نوا: دادگان موسیقی سنتی ایرانی برای تشخیص دستگاه و سازهای اصیل ایرانی

باقر باباعلی[۱]، استادیار؛ آشنا گرگان محمدی[۲]، دانشجوی کارشناسی؛ اسماء فرجی دیزجی[۱]، دانشجوی کارشناسی

۱- دانشکده ریاضی، آمار و علوم کامپیوتر - پردیس علوم - دانشگاه تهران - تهران - ایران - babaali@ut.ac.ir
۲- دانشکده ریاضی، آمار و علوم کامپیوتر - پردیس علوم - دانشگاه تهران - تهران - ایران - ashena.mohammadi@ut.ac.ir
۳- دانشکده ریاضی، آمار و علوم کامپیوتر - پردیس علوم - دانشگاه تهران - تهران - ایران - asma.farajidzji@ut.ac.ir

چکیده: در دنیا پژوهش‌های بسیاری در حوزه پردازش سیگنال موسیقی به منظور بازیابی محتوای موسیقی بر مبنای محتوا انجام شده است. متأسفانه پژوهش‌های انجام شده در زمینه پردازش رایانه‌ای موسیقی سنتی ایرانی بسیار اندک است که یکی از دلایل اصلی آن را می‌توان نبود دادگان برشمرد. در این پژوهش، یک دادگان برای دو مسئله اساسی حوزه موسیقی سنتی ایرانی، یعنی تشخیص دستگاه و تشخیص ساز معرفی شده است که از جامعیت و تنوع لازم برخوردار است و به نام "دادگان نوا" نام‌گذاری شده است. این دادگان شامل صدای پنج ساز متداول سنتی ایرانی در هفت دستگاه است که توسط چهل هنرمند نواخته شده است. در ادامه برای پرداختن به دو مسئله تشخیص ساز و دستگاه، سامانه‌ای پیشنهاد داده شده است که در بخش استخراج ویژگی آن از سیگنال موسیقی ورودی، ضرایب کپسترال بر مبنای مقیاس مل استخراج می‌شود و دنباله بردارهای ویژگی حاصل، به کمک روش بردار هویت به یک بردار با طول ثابت تبدیل می‌شود. در بخش دسته‌بندی، بردار هویت حاصل از موسیقی ورودی به یک ماشین بردار پشتیبان داده می‌شود. بهترین صحت حاصل از تشخیص ساز و دستگاه بر روی دادگان نوا به ترتیب حدود ۹۸ و ۳۴ درصد است که گویای میزان دشواری مسأله تشخیص دستگاه در مقایسه با تشخیص ساز است.

واژه‌های کلیدی: دادگان نوا، موسیقی سنتی ایرانی، تشخیص دستگاه موسیقی، تشخیص ساز موسیقی، بردار هویت، ماشین بردار پشتیبان.

Nava: A Persian Traditional Music Database for the Dastgah and Instrument Recognition Tasks

B. BabaAli[1], Assistant Professor; A. Gorgan Mohammadi[2], BSc Student; A. Faraji Dizaji[3], BSc Student

1- School of Mathematics, Statistics and Computer Science, College of Science, University of Tehran, Tehran, Iran, Email: bagher.babaali@ut.ac.ir
2- School of Mathematics, Statistics and Computer Science, College of Science, University of Tehran, Tehran, Iran, Email: ashena.mohammadi@ut.ac.ir
3- School of Mathematics, Statistics and Computer Science, College of Science, University of Tehran, Tehran, Iran, Email: asma.farajidzji@ut.ac.ir

Abstract: Extensive research has been conducted in the field of music signal processing which targeted context-based music information retrieval. Unfortunately, research on the computer-based processing of the traditional Persian music is rare, which is due to lack of standard databases. In this paper, a database, named Nava, is introduced for two basic tasks of the traditional Persian music field, Dastgah classification and instrument recognition. In terms of instrument, Dastgah and artist, Nava has enough comprehensiveness and variety. It contains the sound of five common traditional instruments played by 40 artists in seven Dastgahs. In order to address the two mentioned basic tasks, a system is proposed which extracts a sequence of Mel frequency cepstral coefficients (MFCC) feature vectors from input music signal and then converts it to a fixed-length feature vector using i-vector technique. In the classification stage, the extracted i-vector is fed into a support vector machine classifier. The best obtained accuracy on the Nava database for the Dastgah classification and instrument recognition are about 34% and 98% respectively, which indicates the difficulty of the former in comparison with the latter.

Keywords: Nava Database, Traditional Persian Music, Music Dastgah Classification, Music Instrument Recognition, Identity vector, Supporting Vector Machine.

تاریخ ارسال مقاله: ۱۳۹۸/۰۵/۰۷
تاریخ اصلاح مقاله: ۱۳۹۸/۰۶/۲۷
تاریخ پذیرش مقاله: ۱۳۹۸/۰۷/۱۵
نام نویسنده مسئول: باقر باباعلی
نشانی نویسنده مسئول: ایران - تهران - میدان انقلاب - دانشگاه تهران - دانشکده ریاضی، آمار و علوم کامپیوتر.

۱- مقدمه

با افزایش روزافزون حجم داده‌های موسیقی ذخیره‌شده در منابع قابل دسترس و گوناگون، نیاز به سیستم‌های خودکار بازیابی و استخراج اطلاعات از موسیقی اهمیت ویژه‌ای یافته است. یک دسته از روش‌های مورد توجه برای نیل به این هدف، بازیابی اطلاعات موسیقی مبتنی بر محتوا است که در دو دهه اخیر یکی از موضوعات پرطرفدار پژوهشی در حوزه پردازش سیگنال موسیقی و بازیابی اطلاعات بوده است و تکنیک‌ها و ابزارهای مختلفی نیز در این زمینه توسعه یافته است. بطور کلی، سیگنال موسیقی به عنوان یک پدیده پیچیده حاوی حجم زیاد و متنوعی از اطلاعات در خصوص ژانر، احساس، هنرمند، ساز و غیره است. تنوع بالای اطلاعات موجود در سیگنال موسیقی، باعث مطرح شدن حیطه گسترده‌ای از مسائل در "بازیابی اطلاعات موسیقی مبتنی بر محتوا"[۱] جهت مطالعه و پژوهش می‌شود که برخی از این مسائل عبارتند از [۱] : قطعه‌بندی یک قطعه موسیقی به بخش‌های آواز و غیرآواز[۲]، شناسایی خواننده[۲] [۳]، دسته‌بندی ژانر[۴]، جستجو با زمزمه[۳][۵]، تشخیص بار احساسی موسیقی [۶]، تشخیص ساز موسیقی [۷]، حاشیه‌نویسی خودکار موسیقی[۴] [۸] و غیره. از کاربردهای مرسوم بازیابی اطلاعات موسیقی مبتنی بر محتوا می‌توان مجموعه موسیقی شخصی‌سازی شده[۵]، سیستم توصیه‌گر موسیقی، دسته‌بندی موسیقی و حفاظت حق چاپ و غیره را نام برد.

