

## AIS-RCA: An Efficient Feature Reduction Method to Improve the Seizure Detection Rate

A. Zare<sup>1\*</sup>, R. Boostani<sup>2</sup>, M. Zolghadr Jahromi<sup>3</sup>

<sup>1</sup> Instructor, Department of Computer Engineering, Shiraz Branch, Islamic Azad University, Shiraz, Iran

<sup>2</sup> Associate Professor, Biomedical Engineering Group, Department of Computer Sciences and Engineering, Shiraz University, Shiraz, Iran, boostani@shirazu.ac.ir

<sup>3</sup> Associate Professor, Department of Computer Sciences and Engineering, Shiraz University, Shiraz, Iran, zjahromi@shirazu.ac.ir

### Abstract

There is a growing interest to improve seizure prediction by online analyzing of electroencephalogram (EEG) signals in epileptic patients. Seizure attack is occurred infrequently and unpredictably; hence, automatic detection of seizure during long-term is highly recommended. In this paper a novel Feature Reduction method namely AIS-RCA which adopted from the immunity system is proposed to improve the seizure detection rate. The automatic seizure detection can be performed in two successive stages: 1) The feature extraction/selection stage from EEG signals and 2) classifying the feature vectors by an efficient classifier. In this study, first, pseudo-Wigner-Ville distribution was applied to each window of the EEG signals and then the extracted features were transformed by AIS-RCA transform to represent the features in a more separable space.

The AIS-RCA transformation matrix is estimated by using chunklets (a chunklet is defined as a subset of points that are known to be same). AIS-RCA using the proposed Artificial Immune System algorithm named Adaptive Distance-AIRS to discover the chunklets in the data space. Finally KNN classifier was applied to the transformed features to classify the seizure and non-seizure windows. The experimental results show that the proposed method yields epileptic detection accuracy rate up to 99.9% which is better than the results achieved by other types of features such as FFT, Wavelet transform, entropy and chaotic measures.

**Key words:** EEG Signal, Time-Frequency Analysis, Spectrum, Epileptic seizures, AIS-RCA, AD-AIRS, RCA, Artificial Immune System (AIS).

\* Corresponding author

Address: Amin Zare, Islamic Azad University- Shiraz Branch, Sadra, P.B 71955-845, Shiraz, Iran

Tel: +98 917 3385218, +98 711 7254377

Fax: +987116473575

E-mail: amin\_zare@cse.shirazu.ac.ir

## AIS-RCA: روشی مؤثر برای بهبود فضای ویژگی‌ها در افزایش دقت تشخیص حملات صرع

امین زارع<sup>۱\*</sup>، رضا بوستانی<sup>۲</sup>، منصور ذوالقدر جهرمی<sup>۳</sup>

<sup>۱</sup> مربی، دانشگاه آزاد اسلامی، واحد شیراز، گروه مهندسی کامپیوتر، شیراز

<sup>۲</sup> دانشیار، گروه مهندسی پزشکی، دانشکده مهندسی، دانشگاه شیراز، boostani@shirazu.ac.ir

<sup>۳</sup> دانشیار، گروه علوم و مهندسی کامپیوتر، دانشکده مهندسی، دانشگاه شیراز، zjahromi@shirazu.ac.ir

### چکیده

پیش‌بینی زمان وقوع حملات صرع در بیماران از جمله موضوعاتیست که مورد توجه محققان است. حملات صرع به طور نامنظم و غیر قابل پیش‌بینی شده‌ای اتفاق می‌افتند. بنابراین تشخیص حملات صرع از روی سیگنال‌های EEG که در بازه زمانی طولانی گرفته می‌شوند؛ بسیار حائز اهمیت است. این امر تشخیصی به دو مرحله مجزای استخراج ویژگی‌ها از قطعات سیگنال EEG و اعمال الگوریتم طبقه‌بندی بر روی بردارهای ویژگی تقسیم می‌شود. به همین منظور در مرحله اول با استفاده از تحلیل زمان-فرکانس بر روی قطعات سیگنال EEG و به دست آوردن صفحه زمان-فرکانس هر قطعه، استخراج ویژگی‌ها از سیگنال‌ها انجام می‌شود. در مرحله دوم با استفاده از الگوریتم نزدیک‌ترین همسایه کار تشخیص حملات صورت می‌گیرد. اما قبل از اعمال الگوریتم طبقه‌بندی، برای اصلاح فضای ویژگی‌ها و یادگیری معیار فاصله، از الگوریتم AIS-RCA استفاده شده است. این الگوریتم برای به دست آوردن ماتریس تبدیل W، داده‌ها را به صورت مجموعه‌ای از دسته‌ها در نظر می‌گیرد و با ارائه الگوریتم جدید AD-AIRS و با الهام گرفتن از سیستم ایمنی بدن دسته‌ها را می‌یابد. آزمایش‌های انجام شده نشان دهنده دقت ۱۰۰٪ و بهبود نتایج در مقایسه با برخی روش‌های انتقال موجک، آنتروپی، معیار بی‌نظمی و تبدیل انتقال فوریه سریع را نشان می‌دهد.

کلیدواژگان: سیگنال EEG، پردازش زمان-فرکانس، صفحه زمان-فرکانس، AIS-RCA، AD-AIRS، RCA، صرع، سیستم ایمنی مصنوعی (AIS).

\*عهده‌دار مکاتبات

نشانی: شیراز، شهرک صدر، پردیس دانشگاه آزاد اسلامی-واحد شیراز، صندوق پستی: ۷۱۹۵۵-۸۴۵  
تلفن: ۰۹۱۷۳۳۸۵۲۱۸-۰۷۱۱۷۲۵۴۳۷۷، دورنگار: ۰۷۱۱۶۴۷۴۶۰۵، پیام‌نگار: amin\_zare@cse.shirazu.ac.ir

## ۱- مقدمه

کننده‌هایی هستند که برای شناسایی سیگنال‌های حملات صرع مورد استفاده قرار گرفته‌اند.

از آنجا که سیگنال‌های حملات صرع از نوع ناپیوسته<sup>۱۱</sup> هستند؛ با نمایش زمان-فرکانس<sup>۱۱</sup> می‌توان اطلاعات زمان و فرکانس را به صورت همزمان نمایش داد. به همین دلیل در بسیاری از مطالعات از پردازش زمان-فرکانس برای استخراج ویژگی‌ها استفاده شده است [۱۶، ۲۰-۲۲].

مرجع [۲۳] از بهترین مطالعاتیست که با استفاده از تحلیل زمان-فرکانس، استخراج ویژگی‌ها از سیگنال‌های EEG را انجام داده است و با به‌کارگیری شبکه عصبی به عنوان طبقه‌بندی کننده به دقت‌های خوبی در مسئله تشخیص خودکار حملات صرع دست یافته است. زالس<sup>۱۲</sup> و همکاران ابتدا قطعات سیگنال‌های EEG را در بازه‌های زمانی ثابت به دست آورده و با به‌کارگیری تبدیل SPWVD<sup>۱۳</sup> صفحه زمان-فرکانس قطعات سیگنال‌های EEG را استخراج نموده‌اند [۲۳]. با تقسیم‌بندی صفحه زمان-فرکانس به پنجره‌های زمانی و ناحیه‌های فرکانسی مختلف، تقسیمات متعددی از صفحه زمان-فرکانس به دست می‌آید و انرژی موجود در هر قسمت، به عنوان بردار ویژگی‌ها در نظر گرفته می‌شود. دقت‌های گزارش شده نیز براساس ترکیبات مختلف پنجره‌های زمانی و نواحی فرکانسی در صفحه زمان-فرکانس است. مطالعه زالس و همکاران از این نظر که در تشخیص خودکار حملات صرع ۴ نوع مسئله طبقه‌بندی در نظر گرفته است؛ یکی از کامل‌ترین کارهای انجام شده بوده و در تمام مسئله‌ها به دقت‌های مطلوبی دست یافته است.

در این مطالعه نیز از روش پیشنهادی در مرجع [۲۳] برای استخراج ویژگی‌هایی از سیگنال‌های EEG استفاده شده است اما تنها یک حالت از تقسیم‌بندی‌های صورت گرفته بر صفحه زمان-فرکانس برای تولید بردارهای ویژگی در نظر گرفته شده است. فرایند تشخیص خودکار حملات صرع را می‌توان به ۴ مرحله تقسیم کرد. در مرحله اول با پردازش زمان-فرکانس بر روی قطعات سیگنال‌های EEG صفحه زمان-فرکانس قطعات به دست می‌آیند و در مرحله دوم با

بیماری صرع یکی از بیماری‌های عصبیست که در اثر سوء عمل ناگهانی و تکراری مغز رخ می‌دهد. یکی از ابزارهای اولیه‌ای که عصب‌روانشناسان برای تشخیص بیماری افراد مظنون به کار می‌گیرند ثبت و بررسی سیگنال‌های EEG<sup>۱</sup> است که توسط دستگاه ثبت امواج الکتریکی مغز گرفته می‌شود.

برای تشخیص بیماری، سیگنال‌های الکتریکی EEG به مدت حدود ۲۰ دقیقه ثبت می‌شوند. بهتر است رفتار خود بیمار نیز در این لحظات در نظر گرفته شود. برای تشخیص دقیق‌تر، سیگنال‌های EEG برای مدت زمان بیشتری (حتی تا چند روز) به همراه رفتار بیمار ضبط و مورد بررسی قرار می‌گیرند. بررسی دقیق حجم بزرگی از داده‌ها و تشخیص نوع سیگنال کاری دشوار است. بنابراین برای ابداع سیستم‌های تشخیص خودکار کامپیوتری تلاش‌های زیادی صورت گرفته است. سیستم تشخیص خودکار حملات صرع، با شناسایی الگوهای مختلف موجود در سیگنال‌های EEG، می‌تواند افراد سالم و ناسالم و ویژگی‌های بیماری را تشخیص دهد. بهبود سیستم‌های تشخیص خودکار حملات صرع از جمله دغدغه‌های محققان است و تاکنون روش‌های مختلفی برای آن ارائه شده است.

