

کاربرد الگوریتم تعدیل یافته BIRCH در ناحیه بندی عملکرد مغز بر اساس داده های fMRI

نوید ولی زاده^۱ (M.Sc.)، سهیلا خداکریم^۲ (Ph.D.)، سید محمد طباطبائی^۳ (M.D.)، اعظم صفار^۱ (Ph.D Student)، علیرضا اکبرزاده باغبان^{۴*} (Ph.D.)

۱- گروه آمار زیستی، دانشکده پیراپزشکی، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران

۲- گروه آمار زیستی، دانشکده بهداشت و ایمنی، دانشکده پیراپزشکی، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران

۳- مرکز توسعه تحقیقات بالینی، بیمارستان امام رضا، دانشگاه علوم پزشکی مشهد، مشهد، ایران

۴- مرکز تحقیقات پروتئومیکس، دانشکده علوم توان بخشی، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران

تاریخ دریافت: ۱۳۹۸/۷/۱۰ تاریخ پذیرش: ۱۳۹۹/۳/۱۷

akbarzad@sbmu.ac.ir

* نویسنده مسئول، تلفن: ۰۹۱۲۲۹۵۵۴۱۱

چکیده

هدف: خوشه بندی نواحی مغز در تشخیص، درمان و پیگیری بیماری تومور مغزی بسیار مفید است. روش های مختلفی برای خوشه بندی نواحی مغز وجود دارد. در این مقاله الگوریتم جدیدی تحت عنوان Balanced Iterative Reducing and Clustering using Hierarchies (BIRCH) تعدیل یافته برای ناحیه بندی مغز معرفی می شود که دارای دقت و سرعت بالایی در خوشه بندی مغز است.

مواد و روش ها: در این پژوهش از داده های اسکن مغزی به عنوان معرفی از عملکرد نواحی مختلف مغز استفاده شد. این داده ها شامل ۷۴ اسکن مغزی متوالی بود. خوشه بندی نواحی مورد نظر با الگوریتم های BIRCH معمولی و BIRCH تعدیل یافته با استفاده از ابزار WFU-PickAtlas نرم افزار Matlab انجام شد و نتایج با اطلس استاندارد TD Lobes مقایسه شد.

یافته ها: الگوریتم BIRCH تعدیل یافته در مواجهه با داده های بسیار حجیمی مانند داده های اسکن مغزی نسبت به الگوریتم BIRCH معمولی دارای مزیت نسبی بود و با افزایش حد آستانه ای سرعت اجرای الگوریتم کاهش یافت. در این الگوریتم، الگوی صعودی یا نزولی خاصی بین فاکتور شاخه بندی و زمان اجرای الگوریتم دیده نشد. ماکزیمم مقدار زمان اجرای الگوریتم مربوط به فاکتور شاخه بندی ۳۰، برابر ۹۴ ثانیه بود که هم ارز حد آستانه بالای الگوریتم BIRCH معمولی بود.

نتیجه گیری: الگوریتم BIRCH تعدیل یافته می تواند تعادلی بین پیچیدگی زمانی و مکانی برقرار کند و نیاز به حافظه کامپیوتری کمی در اجرای خوشه بندی دارد. هم چنین با افزایش هم زمان فاکتور شاخه بندی و حد آستانه ای میزان حساسیت در خوشه بندی آن کاهش یافت و بنابراین انتظار می رود در داده مورد مطالعه و اکسل هایی با شباهت کم تر در یک خوشه قرار گرفته باشند.

واژه های کلیدی: مغز، تحلیل خوشه ای، الگوریتم ها، تصویربرداری از طریق تشدید مغناطیسی

مقدمه

مغز مهم ترین قسمت عصبی مرکزی را تشکیل می دهد. ساختار و عملکرد مغز می تواند با استفاده از تصویربرداری تشدید مغناطیسی مورد بررسی قرار گیرد. به طور کلی فرآیندهای پیچیده پزشکی بدون روش های پردازش تصویر غیر ممکن است [۱]. خوشه بندی تصاویر حاصل از تصویربرداری تشدید مغناطیسی کارکردی (fMRI) (Functional Magnetic Resonance Imaging) در تشخیص پزشکی بسیار حائز اهمیت می باشد. این خوشه بندی می تواند در تشخیص،

برنامه ریزی درمان و پیگیری بیماری تومور مغزی بسیار مفید باشد [۲].

