولوژی روشی جدید به‌منظور پردازش تصویر است. ایده اصلی این روش بر مبنای آنالیز اطلاعات هندسی توسط کاوش یک تصویر با یک الگوی هندسی کوچک میباشد. از جمله مهم‌ترین الگوریتم‌های مورفولوژی ساییدگی (Erosion) و انبساط (Dilation) میباشد. با استفاده از عملگر انبساط اجزا در تصویر شروع به رشد و گسترش میکنند، در مقابل با اعمال عملگر سایش در تصویر، اجزا در تصویر کوتاه و نازک میشوند. در این پژوهش به علت اینکه توده‌ی استخراج شده توسط الگوریتم آستانه‌گذاری به سمت داخل تمرکز یافته است لذا با به‌کارگیری عملگر انبساط شدت روشنایی پیکسل‌های پیرامون لبه به سمت

سرطان خوشخیم و بدخیم بافت سینه است. اندازه اصلی تصاویر 1024×1024 پیکسل و به فرمت DICOM هستند. لازم به ذکر است که مرزبندی اولیه‌ی توده‌های سرطانی در تصاویر این پایگاه داده در ابتدا توسط رادیولوژیست‌ها انجام شده است و گزارشی در مورد خوش‌خیم یا بدخیم بودن توده‌ها نیز در این پایگاه داده برای هر تصویر در دسترس می‌باشد. تصویر شماره ۲ نمونه‌ای از تصاویر پایگاه داده DDSM را نشان میدهد [۱۷].

لازم به ذکر است داده‌های موجود در این پایگاه داده دارای برچسب‌های از پیش تعیین شده‌ای هستند. این برچسب‌ها شامل Lobulated Margins، Oval، Amorphous distribution، Irregular Architectural، Irregular Margins، Pleomorphic distribution و Round margins هستند. در میان این تصاویر توده‌های خوش‌خیم شکلی شبیه به دایره و یا بیضی داشته در حالی که توده‌های بدخیم شکل بی‌نظم و حاوی زائده‌های کوچک و سوزنی دارند. تصویر شماره ۳ انواع توده‌های سرطانی سینه را نشان میدهد [۱۸].

پیش‌پردازش

به علت نوع تصویربرداری ماموگرافی، این تصاویر در دسته‌ی تصاویر نویزدار قرار می‌گیرند. بنابراین قبل از انجام عملیات پردازش تصویر بر روی این تصاویر باید نویزهای موجود در تصویر از بین بروند. فیلتر میانه به‌طور گسترده‌ای در پردازش تصویر کاربرد دارد. این فیلتر یکی از فیلترهای غیرخطی است که برای حذف نویز ضربهای در تصاویر استفاده می‌شود [۱۹]. این فیلتر با استفاده از یک پنجره 3×3 یا 5×5 ، روشنایی پیکسل وسط یا هدف را بر اساس میانه روشنایی پیکسل‌های همسایه تغییر می‌دهد. در نتیجه یک ترکیب از روشنایی در یک محدوده بوجود می‌آید. این فیلتر پیکسل‌هایی که شدت روشنایی بسیار متفاوتی با همسایه‌ها داشته باشند را در نظر نمی‌گیرد. در این پژوهش فیلتر میانه با پنجره‌ای با ابعاد 3×3 روی تصاویر ماموگرافی دیجیتال برای حذف نویز با استفاده از نرم‌افزار MATLAB اعمال شده است.

استخراج توده‌های سرطانی

پس از اعمال حذف نویز و هموارسازی تصاویر، افزایش کیفیت و تعیین ساختار نهایی توده‌ی سرطانی نیاز است. این مرحله

جدول ۱. فهرست ویژگی‌های مورد استفاده

ردیف	ویژگی	توصیف
۱	مساحت	شامل تعداد واقعی پیکسل‌های موجود در ناحیه‌ی استخراج شده است که در واقع مساحت توده سرطانی را مشخص می‌کند. این ویژگی به صورت یک عدد اسکالر است.
۲	Convex Area	یک عدد اسکالر است که بیانگر تعداد پیکسل‌های با درایه ۱ موجود در ویژگی Convex Image است.
۳	Eccentricity	یک عدد اسکالر است که بیانگر خروج از مرکز بیضی مشخص شده‌ای است که به عنوان ناحیه‌ی ثانویه مربوط به ناحیه‌ی توده‌ی مورد نظر است [۱۸].
۴	Equiv Diameter	یک مقدار عددی است که بیانگر قطر یک دایره با همان منطقه بعنوان منطقه مشخص شده است. این ویژگی بصورت جزر عبارت محاسبه می‌شود [۱۸].
۵	Euler Number	یک مقدار عددی است که تعداد اشیاء موجود در منطقه را منهای تعداد سوراخ‌ها (یا حفره‌ها) در آن اشیاء را مشخص می‌کند [۱۸].
۶	شدت	یک ویژگی عددی است که نسبت پیکسل‌ها در منطقه به پیکسل‌ها در کل bounding box را مشخص می‌کند. این محاسبه بصورت تقسیم Area بر ناحیه‌ی bounding box بدست می‌آید.
۷	Filled Area	یک مقدار اسکالر است که مشخص کننده‌ی تعداد پیکسل موجود در Filled Image است. ویژگی Filled Image یک تصویر باینری (منطقی) هم اندازه با Bounding Box مربوط به ناحیه مورد نظر است. در واقع این ویژگی مربوط به پیکسل‌های روشن (on) مربوط به ناحیه مورد نظر با تمامی سوراخ‌های پر شده در آن تعریف می‌شود.
۸	طول کوچکترین محور	بصورت یک مقدار اسکالر طول (در پیکسل‌ها) مربوط به محور کوچک بیضی که بعنوان ناحیه‌ی مرکزی ثانویه بهنجار شده مربوط به ناحیه‌ی مورد نظر است تعریف می‌شود [۱۸].
۹	طول بزرگترین محور	بصورت یک مقدار اسکالر طول (در پیکسل‌ها) مربوط به محور بزرگ بیضی که بعنوان ناحیه‌ی مرکزی ثانویه بهنجار شده مربوط به ناحیه‌ی مورد نظر است تعریف می‌شود [۱۸].
۱۰	جهت	یک مقدار اسکالر زاویه است (در درجه‌های مختلف از ۹۰- درجه تا ۹۰+ درجه) که بین محور X و محور بزرگ بیضی است که همان همان دوم ناحیه است تعریف می‌شود.
۱۱	محیط	یک مقدار اسکالر است که بیانگر مسافت دور مرز ناحیه مورد نظر است. جبهه ابزار regionprops در نرم‌افزار MATLAB توسط محاسبه‌ی فاصله بین یک جفت پیکسل مجاور در اطراف مرز ناحیه مورد نظر به محاسبه‌ی ویژگی محیط می‌پردازد.
۱۲	Solidity	یک مقدار اسکالر مشخص است که نسبت پیکسل‌های در بخش‌های محدب بدنه که در منطقه نیز می‌باشند را مشخص می‌کند.
۱۳	نسبت نازکی توده	ویژگی نسبت نازکی توده توده‌های حلقوی را از دیگر انواع توده‌ها متمایز می‌کند. البته این نسبت را فقط برای توده‌های حلقوی می‌توان بکار برد [۱۸]. $TR = \left(\frac{\text{مساحت توده}}{\text{محیط توده}} \right)^2$
۱۴	کشیدگی	ویژگی کشیدگی (En) معرف تفاوت میان توده‌های بیضی شکل منظم با توده‌های بیضوی نامنظم است [۱۸]. $En = \left(\frac{\text{مساحت توده}}{2 \times (\text{شعاع ماکزیمم})^2} \right)$
۱۵	مدور بودن ۱	ویژگی ۱) Circularity بصورت اندازه ثابت چرخشی و موقعیتی یک توده تعریف می‌شود که نشان می‌دهد تا چه اندازه یک توده شبیه به دایره است [۱۸]. $C1 = \left(\frac{\text{مساحت توده}}{\pi \times (\text{شعاع ماکزیمم})^2} \right)$
۱۶	مدور بودن ۲	ویژگی ۲) Circularity مشخص می‌کند که به چه میزان شباهت میان توده‌ی استخراج شده و توده بیضوی (به عنوان مرجع) وجود دارد [۱۸]. $C2 = \left(\frac{\text{شعاع مینیمم}}{\text{شعاع ماکزیمم}} \right)$

