

## A New Computer-Aided Diagnosis System for Breast Cancer Detection in Thermal Images

Amirhassan Gheytsi<sup>1</sup>, Abdalhossein Rezai<sup>2\*</sup>, Seyedeh Shahrbanoo Falahieh Hamidpour<sup>3</sup> and Farzad Khajeh-Khalili<sup>4</sup>

1- Electrical and Biomedical Engineering Department, ACECR Institute of Higher Education, Isfahan Branch, Isfahan, Iran.

2\*- Department of Engineering, University of Science and Culture, Tehran, Iran.

3- Electrical and Biomedical Engineering Department, ACECR Institute of Higher Education, Isfahan Branch, Isfahan, Iran.

4- Electrical and Biomedical Engineering Department, ACECR Institute of Higher Education, Isfahan Branch, Isfahan, Iran

<sup>1</sup>amirhassangheytsi@gmail.com, <sup>2\*</sup>rezai@usc.ac.ir, <sup>3</sup>hamidpour@jdeihe.ac.ir, <sup>4</sup>khalili.farzad@gmail.com

Corresponding author's address: Abdalhossein Rezai, Department of Engineering, University of Science and Culture, Tehran, Iran.

**Abstract-** Breast cancer is one of the most common causes of death among women around the world, but early and accurate diagnosis of this type of cancer can dramatically improve treatment. Thermal imaging is one of the primary methods of diagnosing breast cancer. The computer diagnosis system can also be used to help physicians to increase the accuracy of interpretation of results. This paper presents an intelligent computer diagnostic system for the detection of breast cancer using thermal imaging. The proposed intelligent computer diagnosis system includes SFTA method for feature extraction and SVM, kNN and D-Tree algorithms for classification of results. The performance of the proposed intelligent computer diagnosis system is evaluated using the DMR-IR and Fluminense Federal University databases and MATLAB2018, when using the cuckoo feature selection algorithm and without using the feature selection algorithm. The results show that the average accuracy, sensitivity and specificity are 99%, 99.5% and 98.03%, respectively, using the cuckoo feature selection algorithm and SVM classification algorithm. Also, the presented computer diagnostic system has advantages compared to other computer diagnosis systems. These results indicate that the use of SFTA feature extraction method, cuckoo feature selection algorithm, SVM classification algorithm and DMR-IR database in the proposed computer diagnosis system can improve the evaluation results.

**Keywords-** Breast cancer, cuckoo feature selection algorithm, thermal imagery, feature extraction and feature selection.

## ارائه یک سیستم تشخیص کامپیوتری هوشمند جدید برای تشخیص سرطان سینه با استفاده از تصاویر حرارتی

امیرحسن قیطاسی<sup>۱</sup>، عبدالحسین رضائی<sup>۲\*</sup>، سیده شهربانو فلاحیه حمیدپور<sup>۳</sup>، فرزاد خواجه خلیلی<sup>۴</sup>

۱- گروه مهندسی برق و مهندسی پزشکی، موسسه آموزش عالی جهاددانشگاهی استان اصفهان، اصفهان، ایران.

۲\*- گروه مهندسی برق، دانشکده فنی و مهندسی، دانشگاه علم و فرهنگ، تهران، ایران.

۳- گروه مهندسی برق و مهندسی پزشکی، موسسه آموزش عالی جهاددانشگاهی استان اصفهان، اصفهان، ایران.

۴- گروه مهندسی برق و مهندسی پزشکی، موسسه آموزش عالی جهاددانشگاهی استان اصفهان، اصفهان، ایران.

<sup>1</sup>amirhassangheytsi@gmail.com, <sup>2\*</sup>rezai@usc.ac.ir, <sup>3</sup>hamidpour@jdeihe.ac.ir, <sup>4</sup>khalili.farzad@gmail.com

\* نشانی نویسنده مسئول: عبدالحسین رضائی، تهران، بلوار اشرفی اصفهان، دانشگاه علم و فرهنگ، دانشکده فنی و مهندسی.

چکیده- سرطان سینه از شایع ترین دلایل مرگ و میر در میان زنان جهان می باشد، اما تشخیص زودهنگام و دقیق این نوع سرطان می تواند درمان را به طرز چشمگیری بهبود بخشد. تصویربرداری حرارتی یکی از روش های اولیه تشخیص سرطان سینه است. همچنین از سیستم تشخیص کامپیوتری می توان برای کمک به پزشکان برای افزایش دقت تفسیر نتایج استفاده کرد. در این مقاله، یک سیستم تشخیص کامپیوتری هوشمند برای تشخیص سرطان سینه با استفاده از تصاویر حرارتی ارائه شده است. سیستم هوشمند تشخیص کامپیوتری ارائه شده شامل روش SFTA برای استخراج ویژگی و الگوریتم SVM، kNN و D-Tree برای طبقه بندی نتایج می باشد. عملکرد سیستم تشخیص کامپیوتری هوشمند ارائه شده با استفاده از پایگاه داده DMR-IR و پایگاه داده دانشگاه Fluminense Federal و MATLAB2018 برای استفاده از الگوریتم انتخاب ویژگی فاخته و بدون الگوریتم انتخاب ویژگی ارزیابی شده است. نتایج نشان می دهد که بهترین میانگین صحت، حساسیت و اختصاصیت به ترتیب ۹۹/۵، ۹۹/۵ و ۹۸/۰۳ درصد می باشد که با کارگیری الگوریتم انتخاب ویژگی فاخته و الگوریتم طبقه بندی کننده SVM بدست آمده است. همچنین سیستم تشخیص کامپیوتری ارائه شده دارای مزایایی نسبت به سایر سیستم های تشخیص کامپیوتری می باشد. این نتایج نشان می دهد که استفاده از روش استخراج ویژگی SFTA، الگوریتم انتخاب ویژگی فاخته، الگوریتم طبقه بندی SVM و داده های پایگاه داده DMR-IR در سیستم تشخیص کامپیوتری پیشنهادی، می تواند باعث بهبود نتایج ارزیابی شود.

واژه های کلیدی: سرطان سینه، الگوریتم انتخاب ویژگی فاخته، تصاویر حرارتی، استخراج ویژگی و انتخاب ویژگی.

### ۱- مقدمه

تصویربرداری حرارتی (ترموگرافی<sup>۱</sup>) یکی از روش های تشخیص

ناهنجاری های سینه با استفاده از تصاویر حرارتی است. در این روش، از یک دوربین حرارتی برای ضبط اشعه مادون قرمز ساطع شده از سینه برای تشخیص الگوهای حرارتی استفاده می شود. بنابراین، تصویربرداری حرارتی یک ابزار غربالگری بدون درد، کم هزینه و غیرتهاجمی است که می تواند برای همه زنان در همه سنین مورد استفاده قرار گیرد [۴-۹].

سرطان سینه رایج ترین نوع سرطان و یکی از مهم ترین دلایل مرگ و میر بین زنان است. مطالعات پاتولوژیک نشان می دهد که بیشتر ناهنجاری های سینه، در مراحل اولیه خوش خیم هستند. بنابراین تشخیص زودهنگام مهم ترین مسئله در درمان این نوع از سرطان می باشد [۳-۱].