روش‌های تشخیص خودکار حملات صرع از دو مرحله اساسی استخراج ویژگی‌ها<sup>۲</sup> از سیگنال‌ها و طراحی طبقه‌بندی کننده تشکیل شده‌اند. از جمله روش‌هایی که برای استخراج ویژگی‌ها از سیگنال‌های EEG استفاده کرده‌اند، می‌توان به روش‌هایی اشاره کرد که با به دست آوردن ضرایب موجک<sup>۳</sup> [۴-۱] این کار را انجام داده‌اند. استفاده از ویژگی‌های فضایی [۵، ۶]، بررسی نسبی دامنه سیگنال [۷]، بعد همبستگی<sup>۴</sup> [۸]، [۹]، توان لیپانوف<sup>۵</sup> [۱۰] و محاسبه انتروپی<sup>۶</sup> [۴، ۱۱، ۱۲] نیز در برخی از کارها مورد استفاده قرار گرفته‌اند. نهایتاً بردارهای ویژگی به دست آمده، به الگوریتم‌های طبقه‌بندی داده می‌شوند. طبقه‌بندی کننده‌های آماری [۱۳، ۱۴]، درخت‌های تصمیم<sup>۷</sup> [۱۵]، شبکه‌های عصبی [۱، ۱۰، ۱۶]، SVM<sup>۸</sup> [۱۷-۱۹]، ANFIS<sup>۹</sup> [۳، ۴، ۱۱] از جمله طبقه‌بندی

<sup>1</sup> Electroencephalograph

<sup>2</sup> Feature Extraction

<sup>3</sup> Wavelet

<sup>4</sup> Correlation dimension

<sup>5</sup> Lyapunov exponents

<sup>6</sup> Entropy

<sup>7</sup> Decision tree

<sup>8</sup> Support Vector Machine

<sup>9</sup> Adaptive Neuro Fuzzy Inference System

<sup>10</sup> Non stationary

<sup>11</sup> Time-Frequency

<sup>12</sup> Tzallas

<sup>13</sup> Smoothed Pseudo-Wigner-Ville distribution

در این مقاله با استفاده از الگوریتم AIS-RCA و به دست آوردن ماتریس تبدیل عمومی  $W$ ، فضای ویژگی‌ها به گونه‌ای انتقال می‌یابد که به ویژگی‌های مربوط وزن بیشتر و به ویژگی‌های نامربوط وزن کمتری تخصیص داده می‌شود [۲۷]. در فضای جدید درهم‌ریختگی داده‌ها کمتر و نمایش داده‌ها به گونه‌ای می‌شود که جداپذیری آنها بهتر صورت می‌پذیرد. تبدیل AIS-RCA به کمک اطلاعات موجود در دسته‌ها<sup>۱۷</sup> تخمین زده می‌شود. دسته‌ها زیر مجموعه‌هایی از داده‌ها هستند که به یکدیگر شبیه بوده و ممکن است تعداد اندکی از داده‌ها را در برگیرند. داده‌های یک دسته معمولاً از یک طبقه در نظر گرفته می‌شوند [۲۸].

در این مطالعه برای پیدا کردن دسته‌ها در داده‌ها الگوریتم AD-AIRS<sup>۱۸</sup> با الهام گرفتن از سیستم ایمنی بدن ارائه شده است. الگوریتم AD-AIRS بهبود یافته الگوریتم AIRS<sup>۱۹</sup> [۲۹-۳۱] به شمار می‌رود. AD-AIRS با الهام گرفتن از عملکرد سیستم ایمنی بدن از روی الگوهای آموزشی، سلول‌های حافظه (الگوهای نماینده) تولید می‌کند. این سلول‌ها برای هر طبقه به طور جداگانه و هر جا که تجمع داده‌ها بیشتر باشند، ساخته می‌شوند. در الگوریتم AD-AIRS هریک از سلول‌های حافظه محدوده دایره‌ای شکل را در بر می‌گیرد که دسته‌های مورد استفاده در الگوریتم AIS-RCA را نمایش می‌دهد.

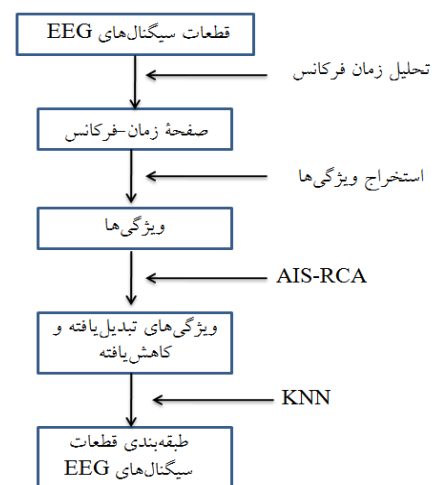
در ادامه مجموعه‌های آزمایش شده؛ پردازش زمان-فرکانس و نحوه به دست آوردن ویژگی‌ها توضیح داده شده‌اند. سپس الگوریتم‌های RCA و AD-AIRS ارائه شده‌اند و الگوریتم AIS-RCA معرفی شده است. کارایی الگوریتم پیشنهادی و نتایج آزمایش‌های انجام شده و مقایسه آن با روش‌های دیگر در بخش پایانی آمده است.

## ۲- روش جمع‌آوری داده‌ها

در این مطالعه از مجموعه داده‌های معرفی شده در مرجع [۳۲] شامل سیگنال‌های افراد سالم و بیماران دچار صرع<sup>۱۹</sup> استفاده شده است. این مجموعه، شامل ۵۰۰ قطعه سیگنال در ۵ مجموعه S, F, N, O, Z که هر مجموعه شامل ۱۰۰ قطعه

گذاشتن پنجره‌های زمانی در ناحیه‌های فرکانسی مختلف، میزان انرژی سیگنال در هر قسمت از صفحه زمان-فرکانس محاسبه و به عنوان ویژگی‌های استخراج شده در نظر می‌شوند. در مرحله سوم برای انتخاب بهترین ویژگی‌ها و بهبود فضای ویژگی‌ها از تبدیل AIS-RCA<sup>۱۴</sup> استفاده شده است و نهایتاً ویژگی‌های بهبود یافته را برای تشخیص سیگنال‌های حملات صرع به طبقه‌بندی کننده نزدیک‌ترین همسایه می‌دهیم. در شکل ۱ این مراحل نشان داده شده است.

هدف نهایی این مقاله تشخیص سیگنال‌های حملات صرع با استفاده از الگوریتم نزدیک‌ترین همسایه است. یکی از اهدافی که دنبال می‌شود بهبود طبقه‌بندی کننده نزدیک‌ترین همسایه از طریق روش‌های یادگیری معیار فاصله<sup>۱۵</sup> است. بسیاری از الگوریتم‌های طبقه‌بندی که از معیار فاصله استفاده می‌کنند؛ به شدت به این معیار وابسته‌اند. با استفاده از روش‌های یادگیری می‌توان این الگوریتم‌ها را بهبود بخشید. این روش‌ها با کم کردن اثر ویژگی‌های بی‌ربط، متقابل<sup>۱۶</sup> و اغتشاشی؛ فضای داده‌ها را به صورت بهتری نمایش می‌دهند. بسیاری از مطالعات سال‌های اخیر به صورت تجربی و تئوری بهبود کارایی الگوریتم‌های طبقه‌بندی و خوشه‌بندی را در صورت استفاده از روش‌های یادگیری معیار فاصله نشان داده‌اند [۲۴-۲۶]. در بسیاری از مواقع انتخاب یک معیار فاصله مناسب مهم‌تر از نوع الگوریتم مورد استفاده است.



شکل ۱- مراحل روش پیشنهادی

<sup>14</sup> Artificial Immune System - Relevant Component Analysis

<sup>15</sup> Distance Metric Learning

<sup>16</sup> Interacting

<sup>17</sup> Chunklets

<sup>18</sup> Adaptive Distance Artificial Immune Recognition System

<sup>18</sup> Artificial Immune Recognition System

<sup>19</sup> Epileptic seizure

جدول ۱- مجموعه داده‌های مورد استفاده

مجموعه داده	نوع	تعداد قطعات
Z	سالم	۱۰۰
O	سالم	۱۰۰
N	بدون صرع	۱۰۰
F	بدون صرع	۱۰۰
S	حمله صرع	۱۰۰

جدول ۲- مسئله‌های طبقه‌بندی افراد سالم و ناسالم

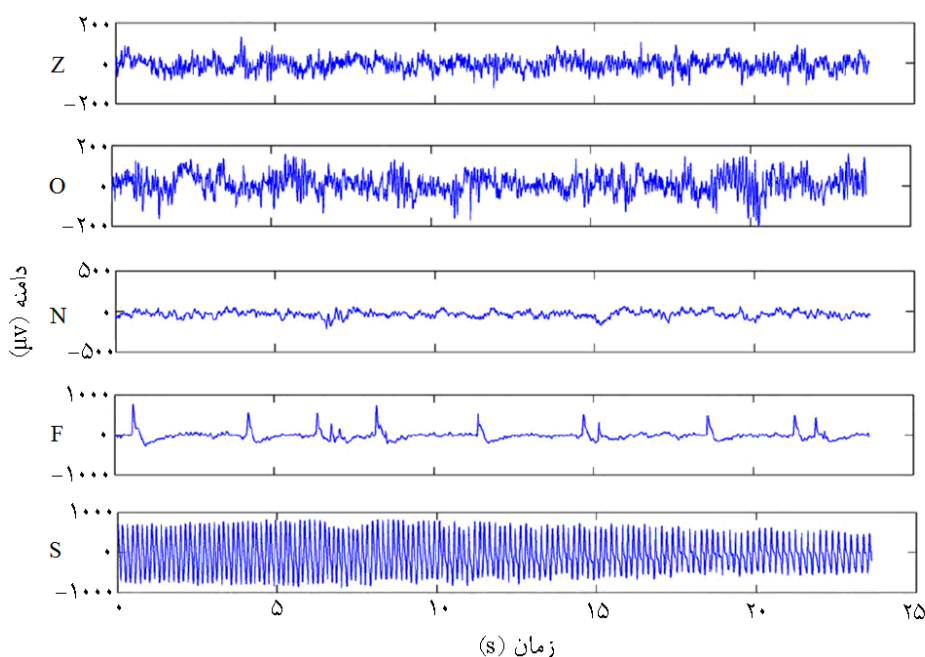
مسئله ۱	Z,O,N,F (طبقه ۱)
	S (طبقه ۲)
مسئله ۲	Z (طبقه ۱)
	S (طبقه ۲)

### ۲-۱- پردازش زمان-فرکانس

روش معمول برای پردازش سیگنال‌های ایستا نمایش فرکانسی سیگنال با استفاده از تبدیل فوریه است. اما این روش برای سیگنال‌های نایستا کافی نیست. در این‌گونه سیگنال‌ها از نمایش زمان-فرکانس سیگنال استفاده می‌شود. برای این منظور سیگنال را در دو بُعد فرکانس و زمان نشان می‌دهیم که تغییرات فرکانسی سیگنال را در بُعد زمان به خوبی نشان می‌دهد.