موسیقی از دیدگاه‌های بسیاری تقسیم‌بندی می‌شود. یکی از این دیدگاه‌ها، ناحیه پیدایش است که موسیقی سنتی ایرانی یک دسته مطرح در تقسیم‌بندی از این دیدگاه محسوب می‌شود. بخش بزرگی از آسیای میانه، افغانستان، پاکستان، جمهوری آذربایجان، ارمنستان، ترکیه، و یونان متأثر از این موسیقی است و هرکدام به سهم خود تأثیراتی در شکل‌گیری این موسیقی داشته‌اند [۹]. علی رغم پژوهش‌های زیادی که در حوزه پردازش و بازیابی اطلاعات موسیقیایی بر مبنای محتوا برای موسیقی سایر نواحی، بویژه موسیقی غربی، صورت گرفته‌اند، پژوهش‌های انجام شده در حوزه پردازش و بازیابی موسیقی سنتی ایرانی بسیار ناچیز است. دلایل متعددی برای کمبود پژوهش در این حوزه قابل طرح است که یکی از مهم‌ترین آن، نبود دادگان معتبر و قابل دسترس برای مسائل مختلف مطرح در این حوزه است. بنابراین یک گام مهم در راستای رونق دادن به پژوهش در حوزه پردازش موسیقی سنتی ایرانی، گردآوری دادگان معتبر و قابل دسترس همگان است. در حوزه موسیقی سایر نواحی و بویژه موسیقی غربی، دادگان‌های متعددی برای هر یک از مسایل پژوهشی وجود دارد[۹].

گردآوری یک دادگان همه منظوره به نحوی که برای همه مسایل پژوهشی حوزه قابل استفاده باشد، سخت و یا تقریباً غیر ممکن است که نبود این چنین دادگانی برای موسیقی غربی دلیلی بر این ادعاست. دو تا از مسائل پایه و اساسی در راستای بازیابی اطلاعات موسیقیایی سنتی ایرانی مبتنی بر محتوا، تشخیص دستگاه و تشخیص ساز است. در موسیقی سنتی ایرانی دستگاه یک توالی از پرده‌های مختلف موسیقی است که به شنونده حس و شور خاصی را منتقل می‌کند، مثلاً دستگاه ماهور طرب‌انگیز و دستگاه سه‌گاه حزن‌انگیز هستند. این موسیقی شامل هفت دستگاه شور، نوا، ماهور، همایون، سه‌گاه، چهارگاه و راست پنج‌گاه است.

تشخیص دستگاه توسط انسان از طریق شنیدن یک قطعه موسیقی کار سختی محسوب می‌شود و تنها افراد مجرب این حوزه می‌توانند از عهده این کار برآیند. بنابراین تشخیص خودکار دستگاه به دلیل کاربردهای متعددی که می‌تواند در زمینه بازیابی موسیقی سنتی ایرانی بر اساس محتوا، آموزش موسیقی و آهنگ‌سازی داشته باشد، به موضوعی مهم در زمینه پردازش موسیقی سنتی ایرانی تبدیل شده است؛ به نحوی که اگرچه کارهای پژوهشی انجام شده در حوزه پردازش رایانه‌ای موسیقی سنتی ایرانی محدود است ولی عمده آن در این حوزه متمرکز بوده است. تشخیص ساز موسیقی نیز از مباحث اساسی در حوزه بازیابی اطلاعات موسیقی مبتنی بر محتوا است که در حوزه موسیقی سنتی ایرانی به نحوه جامع و کاملی به آن پرداخته نشده است. موسیقی می‌تواند شامل صدای یک یا چند ساز باشد. در فرایند تشخیص ساز که یک مسئله برچسب‌زنی توالی[۶] محسوب می‌شود، هر قطعه موسیقی به قطعاتی ریز تقسیم شده و به قطعه برچسب یک یا چند ساز زده می‌شود. به این ترتیب، دو مسئله پایه‌ای و اساسی در حوزه پردازش رایانه‌ای موسیقی سنتی ایرانی، تشخیص دستگاه و تشخیص نوع ساز است. تا جایی که نویسندگان این مقاله مطلع هستند، متأسفانه هیچ دادگان دو منظوره‌ای برای این دو مسئله (تشخیص دستگاه و ساز) وجود ندارد و محدود دادگان‌های معرفی شده برای هر یک از این دو مسئله در مقالات، دارای ایراداتی هستند: ۱-از جامعیت لازم برخوردار نیستند (تعداد محدودی ساز و یا نوازنده دارند). ۲- از تنوع و حجم کافی برای هر ساز یا هر دستگاه برخوردار نیستند. ۳- داده‌های موجود برای سازها و دستگاه‌های مختلف، متعادل نیست. در این پژوهش، با لحاظ ایرادات محدود دادگان‌های موجود برای دو مسئله تشخیص دستگاه و تشخیص ساز ایرانی و با الهام از دادگان‌های استاندارد موجود برای موسیقی غربی، یک دادگان دومنظوره برای دو مسئله ذکر شده گردآوری شده است که "دادگان نوا" نام‌گذاری شده است و شرح مفصل آن در بخش‌های بعدی آمده است. بر خلاف همه قلیل دادگان‌های گزارش شده در مقالات برای موسیقی ایرانی، دادگان نوا برای هر دو مسئله به سه بخش آموزش، اعتبارسنجی و ارزیابی تقسیم‌بندی شده است که این امکان ارزیابی دقیق و قابل مقایسه روش‌های ارائه شده توسط محققین این حوزه را فراهم می‌کند. همچنین با بکارگیری روش استخراج ویژگی بردار هویت و دسته‌بند ماشین بردار پشتیبان برای دو مسئله تشخیص دستگاه و تشخیص ساز ایرانی بر مبنای دادگان نوا، نتایج معیار[۷] گزارش شده است که کار سایر محققین این حوزه را جهت مقایسه نتایج تسهیل خواهد کرد.

در ادامه این مقاله، در بخش بعد به مروری اجمالی بر روی محدود پژوهش‌های انجام شده در حوزه تشخیص دستگاه و ساز موسیقی سنتی ایرانی و دادگان‌های آنها می‌پردازیم. بخش سوم به شرح و توصیف مبسوط دادگان نوا اختصاص داده شده است. در بخش چهارم به شرح مختصر اجزای سامانه بکار گرفته شده در این پژوهش برای دو مسئله تشخیص دستگاه و تشخیص ساز ایرانی پرداخته‌ایم. نتایج آزمایش‌ها و بررسی و تحلیل آن‌ها در بخش پنجم آمده است و در نهایت در بخش ششم، جمع‌بندی ارائه شده است.

۲- مرور کارهای دیگران

همان‌طوری که در بخش مقدمه عنوان شد، پژوهش‌های منتشرشده در حوزه پردازش رایانه‌ای موسیقی سنتی ایرانی بسیار ناچیز است که در این بخش به شرح آنها خواهیم پرداخت. در [۱۰] از یک شبکه عصبی با توابع شعاعی پایه[۸] (RBF) برای تشخیص دستگاه ماهور از سایر دستگاه‌ها برای ساز سه‌تار بهره برده است. دادگان استفاده شده شامل ۱۳۵ قطعه موسیقی است که ۶۰ تای آن در دستگاه ماهور و بقیه در پنج دستگاه دیگر بوده است. بعد از آموزش شبکه RBF با ۷۰درصد داده‌های موجود در دادگان، نهایتاً به دقت حدود ۷۳درصد در تشخیص دستگاه ماهور رسیده است. در پژوهشی دیگر [۱۱]، بر این اساس که نت‌های نواخته شده توسط ساز نقش کلیدی در تشخیص دستگاه‌های موسیقی ایفا می‌کند، سعی بر آن داشته که نت‌های قطعه موسیقی را با دقت بالایی استخراج نماید و در ادامه با مشخص کردن فواصل بین این نت‌ها و با توجه به منحصر به فرد بودن الگوهای این فواصل برای دستگاه‌های مختلف، دستگاه قطعه موسیقی مورد نظر را تشخیص دهد. در این پژوهش برای ارزیابی روش پیشنهادی، از ۴۶ قطعه موسیقی (۴۲ قطعه توسط ساز تار و ۴ قطعه توسط سنتور نواخته شده است.) در پنج دستگاه مختلف استفاده شده و به دقت ۹۳درصد رسیده است.