سیستم تشخیص کامپیوتری خود را با استفاده از داده‌های پایگاه داده DMR-IR ارزیابی کردند. نتایج به دست آمده نشان می‌دهد که صحت طبقه‌بندی‌کننده EBT بهتر از طبقه‌بندی‌کننده AdaBoost است. میانگین صحت به دست آمده برای طبقه‌بندی‌کننده EBT در سیستم تشخیص کامپیوتری آنها ۸۷ درصد گزارش شده است.

احمد و همکاران [۱۳] یک سیستم تشخیص کامپیوتری برای تشخیص سرطان سینه با استفاده از تصاویر حرارتی ارائه نموده‌اند که در آن از الگوریتم‌های بهینه‌سازی کلنی مورچگان ( $ACO^{14}$ ) و بهینه‌سازی ازدحام ذرات ( $PSO^{15}$ ) برای انتخاب ویژگی و از الگوریتم SVM برای طبقه‌بندی استفاده شده است. آنها سیستم تشخیص کامپیوتری خود را با استفاده از مجموعه داده‌های DMR-IR ارزیابی کردند. میانگین صحت به دست آمده برای الگوریتم‌های انتخاب ویژگی ACO و PSO در سیستم تشخیص کامپیوتری آنها به ترتیب ۹۴/۲۹ و ۹۷/۱۴ درصد گزارش شده است. ساتیش و همکاران [۴] با استفاده از تصاویر حرارتی یک سیستم تشخیص کامپیوتری برای تشخیص سرطان سینه ارائه داده‌اند که در آن از الگوریتم‌های ژنتیک و RSFS برای انتخاب ویژگی و از الگوریتم SVM برای طبقه‌بندی استفاده شده است. آنها سیستم تشخیص کامپیوتری خود را با استفاده از پایگاه داده دانشگاه Fluminense Federal ارزیابی کردند. میانگین صحت به دست آمده برای الگوریتم انتخاب ویژگی RSFS در سیستم تشخیص کامپیوتری توسعه یافته آنها ۹۱ درصد گزارش شده است. رودریگز و همکاران [۱۴] با استفاده از تصاویر حرارتی یک سیستم تشخیص کامپیوتری برای تشخیص سرطان سینه ارائه داده‌اند که در آن از الگوریتم‌های PSO و ژنتیک برای انتخاب ویژگی و از الگوریتم SVM برای طبقه‌بندی استفاده شده است. آنها سیستم تشخیص کامپیوتری خود را با استفاده از ۳۳۶ تصویر حرارتی مورد ارزیابی قرار دادند. در این روش، بردار ویژگی برای هر تصویر حرارتی از ۱۶۹ ویژگی به ۶۰ و ۵۷ ویژگی به ترتیب برای الگوریتم‌های انتخاب ویژگی PSO و ژنتیک کاهش یافت. میانگین صحت به دست آمده برای الگوریتم‌های انتخاب ویژگی PSO و ژنتیک در سیستم تشخیص کامپیوتری ارائه شده در [۱۴] به ترتیب ۸۶/۱۶ و ۸۷/۰۸ درصد است.

پیرا و همکاران [۱۵] یک سیستم تشخیص کامپیوتری برای تشخیص سرطان سینه با استفاده از تصاویر حرارتی ارائه داده‌اند که در آن ترکیبی از روش‌های بهینه‌سازی  $ELM^{16}$  و  $ODM^{17}$  برای انتخاب ویژگی و Naïve Bayes، Bayes Net،  $J48$ ، SVM، درخت تصادفی و الگوریتم‌های جنگل تصادفی برای

با وجود اثربخشی تصویربرداری حرارتی در تشخیص سرطان سینه، تشخیص صحیح تمام ناهنجاری‌ها در این تصاویر توسط پزشکان دشوار است. علاوه بر این، تشخیص صحیح و یکسان بین پزشکان مختلف برای تعداد زیادی از تصاویر حرارتی یک چالش می‌باشد. در نتیجه، محققان سیستم‌های تشخیص کامپیوتری<sup>۲</sup> را برای کمک به پزشکان در افزایش دقت در تفسیر تصاویر حرارتی ارائه نمودند [۵، ۱۰، ۱۱]. از این رو، توسعه سیستم تشخیص کامپیوتری هوشمند برای تشخیص سرطان سینه با استفاده از تصاویر حرارتی مورد توجه قرار گرفته و تلاش‌های زیادی برای بهبود عملکرد سیستم‌های تشخیص کامپیوتری با استفاده از تصاویر حرارتی صورت پذیرفته است [۴-۶، ۱۲-۱۷].

در این مقاله، یک سیستم تشخیص کامپیوتری هوشمند برای تشخیص سرطان سینه با استفاده از تصاویر حرارتی ارائه و ارزیابی می‌شود. در سیستم تشخیص کامپیوتری ارائه شده از روش  $SFTA^3$  برای استخراج ویژگی، از الگوریتم بهینه‌سازی فاخته<sup>۴</sup> برای انتخاب ویژگی و از الگوریتم‌های  $k$ -نزدیکترین همسایگی ( $kNN^5$ )، ماشین بردار پشتیبانی ( $SVM^6$ ) و درخت تصمیم‌گیری ( $D-Tree^7$ ) برای مرحله طبقه‌بندی استفاده شده است. عملکرد سیستم تشخیص کامپیوتری ارائه شده با استفاده از پایگاه داده DMR-IR<sup>۸</sup> [۱۸]، پایگاه داده دانشگاه Fluminense Federal و MATLAB 2018 مورد ارزیابی قرار گرفته است. مقدار متوسط صحت، حساسیت و اختصاصیت به دست آمده اثربخشی سیستم تشخیص کامپیوتری هوشمند ارائه شده را نشان می‌دهد.

در ادامه مطالب این مقاله به شرح زیر است: بخش ۲ مروری بر مطالعات گذشته را ارائه می‌دهد. بخش ۳ روش تحقیق را ارائه می‌دهد. بخش ۴ نتایج شبیه‌سازی را ارائه می‌دهد و ارزیابی عملکرد سیستم تشخیص کامپیوتری هوشمند ارائه شده را توضیح می‌دهد. سرانجام، جمع‌بندی این مقاله در بخش ۵ ارائه شده است.

## ۲- مروری بر مطالعات گذشته

بهبود عملکرد سیستم تشخیص کامپیوتری هوشمند می‌تواند نقش مهمی در تشخیص سرطان سینه داشته باشد. بنابراین، محققان زیادی با استفاده از تصاویر حرارتی در این زمینه کار کرده‌اند. ساتیش و همکاران [۱۲] با استفاده از تصاویر حرارتی یک سیستم تشخیص کامپیوتری برای تشخیص سرطان سینه ارائه داده‌اند که در آن الگوریتم‌های  $t$ -test،  $SFS^9$ ،  $SFFS^{10}$ ،  $RSFS^{11}$  و ژنتیک<sup>۱۲</sup> برای انتخاب ویژگی و الگوریتم‌های  $EBT^{13}$  و AdaBoost برای طبقه‌بندی مورد استفاده قرار گرفته است. آنها

عملکرد سیستم تشخیص کامپیوتری پیشنهادی با استفاده از داده-های پایگاه داده DMR-IR مورد ارزیابی قرار گرفته است. طبقه بندی تصاویر حرارتی شامل سرطانی و سالم براساس ۳ نوع شبکه عصبی کانولوشن عمیق شامل google net، resnet18 و vgg16 انجام شده است. میانگین صحت بدست آمده با استفاده از مدل های شبکه های عصبی عمیق از پیش آموزش داده شده google net، resnet18 و vgg16 به ترتیب برابر ۸۵/۰۳ و ۸۳/۸ و ۸۵/۰۳ درصد می باشد.