علی رغم تلاش های بسیار، خوشه بندی دقیق تومورها به دلیل تفاوت در شکل ها، مکان و میزان روشنایی قسمت های مربوطه هنوز سخت و دارای چالش های زیادی است. در fMRI از خاصیت مغناطیسی اکسیژن موجود در خون و ارتباط آن با فعالیت مغزی استفاده شده و در نواحی فعال سیگنال BOLD افزایش می یابد. بر این اساس نواحی فعال مغزی شناسایی می شوند و این شناسایی نواحی فعال پایه بسیاری از مطالعات در زمینه مغز و شناخت است [۳، ۱۶، ۱۵]. تومورها نواحی

آستانه به صورت خودکار و از روی داده‌ها برآورد می‌شود و نیاز نیست تا مقداری از قبل به این منظور در نظر گرفته شود [۹]. روش‌های خوشه‌بندی مذکور مفهومی به نام ویژگی خوشه (Clustering Feature) (CF) دارند که طریقه‌ای جهت نمایش خلاصه اطلاعات خوشه‌ها بوده و باعث می‌شود تا الگوریتم از سرعت مناسب و قابلیت مقیاس‌پذیری خوبی در مواجهه با داده‌های بسیار حجیم از خود نشان دهد. همچنین این الگوریتم‌ها از یک رویکرد خوشه‌بندی چندمرحله‌ای بهره می‌برند که در آن پس از پیمایش اول داده‌ها و خوشه‌بندی مناسب آن‌ها، با پیمایش بعدی در صدد بهبود کیفیت و نتایج حاصل بر می‌آید [۸،۷]. از آنجایی که مطالعات تشخیص ناحیه تومور به سرعت برای شناخت ناحیه آسیب‌دیده برای جراحی نیاز داشته و به علت حجم بالای داده‌های تصویربرداری مغزی الگوریتم‌های رایج بسیار کند عمل می‌کنند و نیاز به سیستم‌های کامپیوتری قوی با حافظه بالا دارند، در این پژوهش استفاده از الگوریتم BIRCH و الگوریتم تعدیل یافته BIRCH برای خوشه‌بندی نواحی مغز و شناسایی استفاده شد تا قابلیت این الگوریتم‌های برای این نوع داده‌ها نشان داده شود و بتوان مسیر هموارتری برای تشخیص و درمان ایجاد نمود. در نهایت نتایج حاصل از دو خوشه‌بندی با هم مقایسه شد.

مواد و روش‌ها

این تحقیق توصیفی مشاهده‌ای در کمیته اخلاق دانشگاه علوم پزشکی شهید بهشتی تایید شده و دارای کد اخلاق IR.SMU.RETECH.IREC.1396.540 می‌باشد. داده‌های مورد استفاده در این پژوهش شامل داده‌هایی با استفاده عام است که از سایت www.NITRC.org بارگذاری شده و مربوط به اسکن مغزی گرفته شده از کودکان است. این مجموعه داده شامل ۷۴ اسکن مغزی متوالی با ابعاد $47 \times 58 \times 49$ (ارتفاع \times عرض \times طول) می‌باشد [۱۰]. الگوریتم قابلیت اجرا بر روی داده‌های دو بعدی را داشت بنابراین داده چهار بعدی با نرم‌افزار متلب به حالت دو بعدی برده شد به این صورت که آدرس طول و عرض و ارتفاع آن به عنوان بعد اول یعنی مکان ذخیره شد و بعد چهارم داده‌ها یعنی زمان به بعد دوم در داده‌های تبدیل یافته انتقال یافت.