در پژوهش[۱۲]، برای دسته‌بندی ردیف میرزا عبدالله از ویژگی-های مختلفی نظیر ناهمگونی[۹]، ضرایب کپسترال بر مبنای مقیاس مل[۱۰]، فرکانس گام، میانگین و انحراف معیار سنترویید طیفی بهره برده شده است. برای دسته‌بندی روش‌های مختلفی نظیر ماشین بردار پشتیبان[۱۱] (SVM)، شبکه عصبی پرسپترون و Kتا نزدیکترین همسایه[۱۲] (KNN)، آزموده شده است که ماشین بردار پشتیبان به دقت بالاتری دست یافته است. دادگان استفاده شده شامل ۱۲۵۰ قطعه موسیقی از سازهای زهی زخمه‌ای تار و سه‌تار توسط چهار استاد معروف ایرانی است و دربرگیرنده هفت دستگاه و شش آواز است.

اخیراً در پژوهشی [۱۳] دیگر، دادگانی با عنوان دادگان موسیقی سنتی ایرانی مریم به منظور تشخیص دستگاه در موسیقی سنتی ایرانی به صورت مستقل از نوع ساز معرفی شده است. این دادگان شامل ۱۱۳۷ قطعه موسیقی است که ۶۳۱ تای آن صدای نی با صدای برخی سازهای دیگر در پس‌زمینه است و در بقیه قطعات، صدای ویولن به عنوان صدای پیش‌زمینه است. قطعات این دادگان در هفت دستگاه است و تعداد قطعات انتخاب شده در هر دستگاه متفاوت است. دستگاه شور با ۴۴۵ قطعه بیشترین و دستگاه سه‌گاه با ۷۴ قطعه کمترین تعداد قطعه در این دادگان را دارد. محققین این پژوهش با انتخاب دو ساز نی و ویولن، سعی بر ارائه دادگانی برای مسئله تشخیص دستگاه به صورت مستقل از ساز داشته‌اند که انتخاب تنها دو ساز برای این ادعا مقعول به نظر نمی‌رسد. در ادامه، این محققین از ۸۰درصد داده‌های هر دستگاه در دادگان مریم برای آموزش و از ۲۰ درصد باقی‌مانده برای آزمون آن استفاده کرده‌اند. ایشان با این ادعا که دستگاه بر مبنای ۱۶ ثانیه موسیقی قابل تشخیص است، هر قطعه موسیقی به قطعات ۱۶ ثانیه‌ای قطعه‌بندی شده است. بر روی هر قطعه ۱۶ ثانیه‌ای تبدیل فوریه زمان کوتاه[۱۳] اعمال شده و ویژگی‌های حاصل به یک شبکه عصبی ژرف (با معماری: پنج لایه پیچشی + دو لایه GRU + دو لایه اتصال کامل) خورانیده شده است و متوسط امتیاز F1 حدود ۸۶ درصد بر روی هفت دستگاه گزارش شده است.

در [۱۴] برای تشخیص دستگاه، با بهره‌گیری از منطق و نظریه مجموعه‌های فازی و با این فرض که هر نت نواخته شده یک مجموعه فازی است؛ هر قطعه موسیقی را یک مجموعه‌ای از مجموعه‌های فازی در نظر می‌گیرد و بر این اساس، به محاسبه شباهت بین دستگاه قطعه موسیقی ورودی و دستگاه‌های مرجع می‌پردازد. دادگان استفاده شده شامل جمعاً ۲۱۰ قطعه موسیقی سنتی ایرانی با ۸۹ قطعه در دستگاه شور و نوا، ۳۰ قطعه در دستگاه سه‌گاه، ۴۱ قطعه در دستگاه ماهور و راست پنجگاه، ۲۶ قطعه در دستگاه همایون، و ۲۴ قطعه در دستگاه چهارگاه است. قطعات عمدتاً شامل آواز سه استاد آواز و تکنوازی چهار ساز تار، سه تار، سنتور و کمانچه است. در این پژوهش نشان داده شده است که در روش پیشنهادی، یک دقیقه موسیقی از هر قطعه برای تشخیص دستگاه آن قطعه لازم و کافی است.

در [۱۵] برای دسته‌بندی هفت دستگاه موسیقی ایرانی، از شبکه عصبی پرسپترون چند لایه استفاده می‌شود. ورودی‌های شبکه عصبی، بیست قله بلندتر از طیف فرکانس هر قطعه موسیقی است. نتایج نشان می‌دهد شبکه می‌تواند دستگاه قطعات آزمون را با دقت حدود ۶۵درصد برای نی، ۷۲درصد برای ویولن و ۵۶درصد برای آواز تشخیص دهد. در [۱۶] نیز از یک شبکه عصبی پرسپترون با یک لایه مخفی برای دسته‌بندی پنج دستگاه شور، ماهور، همایون، سه‌گاه و چهارگاه استفاده شده است. ورودی شبکه، بردارهای باینری ۲۴ مولفه‌ای است که هر مولفه آن گویای یک نت در یک اکتاو است. شبکه به کمک ۱۲۰ الگوی آموزشی تولید شده، آموزش داده شده است و دقت ۱۰۰درصد بر روی این ۱۲۰ الگوی ممکن گزارش شده است.

همان‌طوری که از مرور انجام شده در بالا قابل برداشت است، پژوهش‌های صورت گرفته در حوزه پردازش موسیقی سنتی بسیار ناچیز است. بخش عمده‌ای از این پژوهش‌ها بر روی مسئله تشخیص دستگاه متمرکز بوده است که دادگان‌های استفاده شده آن‌ها به لحاظ حجم، تنوع ساز و دستگاه از جامعیت کافی برخوردار نیستند. در بخش بعد، به معرفی دادگان نوا می‌پردازیم که برای دو مسئله تشخیص

دستگاه و ساز در موسیقی سنتی ایرانی عرضه شده است که از همه دادگان‌های قبلی جامع‌تر و متنوع‌تر است.

۳- شرح دادگان نوا

مجموعه دادگان نوا، اولین دادگان جامع تکنوازی موسیقی سنتی ایرانی است که در دانشگاه تهران جمع‌آوری شده است. این دادگان به دو منظور تشخیص دستگاه و تشخیص ساز طراحی شده است و در طراحی آن سعی شده که از حجم و تنوع کافی به لحاظ نوع ساز، هنرمند و دستگاه برخوردار باشد. همه قطعات موسیقی موجود در این دادگان به فرمت mp3 با فرکانس نمونه‌برداری ۴۴۱۰۰ هرتز و تفکیک-پذیری بیتی[۱۴] ۱۶بیت است.