ردیف	ویژگی	توصیف
۱۷	فشرده‌گی	ویژگی فشرده‌گی (Compactness) بیانگر درجه انحراف یک توده حلقوی کامل است که این درجه انحراف مستقل از تغییر خطی توده است [۱۸]. $CN = \left(\frac{\sqrt{2} \times \text{مساحت توده}}{\text{محیط توده}} \right)$
۱۸	پراکندگی و بی‌نظمی	ویژگی پراکندگی و بی‌نظمی (DP)، خصوصیات یک شکل نامنظم را مشخص می‌کند [۱۸]. $DP = \left(\frac{\text{شعاع ماکزیمم}}{\text{مساحت توده}} \right)$
۱۹	ویژگی شاخص شکل	ویژگی شاخص شکل (SI) بصورت یک شاخص مربوط به اتحنای سطح توده است [۱۸]. $SI = \left(\frac{\text{محیط توده}}{\sqrt{x} \times \text{شعاع ماکزیمم}} \right)$



طبقه‌بندی

پس از استخراج ویژگی از تصاویر و آماده‌سازی بردار ویژگی با استفاده از الگوریتم‌های طبقه‌بندی معرفی شده در این مقاله، به تفکیک داده‌ها به دو دسته‌ی خوشخیم و بدخیم پرداخته شده است. بدین منظور از الگوریتم ماشین بردار پشتیبان (SVM) استفاده شده است. الگوریتم ماشین بردار پشتیبانی یکی از روش‌های یادگیری با نظارت است که از آن برای طبقه‌بندی و تخمین رگرسیون استفاده می‌شود. این روش از جمله روش‌های نسبتاً جدیدی است که در سال‌های اخیر کارایی مناسبی نسبت به روش‌های قدیمی‌تر برای طبقه‌بندی نشان داده است [۲۱]. این الگوریتم جهت جداسازی داده‌ها از یک ابرصفحه استفاده می‌کند و این ابرصفحه به گونه‌ای انتخاب می‌شود که داده‌های هر دسته بیشترین فاصله را با ابرصفحه داشته باشند. در این روش آموزش نسبتاً ساده است و برخلاف شبکه‌های عصبی در ماکزیمم‌های محلی گیر نمی‌افتد. همچنین این روش برای داده‌های با ابعاد بالا تقریباً جواب مناسبی می‌دهد، اما برای اینکه بهترین جواب از این شبکه دریافت شود به یک تابع کرنل مطلوب و انتخاب پارامتر مناسب وابسته است. بنابراین در این مقاله از الگوریتم بهینه‌سازی ملخ Grasshopper Optimization Algorithm (که در ادامه

مرکز گسترش یافته‌اند. برای انجام این عملیات، دایره‌های به شعاع صعودی ساخته شده که با پیمایش تصویر مقادیر شدت روشنایی برابر خود را یافته به حالتی که در هر تکرار شعاع نیز افزایش می‌یابد و به محض یافتن پیکسل هدف، تمامی پیکسل‌های مابین دارای شدت روشنایی جدیدی برابر با شدت روشنایی پیکسل‌های لبه خواهند داشت.

استخراج ویژگی

استخراج ویژگی یکی از مهمترین قسمت‌های سیستم‌های پردازش و تشخیص است. پس از استخراج ناحیه‌ی توده‌های سرطانی سینه که در بخش قبل توضیح داده شد، در این مرحله اقدام به ارائه ویژگی‌های ساختاری از ناحیه‌ی توده‌ی استخراج شده توسط نرم‌افزار MATLAB می‌پردازیم. به منظور طبقه‌بندی داده‌ها به دو دسته خوش‌خیم و بدخیم، ابتدا باید ویژگی‌های مناسب از ناحیه استخراج شده توسط روش قطع‌بندی استخراج گردد در این پژوهش ۱۹ ویژگی ساختاری مورد استفاده قرار گرفته است. در جدول ۱ ویژگی‌های ساختاری به کار گرفته شده در این پژوهش بیان شده است.

