



تغییرات رشدی عملکرد مغز نوزادان شش و دوازده ماهه سالم و LLD

محمد صالح فقفوریان^۱، میلاد یزدانی^۲ و سهیل همایون فرد^۳

^۱ دانشجوی کارشناسی مهندسی مکانیک، دانشگاه صنعتی شریف

^۲ دانشجوی کارشناسی مهندسی برق، دانشگاه صنعتی شریف

^۳ دانشجوی کارشناسی علوم کامپیوتر، دانشگاه صنعتی شریف

چکیده: سیگنال‌های مغزی حاوی اطلاعات مفیدی درباره‌ی عملکرد مغز می‌باشند. در این مقاله داده‌های سیگنال‌های EEG که از نوزادان شش و دوازده ماهه سالم و LLD نمونه‌برداری شده بود، برای عملیات پیش‌پردازش به کار گرفته شد. هدف از این مقاله تبدیل این داده‌های خام به داده‌های بدون نویز با طول‌های یکسان بود. به نحوی که برای ساخت مدلی برای مطالعه‌ی تغییرات این سیگنال‌ها در نوزادان شش و دوازده ماهه سالم و LLD با ابزار یادگیری ماشین مناسب باشند.

کلمات کلیدی: سیگنال‌های EEG، پردازش سیگنال، فیلتر کردن، ICA، Adjust، یادگیری ماشین.

۱ مقدمه

می‌شوند و تا ۳ الی ۵ سالگی ادامه می‌یابند. روش‌های طبقه‌بندی^۱ مبتنی بر یادگیری ماشین به عنوان ابزاری برای تشخیص موارد ذکر شده پیشنهاد شده است. هدف این پروژه پیش‌پردازش داده‌های جمع‌آوری شده از نوزادان شش و دوازده ماهه برای استفاده به عنوان داده‌های آموزشی مدل یادگیری ماشین می‌باشد [۱].

۲ مروری بر مفاهیم

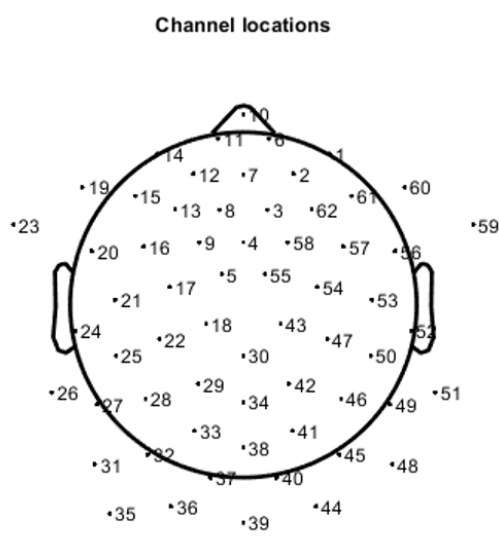
۱-۲ سیگنال‌های مغزی (EEG)

الکتروانسفالوگرافی یا همان نوار مغزی (به اختصار EEG) روشی برای ثبت فعالیت‌های الکتریکی مغز با استفاده از سیگنال‌های دریافت شده از تعدادی الکترود نصب شده روی سر می‌باشد. این روش بر مبنای اندازه‌گیری نوسانات ولتاژ ناشی از جریان‌های یونی ایجاد شده درون نوروهای مغز می‌باشد. سیگنال‌های EEG ترکیبی از چندین سیگنال با فرکانس‌های مختلف می‌باشد که هر کدام از این بازه‌ها نام و مشخصات متفاوتی دارند که در ادامه آورده شده است.

- سیگنال‌های با فرکانس کمتر از ۴ هرتز محدوده‌ی دلتا δ
- بین ۴ تا ۷ هرتز محدوده‌ی تتا θ

