

تشخیص بیماری اختلالات خفیف شناختی بر اساس انتخاب ویژگی‌های داده با استفاده از کدگذاری تُنک

مرتضی تورانداز^{۱*}، محمد یوسف نژاد^۲، صابر نوریان^۳

^۱ دانشجوی کارشناسی ارشد، دانشکده علوم کامپیوتر و هوش مصنوعی، موسسه آموزش عالی روزبهان، ساری، ایران

mtourandaz@rouzbahan.ac.ir

*مسئول مکاتبات

^۲ دستیار پژوهش، دانشکده علوم و تکنولوژی کامپیوتر، دانشگاه هوا و فضای نانجینگ، چین

myousefnezhad@nuaa.edu.cn

^۳ دانشجوی دکتری، دانشکده برق و الکترونیک، دانشگاه علم و صنعت ایران، تهران، ایران

snourian@elec.iust.ac.ir

چکیده

در سال‌های اخیر تشخیص بیماری آلزایمر به ویژه در مراحل اولیه شکل‌گیری آن یعنی اختلالات خفیف شناختی به یکی از مسائل چالش‌برانگیز در علم پزشکی تبدیل شده است. تا کنون روش‌های مختلفی بر اساس استفاده از ویژگی‌های استخراج‌شده از نشانگرهای زیستی مختلف مغز انسان برای تشخیص این بیماری به کار برده شده است. هدف ارائه این مقاله، پیشنهاد روشی جدید جهت بهبود دقت تشخیص بیماری اختلالات خفیف شناختی مبتنی بر استفاده از کدگذاری تُنک در انتخاب ویژگی‌های مفید از میان ویژگی‌های استخراج‌شده از نشانگرهای زیستی مغز انسان است. در پیاده‌سازی‌های انجام‌شده، دقت روش پیشنهادی ۷۶/۶٪ بدست آمده است که این میزان، ۱۰٪ بهبود را نسبت به روش‌های قبلی نشان می‌دهد.

کلمات کلیدی:

طبقه‌بندی، کدگذاری تُنک، ماشین بردار پشتیبان، نشانگر زیستی ام آر آی، نشانگر زیستی پت، نشانگر زیستی مایع مغزی نخاعی

(Bouwman et al, 2009) و (Mattsson et al, 2009) برای

تشخیص اختلالات به وجود آمده در مغز به دلیل ابتلا به بیماری‌های شناختی مانند آلزایمر و اختلالات خفیف شناختی^۱ استفاده می‌شود. در اغلب کارهای انجام‌شده تنها از اطلاعات حاصل از یک نشانگر زیستی برای طبقه‌بندی افراد مبتلا به بیماری‌های شناختی و افراد سالم استفاده شده است (Notestine et al, 2009). این در حالی است که تحقیقات اخیر نشان می‌دهد ترکیب ویژگی‌های حاصل از نشانگرهای زیستی مختلف، قدرت تشخیص بیماری‌های شناختی را به مراتب افزایش می‌دهد. در واقع نشانگرهای زیستی متفاوت، اطلاعاتی تولید می‌کنند که مکمل یکدیگرند (de Leon et al, 2007) و (Apostolova et al, 2010). اخیراً به منظور ترکیب اطلاعات حاصل از نشانگرهای زیستی مختلف، از ابزارهای یادگیری ماشین استفاده شده است (Zhang, Wang, Zhou, Yuan, Shen, 2011) و (Wee et al, 2012). این ابزارها در تشخیص بیماری‌های شناختی با استفاده از اطلاعات حاصل از این نشانگرهای زیستی مختلف بسیار موثر هستند. هدف از این مقاله، ارائه روشی جدید مبتنی بر استفاده از کدگذاری تُنک^۲ به منظور حذف ویژگی‌های غیرمفید جهت بالا بردن دقت

(۱) مقدمه

در حال حاضر بیماری آلزایمر^۱ یکی از دلایل مهم زوال عقل در افراد سالخورده است. تحقیقات به عمل آمده نشان می‌دهد که تعداد افراد مبتلا به این بیماری تا ۲۰ سال آینده دو برابر تعداد فعلی و در سال ۲۰۵۰ از هر ۸۵ نفر یک نفر به این بیماری مبتلا خواهد بود (Brookmeyer, Johnson, Ziegler, Arrighi, 2007). احتمال ابتلا به بیماری آلزایمر در افراد مبتلا به اختلالات خفیف شناختی بسیار بالاست، بنابراین تشخیص ابتلا به اختلالات خفیف شناختی به منظور پیش‌گیری از ابتلا به بیماری آلزایمر بسیار بااهمیت خواهد بود (Thompson, Apostolova, 2007). در حال حاضر از نشانگرهای زیستی مختلفی مانند ۱- مقدار نرون‌های از دست رفته نواحی مختلف مغز و اتصالات تخریب‌شده^۲ بین آن‌ها که از تصاویر ساختاری^۳ ام آر آی مغز قابل استخراج است (de Leon et al, 2007) ۲- هیپومتابولیسیم اندازه‌گیری شده به وسیله تصویربرداری کارکردی^۴ (Morris et al, 2001) و ۳- تعیین مقدار پروتئین مغز به وسیله مایع مغزی نخاعی^۵

