



# ارائه مدل‌های محاسبات نرم مبتنی بر فازی، تکاملی و هوش جمعی در تحلیل تصاویر ماموگرافی

## جهت تشخیص تومورهای سینه

الناز خدادادی<sup>۱</sup>، راحیل حسینی<sup>\*۲</sup> و مهدی مزینانی<sup>۳</sup>

<sup>۱،۲</sup> گروه مهندسی کامپیوتر، واحد شهر قدس، دانشگاه آزاد اسلامی، تهران، ایران

<sup>۳</sup> گروه مهندسی برق الکترونیک، واحد شهر قدس، دانشگاه آزاد اسلامی، تهران، ایران

### چکیده

مدل‌های محاسبات نرم مبتنی بر سامانه‌های هوشمند فازی در تشخیص سرطان سینه، امکان مدیریت عدم قطعیت در فرایند استدلال در سامانه را فراهم می‌کند. در این پژوهش، یک مدل استنتاج فازی به منظور مدیریت عدم قطعیت در داده‌های ورودی طراحی شده است و الگوریتم‌های هابیریدی مبتنی بر فازی جهت تنظیم و بهینه‌سازی پارامترها، به کار برده شده‌اند. هدف، ارائه روش‌های مؤثر جهت تشخیص نوع توده‌های خوش خیم، بدخیم و نرمال سینه است. طبقه‌بندی توده‌ها جهت تشخیص موارد نرمال، خوش خیم و بدخیم با مدل‌های هابیریدی محاسبات نرم و بر اساس تحلیل ویژگی‌ها در تصاویر ماموگرافی انجام شده است. الگوریتم‌های هابیریدی ارائه شده در این پژوهش شامل (۱) فازی-ژنتیک، (۲) فازی-بهینه‌سازی ازدحام ذرات و (۳) فازی-بهینه‌سازی مبتنی بر جغرافیای زیستی است. به منظور سنجش عملکرد سامانه از تحلیل منحنی مشخصه (ROC) و همچنین از روش اعتبارسنجی تقاطعی ده‌بخشی جهت تقسیم‌بندی داده‌ها به بخش‌های آموزش و آزمون برای به دست آوردن نتایج قابل اعتماد و اعتبارسنجی استفاده شده است. نوآوری پژوهش حاضر در ارائه مدل پیشنهادی هابیریدی فازی-بهینه‌سازی مبتنی بر جغرافیای زیستی و بهبود عملکرد مدل طبقه‌بندی جهت تشخیص سرطان سینه است. روش جدید هابیریدی فازی-بهینه‌سازی مبتنی بر جغرافیای زیستی ارائه شده به منظور تشخیص سرطان سینه در این پژوهش، عملکرد بالاتری نسبت به روش‌های موجود بر روی این بانک اطلاعاتی معتبر و معروف جهت تشخیص سرطان سینه داشته است. با توجه به نتایج به دست آمده و مقایسه عملکرد مدل‌های هابیریدی پیشنهادی در این پژوهش، روش هابیریدی فازی مبتنی بر جغرافیای زیستی با میزان صحت ۹۵/۲۵٪ از عملکرد بهینه‌تری نسبت به روش‌های هابیریدی پیشنهادی دیگر جهت تشخیص سرطان سینه برخوردار است. مدل حاضر در مقایسه با سایر مدل‌های پیشنهادی در پژوهش‌های قبلی بهبود یافته است. استفاده از مدل‌های پیشنهادی در این پژوهش، می‌تواند به منظور تشخیص زودهنگام بیماری و ارائه درمان‌های مؤثر امیدبخش باشد.

واژگان کلیدی: سیستم استنتاج فازی، محاسبات نرم، الگوریتم‌های هابیریدی فازی-تکاملی، فازی-بهینه‌سازی ذرات و فازی-جغرافیای زیستی، تومورهای سینه

## Soft Computing Methods based on Fuzzy, Evolutionary and Swarm Intelligence for Analysis of Digital Mammography Images for Diagnosis of Breast Tumors

Elnaz Khodadadi<sup>1</sup>, Rahil Hosseini<sup>\*2</sup> & Mahdi Mazinani<sup>3</sup>

<sup>1&2</sup> Department of Computer Engineering, Shahr-e-Qods Branch, Islamic Azad University, Tehran, Iran

<sup>3</sup> Department of Electrical Engineering, Shahr-e-Qods Branch, Islamic Azad University, Tehran, Iran

\* Corresponding author

\* نویسنده عهده‌دار مکاتبات

## Abstract

Soft computing models based on intelligent fuzzy systems have the capability of managing uncertainty in the image based practices of disease. Analysis of the breast tumors and their classification is critical for early diagnosis of breast cancer as a common cancer with a high mortality rate between women all around the world. Soft computing models based on fuzzy and evolutionary algorithms play an important role in advances obtained in computer aided detection (CAD) systems. Combination of the evolutionary nature of swarm intelligence algorithms in optimization along with the potential of fuzzy models to cope with uncertainty and complex environments.

In this research, a fuzzy inference model has been proposed for managing uncertainty in input data. The main uncertainty issues for classification of the breast tumors were modeled through the linguistic terms, fuzzy variables and fuzzy reasoning processes in the fuzzy inference model. Fuzzy linguist terms and rule sets are valuable to have an intelligent model with the ability to interact with the clinicians. Furthermore, hybrid fuzzy-evolutionary models have been proposed for tuning fuzzy membership functions for diagnosis of malignant and benign breast tumors. The hybrid proposed evolutionary methods are: 1) Fuzzy-Genetic, 2) Fuzzy-Particle swarm intelligence, and 3) Fuzzy-biogeography models. Evolutionary nature inspired combination with the fuzzy inference model (FIM) improves the proficiency of the FIM by adaption to the environment through the tuning process using training and testing datasets. To achieve this, the Genetic Algorithm was applied as a base evolutionary method. Then, the potential of the Particle Swart intelligence algorithm in using local and global experiences of the solutions in the search space. Also, bio-geographical aspects of species in finding an optimum solution lands with the high suitability habitat index has been concentrated in optimization process of the FIM. Evolutionary algorithms perform tuning of the fuzzy membership functions to improve the accuracy of the fuzzy inference model while simplicity and interpretability of the FIM was kept. For performance evaluation, an ROC curve analysis was conducted which is a robust and reliable technique that represents the trades of between classification model benefits and costs. Also, for validation purpose, a 10-fold cross-validation technique was performed for partitioning the dataset into training and testing sets in the evolutionary optimization algorithms. The performance of the proposed methods were evaluated using a dataset including 295 images and extracted features from mammographic image analysis society (MIAS) dataset. The results reveal that the hybrid Fuzzy-biogeography model outperforms the other evolutionary models with an accuracy and area under the ROC curve (AUC) of 95.25%, and 91.43%, respectively. Performance comparison of the hybrid evolutionary models in this study with the related methods for classification of the breast tumors on the MIAS dataset reveals that the fuzzy-biogeography model outperforms the other methods in terms of trades-off between accuracy and interpretability with an area under the ROC curve of 95.25% with four extracted features. The Fuzzy-GA and Fuzzy-Swarm Intelligence models are competitive with the best results of counterpart methods with an accuracy of 93.9% and 94.58% in terms of the AUC, respectively. The proposed fuzzy-evolutionary models in this study are promising for diagnosis of the breast tumors in early stages of the disease and providing suitable treatment.

**Keywords:** Fuzzy Inference System, Soft Computing, Hybrid Evolutionary Algorithms, Breast Tumours

به میزان به روز سرطان سینه در کشورهای مختلف جهان افزوده شده است. با توجه به پایین آمدن میانگین سن ابتلا به این بیماری و اینکه بیشینه جمعیت زنان در ایران جوان هستند، روند رو به رشد این بیماری را در طول ده سال آینده خواهیم داشت. با نرخ رو به رشد وقوع سرطان سینه، تقاضا برای گسترش فناوری و پیشرفت فناوری موجود برای شناسایی سرطان سینه افزایش یافته است. احتمال درمان موفق و بهبود کامل برای بیمار به شناسایی و تشخیص زودهنگام سرطان سینه وابسته است [3]. پیشرفت‌های شگرفی در زمینه تشخیص بیماری با استفاده از مدل‌های هوشمند در زمینه پزشکی به دست آمده است. در این زمینه می‌توان به مقاله آشکارسازی عروق شبکه چشم براساس یک الگوریتم هوشمند جهت تشخیص عروق خونی اشاره کرد [1]. همچنین طبقه‌بندی بیماری پارکینسون بر مبنای شاخص‌های

## ۱- مقدمه

سرطان سینه ناشی از رشد خارج از قاعده سلول‌های غیرطبیعی در سینه است. سرطان سینه شایع‌ترین سرطان در بانوان و یکی از قابل درمان‌ترین بیماری‌ها با ۹۷٪ شانس زنده ماندن است [1-2]. بنابراین شناسایی زودهنگام تومور سینه تنها دلیل، برای کاهش نرخ مرگ و میر سرطان سینه است.

حدود شصت هزار نفر در ایران مبتلا به سرطان سینه‌اند. در ایران یک نفر از هر هشت تا ده زن به این بیماری مبتلا می‌شود و سالانه حدود ده هزار مورد جدید هم به آمار مبتلایان به سرطان سینه اضافه می‌شود. میزان ابتلا به سرطان سینه در زنان آسیایی بیش از زنان آمریکایی و اروپایی است. در شصت سال اخیر به طور متوسط سالانه یک تا چهار درصد

طراحی شوند که قابلیت بیشتری داشته باشند، استفاده شده است. محاسبات نرم، امکان ترکیب روش‌های هوش مصنوعی را با قابلیت مدیریت عدم قطعیت فراهم کرده است. این روش‌ها از تجربیات یاد می‌گیرند و با تغییرات، خود را تطبیق می‌دهند و می‌توانند از فرایند تصمیم‌گیری انسان در شرایط عدم قطعیت تبعیت کنند و تصمیمی با قابلیت اطمینان بالا بگیرند. همچنین، در پژوهش حاضر، استخراج ویژگی‌های ورودی متفاوت از پایگاه داده انجمن تجزیه و تحلیل تصاویر ماموگرافی، اندازه مناسب جهت تقسیم‌بندی کافی میان داده‌های آزمون و آموزش و همچنین استفاده از ویژگی‌های کمینه جهت رسیدن به نرخ طبقه‌بندی بیشینه جهت تشخیص سرطان سینه انجام شده است.

## ۲- مروری بر مدل‌های هوشمند مرتبط با تشخیص سرطان سینه

روش‌های متفاوتی برای مسأله تشخیص سرطان سینه ارائه شده است و از این میان می‌توان به مدل‌های شبکه‌های عصبی، یادگیری ماشین، روش‌های فازی، الگوریتم‌های متاهیوریستیک و ترکیبی از روش‌های بالا اشاره کرد. در این پژوهش روش‌های معرفی شده در مرور مقالات جمع‌آوری شده، از پایگاه داده انجمن تجزیه و تحلیل تصاویر ماموگرافی<sup>۱</sup> (MIAS) استفاده می‌کنند. روش‌های هوشمند مرتبط با سرطان سینه ارائه شده بر روی پایگاه داده انجمن تجزیه و تحلیل تصاویر ماموگرافی در این پژوهش مورد بررسی قرار گرفته است.

در مقاله بررسی شده در [8]، روش انتخاب ویژگی موجه هم‌رخداد با استفاده از سامانه فازی ژنتیک در مسأله طبقه‌بندی تصاویر ماموگرافی مطرح شده است. ویژگی‌های استخراج شده از مرحله استخراج ویژگی شامل انرژی، برآمدگی خوشه، سایه خوشه، مجموع واریانس، مجموع میانگین آنتروپی از ماتریس هم‌رخداد سطح خاکستری<sup>۲</sup> (GLCM) محاسبه شده‌اند و درواقع ویژگی‌های موجه هم‌رخداد هستند. ویژگی‌های موجه هم‌رخداد از جزئیات ضریب‌های موجه در هر سطح از تجزیه از تصاویر به دست آمده است. در چهار سطح از تجزیه ویژگی‌ها انتخاب شده و نرخ‌های میانگین طبقه‌بندی موفق یاد شده است. بهترین نرخ میانگین طبقه‌بندی در سطح چهار از تجزیه با انتخاب شانزده ویژگی از ۷۲ ویژگی ۸۹/۴۷٪ به دست آمده است.