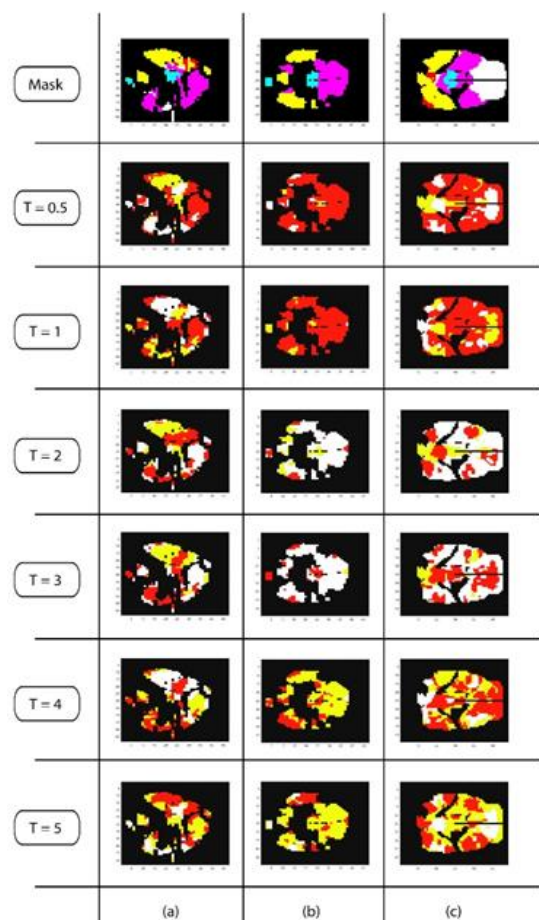
در ابتدا الگوریتم خوشه‌بندی BIRCH روی داده‌ها اجرا شد. لازم بود برای دو پارامتر مهم این الگوریتم یعنی فاکتور شاخه‌بندی (B) و حد آستانه‌ای (T) اعدادی به عنوان پیش فرض منظور شوند. از روی مطالعات مشابه مقادیر فاکتور شاخه‌بندی به ترتیب ۱۰، ۲۰، ۳۰، ۴۰، ۵۰ و مقادیر حد آستانه‌ای نیز ۰،۵، ۱، ۲، ۳، ۴ و ۵ در نظر گرفته شد. سپس الگوریتم خوشه‌بندی

هستند که در آن‌ها مغز فعالیت نداشته یا بسیار کم‌تر از حد متعارف ناحیه فعال باشد و غالباً رگ خونی ندارند بنابراین سیگنال BOLD در آن ناحیه صفر است و در برخورد با محرک‌های مختلف فعالیتی نشان نمی‌دهد.

همواره این علاقه وجود داشته است که مغز انسان را بتوان از نظر کالبدشناسی (Anatomically) و عملکردی (Functionally) به نواحی متمایزی تقسیم کرد. این بحث، بر ارتباط بین واکسل‌ها (Voxel (Volume element)) (کوچک‌ترین جزء سازنده مغز) بنا شده و هدف آن قرار دادن واکسل‌های مرتبط و هم رفتار، در یک ناحیه است [۵، ۶]. یکی از روش‌های رایج در انجام این امر استفاده از روش‌های آماری داده کاوی است که به کمک آن از حجم انبوهی از داده‌ها، بتوان اطلاعات معتبری استخراج نمود [۷].

معمولاً ورودی الگوریتم‌های داده‌کاوی را حجم انبوهی از داده‌ها تشکیل می‌دهد که داده‌های تصویربرداری تشدید عملکردی مغز نیز به همین صورت هستند. تاکنون روش‌های مختلفی برای خوشه‌بندی ناحیه مغز مانند به‌کار گرفته شده است. روش‌هایی مثل K-Means و CLARANS؛ علی‌رغم مفید بودن الگوریتم‌ها در مدت زمان طولانی اجرا شده و تابع تعریف مقادیر مناسبی برای حد آستانه می‌باشند. حد آستانه حداکثر قطر (فاصله) خوشه‌های ذخیره شده پایانی است. قطر نیز به حداکثر فاصله‌ای که میان دو نمونه از خوشه‌ها قرار دارد، اطلاق می‌شود [۴].

الگوریتم خوشه‌بندی BIRCH (Balanced Iterative Reducing and Clustering using Hierarchies) که اولین بار در سال ۱۹۹۶ توسط Zhang ارائه شده از آن دسته از الگوریتم‌ها محسوب می‌شود که با کمک تکنیک‌های خوشه‌بندی سلسله مراتبی و افزایی برای کار بر روی داده‌های با حجم بالا طراحی شده است. در این الگوریتم درخت محور ابتدا بر اساس فواصل و ویژگی‌های داده‌ها خوشه‌های مرتبه پایین ساخته شده و برای مرحله بعد نیازی به همه داده‌ها ندارد و بر اساس مرکز داده‌های حاصل عمل می‌کند و همچنین در محاسبات نیاز به برقراری فرض استقلال در داده‌ها ندارد. اما اشکال ذاتی این الگوریتم آن است که باید برای خوشه‌بندی حد آستانه آن، مقدار مشخصی از قبل تعریف شده باشد [۸]. از آنجا که تخصیص این مقدار اولیه با حدس و گمان انجام می‌شود، می‌تواند با خطا همراه باشد. بنابراین علی‌رغم مزیت زیاد این روش، به الگوریتمی نیاز بود که عاری از این مشکل باشد. الگوریتم جدید به الگوریتم تعدیل یافته BIRCH (modified-BIRCH) یا به اختصار به (m-BIRCH) معروف است که در سال ۲۰۱۵ توسط Madan معرفی شد [۹]. در این الگوریتم حد



