

طبقه‌بندی صداهای طبیعی و غیرطبیعی ضبط‌شده قلب با استفاده از آنالیز زمان – فرکانس سیگنال‌های PCG

هانیه حاضری^۱، کارشناس ارشد؛ قاسم عازمی^۲، دانشیار؛ پگاه زرجام^۳، استادیار

۱- دانشکده مهندسی برق و کامپیوتر - دانشگاه رازی - کرمانشاه - ایران - hazeri.hanie@stu.razi.ac.ir

۲- دانشکده مهندسی برق و کامپیوتر - دانشگاه رازی - کرمانشاه - ایران - g.azemi@razi.ac.ir

۳- دانشکده مهندسی برق و کامپیوتر - دانشگاه رازی - کرمانشاه - ایران - p.zarjam@razi.ac.ir

چکیده: سیگنال صوتی تولید شده ناشی از فعالیت‌های مکانیکی قلب، اطلاعات مفیدی در رابطه با عملکرد دریچه‌های قلبی فراهم می‌کند. اما به دلیل محدودیت شنوایی انسان، ماهیت گذرا و غیر ایستادن سیگنال صدای قلب و انرژی پایین‌تر صداهای پاتولوژیک نسبت به صداهای طبیعی، یافتن نشانه‌های بیماری و تصمیم‌گیری بر مبنای صداهای شنیده‌شده از طریق گوشی پزشکی کار دشواری بوده و نیاز به تمرین و تکرار زیادی دارد. به دلیل محتوای تشخیصی بالای سوافل‌ها در هر دو حوزه زمان و فرکانس، استخراج ویژگی‌های زمان-فرکانس مناسب‌ترین روش برای پردازش این صداهای غیرطبیعی به شمار می‌روند. در این تحقیق به منظور طبقه‌بندی صداهای قلبی، ویژگی‌های حوزه زمان-فرکانس از سیگنال‌های صدای قلب استخراج شده است. در مرحله طبقه‌بندی، از ترکیب دو طبقه‌بند AdaBoost و شبکه عصبی کانولوشن استفاده شده و در نهایت عملکرد روش پیشنهادی با استفاده از روش leave-one-out روش پیشنهادی مورد ارزیابی قرار گرفته است. این روش بر روی پایگاه داده چالش ۲۰۱۶ فیزیونت پیاده‌سازی شده است. نتایج حاصل نشان‌دهنده عملکرد بهتر راهکار پیشنهادی در مقایسه با بهترین روش موجود در چالش ۲۰۱۶ فیزیونت و دست یابی به حساسیت ۹۳/۲۷٪ و اختصاصیت ۸۱/۹۶٪ در طبقه‌بندی صداهای قلبی است در حالی که بهترین روش در چالش ۲۰۱۶ فیزیونت به حساسیت ۹۳/۴۸٪ و اختصاصیت ۸۰/۳۶٪ دست یافته است.

واژه‌های کلیدی: سیگنال PCG، استخراج ویژگی، طبقه‌بندها، بیماری‌های دریچه‌ای، آنالیز زمان-فرکانس.

Classification of Normal/Abnormal Heart Sound Recordings Using Time –Frequency PCG Signal Analysis

H. Hazeri¹, Msc; G. Azemi², Associate Professor; P. Zarjam³, Assistant Professor

1- Faculty of Electrical and Computer Engineering, Razi University, Kermanshah, Iran, Email: hazeri.hanie@stu.razi.ac.ir

2- Faculty of Electrical and Computer Engineering, Razi University, Kermanshah, Iran, Email: g.azemi@razi.ac.ir

3- Faculty of Electrical and Computer Engineering, Razi University, Kermanshah, Iran, Email: p.zarjam@razi.ac.ir

Abstract: The heart's acoustic signal produced by its mechanical activity can provide useful information on the condition of heart valves. The heart sound auscultation, i.e. listening to the heart sounds with a stethoscope, is therefore a primary method for evaluating the cardiovascular function. This method has advantages of being fast, inexpensive, easy to use and noninvasive. On the other hand, due to the transient and non-stationary nature of PCG signals and auscultatory limitations, the correct medical diagnosis based on the heart sound through a stethoscope requires a lot of expertise and needs referral of the patient to a cardiologist. This is not only time-consuming but also imposes a financial burden on the medical system. Thus, automated detection and analysis of the recorded heart sound auscultation has received a lot of attentions in recent years. Even, this was put to the challenge by the PhysioNet/CinC in 2016. This research is based on the non-stationary nature of PCG signals and proposes a new method based on time-frequency analysis of such signals with the aim to classify heart sounds into normal and abnormal sounds. The proposed methodology uses time-frequency features and two classifiers; AdaBoost and CNN. The publicly available 2016 PhysioNet/CinC 2016 Challenge database was used to evaluate the performance of the proposed method using a leave-one-out cross validation. The experimental results show that the proposed method performs very well and has 1.5% higher sensitivity compared to the best existing method.

Keywords: PCG signal, Feature Extraction, Classifiers, Valve diseases, Time –Frequency Analysis.

تاریخ ارسال مقاله: ۱۳۹۷/۰۶/۲۴

تاریخ اصلاح مقاله: ۱۳۹۷/۰۹/۱۲، ۱۳۹۷/۱۱/۱۰ و ۱۳۹۷/۱۲/۱۵

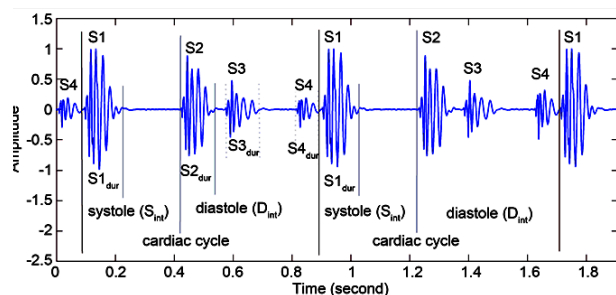
تاریخ پذیرش مقاله: ۱۳۹۸/۰۱/۰۱

نام نویسنده مسئول: قاسم عازمی

نشانی نویسنده مسئول: ایران - کرمانشاه - خیابان دانشگاه - دانشگاه رازی - دانشکده فنی و مهندسی

۱- مقدمه

ناحیه قفسه سینه، در اثر گردش خون در قلب صورت می‌گیرد. سیگنال خروجی این دستگاه، اطلاعات مفیدی را از وضعیت گردش خون نیز به همراه دارد و ظرفیت‌های ویژه‌ای را برای تشخیص بیماری‌های قلب و عروق و امراض مرتبط با آن ارائه می‌دهد. سیگنال‌های صوتی دریافت شده از قلب، به دو بخش مجزای صدای اولیه (S1) و صدای ثانویه (S2) تقسیم می‌شوند. صداهای خارج از این دو سیگنال را با نام سیگنال‌های صوتی غیرطبیعی قلب می‌شناهند که در تشخیص بیماری‌های قلبی مناسب هستند. بسیاری از پزشکان، دستگاه فونوکاردیوگراف را به عنوان یکی از ابزارهای کارآمد در شناخت وضعیت‌های غیرطبیعی قلب می‌دانند؛ به این دلیل که نه تنها اطلاعات کلینیکی مهمی را از بیمار در اختیار متخصص قرار می‌دهد، بلکه استفاده از آن آسان، کم‌هزینه و مؤثر می‌باشد [۴]. صداهای قلب، به دلیل فعالیت‌های مکانیکی قلب و در پاسخ به پالس‌های الکتریکی به وجود آمده در گره سینوسی-دهلیزی رخ می‌دهند. پالس‌های الکتریکی در نتیجه کار عضلات قلب، از طریق فعالیت‌های دهلیزها و بطن‌ها انتقال می‌یابد که مطابق با عملکرد قلب، صداهای متفاوتی در مراحل گوناگون رخ می‌دهد و می‌توان منشأ هر یک را به انقباض و انقباض عضلات قلب و باز یا بسته شدن دریچه‌ها مرتبط دانست. بر این اساس، در مجموع ۴ صدای متفاوت از قلب را می‌توان شنید که با نام‌های صدای اولیه، صدای ثانویه، سیستول (Systole)، فاصله صدای بین حدفاصل صدای اولیه تا ثانویه (Dia) و دیاستول (Diastole)، فاصله صدای بین حدفاصل صدای ثانویه تا اولیه بعدی) شناخته شده‌اند. علاوه بر این، فعالیت‌های مکانیکی قلب نیز ممکن است باعث ایجاد صداهای دیگر مانند صدای سوم قلب (S3) و صدای چهارم قلب (S4) شوند که این صداها به صداهای ضعیف قلبی معروف هستند. در شکل ۱ این صداها و حدفاصل آن‌ها نشان داده شده است [۵].



شکل ۱: سیگنال PCG^۲ و صداهای قلب [۵]. صداهای قلب شامل S2، S1 و S3، S4.

در برخی از پژوهش‌های پیشین، بر روی مدل سازی سیگنال‌های قلبی کار شده است و در این میان مدل سازی سیگنال S2 بیش از S1 مد نظر محققان بوده است. در برخی از این مطالعات، هر مؤلفه از S2 توسط یک سیگنال ویژه مدل شده که در آن سیگنال غیرخطی با پهنای باند باریک در نظر گرفته شده است [۶]. در دیگر روش‌ها، جدا سازی و استخراج میزان هم‌پوشانی مؤلفه‌های موجود در صداهای به وجود آمده در دریچه آئورت و دریچه ریوی قلب مورد مطالعه قرار گرفته است [۷]. استفاده از ضریب هم‌ستگی، تکنیک دیگری برای

در سراسر دنیا، بیماری‌های قلبی-عروقی به عنوان یکی از دلایل عمده در افزایش میزان مرگ و میر شناخته می‌شوند. سازمان جهانی بهداشت در سال ۲۰۰۵ با انتشار آماری، برآورد کرده است که سالیانه جمعیتی بالغ بر ۱۷/۵ میلیون نفر در سراسر دنیا بر اثر بیماری‌های قلبی و عروقی جان خود را از دست می‌دهند. این رقم، ۳۰٪ کل مرگ و میرهای دنیا را به خود اختصاص داده است و پیش‌بینی‌های آماری نشان می‌دهند که در سال ۲۰۳۰ در حدود ۲۳/۶ میلیون نفر بر اثر بیماری‌های قلبی، جان خود را از دست خواهند داد [۱].