سیگنال EEG به مدت ۲۳/۶s است. مجموعه Z و O از ۵ فرد داوطلب سالم با چشم‌های باز و بسته ثبت شده است. مجموعه‌های F و N از ناحیه اپیلتوژنیک<sup>۲۰</sup> و هیپوکمپال<sup>۲۱</sup> در حالت بدون صرع<sup>۲۲</sup> ثبت شده‌اند و مجموعه S شامل سیگنال‌هاییست که در هنگام حمله صرع ضبط شده است (جدول ۱). مجموعه‌های Z و O به صورت خارج مجموعه‌ای<sup>۲۳</sup> و F، N و S به صورت درون مجموعه‌ای<sup>۲۳</sup> ثبت شده‌اند و داده‌ها با فرکانس نمونه‌برداری ۱۷۳/۶۱ نمونه بر ثانیه با دقت ۱۲ بیت دیجیتایز شده‌اند (بازه فرکانس ۰/۵Hz تا ۸۵Hz تغییر می‌کند). نمونه‌ای از این سیگنال‌ها در شکل ۲ نمایش داده شده‌اند.

با استفاده از مجموعه داده‌های فوق، ۲ مسئله طبقه‌بندی مطرح می‌شود. در مسئله اول، مجموعه‌های Z، O، N، F به عنوان طبقه ۱ و مجموعه S به عنوان طبقه دوم در نظر گرفته می‌شود و در مسئله دوم می‌خواهیم سیگنال‌های مجموعه S و Z را از یکدیگر جدا کنیم (جدول ۲). از آنجا که در پزشکی تشخیص این دو گروه از افراد سالم و بیمار مد نظر است در بیشتر تحقیقات انجام شده، فقط دو مسئله طبقه‌بندی تعریف شده‌اند [۲-۴، ۳۳-۳۸]، ولی در برخی از مطالعات مسائل سه طبقه نیز تعریف شده‌اند [۲۳].



شکل ۲- نمونه‌هایی از سیگنال‌های Z، O، N، F، S

<sup>20</sup> Epilitogenic  
<sup>24</sup> Intracranially

<sup>21</sup> Hippocampal

<sup>22</sup> Seizure-Free

<sup>23</sup> Extracranially

تقسیم‌بندی ناحیه فرکانسی به گونه‌ای است که در نواحی از صفحه زمان-فرکانس با انرژی بیشتر، قسمت‌های کوچک‌تری در نظر گرفته شده است.

تعداد ویژگی‌های به‌دست آمده به نحوه تقسیم‌بندی بستگی دارد. در این مطالعه، تعداد تقسیمات برابر ۶۵ (۵ پنجره زمانی و ۱۳ ناحیه فرکانسی) است. همچنین کل انرژی سیگنال به این ویژگی‌ها افزوده می‌شود. بنابراین هر بردار داده دارای ۶۶ ویژگی است.

### ۲-۳- بهبود فضای ویژگی‌ها

در این مطالعه از الگوریتم  $KNN^{26}$  برای طبقه‌بندی داده‌ها استفاده شده است. الگوریتم  $KNN$  با به کارگیری معیار فاصله، کلاس داده آزمون را از روی کلاس نزدیک‌ترین داده‌های آموزشی می‌یابد. این الگوریتم به شدت به معیار فاصله استفاده شده در فضای داده وابسته است. بسیاری از تحقیقات با استفاده از روش‌های یادگیری معیار فاصله به دقت‌های بهتری نسبت به استفاده از فاصله اقلیدسی معمولی دست یافته‌اند [۲۴-۲۶]. از جمله این روش‌ها می‌توان به الگوریتم  $RCA$  اشاره کرد [۴۱].

#### ۲-۳-۱- تحلیل اجزاء وابسته ( $RCA$ )

این الگوریتم ساده و مؤثر با استفاده از مجموع وزن‌دار ماتریس‌های کوواریانس  $^{28}$  دسته‌های موجود در داده‌ها، نوعی فاصله ماهالانوبیس  $^{29}$  را ایجاد می‌کند. این الگوریتم مانند روش‌های  $PCA^{30}$  و  $LDA^{31}$  به ویژگی‌های آماری مرتبه دوم  $^{32}$  داده‌ها وابسته است. برخی از مطالعات بهبود نتایج در استفاده از روش یادگیری  $RCA$  را نشان داده‌اند [۲۸، ۴۲، ۴۳].  $RCA$  با به‌دست آوردن ماتریس تبدیل  $W$ ، تأثیر ویژگی‌های نامربوط را کمتر می‌کند. در این روش فضای ویژگی‌ها با استفاده از تبدیل خطی  $W$  تغییر می‌یابد به گونه‌ای که ویژگی‌های نامربوط در معیار فاصله مورد استفاده در فضای جدید اثر کمتری دارند. معیار فاصله اقلیدسی در فضای جدید شبیه معیار فاصله ماهالانوبیس در فضای اصلی داده‌هاست. به‌دست آوردن ماتریس تبدیل  $W$  از دسته‌های موجود در مجموعه داده استفاده می‌شود. هر دسته، زیرمجموعه‌ای از داده‌هاست که به یکدیگر شبیه‌اند.

پردازش زمان-فرکانس بر اساس نوع تبدیل مورد استفاده به دو دسته خطی و غیرخطی تقسیم می‌شود. در این مطالعه صفحه زمان-فرکانس از روش غیرخطی و با استفاده از تبدیل  $SPWVD$  [۳۹، ۴۰] در رابطه (۱) به‌دست می‌آید.

$$SPWVD_x(t, \omega) = \int_{-\infty}^{+\infty} h(\tau) \left( \int_{-\infty}^{+\infty} g(s-t)x(s+\frac{\tau}{2})x^*(s-\frac{\tau}{2})ds \right) e^{-j2\pi\omega\tau} d\tau \quad (1)$$

در این رابطه  $x(\cdot)$  سیگنال،  $t$  زمان،  $\omega$  فرکانس،  $g(\cdot)$  و  $h(\cdot)$  توابع پنجره‌ای زمان و فرکانس هستند.  $g(\cdot)$  پنجره همینگ  $^{25}$ ، به طول ۶۴ نقطه و  $h(\cdot)$  پنجره فرکانسی به طول ۵۱۲ نقطه است.

### ۲-۲- استخراج ویژگی‌ها

برای استخراج ویژگی‌ها صفحه زمان-فرکانس را قسمت‌بندی می‌کنیم [۲۳] و با استفاده از رابطه (۲)، انرژی نسبی موجود در هر قسمت را محاسبه و به عنوان بردار ویژگی‌ها در نظر می‌گیریم:

$$f(i, j) = \int_{t_i} \int_{\omega_j} SPWVD_x(t, \omega) d\omega dt \quad (2)$$

در این رابطه  $t_i$  نشان دهنده  $i$ امین پنجره زمانی و  $\omega_j$  نشان دهنده  $j$ امین ناحیه فرکانسیست.

برای تقسیم‌بندی صفحه زمان-فرکانس مطابق جدول ۳، بعد زمان را به ۵ قسمت مساوی و بعد فرکانس را به ۱۳ ناحیه فرکانسی تقسیم می‌کنیم.