شکل ۱. نمایش تصویری خوشه‌بندی با استفاده از الگوریتم BIRCH، با در نظر گرفتن فاکتور شاخه‌بندی $B=10$ در حدآستانه‌های مختلف $(T=0.5, 1, 2, 3, 4, 5)$. برای هر حدآستانه در نظر گرفته شده یک خوشه‌بندی حاصل شده که نمایش آن در سه جهت یا برش امکان پذیر است. خوشه‌ها با رنگ‌های متفاوت برای هر خوشه برای هر حدآستانه در طول برش‌های معین در شکل نمایش داده شده‌اند. (a) برش در طول با مقدار $X=25$ ، (b) برش در عرض با مقدار $Y=30$ ، (c) برش در ارتفاع با مقدار $Z=25$

شکل ۳ نتایج خوشه‌بندی حاصل از اجرای الگوریتم m-BIRCH را بر روی داده‌های اسکن مغزی برای فاکتور $B=10$ و مقدار حد آستانه برآورد شده $3/85$ نشان می‌دهد. هر یک از این تصاویر مربوط به فقط یک برش خاص از ابعاد طول و عرض و ارتفاع متناظر با شکل ۱ می‌باشد. لازم به ذکر است که بر خلاف الگوریتم قبلی، در این الگوریتم مقدار حدآستانه‌ای از روی داده‌ها به صورت دقیق‌تر برآورد شده است.

m-BIRCH با فاکتورهای شاخه‌بندی مشابه انجام شده و مقادیر برآورد شده حد آستانه‌ای گزارش و تصاویر خوشه‌های انتخاب شده به صورت گرافیکی نمایش داده شدند. به علت محدودیت‌های موجود تحلیل داده‌ها با سیستم‌های رایج (RAM8,cpu:corei5) صورت گرفت.

داده‌ها با استفاده از ابزار WFU-PickAtlas (Wake Forest University School of Medicine PickAtlas) برنامه Matlab با اطلس استاندارد TD Lobes (یکی از اطلس‌های استاندارد در علوم اعصاب که مغز را بر اساس قسمت‌های فیزیولوژیک اصلی به ۱۲ ناحیه تقسیم می‌کند) نگاهت داده شده و کل حجم مغز به تعداد نواحی اطلس تقسیم شد و به همین جهت تعداد ۱۲ خوشه برای اجرای الگوریتم در نظر گرفته شد.

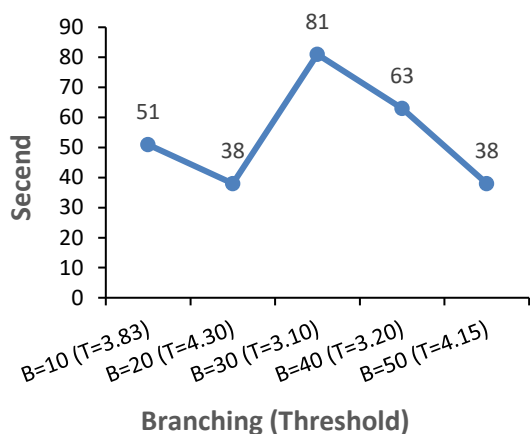
نتایج

شکل ۱ نتایج مربوط به خوشه‌بندی حاصل از اجرای الگوریتم BIRCH بر روی داده‌های اسکن مغزی است که با در نظر گرفتن فاکتور شاخه‌بندی $B=10$ (به عنوان نمونه) در حد آستانه‌های متفاوت $(T=0.5, 1, 2, 3, 4, 5)$ به دست آمده است. برای هر حد آستانه یک تصویر سه بعدی از نواحی دسته‌بندی شده روی مغز طبق الگوریتم مربوطه ارائه شده به دست می‌آید. رنگ‌های متفاوت در تصاویر، نواحی تفکیک شده از مغز را بر اساس میزان فعالیت نشان می‌دهد بنابراین واکنش‌های هر برش (رنگ) از نظر فعالیت مغزی مشابه بودند. در این مطالعه سعی شده الگوریتم‌ها از لحاظ سرعت و توان تفکیک مقایسه شوند تا بتوانند در مطالعات واقعی پزشکی برای شناسایی تومور (به این صورت که ناحیه تفکیک شده‌ای که با هیچ محرکی فعال نشود به عنوان تومور در نظر گرفته می‌شود) به کار روند.

