

آزمون دروغ‌سنجی مبتنی بر تلفیق اطلاعات آشوبناک کانال‌های سیگنال الکتروانسفالوگرام

سکینه رضوی^۱، امین جانقربانی^{۲*}، محمدباقر خدابخشی^۳

*نویسنده مسئول، دریافت: ۱۴۰۰/۰۵/۱۰، بازنگری: ۱۴۰۰/۰۵/۲۹، پذیرش: ۱۴۰۰/۰۶/۰۶

^۱ دانشجوی کارشناسی ارشد، دانشکده علوم و فناوری‌های نوین، دانشگاه سمنان، سمنان، ایران

^۲ استادیار، دانشکده علوم و فناوری‌های نوین، دانشگاه سمنان، سمنان، ایران

^۳ استادیار، گروه مهندسی پزشکی، دانشگاه صنعتی همدان، همدان، ایران

چکیده

در سال‌های اخیر استفاده از مؤلفه P300 در آزمون دانش‌گناهکار، از روش‌های پرکاربرد به شمار می‌رود. در مطالعات مختلف، کانال Pz به عنوان کانال حاوی بیشترین اطلاعات مربوط به این مؤلفه شناخته شده است. با این وجود، پژوهش‌های دیگر نشان داده‌اند که کانال‌های Cz و Fz نیز اطلاعات مفیدی از مؤلفه P300 ارائه می‌دهند. بنابراین، حل چالش استفاده توأمان و بهینه از اطلاعات این سه کانال می‌تواند در بهبود نتایج حاصل از آزمون دروغ‌سنجی مؤثر واقع شود. در این مقاله به منظور استخراج اطلاعات مرتبط با مؤلفه P300، از تجزیه تحلیل کمی بازرخداد سیگنال الکتروانسفالوگرام بهره گرفته شد. از سوی دیگر به منظور تلفیق اطلاعات آشوبناک این سه کانال، از رویکردهای تلفیق اطلاعات در سطح ویژگی و در سطح تصمیم‌گیری استفاده شد. برای تلفیق اطلاعات در سطح ویژگی دو روش (۱) تلفیق تمامی ویژگی‌های سه کانال و تشکیل یک بردار کلی و (۲) انتخاب ویژگی‌های بهینه از بردار مذکور با استفاده از الگوریتم ژنتیک مدنظر قرار گرفت. به منظور تلفیق اطلاعات در سطح تصمیم‌گیری نیز احتمال پسین وزن‌دار هر کلاس بر اساس بردارهای ویژگی و قابلیت اطمینان هر کانال محاسبه شد و برای تشخیص افراد گناهکار و بی‌گناه مورد استفاده قرار گرفت. از میان رویکردهای پیشنهادی، رویکرد تلفیق اطلاعات در سطح تصمیم‌گیری با صحت ۹۰ درصد، حساسیت ۸۶/۶۷ درصد و ویژه بودن ۹۳/۳۳ درصد نشان‌دهنده برتری این روش در قیاس با رویکردهای پیشنهادی دیگر است. علاوه بر این، سرعت اجرای پردازش‌های مربوط به این روش نسبت به روش مبتنی بر الگوریتم ژنتیک بسیار بالاتر است.

کلمات کلیدی: آزمون دروغ‌سنجی، سیگنال الکتروانسفالوگرام، مؤلفه P300، پردازش آشوبناک، تلفیق اطلاعات، الگوریتم ژنتیک

۱- مقدمه

تحقیقات در مورد توانایی کشف فریب طی چند دهه گذشته توجه زیادی را به خود جلب کرده و کاربردهای حقوقی، اخلاقی و بالینی مهمی دارد [۱]. در بسیاری از مطالعات از واکنش‌های فیزیولوژیکی به عنوان معیارهای شناسایی فریب استفاده می‌کنند. از جمله این معیارها می‌توان به معیارهایی همانند نرخ تنفس، ضربان قلب، فشار خون و پاسخ هدایت الکتریکی پوست^۱ اشاره کرد [۲]. رویکرد دیگر در فرآیند تشخیص دروغ، پایش بدون واسطه عملکرد مغز به منظور تشخیص فریب و فرد فریبکار است [۳،۴]. بر اساس این رویکرد در دهه

اخیر تعدادی از سیگنال‌های نوروفیزیولوژیک برای کاربرد احتمالی در تشخیص فریب مورد بررسی قرار گرفته است. از جمله این رویکردها می‌توان به تصویربرداری تشدید مغناطیسی کارکردی^۲ و پتانسیل‌های وابسته به رخداد^۳ اشاره کرد [۴].

۱-۱- آزمون دانش‌گناهکار مبتنی بر الکتروانسفالوگرام

به طور کلی، برخی از محرک‌ها دارای قابلیت ایجاد تغییرات گذرا، معین و قابل تشخیص در سیگنال الکتروانسفالوگرام^۴ هستند. برخی از تغییرات به یک محرک بیرونی وابسته هستند و برخی از آن‌ها بدون وجود محرک بیرونی و در

فرض مهم در روش‌های میانگین‌گیری و آشکارسازی P300 آن است که این مؤلفه‌ها در تک ثبت‌های مختلف ثابت هستند. اما تأخیر، دامنه و ظاهر آن‌ها می‌تواند به طرز چشمگیری در میانگین‌گیری از تک ثبت‌های مختلف، متفاوت باشد. این امر ناشی از متفاوت بودن ویژگی‌های ذکر شده P300 در تک ثبت‌های متفاوت است. در نتیجه، روش میانگین‌گیری می‌تواند منجر به از دست رفتن اطلاعات مذکور شود. علاوه بر این، پیاده‌سازی این روش نیازمند اعمال تعداد زیادی محرک است و از این رو باید زمان بیشتری برای ثبت سیگنال اختصاص داده شود که این امر باعث افزایش خستگی افراد می‌شود [۱۱]. علی‌رغم این که با استفاده از روش‌های بوت‌استرپ تفاوت دامنه و بوت‌استرپ تفاوت ضرایب همبستگی، نسبت سیگنال به نویز مؤلفه P300 قابل افزایش است، اما برخی از افراد گناهکار دارای مؤلفه P300 ضعیفی هستند؛ از این رو تشخیص آن‌ها به عنوان گناهکار یا بی‌گناه دشوار است. بنابراین، منطقی است که از روش‌های مبتنی بر پردازش تک ثبت برای بررسی وجود یا عدم P300 استفاده شود.

در روش پردازش تک ثبت، برای تشخیص فرد گناهکار از بی‌گناه دو رویکرد وجود دارد. یک رویکرد این است که فرآیند تشخیص فرد گناهکار و بی‌گناه بر اساس سیگنال پروب افراد و به روش اعتبارسنجی یک سوژه بیرون^{۱۸} انجام می‌شود. در این رویکرد آموزش طبقه‌بند بر اساس سیگنال پروب تمامی سوژه‌ها بجز یک سوژه انجام می‌شود. سپس آزمون طبقه‌بند با سیگنال پروب سوژه‌ای که در فرآیند آموزش طبقه‌بند استفاده نشده، صورت می‌گیرد [۵]. رویکرد دیگر، تشخیص مبتنی بر سوژه است و طبقه‌بند با سیگنال هدف و غیرمرتبط هر سوژه آموزش می‌بیند و آزمون طبقه‌بند با سیگنال پروب همان سوژه انجام می‌شود. با این رویکرد تنوع فردی پاسخ‌های P300 در فرآیند تشخیص لحاظ می‌شود و سنجش هر فرد بر اساس سیگنال‌های دارای P300 و سیگنال‌های بدون P300 ثبت شده از خود فرد است [۱۲].

چالش اصلی در به دست آوردن تخمین قابل اعتماد از P300ها در سطح پردازش تک ثبت، افزایش نسبت سیگنال به نویز پاسخ‌های P300 به طور مؤثر و مطمئن است. با این حال، این تکنیک توانسته تا حد زیادی موفق عمل کند [۱۳]. بوتا^{۱۹} و همکارانش با استفاده از تحلیل طیف سیگنال نزدیک به مادون قرمز^{۲۰} به صورت تک ثبت به دروغ‌سنجی پرداختند [۱۴]. گائو و همکارانش، به منظور بررسی وجود اختلاف قابل توجه در مقادیر هم‌دوسی موجک^{۲۱} از تک ثبت‌های بخش‌های مختلف مغز در آزمون دروغ‌سنجی استفاده کردند [۱۵]. از دیگر روش‌های تجزیه و تحلیل تک ثبت مؤلفه P300 می‌توان به روش‌های تجزیه زمان-فرکانس، مانند تبدیل موجک و تجزیه حالت عملکردی [۱۶، ۱۷]، اشاره کرد. همچنین از روش پردازش تک ثبت، برای آشکارسازی مؤلفه‌های پتانسیل وابسته به رخداد استفاده می‌شود [۱۸، ۱۹].

از آنجا که P300 یک مؤلفه معین در سیگنال الکتروانسفالوگرام است، وقوع P300 منجر به افزایش تعین و کاهش پیچیدگی سیگنال‌های مغز می‌شود. این بدان معنی است که هنگام ظهور مؤلفه P300 ماهیت آشوبناک سیگنال الکتروانسفالوگرام کاهش می‌یابد؛ در نهایت پیچیدگی سیگنال کاهش یافته که بیانگر تغییرات تعین و دینامیک غیرخطی سیستم است [۵، ۲۰-۲۲]. از دیگر روش‌های پردازش سیگنال به صورت تک ثبت استفاده از روش‌های پردازش مبتنی بر دینامیک غیرخطی سیگنال است. به طور کلی، سیگنال‌های زیستی غیر ایستان، دارای دینامیک آشوبناک، غیرخطی و نویزی هستند. غیر ایستان بودن سیگنال الکتروانسفالوگرام به این معنی است که مشخصات آماری سیگنال با گذشت زمان تغییر می‌کند و فعالیت مغز اساساً غیر ثابت است. بنابراین تجزیه و تحلیل سیگنال الکتروانسفالوگرام با روش‌های پردازشی آشوبناک^{۲۲} می‌تواند مفید واقع شود. یکی از روش‌های تجزیه و تحلیل داده‌های غیرخطی بر اساس بررسی تراژکتوری بازسازی شده در فضای فاز^{۲۳} بر اساس

طی فعالیت عادی در مغز پدیدار می‌شوند. چنین تغییراتی که وابسته به یک تحریک خارجی هستند به عنوان پتانسیل‌های وابسته به رخداد، شناخته می‌شوند. پتانسیل‌های وابسته به رخداد دارای مؤلفه‌های مختلفی است که از بین تمامی آن‌ها، P300 دارای جایگاه ویژه‌ای در آزمون دروغ‌سنجی است. مؤلفه P300 یک موج بالارونده در سیگنال الکتروانسفالوگرام است. این مؤلفه حاصل یک فرآیند طبیعی در مغز است که به دنبال رخداد خاصی مانند مواجهه با یک محرک غیرمعمول در میان تعداد زیادی محرک معمول، در مغز ظاهر می‌شود.

به سازوکار کلی ثبت P300، الگوی تحریک غیرمعمول^۵ می‌گویند. در این سازوکار چندین تحریک به صورت مکرر به مغز اعمال می‌شود که یکی از آن‌ها غیرمعمول است و با احتمال کمتری رخ می‌دهد. انتظار می‌رود که در هنگام مواجهه‌ی فرد با این تحریک، مؤلفه P300 ظاهر شود.

آزمون دانش گناهکار^۶ یک تکنیک مبتنی بر اندازه‌گیری و بررسی تغییرات سیگنال الکتروانسفالوگرام در مواجهه با اطلاعات مربوط به جرم است که می‌تواند در تشخیص دانش پنهان افراد گناهکار استفاده شود. مؤلفه P300 کارایی بالایی از خود در این آزمون نشان داده است [۵-۷]. در آزمون دروغ‌سنجی، سه نوع محرک به فرد ارائه می‌شود:

- محرک پروب^۷: این محرک مربوط به صحنه جرم است که تنها افراد گناهکار آن را می‌شناسند و افراد بی‌گناه آن را نمی‌شناسند. در فرآیند دروغ‌سنجی، محرک پروب برای افراد گناهکار، محرک غیرمعمول و برای افراد بی‌گناه محرک معمول است.
- محرک هدف^۸: هر دو گروه گناهکار و بی‌گناه این محرک را می‌شناسند و از آن‌ها خواسته می‌شود با مشاهده آن واکنش از پیش تعیین‌شده‌ای نشان دهند. در فرآیند دروغ‌سنجی، محرک هدف یک محرک غیرمعمول است و انتظار می‌رود سیگنال الکتروانسفالوگرام ثبت‌شده از اعمال این محرک دارای مؤلفه P300 باشد.
- محرک غیرمرتبط^۹: این محرک به صحنه جرم مربوط نیست و هر دو گروه افراد بی‌گناه و گناهکار با این محرک آشنا نیستند. در فرآیند دروغ‌سنجی محرک غیرمرتبط یک محرک معمول است و انتظار می‌رود سیگنال الکتروانسفالوگرام ثبت‌شده از اعمال آن مؤلفه P300 نداشته باشد.