سانچز-کوس و همکاران [۲۵] یک سیستم تشخیص کامپیوتری برای تشخیص سرطان سینه با ترکیب تصاویر حرارتی از تصاویر با زاویه های مختلف با داده های شخصی و بالینی ارائه دادند. سیستم پیشنهادی آنها یک سیستم چند ورودی بوده و از مزایای شبکه های عصبی کانولوشن برای تجزیه و تحلیل تصویر بهره می برد. آنها ابتدا، سیستم پیشنهادی را تنها با استفاده از تصاویر حرارتی بررسی نمودند و در مرحله بعد، داده های بالینی را با هدف بهبود عملکرد اضافه کردند. میانگین صحت بدست آمده با استفاده از مدل شبکه های عصبی کانولوشن چند ورودی برابر ۹۷ درصد می باشد.

دین و همکاران در [26] مروری بر روش های تشخیص سرطان سینه با استفاده از یادگیری عمیق و چالش های پیش رو داشته اند. آنها در [۲۶] به این نتیجه رسیده اند که مدل های مبتنی بر شبکه عصبی کانولوشن به نتایج امیدوارکننده ای در کاربرد تصویربرداری پزشکی برای تشخیص سرطان دست یافته اند. اما، مدل های مبتنی بر شبکه عصبی کانولوشن به طور گسترده برای تصاویر حرارتی استفاده نشده است.

اگرچه عملکرد سیستم های تشخیص ارائه شده مناسب هستند، اما سیستم های تشخیص کامپیوتری فعلی هنوز برای تشخیص زودهنگام و دقیق سرطان سینه به بهبود بیشتر نیاز دارند. هدف اصلی در این مقاله تجزیه و تحلیل این است که آیا استفاده از الگوریتم بهینه سازی جستجوی فاخته به عنوان الگوریتم انتخاب ویژگی و انتخاب روش مناسب در مراحل استخراج ویژگی و طبقه بندی نتایج، ظرفیت تشخیصی یک مدل برای تشخیص زودهنگام سرطان سینه را در تصاویر حرارتی بهبود می بخشد یا خیر.

### ۳- روش پیشنهادی

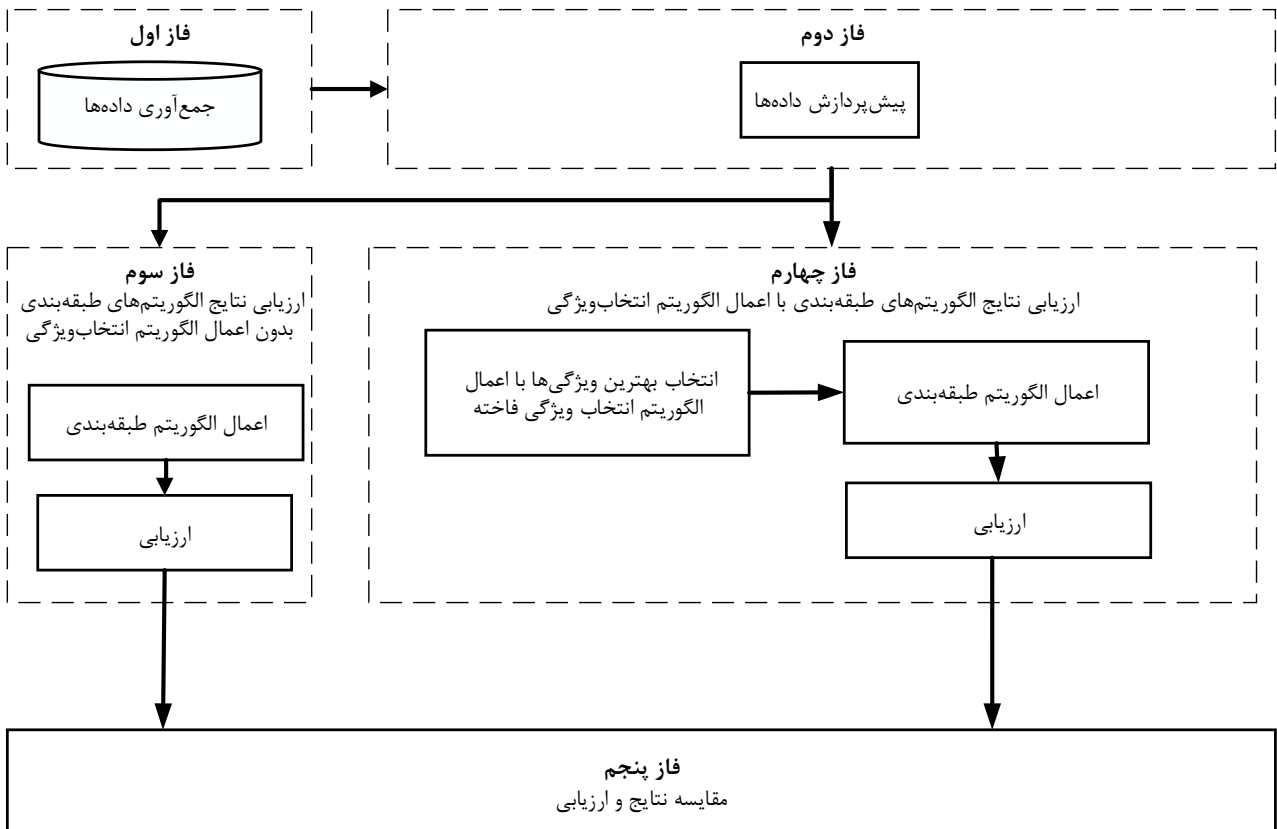
در شکل (۱) فلوچارت سیستم تشخیص کامپیوتری هوشمند پیشنهادی نشان داده شده است.

طبقه بندی استفاده شده است. آنها سیستم تشخیص کامپیوتری خود را با استفاده از ۹۶۸ تصویر ارزیابی کردند. نتایج نشان داد که در این سیستم بردار ویژگی ها حدود ۵۷ درصد کاهش می یابد. میانگین صحت بدست آمده در سیستم تشخیص کامپیوتری ارائه شده در [۱۵] به ترتیب ۴۹/۹۸، ۵۰/۶۴، ۷۸/۷۲، ۶۷/۳۹، ۵۰/۸۱، ۵۶/۲۵ و ۷۰/۱۵ درصد برای Naïve Bayes، Bayes Net، MLP، SVM، درخت تصادفی، J48 و طبقه بندی جنگل تصادفی گزارش شده است. الفایز و همکاران [۱۶] یک سیستم تشخیص کامپیوتری برای تشخیص سرطان سینه با استفاده از تصاویر حرارتی ارائه داده اند که در آن از ELM و MLP برای مرحله طبقه بندی استفاده شده است. آنها سیستم تشخیص کامپیوتری خود را با استفاده از داده های پایگاه داده DMR-IR مورد ارزیابی قرار دادند. میانگین صحت بدست آمده برای طبقه بندی کننده های MLP و ELM به ترتیب ۸۰/۰۴ و ۹۹/۱ درصد گزارش شده است.