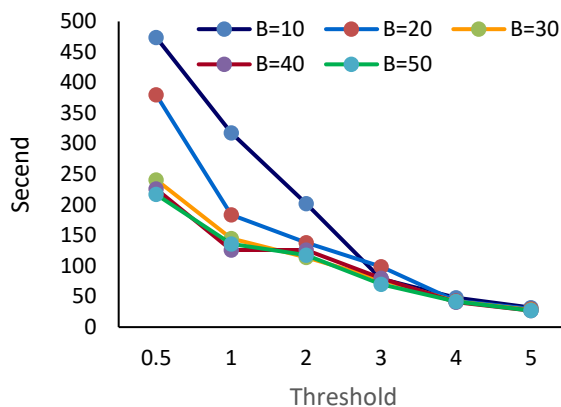
شکل ۲ زمان مورد نیاز برای خوشه‌بندی با الگوریتم BIRCH با در نظر گرفتن حد آستانه‌های مختلف و فاکتور شاخه‌بندی را نشان می‌دهد. همان‌طور که از شکل مشخص است، برای فاکتورهای شاخه‌بندی که مقادیر کوچک‌تری دارند با حد آستانه‌ای‌های کوچک، زمان رسیدن به خوشه‌های نهایی به مراتب بیشتر از حالتی است که فاکتور شاخه‌بندی بزرگ‌تری با حدآستانه‌ای مشابه انتخاب شده باشد. با این حال هر چقدر که حدآستانه‌ای بزرگ‌تری انتخاب شود سرعت اجرای الگوریتم در فاکتورهای شاخه‌بندی مختلف به هم‌دیگر نزدیک می‌شود و به صورت قابل توجهی کاهش می‌یابد.

شکل ۴ زمان های اجرای الگوریتم خوشه بندی m-BIRCH را با در نظر گرفتن مقدار فاکتورهای شاخه بندی برابر با ۱۰، ۲۰، ۳۰، ۴۰ و ۵۰ و مقادیر برآورد شده حد آستانه ای مستخرج از داده ها بر روی داده های اسکن مغزی نشان می دهد.

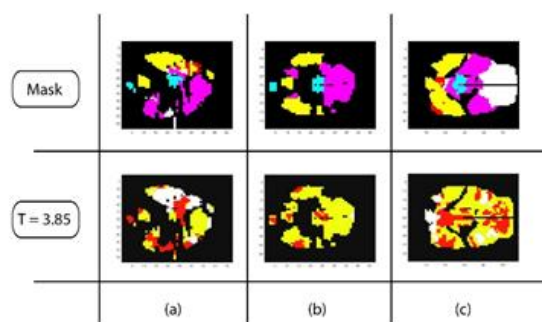
زمان های اجرای الگوریتم BIRCH بر حسب ثانیه در فاکتورهای شاخه بندی و حد آستانه ای های ثابت و هم چنین زمان های اجرای الگوریتم m-BIRCH در فاکتورهای شاخه بندی مذکور و حد آستانه ای برآورد شده در جدول ۱ ارائه شده است.



شکل ۴. زمان اجرای الگوریتم خوشه بندی m-BIRCH با در نظر گرفتن فاکتورهای شاخه بندی مختلف (B=10, 20, 30, 40, 50) در حد آستانه ای های برآورد شده



شکل ۲. زمان اجرای الگوریتم خوشه بندی BIRCH با در نظر گرفتن فاکتورهای شاخه بندی ۱۰، ۲۰، ۳۰، ۴۰ و ۵۰ در حد آستانه ای های متفاوت (T=0.5, 1, 2, 3, 4, 5)



شکل ۳. نمایش تصویری خوشه بندی با استفاده از الگوریتم m-BIRCH، با در نظر گرفتن فاکتور شاخه بندی B=10 در حد آستانه ای ۳.۸۵ (T=3.85). خوشه بندی حاصل شده که نمایش آن در سه جهت صورت گرفته و خوشه ها با رنگ های متفاوت در طول عرش های معین (مشابه حالت قبل) در شکل نمایش داده شده اند. (a) برش در طول با مقدار X=25، (b) برش در عرض با مقدار Y=30، (c) برش در ارتفاع با مقدار Z=25

جدول ۱. زمان های اجرای الگوریتم های BIRCH و m-BIRCH (بر حسب ثانیه) به تفکیک فاکتورهای شاخه بندی و حد آستانه ای