۱-۲- روش‌های پردازش سیگنال در دروغ‌سنجی

دامنه P300 اغلب کوچک‌تر از اندازه سیگنال الکتروانسفالوگرام پس‌زمینه است. بنابراین شناسایی و توصیف P300، به روش‌های افزایش نسبت سیگنال به نویز^{۱۱} وابسته است. روش‌های مختلفی برای افزایش نسبت سیگنال به نویز و آشکارسازی مؤلفه P300 وجود دارد. یکی از رایج‌ترین روش‌ها برای این امر، میانگین‌گیری سنکرون از تعداد زیادی تک ثبت ناشی از اعمال محرک‌های همجنس به سوژه است [۸]. روش بوت‌استرپ تفاوت دامنه^{۱۱} و بوت‌استرپ تفاوت ضرایب همبستگی^{۱۲} از دیگر روش‌های پردازش سیگنال در حوزه دروغ‌سنجی هستند که عملکرد آن‌ها بر اساس میانگین‌گیری از تک ثبت‌های همجنس است [۹]. علاوه بر این گائو^{۱۳} و همکارانش، با استفاده از روش میانگین‌گیری و استخراج ویژگی‌های ریخت‌شناسی^{۱۴}، ویژگی‌های فرکانسی، ویژگی‌های زمان-فرکانسی مبتنی بر تبدیل موجک^{۱۵} به تشخیص افراد گناهکار و بی‌گناه پرداختند [۷]. همچنین سائینی^{۱۶} و همکارانش در سال ۲۰۱۹، از ترکیب‌های مختلف ۴۰ ویژگی استخراج شده از میانگین دادگان [۷]، مانند ویژگی‌های ریخت‌شناسی، ویژگی‌های فرکانسی، ویژگی‌های زمان-فرکانسی مبتنی بر تبدیل موجک و تجزیه حالت عملکردی^{۱۷}، برای طبقه‌بندی افراد گناهکار و بی‌گناه استفاده کردند [۱۰].

با یکدیگر تلفیق شده و یک بردار کلی را تشکیل دادند و در رویکرد دوم به منظور انتخاب ویژگی‌های بهینه از بردار ویژگی مذکور، از الگوریتم ژنتیک استفاده شد. همچنین برای تلفیق اطلاعات در سطح تصمیم‌گیری، احتمال پسین وزن‌دار هر کلاس بر اساس بردارهای ویژگی و قابلیت اطمینان هر کانال مورد استفاده قرار گرفت.

در ادامه این مقاله، در بخش ۲ به کارهای مرتبط پرداخته شده است. سپس، در بخش ۳ روش پیشنهادی بیان شده است. در بخش ۴ روش استخراج ویژگی بر اساس نگاشت بازرخداد، روش‌های تلفیق در سطح ویژگی، انتخاب ویژگی و تلفیق در سطح تصمیم‌گیری شرح داده خواهد شد. بخش ۵ این مقاله به بررسی روش طبقه‌بندی‌های مورد استفاده و اعتبارسنجی اختصاص یافته است. بخش ۶ مربوط به بررسی و تحلیل نتایج کمی است و بخش ۷ به جمع‌بندی مقاله و معرفی پژوهش‌های آتی خواهد پرداخت.

۲- کارهای مرتبط

همانطور که بیان شد، به منظور تلفیق اطلاعات می‌توان از روش‌های تلفیق در سطح ویژگی و تصمیم‌گیری استفاده کرد. در روش تلفیق در سطح ویژگی، ابتدا ویژگی‌های دادگان به صورت مستقل از هر کانال استخراج می‌شوند. سپس، ویژگی‌های استخراج شده با هم ترکیب و یک مجموعه ویژگی جدید را برای استفاده در مرحله طبقه‌بندی می‌سازند. از روش‌های تلفیق ویژگی می‌توان به بهم پیوستن^{۳۱} ویژگی‌های استخراج شده [۳۳] و محاسبه میانگین مکانی از ویژگی‌ها اشاره کرد (رابطه (۱)) [۳۴]:

$$F[n] = \frac{1}{N_c} \sum_{p=1}^{N_c} F_p[n] \quad (1)$$

که N_c تعداد کانال‌ها، $F_p[n]$ ویژگی n ام استخراج شده از کانال p ام است. علیرغم اینکه تلفیق ویژگی‌ها با میانگین‌گیری منجر به کاهش بُعد ویژگی‌ها می‌شود، اما ممکن است با این روش برخی از تغییرات حاوی اطلاعات مفید سیگنال از بین بروند.

لیو^{۳۲} و همکارانش از روش تلفیق ویژگی‌ها به منظور بهبود طبقه‌بندی حالات خستگی استفاده کردند [۳۵]. ساراوانان^{۳۳} از این روش برای تلفیق انرژی‌های زیرباند های آلفا، بتا و گاما استخراج شده از دادگان الکتروانسفالوگرام استفاده کرد. سپس به تشخیص احساسات، تشخیص صرع و تشخیص حالات خواب پرداخت [۳۶]. ابوطالبی و همکارانش با استفاده از روش تلفیق در سطح ویژگی، ویژگی‌های ریخت شناسی، فرکانسی و زمان-فرکانسی مبتنی بر تبدیل موجک استخراج شده از دادگان دروغ‌سنجی را با یکدیگر تلفیق و از الگوریتم ژنتیک به منظور انتخاب ویژگی‌های بهینه از بردار ویژگی تلفیقی استفاده کردند. سپس با طبقه‌بند آنالیز افتراقی خطی^{۳۴} به تشخیص فرد گناهکار از بی‌گناه پرداختند [۳۷].

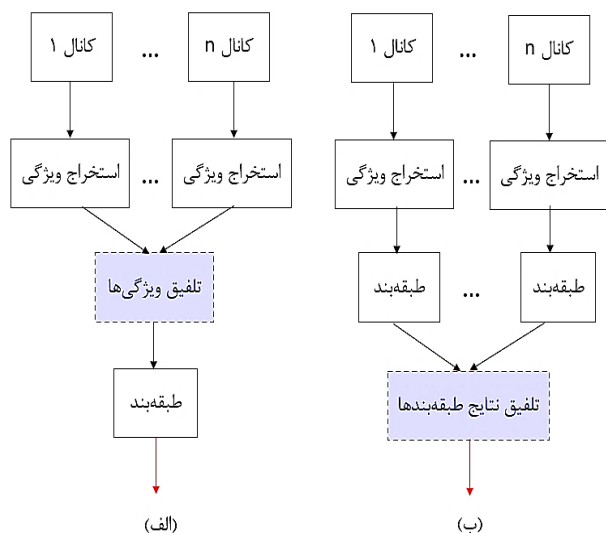
علاوه بر این مهرانم و همکارانش با استفاده از این روش تلفیق اطلاعات به تلفیق ویژگی‌های نگاشت بازرخداد، ریخت شناسی، فرکانسی و زمان-فرکانسی مبتنی بر تبدیل موجک استخراج شده از دادگان دروغ‌سنجی پرداختند. در این مقاله از الگوریتم ژنتیک به منظور انتخاب ویژگی‌های بهینه از بردار ویژگی تلفیقی استفاده و با طبقه‌بند آنالیز افتراقی خطی به طبقه‌بندی افراد گناهکار و بی‌گناه پرداختند [۵]. همچنین آراسته و همکارانش از این روش تلفیق اطلاعات، برای تلفیق ویژگی‌های استخراج شده از دادگان دروغ‌سنجی ثبت شده از کانال‌های Pz، Cz و Fz استفاده کردند. سپس، با استفاده از الگوریتم ژنتیک به عنوان روش انتخاب ویژگی‌های بهینه از مجموعه ویژگی تلفیقی به دست آمده به طبقه‌بندی افراد گناهکار و بی‌گناه پرداختند [۱۷].

نگاشت بازرخداد^{۳۴} انجام می‌شود [۲۱]. نگاشت بازرخداد یک ابزار گرافیکی برای مطالعه سیستم‌های دینامیکی غیرخطی است که اولین بار توسط اکمن^{۳۵} در سال ۱۹۸۷ ارائه شد [۲۳]. این نگاشت فاصله بین حالات تراژکتوری یک سیستم دینامیکی را در زمان‌های مختلف ارزیابی و بر اساس این فاصله، حالات وابسته به یکدیگر تعیین می‌شود. نگاشت بازرخداد و تجزیه و تحلیل کمی بازرخداد^{۳۶} در چندین مطالعه فرآیند سیگنال الکتروانسفالوگرام، مانند بازیابی حافظه [۲۴]، تشخیص مؤلفه‌های پتانسیل وابسته به رخداد [۲۵]، تشخیص احساسات [۲۶، ۲۷]، دروغ‌سنجی [۲۰] و تشخیص صرع [۲۸] با موفقیت بکار گرفته شده‌اند.

۳-۱- چالش و انگیزه پژوهش

برخی مطالعات نشان می‌دهند که مؤلفه P300 دارای دو مؤلفه فرعی P3a و P3b (P300 کلاسیک) است. مؤلفه فرعی P3a در طی فرآیند تمرکز توجه^{۳۷} تولید می‌شود [۲۹]. این مؤلفه فرعی در ناحیه پیشانی-مرکزی دارای حداکثر دامنه است. اگر فعال‌سازی منابع توجه باعث افزایش عملیات حافظه شود، P3b به عنوان زیرمؤلفه دیگر به همراه P3a ظاهر می‌شود. این مؤلفه بازتاب فعالیت مغز در حین استفاده از حافظه است که حداکثر دامنه را در ناحیه جداری^{۳۸} دارد [۳۰، ۳۱]. به طور کلی، رویدادهای غیر معمول باعث ایجاد P3b می‌شوند و هرچه احتمال رخ دادن این رویدادها کمتر باشد دامنه P3b بزرگتر خواهد شد.

بر اساس این شواهد، تمرکز سوژه روی محرک پروب می‌تواند نشانه‌هایی در رابطه با گناهکاری یا بی‌گناهی فرد ایجاد کند. بنابراین استفاده از کانال‌های Fz و Cz که در آن دامنه P3a حداکثر است، به علاوه کانال Pz که در آن دامنه P3b حداکثر است، مفید می‌باشد و می‌تواند منجر به افزایش صحت طبقه‌بندی و قابلیت اطمینان سیستم دروغ‌سنجی شود [۱۷]. ترکیب این دادگان می‌تواند بر اساس تلفیق اطلاعات صورت پذیرد. از روش‌های تلفیق اطلاعات می‌توان به تلفیق در سطح ویژگی^{۳۹} و تلفیق در سطح تصمیم‌گیری^{۳۰} اشاره کرد (شکل ۱) [۳۲].



شکل ۱- نمودار روش‌های مختلف تلفیق کانال، (الف) تلفیق در سطح ویژگی و (ب) تلفیق در سطح تصمیم‌گیری.