بیشتر منابع معاصر، ردیف موسیقی ایرانی را به هفت دستگاه و پنج آواز تقسیم می‌کنند[۱۷]. با این حال در تقسیم‌بندی‌های قدیمی‌تر برخی نظریه‌پردازان حوزه موسیقی دوازده دستگاه برمی‌شمردند، و پنج مورد از آوازهای یادشده را نیز جزو دستگاه‌ها لحاظ می‌کنند. در این پژوهش، گردآوری دادگان بر مبنای پوشش هفت دستگاه شور، نوا، ماهور، همایون، سه‌گاه، چهارگاه و راست پنج‌گاه بنا شد. با وجود آن که تعداد سازهای اصیل موسیقی ایرانی زیاد است ولی اکثرشان منسوخ شده و در محافل موسیقی نواخته نمی‌شوند. در نتیجه، میزان داده موسیقی در دسترس برای این سازها تقریباً نایاب است. با مشورت متخصصین حوزه موسیقی ایرانی و با توجه به میزان داده موجود برای هر ساز، نهایتاً سازهای ایرانی نی، تار، سه‌تار، سنتور و کمانچه به عنوان سازهای ایرانی برای گردآوری این دادگان مورد نظر قرار گرفتند.

مجموعه دادگان گردآوری شده نوا شامل ۱۷۸۶ قطعه تکنوازی از ۴۰ هنرمند ایرانی است. هر قطعه در یکی از هفت دستگاه اصلی موسیقی سنتی و به وسیله یکی از پنج ساز اصیل ایرانی تار، سنتور، سه‌تار، کمانچه و یا نی نواخته شده است. تعداد قطعات هر دستگاه، هر ساز و تعداد قطعات ساز-دستگاه به ترتیب در جداول ۱، ۲ و ۳ آمده است. همانطویکه از جداول ۱ و۲ بر می‌آید، بیشترین تعداد قطعات متعلق به دستگاه شور و ساز سنتور و کمترین آن متعلق به دستگاه راست پنجگاه و ساز کمانچه است.

جدول ۱: تعداد قطعات موجود در دادگان به ازای هر دستگاه

دستگاه	شور	ماهور	چهارگاه	همایون	سه‌گاه	نوا	راست پنجگاه
تعداد قطعات	۲۷۹	۲۷۲	۲۷۱	۲۵۳	۲۵۲	۲۴۲	۲۱۷

جدول ۲: تعداد قطعات موجود در دادگان به ازای هر ساز

ساز	سنتور	نی	سه تار	تار	کمانچه
تعداد قطعات	۴۳۸	۳۹۴	۳۶۴	۳۴۸	۲۴۲

جدول ۳: تعداد قطعات موجود در دادگان به ازای هر ساز-دستگاه

	سنتور	نی	سه تار	تار	کمانچه
شور	۸۸	۴۷	۵۳	۵۷	۳۴
ماهور	۵۰	۴۸	۸۲	۴۲	۵۰
چهارگاه	۷۱	۵۷	۶۷	۴۲	۳۴
همایون	۶۵	۶۹	۴۱	۵۰	۲۸
سه‌گاه	۷۳	۵۰	۳۹	۶۷	۲۳
نوا	۵۳	۵۷	۴۰	۵۱	۴۱
راستپنجگاه	۳۸	۶۶	۴۲	۳۹	۳۲

دادگان نوا حدود ۵۵ ساعت موسیقی دارد که آمار آن به تفکیک هر دستگاه، هر ساز و زوج ساز-دستگاه به ترتیب در جداول ۴، ۵ و ۶ آمده است. همان‌طوری که از جداول ۴ و۵ معلوم است، بیشترین حجم داده متعلق به دستگاه چهارگاه و ساز سنتور و کمترین آن متعلق به دستگاه راست پنجگاه و ساز کمانچه است.

جدول ۴: مدت زمان داده‌های دادگان به ازای هر دستگاه برحسب ساعت

دستگاه	شور	ماهور	چهارگاه	همایون	سه‌گاه	نوا	راست پنجگاه
مدت زمان	۸.۳۳	۸.۴۲	۹.۲۲	۸.۱۲	۸.۰۹	۶.۴	۶.۳۹

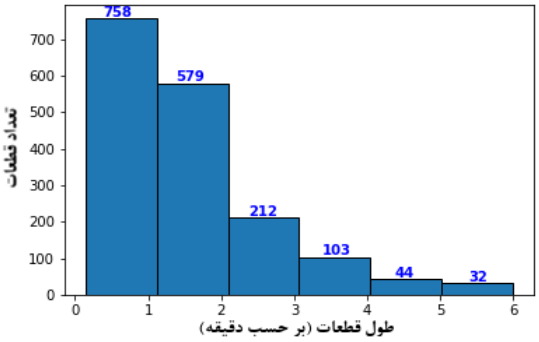
جدول ۵: مدت زمان داده‌های دادگان به ازای هر ساز برحسب ساعت

ساز	سنتور	نی	سه تار	تار	کمانچه
مدت زمان	۱۷.۱	۸.۴۸	۱۲.۰۱	۹.۲۳	۸.۱۶

جدول ۶: مدت زمان داده‌های دادگان به ازای هر ساز-دستگاه برحسب ساعت

	سنتور	نی	سه تار	تار	کمانچه
شور	۳.۳	۱.۲	۱.۶	۱.۲	۱
ماهور	۲.۴	۱	۲.۶	۱.۲	۱.۱
چهارگاه	۲.۸	۱.۲	۲.۶	۱.۲	۱.۴
همایون	۲.۸	۱.۶	۱.۲	۱.۲	۱.۳
سه‌گاه	۲.۵	۱.۱	۱.۳	۲.۱	۱.۱
نوا	۲	۱.۱	۱	۱.۲	۱.۱
راستپنجگاه	۱.۳	۱.۲	۱.۶	۱.۱	۱.۲

در شکل ۱، هیستوگرام طول قطعات موجود در دادگان نوا بر حسب دقیقه نمایش داده شده است. اگر چه ماکزیمم طول قطعه موجود در دادگان حدود ۳۵ دقیقه می‌باشد، ولی طول ۱۷۲۸ قطعه از کل ۱۷۸۶ قطعه (حدود ۹۷درصد) موجود در دادگان کمتر از شش دقیقه است.



شکل ۱: هیستوگرام طول قطعات موجود در دادگان نوا

۳-۱- دادگان نوای کوچک

همان‌طوری که از جداول ۱ تا ۶ برمی‌آید، حجم داده موجود در دادگان نوا برای دستگاه‌ها، سازها و ساز-دستگاه‌های مختلف برابر نیست. بنابراین در صورتی که دادگان نوا برای مسئله تشخیص دستگاه و یا ساز بکار گرفته شود، یک چالش اساسی، داده نامتعادل[۱۵] است. به منظور فراهم کردن شرایطی که فارغ از مسئله داده نامتعادل باشد، بر روی دادگان فعلی یک مجموعه آموزش، معتبرسازی و آزمون تعریف شده است که داده آن متعادل باشد و این بخش از "دادگان نوای کوچک" نام‌گذاری شده است. با توجه به جدول ۶، کمترین داده موجود برای زوج ساز-دستگاه حدود یک ساعت است، بنابراین از هر زوج ساز-دستگاه حدود یک ساعت داده با لحاظ تنوع هنرمند انتخاب شده است. از یک ساعت داده مربوط به هر زوج ساز-دستگاه، حدود ۴۰ دقیقه آن برای آموزش، ۱۰ دقیقه برای مجموعه معتبرسازی و ۱۰ دقیقه برای مجموعه آزمون در نظر گرفته شد. در این بخش‌بندی، سعی بر آن شد که هنرمندان هر سه مجموعه آموزش، معتبرسازی و آزمون متنوع باشد.