جدول ۲. برجسب‌های رفرنس و برجسب‌های الگوریتم پیشنهادی برای محاسبه صحت، حساسیت و اختصاصی بودن

برجسب مشخص شده توسط الگوریتم ارائه شده			
خوش خیم		سرطان	
خوش خیم	خوش خیم - خوش خیم (TP)	خوش خیم - سرطان (FN)	
سرطان	سرطان - خوش خیم (FP)	سرطان - سرطان (TN)	برجسب اصلی داده‌ها





آسیب‌زدن به محصولات تولیدی و کشاورزی به عنوان آفت در نظر گرفته می‌شوند. اگر چه ملخ‌ها معمولاً به‌طور جداگانه در طبیعت دیده می‌شوند، اما آنها در یکی از بزرگترین ازدحامات نسبت به تمام موجودات قرار دارند [۲۲].

به‌طور کلی، الگوریتم‌های الهام گرفته از طبیعت به گونه‌های منطقی فرایند جستجو را به دو گرایش تقسیم می‌کنند: اکتشاف و بهره‌برداری؛ در اکتشاف، عاملان جستجو تشویق می‌شوند که به‌طور ناگهانی حرکت کنند، در حالی که آنها در طول بهره‌برداری تمایل به حرکت محلی دارند. این دو عملکرد، و نیز جستجوی هدف، به‌طور طبیعی توسط ملخ‌ها در طبیعت انجام می‌شود.

مدل ریاضی به‌کار گرفته شده برای شبیه‌سازی رفتار ازدحام ملخ‌ها در فرمول ۵ بیان شده است [۲۲].

۵.

$$X_i = S_i + G_i + A_i$$

در این رابطه، X_i معرف موقعیت ملخ i ام را تعریف می‌کند، S_i فاکتور تعامل اجتماعی است، G_i نیروی جاذبه در ملخ i ام بوده و A_i معرف پیشگویی باد می‌باشد. توجه داشته باشید که برای ارائه‌ی رفتار تصادفی، معادله می‌تواند بصورت $X_i = r_1 S_i + r_2 G_i + r_3 A_i$ باشد، که در آن r_1 ، r_2 و r_3 اعداد تصادفی در بازه $[0, 1]$ هستند. همچنین مطابق فرمول ۶ خواهیم داشت [۲۲].

۶.

$$S_i = \sum_{\substack{j=1 \\ j \neq i}}^N s(d_{ij}) (d_{ij}) \hat{d}_{ij}$$

در این رابطه، d_{ij} فاصله بین ملخ i ام و j ام است و به صورت $d_{ij} = |X_j - X_i|$ محاسبه می‌شود، s یک تابع برای تعریف قدرت نیروی اجتماعی است که در فرمول ۷ نشان داده شده است [۲۲] و $d_{ij} = X_j - X_i / d_{ij}$ معرف یک بردار واحد از ملخ i ام به ملخ j ام است. تابع s که نیروهای اجتماعی را تعریف می‌کند نیز به‌صورت زیر محاسبه می‌شود:

۷.

$$s(r) = fe^{-r} / e^{-r}$$

در این رابطه f نشان‌دهنده‌ی شدت جذب و $1/e^{-r}$ مقیاس طولی جذاب است. در این پژوهش بهینه‌سازی پارامترهای توابع کرنل SVM با به‌کارگیری الگوریتم بهینه‌سازی ملخ انجام شده است. تعداد ملخ‌های استفاده شده در این مقاله ۳ عدد می‌باشد. همچنین عملیات بهینه‌سازی بر روی سه تابع کرنل linear و RBF و poly-nomial از طبقه‌بندی کننده‌ی SVM انجام شده است.

نتایج

در این مطالعه تصاویر پایگاه داده DDSM برای طبقه‌بندی به

شرح داده می‌شود) برای انتخاب بهینه‌ترین پارامترها در الگوریتم SVM، جهت اخذ بهترین نتایج طبقه‌بندی استفاده شده است.

برای ارزیابی اعتباربخشی الگوریتم پیشنهادی از روش اعتبارسنجی ۴ تایی (Fold-۴ Cross Validation) استفاده شده است. به این صورت که کل داده‌های موجود را به چهار بخش تقسیم نموده و در هر مرتبه اجرای برنامه، سه بخش از داده‌ها را به‌عنوان داده‌های آموزشی و یک بخش را به‌عنوان داده‌های اعتبارسنجی به طبقه‌بندی کننده‌های معرفی شده اعمال کردیم. علاوه براین، در این مقاله تعداد ۸۵٪ داده‌ها را به‌عنوان داده‌های آموزشی و ۱۵٪ داده‌ها را به‌عنوان داده‌های تست در نظر گرفته شدند. لازم به ذکر است که مجموعاً ۱۰۰ مرتبه اجرای برنامه انجام شد و در نهایت میانگین و انحراف معیار نتایج بدست آمده گزارش شدند. در جدول ۲ و روابط پس از آن نحوه تعریف معیارهای حساسیت (Sensitivity)، اختصاصی بودن (Speci-ficity) و صحت (Accuracy) در طبقه‌بندی تصاویر ماموگرافی سینه جهت تشخیص نوع توده‌ی سرطانی بیان شده است.

فرمول‌های ۲، ۳ و ۴ بیانگر روابط ریاضی برای محاسبه صحت، حساسیت و اختصاصی بودن می‌باشند که به ترتیب در زیر بیان شده‌اند:

۲.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

۳.