جدول ۳- تقسیم‌بندی ناحیه‌های فرکانسی در بعد فرکانس صفحه

زمان-فرکانس بر اساس روش مرجع [۲۳]

ناحیه‌های فرکانسی (HZ)
۰-۲
۲-۴
۴-۶
۶-۸
۸-۱۰
۱۰-۱۲
۱۲-۱۶
۱۶-۲۰
۲۰-۲۴
۲۴-۲۸
۲۸-۳۲
۳۲-۳۶
۳۶-۴۰

<sup>25</sup> Haming

<sup>26</sup> K-Nearest Neighbor

<sup>27</sup> Relevant Component Analysis

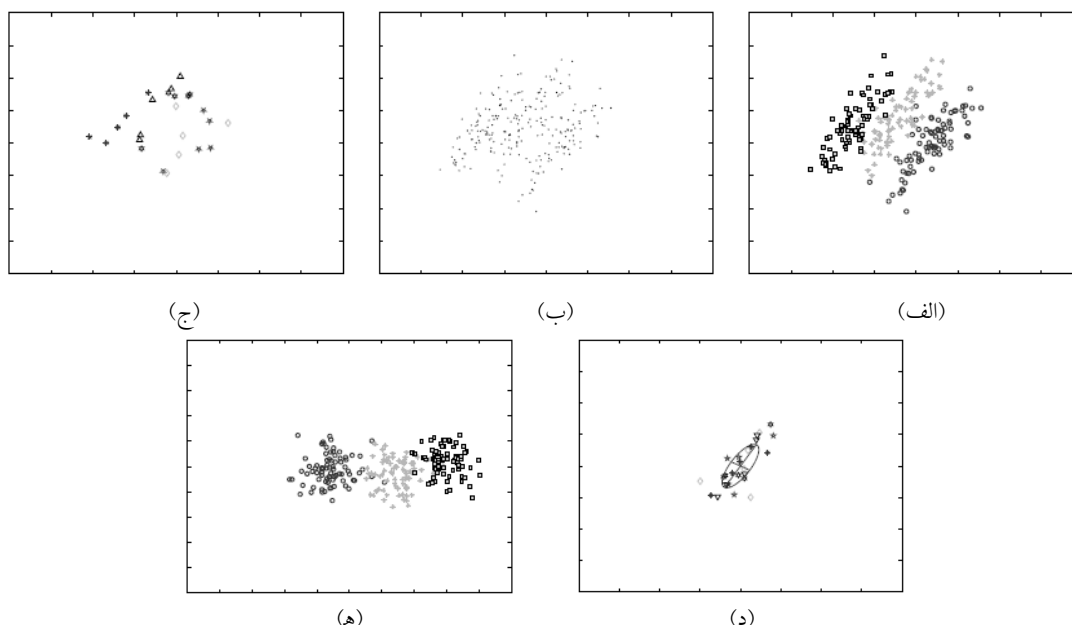
<sup>28</sup> Covariance

<sup>29</sup> Mahalanobis

<sup>30</sup> Principal Component Analysis

<sup>31</sup> Linear Discriminate Analysis

<sup>32</sup> Second-order

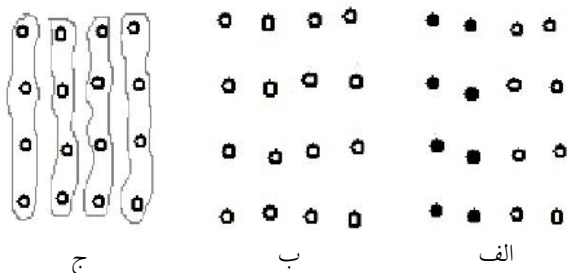


شکل ۳- تغییر فضای داده‌ها با اعمال الگوریتم RCA بر روی داده‌های ۳ طبقه با توزیع گوسی<sup>۳۴</sup>؛ الف) مجموعه داده اولیه؛ ب) حذف برچسب داده‌ها؛ ج) انتخاب دسته‌ها (نقاط یک رنگ و یک شکل یک دسته را نشان می‌دهند)؛ د) انتقال دسته‌ها به مبدا؛ ه) اعمال تبدیل RCA بر داده‌ها

با حذف مرحله ۳ و قرار دادن  $P=I$  (ماتریس واحد)،  $\hat{C}_{new}$  برابر با  $\hat{C}$  می‌شود و تبدیل RCA استاندارد به دست می‌آید. شکل ۳ مراحل الگوریتم RCA استاندارد را به خوبی نشان می‌دهد.

### ۲-۳-۲- به دست آوردن دسته‌ها

برای محاسبه تبدیل RCA باید کوواریانس دسته‌های موجود در داده‌ها را محاسبه کرد. روش‌های مختلفی برای پیدا کردن دسته‌ها در داده‌ها پیشنهاد شده است. بیشتر این روش‌ها داده‌ها را به صورت بدون برچسب در نظر می‌گیرند و سپس با استفاده از روش‌های یادگیری بدون ناظر، دسته‌های موجود در داده‌ها را می‌یابند. تمام داده‌های موجود در یک دسته باید از یک طبقه در نظر گرفته شوند (شکل ۴).



شکل ۴- الف) داده‌های برچسب‌دار؛ ب) داده‌های بدون در نظر گرفتن برچسب؛ ج) دسته‌های موجود در داده‌ها

همچنین می‌توان از روش RCA، مانند PCA، برای کاهش تعداد ویژگی‌های داده‌ها استفاده کرد. در این قسمت الگوریتم RCA به صورت قدم به قدم بر روی مجموعه داده‌های آموزشی به شکل  $\{(X_i, C_i) \mid i = 1, \dots, n\}$  که در آن هر الگو بردار  $X_i = [x_{i1}, x_{i2}, \dots, x_{im}]^T$  با  $m$  ویژگی و  $M$  طبقه بردار  $C_i = [1, 2, \dots, M]$  نشان داده شده است:

- در اولین قدم میانگین هر دسته محاسبه و داده‌های موجود در آن دسته به اندازه میانگین انتقال می‌یابند (تفاضل میانگین هر دسته، از داده‌های موجود در آن دسته)؛
- محاسبه ماتریس کوواریانس داده‌های موجود در دسته‌ها و به دست آوردن ماتریس  $\hat{C}$  با استفاده از رابطه (۳)؛

$$\hat{C} = \frac{1}{P} \sum_{j=1}^k \sum_{i=1}^{n_j} (x_{j,i} - \hat{m}_j)(x_{j,i} - \hat{m}_j)^T \quad (3)$$

در این رابطه  $p$  داده در  $k$  دسته قرار گرفته‌اند و دسته  $j$ ام شامل داده‌های  $\{x_{j,i}\}_{i=1}^{n_j}$  با میانگین  $\hat{m}_j$  است؛

۳. محاسبه مقادیر ویژه  $\hat{C}$  و به دست آوردن ماتریس تبدیل  $P$  با انتخاب  $r$  بردار ویژه  $\hat{C}$  و به دست آوردن  $\hat{C}_{new}$  با استفاده از رابطه (۴)؛

$$\hat{C}_{new} = P^T \hat{C} P \quad (4)$$

۴. محاسبه تبدیل سفیدکننده<sup>۳۳</sup>، به صورت  $W_{RCA} = P \hat{C}_{new}^{-\frac{1}{2}}$  و اعمال آن بر روی کل داده‌ها  $(x_{new} = W_{RCA} x)$ .

<sup>33</sup> Whitening

<sup>34</sup> Gaussian distribution

مهم‌ترین پارامتریست که در این مرحله مقدار می‌گیرد (رابطه ۵).

$$AT = \frac{\sum_{i=1}^{n-1} \sum_{j=i+1}^n d(ag_i, ag_j)}{n(n-1)/2} \quad (5)$$

در این رابطه  $n$  تعداد آنتی‌ژن‌های آموزشی (داده‌های آموزشی)،  $ag_i$  و  $ag_j$  آمین و زامین آنتی‌ژن ورودی هستند و  $d(ag_i, ag_j)$  فاصله اقلیدسی (درجه شباهت) بین دو آنتی‌ژن است و در آن فاصله اقلیدسی الگوی  $Q$  با داده آموزشی  $P_i$  از طریق رابطه ۶ محاسبه می‌شود:

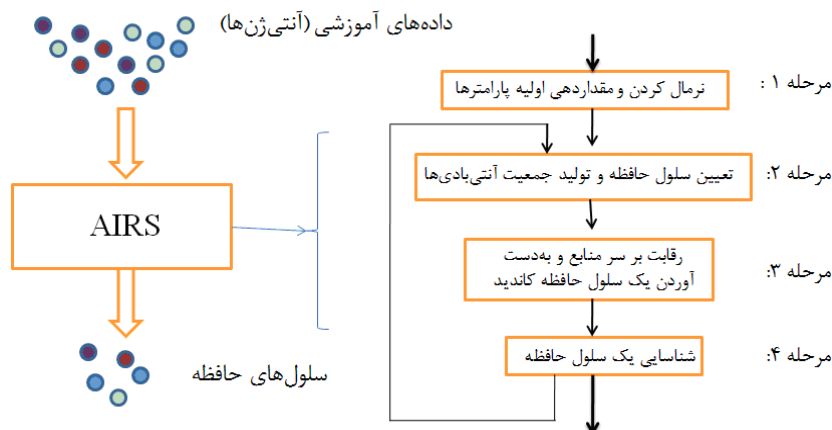
$$d(Q, P_i) = \sqrt{\sum_{k=1}^m (q_k - p_{ik})^2} \quad (6)$$

در این مرحله همچنین، جمعیت سلول‌های حافظه به صورت تصادفی با تعدادی از داده‌های آموزشی مقارنه می‌شود.

در الگوریتم AIRS داده‌های آموزشی به صورت متوالی وارد الگوریتم شده و مراحل ۲ تا ۴ بر اساس شکل ۵ اجرا می‌شوند. در مرحله دوم سلول حافظه‌ای که بیشترین درجه تحریک نسبت به آنتی‌ژن ورودی را دارد از طریق رابطه ۷ به دست می‌آید. این سلول حافظه، سلول حافظه تطبیق شده<sup>۳۵</sup> نامیده می‌شود و با  $mc_{match}$  نشان داده خواهد شد است.

$$Stimulation(ag_i, mc_j) = 1 - \frac{d(ag_i, mc_j)}{d_{max}} \quad (7)$$

در این رابطه  $ag_i$  آنتی‌ژن ورودی و  $mc_j$  یک سلول حافظه و  $d_{max}$  بیشترین فاصله بین دو آنتی‌ژن در فضای ویژگی هاست. بعد از تعیین  $mc_{match}$  با به‌کارگیری فرایند تکثیر و انجام جهش‌هایی متناسب با درجه شباهت مجموعه‌ای از آنتی‌بادی‌ها تولید می‌شود.



شکل ۵- الگوریتم AIRS

در این مطالعه با استفاده از یکی از الگوریتم‌های سیستم ایمنی مصنوعی به نام AD-AIRS، دسته‌های موجود در داده‌ها را می‌یابیم و الگوریتم AIS-RCA را ارائه می‌کنیم.