لیست مجموعه قطعاتی که به ازای هر زوج ساز-دستگاه برای این سه مجموعه در نظر گرفته شده، مشخص شده و ثابت است. با کنار هم قرار دادن قطعات این سه بخش داده برای هفت دستگاه مختلف، به یک دادگان سه بخشی برای مسئله تشخیص ساز می‌رسیم که به ازای هر ساز حدود ۴۰×۷ دقیقه در مجموعه آموزش، حدود ۱۰×۷ دقیقه در مجموعه معتبرسازی و حدود ۱۰×۷ دقیقه در مجموعه آزمون آن داده موجود است. قابل توجه است که استفاده از دادگان نوای کوچک برای مسئله تشخیص ساز به فرمی که گفته شد، از این مزیت برخوردار است که داده‌های هر ساز از هر هفت دستگاه و تقریباً به اندازه برابر داده دارد. به همین سبک، اگر بر روی دادگان نوای کوچک، قطعات زوج ساز-دستگاه را برای سه بخش آموزش، معتبرسازی و آزمون، به ازای پنج ساز موجود کنار هم قرار دهیم به یک دادگانی برای مسئله تشخیص دستگاه می‌رسیم که به ازای هر دستگاه حدود ۴۰×۵ دقیقه در مجموعه آموزش، حدود ۱۰×۵ دقیقه در مجموعه معتبرسازی و حدود ۱۰×۵ دقیقه در مجموعه آزمون آن داده موجود است. در بخش آزمایش‌ها، بر مبنای این دو تعریف مشخص شده بر روی دادگان نوای کوچک، برای دو مسئله تشخیص ساز و دستگاه به ارائه یک سری نتایج می‌پردازیم که می‌تواند به عنوان معیار برای سایر محققین مورد استفاده قرار گیرد. در پیوست الف، اطلاعات مربوط به نحوه نام‌گذاری فایل‌های دادگان، لیست فایل‌های سه مجموعه آموزش، معتبرسازی و آزمون دادگان نوای کوچک، لیست اسامی هنرمندان و جداول نحوه اختصاص شماره به هر دستگاه، ساز و هنرمند آمده است.

۴- سامانه پیشنهادی

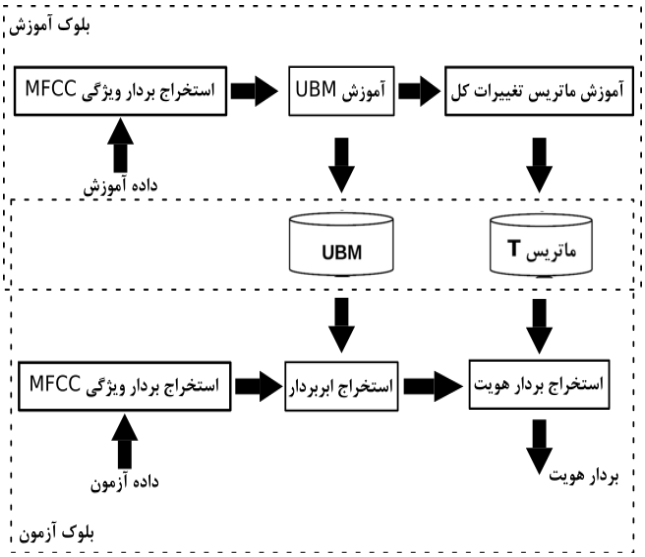
در این بخش، به شرح سامانه پیشنهادی برای دو مسئله تشخیص دستگاه و تشخیص ساز ایرانی بر روی دادگان نوا می‌پردازیم. نمای کلی این سامانه در شکل ۲ آمده است.

شکل ۲: نمای کلی سامانه پیشنهادی تشخیص دستگاه یا نوع ساز

همان‌طوری که در شکل ۲ مشخص است، اولین مرحله، استخراج ویژگی از قطعه موسیقی ورودی است . از هر قطعه موسیقی ورودی به سامانه، دنباله‌ای از بردار ویژگی ضرایب کپسترال بر مبنای مقیاس مل (MFCC) [۱۸] استخراج می‌شود که جزئیات آن در شکل ۳ آمده است؛ به این صورت که سیگنال موسیقی به قاب‌هایی[۱۶] به طور ۲۵ میلی‌ثانیه با هم‌پوشانی ۱۵ میلی‌ثانیه قطعه‌بندی می‌شود. بعد از انجام پیش‌پردازش‌هایی نظیر اعمال فیلتر همینگ؛ از هر قاب تبدیل FFT گرفته شده و در ادامه بانک فیلتر، لگاریتم و تبدیل گسسته کسینوسی اعمال می‌شود و در نهایت به ازای هر قاب، یک بردار ویژگی MFCC بدست می‌آید. در بلاک بعدی که استخراج بردار هویت[۱۷] نامیده شده است، از تمام بردارهای ویژگی MFCC استخراج شده از قطعه موسیقی ورودی، یک بردار ویژگی با طول ثابت به نام بردار هویت استخراج می‌شود. بعد از اعمال تبدیل LDA (به عنوان یک روش با نظارت) به منظور کاهش بعد بردار هویت، بردار کاهش بعد یافته حاصل، توسط

طبقه‌بند ماشین بردار پشتیبان دسته‌بندی می‌شود. با لحاظ این که این سامانه با دادگان نوا آموزش داده شده و ارزیابی می‌شود؛ در صورتی که مسئله تشخیص دستگاه باشد، طبقه‌بند هفت دسته و اگر تشخیص ساز باشد پنج دسته خواهد داشت.

شکل ۳: مراحل استخراج ویژگی MFCC



شکل ۴: نمای کلی فرایند استخراج بردار هویت

بردار هویت که در ابتدا برای کاربرد تشخیص گوینده پیشنهاد شد [۱۹]، در واقع ادامه روش تحلیل عاملی توأم[۱۸] [۲۰] محسوب می‌شود و در سایر کاربردها، نظیر شناسایی زبان [۲۱]، تشخیص لهجه [۲۲]، تشخیص احساس [۲۳]، طبقه‌بندی وقایع صوتی[۱۹] [۲۴]، طبقه‌بندی ژانر موسیقی [۲۵]، و تصدیق برخط امضای دست‌نویس [۲۶]، و غیره هم بکار گرفته شده است. تکنیک بردار هویت را می‌توان به عنوان یک روش برای نگاشت یک توالی از بردارهای ویژگی استخراج شده از یک نمونه داده به یک فضای برداری با ابعاد کم-که فضای تغییرات کل[۲۰] نامیده می‌شود- بر اساس یک تکنیک تحلیل عاملی فرض کرد. به عبارت دیگر، تکنیکی برای ارائه یک نمایش فشرده با طول ثابت از یک توالی از بردارهای ویژگی با طول دلخواه است. در ادامه به شرح مراحل استخراج بردار هویت می‌پردازیم.