درون ناحیه‌ای و بین ناحیه‌ای شبکه حرکتی مغز است. در این روش مدلی از شبکه ارتباطات عملکردی در افراد سالم و بیماران ارائه می‌شود و طبقه‌بندی بین افراد سالم و پارکینسونیسم انجام می‌گیرد [2]. می‌توان مقایسه ویژگی‌های خطی و غیرخطی سیگنال تغییرات نرخ ضربان قلب به منظور کمی‌سازی سطح استرس با استفاده از الگوریتم‌های هوشمند تکاملی و شبکه عصبی را ذکر کرد. دستگاه اعصاب خودکار، دستگاه قلبی-عروقی را در شرایط مختلف کنترل می‌کند. ایستای این سامانه متناسب با حالات روانی مختلف از جمله استرس دست‌خوش تغییرات شده و از این رو می‌توان به وسیله تحلیل سیگنال‌های خروجی آن [32]، فعالیت سامانه اعصاب خودکار را آشکار کرد. جهت تحلیل سیگنال، ویژگی‌های خطی و غیر خطی استخراج می‌شود؛ سپس با استفاده از ترکیب الگوریتم تکاملی و شبکه عصبی، ویژگی‌های بهینه که بیشترین درصد تفکیک بین سطوح مختلف استرس را مهیا می‌سازند، مشخص شده است و در نهایت با استفاده از شبکه عصبی و ویژگی‌های بهینه، تفکیک بین سطوح انجام شده است [4].

تشخیص سرطان سینه در دانش پزشکی با عدم قطعیت همراه است. ورودی‌ها و خروجی‌ها توأم با عدم قطعیت و داده‌ها نادقیق و کیفی هستند. عوامل تأثیرگذار در تشخیص، پیش‌گیری و درمان سرطان سینه مشخص نیست. یکی از روش‌های تشخیص سرطان سینه در دانش پزشکی توسط پزشک متخصص از روی تصاویر ماموگرافی سینه و دانش فرد خبره است که با ابهام و عدم قطعیت فراوان همراه است. به‌طور کلی یک سامانه تحلیل و پردازش تصاویر ماموگرافی از عدم قطعیت در دانش پزشکی و در برخی از روش‌های پردازش تصویر از جمله تشخیص لبه، قطع‌بندی و عدم دقت در منابع تولید تصاویر پزشکی رنج می‌برد [5,6,31].

باتوجه به وجود منابع عدم قطعیت در تشخیص سرطان سینه، روند رو به رشد این بیماری و ویژگی قابل درمان بودن آن در صورت تشخیص زودهنگام، اجرای برنامه‌های تشخیص و پیش‌گیری از سرطان سینه امری ضروری است. بنابراین نقش سامانه‌های هوشمند در تشخیص سرطان می‌تواند بسیار تأثیرگذار باشد. سامانه‌های هوشمند در تشخیص سرطان سینه، امکان مدیریت عدم قطعیت در فرایند استدلال را فراهم می‌کنند.

در این پژوهش جهت تشخیص سرطان سینه و به‌منظور مدیریت منابع عدم قطعیت از روش‌های محاسبات نرم مبتنی بر منطق فازی، با هدف اینکه سامانه‌ها یا مدل‌هایی

<sup>1</sup> Mammography Imaging Analysis Society

<sup>2</sup> Gray Level Co-Occurrence Matrix

این سامانه از یک شناسایی دقیق با استفاده از تکامل تفاضلی مبتنی بر الگوریتم خوشه‌بندی فازی برای تشخیص ضایعه‌ها در تصاویر ماموگرافی استفاده می‌کند. این الگوریتم روی پایگاه داده انجمن تجزیه و تحلیل تصاویر ماموگرافی آزمایش شده و ضایعه‌ها را شناسایی کرده است. عملکرد این روش با ارزش صحیح تصویر که توسط رادیولوژیست به صورت دستی نشانه‌گذاری شده، مقایسه شده است. قطعه‌بندی با الگوریتم تکامل تفاضلی مبتنی بر خوشه‌بندی فازی به نشانه‌گذاری واقعی رادیولوژیست نزدیک است و می‌تواند برای تشخیص پزشکی پیشنهاد شود [9].

در روش ارائه‌شده در مقاله [10]، ابتدا ناحیه مورد نظر سینه توسط الگوریتم بهینه‌سازی ازدحام ذرات استخراج شده است؛ سپس ویژگی‌های شکل و بافت سینه توده‌ها به ترتیب با استفاده از تبدیل فوری و ماتریس هم‌رخداد سطح خاکستری استخراج شده‌اند. طبقه‌بندی توده‌های شناسایی شده توسط ماشین بردار پشتیبان انجام شده است که ناحیه قطعه‌بندی شده را بر پایه ویژگی‌های استخراج شده از روش‌های یاد شده به نرمال و غیرنرمال طبقه‌بندی می‌کند. میزان صحت طبقه‌بندی ۸۷/۸۴٪ به دست آمده است.

یک روش طبقه‌بندی و قطعه‌بندی تمام خودکار برای تصاویر ماموگرافی بر پایه تخمین تراکم سینه و شناسایی عدم تقارن معرفی شده است [11]. ابتدا پیش‌پردازش تصویر و روش‌های قطعه‌بندی، شامل الگوریتم استخراج مرز سینه و یک نوع بهبودیافته از روش قطعه‌بندی ماهیچه قرار گرفته روی سینه، به کار برده شده است. ویژگی‌ها برای طبقه‌بندی تراکم سینه استخراج شده است و شامل یک ویژگی جدید وابسته به بعد می‌باشد. از ماشین بردارهای پشتیبان به منظور طبقه‌بندی استفاده نموده اند که به میزان صحت بالای ۸۵/۷٪ رسیده‌اند. این خصوصیات برای استخراج یک مجموعه از ویژگی‌های آماری به ازای هر سینه استفاده شده است. از تفاوت میان مقدار این ویژگی‌ها جهت شناسایی عدم تقارن بین دو تصویر در هر جفت تصویر ماموگرافی استفاده شده است. در این روش مدل طبقه‌بند یک کلاسه ماشین بردار پشتیبان، به صحت ۸۴/۴۷٪ رسیده است. این روش ترکیبی برای پایگاه داده انجمن تجزیه و تحلیل تصاویر ماموگرافی به کار برده شده و نتایج توسط رادیولوژیست ارزیابی شده و امیدوارکننده است. نتایج، نرخ موفقیت را برابر یا بالاتر از کارهای مرتبط دیگر نشان می‌دهد.

دو روش خودکار برای تشخیص نوع توده‌های خوش‌خیم و بدخیم سینه به کار برده شده است [12]:

(۱) قطعه‌بندی به روش رشد ناحیه‌ای که آستانه آن توسط یک

شبکه عصبی آموزش دیده به دست آمده است. (۲) قطعه‌بندی به وسیله شبکه عصبی سلولی که پارامترهایش توسط الگوریتم ژنتیک به دست آمده‌اند. محاسبات موازی و دسترسی سخت‌افزاری از مزایا و تنوع نتایج روی پایگاه داده از معایب این روش است. میزان حساسیت به تشخیص‌های درست<sup>۱</sup>، میزان حساسیت به تشخیص‌های نادرست، میزان صحت و مقدار سطح زیرمنحنی مشخصه<sup>۲</sup> به ترتیب ۸۰/۷۶٪، ۸۲/۴۰٪، ۸۱/۵۸٪، ۸۸/۱۸٪ برای روش نخست و ۸۷/۹۱٪، ۸۵/۴۰٪، ۸۶/۶۶٪، ۸۸/۱۵٪ برای روش دوم به دست آمده است.

برای طبقه‌بندی تصویرها به موارد نرمال و نامتقارن، عملکرد طبقه‌بند سامانه ایمنی مصنوعی سازگار در این مقاله بررسی شده است. جهت آزمایش کارایی الگوریتم، پایگاه داده انجمن تجزیه و تحلیل تصاویر ماموگرافی در نظر گرفته شده است. نتایج نشان می‌دهد سامانه ایمنی مصنوعی سازگار به بهترین نتایج در مقایسه با طبقه‌بندهای متداول تر می‌رسد. نرخ طبقه‌بندی در این سامانه ۸۵٪ است [13].

یک روش جدید مبتنی بر انرژی آنتروپی موجک و طبقه‌بندی رگرسیون خطی در [14] ارائه شده است. در این روش، ابتدا ناحیه مورد نظر از تصاویر ماموگرافی جدا و در مرحله دوم، آنتروپی انرژی موجک برای تصاویر قطعه‌بندی شده محاسبه و در مرحله آخر، طبقه‌بند رگرسیون خطی جهت طبقه‌بندی موارد نرمال و غیر نرمال استفاده شده است. میزان صحت در این روش، ۹۱/۸۵٪ به دست آمده است. نتایج به دست آمده از مقالات بررسی شده، نشان می‌دهد استفاده از مدل‌های مبتنی بر هوش محاسباتی در سامانه‌های تشخیص سرطان سینه امیدوارکننده بوده است.

### ۳- روش‌های محاسبات نرم مبتنی بر فازی پیشنهادی

در این بخش، روش‌های محاسبات نرم مبتنی بر فازی پیشنهادی شامل الگوریتم‌های هابیریدی فازی-تکاملی پیشنهادی جهت بهینه‌سازی پارامترها، جهت تشخیص سرطان سینه ارائه شده است.

در پژوهش حاضر، ابتدا داده‌ها پیش‌پردازش و توابع عضویت و قوانین سامانه استنتاج فازی با دانش فرد خبره طراحی و سپس پارامترهای توابع عضویت سامانه استنتاج فازی با الگوریتم‌های پیشنهادی تنظیم شده است.

<sup>1</sup> Sensitivity

<sup>2</sup> Area Under The Curve

جدول (۱) آورده شده است. توابع عضویت سامانه استنتاج فازی متناظر با تمام مجموعه‌های فازی به صورت گوسی تعریف شده است. تعدادی از قوانین اگر-آن‌گاه فازی که توسط دانش فرد خبره استخراج شده و در جدول (۲) قابل مشاهده است.

(جدول-۱): متغیرهای ورودی و خروجی سامانه استنتاج فازی  
(Table-1): Input/Output Variables of Fuzzy Inference System

متغیرهای ورودی و خروجی	عبارت‌های زبانی مربوط به هر متغیر
ویژگی‌های بافت زمینه	چربی سان، غده‌های چربی سان، غده‌های متراکم
نوع توده	انباشتگی آهکی، واضح، سوزنی شکل، نامعلوم، کج و کولگی، عدم تقارن، نرمال
اندازه توده	کوچک، متوسط، بزرگ
موقعیت توده	فوقانی داخلی، تحتانی داخلی، فوقانی خارجی، تحتانی خارجی
شدت غیر عادی بودن	بدخیم، خوش‌خیم، نرمال

(جدول-۲): تعدادی از قوانین اگر-آن‌گاه سامانه استنتاج فازی  
(Table-2): Some of the If-Then Rules of Fuzzy Inference System

۱	اگر موقعیت توده فوقانی خارجی باشد پس شدت غیر عادی بودن بدخیم است.
۲	اگر نوع توده واضح و اندازه توده کوچک باشد پس شدت غیرعادی بودن خوش خیم است.
۳	اگر نوع توده نامعلوم و اندازه توده متوسط باشد پس شدت غیرعادی بودن بدخیم است.
۴	اگر نوع توده سوزنی شکل و اندازه توده کوچک باشد پس شدت غیرعادی بودن بد خیم است.

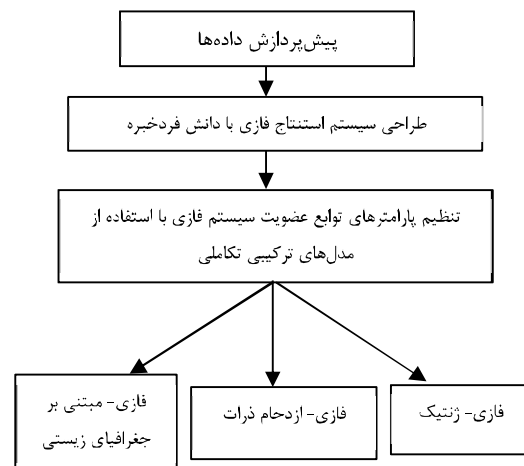
### ۳-۲- الگوریتم‌های هابیریدی مبتنی بر فازی

#### پیشنهادی جهت بهینه‌سازی پارامترها

الگوریتم‌های هابیریدی مبتنی بر فازی پیشنهادی عبارتند از: (۱) الگوریتم فازی- ژنتیک، (۲) الگوریتم فازی- بهینه‌سازی ازدحام ذرات و (۳) الگوریتم فازی- بهینه‌سازی مبتنی بر جغرافیای زیستی است. از الگوریتم‌های فراابتکاری به منظور بهینه‌سازی پارامترهای توابع عضویت در سامانه استنتاج فازی استفاده شده‌اند. مراحل اجرای سامانه‌های هابیریدی مبتنی بر فازی معرفی شده به این صورت است:

- ۱- سامانه استنتاج فازی با دانش فرد خبره طراحی شده است.
- ۲- پارامترهای لازم جهت نمایش توابع عضویت مجموعه‌های فازی ورودی و خروجی تعیین شده است؛

روندنامی مربوط به روند روش‌های محاسبات نرم مبتنی بر فازی استفاده شده در پژوهش مورد بررسی با استفاده از الگوریتم‌های هابیریدی پیشنهادی مبتنی بر فازی، در شکل (۱) معرفی شده است.