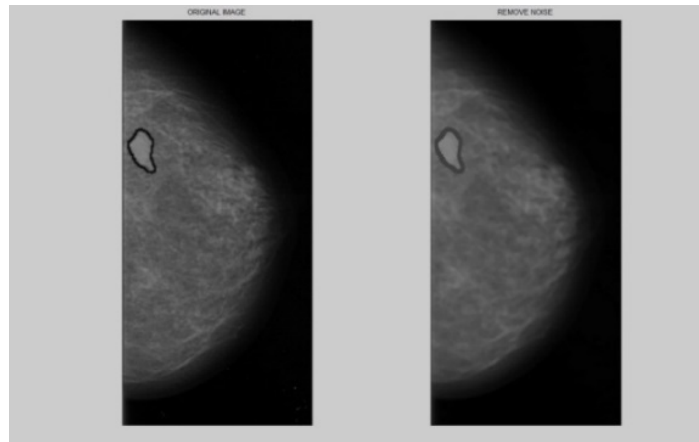
$$Sensitivity = \frac{TP}{TP+FN}$$

۴.

$$Specificity = \frac{TN}{TN+FP}$$

الگوریتم بهینه‌سازی ملخ (GOA)

فرآیند یافتن بهترین مقادیر برای متغیرهای یک مسئله خاص برای به حداقل رساندن یا به حداکثر رساندن یک تابع هدف، بهینه‌سازی نامیده می‌شود. مسائل بهینه‌سازی در زمینه‌های مختلف مطالعات وجود دارد و تاکنون الگوریتم‌های بهینه‌سازی بسیاری نظیر Tabu، simulated annealing، climbing، Genetic Search (TS)، Iterated Local Search (ILS)، Particle Swarm Optimisation (PSO)، Ant Colony Optimisation (ACO)، Differential Evolution (DE)، Dolphin Echolocation (DEL)، Firefly Algorithm (FA)، Bat Algorithm (BA)، Grey Wolf Optimizer (GWO)، جستجو فاخته (CS)، شکار گرگ خاکستری (GWO)، State of Matter Search (SMS) و گرد افشانی گل (FPA) معرفی شده‌اند، اما تاکنون تحقیقات اندکی ازدحام ملخ‌ها را شبیه‌سازی کرده‌اند. ملخ‌ها حشراتی هستند که به علت



همچنین اعتبارسنجی طبقه‌بندی انجام شده به صورت fold-۴ در صورت گرفته و تعداد اجرای برنامه به ازای هر تابع کرنل، ۱۰۰ مرتبه بوده است. همچنین در نتایج بیان شده در جدول ۴ نیز دیتای آموزشی ۸۵٪، طبقه‌بندی بصورت fold-۴ انجام شده، تعداد جمعیت ملخ‌ها در الگوریتم بهینه‌سازی برابر ۳ عدد بوده، تعداد اجرای تکرارهای الگوریتم ملخ برای بهینه‌سازی برای هر تابع کرنل ۱۰۰ تکرار بوده و تعداد اجرای برنامه به ازای هر اجرای هر تابع کرنل طبقه‌بندی کننده ۱۰۰ مرتبه بوده است.

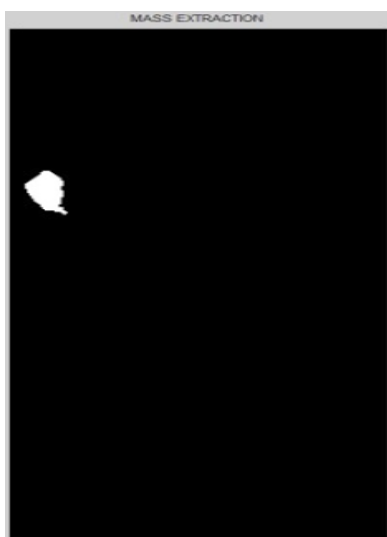
بحث و نتیجه‌گیری

در این مقاله از ویژگی‌های ساختاری بر پایه شکل هندسی توده‌های سرطانی استخراج شده از تصاویر ماموگرافی دیجیتال توسط روش‌های پردازش تصاویر جهت طبقه‌بندی داده‌ها به دو

تصویر ۴. نتیجه‌ی استفاده از بکارگیری و اعمال فیلتر میانه

دو دسته خوشخیم و بدخیم مورد استفاده قرار گرفته است. در این بخش نتایج حاصل از الگوریتم پیشنهاد شده گزارش شده است. تصویر ۴ نتایج حاصل از اعمال فیلتر میانه و تصویر ۵ و ۶ نتایج اعمال قطعه‌بندی توسط روش آستانه‌گذاری و بهبود آن توسط عملگرهای مورفولوژی را نشان می‌دهند.

جدول ۳ درصد مقادیر صحت، حساسیت و اختصاصی بودن (به همراه واریانس آنها) حاصل از به‌کارگیری طبقه‌بندی داده‌ها (به دو دسته‌ی خوش‌خیم و بدخیم) با استفاده از روش طبقه‌بندی SVM را با استفاده از سه تابع کرنل Linear و RBF و Polynomial نشان می‌دهد. همچنین جدول ۴ این نتایج را بعد از اعمال الگوریتم بهینه‌سازی ملخ نشان می‌دهد. در نتایج بیان شده در جدول ۳، دیتای آموزشی بکار گرفته شده ۸۵٪ بوده و ۱۵٪ داده‌ها به‌عنوان داده‌های تست در نظر گرفته شده‌اند،



تصویر ۶. نتیجه‌ی استفاده از روش بهبود قطعه بندی توسط عملگرهای مورفولوژی



تصویر ۵. نتیجه‌ی استفاده از روش قطعه بندی به روش آستانه‌گذاری

جدول ۳. نتایج حاصل از به کارگیری روش طبقه‌بندی کننده‌ی MVS

روش طبقه‌بندی کننده‌ی	درصد		
	SVM (Polynomial)	SVM (Linear)	SVM (RBF)
صحت	۸۳/۳۲±۱/۱۴e-۱۳ %	۹۱/۶۷±۵/۷۱e-۱۴ %	۹۷±۴/۳۶ %
حساسیت	۱۰۰±۰%	۸۰/۰±۰ %	۱۰۰±۰%
اختصاصیت	۸۱/۸۲±۱/۱۴e-۱۳ %	۱۰۰±۰%	۹۶±۵/۸۱ %