شنیدن صداهای مربوط به قلب انسان، یکی از تکنیک‌های سنتی در علم طب است که از دیرباز مورد استفاده پزشکان بوده است. منشأ صداهای قلب، به حرکت‌های مکانیکی قلب بازمی‌گردد. این حرکات امواج صوتی، با فرکانس پایین تولید می‌کنند. به طور کلی، این روش در غربال کردن تشخیص و در درمان اولیه می‌تواند مناسب باشد، چرا که روشی کاملاً غیرتهاجمی بوده و تکنیکی است که در آن غربال‌سازی کم‌هزینه، آسان و مطمئن می‌باشد. از طرفی با استفاده از این روش، اطلاعات مفیدی از عملکرد قلب فراهم می‌آید که با استفاده از آن می‌توان بی‌نظمی‌ها و اختلالات کاری قلب را شناسایی نمود.

گوش دادن به صداهای قلب و تحلیل صحیح عملکرد آن از مهارت‌های بسیار سخت محسوب می‌شود که فرد متخصص تنها در صورت داشتن تجربه چندساله قادر خواهد بود با استفاده از آن، صداهای طبیعی و غیرطبیعی را از یکدیگر متمایز سازد. دلیل اصلی این مسئله از آنجا نشأت می‌گیرد که صداهای مربوط به عملکرد قلب و صداهای تنفس در ریه، بازه زمانی کوتاهی داشته و از این رو در یک بازه زمانی کوتاه، چندین صدای متفاوت از ناحیه قفسه سینه به گوش خواهد رسید [۲].

طبقه‌بندی خودکار سیگنال صدای قلب با استفاده از سیستم‌های مبتنی بر محاسبات کامپیوتری و تحلیل‌های نرم‌افزاری می‌تواند تا حد چشمگیری دقت و صحت تشخیص را افزایش داده و در مقابل، مراجعات مکرر بیمار را به مراکز درمانی به حداقل برساند. بنابراین، ارائه روشی جهت تشخیص خودکار این ناهنجاری‌ها با تحلیل صداهای قلبی ضبط شده از طریق گوشی پزشکی می‌تواند بسیار مفید واقع شود. با توجه به اهمیت و لزوم موارد پیش‌گفته، توسعه الگوریتم‌هایی برای طبقه‌بندی صدای قلب در تشخیص بیماری‌های قلبی به عنوان مسئله‌ای چالش‌برانگیز تبدیل شد و با وجود سابقه تحقیقات در این زمینه، این مسئله به عنوان موضوع چالش ۲۰۱۶ فیزیونت مطرح شد [۳]. در طول سال‌های اخیر نیز محققان و پژوهشگران در سرتاسر دنیا روش‌های متفاوتی را در جهت تشخیص هوشمند و دقیق بیماری‌های قلبی انجام داده‌اند [۳].

شنیدن صداهای قلب معمولاً با استفاده از گوشی‌های طبی و یا ضبط از طریق دستگاه فونوکاردیوگراف انجام می‌شود. فونوکاردیوگراف، دستگاهی است که عملیات ضبط صدا در آن بر اساس نوسانات صوتی

در ادامه این مقاله، تحلیل زمان-فرکانس سیگنال‌های غیرایستنا در بخش دوم و پایگاه داده مورد استفاده در این تحقیق در بخش سوم ارائه خواهند شد. الگوریتم پیشنهادی جهت طبقه‌بندی خودکار صدای ضبط شده قلب در بخش چهارم معرفی می‌گردد. نتایج در بخش پنجم ارائه شده و بخش ششم شامل جمع‌بندی و نتیجه‌گیری است.

۲- تحلیل زمان-فرکانس سیگنال‌های غیرایستنا

تحلیل سیگنال‌های غیرایستنا (که محتوای فرکانسی آن‌ها با زمان تغییر می‌کند) در حوزه زمان-فرکانس صورت می‌گیرد که در آن اطلاعات زمانی و فرکانسی سیگنال توأماً در دسترس می‌باشند [۲۷]. برای انتقال سیگنال زمان گسسته N نقطه‌ای $x[n]$ به حوزه زمان-فرکانس می‌توان از تبدیل زمان-فرکانس درجه دوم (QTFD) استفاده کرد که در رابطه (۱) نشان داده شده است [۲۷].

$$\rho_x[n, k] = 2DFT_{n \rightarrow k} \left\{ G[n, m] * (z_x[n+m] z_x^*[n-m]) \right\} \quad (1)$$

در رابطه بالا $z_x[n] = x[n] + jH\{x[n]\}$ سیگنال تحلیلی متناظر با $x[n]$ و $H\{\}$ عملگر تبدیل هیلبرت است. همچنین $G[n, m]$ کرنل زمان-تأخیر QTFD و $*$ کانولوشن گسسته در حوزه زمان می‌باشد. تبدیل زمان-فرکانس $x[n]$ را با $\rho_x[n, k]$ نمایش می‌دهیم که یک ماتریس $N \times M$ است که M تعداد نقاط FFT مورد استفاده برای محاسبه QTFD است [۲۷]. کرنل‌های مختلف امکان تعریف TFD های مختلف را به ما می‌دهند. برای تحلیل سیگنال‌های غیرایستنا چند مؤلفه‌ای، به QTFD هایی نیاز است که ضمن حفظ رزولوشن مناسب، اثر مؤلفه‌های مختلف روی یکدیگر را در حوزه زمان-فرکانس کم کنند. این QTFD ها را توزیع‌های زمان-فرکانس با تداخل کاسته شده می‌نامند. جدول ۱ برخی از RI-TFDs را در حوزه زمان-تأخیر نشان می‌دهد که در این تحقیق مورد استفاده قرار گرفته‌اند. جزییات بیشتر در مورد آن‌ها را می‌توان در [۲۷] یافت.

جدول ۱: RI-TFDs که در این تحقیق مورد استفاده قرار گرفته‌اند و

کرنل آن‌ها در حوزه زمان-تأخیر [۲۷]

Distribution	$G[n, m]$	Parameters
^{۱۳} WVD	$\delta[n]$	N / A
^{۱۴} SWVD	$\delta[n]w[m]$	$w[n]$: Hamming window $N / 4$: Window length
^{۱۵} CWD	$\frac{\sqrt{\pi\sigma}}{2 m } \exp\left(\frac{-\pi^2\sigma n^2}{4m^2}\right)$	$\sigma = 5$
^{۱۶} SPEC	$w[n+m]w[n-m]$	$w[n]$: Hamming window $N / 4$: Window length
^{۱۷} MBD	$\frac{\cosh^{-2\beta} n}{\sum_n \cosh^{-2\beta} n}$	$\beta = 0.01$
^{۱۸} EMBD	$\frac{\cosh^{-2\beta} n}{\sum_n \cosh^{-2\beta} n} \frac{\cosh^{-2\alpha} n}{\sum_n \cosh^{-2\alpha} n}$	$\beta = 0.01, \alpha = 0.9$

جداسازی سیگنال‌های صدای قلب است که بدون استفاده از دستگاه ECG محقق می‌شود و ایراد اصلی آن عدم کارایی در مواقعی است که فواصل زمانی و طیف سیگنال صوتی با مؤلفه‌های بزرگی همراه باشند [۸]. در [۹]، روشی برای جداسازی سیگنال‌ها با استفاده از تجزیه ضرایب موجک و ساخت مؤلفه‌های S_1 ، S_2 ، دوره انقباض و انقباض، با به‌کارگیری قاعده انرژی شانون و یک ماسک تعریف شده در واحد زمان، ارائه شده است. الگوریتم‌های مبتنی بر استخراج ویژگی، از دیگر روش‌های تحلیل این نوع سیگنال‌ها است. نکته‌ای که باید در هنگام استخراج ویژگی به آن توجه کنیم، ماهیت صدای قلب است که سیگنالی غیرایستنا و دارای تغییرات فرکانسی و زمانی شدیدی است [۵]. براین اساس روش‌های مختلفی برای استخراج ویژگی از سیگنال‌های صدای قلب ارائه شده است که مبتنی بر تجزیه ضرایب موجک [۱۰]، ضرایب کپستروم فرکانس مل [۱۳، ۱۱]، تخمین طیف توان [۱۴]، الگوریتم جستجوی انطباقی [۱۵] و پردازش‌های زمانی و فرکانسی [۱۶، ۱۷] می‌باشد. از سوی دیگر، در زمینه طبقه‌بندی سیگنال‌های صوتی قلب و تقسیم‌بندی آن‌ها در دسته‌های طبیعی و غیرطبیعی، روش‌های مختلفی پیشنهاد شده که از شبکه عصبی مصنوعی (ANN) [۱۸، ۱۹]، ماشین بردار پشتیبان (SVM) [۲۰-۲۲]، نزدیک‌ترین همسایگی (kNN) [۲۳، ۲۴] و شبکه عصبی کانولوشن (CNN) [۱۲، ۱۶] استفاده می‌کنند.