(شکل-۱): روندنامی مربوط به روش‌های محاسبات نرم

استفاده شده در پژوهش حاضر

(Figure-1): Flowchart Of Soft Computing Methods Applied In This Study

### ۳-۱- سامانه استنتاج فازی پیشنهادی

توابع عضویت و قوانین سامانه استنتاج فازی جهت تشخیص سرطان سینه با دانش فرد خبره طراحی شده است. مراحل مورد نیاز برای طراحی یک سامانه استنتاج فازی شامل موارد زیر است:

۱. جمع آوری اطلاعات، که دانشی درخصوص صورت مسأله، داده‌ها و متغیرهای ورودی و خروجی از فرد خبره کسب می‌شود؛
  ۲. متغیرهای ورودی و خروجی و تعداد عبارت‌های زبانی مربوط به هر متغیر شناسایی می‌شود؛
  ۳. مدل استنتاج فازی از نوع سوگنو یا ممدانی مشخص می‌شود؛
  ۴. قوانین با استفاده از دانش فرد خبره استخراج می‌شود؛
  ۵. توابع عضویت و پارامترهای آن طراحی می‌شود.
  ۶. سامانه طراحی شده با داده‌های واقعی ارزیابی می‌شود.
- سامانه استنتاج فازی پیشنهادی شامل چهار متغیر ورودی و یک متغیر خروجی است. متغیرهای ورودی سامانه استنتاج فازی شامل ویژگی‌های بافت زمینه، نوع توده (طبقه غیرعادی بودن توده)، اندازه توده، موقعیت توده و متغیر خروجی، شدت غیرعادی بودن تعریف شده است. متغیرهای ورودی و خروجی و عبارت‌های زبانی مربوط به هر متغیر در



۳- راه حل اولیه با مرتب کردن پارامترهای توابع عضویت هر متغیر ورودی و خروجی با لحاظ کردن محدودیت‌های لازم برای تولید جواب‌های معتبر، تولید شده است؛

۴- بردار اولیه ارزیابی شده است؛

۵- بردار اولیه به عنوان بهترین راه حل به دست آمده در نظر گرفته شده است؛

۶- پارامترهای الگوریتم پیشنهادی جهت بهینه‌سازی، تنظیم شده است؛

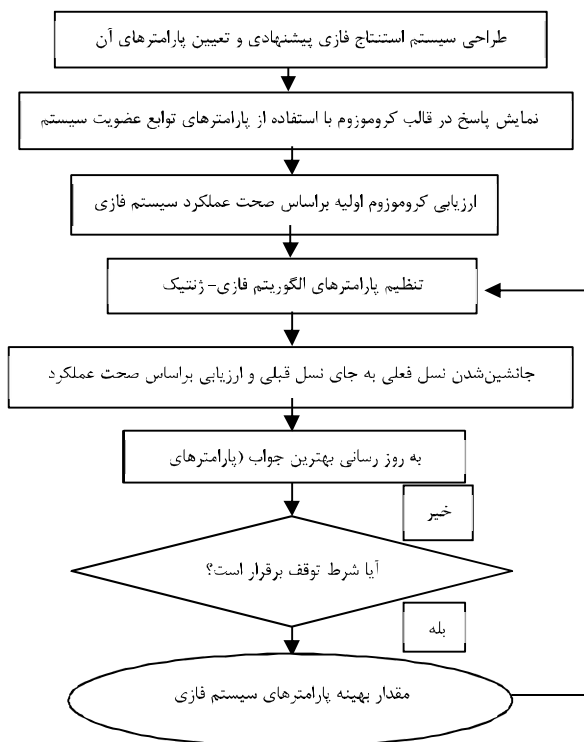
۷- الگوریتم هایبریدی فازی-تکاملی پیشنهادی اجرا شده است. در ادامه جزئیات مربوط به هر کدام از الگوریتم‌های فراابتکاری و مراحل هر الگوریتم آورده شده است.

### ۱-۲-۳- الگوریتم هایبریدی فازی- ژنتیک پیشنهادی

الگوریتم ژنتیک<sup>۱</sup> (GA)، روش بهینه‌سازی الهام گرفته از طبیعت و مبتنی بر هوش دسته‌جمعی است که نخستین بار در دهه هفتاد میلادی توسط جان هلند<sup>۲</sup>، با الهام گرفتن از قوانین تنازع بقا در طبیعت و تکامل داروین برای حل مسائل بهینه‌سازی ریاضی ابداع شد [15]. الگوریتم ژنتیک یک الگوریتم بهینه‌سازی است که از فرایند جستجو در فضای حالت بسیار بزرگ و در مواردی که بررسی تمام حالاتی که در مسئله ممکن است، اتفاق بیفتد ساده نباشد، استفاده می‌کند. این دیدگاه با ایده گرفتن از فرایند تکامل و علم ژنتیک در بحث موجودات زنده مطرح شده است. به صورت تصادفی فضای حالت را جستجو می‌کند و برای این که به این جواب بهینه برسد یک استراتژی دارد. این استراتژی بر اساس زنده ماندن شایسته‌ترین‌ها<sup>۳</sup> است. مراحل طراحی الگوریتم فازی-ژنتیک پیشنهادی به این صورت است:

- ۱- سامانه استنتاج فازی پیشنهادی طراحی می‌شود و پارامترهای آن تعیین می‌شود؛
- ۲- نمایش پاسخ یا راه حل نامزد در قالب کروموزوم با استفاده از پارامترهای توابع عضویت سامانه فازی، به عبارتی تعیین ساختار کروموزوم و ژن‌ها آن در قالب پارامترهای توابع عضویت فازی است؛
- ۳- جمعیت اولیه براساس صحت عملکرد سامانه فازی ارزیابی می‌شود؛
- ۴- پارامترهای الگوریتم فازی-ژنتیک تنظیم می‌شود؛ عملکردهای ژنتیکی شامل انتخاب، ترکیب و جهش، با توجه به پارامترهای مربوط به آن‌ها به کار گرفته می‌شوند [16-17]؛

- ۵- نسل فعلی جایگزین نسل قبلی می‌شود و ارزیابی براساس صحت عملکرد سامانه فازی صورت می‌گیرد؛
- ۶- بر اساس بهترین جواب، پارامترهای سامانه فازی به روزرسانی خواهد شد؛
- ۷- شرط توقف بررسی می‌شود. الگوریتم از مرحله چهار تکرار می‌شود تا زمانی که به شرط توقف برسد؛
- ۸- مقدار بهینه پارامترهای سامانه فازی تعیین می‌شود. روندنمای اجرای الگوریتم فازی-ژنتیک پیشنهادی در شکل (۲) قابل مشاهده است.



(شکل ۲): روندنمای اجرای الگوریتم فازی-ژنتیک پیشنهادی  
(Figure-2): Flowchart of Fuzzy-Genetic Algorithm

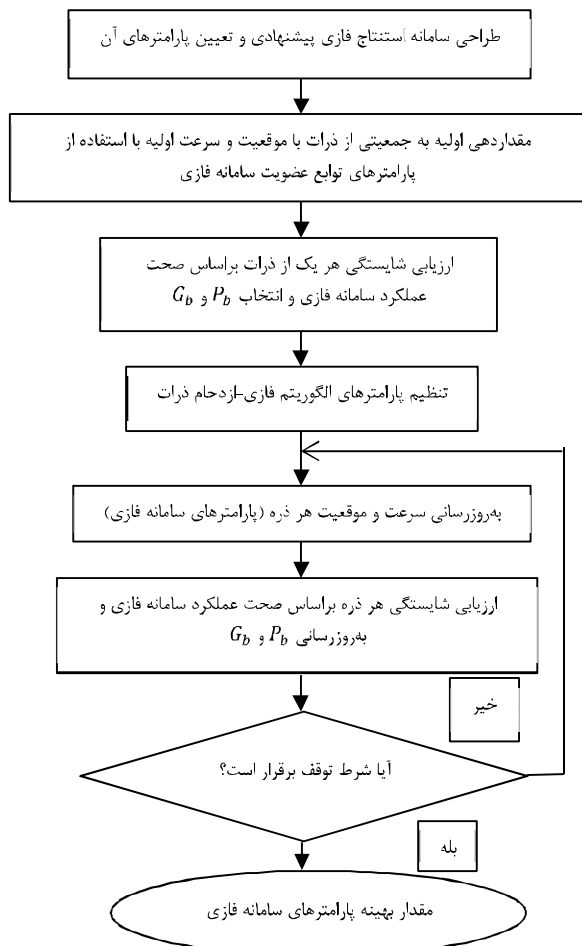
### ۲-۲-۳- الگوریتم هایبریدی فازی- ازدحام ذرات پیشنهادی

بهینه‌سازی ازدحام ذرات<sup>۴</sup> (PSO) یکی دیگر از روش‌های بهینه‌سازی الهام گرفته از طبیعت است. این روش نخستین بار در سال ۱۹۹۵ توسط دو نفر به نام‌های کندی<sup>۵</sup> و ابرهارت<sup>۶</sup> ابداع شده است [18-21]. در این الگوریتم که با الهام‌گیری از زندگی گروهی جانوران، به ویژه پرندگان و ماهی‌ها، ابداع شده است، برای حل یک مسئله بهینه‌سازی، جمعیتی از جواب‌های نامزد با استفاده از یک فرمول ساده به طور تصادفی در دامنه مسئله به حرکت در آورده شده و آن را با هدف یافتن

<sup>۴</sup> Particle Swarm Optimization  
<sup>۵</sup> Kennedy  
<sup>۶</sup> Eberhart

<sup>۱</sup> Genetic Algorithm  
<sup>۲</sup> Holland  
<sup>۳</sup> Survival of the fittest

۹- مقدار بهینه پارامترهای توابع عضویت سامانه فازی تعیین می‌شود.  
روندنامی اجرای الگوریتم فازی- ازدحام ذرات پیشنهادی در شکل (۳) قابل مشاهده است.



(شکل-۳): روندنامی الگوریتم فازی- ازدحام ذرات پیشنهادی  
(Figure-3): Flowchart of the proposed Fuzzy-Swart Intelligence Algorithm

### ۳-۲-۳- الگوریتم هایبریدی فازی-مبتنی بر جغرافیای زیستی پیشنهادی

الگوریتم بهینه‌سازی مبتنی بر جغرافیای زیستی<sup>۶</sup> (BBO)، مطالعه توزیع جغرافیایی موجودات زنده است. ردپای علم جغرافیای زیستی را می‌توان در کارهای طبیعت‌دانان قرن نوزدهم مانند والاس<sup>۷</sup> و داروین<sup>۸</sup> پیدا کرد [22-23]. در بهینه‌سازی مبتنی بر جغرافیای زیستی مسأله‌ای همراه با جمعیتی از جواب‌ها وجود دارد که می‌تواند به‌صورت برداری از اعداد صحیح نشان داده شود. هر عدد صحیح در بردار جواب،

جواب‌های بهینه سراسری مورد کاوش قرار می‌دهند. در الگوریتم بهینه‌سازی ازدحام ذرات هر یک از این جواب‌های نامزد را در اصطلاح یک ذره<sup>۱</sup> می‌نامند و هر ذره به‌عنوان مثال متناظر با یکی از پرندگان موجود در یک دسته پرنده در حال پرواز است [7]. مراحل اجرای الگوریتم فازی- ازدحام ذرات پیشنهادی در زیر آورده شده است:

- ۱- سامانه استنتاج فازی بر اساس دانش فرد خبره پیشنهادی طراحی می‌شود و پارامترهای آن تعیین می‌شود؛
- ۲- جمعیتی از ذرات با موقعیت و سرعت اولیه با استفاده از پارامترهای توابع عضویت سامانه فازی مقداردهی می‌شود. در آغاز اجرای الگوریتم، سرعت تمامی ذرات برابر با صفر در نظر گرفته می‌شود؛
- ۳- شایستگی هر یک از ذرات براساس صحت عملکرد سامانه فازی ارزیابی و  $P_b$  (بردار موقعیت بهینه‌ترین نقطه یافت‌شده توسط هر ذره) و  $G_b$  (بردار موقعیت بهینه‌ترین نقطه یافت‌شده توسط تمامی ذرات) انتخاب می‌شود؛
- ۴- پارامترهای الگوریتم فازی- ازدحام ذرات تنظیم می‌شود؛
- ۵- سرعت ذره  $i$  ام (پارامتر  $i$  ام سامانه فازی) در تکرار  $i$  ام با استفاده از معادله‌های زیر محاسبه به‌روزرسانی می‌شود:

$$V_j(i) = K * (V_j(i-1) + c_1 r_1 [P_{best,j} - X_j(i-1)] + c_2 r_2 [G_{best} - X_j(i-1)]) \quad (1)$$

$$K = 2 / |2 - \phi - \sqrt{\phi^2 - 4\phi}| \quad (2)$$

که در آن،  $c_1$  و  $c_2$  به ترتیب ضرایب یادگیری شناختی<sup>۲</sup> (یا فردی<sup>۳</sup>) و اجتماعی<sup>۴</sup> (یا گروهی<sup>۵</sup>)، و  $r_1$  و  $r_2$  دو عدد تصادفی با توزیع یکنواخت در بازه صفر تا یک هستند.  $K$  تابعی از  $c_1$  و  $c_2$  است که در آن  $\phi > 4$  و  $\phi = \phi_1 + \phi_2$  است.  $C1$  و  $C2$  نیز از معادله زیر به‌دست می‌آید:

$$C1 = K * \phi_1, C2 = K * \phi_2 \quad (3)$$

- ۶- موقعیت ذره  $i$  ام (پارامتر  $i$  ام سیستم فازی) در تکرار  $i$  ام با استفاده از معادله‌های زیر محاسبه به‌روزرسانی می‌شود:

$$X_j(i) = X_j(i-1) + V_j(i); j = 1, 2, \dots, N \quad (4)$$

- ۷- شایستگی هر ذره براساس صحت عملکرد سامانه فازی ارزیابی و  $P_b$  و  $G_b$  به‌روزرسانی می‌شود؛
- ۸- شرط توقف بررسی و الگوریتم از مرحله پنج تکرار می‌شود تا زمانی که به شرط توقف برسد.

<sup>1</sup> Particle

<sup>2</sup> Cognitive

<sup>3</sup> Individual

<sup>4</sup> Social

<sup>5</sup> Group

<sup>6</sup> Biogeography-Based Optimization

<sup>7</sup> Wallace

<sup>8</sup> Darwin

جمعیت (پارامترهای سامانه فازی) قبل می‌شود؛  
 ۱۴- در صورت برآورده نشدن شرایط خاتمه به مرحله پنج باز می‌گردد؛  
 ۱۵- مقدار بهینه پارامترهای سامانه فازی تعیین می‌شود.  
 مراحل اجرای الگوریتم فازی- مبتنی بر جغرافیای زیستی پیشنهادی در شکل (۴) قابل مشاهده است.



(شکل-۴): روندنمای الگوریتم فازی- بهینه سازی مبتنی بر جغرافیای زیستی پیشنهادی

(Figure-4): Flowchart of Fuzzy-Biogeography optimization Model

## ۴- تحلیل نتایج آزمایشی

در این بخش، مقایسه و تحلیل نتایج به دست آمده از روش های هوش محاسباتی پیشنهادی در این پژوهش ارایه می‌شود. نتایج هر روش پیشنهادی با معیارهای ارزیابی یکسان بر روی بانک اطلاعاتی مشترک مقایسه شده است؛ همچنین نتایج با یافته های دیگر پژوهش گران مقایسه و دلایل تفاوت و مزایای سامانه های پیشنهادی بیان می‌شود.

یک شاخص مطلوبیت متغیر<sup>۱</sup> (SIV)، در نظر گرفته می‌شود. همچنین روش هایی برای ارزیابی خوبی جواب وجود دارد. جواب های خوب دارای شاخص مطلوبیت زیستگاه<sup>۲</sup> (HSI) بالا و پاسخ های ضعیف، زیستگاه با شاخص مطلوبیت پایین در نظر گرفته می‌شود [24].

برای اجرای الگوریتم فازی-مبتنی بر جغرافیای زیستی پیشنهادی مراحل زیر به ترتیب انجام می‌شود:

۱- سامانه استنتاج فازی پیشنهادی طراحی و پارامترهای آن تعیین می‌شود؛

۲- مجموعه ای از محل های سکونت تصادفی (جمعیت اولیه) توسط پارامترهای توابع عضویت سامانه فازی تولید می‌شود؛

۳- جمعیت اولیه براساس صحت عملکرد سامانه فازی ارزیابی می‌شود؛

۴- پارامترهای الگوریتم فازی- مبتنی بر جغرافیای زیستی تنظیم می‌شود؛

۵-  $\mu_i$  و  $\lambda_i$  برای هر محل سکونت (هر پارامتر سامانه فازی) تعیین می‌شود. مقدار  $\mu$  متناسب با تعداد محل های سکونت به صورت خطی از یک به صفر افزایش پیدا می‌کند و مقدار  $\lambda$  نیز به صورت خطی از یک به صفر کاهش پیدا می‌کند؛

۶- به ازای هر محل سکونت  $i$ ، (هر پارامتر سامانه فازی) مراحل ۷ تا ۱۱ تکرار می‌شود؛

۷- به ازای هر  $SIV_k$  در محل سکونت  $i$ ، مراحل ۸ تا ۱۱ انجام می‌شود؛

۸- با احتمال  $\lambda_i$  در محل سکونت  $i$  و  $SIV_k$  تغییرات اعمال می‌شود؛

۹- مبدأ مهاجرت  $x_{jk}$  با استفاده از مقادیر  $\mu$  به صورت تصادفی تعیین می‌شود؛

۱۰- مهاجرت از محل سکونت  $j$  به  $SIV_k$  به محل سکونت  $i$  و  $SIV_k$  طبق فرمول زیر انجام می‌شود:

$$x_{ik}^{new} = (1 - \alpha)x_{ik} + \alpha_k x_{jk} \quad (5)$$

۱۱- با احتمال معین بر روی محل سکونت  $i$  و  $SIV_k$  که از مرحله قبل به دست آمده، جهش اعمال می‌شود؛

۱۲- مجموعه پاسخ های جدید به دست آمده (پارامترهای سامانه فازی)، براساس صحت عملکرد سامانه فازی ارزیابی می‌شود؛

۱۳- جمعیت (پارامترهای سامانه فازی) فعلی جایگزین

<sup>1</sup> Suitability Index Variable

<sup>2</sup> Habitat Suitability Index



#### ۴-۱- پایگاه داده‌ها

برای ارزیابی مدل‌های پیشنهادی از اطلاعات ۳۲۲ نمونه تصویر ماموگرافی واقعی بیماران که از پایگاه داده انجمن تجزیه و تحلیل تصاویر ماموگرافی گرفته شده و به صورت برخط برای اهداف علمی در دسترس است، استفاده شده است [25]. هر تصویر شامل اطلاعات ویژگی‌های بافت زمینه، طبقه غیرعادی بودن توده، شعاع از مرکز توده، مختصات X و Y از مرکز توده و شدت غیرعادی بودن توده است که توسط رادیولوژیست تعیین شده است. از این ۳۲۲ نمونه، ۶۲ نمونه خوش خیم، ۵۱ نمونه بدخیم و ۲۰۹ نمونه نرمال است.

در پژوهش حاضر، در هر تصویر موقعیت توده با استفاده از مختصات X و Y از مرکز توده و شعاع توده که توسط رادیولوژیست تعیین شده به صورت دستی در یکی از نواحی چهارگانه سینه تخمین زده شده است. هر پستان شامل چهار ناحیه است که توسط چهار عبارت زبانی-فوقانی-داخلی<sup>۱</sup>، تحتانی-داخلی<sup>۲</sup>، فوقانی-خارجی<sup>۳</sup> و تحتانی-خارجی<sup>۴</sup> تعریف می‌شود. همچنین اندازه توده با استفاده از شعاع از مرکز توده تعیین شده است.

#### ۴-۲- پیش پردازش داده‌ها

عملیات پیش پردازش داده شامل حذف نوفه و نرمال سازی داده‌ها است. ابتدا داده‌ها نرمالایز شده، سپس با استفاده از روش خوشه‌بندی مبتنی بر شناسایی اوتلایر<sup>۶</sup> عملیات حذف نوفه انجام شده است. مراحل روش خوشه‌بندی مبتنی بر شناسایی اوتلایر جهت حذف تعدادی از داده‌های نوفه‌ای به صورت زیر است [26]:

۱- تعداد خوشه‌ها برابر با تعداد طبقه‌های خروجی در نظر گرفته می‌شود. هر خوشه شامل کلیه متغیرهای ورودی مربوط به هر طبقه خروجی است. بنابراین سه خوشه بدخیم، خوش خیم و نرمال وجود دارد؛

۲- برای هر کدام از متغیرهای ورودی در هر خوشه میانگین حساب می‌شود؛

۳- برای هر کدام از متغیرهای ورودی در هر خوشه با استفاده از فرمول زیر کواریانس به دست می‌آید:

$$S_i = \frac{1}{n_i - 1} \sum_{X \in m_i} (X_i - m_i) (X_i - m_i)^T \quad (6)$$

<sup>1</sup> Location Of Abnormality

<sup>2</sup> Upper Inner Quadran

<sup>3</sup> Lower Inner Quadran

<sup>4</sup> Upper Outer Quadran

<sup>5</sup> Lower Outer Quadran

<sup>6</sup> Outlier Detection

که در آن  $X_i$  خوشه  $i$  ام،  $m_i$  میانگین به ازای متغیرهای ورودی مربوط به هر خوشه و  $n_i$  سایز خوشه  $i$  ام است؛  
۴- شعاع نرمال شده آماری داده‌ها، T مربع<sup>۷</sup>، با استفاده از فرمول زیر به دست می‌آید:

$$T_i^2 = (X_i - m_i) S_i^{-1} (X_i - m_i)^T \quad (7)$$

که در آن  $X_i$  خوشه  $i$  ام،  $m_i$  میانگین به ازای متغیرهای ورودی مربوط به هر خوشه و  $S_i^{-1}$  معکوس کواریانس به دست آمده برای متغیرهای ورودی در هر خوشه توسط فرمول مرحله قبل است؛

۵- مقدار  $\alpha$  با توجه به ابعاد مسأله و به نحوی که ماهیت مسأله از دست نرود، تعیین می‌شود. هر چه  $\alpha$  بیشتر باشد، تعداد داده‌های حذفی بیشتر خواهد بود.  $\alpha$  برابر با پنج در نظر گرفته شده است؛

۶- مقادیر به دست آمده برای شعاع توسط فرمول T مربع به ازای هر داده با مقدار  $\alpha$  مقایسه می‌شود و داده‌هایی که مقدارشان بیشتر از  $\alpha$  باشد، به عنوان اوتلایر شناسایی می‌شوند و حذف خواهند شد. با توجه به مقدار تعیین شده برای  $\alpha$ ، ۲۷ مورد داده حذف شده است.

#### ۴-۳- معیارهای ارزیابی

برای سنجش کارایی سامانه‌های پیشنهادی در این پژوهش و مقایسه با مدل‌های دیگر از تحلیل منحنی مشخصه عملکرد سامانه استفاده شده است. منحنی مشخصه عملکرد سامانه معیاری مناسب است، چون موازنه‌ای بین تشخیص‌های درست و نادرست سامانه یا به عبارتی سود و هزینه سامانه را به طور هم‌زمان به تصویر می‌کشد.

یک سامانه طبقه‌بندی، برچسب طبقه یک نمونه را پیش‌بینی می‌کند. هر نمونه یک برچسب واقعی و یک برچسب پیش‌بینی دارد. در بحث تحلیل منحنی مشخصه عملکرد سامانه یک آستانه تصمیم تعیین می‌شود، مقداری که کمتر از آن هستند به عنوان منفی و مقداری که بزرگ‌تر از آستانه هستند به عنوان مثبت در نظر گرفته می‌شوند. در تحلیل منحنی مشخصه عملکرد، معیارهای زیر محاسبه می‌شود: آن‌هایی که مثبت بودند و درست مثبت تشخیص داده شده‌اند<sup>۸</sup> (TP)، مثبت بودند و منفی تشخیص داده شده‌اند<sup>۹</sup> (FN)، منفی بودند و درست منفی تشخیص داده شده‌اند<sup>۱۰</sup> (TN).