دارابی و همکاران [۶] با استفاده از تصاویر حرارتی یک سیستم تشخیص کامپیوتری برای تشخیص سرطان سینه ارائه داده اند که در آن از الگوریتم RSFS و ترکیب الگوریتم mRMR<sup>19</sup> و الگوریتم ژنتیک با الگوریتم RSFS برای انتخاب ویژگی و از الگوریتم های kNN و SVM برای طبقه بندی استفاده شده است. آنها سیستم تشخیص کامپیوتری خود را با استفاده از ۱۲۱ تصویر حرارتی مورد ارزیابی قرار دادند. میانگین صحت به دست آمده برای الگوریتم طبقه بندی kNN و الگوریتم انتخاب ویژگی RSFS، ترکیب الگوریتم های انتخاب ویژگی mRMR و RSFS و ترکیب الگوریتم های انتخاب ویژگی ژنتیک و RSFS به ترتیب ۸۵/۳۶، ۷۷/۴۱ و ۸۳/۸۷ درصد گزارش شده است. پرامانیک و همکاران [۱۷] با استفاده از تصاویر حرارتی یک سیستم تشخیص کامپیوتری برای تشخیص سرطان سینه ارائه داده اند که در آن از مدل BBP<sup>20</sup> برای بهبود تصاویر حرارتی استفاده شده است. سپس، از فیلتر محلی متغیر با اندازه متغیر (DVL<sup>21</sup>) برای استخراج ویژگی تصاویر بهبود یافته حرارتی استفاده نموده اند. سرانجام، از FBN برای طبقه بندی استفاده شده است. آنها سیستم تشخیص کامپیوتری خود را با استفاده از داده های پایگاه داده DMR-IR مورد ارزیابی قرار دادند. میانگین صحت بدست آمده ۹۵/۶ درصد گزارش شده است.

سلیمیان و همکاران [۱۹] ویژگی های موثر در تصاویر حرارتی برای تشخیص سرطان سینه را مورد بررسی قرار داده اند.

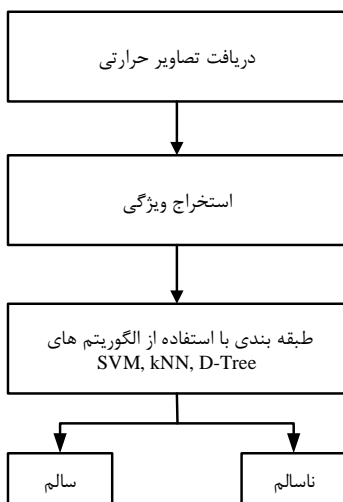
قلی زاده و همکاران [۲۴] به بررسی سیستم های تشخیص کامپیوتری مبتنی بر شبکه های عصبی عمیق از نوع کانولوشن در جهت تشخیص سرطان سینه در تصاویر ترموگرافی پرداخته اند.



شکل (۱): فلوجارت سیستم تشخیص کامپیوتری هوشمند پیشنهادی

### ۳-۳- فاز سوم: ارزیابی نتایج بدون اعمال الگوریتم انتخاب ویژگی

این فاز خود شامل چند مرحله است که در شکل (۲) نشان داده شده است.



شکل (۲): فلوجارت ارزیابی نتایج بدون اعمال الگوریتم انتخاب ویژگی

روند کلی انجام تحقیق شامل پنج فاز متشکل از جمع‌آوری داده‌ها، پیش‌پردازش داده‌ها، ارزیابی نتایج بدون استفاده از الگوریتم انتخاب ویژگی، ارزیابی نتایج با استفاده از الگوریتم انتخاب ویژگی فاخته و تحلیل و بررسی نتایج می‌باشد که در ادامه مورد بررسی قرار می‌گیرند.

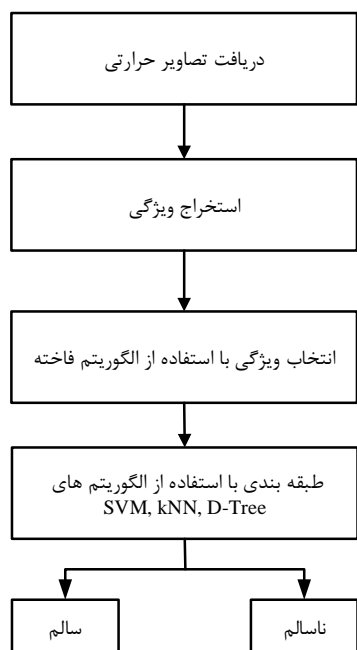
### ۳-۱- فاز اول: جمع‌آوری داده‌ها

در این مقاله از ۱۲۰ تصویر حرارتی سینه از تصاویر موجود در پایگاه داده DMR-IR [۱۸] و ۶۰ تصویر حرارتی سینه از تصاویر موجود پایگاه داده دانشگاه Fluminense Federal استفاده شده است. فرمت تصاویر jpg است.

### ۳-۲- فاز دوم: پیش‌پردازش داده‌ها

تصاویر پایگاه داده DMR-IR [۱۸] و پایگاه داده دانشگاه Fluminense Federal به دو دسته تقسیم شدند: دسته اول شامل تصاویر افراد مبتلا به سرطان و دسته‌ی دوم شامل تصاویر افراد سالم. به منظور پیش‌پردازش، تصاویر سینه‌ی سمت راست و سمت چپ از هم و از پس زمینه جداسازی شدند. تصاویر پس از بخش‌بندی به فرمت سطح خاکستری تبدیل شدند.

SVM و kNN، به دو دسته سالم و ناسالم دسته‌بندی می‌شوند. فلوجارت روش پیشنهادی در شکل (۳) نشان داده شده است.



شکل (۳): فلوجارت روش پیشنهادی

در این فاز، مرحله انتخاب ویژگی با استفاده از الگوریتم بهینه-ساز جستجوی فاخته به عنوان الگوریتم انتخاب ویژگی به مراحل قبل اضافه شده است. فلوجارت الگوریتم فاخته براساس فلوجارت ارائه شده در [۲۰] در شکل (۴) نشان داده شده است.

این الگوریتم مانند سایر الگوریتم‌های تکاملی با یک جمعیت اولیه از فاخته‌ها شروع می‌شود. این فاخته‌های اولیه تخم‌هایی برای قرار دادن در لانه‌های پرندگان میزبان دارند. برخی از این تخم‌ها که شباهت بیشتری به تخم‌های پرنده میزبان دارند، فرصت رشد کردن و تبدیل شدن به فاخته بالغ را دارند. تخم‌های دیگر توسط پرنده‌های میزبان شناسایی و از بین برده می‌شوند. تعداد تخم‌های رشد یافته در هر منطقه، میزان مناسب بودن لانه‌های آن منطقه را نشان می‌دهند. در یک منطقه، هر چه تخم‌های بیشتری زنده بمانند، سود بیشتری در آن منطقه به دست می‌آید. پس مکانی که تخم‌های بیشتری در آن زنده بمانند چیزی است که الگوریتم فاخته در پی بهینه کردن آن است [۲۰].

همانگونه که در شکل (۲) نشان داده شده است، مراحل استخراج ویژگی و طبقه بندی، مراحل مهم این فاز می‌باشند که در ادامه مورد بررسی قرار می‌گیرند.