¹ Alzheimer's disease

² Atrophy

³ Brain structural imaging

⁴ Brain functional imaging

⁵ Cerebrospinal Fluid (CSF)

⁶ Mild Cognitive Impairment (MCI)

⁷ Sparse coding

یادگیری چند هسته‌ای با هم ادغام کرده و به روشی جدید و با دقت بالا برای تشخیص بیماری‌های شناختی رسیده‌اند (Biao et al, 2014).

وی (Wee) و همکاران با به‌کارگیری کدگذاری تُنک روی شبکه‌های کارکردی حاصل از تصاویر کارکردی مغز روشی جدید برای تشخیص بیماری آلزایمر ارائه داده‌اند (Wee, Yap, Zhang, Wang, Shen, 2012). آن‌ها در این روش برای ساخت شبکه کارکردی مغز از کدگذاری تُنک به منظور حذف ارتباطات غیرمفید استفاده کرده‌اند. چنگ یو وی و همکاران نشان داده‌اند که استفاده از کدگذاری تُنک برای ساخت شبکه کارکردی مغز و استفاده از ابزار یادگیری ماشین نهایتاً منجر به بالا رفتن دقت در تشخیص بیماری‌های شناختی می‌شود.

۳) اندیشه جدید

نکته کلیدی در روش پیشنهادی در این مقاله استفاده از کدگذاری تُنک برای انتخاب ویژگی‌های مفید از میان ویژگی‌هایی است که از نشانگرهای زیستی مغز استخراج می‌گردند. با استفاده از ویژگی‌های انتخاب شده و بر اساس یادگیری چند هسته‌ای، ماشین بردار پشتیبان^۴ جهت طبقه‌بندی افراد مبتلا به اختلالات خفیف شناختی و افراد سالم، آموزش داده می‌شود. نتایج نشان می‌دهد که استفاده از کدگذاری تُنک برای انتخاب ویژگی‌ها موجب بالا رفتن دقت این طبقه‌بندی می‌شود.

• طبقه‌بندی چند هسته‌ای

در این بخش یک مرور کلی روی روش‌های یادگیری مبتنی بر هسته خواهیم داشت. این روش‌ها به راحتی قابل‌اعمال به ماشین بردار پشتیبان جهت طبقه‌بندی داده‌ها هستند (Scholkopf, Smola, 2002). قبل از پرداختن به روش‌های چند هسته‌ای، ابتدا ماشین بردار تک هسته‌ای را مورد بررسی قرار می‌دهیم. ایده اصلی ماشین بردار پشتیبان مبتنی بر هسته به صورت خلاصه بدین شرح است: ابتدا داده‌هایی که به صورت خطی جدا پذیر نیستند از فضای اصلی به فضایی با ابعاد بالاتر^۵ که در آن داده‌ها به صورت خطی جدا پذیر باشند نگاشت می‌شود. این نگاشت از طریق توابع هسته^۶ قابل‌اعمال است. سپس مدل مورد نظر با بیشینه مقدار حاشیه برای طبقه‌بندی اطلاعات ساخته می‌شود.

برای ساخت مدل بر اساس ترکیب ویژگی‌های حاصل از نشانگرهای زیستی ام آر آی، پت و مایع مغزی نخاعی جهت تشخیص اختلالات خفیف شناختی از ماشین بردار پشتیبان چند هسته‌ای^۷ استفاده می‌شود (Zhang et al, 2011). برای

طبقه‌بندی باهدف تشخیص بیماری اختلالات خفیف شناختی در افراد است. در این مقاله از داده‌هایی استفاده می‌شود که نشان‌دهنده ویژگی‌های استخراج شده از سه نشانگر زیستی ام آر آی (MRI)، پت (PET) و مایع مغزی نخاعی (CSF)، مربوط به افراد مبتلا به اختلالات خفیف شناختی و افراد سالم است.

در ادامه مقاله ابتدا در بخش دوم به بررسی کارهای انجام شده خواهیم پرداخت، در بخش سوم اندیشه جدید، ابزارهای مورد استفاده در آن، داده‌های مورد استفاده و نتایج حاصل را مورد بررسی و تحلیل قرار داده و نهایتاً در بخش چهارم نتیجه‌گیری و خط و مشی کارهای آتی بیان خواهد شد.