<sup>7</sup> T- square

<sup>8</sup> True positive

<sup>9</sup> False Negative

<sup>10</sup> True Negative

(TN) و آن‌هایی که منفی بودند و مثبت تشخیص داده شده‌اند<sup>۱</sup> (FP). ماتریس کانفیوژن<sup>۲</sup> براساس برچسب واقعی و برچسب پیش‌بینی شده نمونه‌ها و معیارهای TP, FN, TN و FP تشکیل می‌شود.

فرمول میزان حساسیت به تشخیص‌های نادرست:

$$tp\ rate = Specificity = \frac{TN}{FP+} = 1 - fp \quad (۸)$$

فرمول میزان صحت عبارتست از:

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (۹)$$

چنان چه تعداد طبقات یا دسته‌ها برابر  $n$  باشد، ماتریس کانفیوژن، ماتریس  $n \times n$  حاوی  $n$  طبقه‌بندی درست و  $n^2 - n$  خطای احتمالی است. یک راه‌کار جهت پرداختن به مسأله‌ای  $n$  طبقه‌ای، ساخت  $n$  منحنی مشخصه عملکرد سامانه مختلف برای هر کدام از طبقات یا دسته‌ها است. این را فرمول مرجع طبقه می‌نامند. به عبارت مشخص، اگر  $C$ ، مجموعه تمام طبقات باشد، منحنی مشخصه عملکرد سامانه  $i$ ، عملکرد طبقه‌بندی را با استفاده از طبقه  $c_i$  به عنوان طبقه مثبت نشان داده و تمام طبقات دیگر را به عنوان طبقه منفی مشخص می‌کند. براساس تعاریف بالا، در این پژوهش، طبقه خوش‌خیم و بدخیم به عنوان طبقه مثبت و طبقه نرمال به عنوان طبقه منفی در نظر گرفته شده است.

یکی دیگر از معیارهای قابل اطمینان برای گزارش دقت عملکرد طبقه، مساحت زیر منحنی مشخصه عملکرد سامانه است. این روش با پدید آوردن منحنی مشخصه عملکرد سامانه مرجع برای هر طبقه و متعاقباً اندازه‌گیری مساحت زیرمنحنی، مساحت زیرمنحنی‌های مسائل چند طبقه را به دست آورده، سپس مساحت زیر منحنی‌های وزن‌دهی شده بر حسب شیوع طبقه مرجع در داده‌ها را با هم جمع می‌بندد. به عبارت دقیق‌تر، تعریف این روش از مساحت زیر منحنی مجموع بدین صورت است:

$$AUC_{total} = \sum_{c_i \in C} AUC(c_i) \cdot p(c_i) \quad (۱۰)$$

که در این رابطه  $AUC(c_i)$ ، مساحت زیرمنحنی مشخصه عملکرد سامانه مرجع برای طبقه  $c_i$  است [27-30].

#### ۴-۴- عملکرد طبقه‌بندها

روش‌های هوش محاسباتی پیشنهادی در پژوهش انجام شده،

<sup>۱</sup> False Positive

<sup>۲</sup> Confusion Matrix

در نرم‌افزار متلب طراحی و پیاده‌سازی شده‌اند. برای مقایسه روش‌های پیشنهادی در پژوهش‌های انجام شده جهت طبقه‌بندی از پارامترهای ارزیابی یکسان استفاده شده است. منحنی مشخصه عملکرد سامانه برای همه روش‌ها رسم شده است و جهت اعتبارسنجی سامانه‌ها، داده‌ها توسط روش اعتبارسنجی تقاطعی ده بخشی به دو قسمت آموزش و آزمون تقسیم‌بندی شده‌اند. میانگین مقدار سطح زیرمنحنی مشخصه عملکرد سامانه بر روی داده‌های آموزش و آزمون با ده تکرار از اجرای روش اعتبارسنجی تقاطعی ده بخشی با انتخاب تصادفی برای هر بخش به دست آمده است. ماتریس کانفیوژن نیز برای همه روش‌ها جهت مقایسه تشخیص‌های درست و نادرست سامانه به‌ازای خروجی‌ها آورده شده است.

برای ارزیابی مدل‌های پیشنهادی از مقایسه اطلاعات تصاویر ماموگرافی واقعی بیماران که توسط رادیولوژیست زیرنویسی شده‌اند با نتایج سامانه‌ها بهره‌گیری شده است. پارامترهای عملکردهای روش‌های بهینه‌سازی به صورت اکتشافی با آزمایش مقادیر منطقی در محدوده مجاز آن‌ها جهت تنظیم پارامترها و آموزش قوانین انتخاب شده است.

#### ۴-۴-۱- عملکرد مدل طبقه‌بند فازی

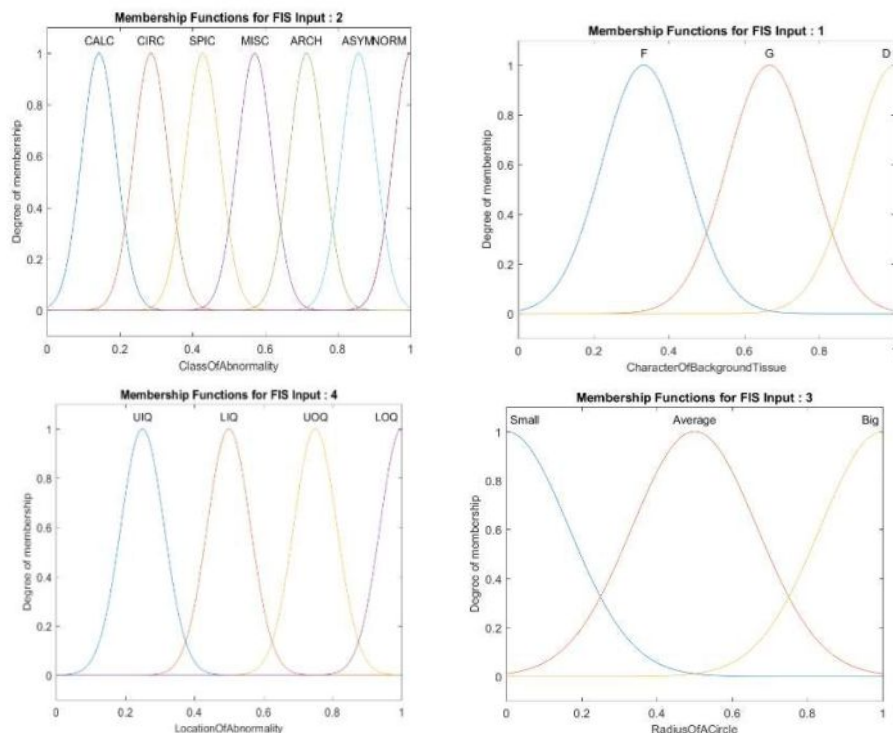
سامانه استنتاج فازی در حالت سوگنو طراحی شده است. پارامترهای تنظیم شده برای طراحی سامانه فازی در جدول (۳) آورده شده است. توابع عضویت متناظر با هر عبارت زبانی ورودی، قبل از فرایند تنظیم پارامترهای توابع عضویت سامانه فازی سوگنو در شکل (۵) قابل مشاهده است.

(جدول-۳): پارامترهای سیستم فازی

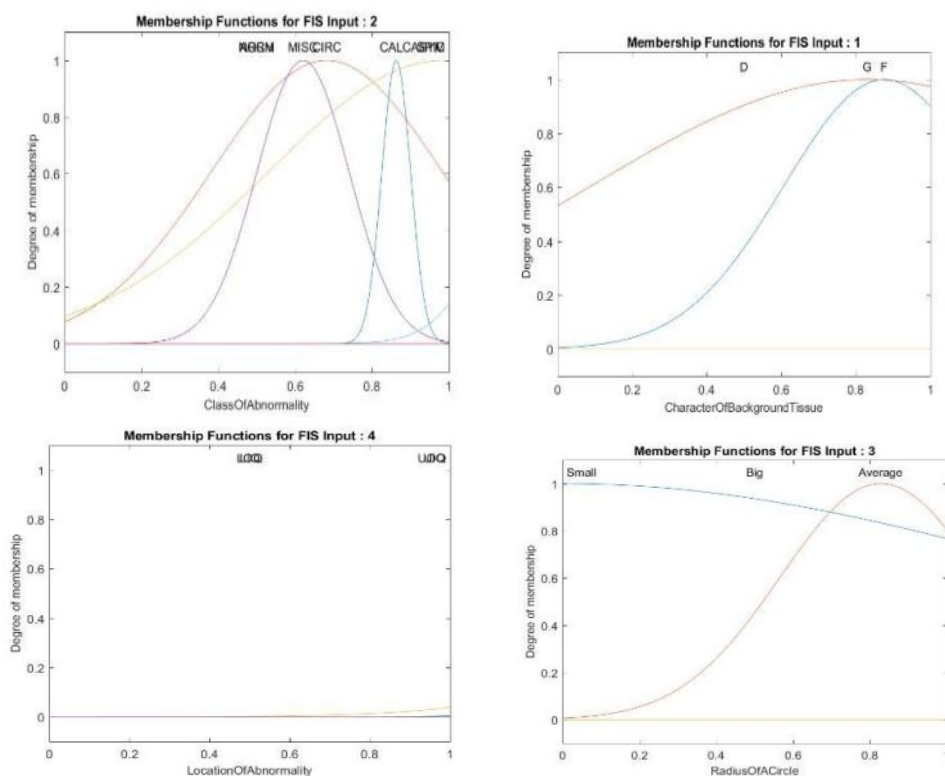
(Table-3): Fuzzy System Parameters

FisName	Breast Cancer Diagnosis
FisType	Sugeno
AndMethod	Min
OrMethod	Max
ImpMethod	Prod
AggMethod	Max
DefuzzMethod	Wtaver

شکل (۵)، شکل توابع عضویت ورودی‌های سامانه استنتاج فازی سوگنو جهت تشخیص سرطان سینه را نشان می‌دهد. نوع خروجی در سامانه استنتاج فازی سوگنو عدد ثابت است. پارامترها و عبارت‌های زبانی مرتبط با هر متغیر از اطلاعات استخراج شده از تصاویر توسط رادیولوژیست تعیین شده است.



(شکل-۵): توابع عضویت سامانه استنتاج فازی سوگنو قبل از فرایند تنظیم  
(Figure-5): Membership Function of Sugeno Fuzzy Inference System before Tuning Process



(شکل-۶): توابع عضویت سامانه فازی سوگنو- ژنتیک پس از فرایند تنظیم پارامترها  
(Figure-6): Membership Function of Fuzzy-Genetic Inference System after Tuning Process

#### ۴-۴-۲- عملکرد طبقه‌بند فازی- ژنتیک

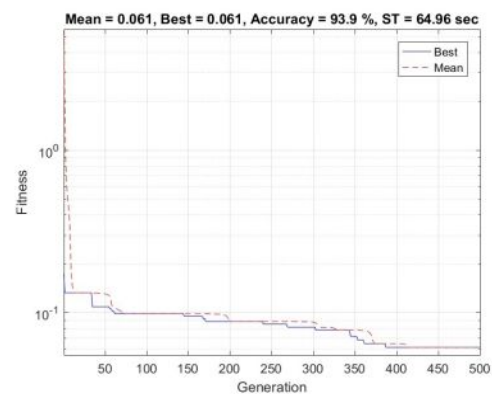
پارامترهای توابع عضویت سامانه استنتاج فازی سوگنو توسط الگوریتم ژنتیک تنظیم شده است و در جدول (۴) آورده شده است. شکل (۶)، شکل توابع عضویت سامانه فازی سوگنو-ژنتیک را پس از فرایند تنظیم پارامترها نشان می‌دهد. در شکل (۷)، بهترین و میانگین شایستگی الگوریتم ژنتیک براساس معیار سطح زیر منحنی ROC جهت بهینه‌سازی پارامترهای سامانه فازی سوگنو قابل مشاهده است.