### ۳-۱-۳-۳ مرحله استخراج ویژگی

یکی از مهم‌ترین مراحل در تشخیص و طبقه‌بندی تصاویر حرارتی، استخراج ویژگی‌های مناسب است. در این مقاله ویژگی‌های تصاویر حرارتی از روش SFTA استخراج می‌شوند.

### ۳-۲-۳-۳ مرحله طبقه بندی

در این مرحله، از الگوریتم‌های مبتنی بر قواعد ریاضی و یا زبانی برای تفکیک کلاس‌های مسئله استفاده می‌شود. ورودی مرحله طبقه‌بندی کننده، ویژگی‌هایی است که در بخش استخراج ویژگی به آن پرداخته شد و خروجی این مرحله برچسب کلاسی است که به هر یک از دسته ویژگی‌ها داده می‌شود. نظر به گستردگی فضای ویژگی و تعداد نسبتاً زیاد نمونه‌های استفاده شده در سیستم تشخیص کامپیوتری هوشمند پیشنهادی، از سه الگوریتم طبقه‌بندی کننده زیر استفاده شده است:

الف- الگوریتم درخت تصمیم‌گیری (DTree)

ب- الگوریتم k-نزدیک‌ترین همسایگی (kNN)

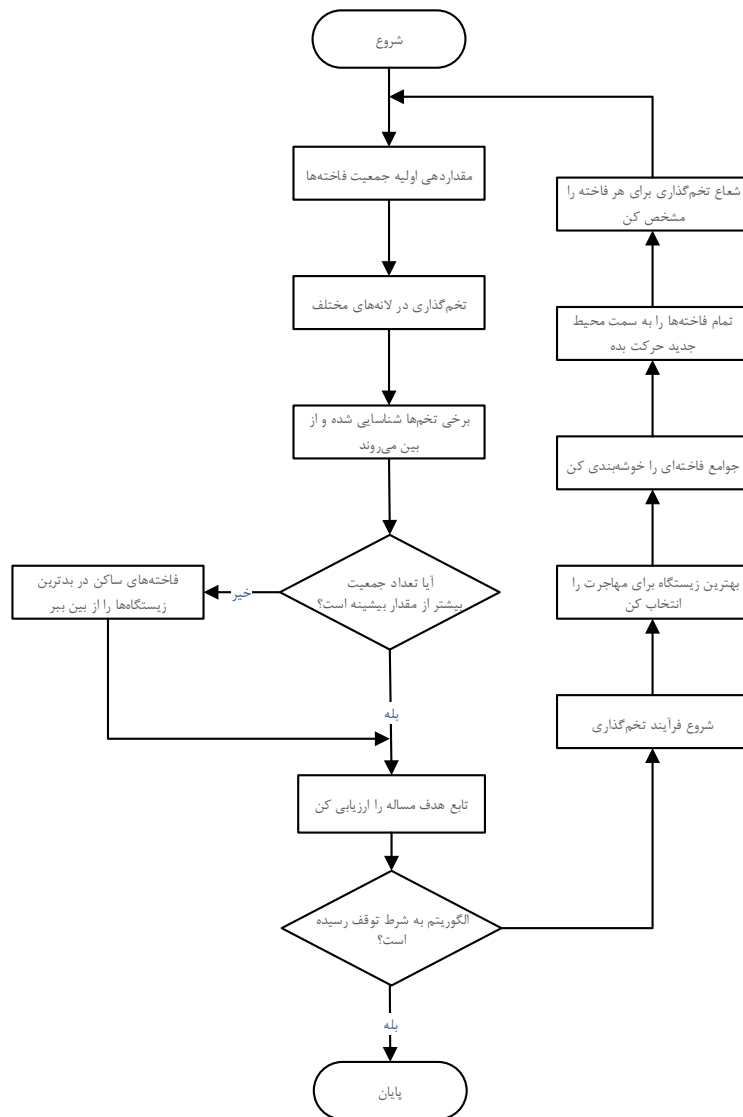
ج- الگوریتم ماشین بردار پشتیبان (SVM)

در ادامه به تشریح این الگوریتم‌های طبقه‌بندی و قوانین حاکم بر آن‌ها پرداخته می‌شود.

### ۳-۴-۳ فاز چهارم: ارزیابی نتایج با اعمال الگوریتم انتخاب

#### ویژگی فاخته

در این مرحله، الگوریتم انتخاب ویژگی فاخته به ویژگی‌های استخراج شده به روش SFTA اعمال می‌شود تا بهترین ویژگی‌ها انتخاب شوند، سپس تصاویر حرارتی براساس ویژگی‌های انتخاب شده و با استفاده از الگوریتم‌های طبقه‌بندی کننده DTree،



شکل (۴): فلوجارت الگوریتم بهینه سازی فاخته براساس الگوریتم ارائه شده در [۲۰]

### ۳-۵- فاز پنجم: مقایسه و ارزیابی

برای ارزیابی سیستم تشخیص کامپیوتری پیشنهادی برای تشخیص سرطان سینه، شش مدل شامل ۳ مدل بدون استفاده از الگوریتم انتخاب ویژگی و ۳ مدل با استفاده از الگوریتم انتخاب ویژگی فاخته ارائه شده است. در این مرحله روش ارزیابی مدل‌های ارائه شده مورد بررسی قرار می‌گیرد. عملکرد سیستم‌های تشخیص کامپیوتری، معمولاً براساس معیارهای صحت<sup>۲۲</sup>، حساسیت<sup>۲۳</sup> و اختصاصیت<sup>۲۴</sup> مورد ارزیابی قرار می‌گیرد که مطابق روابط زیر [۶] محاسبه می‌شوند.

$$\text{Accuracy (\%)} = \frac{TP + TN}{TN + TP + FN + FP} \times 100 \quad (1)$$

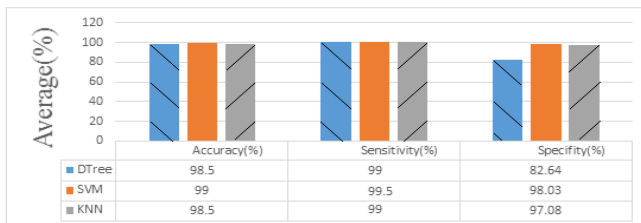
$$\text{Sensitivity (\%)} = \frac{TP}{TP + FN} \times 100 \quad (2)$$

درواقع، فاخته‌ها به دنبال مناسب‌ترین منطقه برای تخم‌گذاری هستند تا نرخ زنده ماندن تخم‌های آنها را بیشتر کند. پس از آنکه تخم‌های باقی مانده رشد کردند و به فاخته بالغ تبدیل شدند، جوامعی را تشکیل می‌دهند. هر جامعه زیست بوم خودش را دارد. بهترین محل زندگی یافت‌شده توسط فاخته‌ها، مقصد فاخته‌ها در جوامع دیگر خواهد بود. آنها در جایی نزدیک بهترین زیست بوم ساکن می‌شوند. با توجه به تعداد تخم‌های هر فاخته و همچنین فاصله فاخته از نقطه مقصد (بهترین زیست بوم) یک شعاع تخم-گذاری به آن داده می‌شود. سپس، فاخته شروع به قراردادن تخم-هایش در لانه‌های تصادفی درون شعاع تخم‌گذاری‌اش می‌کند. این فرآیند ادامه می‌یابد تا بهترین موقعیت با بیشترین سود به دست آید و اکثر جمعیت فاخته‌ها در این مکان جمع شوند [۲۰].