جدول ۴. نتایج حاصل از به کارگیری روش طبقه‌بندی کننده‌ی MVS بهینه‌سازی شده

روش طبقه‌بندی کننده‌ی	درصد		
	SVM (Polynomial)	SVM (Linear)	SVM (RBF)
صحت	۹۱/۶۷±۰/۱۲ %	۱۰۰±۰ %	۹۵/۸۳±۰/۰۶ %
حساسیت	۱۰۰±۰ %	۱۰۰±۰ %	۱۰۰±۰ %
اختصاصیت	۹۰/۰±۰/۱۴ %	۱۰۰±۰ %	۹۴/۴۴±۰/۰۸ %



دقت الگوریتم SVM در فرایند طبقه‌بندی داده‌ها میشود. جدول ۵ نتایج مقایسه الگوریتم پیشنهادی با مطالعات گذشته را نشان میدهد. با توجه به جدول قابل مشاهده است که صحت روش پیشنهادی با استفاده از ویژگی‌های ساختاری مختلف و همچنین بهبود پارامترهای شبکه SVM با سه شاخص صحت، حساسیت و ویژگی (اختصاصیت) ۱۰۰٪ دارای کارایی بالاتری نسبت به سایر پژوهشها است.

با توجه به اینکه روش‌های متعددی برای پیش‌پردازش و بخش‌بندی تصاویر وجود دارد، لذا با مرور برخی الگوریتم‌های مشابه در خصوص آشکارسازی سایر توده‌ها و تومورهای دیگر نقاط بدن، می‌توان به اهمیت به کارگیری این روش در مقاله‌ی حاضر پی برد. به‌عنوان نمونه، در سال ۱۳۹۵، حمیدرضا غفاری و همکاران روشی برای تشخیص ندول‌های ریوی از روی تصاویر سی‌تی‌اسکن ریه با استفاده از طبقه‌بندی کننده‌ی SVM ارائه کردند [۲۴]. محققان در آن پژوهش از گام‌های تبدیل تصویر به مقیاس خاکستری، تبدیل تصویر خاکستری به باینری توسط یک حد آستانه، حذف نواحی نامرتب برای دستیابی به دیواره‌ی ریه، پر کردن حفره‌های تصویر برای دستیابی به پس‌زمینه‌ی تصویر، ارائه‌ی نهایی لبه‌های ریه‌ی استخراج شده، افزودن ماسک به تصویر اصلی و بخش‌بندی ریه و انتخاب نواحی مشکوک توسط به کارگیری الگوریتم رشد ناحیه‌ی آشکارسازی شده صورت پذیرفت [۲۴]. همچنین، در سال ۱۳۹۱، عباس بی‌نیاز و همکاران در مطالعه‌ی به مرور کامل روش‌های اتوماتیک و نیمه‌اتوماتیک بخش‌بندی تصاویر، و به‌طور خاص برای بررسی روش‌های بخش‌بندی تصاویر پزشکی تشدید مغناطیسی مغز انسان پرداختند [۲۵]. محققان در آن مطالعه، انواع روش‌های

دسته خوشخیم و بدخیم استفاده شده است. در این پژوهش پس از پیش‌پردازش تصاویر و حذف نویز آنها توسط فیلتر میانه، مناطق مشکوک به توده‌های سرطانی در بافت سینه با به کارگیری روش طبقه‌بندی آستانه‌گذاری استخراج شدند. در ادامه، توسط الگوریتم آستانه‌گذاری به سمت داخل توده، و با استفاده از عملگر انبساط، شدت روشنایی پیکسل‌های پیرامون لبه به سمت مرکز توده گسترش یافتند. سپس تعدادی ویژگی هندسی و ساختاری از توده‌ی آشکارسازی شده استخراج گردید. توسط این ویژگی‌ها، و با به کارگیری طبقه‌بندی کننده‌ی SVM، داده‌ها به دو دسته طبقه‌بندی شدند. تعداد ۱۹ ویژگی مساحت ناحیه، محیط، Convex Area، Eccentricity، Equiv Diameter، Euler Number، Extent، Filled Area، Major Axis Length، Minor Axis Length، جهت orientation، Solidity، نسبت نازکی توده، کشیدگی، Circular1، Circular2، فشردگی، شاخص شکل، و پراکندگی و بی‌نظمی که از هر تصویر استخراج شده‌اند را جهت تولید بردار ویژگی استفاده شده است. لازم به ذکر است که دلیل استفاده از این ۱۹ ویژگی این است که بهترین ویژگی‌ها از کل ویژگی‌های قابل استخراج از تصاویر توده‌های سرطانی بوده که با استخراج این ویژگی‌ها، نتایج طبقه‌بندی مناسب‌تری حاصل شده است. در مرحله بعد جهت بهینه‌سازی پارامترهای SVM از روش نوین بهینه‌سازی ملخ استفاده شده است. نتایج بهترین صحت حاصل از الگوریتم SVM (در حالت پایه و بدون استفاده از روش بهینه‌سازی و با کرنل خطی) برابر با ۹۱/۶۷٪ و صحت حاصل با به کارگیری روش بهینه‌سازی توسط الگوریتم ملخ برابر ۱۰۰ درصد به دست آمده است. نتایج نشان میدهند که اعمال و استفاده از روش بهینه‌سازی باعث افزایش

