نخستین کنفرانس بین المللی پردازش خط و زبان فارسی
۱۵ و ۱۶ شهریور ۱۳۹۱
دانشگاه سمنان - دانشکده مهندسی برق و کامپیوتر

بررسی ویژگی‌های وابسته به فرکانس پایه لهجه‌های مختلف زبان فارسی

علی قلی‌پور[۱]، محمد حسین صداقی[۲] و موسی شمسی[۳]
[۱]دانشگاه صنعتی سهند تبریز، دانشکده مهندسی برق، a_gholipour@sut.ac.ir
[۲]دانشگاه صنعتی سهند تبریز، دانشکده مهندسی برق، sedaaghi@sut.ac.ir
[۳]دانشگاه صنعتی سهند تبریز، دانشکده مهندسی برق، shamsi@sut.ac.ir

چکیده

این مقاله، به بررسی ویژگی‌های وابسته به زیر و بمی لهجه‌های مختلف زبان فارسی می‌پردازد. زیر و بمی، یکی از زیرمجموعه‌های خصوصیات نوایی گفتار است که با استفاده از کانتور فرکانس پایه مدل‌سازی می‌شود. در این تحقیق، پس از استخراج کانتور فرکانس پایه و هموارسازی آن، برخی از ویژگی‌های آماری منحنی کانتور و مشتق آن، محاسبه می‌شود. سپس الگوریتم انتخاب متوالی پیشرو برای انتخاب ویژگی‌های برتر اجرا می‌شود. نتایج اجرای این الگوریتم نشان می‌دهد که ویژگی‌های مربوط به مشتق کانتور، قدرت تمایز بیشتری برای طبقه‌بندی لهجه‌ها دارند.

در راستای این تحقیق، مجموعه دادگانی از لهجه‌های مختلف زبان فارسی شامل لهجه‌های تهرانی، آذری، کردی، اصفهانی و مازندرانی گردآوری شده است. همچنین برای درک توانایی انسان در تشخیص لهجه‌ها آزمایشی طراحی شده است. نرخ طبقه‌بندی در این پایگاه داده، با استفاده از ویژگی‌های برتر استخراج شده و با بکارگیری طبقه‌بندهای KNN و PNN، ۷۰,۴% حاصل می‌شود که در مقایسه با متوسط قدرت تشخیص انسان (۷۶,۷%) قابل قبول به نظر می‌رسد.

کلمات کلیدی

طبقه‌بندی لهجه‌های زبان فارسی، ویژگی‌های نوایی، فرکانس پایه، بازشناسی گفتار

۱- مقدمه

گفتار، علاوه بر داشتن اطلاعات نوشتاری، مشخصات دیگری نظیر جنسیت، احساسات، سن و لهجه افراد را به همراه دارد. این خصوصیات سبب افزودن اطلاعات اضافی نسبت به اطلاعات نوشتاری و در نتیجه باعث کاهش کارایی سیستم‌های بازشناسی گفتار می‌شوند. رتبه‌بندی عوامل تغییرات گفتار با استفاده از ابزارهای آماری نظیر PCA، نشان می‌دهد که تغییرات گفتار به سبب تغییر در لهجه افراد، پس از عامل جنسیت، دومین عامل کاهش کارایی سیستم‌های بازشناسی است [1]. از این رو درک نحوه تاثیر لهجه‌های مختلف بر پارامترهای گفتار، به منظور بهبود نتایج بازشناسی حیاتی است.

در زبان‌شناسی، به نحوه تلفظ کلمات در یک گروه خاص از افراد یک مکان یا یک ملیت، لهجه می‌گویند. اختلاف بین لهجه‌ها، به طور کلی از دو دیدگاه قابل بررسی است [2]. رویکرد تاریخی، که به بررسی ریشه لهجه‌ها، قوانین تلفظی و نحوه شکل‌گیری این قوانین در طول تاریخ می‌پردازد؛ و رویکرد ساختاری، که اختلاف ساختاری لهجه‌ها را مدل‌سازی می‌کند. بر اساس این مطالعات، دلایل اختلاف بین لهجه‌ها را می‌توان در پنج گروه تقسیم‌بندی کرد:

- اختلاف در تعداد یا هویت مجموعه واج‌ها.
- اختلاف در تلفظ لغات یکسان.
- اختلاف در توزیع آوایی؛ یک واج خاص با توجه به واج‌های اطراف خود ممکن است آواهای متفاوتی داشته باشد.
- اختلاف در ویژگی‌های نوایی[۱] گفتار، شامل اختلاف در الگوی استرس، ضربآهنگ[۲] و زیر و بمی[۳] گفتار.
- اختلاف در تحقق آوایی گفتار، به دلیل اختلاف در پیکربندی اندام‌های تولید اصوات در انسان‌ها.