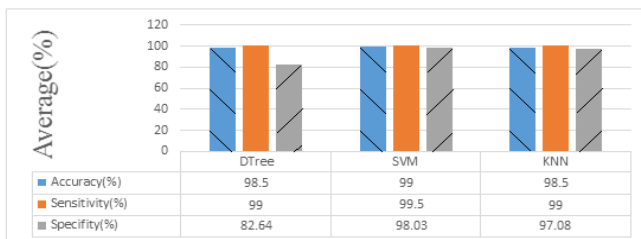
با توجه به نتایج ارائه شده در شکل (۵)، بهترین میانگین صحت و حساسیت به ترتیب با مقادیر ۸۹/۵ و ۹۱/۵۷ درصد مربوط به الگوریتم طبقه‌بندی کننده‌ی SVM و بهترین میانگین اختصاصیت با مقدار ۹۴/۹۳ درصد مربوط به الگوریتم طبقه‌بندی کننده‌ی kNN می‌باشد.

#### ۴-۲- نتایج ارزیابی الگوریتم‌های طبقه‌بندی کننده با استفاده از الگوریتم انتخاب ویژگی فاخته در پایگاه داده DMR-IR

مقدار میانگین صحت، حساسیت و اختصاصیت برای ۳۰ تکرار در نرم‌افزار MATLAB2018، برای الگوریتم‌های طبقه‌بندی کننده DTREE، SVM و kNN با استفاده از الگوریتم انتخاب ویژگی فاخته در شکل (۶) نشان داده شده است.



(الف)



(ب)

شکل (۶): مقدار صحت، حساسیت و اختصاصیت محاسبه شده برای الگوریتم‌های طبقه‌بندی کننده DTREE، SVM و kNN با استفاده از الگوریتم انتخاب ویژگی فاخته (الف) بر اساس الگوریتم‌های طبقه‌بندی کننده (ب) بر اساس معیارهای اندازه‌گیری

با توجه به نتایج ارائه شده در شکل (۶)، بهترین میانگین صحت، حساسیت و اختصاصیت به ترتیب با مقادیر ۹۹، ۹۹/۵ و ۹۸/۰۳ درصد مربوط به الگوریتم طبقه‌بندی کننده SVM می‌باشد.

$$\text{Specificity (\%)} = \frac{TN}{FP + TN} \times 100 \quad (3)$$

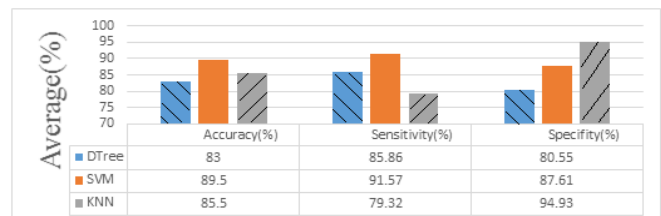
در این روابط TP، FP، TN و FN به ترتیب مثبت واقعی، مثبت کاذب، منفی واقعی و منفی کاذب را نشان می‌دهند.

#### ۴- شبیه‌سازی و مقایسه

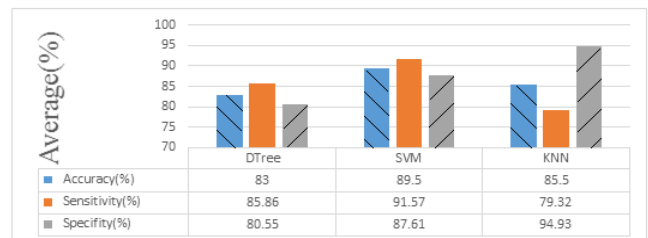
سیستم تشخیص کامپیوتری پیشنهادی با استفاده از نرم‌افزار MATLAB2018 پیاده‌سازی و اجرا شده است. صحت عملکرد سیستم تشخیص کامپیوتری با استفاده از ۱۲۰ تصویر حرارتی سینه موجود در پایگاه داده DMR-IR و ۶۰ تصویر حرارتی سینه تصاویر موجود در پایگاه داده دانشگاه Fluminense Federal مورد ارزیابی قرار گرفته است. تصاویر این پایگاه داده‌ها به دو دسته سالم و ناسالم تقسیم‌بندی شده‌اند. در ادامه نتایج شبیه‌سازی‌ها و ارزیابی‌های حاصل ارائه شده است.

#### ۴-۱- نتایج ارزیابی الگوریتم‌های طبقه‌بندی کننده بدون استفاده از الگوریتم انتخاب ویژگی در پایگاه داده DMR-IR

مقدار میانگین صحت، حساسیت و اختصاصیت برای ۳۰ تکرار در نرم‌افزار MATLAB2018، برای الگوریتم‌های طبقه‌بندی کننده DTREE، SVM و kNN بدون استفاده از الگوریتم انتخاب ویژگی در شکل (۵) نشان داده شده است.



(الف)



(ب)

شکل (۵): مقدار صحت، حساسیت و اختصاصیت محاسبه شده برای الگوریتم‌های طبقه‌بندی کننده DTREE، SVM و kNN بدون استفاده از الگوریتم انتخاب ویژگی (الف) بر اساس الگوریتم‌های طبقه‌بندی کننده (ب) بر اساس معیارهای اندازه‌گیری



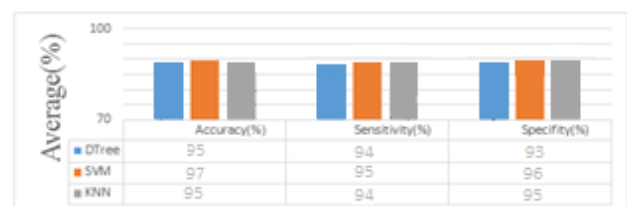
در [۱۶] می‌باشد. علاوه بر این سیستم تشخیص کامپیوتری هوشمند ارائه شده دارای بهترین میانگین حساسیت در مقایسه با سایر سیستم‌های تشخیص کامپیوتری می‌باشد. همچنین سیستم تشخیص کامپیوتری هوشمند ارائه شده در این مقاله دارای میانگین اختصاصیت بهتری در مقایسه با سیستم‌های تشخیص کامپیوتری ارائه شده در [۴، ۱۲، ۱۴، ۱۶، ۱۷، ۲۴] است و تنها سیستم‌های تشخیص کامپیوتری که تاحدودی میانگین اختصاصیت بهتری در مقایسه با سیستم تشخیص کامپیوتری ارائه شده در این مقاله دارند، سیستم تشخیص کامپیوتری ارائه شده در [۱۳، ۲۵] می‌باشند، البته لازم به ذکر است که سیستم تشخیص کامپیوتری ارائه شده در این مقاله دارای میانگین صحت و حساسیت بهتری نسبت به سیستم تشخیص کامپیوتری ارائه شده در [۱۳، ۲۵] می‌باشد. این نتایج اثربخشی سیستم تشخیص کامپیوتری هوشمند ارائه شده در این مقاله را برای بهبود روش تشخیص سرطان سینه نشان می‌دهد.