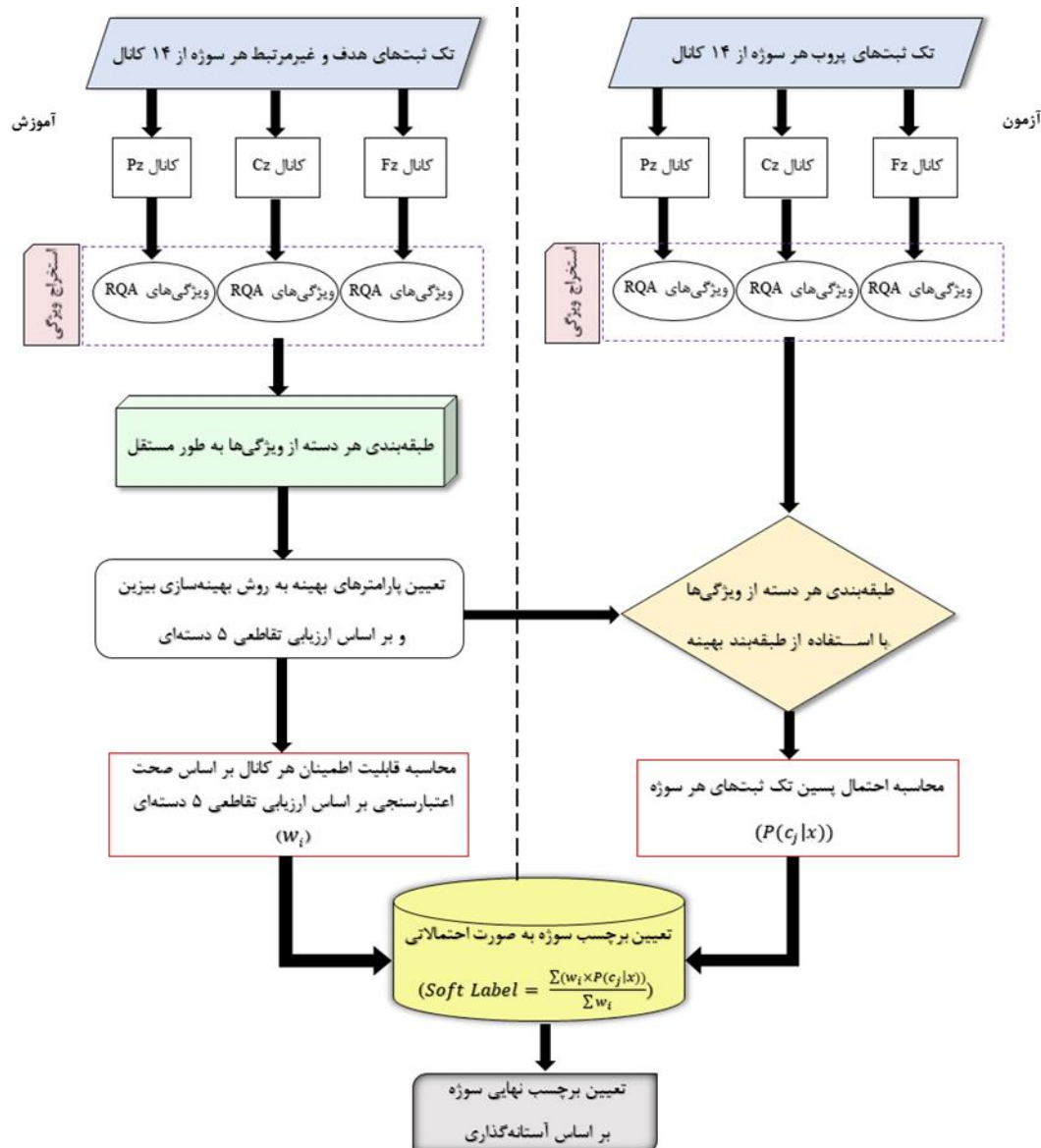
در این مقاله برای تشخیص فرد گناهکار از بی‌گناه و استفاده توأمان از اطلاعات سه کانال Pz، Cz و Fz، از روش‌های تلفیق در سطح ویژگی و تلفیق در سطح تصمیم‌گیری استفاده شد. به منظور تلفیق در سطح ویژگی، دو رویکرد مورد استفاده قرار گرفت. در رویکرد اول تمامی ویژگی‌های حاصل از سه کانال

دادگان و تشخیص فرد گناهکار از بی‌گناه شود. در این مقاله از کانال‌های Pz، Cz و Fz و تلفیق آن‌ها در سطح ویژگی و تلفیق اطلاعات در سطح تصمیم‌گیری استفاده شد. به منظور استخراج ویژگی، روش تجزیه و تحلیل نگاشت بازخداد به صورت تک ثبت مورد استفاده قرار گرفت. برای بررسی این ایده در گام نخست، ویژگی‌های نگاشت بازخداد به طور مستقل از دادگان هر کانال Pz، Cz و Fz استخراج و طبقه‌بندی شدند. نتیجه‌ی حاصل از این طبقه‌بندی که مبتنی بر سوژه انجام شد برای تشخیص افراد گناهکار و بی‌گناه مورد استفاده قرار گرفت. در گام بعدی ابتدا از روش تلفیق در سطح ویژگی به منظور تلفیق ویژگی‌های تجزیه و تحلیل کمی بازخداد حاصل از سه کانال ذکر شده، استفاده شد. سپس این بردار ویژگی تلفیقی برای طبقه‌بندی افراد گناهکار و بی‌گناه بکار رفت. در رویکرد بعدی به منظور انتخاب ویژگی‌های بهینه از بردار ویژگی تلفیقی ذکر شده، الگوریتم ژنتیک مورد استفاده قرار گرفت. سپس ویژگی‌های انتخاب شده با استفاده از الگوریتم ژنتیک، طبقه‌بندی شدند. در رویکرد چهارم که بلوک دیاگرام آن در شکل ۲ نمایش داده شده است، از روش تلفیق ویژگی در سطح تصمیم‌گیری استفاده شد.

در روش تلفیق در سطح تصمیم‌گیری، هر یک از بردارهای ویژگی به طور مستقل توسط طبقه‌بند مربوطه طبقه‌بندی شده سپس با استفاده از خروجی طبقه‌بندها نتیجه نهایی تولید می‌شود. نحوه تولید خروجی نهایی در این روش براساس قوانین ترکیب طبقه‌بندها شامل میانگین، حداکثر، حداقل، ضرب، جمع و رأی اکثریت است [۳۳،۳۸]. بالاگیشنان^{۳۵} و همکارانش از روش تلفیق در سطح تصمیم‌گیری با استفاده از قانون جمع در ترکیب طبقه‌بندها، برای تشخیص احساسات استفاده کردند [۳۳]. همچنین کائور^{۳۶} و همکارانش به منظور بررسی تأثیر افزایش سن در سیستم‌های بیومتریک مبتنی بر سیگنال الکتروانسفالوگرام از روش تلفیق در سطح تصمیم‌گیری مبتنی بر ترکیب طبقه‌بندها با استفاده از قانون رأی اکثریت پرداختند [۳۹]. همچنین مصباح^{۳۷} و همکارانش از روش تلفیق در سطح تصمیم‌گیری مبتنی بر ترکیب طبقه‌بندها با استفاده از قوانین میانگین، حداکثر، حداقل، ضرب، جمع و رأی اکثریت به تشخیص صرع پرداختند [۴۰].

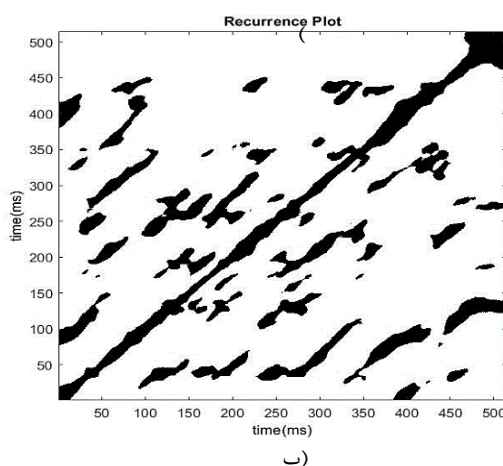
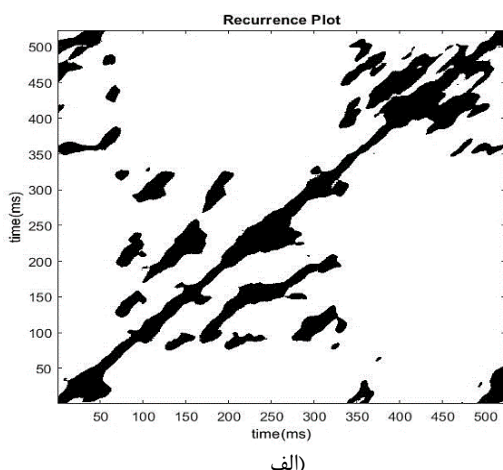
۳- روش پیشنهادی

همانطور که گفته شد، با توجه به اهمیت کانال‌های Pz، Cz و Fz انتظار می‌رود تلفیق اطلاعات این کانال‌ها بتواند منجر به بهبود صحت طبقه‌بندی



شکل ۲- بلوک دیاگرام روش تلفیق در سطح تصمیم‌گیری.

که N تعداد نقاط اندازه‌گیری شده \vec{x}_i ، ε آستانه فاصله، $\theta(\cdot)$ تابع هوی‌ساید^{۴۲} و $\|\cdot\|$ یک نرم^{۴۳} است. نگاشت بازخداد با ترسیم ماتریس بازخداد حالات سیستم به دست می‌آید. این نگاشت به صورت بصری اطلاعات درباره‌ی پویایی و دینامیک سیستم را با استفاده از رنگ‌های سفید و سیاه نمایش می‌دهد؛ به عنوان مثال برای مختصات (i, j) اگر $R_{i,j} \equiv 1$ یک نقطه سیاه است و در صورتی که $R_{i,j} \equiv 0$ یک نقطه سفید است. هر دو محور نگاشت بازخداد محورهای زمانی هستند. این نگاشت همیشه یک خط اصلی مورب سیاه، منطبق بر نیمساز ربع اول صفحه مختصات دارد. علاوه بر این، نگاشت بازخداد نسبت به این نیمساز متقارن است، یعنی $R_{i,j} \equiv R_{j,i}$ [۲۱]. شکل ۳ نشان‌دهنده نگاشت بازخداد برای تک ثبت هدف و غیرمرتبط را نشان می‌دهد.



شکل ۳- نگاشت بازخداد برای تک ثبت‌های (الف) هدف و (ب) غیرمرتبط.

۴-۳- تجزیه و تحلیل کمی بازخداد

به منظور کمی‌سازی اطلاعات بصری بازنمایی شده به وسیله نگاشت بازخداد، چندین معیار کمی مبتنی بر ساختارهای مقیاس کوچک موجود در نگاشت بازخداد معرفی شده‌اند که به عنوان تجزیه و تحلیل کمی بازخداد شناخته می‌شوند [۲۱، ۴۴]. این معیارها بر اساس تراکم نقاط بازخداد و ساختارهای خطی مورب و عمودی نگاشت بازخداد تعریف می‌شوند. در نگاشت بازخداد، خطوط مورب زمانی که بخشی از تراژکتوری در فضای فاز تقریباً به طور موازی با بخش دیگر حرکت می‌کند، ظاهر می‌شوند [۴۵]. بنابراین، خط مورب نشان‌دهنده بازخداد سیستم برای یک دوره است و طول این خط متناسب با مدت زمان بازخداد است. نگاشت بازخداد سیستم‌های آشوبناک، دارای ساختار خطوط مورب بیشتری نسبت به سیستم‌های تصادفی و کمتر نسبت به سیستم‌های معین هستند. تغییر طول این خطوط می‌تواند بیانگر تغییر شدت رفتار آشوبناک یک

براساس این رویکرد، مانند رویکرد اول، ابتدا ویژگی‌های تجزیه و تحلیل کمی نگاشت بازخداد از تک ثبت‌های دادگان هدف و غیرمرتبط هر سوژه از کانال‌های Pz، Cz و Fz استخراج و ویژگی‌های هر کانال به طور جداگانه طبقه‌بندی شدند. صحت اعتبارسنجی طبقه‌بندی ویژگی‌های هر کانال از طریق ارزیابی تقاطعی ۵ دسته‌ای محاسبه و به عنوان قابلیت اطمینان آن کانال در نظر گرفته شد. سپس از طریق میانگین‌گیری وزن‌دار بر اساس قابلیت اطمینان کانال‌ها و احتمال پسین به دست آمده برای دادگان آزمون، برچسب دادگان هر سوژه به صورت احتمالاتی محاسبه شد. در انتها برچسب نهایی هر سوژه از طریق آستانه‌گذاری تعیین شد.

۴- استخراج ویژگی

در این مقاله به منظور تشخیص فرد گناهکار و بی‌گناه رویکرد وابسته به سوژه و پردازش تک ثبت انتخاب گردید. به منظور افزایش نسبت سیگنال به نویز و استخراج اطلاعات مربوط به P300 در تک‌ثبت‌های سیگنال الکتروانسفالوگرام فرآیند استخراج ویژگی بر روی این سیگنال‌ها اعمال می‌شود. در این مقاله از روش‌های پردازش سیگنال آشوبناک به منظور استخراج ویژگی از کانال‌های Pz، Cz و Fz استفاده شد.

۴-۱- بازسازی فضای فاز

برای تجزیه و تحلیل یک سیگنال با استفاده از نظریه دینامیک غیرخطی، ابتدا باید فضای فاز سیگنال بازسازی شود. شناخته‌شده‌ترین روش برای بازسازی فضای فاز استفاده از روش تأخیر زمانی ارائه شده توسط پاکارد^{۴۸} و همکارانش است [۴۱] که توسط تیکنز^{۴۹} توسعه یافت [۴۲]. در این روش یک سری زمانی عددی $[u_1, u_2, \dots, u_L]$ در یک فضای m بُعدی جاسازی می‌شود (رابطه (۲)):

$$X_k = [u_k, u_{k+\tau}, \dots, u_{k+(m-1)\tau}] \quad (2)$$

که τ زمان تأخیر و m بُعد جاسازی و k تعداد نمونه‌های سری زمانی است. در بازسازی فضای فاز انتخاب مقادیر مناسب برای τ و m حائز اهمیت است. اگر τ خیلی کوچک باشد، حالات تراژکتوری در فضای فاز دارای اطلاعات تکراری خواهد بود و اگر τ خیلی بزرگ باشد، نقاط تراژکتوری غیرمرتبط خواهند بود. متداول‌ترین روش برای انتخاب یک تأخیر زمانی مناسب تشخیص اولین حداقل محلی تابع اطلاعات متقابل^{۴۰} است [۴۳]. همچنین در انتخاب بُعد جاسازی، اگر m خیلی کوچک باشد، الگوی تراژکتوری به طور کامل بازسازی نخواهد شد و اگر m خیلی بزرگ باشد، منجر به محاسبات اضافی غیرضروری خواهد شد. برای انتخاب یک بُعد جاسازی مناسب از روش نزدیک‌ترین همسایه کاذب^{۴۱} استفاده می‌شود [۴۳].