#### ۲-۴-۲- الگوریتم سیستم ایمنی مصنوعی AD-AIRS

AD-AIRS یکی از الگوریتم‌های سیستم ایمنی مصنوعی است. این الگوریتم نسخه‌ای از الگوریتم AIRS است که از نوعی معیار فاصله افقی استفاده می‌کند. AD-AIRS با الهام گرفتن از سیستم ایمنی بدن بردارهایی را به عنوان مراکز خوشه‌های درون هر طبقه پیدا می‌کند که در اصطلاح به آنها سلول‌های حافظه (یا الگوهای نماینده) گفته می‌شود. این الگوریتم سعی در پیدا کردن تعداد و مکان بهینه مراکز خوشه‌ها در درون هر کلاس دارد. قبل از ارائه الگوریتم AD-AIRS مراحل مختلف الگوریتم AIRS را که شرح کامل آن در مرجع [۳۱] آمده است، به‌طور خلاصه توضیح می‌دهیم.

#### ۲-۴-۱- الگوریتم AIRS

AIRS یکی از الگوریتم‌های سیستم ایمنی مصنوعی است که با الهام گرفتن از سیستم ایمنی طبیعی مجموعه کوچکی از سلول‌های حافظه (الگوهای نماینده) را از روی داده‌های آموزشی تولید می‌کند. شکل ۵ چارچوب کلی الگوریتم AIRS را نشان می‌دهد.

در مرحله اول الگوریتم ویژگی‌های مجموعه داده‌های آموزشی در بازه [۰، ۱] نرمال می‌شوند و به پارامترهای الگوریتم مقارنه اولیه می‌شود. آستانه شباهت<sup>۳۵</sup> ( $AT$ )

<sup>35</sup> Affinity Threshold

<sup>36</sup> Matched Memory Cell



### ۲-۴-۲- الگوریتم AD-AIRS

در الگوریتم AD-AIRS برای کم کردن حساسیت الگوریتم AIRS به معیار فاصله، نوعی معیار فاصله وفقی محلی در بدنه آن به کار برده شده است. همچنین با استفاده از مفهوم خوشه‌بندی، سلول‌های حافظه را به گونه‌ای تولید می‌کنیم که هر سلول حافظه نمایش دهنده تعدادی از داده‌های شبیه به هم از یک طبقه باشد (مانند یک خوشه دورن طبقه). به این ترتیب الگوریتم AIRS را بهبود بخشیده و سلول‌های حافظه در مکان‌های دقیق‌تری تولید می‌شوند و تعداد آنها نیز کاهش می‌یابد.

الگوریتم AD-AIRS از معیار فاصله وفقی که به وسیله ونگ<sup>۳۸</sup> [۴۴] معرفی شده و در آن با وزن‌دهی به هر یک از الگوهای آموزشی کارایی KNN بهبود یافته، بهره برده است. الگوریتم AD-AIRS مجموعه‌ای از سلول‌های حافظه‌ای وزن‌دار تولید می‌کند. برای وزن‌دهی به سلول حافظه  $(mc_i)$ ، فاصله آن تا نزدیک‌ترین الگوی دشمن (آنتی‌ژنی از کلاس مخالف) محاسبه می‌شود و وزن این سلول حافظه، معکوس این فاصله در نظر گرفته می‌شود. با استفاده از رابطه‌های (۸) و (۹) می‌توان وزن یک سلول حافظه را محاسبه کرد:

$$r_i = \min_k \{d(mc_i, ag_k) / class(mc_i) \neq class(ag_k)\} \quad (8)$$

$$w_i = 1 / r_i \quad (9)$$

با استفاده از این شیوه وزن‌دهی در مرحله تعیین سلول حافظه، تابع تحریک تغییر داده می‌شود به گونه‌ای که الگوهای نماینده در نزدیک مرزها، با احتمال کمتری به مجموعه سلول‌های حافظه اضافه شوند و رابطه (۷) به رابطه (۱۰) تغییر می‌یابد:

$$Weighted\_Stimulation(ag_k, mc_i) = 1 - \frac{w_i \cdot d(ag_k, mc_i)}{d_{max}} \quad (10)$$

الگوریتم AIRS با به کارگیری پارامتر ATS، برای هر سلول حافظه تأثیر یکسانی را در نظر می‌گیرد (ناحیه تحت پوشش ثابت). این در حالی است که الگوریتم AD-AIRS با در نظر گرفتن وزن برای هر یک از سلول‌های حافظه ناحیه تحت پوشش متغیری در نظر گرفته و به آنها درجه اهمیت متفاوتی می‌دهد و به این ترتیب پارامتر ATS از الگوریتم AD-AIRS حذف می‌شود. الگوریتم AD-AIRS برای مکان‌یابی سلول‌های حافظه از مفهوم خوشه‌بندی نیز استفاده

در مرحله سوم مجموعه آنتی‌بادی‌های تولید شده، تحت سازوکاری به نام رقابت بر سر منابع بهبود می‌یابند و نهایتاً یک آنتی‌بادی که نسبت به تحریک آنتی‌ژن ورودی حساس‌تر است بر اساس رابطه ۷، به عنوان سلول حافظه کاندید انتخاب می‌شود و آن را با  $mc_{candid}$  نمایش می‌دهیم.

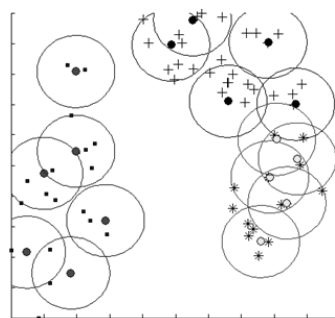
مرحله آخر الگوریتم AIRS، شامل تصمیم‌گیری برای اضافه کردن  $mc_{candid}$  و یا جایگزینی آن با  $mc_{match}$  در جمعیت سلول‌های حافظه است (شکل ۶). در روال شکل ۶، MC نشان دهنده جمعیت سلول‌های حافظه است. در این روال اگر شرط خط ۴ صحیح ارزیابی شود،  $mc_{candid}$  به مجموعه سلول‌های حافظه افزوده می‌شود. حال اگر  $mc_{match}$  در ناحیه تحت پوشش  $mc_{candid}$  باشد (خط ۶)، آن را از مجموعه سلول‌های حافظه حذف می‌کنیم. ناحیه تحت پوشش  $mc_{candid}$  یک آبرکره به مرکز  $mc_{candid}$  و شعاع  $AT \times ATS$  است. پارامتر آستانه شباهت عددی  $ATS$ <sup>۳۷</sup> توسط کاربر تعیین می‌شود ( $0 \leq ATS \leq 1$ )؛ اثرگذاری سلول حافظه ساخته شده را تنظیم می‌کند و اثر مهمی بر روی دقت و تعداد سلول‌های حافظه به دست آمده دارد. در شکل ۷، الگوریتم AIRS بر روی یک مجموعه داده سه طبقه با دو ویژگی اعمال و سلول‌های حافظه به دست آمده و ناحیه تحت پوشش هر سلول حافظه نشان داده شده است. مشاهده می‌شود تمام سلول‌های حافظه ناحیه یکسانی از فضای داده را تحت تأثیر قرار می‌دهند.

```

1. Cand-Stim ← stimulation(agi, mccandid)
2. Match-Stim ← stimulation(agi, mcmatch)
3. Cell-Aff ← d(mccandid, mcmatch)
4. if(Cand-Stim > Match-Stim)
{
5. MC ← MC ∪ mccandid
6. if(Cell-Aff < AT × ATS)
7. remove mcmatch from MC
}

```

شکل ۶- روال شناسایی سلول حافظه در الگوریتم AIRS



شکل ۷- نتیجه اجرای الگوریتم AIRS و سلول‌های حافظه به دست آمده با شعاع همسایگی ثابت

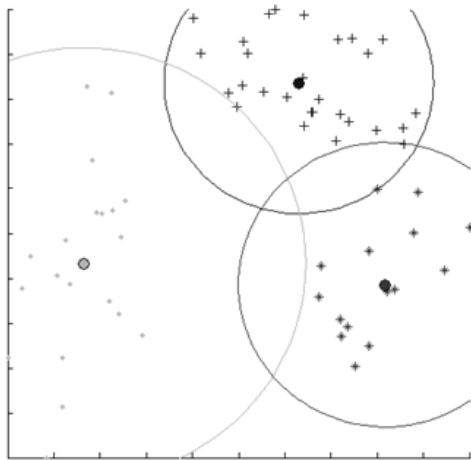
<sup>37</sup> Affinity Threshold Scalar

<sup>38</sup> Wang

می‌کند. مرکز هر خوشه نمایش دهنده مجموعه داده‌های موجود در آن خوشه است. یک الگوریتم خوشه‌بندی (مانند K- میانگین<sup>۳۹</sup>) با اضافه شدن هر عنصر جدید به خوشه، مرکز آن خوشه را به‌روز می‌کند. در الگوریتم AD-AIRS هر سلول حافظه نشان‌دهنده مرکز خوشه‌ای از داده‌های متعلق به آن خوشه است. هنگامی که الگوی جدید  $X_i$  به خوشه  $C_n$  با  $n$  عنصر اضافه می‌شود، مرکز خوشه با استفاده از رابطه (۱۱) به‌روز می‌شود:

$$C_{new} = \frac{n}{n+1} C_n + \frac{1}{n+1} X_i \quad (11)$$

روال الگوریتم AD-AIRS در شکل ۸ به صورت کامل نشان داده شده است. در ابتدا جمعیت سلول‌های حافظه به صورت تصادفی با تعدادی از آنتی‌ژن‌های ورودی مقداردهی اولیه می‌شوند و هر سلول حافظه در ابتدا نشان دهنده خوشه‌ای با یک عضو است ( $K(mc_i)=1, i=1,2,\dots,n$ ). پس از مراحل ۲ تا ۱۰ و پیدا کردن سلول‌های  $mc_{match}$  و  $mc_{candid}$  در خط ۱۱، درجه تحریک وزن‌دار این دو سلول حافظه با هم مقایسه می‌شود و در شرط بعدی در صورت قرارگیری  $mc_{candid}$  در ناحیه تحت پوشش  $mc_{match}$ ؛ مرکز خوشه جدید محاسبه شده؛ به جمعیت سلول‌های حافظه اضافه می‌شود؛