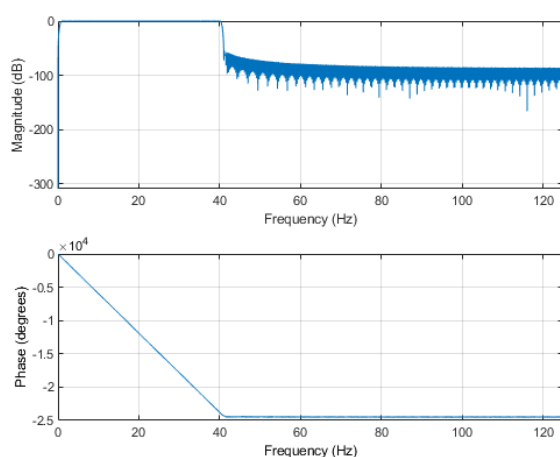
مغز در معرض تغییرات ساختاری و عملکردی بزرگی است. بیش از توسعه‌ی اولیه‌ی پردازش شنوایی سریع، توانایی تشخیص تغییر شنوایی در محدوده‌ی چند ده میلی‌ثانیه برای رمزگشایی جریان گفتار بسیار مهم هستند و جنبه‌های حیاتی گفتار هستند و رشد زبان از بدو تولد شروع می‌شود. این توانایی پیچیده‌ی تشخیص تغییرات ظریف صدا در اوایل دوران نوزادی، به ویژه تغییرات متوالی سریع در دامنه و ترکیب فرکانس مربوط به گفتار، به نظر می‌رسد که در زیرمجموعه‌ای از کودکان به اشتباه می‌افتد. تصور می‌شود که چنین انحراف اولیه از مسیرهای پردازش صوتی هنجاری منجر به اختلالات یادگیری مبتنی بر زبان (LLD) مانند اختلال زبان خاص و اختلال در خواندن می‌شود و به نظر می‌رسد با برخی از انواع اوتیسم همراه باشد. مطالعات ژنتیکی خانوادگی نشان می‌دهد که تقریباً ۳۰ تا ۶۰ درصد از نوزادانی که در خانواده‌هایی با LLD متولد می‌شوند در معرض ابتلا به مشکلاتی از این قبیل هستند. همچنین مطالعات قبلی نشان می‌دهد که LLD با تفاوت‌های قابل تشخیص در ساختار مغز مرتبط هستند که ممکن است قبل از تولد شروع شوند. با توجه به استعداد ژنتیکی شناسایی شده در بسیاری از موارد، تصور می‌شود که این تفاوت‌های کالبدشناختی عاملی در ایجاد LLD هستند. بنابراین این مطالعات از اوایل دوران نوزادی شروع

¹ Classification



62 of 62 electrode locations shown

شکل ۲: مختصات محل کانال‌های نمونه‌برداری



شکل ۳: دیاگرام Bode فیلتر میان‌گذر

۵-۳ ICA

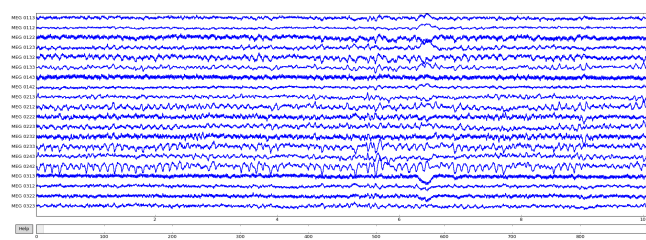
آنالیز مولفه‌های مستقل یا به اختصار ICA^۱ روشی است برای جداسازی سیگنال به مجموع چند سیگنال دیگر به طوری که سیگنال‌های حاصل مستقل و دارای توزیع غیرگوسی باشند. در نتیجه خروجی الگوریتم ICA تعدادی وزن برای سیگنال‌های جدا شده می‌باشد. این وزن‌ها سهم هر سیگنال را نشان می‌دهند. ICA یک حات خاص از جداسازی کور منابع^۲ می‌باشد. ICA سیگنال اولیه را از سیگنال‌های آلوده مانند پلک زدن، حرکت چشم‌ها، عضله‌ها و قلب تمیز می‌کند.

¹Independent Component Analysis

²Blind Source Separation

- بین ۸ تا ۱۲ هرتز محدوده‌ی آلفا α
- بین ۱۳ تا ۳۰ هرتز محدوده‌ی بتا β
- بالای ۳۰ هرتز محدوده‌ی گاما γ

داده‌های جمع‌آوری شده، مربوط به نوزادان شش و دوازده ماهه سالم و LLD می‌باشد که مدت زمان هر داده‌ی نمونه‌برداری ۱۲۰ ثانیه با نرخ نمونه‌برداری ۲۵۰ هرتز می‌باشد.



شکل ۱: یک نمونه سیگنال مغزی EEG

۳ روش

عملیات پیش‌پردازش داده‌ها شامل چند مرحله است که در ادامه به آن‌ها پرداخته می‌شود.

۱-۳ خوانش داده‌ها

در این مرحله، داده‌های جمع‌آوری شده در محیط MATLAB و در جعبه‌ابزار EEGLAB با نرخ داده‌برداری ۲۵۰ هرتز وارد می‌شوند. حالا این داده‌ها آمادگی عملیات پیش‌پردازش هستند.