۲) کارهای انجام شده

برای تشخیص بیماری‌های شناختی مطالعات بسیاری انجام و روش‌های زیادی ارائه شده است. در اکثر تحقیقات، اطلاعات حاصل از تصویربرداری ساختاری و کارکردی مغز برای تشخیص این بیماری‌ها مورد مطالعه قرار گرفته است (Wee et al, 2012). در تحقیقات و کارهای اولیه تشخیص این بیماری، بر اساس استفاده از ویژگی‌های حاصل یک نشانگر زیستی بوده است. اخیراً بر اساس مطالعات انجام گرفته، دانشمندان دریافته‌اند که استفاده از اطلاعات حاصل از ترکیب ویژگی‌های این نشانگرهای زیستی موجب بالا رفتن دقت در تشخیص بیماری آلزایمر می‌شود. وموری (Vemuri) و همکاران، تنها از نشانگرهای زیستی ام آر آی و مایع مغزی نخاعی برای تشخیص بیماری‌های شناختی (Vemuri et al, 2007) و فلگیبل (Fellgiebel) و همکاران از دو نشانگر زیستی پت و مایع مغزی نخاعی برای تشخیص این بیماری‌ها بهره برده‌اند (Fellgiebel, Scheurich, Bartenstein, Muller, 2007). جانگ (Zhang) و همکاران با استفاده از ویژگی‌های استخراج شده از سه نشانگر زیستی ام آر آی، پت و مایع مغزی نخاعی روشی مبتنی بر یادگیری چند هسته‌ای ارائه داده‌اند (Zhang et al, 2011). جانگ و همکاران نشان داده‌اند که نتایج به دست آمده بر اساس ترکیب ویژگی‌های استخراج شده از نشانگرهای زیستی متفاوت به مراتب بهتر از نتایج به دست آمده با استفاده از تنها یک نشانگر زیستی است.

جائه (Biao) و همکاران با استفاده از تصاویر کارکردی مغز و تشکیل شبکه ارتباطی بین نواحی مشخص شده از مغز توانسته‌اند روشی جدید برای تشخیص بیماری‌های شناختی ارائه دهند. آن‌ها با استفاده از روش‌های تحلیل شبکه‌های ارتباطی، ویژگی‌های محلی^۱ و توپولوژیک^۲ شبکه را استخراج و بعد از ساخت ماتریس هسته^۳ هر کدام، آن‌ها را با کمک

⁴ Support Vector Machine (SVM)

⁵ Feature space

⁶ Kernel function

⁷ Multikernel Support Vector Machine (MKLSVM)

¹ Local feature

² Topological feature

³ Kernel matrix

بررسی ماشین بردار پشتیبان چند هسته‌ای فرض می‌کنیم n داده آموزشی و m وجه (به طور مثال نشانگر زیستی) داریم. $X_i^{(m)}$ نشان‌دهنده بردار ویژگی m -امین وجه از i -امین داده آموزشی است و برچسب معادل آن y_i یکی از دو مقدار -1 و 1 است. ماشین بردار پشتیبان چند هسته‌ای تابع هدف در رابطه ۱ را بهینه می‌کند.

$$\min_{w^{(m)}, b, \xi_i} \frac{1}{2} \sum_{m=1}^M \beta_m \|w^{(m)}\|^2 + C \sum_i \xi_i \quad (1)$$

$$s. t. y_i \left(\sum_{m=1}^M \beta_m ((w^{(m)})^T \phi^{(m)}(x_i^{(m)} + b) \right) \geq 1 - \xi_i$$

در رابطه ۱، $w^{(m)}$ ، $\phi^{(m)}$ و $\beta_m \geq 0$ به ترتیب نشان‌دهنده بردار وزن، تابع هسته نگاشت و وزن هر وجه برای ترکیب ویژگی‌هاست. فرم دوگان ماشین بردار پشتیبان همانند فرم اولیه آن به دنبال بهینه کردن تابع معیار در رابطه ۲ آورده شده است.

$$\max_{\alpha} \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{ij} \alpha_i \alpha_j y_i y_j \sum_{m=1}^M \beta_m k^{(m)}(x_i^{(m)}, x_j^{(m)}) \quad (2)$$

$$s. t. \sum_{i=1}^n \alpha_i y_i = 0$$

$$0 \leq \alpha_i \leq C, i = 1, \dots, n$$

برای داده‌ی آزمایشی جدید $x = \{x^{(1)}, x^{(2)}, \dots, x^{(M)}\}$ ابتدا مقدار $k^{(m)}(x_i^{(m)}, x_j^{(m)}) = \phi^{(m)}(x_i^{(m)})^T \phi^{(m)}(x_j^{(m)})$ که نشان‌دهنده مقدار تابع هسته مورد استفاده میان داده آزمایشی و تمام داده‌های آموزشی روی m -امین وجه است، محاسبه می‌گردد. سپس رابطه ۳ برای پیش‌بینی برچسب داده آزمایشی اعمال می‌گردد.