در سال‌های گذشته، تلاش‌های قابل ملاحظه‌ای در زمینه تجزیه و تحلیل سیگنال صدای قلب (PCG) به‌منظور تشخیص خودکار بیماری‌های قلبی انجام شده است. باوجود ارزنده بودن این تلاش‌ها، تمرکز عمده آن‌ها بر روی قطعه‌بندی سیگنال صوتی، حذف نویز و آنالیز محتویات طیفی قطعات مختلف سیگنال PCG شامل صدای اول و صدای دوم قلب و در نتیجه، طبقه‌بندی بیماری‌های قلبی با روش‌های خودکار انجام شده است، که وارد مرحله به‌کارگیری کلینیکی نشده است. یکی از دلایل مهم در این زمینه، نداشتن شناخت و دانش کافی از ساختارهای ذاتی و رفتارهای بنیادی سیگنال PCG است. به عبارتی، کارهای صورت گرفته برای آشکار نمودن ساختار اجزای مختلف سیگنال و استخراج ویژگی‌هایی که برخاسته از ماهیت سیگنال باشند بسیار اندک بوده است به همین دلیل این مسئله به‌عنوان موضوع چالش ۲۰۱۶ فیزیوت برگزیده شد که بهترین روش ارائه شده برای طبقه‌بندی صدای قلب در این چالش به میانگین حساسیت و اختصاصیت، ۸۶/۹۲٪ دست یافته است [۱۶].

روش ارائه شده در این مقاله، با توجه به ماهیت غیرایستنا سیگنال صدای قلب، بر تحلیل سیگنال‌های PCG در حوزه زمان-فرکانس و استخراج ویژگی‌های حوزه زمان-فرکانس از این سیگنال‌ها به منظور تشخیص خودکار ناهنجاری‌های قلبی استوار است. کارایی این ویژگی‌ها در طبقه‌بندی سیگنال‌های غیرایستنا به‌طور کلی و سیگنال‌های بیولوژیکی به‌طور خاص، در مطالعات قبلی نشان داده شده است [۲۱، ۲۷-۲۵].

۲-۱- ویژگی‌های زمان-فرکانس

میزان پیچیدگی توزیع زمان-فرکانس را مشخص می‌کند به‌نحوی که هر چه مقدار آنتروپی Renyi کمتر باشد، نشانگر پیچیدگی کمتر توزیع زمان-فرکانس است. کمتر بودن پیچیدگی یک توزیع زمان-فرکانس را می‌توان معادل تفکیک‌پذیری بالاتر و تمرکز انرژی بیشتر دانست. این آنتروپی در رابطه (۶) داده شده است که در آن عدد صحیح α معمولاً بزرگ‌تر از یک (مرتبه آنتروپی را مشخص می‌کند).

$$RE_x(\alpha) = \frac{1}{1-\alpha} \log_2 \sum_{n=1}^N \sum_{k=1}^M \rho_x^\alpha[n, k] \quad (6)$$

دقت کنیم که این آنتروپی محدودیت آنتروپی شانون را ندارد.

۶- همواری TFD سیگنال:

این معیار که نشان می‌دهد توان سیگنال در حوزه زمان-فرکانس چگونه توزیع شده است در رابطه (۷) آمده است.

$$TFF_x = \frac{\left(\prod_{n=1}^N \prod_{k=1}^M \rho_x[n, k] \right)^{\frac{1}{NM}}}{\sum_{n=1}^N \sum_{k=1}^M \rho_x[n, k]} \quad (7)$$

۷- ویژگی‌های مبتنی بر تجزیه مقدار تکین (SVD)^۹: در تجزیه مقدار تکین، ماتریس توزیع زمان-فرکانس $M \times N$ ، ρ_x به حاصل ضرب سه ماتریس U ، S و V تبدیل می‌شود به‌طوری که:

$$\rho_x = USV^H \quad (8)$$

که U و V ماتریس‌های یکه متعامد و به ترتیب با ابعاد $N \times N$ و $M \times M$ می‌باشد و S یک ماتریس شبه قطری $M \times N$ است. عناصر روی قطر ماتریس S مقادیر ویژه ماتریس ρ_x هستند. آمارگان این مقادیر (مانند بیشینه و یا کمینه مقادیر ویژه) را می‌توان به‌عنوان ویژگی‌های سیگنال در حوزه زمان-فرکانس در نظر گرفت.

۸- ویژگی‌های مبتنی بر پردازش تصویر توزیع‌های زمان-فرکانس:

علاوه بر ویژگی‌های ارائه شده در قسمت‌های قبل، با توجه به این که یک توزیع زمان-فرکانس را می‌توان یک تصویر در نظر گرفت، امکان استفاده از تکنیک‌های پردازش تصاویر برای استخراج ویژگی از یک توزیع زمان-فرکانس استفاده کرد. ابتدا قطعه‌هایی از تصویر توزیع زمان-فرکانس که شامل اطلاعات مهم‌تری هستند را شناسایی شده و به یک تصویر دودویی تبدیل می‌شود و نتیجه یک ماتریس دودویی $\rho'_x[n, k]$ خواهد بود. ویژگی‌های زمان-فرکانس می‌توانند برحسب گشتاور مرتبه p ام، ماتریس $\rho'_{pq}[n, k]$ (محاسبه کردند که در رابطه (۹) آمده است.

ویژگی‌های زمان-فرکانس سیگنال $x[n]$ از QTFD آن که در رابطه (۱) ارائه شده استخراج می‌گردند. مهم‌ترین این ویژگی‌ها در این بخش ارائه می‌شوند. جزئیات بیشتر در مورد این ویژگی‌ها در [۲۱، ۲۶] آمده است.

۱- انرژی سیگنال:

انرژی سیگنال در حوزه زمان-فرکانس در فرمول (۲) داده شده است و می‌تواند به‌عنوان یک ویژگی استخراج شود.

$$E_x = \sum_{n=1}^N \sum_{k=1}^M \rho_x[n, k] \quad (2)$$

۲- ویژگی‌های زمان-فرکانس بر اساس فرکانس لحظه‌ای:

فرکانس لحظه‌ای سیگنال $x[n]$ به‌صورت رابطه (۳) تخمین زده می‌شود.

$$f_{i,x}[n] = \frac{f_s}{2M} \frac{\sum_{k=1}^M k \rho_x[n, k]}{\sum_{k=1}^M \rho_x[n, k]} \quad (3)$$

آمارگان‌های فرکانس لحظه‌ای (مثلاً بیشینه، کمینه و انحراف معیار آن‌را) را می‌توان به‌عنوان ویژگی در نظر گرفت.

۳- معیار تمرکز انرژی:

این معیار، تمرکز انرژی سیگنال $x[n]$ را در حوزه زمان-فرکانس به‌صورت کمی اندازه‌گیری می‌کند. این معیار را می‌توان با استفاده از رابطه (۴) یافت.

$$EC_x = \left(\sum_{n=1}^N \sum_{k=1}^M |\rho_x[n, k]|^2 \right)^{\frac{1}{2}} \quad (4)$$

۴- آنتروپی شانون:

تمرکز انرژی و تفکیک‌پذیری توزیع زمان-فرکانس سیگنال $x[n]$ را ارزیابی می‌کند (به‌نحوی که هر چه مقدار آنتروپی شانون بیشتر باشد، تفکیک‌پذیری و تمرکز انرژی توزیع کمتر است و هر چه مقدار آنتروپی کمتر باشد، تفکیک‌پذیری و تمرکز انرژی توزیع زمان-فرکانس بیشتر است) آنتروپی شانون $\rho_x[n, k]$ از رابطه (۵) قابل محاسبه است:

$$SE_x = - \sum_{n=1}^N \sum_{k=1}^M \rho_x[n, k] \log_2(\rho_x[n, k]) \quad (5)$$

مهم‌ترین محدودیت آنتروپی شانون (با توجه به استفاده از لگاریتم در رابطه (۵) این است که تنها برای توزیع‌های مثبت (مانند اسپکتروگرام) قابل تعریف است.

۵- آنتروپی Renyi:

۴- الگوریتم پیشنهادی جهت طبقه‌بندی سیگنال PCG

الگوریتم پیشنهادی، جهت طبقه‌بندی سیگنال‌های PCG شامل دو رهیافت می‌باشد. رهیافت اول شامل طبقه‌بندی سیگنال‌های PCG مبتنی بر ویژگی‌های زمان-فرکانس با استفاده از سه طبقه‌بند مختلف SVM، AdaBoost و MLP می‌باشد. رهیافت دوم، شامل سیگنال‌های صدای ضبط‌شده قلب با استفاده از شبکه عصبی کانولوشن طبقه‌بندی می‌شوند.

در الگوریتم پیشنهادی، مطابق بلوک‌دیگرام ارائه شده در شکل ۲، در رهیافت اول ابتدا سیگنال‌های صدای قلب ضبط شده، پیش پردازش و قطع‌بندی می‌شوند. در مرحله قطع‌بندی با توجه به این که طول صداهاى ضبط شده قلب در پایگاه داده مورد استفاده در این تحقیق متغیر می‌باشند، تعداد قطعه‌های استخراج شده S1، S2، Sys و Dia هر ضبط متفاوت خواهند بود. برای غلبه بر این مشکل، پس از قطع‌بندی سیگنال PCG و استخراج قطعه‌های S1، S2، Sys و Dia قطعات متناظر با استفاده از یک عملگر (مانند میانگین، میانه، کمینه و بیشینه) با هم ترکیب می‌شوند. بدین ترتیب برای یک ضبط در ورودی سیستم، در خروجی بلوک قطع‌بندی و ترکیب قطعه‌ها، ۴ قطعه S1، S2، Sys و Dia را خواهیم داشت. سپس از هریک از قطعات، ۱۸ ویژگی زمان-فرکانس استخراج می‌شود. در خروجی این بخش یک بردار شامل $72 = 4 \times 18$ ویژگی برای هر ضبط استخراج می‌شود. از آنجا که از توزیع‌های زمان-فرکانس برای انتقال سیگنال PCG به حوزه زمان-فرکانس استفاده کرده‌ایم، عملکرد ۶ کرنل مختلف، برای استخراج ویژگی‌های زمان-فرکانس مورد ارزیابی قرار گرفته است. در مرحله طبقه‌بندی نیز برای انتخاب طبقه‌بند مناسب، از سه طبقه‌بند مختلف SVM، AdaBoost و MLP استفاده شده است و در نهایت با استفاده از روش leave-one-out و معیار میانگین حساسیت و اختصاصیت، طبقه‌بندهای مختلف مورد ارزیابی قرار گرفته است.