جدول ۵. مقایسه بهترین نتایج مطالعه‌ی حاضر با بهترین نتایج مطالعات گذشته

نام محقق	پایگاه داده	روش پژوهش	٪	
			صحت	حساسیت
رباب شیخ پور و همکاران [۱۸]	WBCD WDBC	روش طبقه‌بندی ناپارمتری براساس برآورد چگالی هسته‌ای گوسین مبتنی بر فاصله اقلیدسی؛ و روش k نزدیکترین همسایه	۹۸/۱۷ ۹۷/۹۳	۹۴/۷۴ ۹۳/۱۸
پگاه توکلخواه و همکاران [۱۰]	MIAS	استخراج ویژگی‌های آماری و ویژگی‌های وابستگی مکانی، و بکارگیری روش آستانه‌گذاری فازی مبتنی بر آنتروپی، و طبقه‌بندی با استفاده از شبکه عصبی	-	-
حامد جباری [۹]	DDSM	طبقه‌بندی مبتنی بر قطعه‌بندی تصویر؛ بکارگیری الگوریتم‌های بهینه‌سازی ژنتیک، کلونی مورچگان و ازدحام ذرات جهت بهینه‌سازی فرایند قطعه‌بندی تصویر	۹۱/۰۴	-
سعیده ناصری نوروزانی و همکاران [۸]	SEER پایگاه داده محلی	تشخیص و پیش‌بینی مرحله سرطان پستان براساس سیستم TNM؛ ایجاد مدل پیش‌بینی استفاده از الگوریتم فراالبتکاری کلونی مورچگان، جهت استخراج ویژگی‌های مؤثرتر؛ استفاده از طبقه‌بندی کننده‌های بیزین، Lazy، Functions، Meta، Rules و Trees؛	۹۹/۹۳ ۹۹/۹۱	-
ایمان عباس‌پور کازرونی و همکاران [۷]	MIAS	ارائه مدلی جهت کاهش اطلاعات موجود در تصاویر براساس آنالیز اجزاء اولیه، جهت کاهش داده‌های اضافی موجود در تصاویر ماموگرافی؛ بکارگیری طبقه‌بندی کننده SVM (با کرنل RBF)؛	۸۸	-
علیرضا نیک روان شلمانی [۶]	MIAS	استخراج ویژگی‌های بافتی و هندسی توده‌ها؛ طبقه‌بندی توده‌ها به دو کلاس خوش‌خیم و بدخیم با استفاده از شبکه عصبی پرسپترون	۹۳/۳۰	-
پزشکی و همکاران [۱۲]	MIAS DDSM	استفاده از تکنیک‌های پردازش تصویر جهت استخراج ناحیه تومور بوسیله روش FCM جهت قطعه‌بندی، و سپس استخراج ویژگی‌های (شدت، GLCM، شکل، LBP و FD) از تومور حاصل از تصویر قطعه‌بندی شده‌ی نهایی و طبقه‌بندی تومورها به دو دسته‌ی خوش‌خیم و بدخیم توسط روش SVM	۹۱/۳۷ ۹۳/۲۲	۹۲/۱۰ ۹۲/۰۶
بوشرا موگال [۱۳]	MIAS DDSM	استخراج ویژگی‌های بافتی (مانند انرژی، همبستگی، آنتروپی، کنتراست، تجانس، مجموع واریانس‌ها و مجموع میانگین‌ها مستخرج از GLCM)، و ترکیب تبدیل top-Hat و GLCM با شبکه عصبی بازگشتی جهت انجام طبقه‌بندی (سالم-غیرنرمال؛ خوش‌خیم-بدخیم)	۹۵ : خوش‌خیم و بدخیم ۹۸ : خوش‌خیم و بدخیم	۱۰۰ ۱۰۰

نام محقق	پایگاه داده	روش پژوهش	٪	
			صحت	حساسیت
ویلیام تورس [۱۴]	DDSM	روشی برای کاهش تشخیص مثبت کاذب در طبقه‌بندی تصاویر ماموگرافی، و طبقه‌بندی مناطق مورد نظر بر اساس وجود یا عدم وجود توده با استفاده از روش‌های Random Forest، REP Tree و IBK؛	۹۲/۲۹	۹۰/۱۵
باسما موهام [۱۵]	DDSM	استخراج ویژگی‌های هندسی، بافتی و ویولت؛ طبقه‌بندی جهت تمایز بین توده‌های خوش‌خیم و بدخیم با استفاده از طبقه‌بندی کننده‌های، شبکه عصبی مصنوعی، SVM و k-نزدیک‌ترین همسایه؛	۹۸/۹	۱۰۰
ماریا پرز [۱۶]	DDSM	استخراج ویژگی‌های بافتی جهت طبقه‌بندی تصاویر به سه دسته نرمال، خوش‌خیم و سرطانی با استفاده از شبکه عصبی مصنوعی	۸۴/۷۲	-
ناصر صفدریان و همکاران [۲۳]	DDSM	استخراج ویژگی‌های ساختاری و هندسی جهت طبقه‌بندی تصاویر دارای توده‌های خوش‌خیم و بدخیم با استفاده از طبقه‌بندی کننده‌های بیزین، شبکه‌های عصبی PNN و MLP، و روش‌های k-نزدیکترین همسایه، الگوریتم فازی TSK و روش SVM	۹۷	۱۰۰
بهترین نتایج روش پیشنهاد شده در این مقاله	DDSM	استخراج ویژگی‌های ساختاری و هندسی از توده‌های استخراج شده و طبقه‌بندی توده‌ها توسط بکارگیری طبقه‌بندی کننده‌ی SVM با سه کرنل؛ و سپس بهینه‌سازی پارامترهای SVM توسط الگوریتم بهینه‌سازی ملخ	۱۰۰	۱۰۰



بافت سینه نیز از روش‌های کلی پردازش تصاویر پزشکی تبعیت می‌کند.

علاوه بر موارد فوق، با توجه به جدول ۵، می‌توان نتایج روش مطرح شده در این مقاله را با نتایج روش‌هایی مقایسه کرد که در آنها نیز از پایگاه داده‌ی DDSM استفاده شده است. با توجه به این مورد، پژوهش حامد جباری و همکاران [۹] دارای صحت ۹۱/۰۴ درصد بوده است (شاخص‌های حساسیت و اختصاصیت گزارش نشده‌اند)، که نشان می‌دهد نتایج صحت نهایی الگوریتم ارائه شده در مقاله‌ی حاضر از پژوهش بالاتر بوده است [۹]. علاوه بر این، در پژوهش H. Pezeshki و همکاران [۱۲]، که از دو پایگاه داده‌ی تصاویر ماموگرافی MIAS و DDSM استفاده شده است، بهترین نتایج پارامترهای صحت، حساسیت و اختصاصیت