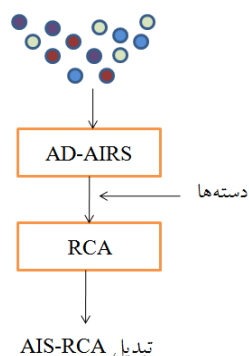


شکل ۹- سلول‌های حافظه به‌دست آمده توسط الگوریتم AD-AIRS

<p><b>Inputs:</b> Training instances or training antigens <math>AG=\{ag_i, i=1,2,\dots,m\}</math>  <b>Outputs:</b> Memory cell population <math>MC=\{mc_i, i=1,2,\dots,p\}</math></p> <ol style="list-style-type: none"> <li>1. Initialize in random the memory cell population with <math>n</math> antigens <math>MC=\{mc_i, i=1,2,\dots,n, n \leq m\}</math> and set <math>K(mc_i)=1, i=1,2,\dots,n</math>.</li> <li>2. <b>for</b> each antigen <math>ag_k</math></li> <li>3.     find <math>mc_{match} = \arg \max_{1 \leq l \leq l} \{weighted\_stimulation(mc_l, ag_k)\}</math>, where <math>l</math> is the number of memory cells</li> <li>4.     find <math>r_{match}</math> using equation (8)</li> <li>5.     generate antibodies by cloning and mutation of <math>mc_{match}</math></li> <li>6.     find <math>mc_{candid}</math> from the generated antibodies according to step 3 of AIRS algorithm</li> <li>7.     find radius <math>r_{candid}</math> using equation (8)</li> <li>8.     Cand-Stim <math>\leftarrow</math> weighted_stimulation(<math>ag_k, mc_{candid}</math>)</li> <li>9.     Match-Stim <math>\leftarrow</math> weighted_stimulation(<math>ag_k, mc_{match}</math>)</li> <li>10.    CellAff <math>\leftarrow</math> d(<math>mc_{candid}, mc_{match}</math>)</li> <li>11.    <b>if</b> (Cand-Stim &gt; Match-Stim)</li> <li>12.       <b>if</b> (Cell-Aff &lt; <math>r_{match}</math>)</li> <li>13.           <math display="block">mc_{new} = \frac{K(mc_{match})}{K(mc_{match})+1} mc_{match} + \frac{1}{K(mc_{match})+1} mc_{candid}</math></li> <li>14.           <math>K(mc_{new})=K(mc_{match})+1</math></li> <li>15.           remove <math>mc_{match}</math> from <math>MC</math></li> <li>16.           <math>MC \leftarrow MC \cup mc_{new}</math></li> <li>17.           find radius <math>r_{new}</math> using equation (8)</li> <li>18.       <b>else</b></li> <li>19.           <math>MC \leftarrow MC \cup mc_{candid}</math></li> <li>20.    <b>return</b> <math>MC</math></li> </ol>
--

شکل ۸- روال شناسایی سلول حافظه در الگوریتم AD-AIRS

<sup>39</sup> K-means



شکل ۱۰- الگوریتم AIS-RCA

جدول ۴- مشخصات مجموعه داده‌ها

تعداد نمونه‌ها	تعداد ویژگی‌ها	تعداد کلاس	
۵۰۰	۶۶	۲	مسئله ۱ Z, O, N, F (طبقه ۱) S (طبقه ۲)
۲۰۰	۶۶	۲	مسئله ۲ Z (طبقه ۱) S (طبقه ۲)

مشخصات مجموعه داده‌های به دست آمده در جدول ۴ نمایش داده شده است. برای دستیابی به دقت طبقه‌بندی کننده از روش ده‌قسمتی<sup>۴۱</sup> استفاده شده است و دقت‌های ارائه شده با ۵ بار تکرار این عمل به دست آمده‌اند.

برای آزمایش الگوریتم ارائه شده، در ابتدای کار باید هر یک از ویژگی‌ها را در بازه صفر و یک نرمالیزه کرد (زیرا در طبقه‌بندی‌کننده نهایی از معیار فاصله استفاده می‌شود) و الگوریتم AD-AIRS را اعمال نمود. پس از پایان اجرای الگوریتم AD-AIRS، دسته‌های موجود در داده‌های آموزشی تعیین شده و از آنها برای محاسبه تبدیل RCA استفاده می‌شود. در مرحله بعدی تبدیل RCA بر روی داده‌ها اعمال می‌شود. تبدیل RCA با حذف و کم کردن اثر برخی از ویژگی‌ها، فضای داده‌ها را به گونه‌ای نمایش می‌دهد که ویژگی‌های نامربوط در معیار فاصله مورد استفاده در فضای جدید اثر کمتری دارند. در پایان با استفاده از طبقه‌بندی‌کننده KNN بردارهای ویژگی به دست آمده را طبقه‌بندی می‌کنیم. آنجا که تبدیل RCA وزن‌دهی به فضای ویژگی‌ها را انجام می‌دهد؛ می‌توان انتظار داشت به ازای  $K=1$  نتیجه بهینه تولید شود و نیازی به یافتن  $K$  بهینه نیست.

## ۲-۵- الگوریتم AIS-RCA

الگوریتم AIS-RCA با استفاده از روش RCA و به کارگیری الگوریتم AD-AIRS، راهی برای بهبود فضای ویژگی‌ها و وزن‌دهی به ویژگی‌ها ارائه کرده است. این الگوریتم چارچوب کلی الگوریتم RCA را دارد با این تفاوت که برای پیدا کردن دسته‌ها در داده‌ها از الگوریتم AD-AIRS استفاده می‌کند.

معمولاً روش‌هایی که برای پیدا کردن دسته‌ها در داده‌ها استفاده می‌شوند برچسب داده‌ها را نادیده می‌گیرند اما AD-AIRS با در نظر گرفتن طبقه داده‌ها، به سراغ پیدا کردن دسته‌های موجود در کلاس‌ها می‌رود و با ساختن سلول‌های حافظه که نماینده گروهی از داده‌ها هستند از اطلاع برچسب داده‌ها نیز استفاده می‌کند. هر سلول حافظه محدوده‌ای به شکل آبرکره ایجاد می‌کند که تعدادی از داده‌ها در آن قرار می‌گیرند. در این الگوریتم دسته‌ها همان داده‌هایی هستند که در محدوده هر سلول حافظه قرار می‌گیرند. شکل ۱۰ چارچوب کلی الگوریتم AIS-RCA را نشان می‌دهد. در این شکل پس از اعمال الگوریتم AD-AIRS بر روی داده‌های آموزشی، دسته‌های موجود در داده‌های آموزشی جستجو می‌شوند. این دسته‌ها به الگوریتم RCA داده می‌شوند تا ماتریس تبدیل مورد نیاز ساخته شود. بعد از اعمال تبدیل به دست آمده بر روی داده‌ها، طبقه‌بندی داده‌ها با استفاده از طبقه‌بندی‌کننده ۱NN<sup>۴۲</sup>، کار طبقه‌بندی داده‌ها انجام می‌شود.

## ۳- بحث و نتایج

### ۳-۱- نتایج

برای ارزیابی و بررسی روش ارائه شده، از دو مسئله طبقه‌بندی که در تشخیص خودکار حملات صرع تعریف شده‌اند؛ استفاده شده است. مجموعه داده‌های مورد استفاده از ۵۰۰ قطعه سیگنال EEG تشکیل شده است که در قالب ۵ مجموعه Z, O, N, F, S دسته‌بندی شده‌اند. با پردازش زمان-فرکانس؛ به دست آوردن صفحه زمان-فرکانس هر قطعه از سیگنال‌ها و تقسیم‌بندی صفحه زمان-فرکانس برای هر قطعه سیگنال، ۶۵ ویژگی استخراج می‌شود که به همراه انرژی کل سیگنال، یک بردار به همراه ۶۶ ویژگی به دست می‌آید.

<sup>40</sup> 1- Nearest Neighbor

<sup>41</sup> 10-fold cross validation

جدیدی برای استخراج ویژگی‌ها ارائه کرده‌اند و برخی دیگر بر بهبود الگوریتم طبقه‌بندی استفاده شده تأکید می‌کنند. در این مطالعه روش نوینی برای بهبود ویژگی‌های استخراج شده از سیگنال‌ها ارائه شده است و سپس با استفاده از یک طبقه‌بندی کننده ساده به دقت‌های خوبی دست یافته‌ایم.

در جدول ۷ مقایسه‌ای بین روش ارائه شده و چند روش دیگر که در مجلات مختلف منتشر شده‌اند انجام گرفته است. از آنجا که در تمام این مطالعات آزمایش‌ها بر روی مجموعه داده‌های یکسانی انجام شده است مقایسه‌های ارائه شده قابل قبول‌اند. در مسئله طبقه‌بندی Z-S، روش ارائه شده در این مقاله به دقت ۹۹/۹٪ دست یافته است. البته قابل توجه است که در سطر دوم جدول ۷ با استفاده از الگوریتم طبقه‌بندی ۱NN و بدون اعمال الگوریتم AIS-RCA به دقت ۱۰۰٪ رسیده‌ایم که به هم بسیار نزدیک‌اند (در آزمایش‌های انجام شده با به‌کارگیری الگوریتم ۱NN، از بردارهای داده با ۶۵ ویژگی استفاده شده است). زالاس و همکاران بعد از کاهش ویژگی‌ها و استفاده از طبقه‌بندی کننده ANN<sup>۴۶</sup> به دقت ۱۰۰٪ دست یافته‌اند [۲۳]. با توجه به جدول ۷ روش ارائه شده در مسئله طبقه‌بندی Z-S به بالاترین دقت دست یافته است.