در رهیافت دوم، ابتدا سیگنال PCG مطابق آنچه در رهیافت اول الگوریتم پیشنهادی گفته شد، پیش‌پردازش می‌شود. سپس سیگنال پیش پردازش شده به ۴ باند فرکانسی (۲۵-۴۵، ۴۵-۸۰، ۸۰-۲۰۰ و ۲۰۰-۴۰۰ هرتز) تقسیم می‌شود و در مرحله بعد مطابق آنچه که در بالا گفته شد، ۴ باند فرکانسی قطع‌بندی و ترکیب می‌شوند. بدین ترتیب برای یک ضبط در ورودی سیستم، در خروجی بلوک قطع‌بندی و ترکیب قطعه‌ها، ۴ قطعه S1، S2، Sys و Dia را خواهیم داشت که ورودی شبکه عصبی کانولوشن خواهد بود. خروجی شبکه عصبی کانولوشن با خروجی حاصل از قسمت اول الگوریتم با استفاده از یک الگوریتم تصمیم‌گیری با هم ترکیب می‌شوند و در نهایت طبیعى یا غیرطبیعی بودن سیگنال ضبط شده مشخص می‌شود. جزئیات بیشتر الگوریتم پیشنهادی در ادامه آمده است. نرم‌افزار استفاده شده برای آنالیز MATLAB می‌باشد و کدهای تهیه شده نیز ارائه شده‌اند.

$$m_{pq} = \sum_n \sum_k n^p k^q \rho'_x[n, k] \quad (9)$$

$$\bar{n} = m_{10}/m_{00}, \bar{k} = m_{01}/m_{00}$$

این تحقیق از ویژگی‌های زمان-فرکانس ارائه شده در این بخش برای طبقه‌بندی صدای قلب استفاده کرده است.

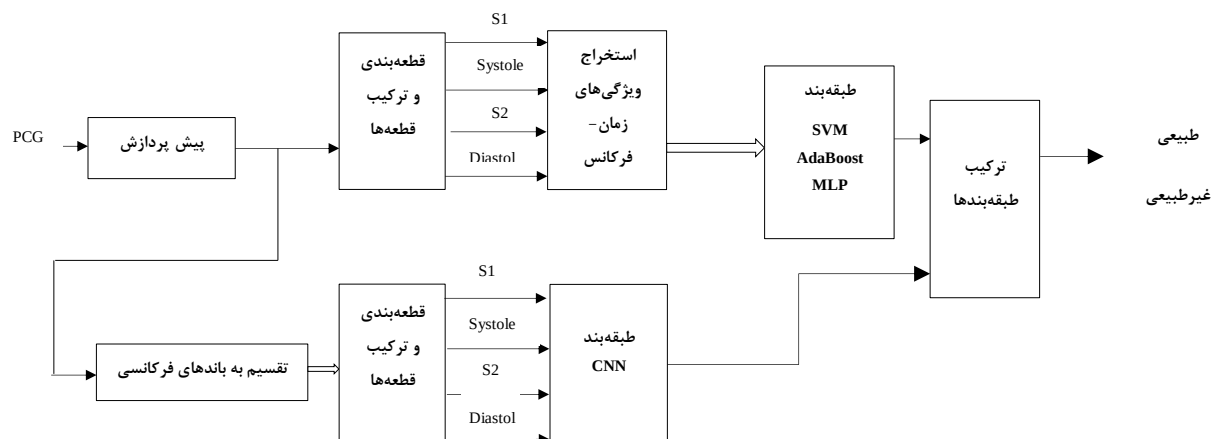
۳- پایگاه داده مورد استفاده در این تحقیق

در این تحقیق از پایگاه داده ارائه شده در چالش ۲۰۱۶، فیزیونت استفاده شده است که یک پایگاه داده جامع است که به‌طور مستقل توسط هفت تیم مختلف تحقیقاتی از هفت کشور در طی یک دوره بیش از یک دهه جمع‌آوری شده است. این پایگاه داده در مجموع شامل تقریباً ۳۰ ساعت صدای ضبط شده که شامل ۲۳۳۵۱۲ صدای قلب از ۱۱۶۸۶۵ ضربان قلب، در ۴۴۳۰ ضبط از ۱۰۷۲ فرد سالم و بیمار با شرایط مختلف (مانند بیماری‌های مربوط به دریچه قلب و انسداد عروق قلب) در سنین مختلف، از مکان‌های مختلف قلب جمع‌آوری شده است. مدت زمان صداهاى ضبط شده بین ۵ ثانیه تا ۱۲۰ ثانیه می‌باشد. داده‌های اضافی ارائه شده عبارت‌اند از موضوع جمعیت‌شناسی (سن و جنس)، اطلاعات ضبط (محل ضبط و مدت زمان صدای ضبط شده) سیگنال‌های ضبط هم‌زمان (مانند ECG)، فرکانس نمونه‌برداری و نوع سنسور استفاده شده است.

تمام صداهاى ضبط شده قلب به دو نوع طبیعى و غیر طبیعى تقسیم شده‌اند. صدای ضبط شده طبیعى مربوط به فرد سالم و صدای ضبط شده غیر طبیعى مربوط به بیمارانی بوده است که به‌طور معمول با نقص دریچه قلب، پرولاپس دریچه میترال، انقباض آئورت، تنگی آئورت و جراحی دریچه‌ای مواجه بوده‌اند. همه ضبط‌ها با فرکانس ۲۰۰۰ هرتز نمونه‌برداری شده‌اند. جزئیات بیشتر در مورد این پایگاه داده در [۲۸، ۲۹] آمده است. همچنین اطلاعات بیشتر در مورد مجموعه‌های آموزش و آزمون این پایگاه داده که در این تحقیق استفاده شده است، در جدول ۲ ارائه شده است.

جدول ۲: جزئیات مجموعه‌های آزمون و آموزش پایگاه داده [۲۸، ۲۹]

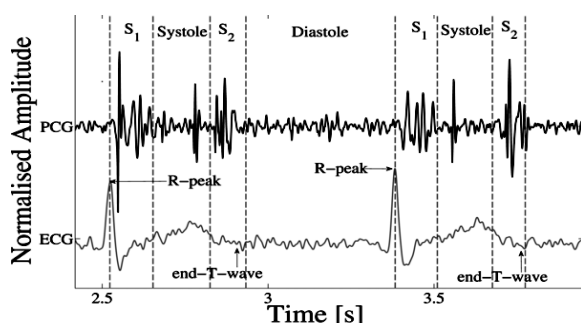
Database	patients	recordings	Proportion of recordings(%) Abnormal	beats
Training-a	121	409	67.5	14559
Training-b	106	490	14.9	3353
Training-c	31	31	64.5	1808
Training-d	38	55	47.3	853
Training-e	356	2141	7.1	59593
Training-f	112	114	27.2	4259
Total	764	3153	18.1	84425
Test-b	45	205	15.6	1269
Test-c	14	14	64.3	853
Test-d	17	24	45.8	260
Test-e	153	883	6.7	26724
Test-g	44	116	18.1	2048
Test-i	35	35	60.0	1286
Total	308	1277	12.0	32440



شکل ۲: بلوک دیاگرام روش پیشنهادی. در این روش از دو رهیافت برای طبقه‌بندی سیگنال‌های PCG استفاده می‌شود.

۴-۱- پیش پردازش

این روش با توجه به همزمانی پیک R کمپلکس QRS با S1 و وقوع S2 بعد از موج T سیگنال ECG، با آشکار سازی R و T از سیگنال می‌توان صداهای S1 و S2 را از سیگنال استخراج کرد. در شکل ۳ سیگنال PCG و ECG که به صورت هم‌زمان ثبت شده است، موقعیت پیک R سیگنال ECG و صدای S1 و موقعیت موج T سیگنال ECG و صدای S2 نشان داده شده است. موقعیت شروع S1، قبل از شروع پیک R در نظر گرفته می‌شود. پایان S1 نیز به طور متناسب با طول سیکل مورد نظر انتخاب می‌شود. موقعیت بیشینه در سیگنال صوتی که در فاصله بعد از موج T قرار دارد به عنوان شروع S2 در نظر گرفته می‌شود. پس از استخراج صدای اول و دوم قلب حد فاصله بین صدای اول و دوم قلب حدود سیستول و حد فاصله بین صدای دوم و اول ضربان قلب بعدی حدود دیاستول را مشخص می‌کند.



شکل ۳: سیگنال PCG ثبت شده هم‌زمان با ECG [۴]

به منظور ترکیب قطعه‌ها، پس از طبقه‌بندی سیگنال PCG و استخراج قطعه‌های S1، S2، Sys و Dia، قطعات متناظر با استفاده از یک عملگر (مانند میانگین، میانه، کمینه و بیشینه) با هم ترکیب می‌شوند. بدین ترتیب برای یک ضبط در ورودی سیستم، در خروجی بلوک طبقه‌بندی و ترکیب قطعه‌ها، ۴ قطعه S1، S2، Sys و Dia را خواهیم داشت که کارایی عملگرهای مختلف در عملکرد پیشنهادی ارزیابی خواهد شد.