		۵۰			۴۰			۳۰			۲۰			۱۰			فاکتور شاخه بندی (B)
		۵	۴/۱۵	۴	۴	۳/۲۰	۳	۴	۳/۱۰	۳	۵	۴/۳۰	۴	۴	۳/۸۵	۳	حد آستانه ای (T)
		۲۷		۴۳	۴۲		۸۴	۴۴		۸۳	۳۲		۵۰	۴۸		۷۹	BIRCH الگوریتم
			۳۸			۶۳			۸۱			۳۸			۵۱		m-BIRCH الگوریتم

نتایج تحقیق ما نشان داد که الگوی صعودی یا نزولی خاصی بین فاکتور شاخه بندی و زمان اجرای الگوریتم در داده های مورد مطالعه وجود ندارد و این امر می تواند به این دلیل باشد که در هر فاکتور شاخه بندی، حد آستانه متفاوتی برآورد می شود. به علاوه ملاحظه شد ماکزیمم مقدار زمانی اجرای الگوریتم m-BIRCH مربوط به فاکتور شاخه بندی ۳۰ می باشد که با حد آستانه ای برآورد شده ۳/۱۰ برابر ۸۱ ثانیه است که هم ارز حد آستانه ای بالای الگوریتم BIRCH بر روی داده ها بود.

بحث و نتیجه گیری

در این تحقیق الگوریتم های داده کاوی BIRCH و m-BIRCH بر روی داده های fMRI به کار برده شد. دیده شد که الگوریتم m-BIRCH بر خلاف الگوریتم BIRCH مقدار حد آستانه را با استفاده از محاسبات پیچیده ای از روی داده ها و به طور خودکار برآورد نمود که این مزیت ارزشمندی در مبحث خوشه بندی برای این روش محسوب می شود.

در سال ۲۰۱۳ مطالعه مشابهی توسط بدناریک و کوواکس به منظور تعیین حد آستانه‌ای مطلوب برای به حداکثر رساندن کارایی خوشه‌بندی بر اساس الگوریتم BIRCH انجام شد. آن‌ها حد آستانه مختلف را انتخاب کرده و برای هر کدام زمان و تعداد مراحل رسیدن به خوشه نهایی محاسبه نموده و بر اساس آن‌ها حد آستانه‌ای مناسب پیشنهاد کردند [۱۳]. اما همان‌طور که قبلاً نیز گفته شد در الگوریتم BIRCH تعدیل یافته که هدف اصلی این مقاله هست، حد آستانه بهینه، طی فرایند خوشه‌بندی از روی داده‌ها برآورد می‌شود.

چونگ و همکارانش در سال ۲۰۱۴ مطالعه‌ی خوشه‌بندی دیگری را بر روی داده‌های مربوط به اسکن‌های مغزی fMRI انجام دادند و با استفاده از نتایج مطالعات پیشین به دنبال یافتن مناسب‌ترین حد آستانه‌ای برای خوشه‌بندی‌های گسترده در تصویربرداری‌های عصبی و هم‌چنین تصویربرداری‌های fMRI بودند و از یک مطالعه شبیه‌سازی برای بررسی حد آستانه‌های مختلف استفاده کردند [۱۴].

انتخاب حد آستانه‌ای مناسب در الگوریتم خوشه‌بندی BIRCH نتیجه مستقیمی در کارایی و سرعت عمل خوشه‌بندی داده‌ها دارد. زمان اجرای الگوریتم m-BIRCH با توجه به حد آستانه‌ای برآورد شده آن متناظر با زمان اجرای الگوریتم BIRCH در آن حد آستانه‌ای است و زمان اجرای الگوریتم آن بین بازه‌های حد آستانه‌ای الگوریتم BIRCH است. پیشنهاد می‌شود در تحقیقات آتی الگوریتم‌های BIRCH و m-BIRCH با منطق فازی که رویکرد جدیدی در پژوهش محسوب می‌شود روی داده‌های fMRI انجام شود.

تشکر و قدردانی

این مقاله حاصل پایان‌نامه کارشناسی ارشد آمار زیستی در دانشکده پیراپزشکی دانشگاه علوم پزشکی شهید بهشتی بوده و طرح مربوطه با حمایت‌های مالی این دانشگاه با کد ۱۰۲۱۶ به تصویب رسیده و اجرا شده است. به این وسیله از مسولین مربوطه تقدیر و تشکر می‌گردد.