۲-۳ مشخص کردن محل کانال‌ها

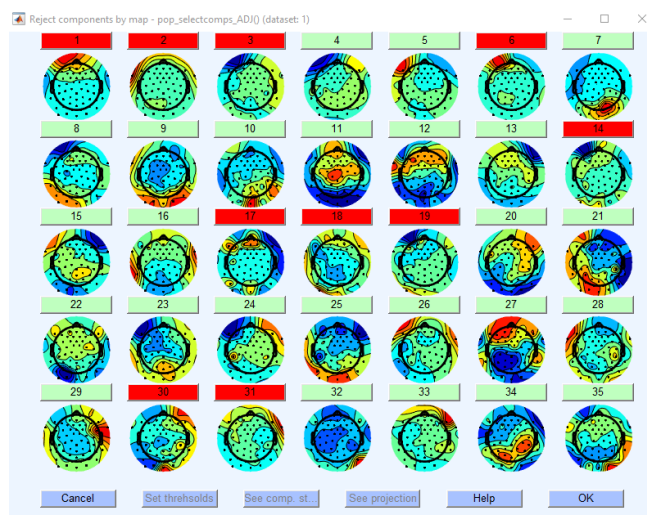
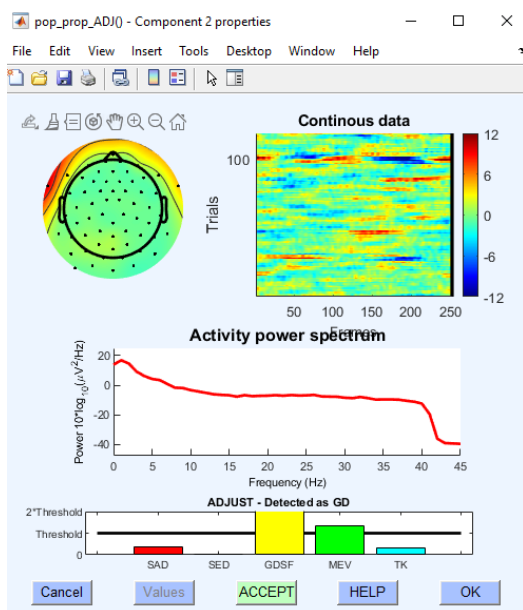
پس از خوانش داده‌ها، بایستی مختصات محل کانال‌ها مشخص شود. این مختصات در واقع همان محل قرارگیری الکترودهایی است که نمونه‌برداری توسط آن‌ها انجام شده است. در محیط EEGLAB و در قسمت Channel Locations فایل مربوط به مختصات محل کانال‌ها وارد می‌شوند. (شکل ۲)

۳-۳ حذف کردن Baseline

در این مرحله baseline داده‌ها (مقدار میانگین داده‌ها) از آن‌ها کم می‌شود تا میانگین داده‌های جدید صفر شود.

۴-۳ فیلتر کردن داده‌ها

داده‌های نمونه‌برداری شده ممکن است دارای نویزهایی باشند. برخی از این نویزها مربوط به سیگنال‌هایی است که فرکانسی بیشتر از ۴۰ هرتز دارند و عملاً در محدوده‌ی سیگنال‌های EEG نیستند. برای این منظور یک فیلتر میان‌گذر روی داده‌ها اعمال می‌شود و فرکانس‌های کمتر از ۱ هرتز و بیشتر از ۴۰ هرتز از داده‌ها حذف می‌شوند. (شکل ۳)