در مرحله پیش پردازش با فیلتر کردن مناسب، اغتشاشات و نویزهای داخل و خارج بدن از سیگنال‌های PCG حذف می‌شوند. در تحلیل‌های مربوط به نویز، نشان داده شده است که اغلب، فرکانس سیگنال اصلی با فرکانس نویز تفاوت دارد [۱]. اگر محیطی بی سر و صدا برای ضبط سیگنال‌های صوتی قلب در نظر گرفته شود، آن‌گاه تنها اغتشاشات باقی مانده در داخل بدن با سیگنال اصلی همراه خواهند شد. بیشتر نویزهای بدن، سیگنال‌هایی با دامنه کم و فرکانس بالا هستند. با توجه به مؤلفه‌های فرکانسی سیگنال صدای قلب، از یک فیلتر میان‌گذر باترورث با فرکانس‌های قطع ۲۵ و ۴۰۰ هرتز برای جدا کردن نویز از سیگنال اصلی استفاده شده است. فرکانس‌های قطع پایین و بالای فیلتر براساس این که بیشترین انرژی الگوها برای کل داده مورد استفاده در این باند فرکانسی وجود دارند، انتخاب شده است [۱۶] و به منظور کاهش بار محاسباتی، فرکانس نمونه‌برداری سیگنال فیلتر شده به ۱۰۰۰ هرتز کاهش یافته است.

۴-۲- طبقه‌بندی و ترکیب قطعه‌ها

بخش‌های فعال سیگنال‌های صوتی قلب S1 و S2 می‌باشند، بنابراین یافتن مکان و موقعیت بخش‌های S1 و S2 از اهمیت زیادی برخوردار است. به منظور طبقه‌بندی سیگنال PCG و استخراج صداهای S1، S2، Sys و Dia معمولاً از سیگنال کمکی ECG که به طور هم‌زمان با سیگنال PCG ثبت شده استفاده می‌شود.

در این مرحله با توجه به وجود سیگنال کمکی ECG در پایگاه داده از الگوریتم ارائه شده در [۴] که از داده‌های این چالش برای طبقه‌بندی سیگنال PCG کمک گرفته است، برای طبقه‌بندی استفاده کرده‌ایم. در این الگوریتم از سیگنال کمکی ECG پایگاه داده a برای آموزش الگوریتم استفاده شده است که بعد از آموزش الگوریتم، نیاز به سیگنال کمکی ECG ندارد. جزئیات دقیق این الگوریتم در [۴] آمده است. در

۴-۳- استخراج ویژگی

استخراج ویژگی مهم‌ترین مرحله در پردازش یک سیگنال به منظور طبقه‌بندی صدای قلب است. در این مرحله مشخصاتی از سیگنال‌ها استخراج می‌شوند که نسبت به کلاس‌های مورد نظر حساسیت و اختصاصیت داشته باشند. در این تحقیق با توجه به ماهیت غیرایستادن سیگنال‌های PCG ویژگی‌های حوزه زمان-فرکانس از سیگنال‌های قطعه‌بندی و ترکیب شده استخراج شده است. ابتدا با استفاده از معادله (۱) سیگنال PCG به حوزه زمان-فرکانس انتقال داده می‌شود و سپس (با توجه به روابط بخش ۲-۱-۱ تا ۲-۱-۸) ویژگی ارائه شده در جدول ۳ از هر قطعه سیگنال PCG استخراج شده است. در خروجی این بخش یک بردار شامل $72 \times 18 \times 4$ ویژگی برای هر ضبط استخراج شده است. کرنل‌های مختلف امکان تعریف TFD های مختلف را می‌دهند. به منظور بررسی تأثیر کرنل‌های مختلف، ۶ کرنل ارائه شده در جدول ۱ برای انتخاب بهترین کرنل مورد ارزیابی قرار گرفته است.

جدول ۳: از هر قطعه سیگنال ضبط شده صدای قلب (S1, S2, Sys و Dia) ۱۸ ویژگی زمان-فرکانس استخراج می‌شود که در مجموع ۷۲ ویژگی زمان-فرکانس از هر سیگنال PCG استخراج می‌شود. جزئیات بیشتر این ویژگی‌ها را می‌توان در [۲۷] و بخش ۲-۱-۱ یافت.

Name TF features	TF features
Energy	$TF_1 = E_x = \sum_{n=1}^N \sum_{k=1}^M \rho_x[n, k]$
mean IF	$TF_2 = \frac{1}{N} \sum_{n=1}^N f_{i,x}[n]$
deviation IF	$TF_3 = \max(f_{i,x}[n]) - \min(f_{i,x}[n])$
Energy concentration measure	$TF_4 = EC_x = \left(\sum_{n=1}^N \sum_{k=1}^M \rho_x[n, k] ^2 \right)^{\frac{1}{2}}$
Shannon entropy	$TF_5 = SE_x = - \sum_{n=1}^N \sum_{k=1}^M \rho_x[n, k] \log_2(\rho_x[n, k])$
Renyi entropy	$TF_6 = RE_x(\alpha) = \frac{1}{1-\alpha} \log_2 \sum_{n=1}^N \sum_{k=1}^M \rho_x^\alpha[n, k]$
mean Spectral flatness	$TF_7 = \frac{1}{N} \sum_{n=1}^N TFF_x$
deviation Spectral flatness	$TF_8 = \max(TFF_x) - \min(TFF_x)$
mean SVD	$TF_9 = \max(\rho_x)$
variance SVD	$TF_{10} = (\max(\rho_x) - \min(\rho_x))^2$
Area or Convex Hull	$TF_{11} = m_{00} = \sum_n \sum_k \rho'_x[n, k]$
Perimeter	$TF_{12} = p = (m_{30} + m_{12})^2 + (m_{03} + m_{21})^2$
Compactness	$TF_{13} = p^2 / m_{00}$
Coordinates of the centroid for the segmented	$TF_{14} = m_{10} / m_{00}$ $TF_{15} = m_{01} / m_{00}$
Rectangularity	$TF_{16} = (m_{20} - m_{02})^2 + 4m_{11}^2$
aspect ratio	$TF_{17} = m_{20} - m_{02}$
Form Factor	$TF_{18} = 4\pi / TF_{13}$

۴-۴- طبقه‌بندی

الگوریتم پیشنهادی، همان‌طور که در بخش ۴ مطرح شد، جهت طبقه‌بندی سیگنال‌های PCG از دو رهیافت استفاده می‌کند. رهیافت اول الگوریتم پیشنهادی شامل طبقه‌بندی سیگنال‌های PCG مبتنی بر ویژگی‌های زمان-فرکانس می‌باشد. پس از استخراج ۷۲ ویژگی از قطعه‌های سیگنال PCG، برای بررسی تأثیر انتخاب طبقه‌بند مناسب از سه طبقه‌بند مختلف AdaBoost, SVM و MLP استفاده شده است.

در رهیافت دوم الگوریتم پیشنهادی طبقه‌بندی سیگنال‌های PCG مبتنی بر CNN می‌باشد. در این قسمت پس از قطعه‌بندی سیگنال PCG، قطعات استخراج شده وارد شبکه عصبی کانولوشن شده و در نهایت طبقه‌بندی می‌شوند.

خروجی دو رهیافت الگوریتم پیشنهادی، با استفاده از عملگر یا با هم ترکیب می‌شوند. به این صورت که اگر خروجی یکی از رهیافت‌ها غیر طبیعی باشد، سیگنال PCG ورودی به‌عنوان صدای غیر طبیعی شناخته می‌شود و تنها در صورتی که نتیجه هر دو رهیافت طبیعی باشد، سیگنال PCG ورودی به‌عنوان صدای طبیعی شناخته می‌شود. دلیل اصلی استفاده از عملگر یا در ترکیب خروجی دو رهیافت الگوریتم پیشنهادی آن است که در کاربردهای پزشکی داشتن درصد کم منفی کاذب (تشخیص سالم بودن وضعیت قلبی توسط پزشک برای شخصی با ناهنجاری قلبی) و محدود کردن مقدار مثبت کاذب مطلوب است. بنابراین طبقه‌بندی مثبت که به معنی تشخیص احتمال وجود مشکل قلبی توسط پزشک است، منجر به مراجعه بیمار به متخصص قلب جهت بررسی بیشتر خواهد شد که در تشخیص ناهنجاری‌های قلبی بسیار حائز اهمیت است. در در بخش بعد طبقه‌بندها و معیارهای ارزیابی عملکرد طبقه‌بندها به‌طور خلاصه توصیف شده‌اند و جزئیات مربوط به آن‌ها در مراجع ذکر شده قابل دسترسی هستند.