#### ۵- جمع‌بندی

تشخیص زود هنگام و دقیق سرطان سینه یک مسئله چالش برانگیز مهم در سلامت عمومی در جهان است. پزشکان به سیستم تشخیص کامپیوتری نیاز دارند تا به تشخیص آنها به طور موثر کمک کند. از این رو سیستم‌های تشخیص کامپیوتری زیادی برای کمک به پزشکان در این مسئله چالش برانگیز ارائه شده‌است. با این وجود، سیستم‌های تشخیص کامپیوتری فعلی هنوز نیاز به بهبود بیشتری برای تشخیص زود هنگام و دقیق سرطان سینه دارند. در این مقاله، یک سیستم تشخیص کامپیوتری جدید و هوشمند برای تشخیص سرطان سینه با استفاده از تصاویر حرارتی ارائه شده‌است. در سیستم تشخیص کامپیوتری هوشمند ارائه شده، از روش SFTA برای استخراج ویژگی و از الگوریتم‌های SVM، kNN و D-Tree برای طبقه‌بندی نتایج، استفاده شده‌است. عملکرد سیستم تشخیص کامپیوتری ارائه شده با استفاده از پایگاه داده DMR-IR و پایگاه داده دانشگاه Fluminense Federal و MATLAB 2018 برای هر دو حالت بدون استفاده از الگوریتم انتخاب‌ویژگی و با استفاده از الگوریتم انتخاب‌ویژگی فاخته نشان داد که بهترین میانگین صحت، حساسیت و اختصاصیت به ترتیب با مقادیر ۹۹/۵، ۹۹/۵ و ۹۸/۰۳ درصد با بکارگیری الگوریتم انتخاب‌ویژگی فاخته و الگوریتم طبقه‌بندی کننده SVM بدست می‌آید. همچنین سیستم تشخیص کامپیوتری ارائه شده دارای عملکرد بهتری نسبت به سایر سیستم‌های تشخیص کامپیوتری می‌باشد. این نتایج نشان می‌دهد که استفاده از روش استخراج ویژگی SFTA،

#### ۳-۴- نتایج ارزیابی الگوریتم‌های طبقه‌بندی کننده با استفاده از الگوریتم انتخاب ویژگی فاخته در پایگاه داده دانشگاه Fluminense Federal

مقدار میانگین صحت، حساسیت و اختصاصیت در این حالت برای ۳۰ تکرار در در نرم‌افزار MATLAB2018، برای الگوریتم‌های طبقه‌بندی کننده DTree، SVM و kNN با استفاده از الگوریتم انتخاب ویژگی فاخته در شکل (۷) نشان داده شده است. با توجه به نتایج ارائه شده در شکل (۷)، بهترین میانگین صحت، حساسیت و اختصاصیت به ترتیب با مقادیر ۹۷، ۹۵ و ۹۶ درصد مربوط به الگوریتم طبقه‌بندی کننده SVM می‌باشد.



شکل (۷): مقدار صحت، حساسیت و اختصاصیت محاسبه شده برای الگوریتم‌های طبقه‌بندی کننده DTree، SVM و kNN با استفاده از الگوریتم انتخاب ویژگی فاخته در پایگاه داده دانشگاه

#### Fluminense Federal

این نتایج نشان می‌دهد که استفاده از الگوریتم انتخاب ویژگی فاخته، باعث بهبود معیارهای ارزیابی شده است. همچنین استفاده از الگوریتم طبقه‌بندی SVM برای روش پیشنهادی با استفاده از داده‌های پایگاه داده DMR-IR نتایج مناسب‌تری ارائه می‌نماید.

#### ۴-۴- مقایسه با روش‌های پیشین

جدول ۱ نتایج میانگین صحت، حساسیت و اختصاصیت را برای روش‌های مختلف براساس بهترین میانگین صحت نشان می‌دهد.

براساس تجزیه و تحلیل انجام شده که نتایج آن در جدول ۱ نشان داده شده‌است، سیستم تشخیص کامپیوتری هوشمند ارائه شده در این مقاله دارای میانگین صحت بهتری در مقایسه با سیستم‌های تشخیص کامپیوتری ارائه شده در [۴، ۶، ۱۲، ۱۴-۱۷، ۲۱، ۲۳-۲۵] است و تنها سیستم تشخیص کامپیوتری که تاحدودی میانگین صحت بهتری در مقایسه با سیستم تشخیص کامپیوتری ارائه شده در این مقاله دارد، سیستم تشخیص کامپیوتری ارائه شده در [۱۶] است، البته لازم به ذکر است که سیستم تشخیص کامپیوتری ارائه شده در این مقاله دارای میانگین حساسیت و اختصاصیت بهتری نسبت به سیستم تشخیص کامپیوتری ارائه شده

الگوریتم انتخاب ویژگی فاخته، الگوریتم طبقه‌بندی SVM و داده-های پایگاه داده DMR-IR در سیستم تشخیص کامپیوتری پیشنهادی، می‌تواند باعث بهبود نتایج ارزیابی شود.

جدول ۱: میانگین نتایج براساس بهترین میانگین صحت در روش‌های مختلف

اختصاصیت (%)	حساسیت (%)	صحت (%)	الگوریتم طبقه‌بندی - کننده	الگوریتم انتخاب ویژگی	پایگاه داده	مرجع (سال)
-	-	۹۸/۷۷	SVM	WOA	WDBC	[23] (2016)
-	-	98/96	SVM	GA	WDBC	[21] (2017)
۹۰/۶	۸۳	۸۷	EBT	RSFS	DMR-IR	[12] (2018)
۹۸/۶	۹۸	۹۷/۱۴	SVM	PSO	DMR-IR	[13] (2019)
۹۴/۲۹	۹۲/۹۸	۸۷/۰۸	SVM	ژنتیک	A set of 336 cases	[14] (2019)
-	-	۷۸/۷۲	MLP	ترکیب ELM and ODM	A set of 968 cases	[15] (2020)
۹۷/۰۳	۹۸/۰۵	99/۱	ELM	-	DMR-IR	[16] (2020)
-	۹۴/۱۱	۸۵/۳۶	kNN	RSFS	A set of 121 cases	[6] (2021)
۹۷	۹۴	۹۵/۶	FBN	DVLF	DMR-IR	[17] (2021)
۹۴/۳۴	۸۷/۲۳	۹۱	SVM	RSFS	دانشگاه Fluminense Federal	[4] (2019)
۹۱/۰۱	۶۶/۴۳	۸۵/۰۳	Vgg16	Vgg16	DMR-IR	[۲۴] (2019)
100	83	97	Multi-input CNN	Multi-input CNN	DMR-IR	[25] (2019)
۹۶	۹۵	۹۷	SVM	فاخته	دانشگاه Fluminense Federal	این مقاله
۸۷/۶۱	۹۱/۵۷	۸۹/۵	SVM	بدون فاخته	DMR-IR	
۹۸/۰۳	۹۹/۵	۹۹	SVM	فاخته		

infrared thermal images," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, pp. 1-7, 2021.