منابع

- [1] Selvy P, Palanisamy V, Purusothaman T. Performance analysis of clustering algorithms in brain tumor detection of MR images. *Eur J Sci Res* 2011; 62: 321-330.
- [2] Saha M, Panda C. A review on various image segmentation techniques for brain tumor detection. *Int Confer Electron Commun Aerospace Technol* 2018.
- [3] Mayer AR, Bellgowan PS, Hanlon FM. Functional magnetic resonance imaging of mild traumatic brain injury. *Neurosci Biobehav Rev* 2015; 49: 8-18.
- [4] Anuradha R, Sushila R, Pranav M. Performance analysis of K-means and CLARANS clustering algorithm. *Int J Manag Technol Engin* 2018; 8: 374-378.

نتایج خوشه‌بندی BIRCH روی داده‌های تحقیق نشان داد که با افزایش حد آستانه‌ای سرعت اجرای الگوریتم کاهش می‌یابد. دلیل این کاهش سرعت قرار گرفتن تعداد واکسل‌های بیش‌تری درون یک خوشه است که این مساله موجب می‌شود تعداد خوشه‌ها کم‌تر شده و زمان اجرای الگوریتم خوشه‌بندی بیش‌تر شود. هم‌چنین با افزایش هم‌زمان فاکتور شاخه‌بندی و حد آستانه‌ای میزان حساسیت در خوشه‌بندی کاهش یافته و بنابراین انتظار می‌رود در داده مورد مطالعه واکسل‌هایی با شباهت کم‌تر در یک خوشه قرار گرفته باشند.

از منظر وسعت داده‌ها، تصاویر اسکن مغزی دارای ابعاد $47 \times 58 \times 49$ (ارتفاع \times عرض \times طول) بود و در ۷۴ فاصله زمانی متوالی ثبت شده بودند. در نتیجه حجم داده مورد استفاده برابر با نه میلیون و هشتصد و هشتاد و چهار هزار و چهارصد و هفتاد و شش (۹۸۸۴۷۶) واکسل بوده است. همان‌طور که انتظار می‌رفت با وجود حجم بسیار زیاد داده‌ها، الگوریتم‌های مذکور با سرعت بالایی و در زمان اندکی اجرا شده و خوشه‌ها را مشخص کردند.

ویندیشبرگر و همکاران در سال ۲۰۰۳ تحلیل خوشه‌بندی فازی را روی داده‌های fMRI مورد بررسی قرار دادند. در این مطالعه خوشه‌بندی بر اساس Bold بر روی داده‌های fMRI انجام شد و تاثیر انتخاب پارامترهای خوشه‌بندی با استفاده از منحنی Roc بر روی داده‌های شبیه‌سازی شده صورت گرفت. از مزایای این مطالعه استفاده از منطق فازی در خوشه‌بندی داده‌های fMRI بود به این صورت که خوشه‌ها می‌توانستند هم‌پوشانی داشته باشند که با ماهیت داده‌های مغزی سازگار است [۱۱]. در پژوهش ما از منطق کلاسیک برای خوشه‌بندی استفاده شد و لذا از این منظر نتایج حاصله با نتایج مقاله ویندیشبرگر و همکاران قابل مقایسه نیست.

در مطالعه هلر و همکاران در سال ۲۰۰۶ تحلیلی مبتنی بر خوشه‌بندی برای داده‌های fMRI ارائه شد که در آن فعالیت واکسل‌ها (کوچک‌ترین جزء سازنده مغز) به عنوان اجزای خوشه‌ها مورد آزمون قرار گرفتند. از مزایای آن در نظر گرفتن هم‌بستگی مکانی واکسل‌های فعال‌شده با یک‌دیگر بود. به این گونه که ابتدا هم‌بستگی هر واکسل با همسایه‌های محاسبه شد. سپس برای هر واکسل همسایه‌ای که بیش‌ترین هم‌بستگی را داشت مشخص شد. آن‌ها در این پژوهش به این نتیجه رسیدند که در تحلیل fMRI بهتر است که به جای در نظر گرفتن واکسل‌ها از خوشه‌ای از واکسل‌ها استفاده شود [۱۲]. اما در تحقیق ما واحدهای پژوهش واکسل‌ها بودند و با وجود حجم زیاد آن‌ها، خوشه‌بندی در زمان کم و با سرعت مناسبی انجام شد.