۴-۲- نگاشت بازخداد

اگر بُعد جاسازی سیگنال بیش از سه باشد، ترسیم فضای فاز آن عملاً غیرممکن است. در این شرایط نگاشت بازخداد ابزاری مفید است که فضای فاز را بر اساس همسایگی حالت‌ها و بازگشت تراژکتوری سیستم به حالت‌های هم‌جوار، به یک ماتریس قابل نمایش در صفحه دو بُعدی تبدیل می‌کند. بررسی نگاشت بازخداد، ویژگی‌های مرتبط با دینامیک سیستم را نشان می‌دهد و می‌تواند به پیش‌بینی رفتارهای آینده آن کمک کند. در واقع این نگاشت ابزاری است که می‌تواند بازخداد حالات یک تراژکتوری $\vec{x}_i \in \mathbb{R}^d$ را در فضای فاز نشان دهد. خروجی این روش به صورت نقاط منفرد، خطوط مورب و خطوط عمودی یا افقی بروز پیدا می‌کند. به ساختار حاصل از این خطوط و نقاط، در مقیاس بزرگ، بافت گفته می‌شود. ماتریس نگاشت بازخداد به صورت رابطه (۳) تعریف می‌شود:

$$R_{i,j}(\varepsilon) = \theta(\varepsilon - \|\vec{x}_i - \vec{x}_j\|), i, j = 1, \dots, N \quad (3)$$

که در آن $P(v)$ هیستوگرام از خطوط عمودی به طول v است.

- زمان به دام افتادن

زمان به دام افتادن به متوسط طول ساختارهای خطوط عمودی اشاره دارد که متوسط زمانی را که یک سیستم در یک حالت خاص بدون تغییر باقی مانده است، تخمین می‌زند (رابطه (۱۰)).

$$TT = \frac{\sum_{v=v_{min}}^N vP(v)}{\sum_{v=v_{min}}^N P(v)} \quad (10)$$

- حداکثر طول خط عمودی

حداکثر طول خط عمودی بیانگر طولانی‌ترین خط عمودی در نگاشت بازخداد است. این معیار را می‌توان به صورت رابطه (۱۱) توصیف کرد:

$$v_{max} = \max\{v_i\}_{i=1}^N \quad (11)$$

- زمان بازخداد نوع اول

زمان بازخداد نوع اول مقدار میانگین همه زمان‌های بازخداد را نشان می‌دهد (رابطه (۱۲)).

$$T_1 = \frac{1}{N} \sum_{i=1}^N T_i^{(1)} \quad (12)$$

که $T_i^{(1)}$ میانگین حداقل اختلاف زمانی بین نقاط در همسایگی یک نقطه i در تراژکتوری بازسازی شده می‌باشد [۲۲].

- زمان بازخداد نوع دوم

زمان بازخداد نوع دوم شامل اطلاعاتی در مورد فاصله زمانی بین بازخداد فعلی و شروع ساختار بازخداد بعدی است (رابطه (۱۳)).

$$T_2 = \frac{1}{N} \sum_{i=1}^N T_i^{(2)} \quad (13)$$

که $T_i^{(2)}$ میانگین حداقل اختلاف زمانی بین نقاط بازخداد در همسایگی نقطه i در تراژکتوری بازسازی شده، به استثنای تمام نقاط زمانی پی در پی است [۲۲].

- آنتروپی تراکم دوره بازخداد

آنتروپی تراکم دوره بازخداد حاوی اطلاعاتی در مورد ویژگی‌های تناوبی سیگنال در سیستم دینامیکی است. این اندازه‌گیری به ویژه برای تشخیص تکرارهای توالی یکسان یک سری زمانی در فضای فاز سیستم مناسب است (رابطه (۱۴)).

$$RPDE = \lim_{\epsilon \rightarrow \infty} \lim_{m \rightarrow \infty} \frac{1}{\tau} \ln \frac{c^m(\epsilon)}{c^{m+1}(\epsilon)} \quad (14)$$

- ضریب خوشه‌بندی

برای محاسبه معیار ضریب خوشه‌بندی ابتدا ضریب خوشه‌بندی محلی برای هر گره محاسبه می‌شود سپس، ضریب خوشه‌بندی، میانگین ضریب خوشه‌بندی محلی همه گره‌ها می‌باشد (رابطه (۱۵)).

$$Clust = \sum_{i=1}^N \frac{\sum_{j,k=1}^N A_{i,j} A_{j,k} A_{k,i}}{RR_i} \quad (15)$$

که $RR_i = \sum_{j=1}^N A_{i,j}$ ماتریس بازخداد محلی است. ماتریس مجاورت A که در تعاریف به کار رفته، ماتریس بازخداد است که ماتریس واحد از آن کم می‌شود.

- انتقال پذیری

انتقال پذیری یک شبکه پیچیده به این احتمال مربوط است که دو همسایه هر حالت نیز با هم همسایه هستند و این معیار نشان می‌دهد که یک شبکه چگونه به صورت محلی خوشه‌بندی شده‌است (رابطه (۱۶)).

سیستم با دینامیک غیرخطی باشد [۴۵]. همچنین ساختارهای خطوط عمودی بیانگر مدت زمانی هستند که در آن، حالت سیستم ثابت یا با تغییرات اندک است. همانطور که در این شکل ۳ مشاهده می‌شود، وجود مؤلفه $P300$ در سیگنال الکتروانسفالوگرام محرک هدف، در ساختار نگاشت بازخداد آن، باعث ایجاد خطوط مورب طولانی در حدود ۳۰۰ میلی‌ثانیه شده‌است. اما، در ساختار نگاشت بازخداد سیگنال محرک غیرمرتبط خطوط مورب کوتاه‌تری دیده می‌شود که بیانگر رفتار آشوبناک‌تر در سیستم است.

در این مقاله ۱۳ معیار تجزیه و تحلیل کمی بازخداد استفاده شده‌است. این ویژگی‌ها به شرح زیر است:

- نرخ بازخداد

نرخ بازخداد^{۴۴} که مربوط به تراکم نقاط بازخداد است بر اساس رابطه (۴) محاسبه می‌شود و درصد نقاط بازگشتی در نگاشت بازخداد را کمی می‌کند و میانگین تعداد همسایگان هر یک از عناصر فضای فاز را در همسایگی خود نشان می‌دهد.

$$RR = \frac{1}{N^2} \sum_{i,j=1}^N R_{i,j} \quad (4)$$

که N بُعد ماتریس بازخداد است.

- تعیین

برای محاسبه تعیین، نسبت تعداد نقاط بازخداد که به صورت خطوط مورب در این نگاشت تشکیل می‌شوند به تعداد کل نقاط بازخداد محاسبه می‌شود (رابطه (۵)).

$$DET = \frac{\sum_{l=l_{min}}^N lP(l)}{\sum_{l=1}^N lP(l)} \quad (5)$$

که در آن l_{min} حداقل طول خط مورب و $P(l)$ هیستوگرام خطوط مورب با طول l است.

- میانگین طول خط مورب

این معیار بیانگر متوسط زمانی است که دو بخش از تراژکتوری در فضای فاز نزدیک به هم حرکت می‌کنند [۲۱]. بنابراین می‌تواند به عنوان میانگین زمان پیش‌بینی پذیری یک سیستم آشوبناک معرفی شود. رابطه (۶) بیانگر روش به دست آوردن این ویژگی است.

$$L = \frac{\sum_{l=l_{min}}^N lP(l)}{\sum_{l=1}^N P(l)} \quad (6)$$

- حداکثر طول خط مورب

حداکثر طول خط مورب، بیانگر طولانی‌ترین خط مورب در نگاشت بازخداد است که به صورت رابطه (۷) بیان می‌شود:

$$L_{max} = \max\{L_i\}_{i=1}^N \quad (7)$$

که در آن N_l تعداد کل خطوط مورب است.

- آنتروپی خطوط مورب

این معیار پیچیدگی نگاشت بازخداد را در رابطه با خطوط مورب اندازه‌گیری می‌کند و به صورت رابطه (۸) تعریف می‌شود:

$$ENTR = - \sum_{l=l_{min}}^N P(l) \ln P(l) \quad (8)$$

- آرامی

برای محاسبه معیار آرامی تعداد نقاط بازخداد که به صورت خطوط عمودی در نگاشت بازخداد (با حداقل طول v_{min}) تشکیل می‌شوند بر تعداد کل نقاط بازخداد تقسیم می‌شود (رابطه (۹)).

$$LAM = \frac{\sum_{v=v_{min}}^N vP(v)}{\sum_{v=1}^N vP(v)} \quad (9)$$

ویژگی نام انتخاب می‌شود و اگر بیت نام کروموزوم ۰ باشد، ویژگی نام حذف می‌شود.

بنابراین در هر نسل، بر اساس بیت‌های باینری هر کروموزوم، مجموعه ویژگی‌های انتخاب شده از بردار ویژگی اصلی، طبقه‌بندی شده و صحت ارزیابی چند دسته‌ای طبقه‌بندی آنالیز افتراقی خطی به عنوان خروجی تابع برازندگی هر کروموزوم تعیین می‌شود و برای تولید نسل بعدی مورد استفاده قرار می‌گیرد. در نهایت متناسب‌ترین کروموزوم، کروموزومی است که دقت طبقه‌بندی بر اساس بردار ویژگی مرتبط با آن به حداکثر برسد.

روند این الگوریتم بسیار زمانبر است و زمان آن مستقیماً به جمعیت اولیه و تعداد نسل‌های بعدی ارتباط دارد. اگر جمعیت اولیه و تعداد نسل‌ها افزایش یابد، زمان اجرای الگوریتم بالا خواهد بود اما، احتمال یافتن یک راه حل کارآمد نیز بیشتر می‌شود [۱۷]. جدول ۱ نشان‌دهنده پارامترهای الگوریتم ژنتیک مورد استفاده در این مقاله، به منظور شناسایی ویژگی‌های بهینه از بردار ویژگی ترکیبی به دست آمده از روش تلفیق ویژگی‌هاست.

جدول ۱- پارامترهای مورد استفاده در الگوریتم ژنتیک در فرآیند انتخاب ویژگی

پارامترهای الگوریتم ژنتیک	مقدار
جمعیت اولیه	۱۰۰
تعداد تولید نسل	۲۰۰
تعداد تولید مثل	۷
تعداد تقاطع	۹۰
تعداد جهش	۳
روش تقاطع	تقاطع دو نقطه‌ای ^{۵۰}
تابع جهش	تابع گوسی

۴-۵- تلفیق در سطح تصمیم‌گیری

بر اساس مطالب بیان شده در مطالعات پیشین هر سه کانال Pz، Cz و Fz حاوی اطلاعات شناختی درباره آشنایی فرد با محرک اعمالی است. اما بر اساس نتایج به دست آمده میزان قابلیت اطمینان اطلاعات کانال‌های مختلف متفاوت است. یکی از رویکردهای تلفیق اطلاعات، تلفیق اطلاعات در سطح تصمیم‌گیری است. در این روش، هر یک از بردارهای ویژگی به طور مستقل توسط طبقه‌بند مربوطه طبقه‌بندی شده و خروجی‌های آن‌ها برای تولید نتیجه نهایی ترکیب می‌شوند. ترکیب در این روش می‌تواند به روش‌های مختلف صورت گیرد که از جمله این روش‌ها می‌توان به رأی‌گیری به عنوان ساده‌ترین روش اشاره کرد. این نوع تلفیق اطلاعات قابلیت اطمینان کانال‌ها را در نظر نمی‌گیرد و همه کانال‌ها نقش یکسانی در ایجاد خروجی نهایی دارند. با توجه به این که در کاربرد دروغ‌سنجی قابلیت اطمینان کانال‌ها از درجه اهمیت بالایی برخوردار است، در این مقاله از روش تلفیق در سطح تصمیم‌گیری با استفاده از یک طرح میانگین وزن‌دار برای تشخیص و طبقه‌بندی دادگان هر سوژه استفاده شده‌است.