- [6] N. Darabi, A. Rezai, and S. S. Falahieh Hamidpour, "Breast cancer detection using RSFS-based feature selection algorithms in thermal images," *Biomedical Engineering: Applications, Basis and Communications*, 2021.
- [7] T. Sarigoz and T. Ertan, "Role of dynamic thermography in diagnosis of nodal involvement in patients with breast cancer: A pilot study," *Infrared Physics & Technology*, vol. 108, p. 103336, 2020/08/01/2020, doi: <https://doi.org/10.1016/j.infrared.2020.103336>.
- [8] A. Lozano and F. Hassanipour, "Infrared imaging for breast cancer detection: An objective review of foundational studies and its proper role in breast cancer screening," *Infrared Physics & Technology*, vol. 97, pp. 244-257, 2019/03/01/2019, doi: <https://doi.org/10.1016/j.infrared.2018.12.017>.
- [9] U. Raghavendra, A. Gudigar, T. N. Rao, E. J. Ciacchio, E. Y. K. Ng, and U. Rajendra Acharya, "Computer-aided diagnosis for the identification of breast cancer using thermogram images: A comprehensive review," *Infrared Physics & Technology*, vol. 102, p. 103041, 2019/11/01/2019, doi: <https://doi.org/10.1016/j.infrared.2019.103041>.
- [10] A. S. Eltrass and M. S. Salama, "Fully automated scheme for computer-aided detection and breast cancer diagnosis using digitised

## مراجع

- [1] A. Lashkari, F. Pak, and M. Firouzmand, "Full intelligent cancer classification of thermal breast images to assist physician in clinical diagnostic applications," *Journal of medical signals and sensors*, vol. 6, no. 1, p. 12, 2016.
- [2] T. Andreadis, C. Emmanouilidis, S. Goumas, and D. Koulouriotis, "Development of an intelligent CAD system for mass detection in mammographic images," *IET Image Processing*, vol. 14, no. 10, pp. 1960-1966, 2020.
- [3] M. Navid, S. S. F. Hamidpour, F. Khajeh-Khalili, and M. Alidoosti, "A novel method to infrared thermal images vessel extraction based on fractal dimension," *Infrared Physics & Technology*, vol. 107, p. 103297, 2020/06/01/2020, doi: <https://doi.org/10.1016/j.infrared.2020.103297>.
- [4] D. Sathish, S. Kamath, K. Prasad, and R. Kadavigere, "Role of normalization of breast thermogram images and automatic classification of breast cancer," *The Visual Computer*, vol. 35, no. 1, pp. 57-70, 2019/01/01 2019, doi: 10.1007/s00371-017-1447-9.
- [5] M. Zarei, A. Rezai, and S. S. Falahieh Hamidpour, "Breast cancer segmentation based on modified Gaussian mean shift algorithm for

- [18] L. Silva *et al.*, "A new database for breast research with infrared image," *Journal of Medical Imaging and Health Informatics*, vol. 4, no. 1, pp. 92-100, 2014.
- [19] M. Salimian, A. Rezai, S. Hamidpour, and F. Khajeh-Khalili, "Effective Features in Thermal Images for Breast Cancer Detection," presented at the 2nd National Conference on New Technologies in Electrical and Computer Engineering, 2019.
- [20] R. Rajabioun, "Cuckoo optimization algorithm," *Applied soft computing*, vol. 11, no. 8, pp. 5508-5518, 2011.
- [21] E. Aličković and A. Subasi, "Breast cancer diagnosis using GA feature selection and Rotation Forest," *Neural Computing and Applications*, 2017.
- [23] G. I. Sayed, A. Darwish, A. E. Hassanien, and J.-S. Pan, "Breast cancer diagnosis approach based on meta-heuristic optimization algorithm inspired by the bubble-net hunting strategy of whales," in *International Conference on Genetic and Evolutionary Computing*, 2016.
- [24] M.H Golizadeh " The Investigation of Deep Convolutional Neural Network for Diagnosing Breast Cancer in Thermographic Images [jsmj.ajums.ac.ir/article\\_104943.html](http://jsmj.ajums.ac.ir/article_104943.html)
- [25] R. Sánchez-Cauce, J. Pérez-Martín, M. Luquel, "Multi-input convolutional neural network for breast cancer detection using thermal images and clinical data," *Computer Methods and Programs in Biomedicine*, vol. 204, 106045, 2021.
- [26] N. Din, R. Dar, M. Rasool, A. Assad, Breast cancer detection using deep learning: Datasets, methods, and challenges ahead, *Computers in Biology and Medicine*, vol.149, 106073, 2022.
- mammograms," *IET Image Processing*, vol. 14, no. 3, pp. 495-505, 2020.
- [11] N. Ahmad, S. Asghar, and S. A. Gillani, "Transfer learning-assisted multi-resolution breast cancer histopathological images classification," *The Visual Computer*, 2021/05/13 2021, doi: 10.1007/s00371-021-02153-y.
- [12] D. Sathish and S. Kamath, "Detection of Breast Thermograms using Ensemble Classifiers," *Journal of Telecommunication, Electronic and Computer Engineering (JTEC)*, vol. 10, no. 3-2, pp. 35-39, 2018.
- [13] A. Ahmed, M. Ali, and M. Selim, "Bio-inspired based techniques for thermogram breast cancer classification," *International Journal of Intelligent Engineering and Systems*, vol. 12, no. 2, pp. 114-124, 2019.
- [14] A. L. Rodrigues *et al.*, "Identification of mammary lesions in thermographic images: feature selection study using genetic algorithms and particle swarm optimization," *Research on Biomedical Engineering*, vol. 35, no. 3, pp. 213-222, 2019/12/01 2019, doi: 10.1007/s42600-019-00024-z.
- [15] J. M. S. Pereira, M. Santana, W. Silva, R. Lima, S. Lima, and W. Santos, "Dialectical optimization method as a feature selection tool for breast cancer diagnosis using thermographic images," *Understanding a Cancer Diagnosis*, pp. 95-118.
- [16] F. AlFayez, M. W. A. El-Soud, and T. Gaber, "Thermogram Breast Cancer Detection: A Comparative Study of Two Machine Learning Techniques," *Applied Sciences*, vol. 10, no. 2, 2020, doi: 10.3390/app10020551.
- [17] S. Pramanik, D. Bhattacharjee, and M. Nasipuri, "Breast Abnormality Detection Using Texture Feature Extracted by Difference-Based Variable-Size Local Filter (DVLF)," in *Proceedings of Research and Applications in Artificial Intelligence*, Singapore, I. Pan, A. Mukherjee, and V. Piuri, Eds., 2021// 2021: Springer Singapore, pp. 111-120.

## پاورقی ها:

Ensemble Bagged Trees	<sup>13</sup>	Thermography	<sup>1</sup>
Ant Colony Optimization	<sup>14</sup>	Computer-aided diagnosis systems	<sup>7</sup>
Particle Swarm Optimization	<sup>15</sup>	Segmentation-based Fractal Texture Analysis	<sup>3</sup>
Extreme Learning Machine	<sup>16</sup>	Cuckoo optimization algorithm	<sup>4</sup>
Optimal Multiuser Detection	<sup>17</sup>	k-Nearest Neighbors	<sup>5</sup>
Multi-Layer Perceptron	<sup>18</sup>	Support Vector Machines	<sup>6</sup>
minimal Redundancy Maximal Relevance	<sup>19</sup>	Decision Tree	<sup>7</sup>
Breast Blood Perfusion	<sup>20</sup>	Database for Mastology Research with Infrared Image	<sup>8</sup>
Difference-based Variable-size Local Filter	<sup>21</sup>	Sequential Forward Selection	<sup>9</sup>
Accuracy	<sup>22</sup>	Sequential Floating Forward Selection	<sup>10</sup>
Sensitivity	<sup>23</sup>	Resilient Sequential Forward Selection	<sup>11</sup>
Specificity	<sup>24</sup>	Genetic	<sup>12</sup>