ناحیه‌بندی تصاویر را به روش دستی، روش بدون سرپرست یا اتوماتیک و روش با سرپرست یا نیمه‌اتوماتیک تقسیم‌بندی نمودند. علاوه بر این، آنها روش‌های ناحیه‌بندی ساختاری را که بصورت استفاده از ویژگی‌های مکانی تصویر یا محاسبه‌ی پارامترهای آماری برای ناحیه‌بندی تصاویر است را در دو دسته‌ی کلی بخش‌بندی به کمک ویژگی‌های ساختاری (که دارای زیربخش‌های روش‌های وابسته به لبه، روش‌های وابسته به ناحیه و روش‌های وابسته به پیکسل) و روش‌های بخش‌بندی آماری معرفی کردند [۲۵]. لذا با توجه به بررسی این منابع درخصوص بخش‌بندی تصاویر جهت تشخیص تومورها و ضایعات موجود در تصاویر پزشکی، نتیجه گرفته می‌شود که روش به کار گرفته شده در این مقاله برای آشکارسازی توده‌های سرطانی موجود در

از این پژوهش نشان‌دهنده‌ی عملکرد بالای روش پیشنهاد شده نسبت به سایر پژوهش‌های پیشین انجام شده در این زمینه است.

ملاحظات اخلاقی

پیروی از اصول اخلاق پژوهش

در این پژوهش از تصاویر پایگاه داده‌ی DDSM استفاده شده است. لازم به ذکر است که تصاویر موجود در این پایگاه داده بصورت رایگان جهت دسترسی وجود دارند.

حامی مالی

این مقاله حامی مالی ندارد.

مشارکت نویسندگان

مفهوم‌شناسی، تحقیق و روش‌شناسی: ناصر صفدریان؛ جمع‌آوری اطلاعات، منابع و نگارش مقاله: شادی یوسفیان دزفولی‌نژاد؛ ویراستاری، اصلاح نگارشی و نهایی سازی مقاله: ناصر صفدریان؛

تعارض منافع

طبق نظر نویسندگان هیچ‌گونه تضاد منافی در پژوهش حاضر وجود ندارد.

تشکر و قدردانی

نویسندگان مقاله از باشگاه پژوهشگران جوان دانشگاه آزاد اسلامی واحد تبریز کمال تشکر و قدردانی را دارند.

نیز پائین‌تر از نتایج گزارش شده در مقاله‌ی حاضر (حتی بدون اعمال الگوریتم بهینه‌سازی ارائه شده) بوده است که این موضوع نیز نشان‌دهنده‌ی دقت بالای روش ارائه شده در این مقاله است. تحلیل مشابهی در مطالعه‌ی Bushra Mughal و همکاران [۱۳] را نیز می‌توان مطرح کرد. در میان تمامی پژوهش‌های پیشین ارائه شده در این حوزه، روش ارائه‌شده توسط Basma A. Mo-hamed و همکاران [۱۵] دارای عملکرد بسیار بهتری در مقایسه با پژوهش‌های پیشین بوده است. با این حال، روش ارائه شده در مقاله‌ی حاضر با دقت مناسب‌تری به تشخیص و طبقه‌بندی توده‌های سرطانی سینه پرداخته است.

از نوآوری‌های موجود در این تحقیق می‌توان به تعداد بالای ویژگی‌های استخراج شده از توده‌های مستخرج از تصاویر ماموگرافی و همچنین استفاده از الگوریتم بهینه‌سازی ملخ جهت تعیین دقیق‌تر نوع توده‌های سرطانی بافت سینه که سعی شده است با دقت بالا صورت گیرد اشاره نمود. علاوه بر این، از نقاط قوت روش ارائه شده در این مقاله می‌توان به سرعت و دقت بسیار بالای الگوریتم ارائه شده اشاره نمود. به اینصورت که مرزهای تومورهای سرطانی با دقت بالا استخراج شده، و در نهایت عملیات طبقه‌بندی با بکارگیری ویژگی‌های مورفولوژیکی ساده صورت پذیرفته است. همچنین، با بررسی مطالعات پیشین که خلاصه‌ی آنها در **جدول ۵** ارائه شده است، مشخص است که تاکنون در پژوهش‌های پیشین اقدام به استفاده از روش‌های بهینه‌سازی در مرحله‌ی طبقه‌بندی نهایی (یا بهینه‌سازی طبقه‌بند بکار گرفته شده) نشده است. لذا بکارگیری روش بهینه‌سازی ملخ در بهینه کردن پارامترهای کرنل‌های مختلف طبقه‌بندی‌کننده‌ی SVM برای اولین بار در این مقاله ارائه شده است که می‌توان آغاز کننده‌ی روندهای جدید جهت بهبود فرایندهای مختلف طبقه‌بندی در انواع تشخیص‌های پزشکی باشد.

به عنوان ارائه‌ی نتیجه‌گیری کلی از این مقاله، می‌توان به این نکته اشاره نمود که پس از آشکارسازی و تشخیص توده‌های سرطانی سینه که با دقت بالا در این مطالعه انجام شده است، با توجه به ویژگی‌های مورفولوژیکی و ساده‌ی استخراج شده از توده‌های سرطانی، عملیات طبقه‌بندی بخوبی و با دقت بالا انجام پذیرفته است. در نتایج اولیه‌ی طبقه‌بندی‌کننده‌ی SVM، شاخص‌های صحت، حساسیت و اختصاصی بودن (یا شاخص ویژگی) مقادیر نسبتاً مناسبی را به دست آوردند. اما پس از بهینه‌سازی پارامترهای طبقه‌بندی‌کننده، مقادیر شاخص‌های ذکر شده افزایش قابل توجهی از خود نشان دادند. بر اساس مدل پیشنهاد شده در این مقاله، ویژگی‌های مبتنی بر شکل هندسی و ساختار توده‌های استخراج شده از تصویر ماموگرافی دیجیتال بافت پستان، جهت آموزش مدل و تشخیص نوع توده در سرطان پستان دارای کارایی بالایی هستند، و همچنین روش بهینه‌سازی ملخ با استفاده از بهینه‌سازی پارامترها، دقت کلی تشخیص سیستم ارائه شده را بهبود بخشیده است. در نهایت نتایج حاصل