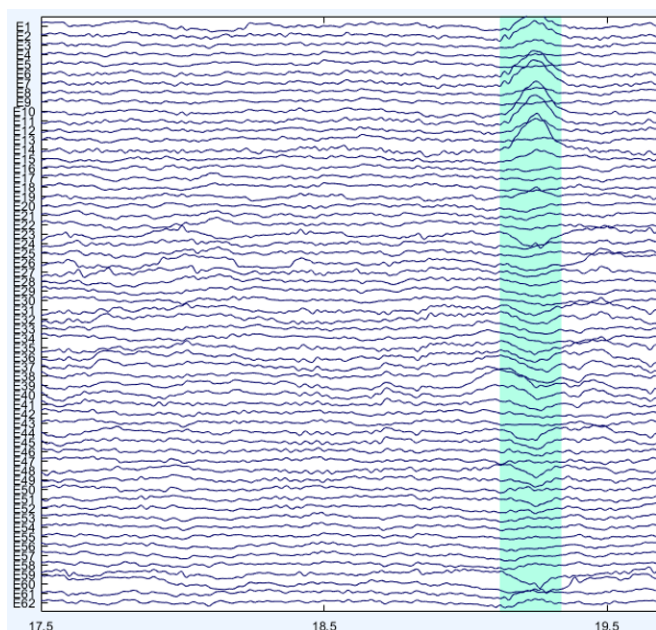


شکل ۴: Adjust

شکل ۵: یک نمونه از تشخیص سیگنال‌ها توسط الگوریتم Adjust

۷-۳ حذف دستی نویزها

در این بخش داده‌ها نمایش داده شده و سپس نویزهای سیگنال (مانند دامنه‌های شدید سیگنال^۵) مشخص می‌شوند. این نویزها می‌توانند به دلایل متعددی مانند حرکت سر نوزاد، جابجایی الکترودها و یا سالم نبودن الکترودها ایجاد شده باشند (شکل‌های ۶ و ۷). پس از حذف نویزها از سیگنال، مجدداً عملیات حذف baseline مجدداً انجام شده است.



شکل ۶: peakها در کانال‌های متوالی

ICA زمانی بهترین نتیجه را می‌دهد که سیگنال‌ها چند کاناله باشند. الگوریتم ICA برای تشخیص مفهوم کلمه و جایگاه آن در مغز مناسب می‌باشد. نتایج این آزمایش با نتایج به دست آمده از روش‌هایی مانند تصویربرداری تشدید مغناطیسی عملکردی^۳ و روش‌های مبتنی بر برق‌نگاری سیگنال‌های مغزی در تصور واکه‌ای در گفتار خاموش یکسان می‌باشد.

برای انجام این مرحله، گزینه‌ی "Decompose data by ICA" در محیط EEGLAB اجرا می‌شود. این مرحله طولانی‌ترین مرحله‌ی پیش‌پردازش داده‌ها می‌باشد. به دلیل اینکه باید در چند گام (معمولاً نزدیک به ۳۰۰ گام) انجام شود. خروجی این مرحله، چند ماتریس مانند ماتریس $n \times n$ ICA-weights می‌باشد که n تعداد کانال‌ها است.

۶-۳ حذف سیگنال‌های مصنوعی

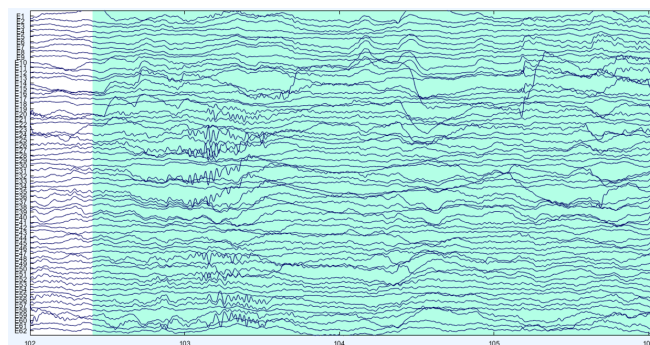
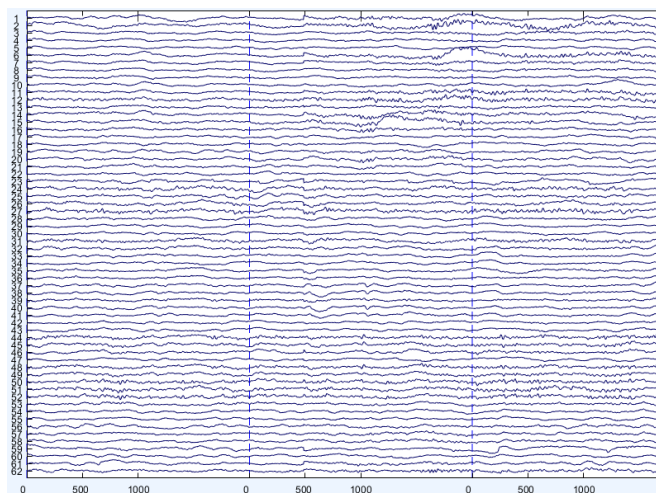
الگوریتم Adjust یک الگوریتم اتوماتیک برای تشخیص سیگنال‌های مصنوعی و حذف آن‌ها بر پایه‌ی ICA است^۴. اجرای Adjust کاملاً به کاربر مربوط است و الگوریتم آن برای تشخیص مولفه‌های مستقل مصنوعی به این صورت کار می‌کند که تمام سیگنال‌های تقلید شده از مولفه‌های مصنوعی را با هم ترکیب می‌کند و از طریق آن مولفه‌های مصنوعی را تشخیص می‌دهد.