(جدول-۴): پارامترهای الگوریتم ژنتیک جهت تنظیم پارامترهای

توابع عضویت

(Table-4): Parameters of Genetic Algorithm for Tuning Membership Functions

نحوه نمایش کروموزوم	اعداد حقیقی (پیوسته)
تعداد ژن‌ها	۳۷
تعداد جمعیت	۱۰۰
تعداد نسل‌ها	۵۰۰
نرخ ترکیب	۰/۸
نرخ جهش	۰/۳
شرط توقف	پایان نسل‌ها
استراتژی انتخاب	استدلال نخیه‌گرایی



(شکل-۷): بهترین و میانگین شایستگی الگوریتم ژنتیک جهت

بهینه‌سازی پارامترهای سامانه فازی سوگنو

(Figure-7): The best and average Fitness of the GA for Optimization of Sugeno Fuzzy Inference Model

#### ساختار کروموزوم

ساختار کروموزوم به شیوه نمایش اعداد حقیقی پیوسته در سامانه فازی ژنتیک طراحی شده است. ژن‌های کروموزوم پارامترهای توابع عضویت تعریف‌شده به‌ازای هر عبارت زبانی برای هر متغیر ورودی و خروجی به‌دست می‌آید. توابع عضویت سامانه استنتاج فازی سوگنو، گوسین است و شامل دو پارامتر میانگین و انحراف معیار است. متغیر ورودی نخست، دوم، سوم و چهارم به‌ترتیب شامل ۳، ۷، ۳ و ۴ عبارت زبانی است. تعداد کل ژن‌ها که پارامترهای توابع عضویت به‌ازای عبارت‌های زبانی مختلف شامل ۳۴ ژن است. در سامانه استنتاج سوگنو خروجی عدد ثابت در نظر گرفته می‌شود؛ پس به‌ازای سه

عبارت زبانی برای متغیر خروجی، سه ژن تعریف شده است. بنابراین تعداد ژن‌های به‌ازای متغیرهای ورودی و خروجی در ۳۷ در نظر گرفته شده است.

#### تابع برازندگی

تابع برازندگی در الگوریتم فازی-ژنتیک، صحت عملکرد (Accuracy)، بر اساس فرمول شماره ۹، در سامانه استنتاج فازی بر اساس پارامترهای تعیین‌شده در ساختار کروموزوم در نظر گرفته شده است.

#### ۴-۴-۳- عملکرد طبقه‌بند فازی- ازدحام ذرات

پارامترهای توابع عضویت سیستم استنتاج فازی سوگنو توسط الگوریتم بهینه‌سازی ازدحام ذرات تنظیم و در جدول (۵) آورده شده است. شکل (۸)، شکل توابع عضویت سامانه فازی سوگنو- ازدحام ذرات را پس از فرایند تنظیم پارامترها نشان می‌دهد. در شکل (۹)، بهترین و میانگین شایستگی الگوریتم ازدحام ذرات براساس معیار سطح زیرمنحنی ROC جهت بهینه‌سازی پارامترهای سامانه فازی سوگنو قابل مشاهده است.

#### ساختار ذرات

در پژوهش حاصل، ساختار ذرات، به‌صورت برداری از پارامترهای توابع عضویت سامانه فازی (میانگین و انحراف معیار تابع عضویت گوسی) متناسب با هر متغیر زبانی (متناظر با سامانه فازی) تعریف شده است.

#### تابع برازندگی

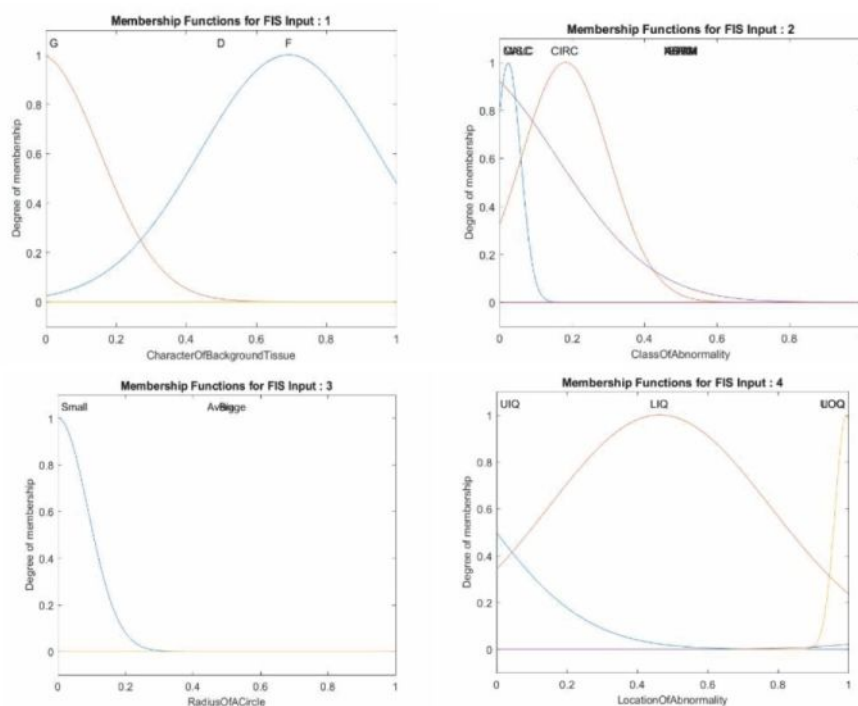
تابع برازندگی در الگوریتم فازی-ازدحام ذرات، صحت عملکرد (Accuracy)، بر اساس فرمول شماره ۹، در سامانه استنتاج فازی بر اساس پارامترهای تعیین‌شده در بردار ذرات در نظر گرفته شده است.

#### ساختار ذرات

در پژوهش حاصل، ساختار ذرات، به‌صورت برداری از پارامترهای توابع عضویت سامانه فازی (میانگین و انحراف معیار تابع عضویت گوسی) متناسب با هر متغیر زبانی (متناظر با سامانه فازی) تعریف شده است.

#### تابع برازندگی

تابع برازندگی در الگوریتم فازی-ازدحام ذرات، صحت عملکرد (Accuracy)، بر اساس فرمول شماره ۹، در سامانه استنتاج فازی بر اساس پارامترهای تعیین‌شده در بردار ذرات در نظر گرفته شده است.



(شکل-۸): توابع عضویت سامانه فازی سوگنو- ازدحام ذرات پس از فرایند تنظیم پارامترها  
(Figure-8): Membership Functions of Fuzzy-PSO Algorithm after Tuning Parameters

### ساختار ذرات

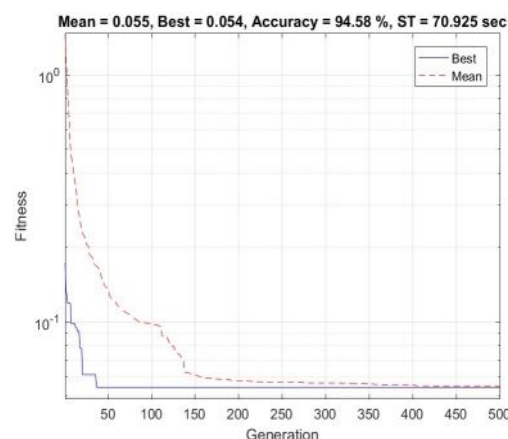
در پژوهش حاصل، ساختار ذرات، به صورت برداری از پارامترهای توابع عضویت سامانه فازی (میانگین و انحراف معیار تابع عضویت گوسی) متناسب با هر متغیر زبانی (متناظر با سامانه فازی) تعریف شده است.

### تابع برازندگی

تابع برازندگی در الگوریتم فازی-ازدحام ذرات، صحت عملکرد (Accuracy)، بر اساس فرمول شماره ۹، در سامانه استنتاج فازی بر اساس پارامترهای تعیین شده در بردار ذرات در نظر گرفته شده است.

### ۴-۴-۴ عملکرد طبقه‌بند فازی- مبتنی بر جغرافیای زیستی

پارامترهای توابع عضویت سامانه استنتاج فازی سوگنو توسط الگوریتم مبتنی بر جغرافیای زیستی تنظیم شده است. تعداد جمعیت، تعداد نسل و ضریب تأثیر در مهاجرت به عنوان پارامترهای الگوریتم در جدول (۶) آورده شده است. شکل (۱۰)، شکل توابع عضویت سامانه فازی سوگنو- مبتنی بر جغرافیای زیستی را پس از فرایند تنظیم پارامترها نشان می‌دهد. در شکل (۱۱)، بهترین و میانگین شایستگی الگوریتم



(شکل-۹): بهترین و میانگین اجرای الگوریتم ازدحام ذرات

جهت بهینه‌سازی پارامترهای سیستم فازی سوگنو

(Figure-9): The Best and Average Fitness of the PSO for Optimization of Sugeno Fuzzy Inference Model

(جدول-۵): پارامترهای الگوریتم بهینه‌سازی ازدحام ذرات

جهت تنظیم پارامترهای توابع عضویت

(Table- 5): Parameters of Particle Swarm Intelligence Algorithm for Tuning Membership Functions

۱۰۰	تعداد جمعیت
۰۰	تعداد نسل‌ها
۰/۷۲	اینرسی وزن
۱/۴۹	ضریب یادگیری شخصی ذرات
۱/۴۹	ضریب یادگیری جمعی ذرات

مبتنی بر جغرافیای زیستی براساس معیار سطح زیر منحنی ROC جهت بهینه‌سازی پارامترهای سامانه فازی سوگنو قابل مشاهده است.

### ساختار زیستگاه

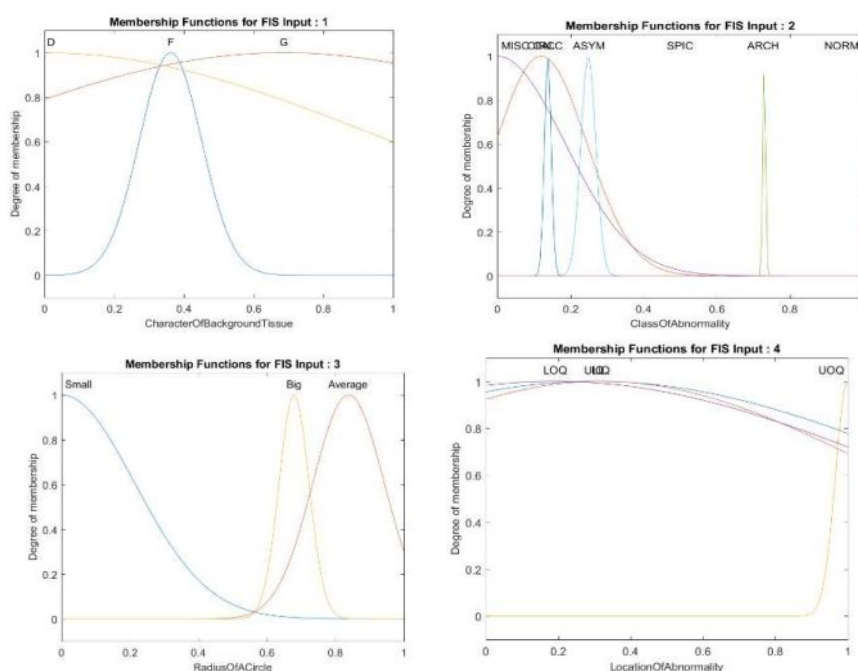
ساختار زیستگاه به صورت برداری از پارامترهای توابع عضویت سامانه فازی (میانگین و انحراف معیار) متناسب با هر متغیر زبانی (در سامانه فازی) تعریف شده است. هر عدد صحیح در بردار جواب، یک شاخص مطلوبیت متغیر در نظر گرفته می‌شود.

### تابع برازندگی

تابع برازندگی در الگوریتم فازی-جغرافیای زیستی، صحت عملکرد (Accuracy)، بر اساس فرمول شماره ۹، بر اساس پارامترهای تعیین‌شده در توابع عضویت سامانه فازی در نظر گرفته شده است.

(جدول-۶): پارامترهای الگوریتم بهینه‌سازی مبتنی بر جغرافیای زیستی جهت تنظیم پارامترهای توابع عضویت (Table-6): Parameters of Biogeography Algorithm for Tuning Membership Functions

تعداد جمعیت	۵۰
تعداد نسل‌ها	۵۰۰
ضریب تأثیر در مهاجرت	۰/۷



(شکل-۱۰): سیستم فازی سوگنو- مبتنی بر جغرافیای زیستی پس از فرایند تنظیم پارامترها (Figure-10): Membership Functions of Fuzzy-Biogeography Algorithm after Tuning Parameters

### ۴-۵- مقایسه کارایی روش‌های پیشنهادی در

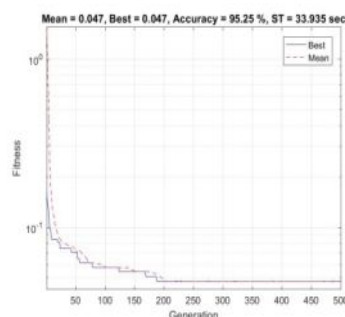
#### پژوهش حاضر

مقایسه تحلیل نتایج روش‌های پیشنهادی با معیارهای منحنی مشخصه عملکرد سامانه در فرایند بهینه‌سازی پارامترهای توابع عضویت سامانه فازی، روی کل داده‌ها به تعداد ۲۹۵ نمونه در جدول (۷) قابل مشاهده است.