نمای کلی فرایند استخراج بردار هویت در شکل ۴ نمایش داده شده است. با استفاده از یک مدل پس‌زمینه جهانی[۲۱] مبتنی بر مدل مخلوط گوسی[۲۲] (GMM-UBM)، دنباله ویژگی‌های MFCC استخراج شده از قطعه موسیقی به یک فضا به ابعاد بالا به منظور ارائه یک بازنمایی بهتر نگاشت می‌شود. در ادامه با استفاده از مدل تغییرات کل (TVM) به بردار هویت می‌رسیم. در GMM-UBM که یک روش برای نگاشت به فضای بعد بالا محسوب می‌شود، از مدل مخلوط گوسی برای تخمین توزیع احتمال بردارهای ویژگی استخراج شده از موسیقی استفاده می‌شود. با فرض این که x یک بردار ویژگی باشد، تابع چگالی احتمال آن بصورت زیر خواهد بود:

$$P(x|\Theta)=\sum_{i=1}^{K} w_i p_i(x|w_i,\mu_i,\Sigma_i) \qquad (۱)$$

که K تعداد گوسی‌ها و w_i ، μ_i Σ_i به ترتیب ماتریس کواریانس، بردار میانگین و وزن گوسی iام و Θ پارامترهای مدل مخلوط گوسی است. رابطه هر گوسی p_i به صورت زیر است:

$$P_i(x|w_i,\mu_i,\Sigma_i) = \frac{1}{(2\pi)^{D/2}\Sigma_i^{1/2}} e^{-1/2(x-\mu_i)^T\Sigma_i^{-1}(x-\mu_i)} \qquad (۲)$$

که D بعد بردار ویژگی x است. پارامترهای مدل مخلوط گوسی Θ که شامل همه پارامترهای K تا گوسی تشکیل دهنده آن است، به روش بیشینه شباهت و به کمک الگوریتم تکراری EM تخمین زده می‌شوند. فرض کنید که $X=x_1,x_2,\ldots x_N$ دنباله بردارهای ویژگی استخراج شده از یک نمونه آموزشی باشد، آنگاه احتمال پسین بردار ویژگی x_k با فرض گوسی iام با استفاده از رابطه زیر تخمین زده می‌شود:

$$Y_i(x_k)=p(i|x_k)=\frac{w_i p_i(x_k)}{p(x_k|\Theta)}. \qquad (۳)$$

در حین آموزش مدل مخلوط گوسی، آمارگان مرتبه صفرام، یکم و دوم بام-ولش گوسی iام با استفاده از روابط زیر محاسبه می‌شوند:

$$B_i^0=\sum_{k=1}^{N} Y_i(x_k) \quad , \qquad (۴)$$

$$B_i^1 = \sum_{k=1}^{N} Y_i(x_k)x_k \quad , \quad B_i^2 = \sum_{k=1}^{N} Y_i(x_k)x_k x_k^T \quad ,$$

و به شکل زیر مجموعه پارامترهای GMM بروز رسانی می‌شود:

$$w_i = \frac{B_i^0}{N} \quad , \quad \mu_i = \frac{B_i^1}{B_i^0} \quad , \quad \Sigma_i = \frac{B_i^2}{B_i^0} - \mu_i\mu_i^T \qquad (۵)$$

در نهایت، بعد از اتمام فرایند آموزش UBM-GMM، از کنار هم قرار دادن میانگین تمام گوسی‌ها به یک بردار در فضای بعد بالا می‌رسیم که آبَر بردار[۲۳] نامیده می‌شود:

$$m = \left[\mu_1^T, \mu_2^T, \ldots, \mu_N^T\right]^T . \qquad (۶)$$

بعد بالای آبَر بردار، آموزش یک مدل را با مشکل مواجه می‌کند، بنابراین بحث کاهش بعد مطرح می‌شود. برای انجام این کار از روش ماتریس تغییرات کل (TVM) بهره گرفته می‌شود. فرض کنید که M آبَر بردار مربوط به یک قطعه موسیقی باشد. آنگاه با اعمال تحلیل عاملی روی فضای آبَر بردار، این آبَر بردار M به صورت خطی روی فضایی با ابعاد کم نگاشت می‌شود:

$$M = m + Tw \qquad (۷)$$

که m میانگین بردار M است و از مدل UBM (کنار هم قرار دادن میانگین تمام گوسی‌ها) بدست می‌آید و $\mathbf{T}$ یک ماتریس مستطیلی شکل مرتبه پایین و w بردار هویت یا i-vector نامیده می‌شود. در این مدل‌سازی فرض شده است که آبَر بردار M دارای توزیع گوسی با میانگین M و ماتریس کواریانس $\mathbf{T}.\mathbf{T}^T$ است. از آنجا که آموزش ماتریس $\mathbf{T}$ به صورت بدون نظارت انجام می‌شود، حاوی اطلاعات تمایزی کافی در مورد دسته‌ها نیست؛ بنابراین در ادامه با بکارگیری تحلیل تمایزی خطی (LDA) که یک روش کاهش ابعاد ویژگی بانظارت محسوب می‌شود، ابعاد بردار هویت مجدداً کاهش می‌یابد.

۵- آزمایش‌ها و نتایج

در این بخش، به ارائه و بررسی نتایج حاصل از آزمایش‌های انجام شده برای دو مسئله تشخیص دستگاه و تشخیص ساز بر روی دادگان نوای کوچک که یک زیرمجموعه متعادل از دادگان نوا است، می‌پردازیم.

در کلیه آزمایش‌های انجام شده، برای استخراج ویژگی MFCC، طول قاب‌ها ۲۵میلی‌ثانیه و هم‌پوشانی بین آنها ۱۵میلی‌ثانیه، تعداد ۲۳ فیلتر بر مبنای مقیاس مل بوده است و ۱۳ ضریب کپستروم به همراه مشتقات اول و دوم آن استخراج شده است. در بخش استخراج بردار هویت، تعداد گوسی‌های مدل پس‌زمینه جهانی ۲۵۶ و طول بردار هویت استخراج شده هم ۲۵۶ درنظر گرفته شده است. همچنین، در خور توجه است که برای انجام آزمایش‌ها، ابتدا تمام قطعات موسیقی موجود در سه مجموعه آموزش، اعتبارسنجی و آزمون به قطعاتی با طول مشخص و ثابت قطعه‌بندی شده و بردار هویت از این قطعات با طول ثابت استخراج شده است.

نتایج آمده در ادامه این مقاله، بر مبنای دو معیار صحت[۲۴] و امتیاز F1 گزارش شده است که صحت بر مبنای رابطه:

$$\text{Accuracy} = \frac{Number\ of\ True\ Classified\ samples}{Number\ of\ samples} \qquad (۸)$$

و امتیاز F1 برای هر کلاس مطابق رابطه محاسبه می‌شود:

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall} \qquad (۹)$$

با توجه این که در این مقاله، تشخیص ساز یک مسئله پنج کلاسه و تشخیص دستگاه یک مسئله هفت کلاسه است؛ میانگین امتیاز F1 برای تمام کلاس‌های مسئله مورد نظر گزارش شده است.

اولین آزمایش به منظور تشخیص ساز به ازای قطعه‌بندی هر قطعه موسیقی به قطعاتی با طول‌های مختلف انجام شد که نتایج آن در جدول ۷ آمده است. همان‌طوری که از جدول بر می‌آید، با کاهش طول قطعات موسیقی، کار تشخیص نوع ساز دشوارتر شده و صحت کاهش می‌یابد که منطقی هم به نظر می‌رسد. اگر طول قطعات ۵ ثانیه باشد، بر مبنای روش پیشنهادی، صحت تشخیص بر روی مجموعه اعتبارسنجی و آزمون به ترتیب حدود ۸۸ و ۸۴ درصد خواهد بود. از آنجایی که دستیابی به صحت تشخیص بالا به ازای قطعات کوچک موسیقی مطلوب است، بنابراین یکی از چالش‌های مطرح بر روی دادگان نوا، بهبود صحت تشخیص ساز در مواقعی است که طول قطعات موسیقی کوتاه است. بهترین صحت تشخیص ساز به ازای قطعات ۶۰ ثانیه‌ای حاصل شده است که ماتریس ابهام[۲۵] آن بر روی مجموعه آزمون در جدول ۸ آمده است. بزرگترین خطای قابل مشاهده در جدول ابهام، تشخیص اشتباه ساز سه‌تار بجای تار است که ۹ بار اتفاق افتاده است و این اشتباه چندان هم غیرمنطقی نیست.