۴-۴-۱- طبقه‌بندی سیگنال PCG مبتنی بر ویژگی‌های زمان-فرکانس

۱- ماشین بردار پشتیبان (SVM)

روش SVM یکی از الگوریتم‌های مبتنی بر شبکه عصبی برای طبقه‌بندی داده‌ها است. در این روش، هدف، یافتن خط مستقیم (ابرفضا) حائل بین دو کلاس مختلف با استفاده از روابط بهینه‌سازی، برنامه‌ریزی درجه دو (QP-SVM) و نسخه کم‌ترین مربعات (LS-SVM) می‌باشد. ولی برای بهبود دقت، یک حاشیه امن در دو سمت این خط (ابرفضا) در نظر گرفته می‌شود که به بردارهای روی این خطوط بردار پشتیبان می‌گویند. اما این خطوط معمولاً تفکیک‌کننده‌های دقیقی نیستند. بنابراین، ما ابتدا با استفاده از تابع هسته، داده‌ها را ابتدا به فضای دیگری می‌بریم که در این فضا بتوان با خط راست دو کلاس را از هم جدا کرد. در این مقاله، به‌منظور طبقه‌بندی ویژگی‌های زمان-فرکانس استخراج شده، از QP-SVM و

به‌طور معمول پس از هر لایه کانولوشن یک لایه پولینگ قرار می‌گیرد که با بکارگیری این لایه ضمن کاهش محاسبات مورد نیاز در بخش آموزش، باعث کنترل فرایادگیری احتمالی در شبکه نیز می‌گردد. لایه کاملاً متصل، لایه آخر یک شبکه عصبی کانولوشن محسوب می‌شود و مشابه شبکه عصبی معمولی اتصالات کاملی با خروجی لایه قبلی ایجاد می‌کند. این لایه ورودی را دریافت می‌کند و خروجی آن، کلاس ورودی را مشخص می‌کند. ساختار شبکه عصبی کانولوشن مورد استفاده در این مقاله متشکل از دو لایه کانولوشن، دو لایه پولینگ و یک لایه کاملاً متصل است. که ورودی شبکه عصبی کانولوشن با توجه به سیگنال PCG ورودی و قرار گرفتن سیگنال در ۴ باند فرکانسی و استخراج ۴ قطعه سیگنال PCG، سه بعدی در نظر گرفته شده است. لایه‌های کانولوشن اول و دوم به ترتیب دارای ۸ و ۴ فیلتر به طول ۵ و با اندازه گام ۵ می‌باشند. هر دو لایه کانولوشن از دنباله‌گذاری صفر با اندازه ۲ و تابع فعالیت یکسو ساز خطی استفاده می‌کنند. بعد از هر لایه کانولوشن، یک لایه پولینگ با اندازه گام ۲ قرار دارد، خروجی این لایه‌ها، یک شبکه MLP شامل یک لایه ورودی با ۱۶ نورون، یک لایه مخفی با ۲۰ نورون می‌باشد و لایه خروجی با ۱ نورون که لایه کلاس بندی محسوب می‌شود. تابع فعالیت لایه مخفی تابع یکسو ساز خطی و تابع فعالیت لایه خروجی تابع سیگموئید می‌باشد. در خصوص اجتناب از پدیده فرایادگیری، شبکه‌های عصبی کانولوشن به دلیل این که تعداد پارامترهایی که آموزش می‌بینند بسیار زیاد است و همچنین این شبکه‌ها دارای لایه‌های عمیق هستند، لذا در طی زمان آموزش، مستعد فرایادگیری می‌باشند. برای حل نمودن این مشکل راه‌های متفاوتی پیشنهاد شده است. یکی از روش‌های معروف، روش افزایش تعداد داده‌های آموزش و روش دیگر، روش اضافه نمودن لایه dropout مناسب در لایه کاملاً متصل برای رفع این معضل است. در این مقاله برای حل این مشکل از روش اضافه نمودن لایه dropout با نسبت ۲۵٪ بعد از لایه Max-pooling در لایه کانولوشن دوم و با نسبت ۵۰٪ به لایه مخفی شبکه MLP استفاده شده است [۳۲-۳۴].

۴-۳-۴-۳ معیارهای ارزیابی عملکرد

به‌منظور فراهم کردن امکان مقایسه عملکرد روش پیشنهادی در این مقاله با روش‌های موجود، طبقه‌بندی کننده‌های مختلف با استفاده از معیارهای حساسیت، اختصاصیت و میانگین این دو ارزیابی شده است. این معیارها با استفاده از فرمول‌های زیر محاسبه می‌شوند:

$$Sen = \frac{TP}{TP + FN} \quad (13)$$

$$Spe = \frac{TN}{TN + FP} \quad (14)$$

$$MAcc = \frac{Sensitivity + Specificity}{2} \quad (15)$$

در روابط فوق، مثبت صحیح (TP) به تعداد صداهاى طبیعى که (به‌عنوان کلاس مثبت) که به درستی طبقه‌بندی شده‌اند اشاره می‌کند، منفی صحیح (TN) تعداد صداهاى غیر طبیعى (به‌عنوان کلاس

LS-SVM با کرنل چند جمله‌ای درجه دو استفاده شده‌است. این طبقه‌بند کارایی مناسبی در طبقه‌بندی سیگنال‌ها در کاربردهای مختلف نشان داده است [۳۰ و ۳۱].

۲-طبقه‌بند AdaBoost

در این الگوریتم، مجموعه‌های وزن‌دار از تعداد زیادی طبقه‌بندی کننده ضعیف را که هر یک معمولاً برحسب برخی ویژگی‌ها در فضای ویژگی‌های مسئله تعریف می‌گردند، به‌عنوان طبقه‌بندی کننده نهایی انتخاب می‌نماید. به‌طور کلی می‌توان گفت که در واقع Adaboost تعدادی طبقه‌بند می‌سازد و خروجی پیش بینی شده توسط طبقه‌بندها را از طریق رأی حداکثر وزن‌دار با یکدیگر ترکیب می‌نماید. طبقه‌بندها از طریق آموزش یک طبقه‌بند ضعیف، بر روی توزیعی از داده‌های آموزش که در هر مرحله بروز می‌شود به‌دست می‌آیند. توزیع بروز شده بدین صورت عمل می‌کند که نمونه‌هایی که در مراحل قبل به اشتباه طبقه‌بندی شده‌اند با احتمال بیشتری برای آموزش طبقه‌بند بعدی مورد استفاده قرار می‌گیرند. از این‌رو، داده‌های آموزش طبقه‌بندهای متوالی طراحی شده، به سمت نمونه‌هایی که به سختی طبقه‌بندی می‌شوند کشیده می‌شود [۳۲-۳۴].

۳-طبقه‌بند MLP

شبکه‌های عصبی پرسپترون چند لایه، شناخته شده‌ترین شبکه عصبی هستند که دارای بیشترین کاربرد در میان انواع مختلف طبقه‌بندهایی می‌باشند که سیگنال‌های ورودی را براساس ویژگی‌هایی که از آن‌ها استخراج شده‌اند، طبقه‌بندی می‌کنند. این شبکه‌ها از سه مجموعه لایه، یعنی: لایه ورودی، لایه مخفی (میانی) و لایه خروجی تشکیل شده است. خروجی هر لایه پس از تأثیر گذار شدن، تابع محرک ورودی لایه بعدی خواهد شد. در این مقاله، طبقه‌بندی با استفاده از یک شبکه عصبی پرسپترون چندلایه انجام شده است که دارای ۷۲ نورون در لایه ورودی، یک لایه مخفی با ۷۲ نورون و یک لایه خروجی با ۱ نورون می‌باشد.

۴-۲-۴-۲ طبقه‌بندی سیگنال PCG با استفاده از CNN

۱-شبکه عصبی کانولوشن (CNN)

شبکه‌های عصبی کانولوشن مدلی از شبکه‌های عصبی چند لایه هستند که معمولاً شامل لایه‌های کانولوشن، پولینگ و تماماً متصل می‌باشند. لایه کانولوشن به‌عنوان اصلی‌ترین لایه در این دسته از شبکه‌های عصبی محسوب شده که وظیفه آن استخراج ویژگی‌ها می‌باشد. این لایه عملیات کانولوشن را بر روی داده ورودی اعمال می‌کند و خروجی‌هایی به نام نقشه ویژگی از این لایه به دست می‌آید. به‌کارگیری توابع فعالیت در لایه کانولوشن، باعث افزایش خصوصیات غیرخطی در خروجی می‌شوند. توابع فعالیت انواع متفاوتی دارند. چند نمونه از معروف‌ترین توابع فعالیت، تابع سیگموئید، تابع تانژانت هایپربولیک، تابع یکسوساز خطی و تابع خطی نمای است. در یک شبکه عصبی کانولوشن

جدول ۵: نتایج استفاده از انواع کرنل‌های توزیع‌های زمان-فرکانس و**طبقه‌بند LS-SVM**

Kernel	MAcc
MBD	٪۸۲/۱۵
SWVD	٪۸۲/۴۷
EMBD	٪۸۳/۳۶
SPEC	٪۸۳/۴۸
WVD	٪۸۲/۹۷
CWD	٪۸۲/۵۳

جدول ۶: نتایج استفاده از انواع کرنل‌های توزیع‌های زمان-فرکانس و**طبقه‌بند QP-SVM**

Kernel	MAcc
MBD	٪۸۲/۶۳
SWVD	٪۸۲/۴۳
EMBD	٪۸۳/۱۷
SPEC	٪۸۴/۱۶
WVD	٪۸۳/۴۵
CWD	٪۸۲/۷۹

همان‌طور که از نتایج به دست آمده در جدول ۵ و ۶ مشخص است به ازای همه ویژگی‌ها کرنل اسپکتروگرام نسبت به بقیه کرنل‌ها عملکرد بهتری در دو روش ارائه شده در طبقه‌بند SVM دارد. بنابراین کرنل اسپکتروگرام به عنوان بهترین کرنل برای توزیع‌های زمان-فرکانس و روش برنامه‌ریزی درجه دو ماشین بردار پشتیبان به عنوان بهترین نتیجه طبقه‌بند ماشین بردار پشتیبان انتخاب می‌شود. به منظور تعیین انتخاب طبقه‌بند مناسب، طبقه‌بند MLP و AdaBoost استفاده شده است که نتایج حاصل از این طبقه‌بندها در جدول ۷ نشان داده شده است.

جدول ۷: نتایج حاصل از طبقه‌بند MLP و AdaBoost

طبقه‌بند	MAcc
AdaBoost	٪۸۶/۱۲
TNN	٪۸۳/۷۳

همان‌طور که جدول ۷ نشان می‌دهد، طبقه‌بند AdaBoost نسبت به سایر طبقه‌بندها عملکرد بهتری دارد. در نتیجه بهترین پارامترها برای رهیافت اول الگوریتم پیشنهادی شامل عملگر، میانگین، کرنل اسپکتروگرام و طبقه‌بند AdaBoost می‌باشد.