در این روش ابتدا ویژگی‌های تجزیه و تحلیل کمی بازخداد به دست آمده از تک ثبت‌های هدف و غیرمرتبط هر سوژه به عنوان دادگان آموزش به طبقه‌بندها داده می‌شود. نرخ خطای هر طبقه‌بند از طریق ارزیابی تقاطعی ۵ دسته‌ای محاسبه می‌شود. سپس، بر اساس این خطا مقدار صحت آموزش طبقه‌بند برای هر کانال به دست می‌آید (رابطه (۱۸)):

$$ACC = 1 - Error Rate \quad (18)$$

این فرآیند برای هر سه کانال Pz، Cz و Fz انجام شده و صحت به دست آمده از طبقه‌بندی دادگان آموزش هر کانال، به عنوان قابلیت اطمینان همان کانال در نظر گرفته می‌شود. سپس، ویژگی‌های به دست آمده از تک ثبت‌های پروب هر سوژه به

$$Trans = \sum_{i=1}^N \frac{\sum_{j,k=1}^N A_{ij} A_{jk} A_{ki}}{\sum_{j,k=1}^N A_{ij} A_{ki}} \quad (16)$$

۴-۴- تلفیق در سطح ویژگی و انتخاب ویژگی

در این روش، ابتدا ۱۳ ویژگی تجزیه و تحلیل کمی بازخداد به طور مستقل از هر یک از کانال‌های Pz، Cz و Fz استخراج شدند. سپس تلفیق ویژگی از طریق بهم پیوستن بردارهای ویژگی استخراج شده از تک ثبت‌های هر سوژه از کانال‌های Pz، Cz و Fz انجام شده و یک ماتریس ویژگی ترکیبی، F_c ، به دست می‌آید (رابطه (۱۷)).

$$F_c = \begin{bmatrix} F_{11} & F_{12} & \dots & F_{1M} \\ F_{21} & F_{22} & \dots & F_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ F_{d1} & F_{d2} & \dots & F_{dM} \end{bmatrix} \quad (17)$$

که d بُعد بردار ویژگی F و M تعداد کانال‌ها می‌باشد. با استفاده از روش تلفیق ویژگی و بهم پیوستن ۱۳ ویژگی نگاشت بازخداد هر کانال، بردار ویژگی ترکیبی به طول ۳۹ از بردارهای ویژگی اولیه ساخته شد.

مزیت این نوع تلفیق ویژگی نسبت به تلفیق ویژگی از طریق میانگین‌گیری در این است که با این روش تمام ویژگی‌های استخراج شده از کانال‌ها در بردار ویژگی ترکیبی وجود دارد و هیچ اطلاعاتی حذف نمی‌شود. اما، در مواردی که تعداد کانال‌ها زیاد باشد، بُعد بردار ویژگی افزایش می‌یابد و ممکن است در فرآیند آموزش طبقه‌بندی با تعداد دادگان محدود، باعث ایجاد چالش شود.

به منظور شناسایی و انتخاب بهترین زیر مجموعه از ویژگی‌های استخراج شده، چندین روش می‌تواند مورد استفاده قرار گیرد. در این مقاله، برای انتخاب ویژگی‌های بهینه از بردار ویژگی حاصل از تلفیق همه ویژگی‌ها، از الگوریتم ژنتیک^{۴۵} استفاده شد. الگوریتم ژنتیک یک تکنیک بهینه‌سازی جستجوگرانه و مبتنی بر اصول علم ژنتیک است [۴۶، ۴۷]. این روش اغلب برای یافتن راه حل‌های بهینه برای مسائل بهینه‌سازی مورد استفاده قرار می‌گیرد. در این روش، تعدادی رشته عددی به عنوان کروموزوم در نظر گرفته می‌شوند که هر کروموزوم حاوی چند ژن است. هر ژن کروموزوم حاوی یک عدد رمزگذاری شده است که معادل یک ویژگی راه حل مساله می‌باشد. در مجموع، ژن‌های یک کروموزوم بیانگر ویژگی‌های منحصر به فرد راه حل متناظر با آن کروموزوم است.

تابع برازندگی^{۴۶} معیاری است که میزان بهینه بودن راه حل متناظر با هر کروموزوم را تعیین می‌کند. مقدار بیشتر تابع برازندگی نشان‌دهنده بهینه‌تر بودن راه حل است. پس از محاسبه تابع برازندگی برای کروموزوم‌ها، در یک فرآیند خاص و با استفاده از برخی عملگرها، نسل جدیدی از کروموزوم‌ها از نسل قبلی تولید می‌شود. برخی از عملگرهای الگوریتم ژنتیک به شرح ذیل می‌باشند:

- تولید مثل^{۴۷}: عملگری است که در آن مناسب‌ترین کروموزوم‌ها (آن‌هایی که بهترین راه حل‌های مربوطه را دارند) به طور مستقیم به نسل بعدی منتقل می‌شوند.
- تقاطع^{۴۸}: یکی از عملگرهایی است که در آن بخشی از یک کروموزوم با بخشی از کروموزوم دیگر ترکیب می‌شود تا کروموزوم جدیدی تولید شود.
- جهش^{۴۹}: عملگری است که در آن چند کروموزوم تصادفی تولید می‌شود.

برای استفاده از الگوریتم ژنتیک در حل مساله انتخاب ویژگی کروموزوم‌ها به صورت باینری کد می‌شوند و هر بیت وجود یا عدم وجود یک ویژگی را در بردار ویژگی انتخاب شده نشان می‌دهد. به عنوان مثال؛ اگر بیت نام کروموزوم ۱ باشد،

نزدیک‌ترین همسایگی و نوع فاصله مورد استفاده وجود دارد. از جمله فواصل قابل استفاده در این الگوریتم می‌توان به فواصل منتهی^{۵۲}، ماهالانویس^{۵۳}، چیشیف^{۵۴} و مینکوفسکی^{۵۵} اشاره کرد.

- درخت تصمیم:

الگوریتم درخت تصمیم به خانواده الگوریتم‌های یادگیری ماشین نظارت شده تعلق دارد. این الگوریتم هم برای مسائل طبقه‌بندی و هم برای مسائل رگرسیون قابل استفاده است. درخت تصمیم یکی از طبقه‌بندیهایی است که فضای ویژگی را به چند قسمت تقسیم‌بندی می‌کند و براساس ویژگی‌ها و پراکندگی آماری کلاس‌های مختلف به هر ناحیه از فضای ویژگی برچسب یک کلاس را اختصاص می‌دهد [۵۴].

در درخت تصمیم هر گره داخلی نشان‌دهنده یک پرسش یا آزمون در مورد یک ویژگی و هر شاخه نشان‌دهنده نتیجه آزمایش و پاسخ‌های متصور برای آن سوال است. در نهایت برچسب کلاس توسط هر گره برگ (یا گره انتهایی) نشان داده می‌شود. درخت تصمیم را می‌توان در مقایسه با سایر روش‌های طبقه‌بندی نسبتاً سریع دانست. طبقه‌بند درخت تصمیم در مقایسه با سایر روش‌های طبقه‌بندی می‌تواند دقت مشابه یا بهتری را به دست آورد. از جمله ابرپارامترهای این طبقه‌بند می‌توان به عمق درخت اشاره کرد که تأثیر زیادی بر روی تعمیم‌پذیری درخت تصمیم دارد.

۵-۲- اعتبارسنجی

همانگونه که ذکر شد تنوع پاسخ‌های P300 در هر سوژه، در مطالعات مختلف گزارش شده‌است. بنابراین در نظر گرفتن این تفاوت‌های بین سوژه‌ای می‌تواند منجر به بهبود عملکرد سیستم گردد. بر همین اساس در این مقاله رویکرد تجزیه و تحلیل داده‌های دروغ‌سنجی به صورت مبتنی بر سوژه انتخاب شد. در این رویکرد، از داده‌های هدف و غیرمرتبط هر سوژه برای آموزش طبقه‌بند مربوط هر سوژه استفاده گردید. بر این اساس دادگان پروب به عنوان مجموعه داده آزمون مستقل که در هیچ یک از مراحل آموزش دخالت داده نمی‌شود در نظر گرفته می‌شود. لازم به ذکر است تمامی ابرپارامترهای طبقه‌بندهای فوق‌الذکر از طریق روش‌های جستجوگرانه‌ی بهینه‌سازی بیزین و ارزیابی تقاطعی ۵ دسته‌ای، تعیین شده است [۵۵،۵۶].

۶- نتایج آزمایشات تجربی

۶-۱- پروتکل ثبت داده

دادگان مورد استفاده در این مقاله، دادگان ثبت شده در مقاله [۷] است. بر اساس پروتکل ثبت دادگان این مقاله، در ابتدا شرکت‌کنندگان به صورت تصادفی به دو گروه گناهکار و بی‌گناه تقسیم شدند و از شش شیء برای انجام روند آزمایش استفاده شد. جعبه‌ای حاوی یک شیء (برای بی‌گناهان) و دو شیء (برای گناهکاران) به شرکت‌کنندگان داده شد. به آن‌ها گفته شد که جعبه را باز کرده و جزئیات شیء را بخاطر بسپارند. به گروه گناهکار گفته شد که فقط یک شیء را سرقت کنند، تصویر این شیء در روند ثبت داده به عنوان محرک پروب برای گناهکار نقش دارد. شیء باقیمانده در جعبه محرک هدف برای فرد گناهکار است و چهار شیء باقیمانده محرک‌های غیرمرتبط می‌باشند. شیء موجود در جعبه برای شخص بی‌گناه به عنوان محرک هدف است. سپس، از پنج شیء باقیمانده، یک شیء به طور تصادفی انتخاب شده و به عنوان محرک پروب تنظیم شده و چهار شیء باقیمانده به عنوان محرک غیرمرتبط در نظر گرفته شده‌است.

به همه افراد دستور داده شد که اطلاعات مربوط به اشیاء موجود در جعبه مانند سبک و رنگ اشیاء را به خاطر بسپارند. تصاویر این اشیاء به عنوان محرک به صورت تصادفی روی صفحه، نمایش داده شدند. شکل ۴ طرح پروتکل محرک‌ها را

عنوان دادگان آزمون به طبقه‌بندهای آموزش دیده داده شده و احتمال پسین تعلق تک ثبت‌های هر سوژه به کلاس‌های بی‌گناه (۰) و گناهکار (۱) تعیین می‌شود. براساس این احتمال و با استفاده از قابلیت اطمینان کانال‌های Pz، Cz و Fz، برچسب تک ثبت‌های هر سوژه به صورت احتمالاتی با استفاده از رابطه (۱۹) مشخص می‌شود.

$$Soft\ Label = \frac{\sum(w_i \times P(c_j|x))}{\sum w_i} \quad (19)$$

که w_i قابلیت اطمینان هر کانال است که همان صحت اعتبارسنجی هر طبقه‌بند برای هر کانال می‌باشد و $P(c_j|x)$ احتمال تعلق مشاهده x به کلاس c_j است. در نهایت با استفاده از آستانه‌گذاری بر روی این احتمال، گناهکار یا بی‌گناه بودن تک ثبت‌های سوژه تعیین می‌شود. در این مقاله میزان آستانه ۰/۵ در نظر گرفته شد که به ازای احتمال کمتر از آستانه تک ثبت‌های سوژه بی‌گناه و به ازای احتمال بیشتر از آستانه سوژه گناهکار است. سپس برچسب نهایی هر سوژه بر اساس روش رأی اکثریت بر روی تک ثبت‌های آن سوژه تعیین می‌شود.

۵- طبقه‌بندی و اعتبارسنجی

۵-۱- طبقه‌بند

برای بررسی و مقایسه بهتر روش‌های تلفیق اطلاعات، در این مقاله از سه طبقه‌بند ماشین بردار پشتیبان، k-نزدیکترین همسایگی و درخت تصمیم به منظور طبقه‌بندی دادگان استفاده شد.