$$f(x^{(1)}, x^{(2)}, \dots, x^{(M)}) = \text{sign} \left(\sum_{i=1}^n \alpha_i y_i \sum_{m=1}^M \beta_m k^{(m)}(x_i^{(m)}, x_j^{(m)} + b) \right) \quad (3)$$

گ α ر $k(x_i, x_j) = \sum_{m=1}^M \beta_m k^{(m)}(x_i^{(m)}, x_j^{(m)})$ به عنوان تابع هسته‌ی ادغام‌شده بین دو داده‌ی آموزشی x_i و x_j و $k(x_i, x) = \sum_{m=1}^M \beta_m k^{(m)}(x_i^{(m)}, x^{(m)})$ به عنوان تابع هسته‌ی ادغام‌شده بین داده‌ی آموزشی x_i و داده‌ی آزمایشی x در نظر بگیریم، آنگاه می‌توان ماشین بردار پشتیبان چند هسته‌ای را به وسیله‌ی ماشین بردار تک هسته‌ای پیاده‌سازی کرد. در رابطه‌های فوق این محدودیت وجود دارد که $\sum_m \beta_m = 1$ می‌توان با قرار دادن مقادیر مختلف برای هر β_m و

در نظر گرفتن محدودیت ذکرشده و با استفاده از اعتبارسنجی k -تکه‌ای^۱ روی داده‌های آموزشی بهترین مقادیر را برای این متغیرها یافت. در این مقاله برای آموزش بردار پشتیبان از ابزار LIBSVM (Chang, Lin, 2001) استفاده شده است.

• کدگذاری تُنک

نکته کلیدی در این مقاله استفاده از کدگذاری تُنک جهت انتخاب ویژگی‌های مفید برای تشخیص اختلالات خفیف شناختی در افراد است. انتخاب ویژگی‌های مفید با بهینه کردن تابع معیار رابطه ۴ امکان پذیراست.

$$\min_{w, b} \frac{1}{2} \sum_i^N (y_i - w^{(T)} x_i - b)^2 + \lambda \|w\|_1 \quad (4)$$

در رابطه ۴، x_i نشان‌دهنده بردار ویژگی، y_i نشان‌دهنده برچسب کلاسی است که داده‌ی x_i به آن تعلق دارد، b مقدار عرض از مبدا داده x_i ، w ضرایب رگرسیون بردار ویژگی داده x_i ، λ پارامتر تنظیم و N تعداد داده آموزشی است. $\|w\|_1$ ، تنظیم‌کننده نُرم L_1 است که برای کدگذاری تُنک در فضای ویژگی مورد استفاده قرار می‌گیرد. با استفاده از این روش ویژگی‌های معادل با ضرایبی که با بهینه کردن رابطه ۴ صفر می‌شوند، از بردار ویژگی حذف می‌گردند. بهترین مقدار برای پارامتر تنظیم ($\lambda > 0$) در رابطه ۴ با استفاده از عمل اعتبارسنجی k -تکه‌ای روی داده‌های آموزشی قابل محاسبه است. در این مقاله از ابزار LASSO (Tibshirani, 1996) برای بهینه کردن رابطه ۴ استفاده می‌شود.

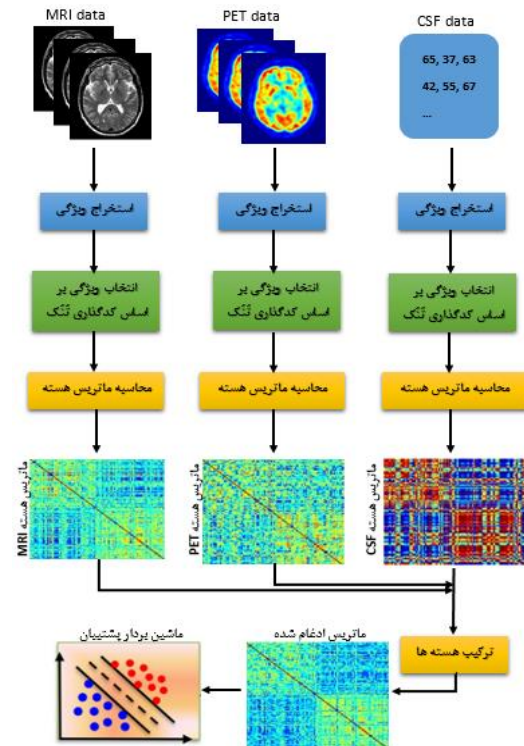
در شکل ۱ شمای کلی روش پیشنهادی به تصویر کشیده شده است. همان طور که در این شکل می‌بینید ابتدا ویژگی‌های مورد نظر از نشانگرهای زیستی ام آر آی، پت و مایع مغزی نخاعی، استخراج شده و در مرحله بعد با استفاده از روش کدگذاری تُنک، ویژگی‌های غیرمفید برای تشخیص اختلالات خفیف شناختی حذف می‌گردند، سپس با استفاده از ویژگی‌های باقیمانده ماتریس هسته‌ی مربوط به هر کدام از نشانگرهای زیستی فوق، محاسبه و باهم ادغام می‌شوند. در نهایت با استفاده از ماشین بردار پشتیبان و بر اساس ماتریس‌های ترکیب‌شده عمل طبقه‌بندی روی افراد مبتلا به اختلالات خفیف شناختی و افراد سالم به منظور بالا بردن دقت تشخیص ابتلای افراد به اختلالات خفیف شناختی انجام می‌پذیرد.