جدول ۷: نتایج تشخیص ساز بر روی دادگان نوای کوچک

طول قطعه (ثانیه)	مجموعه اعتبارسنجی		مجموعه آزمون	
	صحت (٪)	امتیاز F1 (٪)	صحت (٪)	امتیاز F1 (٪)
۵	۸۸.۱۹	۸۸	۸۴.۷۵	۸۴
۱۰	۹۵.۹۰	۹۶	۹۴.۰۳	۹۴
۱۵	۹۶.۴۲	۹۶	۹۳.۴۱	۹۳
۳۰	۹۷.۹۳	۹۸	۹۶.۸۴	۹۷
۴۵	۹۷.۳۴	۹۷	۹۵.۳۵	۹۵
۶۰	۹۷.۹۷	۹۸	۹۷.۱۷	۹۷
۷۵	۹۷.۴۰	۹۷	۹۵.۵۴	۹۵
۹۰	۹۵.۳۷	۹۵	۹۴.۶۰	۹۴

نتایج تشخیص دستگاه بر روی دادگان نوای کوچک در جدول ۹ آمده است. از مقایسه مقادیر جداول ۷ و ۹ برمی‌آید که صحت تشخیص دستگاه به مراتب از صحت تشخیص ساز پایین‌تر است. مسئله تشخیص دستگاه ذاتاً مسئله دشواری محسوب می‌شود و نتایج بدست آمده چندان دور از انتظار نبوده است. با افزایش طول قطعات، صحت بهبود پیدا کرده که منطقی به نظر می‌رسد زیرا برای خبره‌های این

حوزه هم تشخیص دستگاه با شنیدن قطعات موسیقی بلندتر، راحت‌تر است. به این ترتیب، چالش اصلی بر روی دادگان نوای کوچک، مسئله تشخیص دستگاه است. بهترین صحت تشخیص دستگاه بر روی مجموعه اعتبارسنجی به ازای قطعات ۴۵ ثانیه‌ای بدست آمده است که به ازای این طول قطعه صحت بر روی مجموعه آزمون حدود ۲۲درصد است.

جدول ۸: ماتریس ابهام نتایج حاصل از تشخیص ساز بر روی مجموعه آزمون به ازای طول قطعه ۶۰ ثانیه‌ای

	تار	کمانچه	سنتور	سه‌تار	نی
تار	۶۷	۰	۰	۹	۰
کمانچه	۰	۸۲	۰	۰	۰
سنتور	۰	۰	۷۳	۰	۰
سه‌تار	۰	۰	۲	۷۶	۰
نی	۰	۰	۰	۰	۸۱

جدول ۹: نتایج تشخیص دستگاه بر روی دادگان نوای کوچک

طول قطعه (ثانیه)	مجموعه اعتبارسنجی		مجموعه آزمون	
	صحت (٪)	امتیاز F1 (٪)	صحت (٪)	امتیاز F1 (٪)
۱۵	۲۶.۸۰	۲۶	۱۹.۴۶	۱۹
۳۰	۲۸.۰۹	۲۷	۲۰.۱۳	۱۹
۴۵	۳۴.۸۹	۳۴	۲۲.۱۳	۲۱
۶۰	۳۴.۳۴	۳۲	۲۲.۵۶	۲۱
۷۵	۳۰.۲۵	۲۹	۲۳.۷۳	۲۳
۹۰	۳۰.۸۶	۲۹	۲۱.۲۶	۱۹

۶- جمع‌بندی

در این مقاله بعد از بررسی قلیل کارهای پژوهشی انجام شده در حوزه موسیقی اصیل ایرانی به معرفی دادگان نوا، به عنوان اولین دادگان موسیقی ایرانی برای دو مسئله اساسی در حوزه موسیقی، یعنی تشخیص دستگاه و تشخیص ساز پرداختیم. در طراحی و گردآوری این دادگان سعی بر آن بوده که از تنوع خوبی به لحاظ ساز، دستگاه و نوازنده برخوردار باشد. به منظور فراهم آوردن شرایط مقایسه برای محققینی که در ادامه این دادگان را در پژوهش‌های خود بکار خواهند گرفت، برای دو مسئله مورد نظر (تشخیص ساز و دستگاه) بر مبنای یک سامانه پیشنهادی نتایجی ارائه شد. بر مبنای نتایج بدست آمده کاملاً مشهود است که کار تشخیص دستگاه به نسبت تشخیص ساز به مراتب دشوارتر است و نیازمند پژوهش بیشتری است.

گردآوری دادگان نوا به عنوان گامی ابتدایی و در عین حال پایه‌ای در حوزه پردازش موسیقی ایرانی است. دادگان به صورت تک‌نوازی است که با توجه به واقعیت دنیای موسیقی که عمده موسیقی‌ها به صورت چندنوازی است، گردآوری یک دادگان چندنوازی امری الزامی در ادامه این راه به نظر می‌رسد. مشابه موسیقی غربی، مسائل پژوهشی متعددی در حوزه موسیقی ایرانی قابل طرح است، نظیر تشخیص احساس موسیقی که پرداختن به آنها نیازمند گردآوری دادگان است.

سپاس‌گزاری

نویسندگان مقاله برخود لازم می‌دانند که از راهنمایی‌های آقای دکتر هومان اسعدی، عضو هیئت علمی گروه موسیقی دانشکده هنرهای زیبا دانشگاه تهران و آقای حبیب‌پور فارغ التحصیل مقطع کارشناسی دانشکده هنرهای نمایشی و موسیقی پردیس هنرهای زیبای دانشگاه تهران و مدیر فنی سایت بیپ‌تونز بخاطر همکاری و حمایت بی‌دریغ‌شان و شرکت حامی سایت بیپ‌تونز به خاطر به اشتراک گذاشتن منابع مورد نیاز برای تهیه این دادگان، صمیمانه و خاضعانه تشکر کنند.

پیوست الف

نام‌گذاری فایل‌های دادگان به این صورت پس از حذف ابرداده‌ها از دادگان، به هر فایل، نامی نسبت دادیم که فرمت کلی آن به این صورت است: X_Y_ZZ_WW که X بیانگر کد ساز، Y بیانگر کد دستگاه، ZZ بیانگر کد هنرمند قطعه و WW اندیس قطعه در بین تمام قطعات آن هنرمند با ساز و دستگاه مشخص است. جدول‌های مرتبط با کدگذاری ساز، دستگاه و هنرمندان، در زیر آمده است.