نمودارهای هم‌زمان منحنی مشخصه عملکرد سامانه مربوط به روش‌های هایبریدی پیشنهادی جهت بهینه‌سازی پارامترهای توابع عضویت سامانه استنتاج فازی سوگنو پیشنهادی در شکل (۱۲) مصور شده است.

(شکل-۱۱): بهترین و میانگین اجرای الگوریتم مبتنی بر جغرافیای زیستی جهت بهینه‌سازی پارامترهای سامانه فازی سوگنو

(Figure-11): The Best and Average Fitness of the Biogeography Algorithm for Optimization of Sugeno Fuzzy Inference Model





عملکرد بهتری نسبت به روش‌های پیشنهادی دیگر جهت تشخیص سرطان سینه است. زیرا تعداد تشخیص‌های درست در این روش نسبت به روش‌های دیگر بیشتر است.

جدول (۸)، مقایسه تحلیل نتایج ماتریس کانفیوژن روش‌های هایبریدی پیشنهادی در فرایند بهینه‌سازی پارامترهای توابع عضویت سامانه فازی سوگنو را نشان می‌دهد. نتایج به‌دست‌آمده از ماتریس کانفیوژن نشان می‌دهد روش هایبریدی فازی-مبتنی بر جغرافیای زیستی، پیشنهادی دارای

(جدول-۷): مقایسه نتایج روش‌های پیشنهادی با معیارهای منحنی مشخصه عملکرد سیستم در فرایند بهینه‌سازی پارامترهای توابع عضویت سیستم فازی

(Table-7): Comparison of the Overall Results of Proposed Method using ROC Curve Analysis for Tuning Membership Function

روش‌ها	Accuracy%	AUC%	Specificity%	CI%	SE	پیچیدگی زمانی (ثانیه)
فازی	۸۲/۷۱	۶۸/۹۳	۵۵/۱۵	[۶۰,۳۳ ۷۷,۵۴]	۰/۰۴	-
فازی-ژنتیک	۹۳/۹۰	۸۸/۹۸	۸۴/۰۷	[۷۸,۹۷ ۹۸,۹۹]	۰/۰۵	۷۴/۹۶
فازی-ازدحام ذرات	۹۴/۵۸	۹۰/۲۳	۸۵/۸۸	[۸۰,۱۴ ۱۰۰]	۰/۰۵	۷۰/۹۲۵
فازی-مبتنی بر جغرافیای زیستی	۹۵/۲۵	۹۱/۴۳	۸۷/۶۱	[۸۱,۲۴ ۱۰۰]	۰/۰۵	۳۳/۹۳۵

از روش اعتبارسنجی تقاطعی ده‌بخشی به‌منظور اعتبارسنجی در مرحله آموزش و آزمون روش‌های پیشنهادی استفاده شده است. میانگین نتایج آموزش و آزمون حاصل از اجرای آن برای روش‌های هایبریدی پیشنهادی پس از بهینه‌سازی پارامترهای توابع عضویت سامانه فازی سوگنو در جدول (۹) نشان داده شده است.

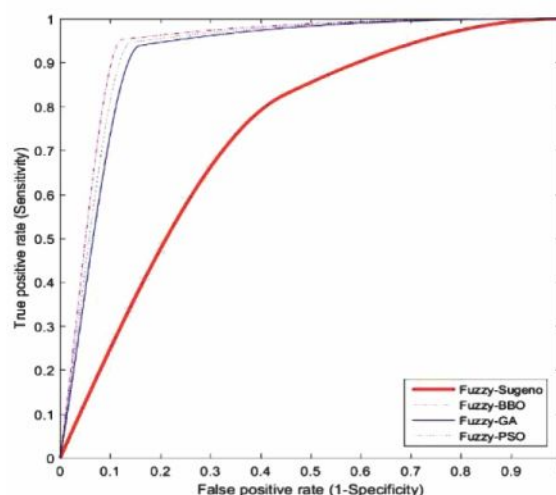
همان‌طور که در جداول (۷ و ۸) قابل مشاهده است و با توجه به مقدار سطح زیرنمودار در شکل (۱۲)، نتایج حاصل از اجرای معیارهای منحنی مشخصه عملکرد سامانه و اعتبارسنجی تقاطعی ده‌بخشی نشان می‌دهد که روش هایبریدی فازی-مبتنی بر جغرافیای زیستی پیشنهادی از عملکرد بهتر نسبت به روش‌های هایبریدی پیشنهادی دیگر جهت تشخیص سرطان سینه در بهینه‌سازی پارامترهای توابع عضویت سامانه استنتاج فازی پیشنهادی برخوردار است. گفتنی است که هر یک از مدل‌های پیشنهادی به نحوی سامانه را بهبود بخشیده‌اند.

روش‌های ارائه‌شده در تشخیص سرطان سینه بر روی پایگاه داده انجمن تجزیه و تحلیل تصاویر ماموگرافی توسط دیگران با روش‌های برتر هوش محاسباتی پیشنهادی در این پژوهش در جدول (۱۰) مقایسه شده است. براساس مطالعات انجام‌شده بر روی پایگاه داده انجمن تجزیه و تحلیل تصاویر ماموگرافی، روش‌های بررسی‌شده از ماشین بردارهای پشتیبان، روش‌های رگرسیون، یادگیری ماشین و سامانه ایمنی مصنوعی جهت طبقه‌بندی توده‌های سینه استفاده

(جدول-۸): مقایسه نتایج ماتریس کانفیوژن روش‌های هایبریدی پیشنهادی در فرایند بهینه‌سازی پارامترهای توابع عضویت سیستم فازی

(Table-8): Comparison of the Results of Proposed hybrid Method using Confusion Matrix Tuning Membership Function

روش‌ها	TP	TN	FP	FN
فازی	۳۵	۲۰۹	۵۱	.
فازی-ژنتیک	۶۸	۲۰۹	۱۸	.
فازی-ازدحام ذرات	۷۰	۲۰۹	۱۶	۰



(شکل-۱۲): نمودارهای منحنی مشخصه عملکرد مربوط به روش‌های هایبریدی پیشنهادی جهت بهینه‌سازی پارامترهای توابع عضویت سیستم استنتاج فازی سوگنو پیشنهادی

(Figure-12): The ROC Curve of the Hybrid Proposed Models for Optimization of MFs of fuzzy System

کرده‌اند. متغیرهای ورودی و خروجی و تعداد داده‌های استفاده‌شده پس از مرحله پیش‌پردازش در این روش‌ها مشخص شده است. در بیشتر روش‌ها ارزیابی براساس میزان صحت انجام شده است.

(جدول-۹): میانگین نتایج آموزش و آزمون روش اعتبار سنجی تقاطعی ده‌بخشی برای روش‌های هایبریدی پیشنهادی پس از بهینه‌سازی

پارامترهای توابع عضویت سامانه فازی سوگنو

(Table-9): Average Result of the Proposed hybrid MF Optimization Algorithms using 10-Fold Cross Validation Techniques on Training and Testing Dataset

SE		CI%		Specificity%		AUC%		Accuracy%		روش‌ها
Test	Train	Test	Train	Test	Train	Test	Train	Test	Train	
۰/۱۴	۰/۰۵	[۶۱,۳۶ ۹۹,۹۱]	[۷۸,۴۲ ۹۹,۳۲]	۸۴/۱۳	۸۴/۰۶	۸۹.۰۲	۸۸/۹۸	۹۳/۹۱	۹۳.۸۹	فازی- ژنتیک
۰/۱۳	۰/۰۵	[۶۳,۳۴ ۹۹,۹۸]	[۷۹,۵۸ ۹۹,۹۵]	۸۵.۸۷	۸۵/۸۷	۹۰.۲۲	۹۰/۲۲	۹۴/۵۷	۹۴.۵۷	فازی- ازدحام ذرات
۰/۱۳	۰/۰۵	[۶۴,۵۸ ۱۰۰]	[۸۰,۶۸ ۱۰۰]	۸۷/۵۶	۸۷/۶۰	۹۱.۴۱	۹۱/۴۲	۹۵/۲۵	۹۵.۲۵	فازی- مبتنی بر جغرافیای زیستی

(جدول-۱۰): مقایسه‌ای از روش‌های ارائه‌شده در تشخیص سرطان سینه بر روی پایگاه داده انجمن تجزیه و تحلیل تصاویر

ماموگرافی دیگران با تکنیک‌های برتر محاسبات نرم پیشنهادی در این پژوهش

(Table-10): Comparison of the proposed methods with the current soft computing methods applied to classification of the MIAS mammography images

ارزیابی (%)	متغیرها	روش
۸۳/۸۷ (براساس میزان صحت)	ورودی: ۱۰۷ ویژگی استخراج شده از مرحله استخراج ویژگی خروجی: نرمال و غیر نرمال	قطعه بندی تصاویر ماموگرافی توسط روش بهینه‌سازی ازدحام ذرات و طبقه‌بندی توده‌های نرمال و غیر نرمال توسط ماشین بردار پشتیبان [10]
۸۹/۴۷ (براساس میزان صحت)	ورودی: ۱۶ ویژگی استخراج شده از مرحله استخراج ویژگی خروجی: ۱. انباشتگی آهکی ۲. واضح ۳. سوزنی شکل ۴. نا معلوم ۵. کج و کولگی ۶. عدم تقارن ۷. نرمال	سامانه فازی ژنتیک مبتنی بر انتخاب ویژگی موزجک هم رخداد در طبقه‌بندی ماموگرام برای تشخیص سرطان سینه [8]
گزارش نشده است	ورودی: تصویر خروجی: شناسایی توده در تصویر	شناسایی ضایعه در تصاویر ماموگرافی با استفاده از تکامل تفاضلی مبتنی بر خوشه‌بندی اتوماتیک فازی [9]
۹۱/۸۵ (براساس میزان صحت)	ورودی: ویژگی‌های استخراج شده از مرحله استخراج ویژگی خروجی: نرمال و غیر نرمال	استخراج ویژگی توسط روش آنتروپی اثرزی موزجک و طبقه‌بندی توسط رگرسیون خطی [14]
۸۵ (براساس میزان صحت)	ورودی: ویژگی‌های استخراج شده از مرحله استخراج ویژگی خروجی: نرمال و نا متقارن	مدل طبقه‌بندی سازگار مبتنی بر سیستم ایمنی مصنوعی برای شناسایی سرطان سینه [13]
۹۳/۹۰ (براساس میزان صحت)	ورودی: ۴ ویژگی استخراج شده از تصویر توسط رادیولوژیست خروجی: خوش خیم، بدخیم و نرمال	سامانه فازی سوگنو- ژنتیک پیشنهادی
۹۴/۵۸ (براساس میزان صحت)	ورودی: ۴ ویژگی استخراج شده از تصویر توسط رادیولوژیست خروجی: خوش خیم، بدخیم و نرمال	سامانه فازی سوگنو- ازدحام ذرات پیشنهادی
۹۵/۲۵ (براساس میزان صحت)	ورودی: ۴ ویژگی استخراج شده از تصویر توسط رادیولوژیست خروجی: خوش خیم، بدخیم و نرمال	سامانه فازی سوگنو - مبتنی بر جغرافیای زیستی پیشنهادی

مراحل اولیه و با میزان صحت بالا تشخیص داده شود، شانس زنده ماندن فرد افزایش پیدا می‌کند. از طرفی با توجه به وجود منابع عدم قطعیت در سامانه‌های واقعی پزشکی، در این پژوهش جهت تشخیص سرطان سینه، روش‌های هوش

## ۵- نتیجه‌گیری و پیشنهادهای آینده

در تشخیص سرطان سینه با روش‌های هوش مصنوعی، میزان صحت اهمیت ویژه‌ای دارد؛ زیرا اگر این بیماری در