References

- [1] World Health Organization. Cancer: Early diagnosis and screening [Internet]. 2020 [Retrieved 6 Aug 2020]. Available at: <https://www.who.int/cancer/prevention/diagnosis-screening/en>
- [2] Center of Disease Control. [Report of registered cancer cases (Persian)]. Tehran: Center of Disease Control; 2004.
- [3] Baines CJ, McFarlane DV, Miller AB. The role of the reference radiologist. Estimates of inter-observer agreement and potential delay in cancer detection in the national breast screening study. *Investigative radiology*. 1990; 25(9):971-6. [DOI:10.1097/00004424-199009000-00002] [PMID]
- [4] Wallis MG, Walsh MT, Lee JR. A review of false negative mammography in a symptomatic population. *Clin Radiol*. 1991; 44(1):13-5. [DOI:10.1016/S0009-9260(05)80218-1]
- [5] Behnam H, Zakeri F, Gifani P, Torkashvand P, Shalbaf A, [Ultrasound Imaging Processing (Persian)]. Tehran: Ishraqiya Publishing; 2011.
- [6] Nick Ravan Shalmani A, Karami Mohammadi M. [Diagnosis of breast cancer masses in computer aided mammography images (Persian)]. The 3rd International Conference on Recent Innovations in Electrical and Computer Engineering, 9 September 2016; Tehran, Iran.
- [7] Abbaspour Kazerouni I, Haddad Nia J. [Introducing a precise intelligent system for mammographic image separation based on density of tissues and masses (Persian)]. *Iran J Breast Dis*. 2013; 6(1).
- [8] Naseri Noroozani S, Shayegan MA. [Clinical stage detection of breast cancer patients using tnm system and ant colony algorithm (Persian)]. *Iran J Breast Dis*. 2018; 11(3):56-70.
- [9] Jabbari H, Bigdeli N, Khadem A. [A new hybrid approach to segmentation and diagnosis of tumors in breast mammography images (Persian)]. *Iran J Breast Dis*. 2016; 9(3):14-24.
- [10] Tavakkoloh P, Safabakhsh R., [A new approach to classifying and classifying breast cancer masses (Persian)] Third Information and Knowledge Technology Conference. 6-8 December 2007: Tehran; Iran.
- [11] Sheikhpour Ro, Sheikhpour Ra. [Diagnosis of breast cancer using non-parametric estimation of nuclear methods-based probability density (Persian)]. *Razi J Med Sci*. 2016; 23(144):30-40.
- [12] Pezeshki H, Rastgarpour M, Sharifi A, Yazdani S. Extraction of spiculated parts of mammogram tumors to improve accuracy of classification. *Multimed Tools Appl*. 2019; 78:1-25. [DOI:10.1007/s11042-019-7185-4]
- [13] Mughal B, Sharif M, Muhammad N, Saba T. A novel classification scheme to decline the mortality rate among women due to breast tumor. *Microsc Res Tech*. 2018; 81(2):171-80. [DOI:10.1002/jemt.22961] [PMID]
- [14] Torres W, Oseas A, Sousa A, Silva FA. Functional Diversity applied to the false positive reduction in breast tissues based on digital mammography. 2018 IEEE Symposium on Computers and Communications (ISCC), Natal. 2018; 25:1120-25. [DOI:10.1109/ISCC.2018.8538658]
- [15] Mohamed BA, Salem NM. Automatic classification of masses from digital mammograms. 2018, 35th National Radio Science Conference (NRSC). 2018 Mar 20; 495-502. IEEE. [DOI:10.1109/NRSC.2018.8354408]
- [16] Pérez M, Benalcázar ME, Tusa E, Rivas W, Conci A. Mammogram classification using back-propagation neural networks and texture feature descriptors. 2017 IEEE Second Ecuador Technical Chapters Meeting (ETCM). 2017 Oct 16; 1-6. [DOI:10.1109/ETCM.2017.8247515]
- [17] Digital Database for Screening Mammography. Available at: <http://marathon.csee.usf.edu/Mammography/Database.html>
- [18] Vadivel A, Surendiran B. A fuzzy rule-based approach for characterization of mammogram masses into BI-RADS shape categories. *Comput Biol Med*. 2013; 43(4):259-67. [DOI:10.1016/j.combiomed.2013.01.004] [PMID]
- [19] Huang T, Yang GJ, Tang G. A fast two-dimensional median filtering algorithm. *IEEE Trans Acoust Speech Signal Process*. 1979; 27(1):13-8. [DOI:10.1109/TASSP.1979.1163188]
- [20] Massodi P, Safdarian N, Kalantar B. [Detection and classification of breast cancer using feature extraction from mammography images by image processing technique and neural network (Persian)]. International Conference on Non-Linear System & Optimization in Computer & Electrical Engineering. May 2015: Dubai; United States of Emirates.
- [21] Fradkin D, Muchnik I. Support vector machines for classification. DI-MACS series in discrete mathematics and theoretical computer science. 2006; 70:13-20.
- [22] Saremi S, Mirjalili S, Lewis A. Grasshopper optimisation algorithm: Theory and application. *Adv Eng Softw*. 2017;105:30-47. [DOI:10.1016/j.advengsoft.2017.01.004]
- [23] Safdarian N, Hedyezadeh M. Detection and classification of breast cancer in mammography images using pattern recognition methods. *Multidiscip Cancer Investig*. 2019; 3(4):13-24 URL: <http://mcjournal.com/article-1-229-en.html> [DOI:10.30699/acadpub.mci.3.4.13]
- [24] Ghaffari H, Mostashari M, Mahmoodi M. [Designing a system for detection of pulmonary nodules in lung ct images using support vector machine classifier (Persian)]. *J Health Biomed Informatics, Med Informatics Res Center*. 2017; 3(4):300-9.
- [25] Biniiaz A, Abbasi A, Shamsi M. [A comprehensive overview of segmentation of medical images of magnetic resonance in the human brain: methods, clinical applications, advantages and disadvantages (Persian)]. *Majlis Electrical Engineering Conference*. 2012: New Majlis City: Iran.

This Page Intentionally Left Blank