۵-۲- نتایج حاصل از طبقه‌بندی سیگنال PCG مبتنی بر CNN

در رهیافت دوم الگوریتم پیشنهادی پس از پیش‌پردازش‌های لازم، سیگنال پیش‌پردازش شده به ۴ باند فرکانسی (۲۵-۴۵، ۴۵-۸۰، ۸۰-۱۸۰ و ۲۰۰-۴۰۰ هرتز) تقسیم می‌شود و در مرحله بعد طبقه‌بندی

منفی) که به درستی تشخیص داده شده‌اند را نشان می‌دهند. مثبت نادرست و منفی نادرست نیز به ترتیب با FP^{22} و FN^{23} نمایش داده شده‌اند.

۵- نتایج**۵-۱- نتایج حاصل از طبقه‌بندی سیگنال PCG مبتنی بر****ویژگی‌های زمان-فرکانس**

پس از انجام پیش‌پردازش‌های لازم بر روی داده‌ها و قطع‌بندی سیگنال PCG در رهیافت اول الگوریتم پیشنهادی ویژگی‌های حوزه زمان-فرکانس از هر قطعه سیگنال صدای قلب را استخراج شده است. از آنجا که از توزیع‌های زمان-فرکانس مبتنی بر هسته استفاده شده، به منظور بررسی تأثیر کرنل‌های مختلف، ۶ کرنل مختلف برای استخراج ویژگی‌های زمان-فرکانس را به کار گرفته شده است. در مرحله طبقه‌بندی نیز برای بررسی تأثیر انتخاب طبقه‌بند مناسب از سه طبقه‌بند مختلف SVM، MLP و AdaBoost مبتنی بر روش leave-one-out استفاده شده تا همه حالات مختلف پوشش داده شود. در این روش مجموعه داده‌ها به ۶ دسته جداگانه، a تا f تقسیم می‌شوند که ۵ دسته برای آزمون و یک دسته برای آموزش استفاده می‌شوند، که در نهایت میانگین وزنی نتایج بدست آمده محاسبه می‌شود. سپس معیارهای حساسیت، اختصاصیت و میانگین حساسیت و اختصاصیت برای همه طبقه‌بندها محاسبه می‌شود.

به منظور یکسان‌سازی تعداد قطعه‌های استخراج شده از هر بخش سیگنال‌های صدای قلب اپراتورهای حداقل، حداکثر، میانگین و میانه مورد ارزیابی قرار گرفته است. نتایج حاصل از این اپراتورها با استفاده از کرنل اسپکتروگرام و طبقه‌بند SVM در جدول ۴ نشان داده شده است. که نتایج به دست آمده نشان می‌دهد که اپراتور میانگین عملکرد بهتری داشته است.

جدول ۴: نتایج استفاده از کرنل اسپکتروگرام، طبقه‌بند QP-SVM

با کرنل چندجمله‌ای درجه دوم برای اپراتورهای میانگین، میانه،

بیشینه و کمینه

Operator	MAcc
Mean	٪۸۴/۱۶
Median	٪۸۳/۳۹
^{24}Max	٪۸۲/۷۳
^{25}Min	٪۸۳/۲۸

به منظور تعیین بهترین کرنل توزیع‌های زمان-فرکانس از طبقه‌بند SVM استفاده نموده‌ایم. نتایج حاصل از این طبقه‌بند با استفاده از روش برنامه‌ریزی درجه دو و نسخه کم‌ترین مربعات ماشین برای ۷۲ ویژگی استخراج شده از هر ضبط در جدول ۵ و ۶ نشان داده شده است.

در ادامه، عملکرد الگوریتم ارائه شده در این مقاله با عملکرد بهترین روش‌های ارائه شده در چالش ۲۰۱۶ فیزیوت (که بالاترین امتیازها را در چالش کسب کرده بودند)، با اعمال آن‌ها روی پایگاه داده مورد استفاده در این مقاله، مقایسه شد. نتایج این مقایسه در جدول ۱۰ آمده است. نتایج گزارش شده در جدول، نشانگر عملکرد بهتر روش ارائه شده در مقایسه با بهترین روش‌های موجود است.

۶- بحث و نتیجه‌گیری

در این مقاله، روشی جهت طبقه‌بندی صداهای طبیعی و غیرطبیعی ضبط شده قلب با استفاده از آنالیز زمان-فرکانس سیگنال‌های PCG ارائه شد. این موضوع به دلیل اهمیت، در سال ۲۰۱۶ به‌عنوان یک چالش در بزرگترین و معتبرترین گروه بین‌المللی پردازش سیگنال‌های حیاتی مطرح گردید. روش پیشنهادی از ویژگی‌های زمان-فرکانس استخراج شده از سیگنال‌های PCG جهت طبقه‌بندی آن‌ها استفاده می‌کند. نتایج پیاده‌سازی این روش بر روی پایگاه داده چالش ۲۰۱۶ فیزیوت نشان‌دهنده عملکرد بهتر راهکار پیشنهادی در مقایسه با بهترین روش موجود در چالش ۲۰۱۶ فیزیوت و دست‌یابی به حساسیت ۹۳/۲۷٪ و اختصاصیت ۸۱/۹۶٪ در طبقه‌بندی صداهای قلبی است. از مجموع روش‌های ارائه شده در این چالش، ۴۸ الگوریتم پذیرفته شده‌اند که از بین آن‌ها روش ارائه شده در [۱۶] بالاترین امتیاز را کسب کرده است البته توجه به این نکته ضروری است که فاصله امتیاز نهایی بین روش‌های برتر در چالش خیلی کم و در حد چند صدم است در نتیجه هر چند بهبود عملکرد روش پیشنهادی در این مقاله نسبت به بهترین روش موجود کم است، اما با اهمیت است. نتایج بدست آمده از اعمال روش پیشنهاد داده شده به سیگنال‌های واقعی ثبت شده نشان داد که ویژگی‌های زمان-فرکانس عملکرد بسیار خوبی در طبقه‌بندی صداهای ضبط شده قلب نشان می‌دهند و عملکرد بهتری نسبت به متداول‌ترین ویژگی حوزه زمان و فرکانس استفاده شده برای تشخیص این‌گونه کارهای انجام شده پیشین دارد. در ادامه کار، بهبود عملکرد روش پیشنهادی با بررسی امکان استخراج دیگر ویژگی‌های زمان-فرکانس، ترکیب نتایج هر ۴ طبقه‌بند ارائه شده در این مقاله، ترکیب نتایج طبقه‌بندی بر اساس ویژگی‌های استخراج شده و همچنین ترکیب دیگر طبقه‌بندهای کلاسیک با شبکه عصبی کانولوشن مورد توجه خواهد بود.

مراجع

- [1] K. Rezaee and J. Haddadnia, "Design and performance evaluation of intelligent system to segregate and classify the phonocardiograph abnormalities using matched filter and multilayer perceptron-back propagation neural networks, " Pajoohandeh Journal, vol. 18, no. 5, pp. 277-286, 2013.
- [2] F. Javed, P. Venkatachalam, and A. F. MH, "A signal processing module for the analysis of heart sounds and

می‌شوند. سیگنال قطعه‌بندی شده وارد شبکه عصبی کانولوشن می‌شود، خروجی حاصل از این مرحله در جدول ۸ نشان داده شده است.

جدول ۸: نتایج حاصل از طبقه‌بند CNN

طبقه‌بند	MAcc
CNN	۸۵/۳۵٪

با توجه به نتایج حاصل از رهیافت اول الگوریتم پیشنهادی طبقه‌بند AdaBoost با بیشترین درصد MAcc با طبقه‌بند CNN در رهیافت دوم الگوریتم پیشنهادی با استفاده از الگوریتم ارائه شده در قسمت ذیل با هم ترکیب می‌شوند. نتایج حاصل از ترکیب این دو طبقه‌بند در جدول ۹ ارائه شده است.

جدول ۹: نتایج حاصل از ترکیب دو طبقه‌بند CNN&AdaBoost

طبقه‌بند	MAcc
CNN&AdaBoost	۸۷/۶۱٪

همان‌طور که جدول ۹ نشان می‌دهد که ترکیب دو طبقه‌بند CNN و AdaBoost نسبت به سایر روش‌ها عملکرد بهتری دارد. در نتیجه بهترین پارامترها برای الگوریتم پیشنهادی شامل عملگر، میانگین، کرنل اسپکتروگرام و ترکیب دو طبقه‌بند AdaBoost&CNN می‌باشد.

جدول ۱۰: مقایسه نتیجه‌ی روش پیشنهادی و بهترین روش‌های ارائه شده در چالش ۲۰۱۶ فیزیوت

روش‌ها	MAcc	حساسیت	اختصاصیت
روش پیشنهادی	۸۷/۶۱٪	۹۳/۲۷٪	۸۱/۹۶٪
روش مبتنی بر استخراج ویژگی‌های زمان و فرکانس و طبقه‌بندی AdaBoost&CNN [۱۶]	۸۶/۹۲٪	۹۳/۴۸٪	۸۰/۳۶٪
روش مبتنی بر استخراج ویژگی‌های زمان و فرکانس و طبقه‌بندی با استفاده از مجموعه‌ای از شبکه‌های عصبی [۱۲]	۸۵/۴۳٪	۸۵/۹۳٪	۸۴/۹۳٪
روش مبتنی بر MFCCs و Wavelets برای استخراج ویژگی و طبقه‌بندی با استفاده از Regularized Neural Network [۳]	۸۵/۰۹٪	۸۶/۲۷٪	۸۳/۹۱٪
روش مبتنی بر MFCCs، Wavelets و Tensors برای استخراج ویژگی و طبقه‌بند KNN [۳]	۸۴/۶۷٪	۸۵/۴۶٪	۸۳/۸۸٪
روش مبتنی بر استخراج ویژگی‌های زمان و فرکانس و طبقه‌بندی با استفاده از Random Forest و LogitBoost [۱۷]	۸۴/۲۳٪	۸۰/۹۰٪	۸۷/۵۶٪
روش مبتنی بر Probability-distribution [۳]	۸۳/۹۸٪	۹۰/۸۵٪	۷۷/۱۱٪
روش مبتنی بر استفاده از MFCCs و طبقه‌بند CNN [۱۳]	۸۳/۵۷٪	۹۳/۵۰٪	۷۳/۶۴٪