در مسئله طبقه‌بندی دوم (Z, O, N, F-S) نتیجه به‌دست آمده با روش ارائه شده در این مقاله (سطر ۱۲) بسیار نزدیک به بالاترین دقت‌های به‌دست آمده در سایر کارهای انجام شده است.

نکته مهم در تشخیص خودکار بیماری‌ها درصد حساسیت<sup>۴۲</sup> بالای طبقه‌بندی کننده است، حتی اگر درصد تشخیص نادرست<sup>۴۳</sup> بالا باشد. به همین دلیل علاوه بر دقت کل<sup>۴۴</sup> (رابطه ۱۲)، نتایج حساسیت و خصوصیت<sup>۴۵</sup> با استفاده از روابط (۱۳) و (۱۴) محاسبه شده است.

$$(12) \text{ دقت کل}_i = \frac{\text{تعداد الگوهای طبقه } i \text{ که در طبقه } i \text{ دسته‌بندی شده‌اند}}{\text{تعداد کل الگوها}}$$

$$(13) \text{ حساسیت}_i = \frac{\text{تعداد الگوهای طبقه } i \text{ که در طبقه } i \text{ دسته‌بندی شده‌اند}}{\text{تعداد الگوهای طبقه } i}$$

$$(14) \text{ خصوصیت}_i = \frac{\text{تعداد الگوهایی که در طبقه } i \text{ نیستند و در طبقه } i \text{ دسته‌بندی نشده‌اند}}{\text{تعداد الگوهایی که در طبقه } i \text{ نیستند}}$$

جداول ۵ و ۶ نتایج آزمایش‌های انجام شده را در دو مسئله طبقه‌بندی تشخیص خودکار حملات صرع نشان می‌دهند. در این جداول علاوه بر دقت کل، حساسیت، خصوصیت، انحراف معیار و تعداد اولیه و تعداد کاهش یافته ویژگی‌ها بعد از اعمال تبدیل RCA ارائه شده‌اند. طی آزمایش‌های انجام شده مشاهده شد که بهترین نتایج با تعداد ۸ ویژگی به‌دست می‌آید.

### ۲-۳- مقایسه‌ای با سایر کارهای انجام شده

تاکنون روش‌های گوناگونی برای حل مسئله تشخیص خودکار حملات صرع ارائه شده‌اند. برخی از آنها روش

جدول ۵- نتایج به‌دست آمده مسئله طبقه‌بندی ZONF-S

(در ستون دقت کل مقادیر به ترتیب از چپ به راست دقت به درصد، انحراف معیار و تعداد ویژگی‌های کاهش یافته به کل ویژگی‌ها هستند)

دقت کل	حساسیت		خصوصیت		مسئله ۱
	Z, O, N, F	S	Z, O, N, F	S	
۹۶/۳۱ (±۰/۰۲۳) ۸/۶۶	۹۷/۲۱	۹۱/۴۰	۹۵/۴۴	۹۷/۱۸	Z, O, N, F (طبقه ۱) S (طبقه ۲)

جدول ۶- نتایج به‌دست آمده مسئله طبقه‌بندی Z-S

(در ستون دقت کل مقادیر به ترتیب از چپ به راست دقت به درصد، انحراف معیار و تعداد ویژگی‌های کاهش یافته به کل ویژگی‌ها هستند)

دقت کل	حساسیت		خصوصیت		مسئله ۲
	Z	S	Z	S	
۹۹/۹۰ (۰/۰۰۱) ۸/۶۶	۱۰۰	۹۹/۸۲	۹۹/۹۰	۱۰۰	Z (طبقه ۱) S (طبقه ۲)

<sup>47</sup> Sensitivity

<sup>43</sup> False Detection

<sup>44</sup> Total Accuracy

<sup>45</sup> Specificity

<sup>46</sup> Artificial Neural Network

جدول ۷- مقایسه کارهای انجام شده

مؤلفین	روش	مجموعه داده	دقت کل
این مطالعه	پردازش زمان فرکانس، ۵ پنجره زمانی $\times 13$ ناحیه فرکانسی با قابلیت تفکیک <sup>۴۷</sup> فرکانسی ۵۱۲، AIS-RCA	Z,S	۹۹/۹۰
این مطالعه	پردازش زمان فرکانس، ۵ پنجره زمانی $\times 13$ ناحیه فرکانسی با قابلیت تفکیک فرکانسی ۵۱۲، 1NN	Z,S	۱۰۰
نیگام <sup>۴۸</sup> و همکاران [۴۵]	پیش پردازش غیر خطی فیلتر، شبکه عصبی تشخیصی	Z,S	۹۷/۲
سرینواسام <sup>۴۹</sup> و همکاران [۳۳]	ویژگی های زمانی و فرکانسی، شبکه عصبی بازگشتی	Z,S	۹۹/۶
کاناتال <sup>۵۰</sup> و همکاران [۱۴]	معیار آنتروپی، سیستم استنتاجی عصبی - فازی وفقی (ANFIS)	Z,S	۹۲/۲۲
کاناتال و همکاران [۱۱]	معیار بی نظمی <sup>۵۱</sup> ، پردازش داده جایگزین <sup>۵۲</sup>	Z,S	۹۰~
پولات <sup>۵۳</sup> و همکاران [۱۵]	تبدیل فوریه سریع، درخت تصمیم	Z,S	۹۸/۷۲
سوباسی <sup>۵۴</sup> [۲]	تبدیل موجک گسسته، سیستم خبره ترکیبی	Z,S	۹۵
زالاس و همکاران [۲۳]	پردازش زمان فرکانس، ۵ پنجره زمانی $\times 13$ ناحیه فرکانسی با قابلیت تفکیک فرکانسی ۵۱۲، شبکه عصبی مصنوعی	Z,S	۱۰۰
جنو <sup>۵۵</sup> و همکاران [۳۴]	تبدیل موجک گسسته و انرژی نسبی موجک، MLPNN <sup>۵۶</sup>	Z,S	۹۵/۲
جنو و همکاران [۳۵]	تبدیل موجک گسسته - ویژگی طول خط <sup>۵۷</sup> ، MLPNN	Z,S	۹۹/۶
این مطالعه	پردازش زمان فرکانس، ۵ پنجره زمانی $\times 13$ ناحیه فرکانسی با قابلیت تفکیک فرکانسی ۵۱۲، AIS-RCA	(Z,O,N,F),S	۹۶/۳۱
این مطالعه	پردازش زمان فرکانس، ۵ پنجره زمانی $\times 13$ ناحیه فرکانسی با قابلیت تفکیک فرکانسی ۵۱۲، 1NN	(Z,O,N,F),S	۹۴/۳
زالاس و همکاران [۲۳]	پردازش زمان فرکانس، ۵ پنجره زمانی $\times 13$ ناحیه فرکانسی با قابلیت تفکیک فرکانسی ۵۱۲، شبکه عصبی مصنوعی	(Z,O,N,F),S	۹۳/۲۰
زالاس و همکاران [۲۳]	پردازش زمان فرکانس، ۳ پنجره زمانی $\times 5$ ناحیه فرکانسی با قابلیت تفکیک فرکانسی ۵۱۲، شبکه عصبی مصنوعی	(Z,O,N,F),S	۹۷/۷۳
جنو و همکاران [۳۵]	تبدیل موجک گسسته - ویژگی طول خط، MLPNN	(Z,O,N,F),S	۹۷/۷۵

الگوریتم AIS-RCA ترکیبی از الگوریتم RCA و AD-AIRS است. RCA تبدیلی است که با استفاده از دسته های موجود در داده ها کار وزن دهی به ویژگی ها و کاهش ویژگی ها را انجام می دهد. در این مقاله برای اولین بار الگوریتم AD-AIRS ارائه شده و از آن برای یافتن دسته ها استفاده شده است. AD-AIRS الگوریتمی تکاملیست که از سیستم ایمنی بدن الهام گرفته است و از روی داده ها ورودی الگوهای جدیدی به نام سلول های حافظه تولید می کند. این سلول ها جایی ساخته می شوند که تجمع داده ها بیشتر باشد. برای نشان دادن توانایی الگوریتم AIS-RCA، از آن برای بهبود ویژگی های استخراج شده از سیگنال های EEG استفاده شده است. استخراج ویژگی ها به کمک پردازش زمان-فرکانس و تقسیم بندی صفحه زمان-فرکانس قطعات سیگنال های EEG، انجام شده است و به این ترتیب برای هر

در این مسئله 1NN نتوانسته است به دقت خوبی دست یابد. از آن جا که در روش پیشنهادی در این مقاله از ۵ پنجره زمانی و 13 ناحیه فرکانسی استفاده شده نتیجه آزمایش های زالاس و همکاران [۲۳] در این حالت نیز گزارش شده است. آنها با تغییر تعداد پنجره های زمانی و نواحی فرکانسی به دقت 97/73٪ دست یافته اند. با توجه به دقت های ارائه شده در جدول ۷ و مقایسه با سایر کارهای انجام شده، روش ارائه شده در این مقاله در هر دو مسئله طبقه بندی Z-S و (N, F)-S، (Z, O)، به دقت های بسیار نزدیک به بالاترین نتایج به دست آمده توسط سایر تحقیقات دست یافته است.

#### ۴- نتیجه گیری

در این مقاله با ارائه الگوریتم جدید AIS-RCA طبقه بندی سیگنال های EEG شامل الگوهای صرع انجام شده است.