در این مقاله، با مدل‌سازی پارامتری زیر و بمی گفتار، به بررسی ویژگی‌های این مدل برای لهجه‌های مختلف زبان فارسی می‌پردازیم.

مطالعه در زمینه لهجه‌ی گفتار، قدمت چندانی ندارد و غیر از زبان انگلیسی، در زبان‌های دیگر به ندرت مرجعی یافت می‌شود. اکثر مطالعات انجام گرفته در این زمینه، از ویژگی‌های آکوستیک گفتار، نظیر ضرایب MFCC[۴]، انرژی و فرکانس‌های فرمنت[۵] استفاده کرده‌اند و با به کارگیری ابزارهایی مانند HMM، SVM و شبکه‌های عصبی به طبقه‌بندی لهجه‌های مختلف یک زبان پرداخته‌اند [3,4,5].

در تحقیقی دیگر در زبان انگلیسی، [2] با استفاده از اطلاعات آماری نظیر میانگین طول صامت‌ها و مصوت‌ها و همچنین نرخ گفتار (بر حسب واج بر ثانیه) به طبقه‌بندی برخی از لهجه‌های زبان انگلیسی پرداخته است.

ثابت شده است، گویندگان هنگام صحبت به زبان دوم، همچنان از ویژگی‌های نوایی زبان مادری خود استفاده می‌کنند [6,7]. از این رو ویژگی‌های نوایی گفتار نقش مهمی در تشخیص زبان مادری یا لهجه گویندگان دارند. مرجع [8] از جمله تحقیقاتی است که از ویژگی‌هایی نظیر ضرب‌آهنگ و زیروبمی گفتار برای طبقه‌بندی لهجه فرانسه زبانان با زبان‌های مادری مختلف استفاده کرده است.

با توجه به لهجه‌های مختلفی که در زبان فارسی وجود دارد، طراحی سیستمی که بتواند منطبق با انواع لهجه‌های این زبان باشد، ضروری به نظر می‌رسد. لازمه این کار آشنایی با خصوصیات هر لهجه است.

این مقاله توسعه‌ای بر تحقیقات گذشته‌ی ما [۹] در زمینه طبقه‌بندی لهجه‌های مختلف زبان فارسی است. این لهجه‌ها شامل لهجه‌ی گویندگان تهرانی، آذری، کردی، اصفهانی و مازندرانی می‌باشند که به زبان فارسی صحبت می‌کنند و در مجموع بیشترین جامعه آماری را در کشور دارند. در این مقاله، ویژگی‌های وابسته به کانتور فرکانس پایه[۶]، مربوط به این لهجه‌ها بررسی می‌شود.

کانتور فرکانس پایه، مدل پارامتری زیر و بمی گفتار است. زیر و بمی، یکی از زیرمجموعه‌های ویژگی‌های نوایی گفتار می‌باشد. در این تحقیق، پس از استخراج کانتور فرکانس پایه و هموارسازی آن، برخی از ویژگی‌های آماری منحنی کانتور هموارشده و مشتق آن نظیر میانگین، کمینه، بیشنیه، واریانس و ...، استخراج می‌شوند. سپس با استفاده از الگوریتم انتخاب متوالی پیشرو[۷] و به کارگیری طبقه‌بندهای KNN و PNN، ویژگی‌های موثر انتخاب شده و برای طبقه‌بندی از این ویژگی‌ها استفاده می‌شود.

ادامه مقاله به صورت زیر سازماندهی شده است: در بخش ۲، مجموعه دادگان گردآوری شده معرفی می‌شود. همچنین در آزمایشی، توانایی تشخیص لهجه‌های مختلف، توسط انسان سنجیده می‌شود. الگوریتم استخراج ویژگی در بخش ۳ توصیف می‌شود. در بخش ۴، نحوه انتخاب ویژگی‌های برتر با استفاده از الگوریتم انتخاب متوالی پیشرو توضیح داده می‌شود. بخش ۵، مراحل ارزیابی و نتایج را بیان می‌کند و در انتها نتیجه‌گیری مقاله در بخش ۶ مطرح می‌گردد.