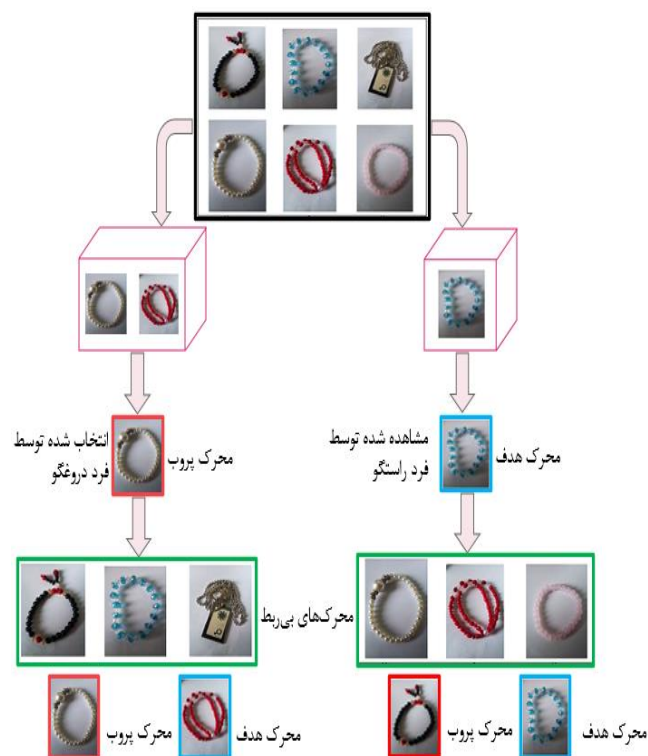
- ماشین بردار پشتیبان:

ماشین بردار پشتیبان یک الگوریتم نظارت‌شده یادگیری ماشین است که هم برای مسائل طبقه‌بندی و هم مسائل رگرسیون قابل استفاده است؛ با این حال از آن بیشتر در مسائل طبقه‌بندی استفاده می‌شود. ماشین بردار پشتیبان می‌تواند داده‌های خطی و غیرخطی را طبقه‌بندی کند [۴۸]. این الگوریتم یک طرح یادگیری ریاضی مبتنی بر ابر صفحه است [۴۹] که دو کلاس را در حداکثر فاصله تا نزدیکترین نقطه داده از هر کلاس جدا کند [۵۰]. این استراتژی طبقه‌بندی، خطاهای طبقه‌بندی داده‌های آموزش را به حداقل می‌رساند و توانایی تعمیم‌پذیری بهتری به دست می‌آورد [۵۱]. در واقع وجود این حاشیه توانایی طبقه‌بندی دقیق را افزایش و احتمال طبقه‌بندی غلط را کاهش می‌دهد. همچنین برای دادگان جدایی‌ناپذیر خطی ماشین بردار پشتیبان با اعمال یک کرنل غیرخطی بر روی دادگان تبدیل فضای داده‌های اصلی به فضای چند بُعدی و ساخت ابرصفحه خطی در ابعاد بالاتر طبقه بندی دادگان را انجام می‌دهد. از جمله ابرپارامترهای این طبقه‌بند می‌توان به نوع کرنل و پارامترهای مرتبط با آن و ضریب محدودیت جعبه اشاره کرد. تنظیم بهینه این ابرپارامترها تأثیر زیادی بر روی عملکرد این طبقه‌بند دارد.

- k-نزدیکترین همسایگی:

الگوریتم k-نزدیکترین همسایگی یک روش طبقه‌بندی غیر پارامتری است که برای طبقه‌بندی و رگرسیون استفاده می‌شود. این طبقه‌بند به عنوان یکی از طبقه‌بندهای پرکاربرد و کارآمد در بازشناسی آماری الگو به شمار می‌رود [۵۲]. الگوریتم یادگیری در این طبقه‌بند در دسته روش‌های تنبل یادگیری قرار می‌گیرد که بر اساس آن تمامی داده‌های مجموعه داده آموزش و برچسب متناظر با آن در نظر گرفته می‌شود با ورود نمونه آزمون جدید فاصله آن نمونه تا نمونه‌های مجموعه‌ی آموزش محاسبه می‌شود. سپس k نزدیک‌ترین همسایه‌ی نمونه جدید تعیین شده و براساس برچسب آن‌ها و با استفاده از استراتژی رأی اکثریت برچسب نمونه جدید تعیین می‌گردد [۵۳]. در این الگوریتم ابرپارامترهایی^{۵۱} مانند تعداد

نمایش می‌دهد. هر محرک به مدت ۱/۱ ثانیه با ۳۰ تکرار در یک جلسه نشان داده شد و جلسه ثبت حدود ۵ دقیقه با ۲ دقیقه زمان استراحت به طول انجامید. فاصله بین هر دو محرک ۰/۵ ثانیه بود.



شکل ۴- طرح پروتکل محرک‌ها [۷].

جدول ۲- نتایج حاصل از طبقه‌بندی ویژگی‌های سه کانال به صورت مستقل

کانال	طبقه‌بند	صحت (%)	حساسیت (%)	ویژه بودن (%)
Pz	SVM	۷۶/۶۷	۸۰	۷۳/۳۳
	KNN	۶۶/۶۷	۷۳/۳۳	۶۰
	Decision Tree	۷۰	۷۰	۸۰
Cz	SVM	۸۳/۳۳	۶۰	۸۶/۶۷
	KNN	۷۶/۶۷	۸۰	۷۳/۳۳
	Decision Tree	۷۶/۶۷	۷۳/۳۳	۸۰
Fz	SVM	۸۰	۹۳/۳۳	۶۶/۶۷
	KNN	۷۶/۶۷	۸۶/۶۷	۶۶/۶۷
	Decision Tree	۷۳/۳۳	۶۶/۶۷	۸۰

همانطور که مشاهده می‌شود بهترین نتایج در این رویکرد برای هر کانال متعلق به طبقه‌بند ماشین بردار پشتیبان می‌باشد.

در گام بعدی ویژگی‌های استخراج شده از هر سه کانال از طریق بهم پیوستن با یکدیگر تلفیق شدند. نتایج طبقه‌بندی این روش در جدول ۳ نمایش داده شده‌است.

جدول ۳- نتایج حاصل از طبقه‌بندی تلفیق ویژگی‌های سه کانال بدون انتخاب ویژگی

رویکرد تلفیقی	طبقه‌بند	صحت (%)	حساسیت (%)	ویژه بودن (%)
تلفیق در سطح ویژگی بدون انتخاب ویژگی	SVM	۷۶/۶۷	۸۶/۶۷	۶۶/۶۷
	KNN	۷۰	۶۰	۸۰
	Decision Tree	۸۰	۷۳/۳۳	۸۶/۶۷

همانطور که در جدول ۳ مشاهده می‌شود بهترین صحت در روش تلفیق ویژگی‌های سه کانال بدون انتخاب ویژگی، مربوط به طبقه‌بند درخت تصمیم با صحت ۸۰٪ می‌باشد.

علاوه بر این، عملکرد طبقه‌بندی مبتنی بر بردار ویژگی ترکیبی حاصل از تلفیق ویژگی‌های سه کانال Pz، Cz و Fz، با استفاده از الگوریتم ژنتیک به عنوان روش انتخاب ویژگی مورد بررسی قرار گرفت. نتایج این مرحله در جدول ۴ نمایش داده شده‌است.

جدول ۴- نتایج حاصل از طبقه‌بندی ویژگی‌های سه دست آمده از تلفیق ویژگی‌های سه کانال با استفاده از الگوریتم ژنتیک

رویکرد تلفیقی	طبقه‌بند	صحت (%)	حساسیت (%)	ویژه بودن (%)
تلفیق در سطح ویژگی و انتخاب ویژگی با الگوریتم ژنتیک	SVM	۸۳/۳۳	۸۰	۸۶/۶۶
	KNN	۷۶/۶۶	۸۰	۷۳/۳۳
	Decision Tree	۶۶/۶۷	۶۶/۶۷	۶۶/۶۷

بهترین صحت در رویکرد تلفیق ویژگی‌های تجربه و تحلیل کمی بازخداد استخراج شده از کانال‌های Pz، Cz و Fz و انتخاب ویژگی با استفاده از الگوریتم ژنتیک متعلق به طبقه‌بند ماشین بردار پشتیبان می‌باشد که برابر با ۸۳/۳۳٪ است.

۲-۶- مجموعه داده‌ها

این دادگان از ۳۰ فرد سالم ثبت شده‌است. دوازده الکتروود (Fp1، Fp2، F3، Fz، Cz، C3، C4، P3، Pz، P4، Oz) و سیستم بین‌المللی ۱۰-۲۰ استفاده شد. سیگنال الکترواوکولوگرام عمودی (VEOG) از چشم راست (۲/۵ سانتی‌متر زیر و بالاتر از مردمک) و سیگنال الکترواوکولوگرام افقی (HEOG) از قسمت بیرونی گوشه چشم ثبت شده‌است. سیگنال‌های الکتروانسفالوگرام و الکترواوکولوگرام با یک فیلتر میان‌گذر از ۰/۱ تا ۳۰ هرتز به صورت آنالین فیلتر شدند و با فرکانس نمونه‌برداری ۵۰۰ هرتز رقمی شدند. در این ثبت مرجع همگی الکتروودها، الکتروود متصل به گوش سمت راست در نظر گرفته شد.

۳-۶- نتایج آزمایشات

همانطور که گفته شد استفاده از تلفیق اطلاعات کانال‌های Cz و Fz که در آن دامنه P3a حداکثر است، به علاوه کانال Pz که در آن دامنه P3b حداکثر است، می‌تواند منجر به بهبود در تشخیص فرد گناهکار از بی‌گناه شود و نتایج قابل اطمینان‌تری را به دست آورد. به منظور بررسی این امر از کانال‌های Pz، Cz و Fz و دو روش تلفیق اطلاعات در سطح ویژگی و تلفیق اطلاعات در سطح تصمیم‌گیری استفاده شد. همچنین از روش تجزیه و تحلیل کمی بازخداد برای استخراج ویژگی‌های آشوبناک سیگنال الکتروانسفالوگرام کانال‌های Pz، Cz و Fz، استفاده شد. در گام نخست ویژگی‌های استخراج شده از هر کانال، به صورت مستقل با طبقه‌بندهای ذکر شده طبقه‌بندی شدند. نتایج حاصل از این فرآیند در جدول ۲ نمایش داده شده‌است.

(جدول ۲). در گام بعدی روش تلفیق ویژگی‌های استخراج شده از سه کانال مورد بررسی قرار گرفت که بهترین نتایج، مربوط به طبقه‌بند درخت تصمیم با صحت ۸۰٪ می‌باشد (جدول ۳).

همچنین نتایج عملکرد طبقه‌بندی مبتنی بر تلفیق در سطح تصمیم‌گیری و استفاده از طرح میانگین وزن‌دار پیشنهادی به منظور تعیین برچسب هر سوژه به طور مستقل در جدول ۵ نمایش داده شده‌است.

جدول ۶- میانگین صحت تشخیص تک ثبت‌ها در رویکردهای پیاده‌سازی شده

رویکرد پیاده‌سازی شده	طبقه‌بند	میانگین درصد صحت تک ثبت‌ها
طبقه‌بندی ویژگی‌های کانال Pz	SVM	۵۷/۴۶
	KNN	۵۴/۳۶
	Decision Tree	۵۳/۰۴
طبقه‌بندی ویژگی‌های کانال Cz	SVM	۵۸/۴۲
	KNN	۵۷/۶۸
	Decision Tree	۵۷/۱۴
طبقه‌بندی ویژگی‌های کانال Fz	SVM	۵۶/۶۲
	KNN	۵۴/۷
	Decision Tree	۵۵/۴۲
تلفیق در سطح ویژگی بدون انتخاب ویژگی	SVM	۵۶/۵۸
	KNN	۵۷/۱۲
	Decision Tree	۵۷/۶۲
تلفیق در سطح ویژگی و انتخاب ویژگی با الگوریتم ژنتیک	SVM	۵۸/۹۶
	KNN	۶۰/۵۷
	Decision Tree	۵۵/۷۴
تلفیق در سطح تصمیم‌گیری	SVM	۶۲/۹۳
	KNN	۵۹/۴
	Decision Tree	۵۸/۶۸

جدول ۵- نتایج حاصل از طبقه‌بندی مبتنی بر روش تلفیق در سطح تصمیم‌گیری

و بر اساس قابلیت اطمینان هر کانال				
رویکرد تلفیقی	طبقه‌بند	صحت (%)	حساسیت (%)	ویژه بودن (%)
تلفیق در سطح تصمیم‌گیری	SVM	۹۰	۸۶/۶۷	۹۳/۳۳
	KNN	۸۶/۶۷	۸۶/۶۷	۸۶/۶۷
	Decision Tree	۸۰	۸۰	۸۰

بهترین نتیجه در روش تلفیق در سطح تصمیم‌گیری متعلق به طبقه‌بند ماشین بردار پشتیبان با صحت ۹۰٪ می‌باشد. شکل ۵ نمودار صحت روش‌های پیاده‌سازی شده در این مقاله با طبقه‌بندهای ماشین بردار پشتیبان، K-نزدیک‌ترین همسایگی و درخت تصمیم را نشان می‌دهد.

همچنین بررسی نتایج درصد تک ثبت‌های درست برای هر سوژه در جدول ۶ نشان داده شده‌است. همانطور که مشاهده می‌شود میانگین صحت تشخیص تک ثبت‌ها در رویکرد تلفیق در سطح تصمیم‌گیری ۶۲/۹۳٪ است که بیانگر برتری این رویکرد نسبت به سایر روش‌های پیاده‌سازی است.

۴-۶- تحلیل نتایج

در این مقاله از چهار رویکرد به منظور طبقه‌بندی ویژگی‌های تجزیه و تحلیل کمی بازرخداد استخراج شده از کانال‌های Pz، Cz و Fz استفاده شد. همانطور که مشاهده شد، در رویکرد طبقه‌بندی ویژگی‌های استخراج شده از هر کانال به طور مستقل، بهترین صحت متعلق به طبقه‌بندی ویژگی‌های استخراج شده از کانال Cz با استفاده از طبقه‌بند ماشین بردار پشتیبان می‌باشد که برابر با ۸۳/۳۳٪ است



شکل ۵- نمودار صحت روش‌های پیاده‌سازی شده در این مقاله با طبقه‌بندهای ماشین بردار پشتیبان، K-نزدیک‌ترین همسایگی و درخت تصمیم.