<sup>47</sup> Resolution

<sup>51</sup> Chaotic measure

<sup>55</sup> Guo

<sup>48</sup> Nigam

<sup>52</sup> Surrogate data analysis

<sup>56</sup> Multi-Layer Perceptron Neural Network

<sup>49</sup> Srinivasan

<sup>53</sup> Polat

<sup>57</sup> Line length feature

<sup>50</sup> Kannathal

<sup>54</sup> Subasi

- a nonlinear viewpoint; *Computer Methods and Programs in Biomedicine*, 2005; 79(2): 151-159.
- [14] Kannathal N., Acharya U.R., Lim C.M., Sadasivan P.K., Characterization of EEG—a comparative study; *Computer Methods and Programs in Biomedicine*, 2005; 80(1): 17-23.
- [15] Polat K., Güneş S., Classification of epileptic form EEG using a hybrid system based on decision tree classifier and fast Fourier transform; *Applied Mathematics and Computation*, 2007; 187(2): 1017-1026.
- [16] Mohseni H., Maghsoudi A., Kadbi M., Hashemi J., Ashourvan A., Automatic detection of epileptic seizure using time-frequency distributions; In: *IET 3rd International Conference on Advances in Medical Signal and Information Processing MEDSIP*, 2006: 1-4.
- [17] Gardner A.B., Krieger A.M., Vachtsevanos G., Litt B., One-class novelty detection for seizure analysis from intracranial EEG; *Journal of Machine Learning Research*, 2006, 7: 1025-1044.
- [18] Übeyli E.D., Support vector machines for detection of electrocardiographic changes in partial epileptic patients; *Engineering Applications of Artificial Intelligence*, 2008; 21(8): 1196-1203.
- [19] Temko A., Thomas E., Marnane W., Lightbody G., Boylan G., EEG-based neonatal seizure detection with Support Vector Machines; *Clinical Neurophysiology*, 2011; 122(3): 464-473.
- [20] Boashash B., Mesbah M., Golditz P., Time frequency detection of EEG abnormalities; in *Time-Frequency Signal Analysis and Processing: A Comprehensive Reference*, 2003; 15: 663-670.
- [21] Hassanpour H., Mesbah M., Boashash B., Time frequency based newborn EEG seizure detection using low and high frequency signatures; *Physiological Measurement*, 2004; 25(4): 935-944.
- [22] Boashash B., Mesbah M., Time-frequency methodology for newborn electroencephalographic seizure detection; in *Applications in Time-Frequency Signal Processing*, 2003; 9: 339-369.
- [23] Tzallas A.T., Tsipouras M.G., Fotiadis D.I., Automatic Seizure Detection Based on Time-Frequency Analysis and Artificial Neural Networks; *Computational Intelligence and Neuroscience*, 2007: 1-13.
- [24] He X., King O., Ma W.Y., Li M., Zhang H. J., Learning a semantic space from user's relevance feedback for image retrieval; *IEEE Trans on Circuits and Systems for Video Technology*, 2003; 13(1): 39-48.
- [25] Muller H., Pun T., Squire D., Learning from user behavior in image retrieval: Application of market basket analysis; *International Journal of Computer Vision*, 2004; 56(1-2): 65-77.
- [26] Hastie T., Tibshirani R., Discriminant adaptive nearest neighbor classification; *IEEE Pattern Analysis and Machine Intelligence*, 1996; 18(6): 607-616.
- [27] Tenenbaum J., Freeman W., Separating Style and content with bilinear models; *neural computation*, 2000; 12(6): 1247-1283.
- [28] Bar-Hillel A., Hertz T., Shental N., Weinshall D., Learning Distance Functions using Equivalence Relations; in: *Proceedings of the Twentieth International Conference on Machine Learning*, 2003: 11-18.
- [29] Watkins A., Timmis J., Artificial Immune Recognition System (AIRS): Revisions and Refinements; *International Conference on Artificial Immune Systems*, 2002: 173-181.

قطعه سیگنال، ۶۶ ویژگی استخراج شده است. AIS-RCA با جستجوی دسته‌ها در بردارهای به‌دست آمده، تبدیلی را به‌دست می‌آورد که ضمن وزن‌دهی به ویژگی‌ها، باعث کاهش تعداد ویژگی‌ها نیز می‌شود. در مرحله نهایی به کمک طبقه‌بندی کننده ۱NN تشخیص سیگنال‌های افراد سالم و بیمار صورت می‌گیرد.

## مراجع

- [1] Subasi A., Alkan A., Koklukaya E., Kiyimik M.K., Wavelet neural network classification of EEG signals by using AR model with MLE preprocessing; *Neural Networks*, 2005; 18(7): 985-997.
- [2] Subasi A., EEG signal classification using wavelet feature extraction and a mixture of expert model; *Expert Systems with Applications*, 2007; 32(4): 1084-1093.
- [3] Sadati N., Mohseni H.R., Maghsoudi A., Epileptic seizure detection using neural fuzzy networks; in *Proceedings of IEEE International Conference on Fuzzy Systems (FUZZY '06)*, 2006: 596-600.
- [4] Guo L., Rivero D., Pazos A., Epileptic seizure detection using multiwavelet transform based approximate entropy and artificial neural networks; *Journal of Neuroscience Methods*, 2010; 193(1): 156-163.
- [5] Tzallas A.T., Karvelis P.S., Katsis C.D., Fotiadis D.I., Giannopoulos S., Konitsiotis S., A method for classification of transient events in EEG recordings: application to epilepsy diagnosis; *Methods of Information in Medicine*, 2006; 45(6): 610-621.
- [6] Argoud F.I.M., de Azevedo F.M., Neto J.M., Grillo E., SADE3: an effective system for automated detection of epileptic form events in long-term EEG based on context information; *Medical and Biological Engineering and Computing*, 2006; 44(6): 459-470.
- [7] Murro A.M., King D.W., Smith J.R., Gallagher B.B., Flanigin H.F., Meador K., Computerized seizure detection of complex partial seizures; *Electroencephalography and Clinical Neurophysiology*, 1991; 79(4): 330-333.
- [8] Lerner D.E., Monitoring changing dynamics with correlation integrals: case study of an epileptic seizure; *Physica D: Nonlinear Phenomena*, 1996; 97(4): 563-576.
- [9] Lehnertz K., Elger C.E., Spatio-temporal dynamics of the primary epileptogenic area in temporal lobe epilepsy characterized by neuronal complexity loss; *Electroencephalography and Clinical Neurophysiology*, 1995; 95(2): 108-117.
- [10] Güler N.F., Ubeyli E.D., Güler I., Recurrent neural networks employing Lyapunov exponents for EEG signals classification; *Expert Systems with Applications*, 2005; 29(3): 506-514.
- [11] Kannathal N., Choo M.L., Acharya U.R., Sadasivan P.K., Entropies for detection of epilepsy in EEG; *Computer Methods and Programs in Biomedicine*, 2005; 80(3): 187-194.
- [12] Ocak H., Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy; *Expert Systems with Applications*, 2009; 36(2): 2027-2036.
- [13] Paivinen N., Lammi S., Pitkanen A., Nissinen J., Penttonen M., Gronfors T., Epileptic seizure detection:

- determination in intracranial EEG; *Clinical Neurophysiology*, 2008; 119 (12): 2687–2696.
- [38] Mohseni H.R., Maghsoudi A., Shamsollahi M.B., Seizure Detection in EEG signals: A Comparison of Different approaches; *Engineering in Medicine and Biology Society EMBS '06*, 2006: 6724 – 6727.
- [39] Allen R.L., Mills D.W., *Signal Analysis: Time, Frequency, Scale, and Structure*; Wiley-IEEE Press 2004.
- [40] Auger F., Flandrin P., Goncalv'es P., Lemoine O., *Time-Frequency Toolbox*; Rice University and CNRS France, 1996.
- [41] Shental N., Hertz T., Weinshall D., Pavel M., Adjustment learning and relevant component analysis; in *Proc. of the European Conference on Computer Vision*, 2002:776–792.
- [42] BarHillel A., Hertz T., Shental N., Weinshall D., Learning distance functions using equivalence relations; in *Proc. International Conference on Machine Learning*, 2003: 11-18.
- [43] Shental N., Hertz T., Weinshall D., Pavel M., Adjustment Learning and Relevant Component Analysis; *Proceedings of the 7th European Conference on Computer Vision*, 2002: 776-792.
- [44] Wang J., Neskovic P., Cooper L., Improving nearest neighbor rule with a simple adaptive distance measure; *Pattern Recognition Letters*, 2007; 28(2): 207–213.
- [45] Nigam V., Graupe D., A neural-network-based detection of epilepsy; *Neurological Research*, 2004; 26(1): 55–60.
- [30] Watkins A., AIRS: A resource limited artificial immune classifier; M.S. thesis, 2001.
- [31] Watkins A., Boggess L., A new classifier based on resource limited artificial immune systems; *Proceedings of the Evolutionary Computation*, 2002; 2: 1546-1551.
- [32] Andrzejak R.G., Lehnertz K., Mormann F., Rieke C., David P., Elger C.E., Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: dependence on recording region and brain state; *Physical Review E*, 2001; 64(6): 8-23.
- [33] Srinivasan V., Eswaran C., Sriraam A.N., Artificial neural network based epileptic detection using time-domain and frequency-domain features; *Journal of Medical Systems*, 2005; 29(6): 647–660.
- [34] Guo L., Rivero D., Seoane J., Pazos A., Classification of EEG signals using relative wavelet energy and artificial neural networks; In: *Proceedings of the first ACM/SIGEVO Summit on Genetic and Evolutionary Computation*, 2009: 177–84.
- [35] Guo L., Rivero D., Dorado J., Rabunal J.R., Pazos A., Automatic epileptic seizure detection in EEGs based on line length feature and artificial neural networks; *Journal of Neuroscience Methods*, 2010; 191(1): 101–109.
- [36] Sivasankari N., Thanushkodi K., Automated Epileptic Seizure Detection in EEG Signals Using FastICA and Neural Network; *Advanced Soft Computing and Application*, 2009; 1(2): 91-104.
- [37] Chan M., Sun T., Boto H., Wingeier M., Automated seizure onset detection for accurate onset time