در شکل ۲ شبه کد مربوط به روش پیشنهادی آورده شده است. ورودی تابع تعریف‌شده ویژگی‌های استخراج‌شده از نشانگرهای زیستی ام آر آی، پت و مایع مغزی نخاعی و خروجی آن ماشین بردار پشتیبان آموزش داده‌شده جهت تشخیص بیماری اختلالات خفیف شناختی است.

¹ K-fold cross validation

• اعتبارسنجی

برای سنجیدن کارایی ماشین بردار پشتیبان آموزش داده شده از اعتبارسنجی k -تکه‌ای به منظور محاسبه دقت طبقه‌بندی استفاده می‌کنیم. برای این کار مجموعه داده‌های موجود را به ۱۰ زیرمجموعه افراز می‌کنیم. در هر بار از ۹ زیرمجموعه برای آموزش ماشین بردار پشتیبان و از یک زیرمجموعه باقیمانده برای سنجش ماشین بردار پشتیبان آموزش داده شده استفاده می‌کنیم. این کار به تعداد ۱۰ بار به منظور ساختن زیرمجموعه‌هایی با ترکیب داده‌های مختلف و اطمینان از رخ ندادن Overfitting مدل، تکرار می‌شود.



شکل ۱) نمایش شماتیک روند کلی کار برای طبقه‌بندی داده‌های چندوجهی^۱

برای یافتن وزن بهینه ماتریس هسته هر کدام از m نشانگر زیستی (β_m) به منظور ترکیب هسته‌ها، از داده‌های آموزشی استفاده می‌کنیم. بدین ترتیب که با قرار دادن مقادیر مختلف از ۰ تا ۱ و اندازه گام ۰,۱ برای β_m ها و آموزش مدل بر اساس این مقادیر (با در نظر گرفتن محدودیت $\sum \beta_m = 1$), دقت مدل را بر حسب داده‌های آزمایشی می‌سنجیم و وزن‌هایی که موجب بیشترین دقت برای مدل شده‌اند را به عنوان وزن هسته‌ها در نظر می‌گیریم.

• تجزیه و تحلیل داده‌ها

داده‌های مورد استفاده در این مقاله از پایگاه‌های داده آزمایشگاه تصویربرداری عصبی بیماری آلزایمر^۲ (ADNI) استخراج شده است.^۳ افرادی که در این آزمایشگاه مورد آزمایش قرار می‌گیرند دارای سنی بین ۵۵ و ۹۰ سال بوده و عبارت‌اند از ۱- افراد سالم^۴، ۲- افراد مبتلا به اختلالات خفیف شناختی و ۳- افراد مبتلا به بیماری آلزایمر. در این مقاله از اطلاعات مربوط به ۱۵۱ فرد مورد آزمایش در ADNI که هر سه نشانگر زیستی ام آر آی، پت و مایع مغزی نخاعی برای آن‌ها محاسبه شده، استفاده می‌کنیم که عبارت‌اند از: ۹۹ فرد مبتلا به اختلالات خفیف شناختی و ۵۲ فرد سالم.

Pseudocode for the proposed method

Inputs: Feature vector of MRI biomarker
Feature vector of PET biomarker
Feature vector of CSF biomarker

Output: SVM model for diagnose Alzheimer's disease
Function AD_HC_SVM (Inputs)

1. Select useful features with sparse coding
2. Compute kernel matrices
3. Compute weight of each kernel matrix
4. Combine all matrices
5. Train SVM model for diagnose MCI disease via K-fold cross validation
6. Output=SVM model

End function

شکل ۲) شبه کد روش پیشنهادی

برای استخراج ویژگی‌های مربوط به نشانگرهای زیستی ام آر آی و پت، تصاویر مغز حاصل از تصویربرداری ساختاری و کارکردی به ۹۳ ناحیه^۵ تقسیم می‌شود. برای تصاویر ساختاری، سطح ماده خاکستری هر ناحیه از مغز به عنوان ویژگی آن در نظر گرفته شده است. برای تصاویر کارکردی بعد از اعمال پیش‌پردازش‌هایی همچون حذف حرکات اضافی سر^۶، اصلاح زمانی سری زمانی تصاویر^۷ برای هر فرد و اصلاح مکانی^۸ هر کدام از تصاویر موجود در سری زمانی نسبت به اولین تصویر موجود در این سری زمانی برای هر فرد، مقدار میانگین شدت هر ناحیه به عنوان ویژگی آن در نظر گرفته می‌شود. برای نشانگر زیستی مایع مغزی نخاعی نیز سه مقدار CSF A β 42, CSF t-tau, CSF p-tau استفاده شده است (Zhang et al, 2011); بنابراین برای هر فرد در مجموع ۱۸۹ ویژگی در نظر گرفته می‌شود.