- [12] Heller R, Stanley D, Yekutieli D, Benjamini Y. Cluster-based analysis of fMRI data. *NeuroImage* 2006; 33: 599-608.
- [13] Bednarik L, Kovacs L. Efficiency analysis of quality threshold clustering algorithms. *Produc Syst Inform Engin* 2013; 6: 15-26.
- [14] Woo CW, Krishnan A, Wager TD. Cluster-extent based thresholding in fMRI analyses: pitfalls and recommendations. *Neuroimage* 2014; 91: 412-419.
- [15] Borumandnia N, Alavi Majd H, Zayeri F, Baghestani A R, Faeghi F, Tabatabaei SM. Bayesian spatiotemporal model for detecting of active areas in brain for analyzing of fMRI data. *Koomesh* 2017; 19: 845-851. (Persian).
- [16] zolghadr Z, Alavi Majd H, Faeghi F, Niaghi F, Hajizadeh N. Classification of brain stem glioma tumor grade based on MRI findings using support vector machine. *Koomesh* 2017; 19: 584-590. (Persian).
- [5] Filler A. The history, development and impact of computed imaging in neurological diagnosis and neurosurgery: CT, MRI, and DTI. *Int J Neurosurg* 2010; 7: 5-35.
- [6] Lindquist M. The statistical analysis of fMRI data. *Stat Sci* 2008; 23: 439-464.
- [7] Kantardzic M. *Data mining: concepts, models, methods, and algorithms*. John Wiley & Sons 2011.
- [8] Zhang T, Ramakrishnan R, Livny M. BIRCH: an efficient data clustering method for very large databases. *ACM Sigmod Record* 1996.
- [9] Madan S, Dana KJ. Modified balanced iterative reducing and clustering using hierarchies (m-BIRCH) for visual clustering. *Pattern Anal Appl* 2016; 19: 1023-1040.
- [10] www.nitrc.org. last visit: 15 May 2019.
- [11] Windischberger C, Barth M, Lamm C, Schroeder L, Bauer H, Gur RC, Moser E. Fuzzy cluster analysis of high-field functional MRI data. *Artif Intell Med* 2003; 29: 203-223.

Application of modified balanced iterative reducing and clustering using hierarchies algorithm in parceling of brain performance using fMRI data

Navid Valizadeh (M.SC)¹, Soheila Khodakarim (Ph.D)², Seyyed Mohammad Tabatabaei (M.D)³, Azam Saffar (Ph.D Student)¹, Alireza Akbarzadeh Baghban (Ph.D)^{*4}

1 - Department of Biostatistics, School of Paramedical Science, Shahid Beheshti University of Medical Science, Tehran, Iran

2 - Department of Biostatistics, School of Paramedical Science, Shahid Beheshti University of Medical Sciences, Tehran, Iran

3 - Clinical research unit, Imam Reza hospital, Mashhad University of Medical Sciences, Mashhad, Iran

4 - Proteomics Research Center, School of Rehabilitation, Shahid Beheshti University of Medical Sciences, Tehran, Iran

* Corresponding author. +98 09122955411 akbarzad@sbmu.ac.ir

Received: 2 Oct 2019; Accepted: 6 Jun 2020

Introduction: Clustering of human brain is a very useful tool for diagnosis, treatment, and tracking of brain tumors. There are several methods in this category in order to do this. In this study, modified balanced iterative reducing and clustering using hierarchies (m-BIRCH) was introduced for brain activation clustering. This algorithm has an appropriate speed and good scalability in dealing with very large data using a new concept of Clustering Feature.

Materials and Methods: In this study, data from the brain scan had been used. This dataset consisted of 74 consecutive brain scans. After data preprocessing, brain scan images were clustered through the BIRCH and m-BIRCH algorithms. Data were analyzed using WFU-PickAtlas in Matlab software and were compared with the TD Lobes Standard Atlas.

Results: The speed of implementation of the m-BIRCH algorithm decreased as threshold limit increased. The m-BIRCH clustering algorithm showed that there was no specific ascending or descending pattern between branch factor and the run-time of the algorithm. The maximum runtime value of the algorithm was related to the branching factor of 30 which was 94 seconds, equivalent to the upper threshold limit of the BIRCH algorithm.

Conclusion: Applying the m-BIRCH algorithm on high-dimension data set such as brain scan images has relative advantages and provides a tradeoff between time and space complexity. By simultaneously increasing the branching factor and threshold limit, the sensitivity of clustering will be decreased.

Keywords: Brain, Cluster Analysis, Algorithms, Magnetic Resonance Imaging