پیش‌پردازشی برای افزایش نسبت سیگنال به نویز در تک ثبت‌ها انجام نشده است.

در این مقاله اطلاعات استخراج شده از کانال‌های مختلف، همجنس و در حوزه آشوب بود. اضافه شدن منابع اطلاعاتی غیرهمجنس از کانال‌های دیگر ممکن است منجر به غنی شدن اطلاعات و کاهش تکرار اطلاعات گردد. در نتیجه در مطالعات آینده می‌توان علاوه بر اطلاعات آشوبناک، براساس روش‌های دیگر استخراج ویژگی مانند روش‌های حوزه فرکانس و زمان-فرکانس ویژگی‌ها و اطلاعات جدیدتری استخراج نمود و رویکرد تلفیق این اطلاعات در کانال‌های مختلف را براساس قابلیت اطمینان کانال‌ها مدنظر قرار داد. علاوه بر این، رویکرد استفاده از طبقه‌بندهای تجمعی یکی دیگر از روش‌های تلفیق اطلاعات است که می‌تواند به منظور بهبود عملکرد سیستم دروغ‌سنجی مورد بررسی قرار گیرد.

۸- مراجع

- [1] M. Gamer and S. Berti, "Task relevance and recognition of concealed information have different influences on electrodermal activity and event-related brain potentials," *Psychophysiology*, vol. 47, no. 2, pp. 355-364, 2010.
- [2] National Research Council. "Committee to review the scientific evidence on the polygraph." *The polygraph and lie detection*. 2003
- [3] L. A. Farwell and E. Donchin, "The Truth Will Out: Interrogative Polygraphy ('Lie Detection') With Event-Related Brain Potentials," *Psychophysiology*, vol. 28, no. 5, pp. 531-547, 1991.
- [4] J. P. Rosenfeld, J. Ellwanger, and J. Sweet, "Detecting simulated amnesia with event-related brain potentials," *International Journal of Psychophysiology*, vol. 19, no. 1, pp. 1-11, 1995.
- [5] A. H. Mehrnam, A. M. Nasrabadi, M. Ghodousi, A. Mohammadian, and S. Torabi, "Reprint of 'A new approach to analyze data from EEG-based concealed face recognition system,'" *International Journal of Psychophysiology*, vol. 122, no. January 2016, pp. 17-23, 2017.
- [6] V. Abootalebi, M. H. Moradi, and M. A. Khalilzadeh, "A new approach for EEG feature extraction in P300-based lie detection," *Computer Methods and Programs in Biomedicine*, vol. 94, no. 1, pp. 48-57, 2009.
- [7] J. Gao, H. Tian, Y. Yang, X. Yu, C. Li, and N. Rao, "A novel algorithm to enhance P300 in single trials: Application to lie detection using F-score and SVM," *PLoS ONE*, vol. 9, no. 11, 2014.
- [8] M. A. Mubeen and K. H. Knuth, "Evidence-Based Filters for Signal Detection: Application to Evoked Brain Responses," Jul. 2011, Accessed: Jul. 08, 2020. [Online]. Available: <http://arxiv.org/abs/1107.1257>
- [9] V. Abootalebi, M. H. Moradi, and M. A. Khalilzadeh, "A comparison of methods for ERP assessment in a P300-based GKT," *International Journal of Psychophysiology*, vol. 62, no. 2, pp. 309-320, 2006.
- [10] N. Saini, S. Bhardwaj, and R. Agarwal, "Classification of EEG signals using hybrid combination of features for lie detection," *Neural Computing and Applications*, vol. 32, no. 8, pp. 3777-3787, 2020.
- [11] V. Abootalebi, "Computer Brain Communication (BCI) using ERP cognitive components," 2003. [Online]. Available: <https://civilica.com/doc/55111/>
- [12] A. Akhavan and M. H. Moradi, "Detection of Concealed Information Using Multichannel Discriminative Dictionary and Spatial Filter Learning," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 10, pp. 2616-2627, 2018.

در رویکرد بعدی، مشابه مقالات [۵-۶، ۱۷]، از روش تلفیق در سطح ویژگی از طریق بهم پیوستن ویژگی‌های استخراج شده از هر کانال و انتخاب ویژگی از طریق الگوریتم ژنتیک استفاده شد. در نهایت فرآیند طبقه‌بندی با طبقه‌بندهای ذکر شده انجام شد. در این رویکرد بهترین نتایج متعلق به طبقه‌بند ماشین بردار پشتیبان با صحت ۸۳/۳۳٪ بود که نشان‌دهنده بهبود نتایج نسبت به روش تلفیق ویژگی بدون استفاده از روش‌های انتخاب ویژگی می‌باشد (جدول ۴). اما، این نتایج در مقایسه با نتایج طبقه‌بندی ویژگی‌های استخراج شده از کانال‌های Cz، Pz و Fz به طور مستقل، برتری ملموسی را نشان نمی‌دهد.

در رویکرد چهارم، تلفیق در سطح تصمیم‌گیری از طریق میانگین‌گیری وزن‌دار بر اساس قابلیت اطمینان کانال‌های مختلف انجام شد. یک مزیت تلفیق در سطح تصمیم‌گیری نسبت به تلفیق در سطح ویژگی این است که در این فرآیند ویژگی‌های استخراج شده از هر کانال به طور مستقل پردازش و طبقه‌بندی می‌شوند. این رویکرد باعث کاهش بُد بردار ویژگی مورد استفاده در فرآیند طبقه‌بندی و در نتیجه بهبود فرآیند آموزش و اجتناب از کاهش تصمیم‌پذیری طبقه‌بند می‌شود. علاوه بر این می‌توان قابلیت اطمینان هر کانال در فرآیند طبقه‌بندی را با توجه به میزان صحت ارزیابی تقاطعی طبقه‌بندی ویژگی‌های کانال‌ها تعیین کرد. بر اساس نتایج که در جدول ۵ مشاهده می‌شود، این روش در قیاس با سایر روش‌های بررسی شده در این مقاله به صورت قابل توجهی نتایج بهتر، با صحت و حساسیت بالاتری را از خود نشان می‌دهد. بهترین نتیجه حاصل از پیاده‌سازی این روش متعلق به طبقه‌بند ماشین بردار پشتیبان برابر با صحت ۹۰٪ می‌باشد. علاوه بر این، عملکرد بهتر این روش در میانگین صحت تشخیص تک ثبت‌های سوژه‌های مختلف، منجر به افزایش قابلیت اطمینان سیستم دروغ‌سنجی مبتنی بر طبقه‌بندی تک ثبت‌ها شده است. این برتری در نتایج با توجه به ماهیت مساله دروغ‌سنجی از جمله امتیازات مهم این روش است.

۷- نتیجه‌گیری و پژوهش‌های آینده

در این مقاله بکارگیری توأمان اطلاعات کانال‌های Cz، Pz و Fz به منظور دستیابی به قابلیت اطمینان و صحت بالاتر در طبقه‌بندی افراد گناهکار و بی‌گناه مد نظر قرار گرفت. برای تحقق این هدف از دو رویکرد تلفیق اطلاعات در سطح ویژگی و تلفیق اطلاعات در سطح تصمیم‌گیری استفاده شد. همچنین به منظور مقایسه بهتر، طبقه‌بندی ویژگی‌های تجزیه و تحلیل کمی بازرخداد کانال‌ها به طور مستقل و ویژگی‌های تلفیق شده از سه کانال بدون استفاده از روش انتخاب ویژگی، پیاده‌سازی شدند. همانطور که مشاهده شد نتایج حاصل از رویکرد تلفیق در سطح تصمیم‌گیری و استفاده از میانگین وزن‌دار بر اساس قابلیت اطمینان کانال‌ها توانسته بهتر از سایر روش‌های مورد استفاده، عمل کند. بالا بودن صحت و حساسیت طبقه‌بندی و میانگین صحت تشخیص تک ثبت‌های رویکرد تلفیق در سطح تصمیم‌گیری نسبت به رویکرد تلفیق در سطح ویژگی و سایر رویکردهای پیاده‌سازی، از جمله برتری‌های قابل توجه روش ارائه شده در این مقاله است.

از دیگر نقاط قوت این مقاله رویکرد مبتنی بر سوژه آن در تشخیص فرد گناهکار و بی‌گناه است. در این رویکرد برای طبقه‌بندی یک سوژه صرفاً نیاز به دادگان ثبت شده در آزمون دروغ‌سنجی همان سوژه است در حالی که در مطالعات صورت گرفته بر روی همین مجموعه داده بر اساس ویژگی‌های استخراج شده از میانگین تک ثبت‌ها، برای تشخیص فرد گناهکار از بی‌گناه نیاز به مجموعه دادگان مناسب از افراد گناهکار و بی‌گناه است. علاوه بر این در روش معرفی شده در این مقاله در مقایسه با مقالات [۷، ۱۰] هیچ‌گونه

- Biomedical Engineering/Biomedizinische Technik*, vol. 65, no. 5, pp. 507–520, 2020.
- [28] I. Gruszczyńska, R. Mosdorf, P. Sobaniec, M. Żochowska-Sobaniec, and M. Borowska, "Epilepsy identification based on EEG signal using RQA method," *Advances in Medical Sciences*, vol. 64, no. 1, pp. 58–64, 2019.
- [29] J. Polich, *Detection of change: event-related potential and fMRI findings.*, Norwell, MA.: Kluwer Academic Publishers, 2003.
- [30] J. Polich, "Updating P300: An integrative theory of P3a and P3b," *Clinical Neurophysiology*, vol. 118, no. 10. NIH Public Access, pp. 2128–2148, Oct. 2007.
- [31] J. Polich and J. R. Criado, "Neuropsychology and neuropharmacology of P3a and P3b," *International Journal of Psychophysiology*, vol. 60, no. 2, pp. 172–185, 2006.
- [32] T. Zhu, W. Luo, and F. Yu, "Multi-branch convolutional neural network for automatic sleep stage classification with embedded stage refinement and residual attention channel fusion," *Sensors (Switzerland)*, vol. 20, no. 22, pp. 1–15, 2020.
- [33] M. BalaKrishnan, P. Colditz, and B. Boashash, "A Multi-Channel Fusion Based Newborn Seizure Detection," *Journal of Biomedical Science and Engineering*, vol. 07, no. 08, pp. 533–545, 2014.
- [34] B. Boashash and S. Ouelha, "Automatic signal abnormality detection using time-frequency features and machine learning: A newborn EEG seizure case study," *Knowledge-Based Systems*, vol. 106, pp. 38–50, 2016.
- [35] Q. Liu, Y. Liu, K. Chen, L. Wang, Z. Li, Q. Ai, & L. Ma, "Research on channel selection and multi-feature fusion of EEG signals for mental fatigue detection," *Entropy*, vol. 23, no. 4, 2021.
- [36] S. Saravanan and S. Govindarajan, "Novel feature extraction of EEG signal for accurate event detection," *International Journal of Medical Engineering and Informatics*, vol. 12, no. 4, pp. 336–344, 2020.
- [37] V. Abootalebi, M. H. Moradi, and M. A. Khalilzadeh, "A new approach for EEG feature extraction in P300-based lie detection," *Computer Methods and Programs in Biomedicine*, vol. 94, no. 1, pp. 48–57, 2009.
- [38] J. Kittler, M. Hatef, R. P. W. Duin, and J. Matas, "On combining classifiers," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 3, pp. 226–239, 1998.
- [39] B. Kaur, P. Kumar, P. P. Roy, and D. Singh, "Impact of ageing on EEG based biometric systems," *Proceedings - 4th Asian Conference on Pattern Recognition, ACPR 2017*, pp. 459–464, 2018.
- [40] M. Mesbah, M. Balakrishnan, P. B. Colditz, and B. Boashash, "Automatic seizure detection based on the combination of newborn multi-channel EEG and HRV information Advances in Nonstationary Electrophysiological Signal Analysis and Processing," *Eurasip Journal on Advances in Signal Processing*, vol. 2012, no. 1, pp. 1–14, 2012.
- [41] G. P. King and I. Stewart, "Phase space reconstruction for symmetric dynamical systems," *Physica D: Nonlinear Phenomena*, vol. 58, no. 1–4, pp. 216–228, 1992.
- [42] F. Takens, "Detecting strange attractors in turbulence," in *Dynamical systems and turbulence, Warwick 1980*, Springer, 1981, pp. 366–381, [Online]. Available: <http://www.springer.com/gp/book/9783540111719>
- [43] S. Wallot and D. Mønster, "Calculation of Average Mutual Information (AMI) and false-nearest neighbors (FNN) for the estimation of embedding parameters of multidimensional time series in matlab," *Frontiers in Psychology*, vol. 9, no. SEP, Sep. 2018.
- [44] N. Marwan, J. F. Donges, Y. Zou, R. v. Donner, and J. Kurths, "Complex network approach for recurrence analysis of time
- [13] A. Turnip, M. F. Amri, H. Fakrurroja, A. I. Simbolon, M. A. Suhendra, and D. E. Kusumandari, "Deception detection of EEG-P300 component classified by SVM method," *ACM International Conference Proceeding Series*, pp. 299–303, 2017.
- [14] M. R. Bhutta, M. J. Hong, Y. H. Kim, and K. S. Hong, "Single-trial lie detection using a combined fNIRS-polygraph system," *Frontiers in Psychology*, vol. 6, no. JUN, pp. 1–9, 2015.
- [15] J. F. Gao, Y. Yang, W. T. Huang, P. Lin, S. Ge, H. M. Zheng, L. Y. Gu, H. Zhou, C. H. Li, & N. N. Rao, "Exploring time-and frequency-dependent functional connectivity and brain networks during deception with single-trial event-related potentials," *Scientific Reports*, vol. 6, no. November, pp. 1–13, 2016.
- [16] A. Arasteh, A. Janghorbani, and M. H. Moradi, "A survey on EMD sensitivity to SNR for EEG feature extraction in BCI application," 2010.
- [17] A. Arasteh, M. H. Moradi, and A. Janghorbani, "A Novel Method Based on Empirical Mode Decomposition for P300-Based Detection of Deception," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 11, pp. 2584–2593, 2016.
- [18] L. Hu, A. Mouraux, Y. Hu, and G. D. Iannetti, "A novel approach for enhancing the signal-to-noise ratio and detecting automatically event-related potentials (ERPs) in single trials," *NeuroImage*, vol. 50, no. 1, pp. 99–111, Mar. 2010.
- [19] C. Saavedra, R. Salas, and L. Bougrain, "Wavelet-based semblance methods to enhance the single-trial detection of event-related potentials for a BCI spelling system," *Computational Intelligence and Neuroscience*, vol. 2019, 2019.
- [20] A. H. Mehrnam, A. M. Nasrabadi, A. Mohammadian, and S. Torabi, "Concealed face recognition analysis based on recurrence plots," *2011 18th Iranian Conference of Biomedical Engineering, ICBME 2011*, no. December, pp. 1–4, 2011.
- [21] N. Marwan, M. Carmen Romano, M. Thiel, and J. Kurths, "Recurrence plots for the analysis of complex systems," *Physics Reports*, vol. 438, no. 5–6, pp. 237–329, 2007.
- [22] S. Martín-González, J. L. Navarro-Mesa, G. Juliá-Serdá, G. M. Ramírez-Ávila, and A. G. Ravelo-García, "Improving the understanding of sleep apnea characterization using Recurrence Quantification Analysis by defining overall acceptable values for the dimensionality of the system, the delay, and the distance threshold," *PLoS ONE*, vol. 13, no. 4, Apr. 2018.
- [23] J. P. Eckmann, O. Oliffson Kamphorst, and D. Ruelle, "Recurrence plots of dynamical systems," *Epl*, vol. 4, no. 9, pp. 973–977, 1987.
- [24] N. Talebi and A. M. Nasrabadi, "Recurrence plots for identifying memory components in single-trial EEGs," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 6334 LNAI, pp. 124–132, 2010.
- [25] S. Schinkel, N. Marwan, and J. Kurths, "Brain signal analysis based on recurrences," *Journal of Physiology Paris*, vol. 103, no. 6, pp. 315–323, 2009.
- [26] F. Bahari and A. Janghorbani, "EEG-based emotion recognition using Recurrence Plot analysis and K nearest neighbor classifier," *2013 20th Iranian Conference on Biomedical Engineering, ICBME 2013*, no. Icbme, pp. 228–233, 2013.
- [27] M. B. Khodabakhshi and V. Saba, "A nonlinear dynamical approach to analysis of emotions using EEG signals based on the Poincare map function and recurrence plots,"