• نتایج

برای انجام مقایسه و سنجیدن دقت روش پیشنهادی، روش‌های درخت تصمیم^۹، شبکه عصبی^{۱۰}، k -نزدیک‌ترین

² Alzheimer's Disease Neuroimaging Initiative

³ www.loni.ucla.edu/ADNI

⁴ Healthy subjects

⁵ Region of interest

⁶ Motion correction

⁷ Slice timing

⁸ Realignment

⁹ Decision Tree (DT)

¹⁰ Neural Network (NN)

¹ Multimodal

جدول ۱) نتایج حاصل از اجرای روش‌های طبقه‌بندی روی داده‌ها بر حسب میانگین دقت \pm خطای استاندارد در ۱۰ بار اجرای مستقل

روش پیشنهادی	ماشین بردار پشتیبان	K-نزدیک‌ترین همسایه	شبکه عصبی	درخت تصمیم	نشان گریستی
۸۱/۰۲±۰/۳۵	۶۷/۵۷±۱/۱	۶۴/۱۱±۰/۵۳	۴۶/۷۲±۰/۸۷	۶۵/۵۵±۰/۹۵	MRI
۷۸/۱۶±۰/۲۸	۶۹/۱۲±۰/۶۴	۶۰/۰۵±۰/۸۹	۴۸/۳۷±۰/۹۴	۶۰/۹۱±۰/۷۹	PET
۶۹/۸۱±۰/۲	۶۹/۱۴±۰/۲۹	۶۴/۴۶±۰/۵۴	۴۶/۲۵±۱/۱	۶۲/۸۸±۱/۰۲	CSF
۷۸/۱۶±۰/۳۲	۶۵/۴۵±۰/۶۵	۶۴/۶۳±۰/۹۵	۵۰/۱۵±۰/۷۸	۶۲/۷۵±۰/۷۸	MRI+PET
۶۵/۶۲±۰/۰۲	۶۲/۵۷±۰/۷۱	۶۳/۶۲±۰/۸۲	۴۹/۶۳±۰/۸۵	۶۲/۴۳±۰/۵۳	MRI+CSF
۷۸/۴۷±۰/۳۷	۶۲/۸۱±۱/۱۸	۶۳/۴۲±۱/۰۴	۴۸/۲±۰/۶۸	۶۰/۴۶±۱/۳۸	CSF+PET
۸۵/۱±۰/۳	۷۰/۹۹±۰/۶۵	۶۴/۹۳±۰/۷۱	۵۱/۹۵±۰/۷۳	۶۲/۷۴±۰/۸۲	MRI+PET+CSF

در روش‌های اولیه از نشانگرهای زیستی مختلف به صورت مجزا برای انجام عمل طبقه‌بندی به منظور تشخیص بیماری‌های شناختی در افراد همچون اختلالات خفیف شناختی استفاده شده است، این در حالی است که نتایج تحقیقات اخیر نشان می‌دهد ترکیب اطلاعات حاصل از نشانگرهای زیستی مختلف مغز موجب بهبود دقت تشخیص بیماری‌های شناختی در افراد می‌گردد.

کدگذاری تَنک روشی است که ما در این مقاله به منظور انتخاب ویژگی‌های مفید از میان ویژگی‌های استخراج شده از نشانگرهای زیستی مغز به منظور اکتساب اطلاعات موثر از این نشانگرها و به دنبال آن ترکیب این اطلاعات جهت طبقه‌بندی افراد مبتلا به اختلالات خفیف شناختی و افراد سالم به منظور بالا بردن دقت تشخیص بیماری اختلالات خفیف شناختی، استفاده کرده‌ایم.

تاکنون مطالعات بسیار زیادی برای تشخیص بیماری‌های شناختی انجام و روش‌های بسیار متنوعی بر اساس نشانگرهای مختلف زیستی مغز برای بالا بردن قدرت تشخیص این بیماری‌ها ارائه شده است. مطالعات اخیر نشان می‌دهد که شبکه ارتباطی کارکردی مغز^۲ که با استفاده از تصاویر کارکردی مغز ساخته می‌شود و نشان‌دهنده چگونگی تعامل نواحی مختلف مغز با یکدیگر در شرایط گوناگون است، اطلاعات بسیار مفیدی برای تشخیص بیماری‌های شناختی فراهم می‌آورد. ما در کارهای آتی به بررسی روش‌های مختلف استخراج ویژگی با استفاده از مفاهیم موجود در تحلیل شبکه‌های ارتباطی جهت استخراج ویژگی‌های موثر از شبکه‌های کارکردی مغز به منظور بالا بردن دقت در تشخیص بیماری‌های شناختی خواهیم پرداخت.