این الگوریتم برای تشخیص حرکت‌های چشم و پلک زدن و ناپیوستگی‌ها به صورت بهینه عمل می‌کند. پس از تشخیص مولفه‌های مستقل، می‌توان آن‌ها را بدون آنکه به سیگنال‌ها آسیبی برسد به راحتی حذف کرد. در واقع Adjust یک روش سریع، بهینه و اتوماتیک بر مبنای الگوریتم ICA برای حذف سیگنال‌های مصنوعی فراهم می‌کند (شکل‌های ۴ و ۵)

^۳ FMRI

^۴ Artifact Removal

^۵ Peak



شکل ۷: peakهای شدید

۳-۸ Segmentation

پس از انجام مراحل فوق و حذف تمام نویزها از سیگنال، برای سادگی کار با داده‌ها، می‌توان عملیات Segmentation را بر روی آن‌ها انجام داد. به این صورت که داده‌ها را در قطعه‌های مجزا^۶ با طول‌های یکسان دسته‌بندی کرد. از آنجا که زمان هر نمونه حدود ۱۲۰ ثانیه با نرخ نمونه‌برداری ۲۵۰ هرتز می‌باشد، داده‌ها را با طول‌های ۲ ثانیه (قطعه‌هایی با طول ۵۰۰) ذخیره می‌کنیم و به این ترتیب پیش‌پردازش داده‌های اولیه به پایان می‌رسد.

۴ نتیجه‌گیری

با انجام مراحل فیلتر، استفاده از الگوریتم ICA، حذف سیگنال‌های مصنوعی مبتنی بر ICA، حذف دستی نویزهای باقی‌مانده و قطعه‌سازی داده‌ها می‌توان داده‌هایی بدون نویز آماده کرد. این داده‌ها قابلیت این را خواهند داشت تا با استخراج ویژگی‌های توان و گراف آن‌ها^۱، مدلی برای مقایسه‌ی تغییرات مغز برای نوزادان شش و دوازده ماهه بدست آورد (شکل ۸).

۵ استخراج ویژگی‌های توان

در ادامه‌ی کار با داده‌ها، پس از پیش‌پردازش بایستی ویژگی‌های توان را استخراج کرد. در واقع در مسائل یادگیری ماشین^۱ نیاز است تا پارامترهای موثر ورودی‌های مسئله مشخص باشند تا بر اساس اثر آن ویژگی‌ها، مدلی برای تخمین رفتار سیستم محاسبه شود. با توجه به اینکه این ویژگی‌ها در سیگنال‌های مغزی مشخص نیستند، می‌توان تعدادی از ویژگی‌های توان سیگنال را محاسبه کرد. در ادامه از میان این ویژگی‌ها، آن ویژگی‌هایی که تاثیر مستقیم در رفتار مطلوب دارند انتخاب شوند. برخی ویژگی‌های بدست آمده در ادامه ذکر شده‌اند:

- توان مطلق^۲ برای هر محدوده‌ی سیگنال $(\gamma, \beta, \alpha, \theta, \delta)$

^۶Segment

^۱Feature

^۱Machine Learning

^۲Absolute Power

شکل ۸: یک نمونه سیگنال پس از انجام مراحل پیش‌پردازش

- نسبت توان مطلق^۳ برای هر محدوده‌ی سیگنال: نسبت توان هر سیگنال به توان کل
- نسبت ویژگی‌های استخراج شده^۴: نسبت تمام توان‌های استخراج شده

۶ پژوهش‌های آینده

پس از استخراج این ویژگی‌ها با بررسی آماری آن‌ها (به عنوان مثال با مطالعه‌ی واریانس آن‌ها برای بررسی پراکندگی هر ویژگی) می‌توان ویژگی‌های تاثیرگذار را بدست آورد. پس از این مرحله می‌توان با استفاده از داده‌ها به همراه ویژگی‌های بدست آمده، مدلی مبتنی بر یادگیری ماشین (به طور مثال با استفاده از الگوریتم SVM) برای تخمین رفتار سیگنال ارائه داد. واضح است که هر چه ویژگی‌های بیشتری استخراج شده باشند، ویژگی‌های موثرتری انتخاب شده و مدل نهایی، تخمین بهتری از رفتار سیگنال خواهد بود.

پیوست‌ها

موارد زیر در فایل‌های پیوست موجود هستند.

- یک نمونه داده‌ی نمونه برداری شده‌ی خام (Data)
- کد MATLAB مربوط به پیش‌پردازش داده (Preprocessor)
- خروجی سیگنال پیش‌پردازش شده (Preprocessed Data)

مراجع

- [1] M. Zare, Z. Rezvani, and A. A. Benasich, "Automatic classification of 6-month-old infants at familial risk for language-based learning disorder using a support vector machine," *Clinical Neurophysiology*, vol.127, pp.2695–2703, July 2016.

^۳Absolute Ratio Power

^۴Feature Ratio