نام ساز	کد
تار	۰
کمانچه	۱
سنتور	۲
سه‌تار	۳
نی	۴

نام دستگاه	کد
شور	۰
سه‌گاه	۱
ماهور	۲
همایون	۳
راست پنجگاه	۴
نوا	۵
چهارگاه	۶

نام هنرمند	کد	نام هنرمند	کد
سیامک جهانگیری	۰	سامان ضرابی	۲۰
ابوالحسن صبا	۱	علیرضا جواهری	۲۱
داریوش ثقفی	۲	داریوش طلایی	۲۲
سلیمان روح افزا	۳	مینا نی زاری	۲۳
حسن خان	۴	رضا ورزنده	۲۴
بابک بوبان	۵	عبدالنقی افشارنیا	۲۵
مرتضی گودرزی	۶	محمدرضا شجریان	۲۶

جلیل شهناز	۷	پیمان جهانمانی	۲۷
علی اصغر بهاری	۸	سامر حبیبی	۲۸
انوش جهانشاهی	۹	لطف الله مجد	۲۹
پرویز مشکاتیان	۱۰	حسن کسایی	۳۰
احمد عبادی	۱۱	نیما فریدونی	۳۱
حسن کیانی نژاد	۱۲	علی اکبر شکارچی	۳۲
میرزا علی اکبر شاهی	۱۳	داریوش صفوت	۳۳
حسین صبا	۱۴	پیام جهانمانی	۳۴
فرامرز پایور	۱۵	بهاره فیاضی	۳۵
محمد موسوی	۱۶	علی کاظمی	۳۶
حسین علیزاده	۱۷	ارشد طهماسبی	۳۷
نورعلی برومند	۱۸	عطا جنگوک	۳۸
ارژنگ کامکار	۱۹	رسول اکبری	۳۹

مراجع

[1] Y. V. Murthy, and S.G. Koolagudi, "Content-based music information retrieval (cb-mir) and its applications toward the music industry: A review," *ACM Computing Surveys*, vol. 51, no. 3, 2018.

[2] A. Bonjyotsna and M. Bhuyan, "Signal processing for segmentation of vocal and non-vocal regions in songs: A review," *International Conference on Signal Processing Image Processing and Pattern Recognition, Los Alamitos*, CA, pp. 87–91, 2013.

[3] T. Zhang, "Automatic singer identification," *International Conference on Multimedia and Expo.*, vol. 1, pp. 1-33, 2003.

[4] N. Scaringella, G. Zoia, and D. Mlynek, "Automatic genre classification of music content: A survey," *IEEE Signal Process. Mag.* vol. 23, no. 2, pp. 133–141, 2006.

[5] A. Kotsifakos, P. Papapetrou, J. Hollmen, D. Gunopulos, and V. Athitsos, "A survey of query-by-humming similarity methods, " *In Proceedings of the 5th International Conference on Pervasive Technologies Related to Assistive Environments. ACM*, New York, NY, 5. 2012

[6] B. Y. Chua, *Automatic Extraction of Perceptual Features and Categorization of Music Emotional Expressions from Polyphonic Music Audio Signals*, Ph.D. Thesis, Monash University, 2008.

[7] A. Eronen, *Automatic Musical Instrument Recognition*, Master's Thesis, Department of Information Technology, Tampere University of Technology, 2001

[8] Z. Fu, G. Lu, K. M. Ting, and D. Zhang, , "A survey of audio-based music classification and annotation," *IEEE Trans. Multimed.* Vol. 13, no. 2, pp. 303-319, 2011

[9] L. Miller, *Music and Song in Persia (RLE Iran B): The Art of Avaz.* Routledge, 2012.

[۱۰] سار محمودان، ایوب بنوشی، «دسته‌بندی خودکار گام ماهور موسیقی ایرانی توسط یک شبکه عصبی مصنوعی»، *دومین کنفرانس بین المللی آکوستیک و ارتعاشات*، دانشگاه صنعتی شریف، دیماه ۱۳۹۱

[۱۱] صابر عبداله زادگان، شهرام جعفری، مرتضی دیرند، «تشخیص خودکار دستگاه و گام موسیقی سنتی ایرانی مبتنی بر تکنوازی سازهای تار و سنتور به وسیله استخراج نت هوشمند»، *بیستمین کنفرانس ملی سالانه انجمن کامپیوتر ایران*، دانشگاه فردوسی مشهد، اسفند ۱۳۹۳

[12] M. A. Layegh, S. Haghipour, & Y. N. Sarem, "Classification of the Radif of Mirza Abdollah a canonic repertoire of Persian music using SVM method," *Gazi University Journal of Science Part A: Engineering and Innovation*, vol. 1, no. 4, pp. 57-66, 2013.

[13] S. R. Azar, A. Ahmadi, S. Malekzadeh, and M. Samami, "Instrument-Independent Dastgah Recognition of Iranian Classical Music Using AzarNet," *arXiv preprint arXiv:1812.07017*, 2018.

[14] S. Abdoli, "Iranian Traditional Music Dastgah Classification," *In ISMIR*, pp. 275-280, 2011.

[15] B. Beigzadeh, M. Belali Koochesfahani, "Classification of Iranian traditional musical modes (DASTGÄH) with artificial neural network," *Journal of Theoretical and Applied Vibration and Acoustics*, vol. 2, no. 2, pp.107-118, 2016.

[16] H. Hajimolahoseini, R. Amirfattahi, and M. Zekri, "Real-time classification of Persian musical Dastgahs using artificial neural network," *In The 16th CSI International Symposium on Artificial Intelligence and Signal Processing*, pp. 157-160, 2012.

[17] H. Farhat, *The dastgah concept in Persian music*, *Cambridge University Press*, 2004.

[18] B. Logan, "Mel Frequency Cepstral Coefficients for Music Modeling," *In ISMIR*, vol. 270, pp. 1-11, 2000

[19] N. Dehak, P. Kenny, R. Dehak, and et al. "Front-end factor analysis for speaker verification," *IEEE Trans. Audio Speech Lang. Process.*, vol. 19, no. 4, pp. 788–798, 2011.

[20] P. Kenny, G. Boulianne, P. Ouellet, and et al. "Joint factor analysis versus eigenchannels in speaker recognition," *IEEE Trans. Audio Speech Lang. Process.*, vol 15, no. 4, pp. 1435–1447, 2007.

[21] N. Dehak, P. A. Torres-Carrasquillo, D. A. Reynolds, and et al. "Language recognition via i-vectors and dimensionality reduction," *InterSpeech*, pp. 857–860, 2011.

[22] M. H. Bahari, R. Saeidi, D. Van Leeuwen, and et al. "Accent recognition using i-vector, Gaussian mean supervector and Gaussian posterior probability supervector for spontaneous telephone speech," *IEEE Int. Conf. Acoustics, Speech and Signal Processing (ICASSP)*, pp. 7344–7348, 2013.

[23] R. Xia, Y. Liu, "Using i-vector space model for emotion recognition," *InterSpeech*, 2012.

[24] H. Eghbal-zadeh, B. Lehner, M. Dorfer, and et al. "CP-JKU submissions for DCASE-2016: a hybrid approach using binaural i-vectors and deep convolutional neural networks," 2016.

[25] X. Wei, and L. Wenju, "Multilingual I-Vector based Statistical Modeling for Music Genre Classification," *InterSpeech*, 2017.

[26] H. Zeinali, B. BabaAli, and H. Hadian,"Online signature verification using i-vector representation. *IET Biometrics*," vol. 7, no. 5, pp. 405-414, 2017.

زیرنویس‌ها

۱ Context-Based Music Information Retrieval

۲ Singer Identification

۳ Query By Humming

[۴] Music Annotation
[۵] Personalized Music Collection
[۶] Sequence Labeling
[۷] Benchmark
[۸] Radial Basis Functions
[۹] Inharmonicity
[۱۰] Mel Frequency Cepstral Coefficients
[۱۱] Support Vector Machine
[۱۲] K-Nearest Neighbor
[۱۳] Short Time Fast Fourier Transform
[۱۴] Bit Resolution
[۱۵] Imbalance Data
[۱۶] Frames
[۱۷] Identity vector (i-vector)
[۱۸] Joint Factor Analysis
[۱۹] Audio Scene Classification
[۲۰] Total Variability Space
[۲۱] Universal Background Model
[۲۲] Gaussian Mixture Model
[۲۳] Super-vector
[۲۴] Accuracy
[۲۵] Confusion Matrix