همسایه^۱ و ماشین بردار پشتیبان را نیز به منظور طبقه‌بندی داده‌های مورد نظر اجرا کرده‌ایم که نتایج در جدول ۱ نشان داده شده است. روش پیشنهادی در محیط MATLAB پیاده‌سازی و نتایج آزمایش روی میانگین و خطای استاندارد ۱۰ بار آموزش و آزمایش مدل در جدول ۱ گزارش شده است. آموزش مدل با استفاده از اعتبارسنجی ۱۰-تکه‌ای انجام شده است. سطرهای ۱ تا ۳ این جدول نتایج به دست آمده از اجرای روش‌های گفته شده با استفاده از هر یک از نشانگرهای زیستی، در سطرهای چهارم تا ششم، نتایج به دست آمده از طبقه‌بندی داده‌های مورد استفاده در این مقاله با استفاده از ترکیب‌های دوتایی مختلف از ۳ نشانگر زیستی و در سطر آخر نیز نتایج بر اساس ترکیب هر سه این نشانگرها نشان داده شده است. نتایج حاصل از روش پیشنهادی در آخرین ستون آورده شده است. این نتایج نشان می‌دهد که روش پیشنهادی موجب افزایش چشمگیری در تشخیص اختلالات خفیف شناختی می‌شود.

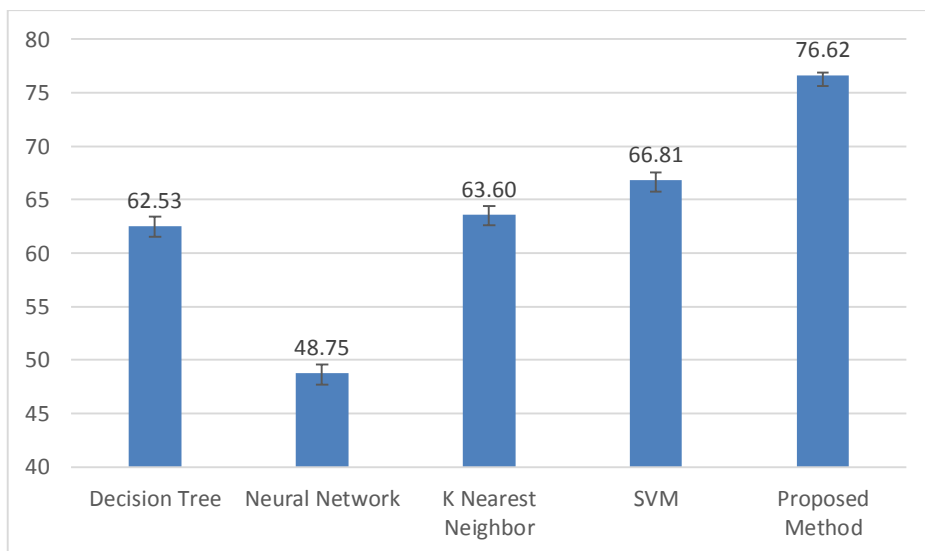
در شکل ۳ نمودار مربوط به میانگین دقت و خطای استاندارد به دست آمده در پیاده‌سازی روش‌های گفته شده آورده شده است. این نمودار نشان می‌دهد که روش پیشنهادی دارای بیشترین مقدار میانگین دقت و کمترین مقدار میانگین خطای استاندارد در میان سایر روش‌هاست.

۴) نتیجه‌گیری

در این مقاله طرحی جدید مبتنی بر استفاده از کدگذاری تَنک به منظور انتخاب ویژگی‌های مفید از میان ویژگی‌های استخراج شده از نشانگرهای زیستی مغز انسان، برای بالا بردن دقت تشخیص بیماری اختلالات خفیف شناختی ارائه شده است. این روش موجب بالا رفتن دقت تشخیص بیماری اختلالات خفیف شناختی نسبت به روش‌های قبلی می‌شود.