امین جانقربانی مدرک کارشناسی خود را در رشته مهندسی پزشکی در سال ۱۳۸۷ از دانشکده فنی و مهندسی دانشگاه شاهد و مدرک کارشناسی ارشد و دکتری تخصصی خود را در سال‌های ۱۳۸۹ و ۱۳۹۶



در رشته مهندسی پزشکی- بیوالکتریک از دانشکده مهندسی پزشکی دانشگاه صنعتی امیرکبیر اخذ کرد. در حال حاضر وی به عنوان عضو هیات علمی گروه بیوالکتریک دانشکده بیوفناوری دانشگاه سمنان مشغول به فعالیت است. زمینه‌های تحقیقاتی مورد علاقه او مدل‌سازی و مدیریت عدم قطعیت در دانش پزشکی، پردازش سیگنال‌های بیولوژیکی و کاربرد بازشناسی الگو و هوش مصنوعی در پزشکی است. آدرس پست الکترونیکی ایشان عبارت است از:

a.janghorbani@semnan.ac.ir

محمدباقر خدابخش مدرک کارشناسی خود را در رشته مهندسی پزشکی از دانشکده فنی و مهندسی دانشگاه شاهد در سال ۱۳۸۷ اخذ نمود. همچنین وی مدارک تحصیلی کارشناسی ارشد و دکتری تخصصی خود را در سال‌های ۱۳۹۰ و ۱۳۹۶ به ترتیب از دانشگاه علم و صنعت ایران و دانشگاه صنعتی امیرکبیر دریافت نمود. هم‌اکنون وی به عنوان عضو هیات علمی گروه مهندسی پزشکی-بیوالکتریک و سرپرست کارگاه تجهیزات پزشکی در دانشگاه صنعتی همدان مشغول فعالیت است. زمینه‌های تحقیقاتی مورد علاقه ایشان ابزار دقیق پزشکی، پردازش سیگنال‌های بیولوژیکی، شبکه‌های عصبی عمیق و کاربرد آن‌ها در حوزه پزشکی است. آدرس پست الکترونیکی ایشان عبارت است از:



mb.khodabakhshi@hut.ac.ir

¹ Galvanic Skin Response (GSR)

² Functional Magnetic Resonance Imaging (fMRI)

³ Event-Related Potential (ERP)

⁴ Electroencephalogram

⁵ Oddball

⁶ Guilty Knowledge Test (GKT)

⁷ Probe

⁸ Target

⁹ Irrelevant

¹⁰ Signal-to-Noise Ratio (SNR)

¹¹ Bootstrapped Amplitude Difference (BAD)

¹² Bootstrapped Amplitude Difference (BCD)

¹³ Gao

¹⁴ Morphological

¹⁵ Wavelet Transform (WT)

¹⁶ Saini

¹⁷ Empirical Mode Decomposition (EMD)

¹⁸ Leave One Subject Out

¹⁹ Bhutta

²⁰ Functional Near-Infrared Spectroscopy (fNIRS)

²¹ Wavelet Coherence (WC)

²² Chaotic

²³ Phase Space

- series," *Physics Letters, Section A: General, Atomic and Solid State Physics*, vol. 373, no. 46, pp. 4246-4254, 2009.
- [45] N. Marwan, "How to avoid potential pitfalls in recurrence plot based data analysis," *International Journal of Bifurcation and Chaos*, vol. 21, no. 4, pp. 1003-1017, 2011.
- [46] D. A. Coley, "An Introduction to Genetic Algorithms for Scientists and Engineers," *World Scientific Publishing Co.Pte.Ltd*, 1999.
- [47] D. E. Goldberg and J. H. Holland, "Genetic Algorithms and Machine Learning," *Mach. Learn.*, vol. 3, no. 2, pp. 95-99, 1988.
- [48] M. Arun Kumar and M. Gopal, "Least squares twin support vector machines for pattern classification," *Expert Systems with Applications*, vol. 36, no. 4, pp. 7535-7543, 2009.
- [49] N. H. Farhat, "Photonit neural networks and learning machines the role of electron-trapping materials," *IEEE Expert-Intelligent Systems and their Applications*, vol. 7, no. 5, pp. 63-72, 1992.
- [50] D. Morariu, L. N. Vintan, and V. Tresp, "Feature Selection Methods for an Improved SVM Classifier," *International Journal of Intelligent*, no. 4, pp. 575-585, 2006, [Online]. Available: http://webSPACE.ulbsibiu.ro/daniel.morariu/html/docs/ICI_S2006Prague.pdf
- [51] J. Cervantes, F. Garcia-Lamont, L. Rodríguez-Mazahua, and A. Lopez, "A comprehensive survey on support vector machine classification: Applications, challenges and trends," *Neurocomputing*, vol. 408, pp. 189-215, 2020.
- [52] A. Janghorbani and M. H. Moradi, "Fuzzy Evidential Network and Its Application as Medical Prognosis and Diagnosis Models," *Journal of Biomedical Informatics*, vol. 72, pp. 96-107, 2017.
- [53] W. Hou, D. Li, C. Xu, H. Zhang, and T. Li, "An advanced k nearest neighbor classification algorithm based on KD-tree." In *2018 IEEE International Conference of Safety Produce Informatization (IICSPI)*, pp. 902-905. IEEE, 2018.
- [54] M. Batra and R. Agrawal, "Comparative analysis of decision tree algorithms," *Advances in Intelligent Systems and Computing*, vol. 652, pp. 31-36, 2018.
- [55] M. Injadat, F. Salo, A. B. Nassif, A. Essex, and A. Shami, "Bayesian Optimization with Machine Learning Algorithms Towards Anomaly Detection," *2018 IEEE Global Communications Conference, GLOBECOM 2018 - Proceedings*, 2018.
- [56] M. Pelikan, "Hierarchical Bayesian optimization algorithm: toward a new generation of evolutionary algorithms," *SICE 2003 Annual Conference in Fukui*, pp. 547-552, 2003, [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1323811

سکینه رضوی مدرک کارشناسی خود را در رشته

مهندسی پزشکی-بیوالکتریک از دانشگاه ایوانکی در سال

۱۳۹۸ اخذ کرد و در حال حاضر به عنوان دانشجوی

کارشناسی ارشد رشته بیوالکتریک در دانشگاه سمنان

مشغول به تحصیل است. زمینه‌های تحقیقاتی مورد علاقه او پردازش

سیگنال‌های بیولوژیکی، کاربرد هوش مصنوعی و بازشناسی الگو در

پزشکی و مدل‌سازی سیستم‌های بیولوژیکی است. آدرس پست

الکترونیکی ایشان عبارت است از:

s_razavi@semnan.ac.ir



-
- 24 Recurrence Plot (RP)
 - 25 Eckmann
 - 26 Recurrence Quantification Analysis (RQA)
 - 27 Attentional Focus
 - 28 Parietal
 - 29 Feature-level Fusion
 - 30 Decision-level Fusion
 - 31 Concatenating
 - 32 Liu
 - 33 Saravanan
 - 34 Linear discriminant analysis
 - 35 BalaKrishnan
 - 36 Kaur
 - 37 Mesbah
 - 38 Packard
 - 39 Takens
 - 40 Mutual information (MI)
 - 41 False Nearest Neighbor (FNN)
 - 42 Heaviside
 - 43 Norm
 - 44 Recurrence Rate (RR)
 - 45 Genetic Algorithm (GA)
 - 46 Fitness Function
 - 47 Reproduction
 - 48 Crossover
 - 49 Mutation
 - 50 crossover two point
 - 51 Hyper-parameters
 - 52 Manhattan
 - 53 Mahalanobis
 - 54 Chebyshev
 - 55 Minkowski

Guilty Knowledge Test Based on the Fusion of EEG Channels Chaotic Information

Sakineh Razavi¹, Amin Janghorbani^{2*}, Mohammad Bagher Khodabakhshi³

^{1,2} Department of New Sciences and Technologies, Semnan University, Semnan, Iran

³ Department of Biomedical Engineering, Hamedan University of Technology, Hamedan, Iran

Abstract

In the recent decade, the P300 component has been applied frequently in the guilty knowledge test (GKT). In various studies, the Pz channel is introduced as the most informative channel about P300 components. However, other studies have shown that Fz and Cz channels also provide helpful information about P300. Therefore, solving the challenge of optimal fusion of the information of these three channels can improve the performance of GKT. In this study, recurrent quantification analysis of EEG signal was applied to extract information about P300. Also, information fusion techniques, feature level fusion and decision level fusion approaches, were used to solve the information fusion challenge. Feature level fusion is performed based on two methods; (1) concatenating feature vectors of the three channels to form a new overall feature vector and (2) feature selection from this new feature vector using a genetic algorithm. In decision-level fusion, the posterior probability of each class is calculated based on feature vectors and reliability of different channels and applied for discrimination of guilty and innocents subjects. Among different proposed approaches, decision-level fusion outperforms other methods with an accuracy of 90%, a sensitivity of 86.67%, and a specificity of 93.33%. In addition, the speed of this method is higher than genetic algorithm-based information fusion.

Keywords: Guilty Knowledge Test; Electroencephalogram; Information Fusion; Channel Reliability; Genetic Algorithm