- [15] S. Jabbari and H. Ghassemian, "Modeling of heart systolic murmurs based on multivariate matching pursuit for diagnosis of valvular disorders," *Computers in biology and medicine*, vol. 41, no. 9, pp. 802-811, 2011.
- [16] C. Potes, S. Parvaneh, A. Rahman, and B. Conroy, "Ensemble of feature-based and deep learning-based classifiers for detection of abnormal heart sounds," in *Computing in Cardiology Conference (CinC)*, 2016, 2016, pp. 621-624: IEEE.
- [17] M. N. Homsy *et al.*, "Automatic heart sound recording classification using a nested set of ensemble algorithms," in *Computing in Cardiology Conference (CinC)*, 2016, 2016, pp. 817-820: IEEE.
- [18] D. BARSCHDORFF, S. ESTER, and E. MOST, "Phonocardiogram analysis of congenital and acquired heart diseases using artificial neural networks," in *Comparative Approaches To Medical Reasoning: World Scientific*, 1995, pp. 271-288.
- [19] I. Grzegorzczuk *et al.*, "PCG classification using a neural network approach," in *Computing in Cardiology Conference (CinC)*, 2016, 2016, pp. 1129-1132: IEEE.
- [20] M. A. Goda and P. Hajas, "Morphological determination of pathological PCG signals by time and frequency domain analysis," in *Computing in Cardiology Conference (CinC)*, 2016, 2016, pp. 1133-1136: IEEE.
- [21] B. Boashash, G. Azemi, and N. A. Khan, "Principles of time-frequency feature extraction for change detection in non-stationary signals: Applications to newborn EEG abnormality detection," *Pattern Recognition*, vol. 48, no. 3, pp. 616-627, 2015.
- [22] J. J. G. Ortiz, C. P. Phoo, and J. Wiens, "Heart sound classification based on temporal alignment techniques," in *Computing in Cardiology Conference (CinC)*, 2016, 2016, pp. 589-592: IEEE.
- [23] N. Verbiest, C. Cornelis, and R. Jensen, "Fuzzy rough positive region based nearest neighbour classification," in *Fuzzy Systems (FUZZ-IEEE)*, 2012 IEEE International Conference on, 2012, pp. 1-7: IEEE.
- [24] J. M. Keller, M. R. Gray, and J. A. Givens, "A fuzzy k-nearest neighbor algorithm," *IEEE transactions on systems, man, and cybernetics*, no. 4, pp. 580-585, 1985.
- [25] W.-C. Kao and C.-C. Wei, "Automatic phonocardiograph signal analysis for detecting heart valve disorders," *Expert Systems with Applications*, vol. 38, no. 6, pp. 6458-6468, 2011.
- [26] B. Boashash, G. Azemi, and J. M. O'Toole, "Time-frequency processing of nonstationary signals: Advanced TFD design to aid diagnosis with highlights from medical applications," *IEEE Signal Processing Magazine*, vol. 30, no. 6, pp. 108-119, 2013.
- [27] B. Boashash, *Time-frequency signal analysis and processing: a comprehensive reference*. Academic Press, 2015.
- [28] C. Liu, D. Springer, and G. D. Clifford, "Performance of an open-source heart sound segmentation algorithm on eight independent databases," *Physiological measurement*, vol. 38, no. 8, p. 1730, 2017.
- heart murmurs," in *Journal of Physics: Conference Series*, 2006, vol. 34, no. 1, p. 1098: IOP Publishing.
- [3] G. D. Clifford *et al.*, "Classification of normal/abnormal heart sound recordings: The PhysioNet/Computing in Cardiology Challenge 2016," in *Computing in Cardiology Conference (CinC)*, 2016, 2016, pp. 609-612: IEEE.
- [4] D. B. Springer, L. Tarassenko, and G. D. Clifford, "Logistic regression-hsmm-based heart sound segmentation," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 4, pp. 822-832, 2016.
- [5] M. Nabih-Ali, E.-S. A. El-Dahshan, and A. S. Yahia, "Heart diseases diagnosis using intelligent algorithm based on PCG signal analysis," *Circuits Syst*, vol. 8, no. 7, pp. 184-190, 2017.
- [6] J. Xu, L. Durand, and P. Pibarot, "Nonlinear transient chirp signal modeling of the aortic and pulmonary components of the second heart sound," *IEEE Transactions on Biomedical Engineering*, vol. 47, no. 10, pp. 1328-1335, 2000.
- [7] J. Xu, L.-G. Durand, and P. Pibarot, "Extraction of the aortic and pulmonary components of the second heart sound using a nonlinear transient chirp signal model," *IEEE Transactions on Biomedical Engineering*, vol. 48, no. 3, pp. 277-283, 2001.
- [8] T. Leung, P. White, W. Collis, A. Salmon, and E. Brown, "Time-frequency methods for analysing paediatric heart murmurs," *Appl Sign Process*, vol. 4, no. 3, pp. 154-167, 1997.
- [9] L. Huiying, L. Sakari, and H. Iiro, "A heart sound segmentation algorithm using wavelet decomposition and reconstruction," in *Engineering in Medicine and Biology Society*, 1997. Proceedings of the 19th Annual International Conference of the IEEE, 1997, vol. 4, pp. 1630-1633: IEEE.
- [10] H. Liang and I. Nartimo, "A feature extraction algorithm based on wavelet packet decomposition for heart sound signals," in *Time-Frequency and Time-Scale Analysis*, 1998. Proceedings of the IEEE-SP International Symposium on, 1998, pp. 93-96: IEEE.
- [11] T. Nilanon, J. Yao, J. Hao, S. Purushotham, and Y. Liu, "Normal/abnormal heart sound recordings classification using convolutional neural network," in *Computing in Cardiology Conference (CinC)*, 2016, 2016, pp. 585-588: IEEE.
- [12] M. Zabihi, A. Bahrami Rad, S. Kiranyaz, M. Gabbouj, and A. Katsaggelos, "Heart sound anomaly and quality detection using ensemble of neural networks without segmentation," in *Computing in Cardiology Conference (CinC)*, 2016, 2016, pp. 813-816: IEEE.
- [13] J. Rubin, R. Abreu, A. Ganguli, S. Nelaturi, L. Matei and K. Sircharan, "classifying heart sound recordings using deep convolutional neural networks and mel-frequency cepstral Coefficients," in *Computing in Cardiology Conference (CinC)*, 2016, 2016, pp. 1141-1144: IEEE.
- [14] H. Shino, H. Yoshida, K. Yana, K. Harada, J. Sudoh, and E. Harasewa, "Detection and classification of systolic murmur for phonocardiogram screening," in *Engineering in Medicine and Biology Society*, 1996. Bridging Disciplines for Biomedicine. Proceedings of the 18th Annual International Conference of the IEEE, 1996, vol. 1, pp. 123-124: IEEE.

مهندسی برق دانشگاه تبریز، دوره ۴۷، شماره ۴، صص ۱۳۱۷-۱۳۲۷، زمستان ۱۳۹۶.

- [32] L. G. Valiant, "A theory of the learnable," Communications of the ACM, vol. 27, no. 11, pp. 1134-1142, 1984.
- [33] S. Abney, R. E. Schapire, and Y. Singer, "Boosting applied to tagging and PP attachment," in 1999 Joint SIGDAT Conference on Empirical Methods in Natural Language Processing and Very Large Corpora, 1999.
- [34] F. Smeraldi, M. Defoin-Platel, and M. Saqi, "Handling Missing Features with Boosting Algorithms for Protein-Protein Interaction Prediction," in International Conference on Data Integration in the Life Sciences, 2010, pp. 132-147: Springe.

[29] C. Liu *et al.*, "An open access database for the evaluation of heart sound algorithms," Physiological Measurement, vol. 37, no. 12, p. 2181, 2016.

[۳۰] مرتضی خرم کشکولی، مریم دهقانی، «تشخیص، شناسایی و جداسازی عیب توربین گاز پالایشگاه دوم پارس جنوبی با استفاده از روش های ترکیبی داده کاوی، k-means، تحلیل مؤلفه‌های اصلی (PCA) و ماشین بردار پشتیبان (SVM)»، مجله مهندسی برق دانشگاه تبریز، دوره ۴۷، شماره ۲، صص ۵۰۱-۵۱۵، تابستان ۱۳۹۶.

[۳۱] صفر ایراندوست پاکچین، سعید مشگینی و سجاد نصیرزاده، «باز شناسی چهره با استفاده از آنالیز تفکیک خطی بر پایه موجک‌های هار و گابور و ماشین بردار پشتیبان»، مجله

زیرنویس‌ها

- ¹ Systole
- ² Diastole
- ³ Phonocardiogram
- ⁴ Electrocardiogram
- ⁵ Artificial Neural Network
- ⁶ Support Vector Machine
- ⁷ K-Nearest Neighbors
- ⁸ Convolutional Neural Network
- ⁹ Quadratic Time-Frequency Distributions
- ¹⁰ Fast Fourier Transform
- ¹¹ Time-Frequency Distributions
- ¹² Nonlinear energy operator
- ¹³ Wigner Ville Distribution
- ¹⁴ Smoothed Wigner-Ville Distribution
- ¹⁵ Choi-Williams distribution
- ¹⁶ Spectrogram
- ¹⁷ Modified B-Distribution
- ¹⁸ Extended Modified B-Distribution
- ¹⁹ Singular Value Decomposition
- ²⁰ True Positive
- ²¹ True Negative
- ²² False Positive
- ²³ False Negative fraction
- ²⁴ Maximum
- ²⁵ Minimum