² Brain functional connectivity network

¹ K-nearest neighbor (KNN)



شکل ۳) نمودار میانگین درصد صحت و خطای استاندارد روش‌های مختلف

- Hampel, H., Scheltens, P., Pirtila, T., Wallin, A., Jonhagen, M. E., Minthon, L., Winblad, B., Blennow, K. (2009). "CSF biomarkers and incipient Alzheimer disease in patients with mild cognitive impairment". *Jama*.302:385–393.
- Morris, J. C., Storandt, M., Miller, J. P., McKeel D. W., Price, J. L., Rubin, E. H., Berg, L. (2001). "Mild Cognitive Impairment Represents Early-Stage Alzheimer Disease" *Archives of Neurology*.58:397–405.
- Notestine, C., Roddey, J. C., Hagler, D. J. Jr., Holland, D., Karow, D. S, Pung, C. J., Brewer, J. B., Dale, A. M. (2009) "Alzheimer disease: quantitative structural neuroimaging for detection and prediction of clinical and structural changes in mild cognitive impairment". *Radiology*. 2009;251:195–205.
- Scholkopf, B., Smola, A. (2002) "Learning with Kernels". The MIT Press.
- Thompson, P. M., Apostolova L. G. (2007), "Computational anatomical methods as applied to ageing and dementia," *Brit. J. Radiol.*, vol. 80, pp. S78–S91.
- Tibshirani, R. (1996). "Regression shrinkage and selection via the Lasso," *J. Royal Statistical Soc. Series B Methodol.*, vol. 58, pp. 267–288.
- Vemuri, P., Wiste, H. J., Weigand, S. D., Shaw, L. M., Trojanowski, J. Q., Weiner, M. W., Knopman, D. S., Petersen, R. C., Jack, C. R. (2007). "MRI and CSF biomarkers in normal, MCI, and AD subjects", *AAN Enterprises*.
- Wee, C. Y., Yap, P. T., Zhang D. Q., Denny, K., Browndyke, J. N., Potter, G. G., Welsh-Bohmer, K. A., Wang, L. H., Shen, D. G. (2012) "Identification of MCI individuals using structural and functional connectivity networks," *Neuroimage*, vol. 59, pp. 2045–2056, Feb. 1.
- Wee, C. Y., Yap, P. T., Zhang, D. Q., Wang, L., Shen, D. (2012). "Constrained Sparse Functional Connectivity Networks for MCI Classification". *MICCAI 2012, Part II, LNCS 7511*, pp. 212–219.
- Zhang, D., Wang, Y., Zhou, L., Yuan, H., Shen, D. (2011). "Multimodal classification of Alzheimer's disease and mild cognitive impairment," *Neuroimage*, vol. 55, pp. 856–867, Apr. 1.
- Apostolova, L. G., Hwang, K. S., Andrawis, J. P., Green, A. E., Babakchianian, S., Morra, J. H., Cummings, J. L., Toga, A. W., Trojanowski, J. Q., Shaw, L. M., Jack, C. R. Jr., Petersen, R. C., Aisen, P. S., Jagust, W. J., Koeppe, R. A., Mathis, C. A., Weiner, M. W., Thompson, P. M. (2010). "3D PIB and CSF biomarker associations with hippocampal atrophy in ADNI subjects. *Neurobiol Aging*". 31:1284–1303.
- Brookmeyer, R., Johnson, E., Ziegler-Graham, K., and Arrighi, H. M. (2007). "Forecasting the global burden of Alzheimer's disease," *Alzheimer's Dementia*, vol. 3, pp. 186–191.
- Bouwman, F. H., van der Flier, W. M., Schoonenboom, N. S., van Elk, E. J., Kok, A., Rijmen, F., Blankenstein, M. A., Scheltens, P. (2007). "Longitudinal changes of CSF biomarkers in memory clinic patients. *Neurology*".69:1006–1011.
- Biao, J., Zhang, D., Gao, W., Wang, Q., Wee, C. Y., Shen, D. (2014). "Integration of Network Topological and Connectivity Properties for Neuroimaging Classification" *IEEE transactions on biomedical engineering*, vol. 61, no. 2.
- Chang, C. C., Lin, C. J. (2001). "LIBSVM: a library for support vector machines".
- de Leon, M. J., Mosconi, L., Li, J., De Santi, S., Yao Y, Tsui, W. H., Pirraglia, E., Rich, K., Javier, E., Brys, M., Glodzik, L., Switalski, R., Saint Louis, L., A., Pratico, D. (2007). "Longitudinal CSF isoprostane and MRI atrophy in the progression to AD. *J Neurol*". 254:1666–1675.
- Fellgiebel, A., Scheurich, A., Bartenstein, P., Muller, M. J. (2007). "FDG-PET and CSF phospho-tau for prediction of cognitive decline in mild cognitive impairment. *Psychiatry Res*". 155:167–171.
- Mattsson, N., Zetterberg, H., Hansson, O., Andreasen, N., Parnetti, L., Jonsson, M., Herukka, S. K., van der Flier, W. M., Blankenstein, M. A., Ewers, M., Rich, K., Kaiser, E., Verbeek, M., Tsolaki, M., Mulugeta, E., Rosen, E., Aarsland, D., Visser, P. J., Schroder, J., Marcusson, J., de Leon, M.,

منابع