## 6- References

- [1] Ng. Eyk, "A review of thermography as promising non-invasive detection modality for breast tumor", *International Journal of Thermal Sciences*, vol.48, pp.849-859, 2009.
- [2] M.Gautherie, "Thermobiological assessment of benign and malignant breast diseases", *American Journal of Obstetrics and Gynecology*, vol.147, pp.861-869, 1983.
- [3] R. Hosseini and M. Mazinani, "A Mamdani Fuzzy Inference System for Breast Cancer Diagnosis in Intelligent Computer Aided Diagnosis System", 9<sup>th</sup> Symposium of Advances in Science and Technology, National Conference on Science and Computer Engineering, Mashhad, Iran, 2014.
- [4] M. Grgic, K. Delac, M. Ghanbari, "Recent advances in multimedia signal processing and communications", *A survey of image processing algorithms in digital mammography*, Berlin Heidelberg: Springer, 2009, PP. 631-657.
- [5] R. Hosseini and M. Mazinani, "Classification of Uncertainty sources in Intelligent Medical image analysis and understanding Applications", 9<sup>th</sup> Symposium of Advances in Science and Technology, National Conference on Science and Computer Engineering, Mashhad, Iran, 2014.
- [6] R. Hosseini, S.D. Qanadli, S. Barman, M. Mazinani, T. Ellis, and J. Dehmeshki, "Automatic approach for learning and tuning Gaussian interval type-2 fuzzy membership functions applied to lung CAD classification system", *IEEE Transactions on Fuzzy Systems*, vol.20, pp.224-234, 2012.
- [7] F. Merikh Bayat, "Meta-Heuristic Optimization Algorithms", Second Edition, Jahad Daneshgahi Tehran, 2014.
- [8] M. Pawar and S. Talbar, "Genetic Fuzzy System (GFS) based wavelet co-occurrence feature selection in mammogram classification for breast cancer diagnosis", *Perspectives in Science*, vol. 8, pp.247-250, 2016.
- [9] A. Srikrishna, E. Reddy, and V.S. Srinivas, "Detection of Lesion in Mammogram Images Using Differential Evolution Based Automatic Fuzzy Clustering", *Computational Intelligence Techniques in Health Care*, pp. 61-68, 2016.
- [10] K.B. Soulami, M.N. Saidi, A. Tamtaoui, "A CAD System for the Detection of Abnormalities in the Mammograms Using the Metaheuristic Algorithm Particle Swarm Optimization", *Advances in Ubiquitous Networking*, vol. 397, pp. 505-517. 2017.
- [11] D. Stylianos, Tzikopoulos, E. Michael Mavroforakis, V. Georgiou, Nikos Dimitropoulos, and Sergios Theodoridis, "A fully automated scheme for mammographic

محاسباتی مبتنی بر فازی با هدف قابلیت بیشتر در مدیریت منابع عدم قطعیت با تمرکز بر تشخیص با میزان صحت بالاتر پیشنهاد شده است. در پژوهش حاضر، طبقه‌بندی توده‌ها جهت تشخیص موارد نرمال، خوش‌خیم و بدخیم در سامانه‌های پیشنهادی توسط روش‌های متفاوت محاسبات نرم مبتنی بر فازی انجام شده است. همان‌طور که در نتایج قابل مشاهده است، مدل فازی-مبتنی بر جغرافیای زیستی توانایی بالایی در تشخیص سرطان سینه دارد. از مزایای اصلی این پژوهش در مقایسه با دیگر مطالعات انجام‌شده، تعداد ویژگی‌های ورودی کمینه در سامانه پیشنهادی است که این امر از پیچیدگی سامانه کاسته است. در روش‌های پیشنهادی از روش‌های مختلف بهینه‌سازی جهت بهبود کارایی مدل‌های هوشمند استفاده شده است. از روش‌های با قابل اعتماد بالاتر مانند تحلیل منحنی ROC و اعتبارسنجی ده‌بخشی جهت سنجش کارایی سامانه‌ها استفاده شده است. نتایج به‌دست‌آمده نشان می‌دهد روش‌های پیشنهادی در این پژوهش به نرخ صحت طبقه‌بندی قابل قبول قابل رقابت با سایر روش‌های قبلی ارایه‌شده دست یافته‌اند و روش هابیریدی فازی-جغرافیای زیستی صحت عملکرد طبقه‌بندی در تشخیص سرطان سینه را نسبت به سایر روش‌های قبلی ارایه‌شده بهبود داده است.

با توجه به این که ویژگی‌های ورودی در سامانه‌های پیشنهادی کارا توسط رادیولوژیست از اطلاعات تصاویر پزشکی استخراج شده است، پیشنهاد می‌شود جهت خودکار کردن فرآیند طبقه‌بندی سامانه، از روش‌های پردازش تصاویر به‌منظور استخراج ویژگی از تصاویر پزشکی، استفاده شود و نتایج با سامانه‌های معرفی‌شده ارزیابی شود. همچنین سامانه‌های پیشنهادی را می‌توان بر روی پایگاه داده‌های جمع‌آوری‌شده از مراکز معتبر درمانگاهی دیگر اجرا کرد.

باتوجه به مختصات تعیین‌شده از مرکز توده‌ها در تصاویر ماموگرافی توسط رادیولوژیست، می‌توان موقعیت توده‌ها در نواحی چهار گانه سینه را با روش‌های خودکار به‌دست آورد. از طرفی به‌منظور بهبود عملکرد سامانه، می‌توان قوانین را نیز با استفاده از الگوریتم‌های بهینه‌سازی، آموزش داد. همچنین می‌توان از توابع عضویت مختلف و ترکیبی در جهت بهبود عملکرد سامانه استفاده کرد و نوع توابع عضویت را نیز با استفاده از الگوریتم‌های بهینه‌سازی به‌دست آورد.

- [23] C. Darwin, The Origin of Species. New York: Gramercy, 1995.
- [24] D. Simon, "Biogeography-Based Optimization", *IEEE Transactions on Evolutionary Computation*, vol. 12, pp.702-713, 2008.
- [25] MIAS Database, Available from: <http://peipa.essex.ac.uk/info/mias.html>, Updated on: Dec 2012.
- [26] G. S. David Sam Jayakumar, Bejoy John Thomas, "A New Procedure Of Clustering Based on Multivariate Outlier Detection", *Journal of Data Science*, vol.11, pp.69-84, 2013.
- [27] AP. Bradley, "The use of the area under the ROC curve in the evaluation of machine learning algorithms", *Pattern Recognition*, vol.30, pp.1145-1159. 1997.
- [28] T. Fawcett, "An introduction to ROC analysis", *Pattern Recognition Letters*, vol.27, pp.861-874. 2006.
- [29] J. Ha, M. Kamber, and J. Pei, "Data Mining Concepts and Techniques", Chapter (5). 2006
- [30] PN. Tan, M. Steinbach, and V. Kumar, "Introduction to Data Mining". Pearson Addison Wesley, Chapter (5), 2006.
- [31] M. Mahdavi, H. Ahaki, and B. NaserSharif, "Design of a Currency Recognition based on Neural Network and Image Texture and Color", *Signal and Data Processing Journal*, Issue 7, No 2, Feb. 2011.
- [32] N. Gohariyan, S.Moghimi, and H.Kalani, "Using a Combination of Genetic Algorithm and Artificial Neural Network for Predicting Power of Gas from Electromyogram Signal", *Signal and Data Processing Journal*, Issue 14, No. 1, Jun 2017.
- [33] segmentation and classification based on breast density and asymmetry", *Computer Methods and Programs in Biomedicine*. vol.102, pp. 47-63, 2011.
- [12] R. Rouhi, M. Jafari, Sh. Kasaei, and P. Keshavarzian, "Benign and malignant breast tumors classification based on region growing and CNN segmentation", *Expert Systems with Applications*, vol. 42, pp.990-1002. 2015.
- [13] G. Magna, S. Velappa Jayaraman, P. Casti P, A. Mencattini, C. Di Natale, and E. Martinelli, "Adaptive classification model based on artificial immune system for breast cancer detection", *AISEM Annual Conference*, Trento, Feb 3-5. 2015, pp.1-4.
- [14] Y. Chen, Y. Zhang, HM. Lu, XQ. Chen, JW. Li, and SH. Wang, "Wavelet energy entropy and linear regression classifier for detecting abnormal breasts", *Multimedia Tools And Applications*, 2016, pp. 1-20.
- [15] JH. Holland, "Adaption in natural and artificial systems", University of Michigan Press, pp.232, 1975.
- [16] AE. Eiben and JE. Smith, "Introduction to Evolutionary computing genetic algorithms", First Edition, Verlag Berlin Heidelberg: Springer, pp.300, 2003.
- [17] T. Back, "Selective pressure in evolutionary algorithms: A characterization of selection mechanisms", *Proceedings of the First IEEE Conference on Evolutionary Computation, IEEE World Congress on Computational Intelligence*, Orlando, FL, June 27-29, 1994, pp. 57-62.
- [18] J. Kennedy and RC. Eberhart, "Particle Swarm Optimization", *Proceedings of the IEEE International Conference on Neural Networks*, 1995, vol. 4, pp. 1942-1948.
- [19] Y. Shi and RC. Eberhart, "A Modified Particle Swarm Optimizer", *Proceedings of the IEEE International Conference on Evolutionary Computation*. Anchorage, AK, May 4-9, 1998, pp. 69-73.
- [20] M. Nabab Alam, "Particle Swarm Optimization", Algorithms and its Codes in MATLAB, 2016.
- [21] Y. Shi and RC. Eberhart, "Comparing Inertia Weights and Construction Factors in Particle Swarm Optimization", *Proceedings of the IEEE International Conference on Evolutionary Computation*. La Jolla, CA, July 16-19, 2000, PP.84-88.
- [22] A. Wallace, "The Geographical Distribution of Animals", Boston, MA: Adamant Media Corporation, 2005.

[۳۱] مهرگان مهدوی، حبیب آهکی، بابک ناصرشریف، طراحی یک سیستم تشخیص اسکناس مبتنی بر شبکه عصبی با استفاده از مشخصه‌های بافت و رنگ تصویر، مجله پردازش علایم و داده‌ها، دوره ۷، شماره ۲ - ۱۲ - ۱۳۸۹.

[۳۲] نازنین گوهریان، سحر مقیمی، هادی کلانی، استفاده از ترکیب الگوریتم ژنتیک و شبکه‌های عصبی مصنوعی برای پیش‌بینی نیروی گاز گرفتن از روی سیگنال الکترومایوگرام، مجله پردازش علایم و داده‌ها، دوره ۱۴، شماره ۱ - (۳-۱۳۹۶).

[32] N. Gohariyan, S.Moghimi, and H.Kalani, "Using a Combination of Genetic Algorithm and Artificial Neural Network for Predicting Power of Gas from Electromyogram Signal", *Signal and Data Processing Journal*, Issue 14, No. 1, Jun 2017.



**الناز خدادادی** فارغ‌التحصیل رشته مهندسی کامپیوتر هوش مصنوعی در مقطع کارشناسی ارشد از دانشگاه آزاد اسلامی واحد شهرقدس، تهران است. ایشان در دوره کارشناسی ارشد به‌عنوان دانشجوی ممتاز دوره بوده‌اند. نشانی رایانامه ایشان عبارت است از: [elnazkhodadadi88@gmail.com](mailto:elnazkhodadadi88@gmail.com)



**راحیل حسینی** فارغ‌التحصیل رشته سامانه‌های هوشمند در مقطع دکترا در دانشگاه کینگستون لندن هستند. ایشان عضو گروه پژوهشی تحلیل و پردازش تصاویر پزشکی دانشگاه کینگستون لندن بوده‌اند. تحصیلات خود را در مقطع کارشناسی در دانشگاه خوارزمی تهران در سال ۲۰۰۱ و کارشناسی ارشد در دانشگاه آزاد اسلامی واحد تهران جنوب با رتبه ممتاز در سال ۲۰۰۴ به اتمام رسانده‌اند. موضوعات پژوهشی مورد علاقه ایشان سامانه‌های هوشمند فازی، پردازش و تحلیل تصاویر پزشکی، داده‌کاوی، شناسایی الگو در سامانه‌های پیچیده مهندسی دنیای واقعی از جمله کاربردهای پزشکی و مطالعات سرطان و شناسایی پدیده‌های طبیعی است. ایشان تعداد زیادی مقاله و انتشارات در این زمینه دارند. نشانی رایانامه ایشان عبارت است از: [rahil.hosseini@qodsiau.ac.ir](mailto:rahil.hosseini@qodsiau.ac.ir)



**مهدی مزینانی** فارغ‌التحصیل رشته پردازش تصویر در مقطع دکترا در دانشگاه کینگستون لندن هستند. ایشان عضو گروه پژوهشی تحلیل و پردازش تصاویر پزشکی دانشگاه کینگستون لندن بوده‌اند. تحصیلات خود را در مقطع کارشناسی و کارشناسی ارشد در دانشگاه سمنان و کارشناسی ارشد را در سال ۲۰۰۲ به اتمام رسانده‌اند. موضوعات پژوهشی مورد علاقه ایشان پردازش و تحلیل تصاویر پزشکی، شناسایی الگوهای در تصاویر و اندازه‌گیری تصاویر دیجیتال است. ایشان تعداد زیادی مقاله و انتشارات در این زمینه دارند. نشانی رایانامه ایشان عبارت است از: [Mahdi.mazinani@qodsiau.ac.ir](mailto:Mahdi.mazinani@qodsiau.ac.ir)