۲- مجموعه دادگان

۲-۱- معرفی

در راستای این تحقیق، مجموعه‌ای از گفتار با لهجه‌های مختلف زبان فارسی، در دانشگاه صنعتی سهند گردآوری شده است. این پایگاه داده تحت عنوان Sahand Accented Speech (SAS) معرفی می‌شود. این مجموعه با همکاری ۴۰ نفر و در ۵ لهجه مختلف زبان فارسی شامل تهرانی، آذری، کردی، اصفهانی و مازندرانی گردآوری شده است (در هر لهجه ۸ نفر، نصف مرد و نصف زن). هر گوینده ۲۴ جمله معین را بیان کرده و در کل این مجموعه شامل ۹۶۰ نمونه می‌باشد.

۲-۲- محتوای جملات و شرایط گویندگان

در این مجموعه، ۲۴ جمله مشخص توسط افراد ادا شده است. محتوای این جملات به صورتی است که تمام واج‌های زبان فارسی را پوشش می‌دهد. از جمله کوتاهترین این عبارات، می‌توان به "سلام علیکم" و "خداحافظ" اشاره کرد. همچنین طولانی‌ترین جمله در این مجموعه، عبارت زیر است:

"ترکیب‌بندی کامل تاریخ خلیج فارس را در کتاب آقای دکتر یغمایی می‌توان یافت. وی به تاریخچه خلیج فارس و تغییر این نام پرداخته است."

این مجموعه با همکاری دانشجویان جمع‌آوری شده است. تا حد امکان، سعی شده است از دانشجویان ورودی جدید استفاده شود. گویندگان به تعداد برابر مرد و زن انتخاب شده‌اند و دارای شرایط سنی بین ۱۸-۲۸ سال می‌باشند. داده‌ها در یک اتاق آکوستیک ضبط و با فرکانس ۸ کیلوهرتز نمونه‌برداری شده است.

۲-۳- ارزیابی

در این مرحله توانایی انسان در تشخیص لهجه‌ها، مورد ارزیابی قرار گرفته است. هنوز به طور کامل مشخص نیست، انسان چگونه می‌تواند زبان مادری یک گوینده را تشخیص دهد [8].

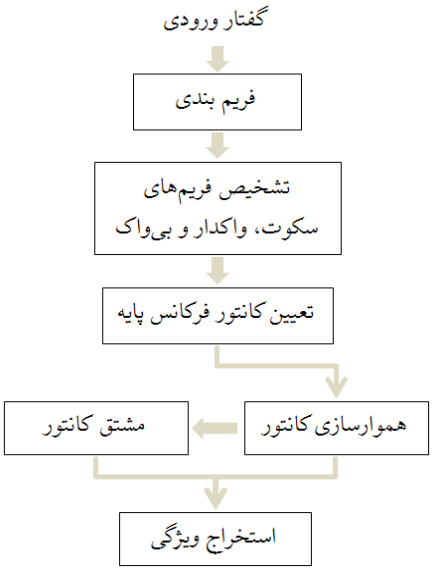
برای درک توانایی انسان در تشخیص لهجه‌ها، آزمایشی طراحی شده است. در این آزمایش از ۱۸ شنونده استفاده می‌شود. برای انجام آزمایش، همه نمونه‌های پایگاه داده، یعنی ۹۶۰ نمونه گفتار، به صورت تصادفی برای هر شنونده پخش شده و از آنها خواسته شده است تا لهجه گفتار پخش شده را تشخیص دهند. پس از پخش هر نمونه، آن نمونه از مجموعه حذف می‌شود تا هر شنونده تنها یک بار شانس شنیدن هر نمونه را داشته باشد. جدول (۱) ماتریس سردرگمی نتایج این آزمایش را نشان می‌دهد.

جدول (۱): ماتریس سردرگمی نتایج ارزیابی انسان(%)

	آذری	مازنی	کردی	تهرانی	اصفهانی
آذری	**۸۸,۲**	۲,۹	۲,۶	۴,۵	۱,۶
مازنی	۲,۳	**۷۵,۶**	۱۵	۲,۹	۳,۹
کردی	۸,۶	۱۹,۲	**۶۸,۳**	۲,۶	۱,۱
تهرانی	۸,۴	۴,۳	۲,۹	**۷۹,۸**	۴,۳
اصفهانی	۶,۱	۶,۸	۳,۶	۱۱,۵	**۷۱,۷**
مقدار متوسط	**۷۶,۷**				

۳- استخراج ویژگی

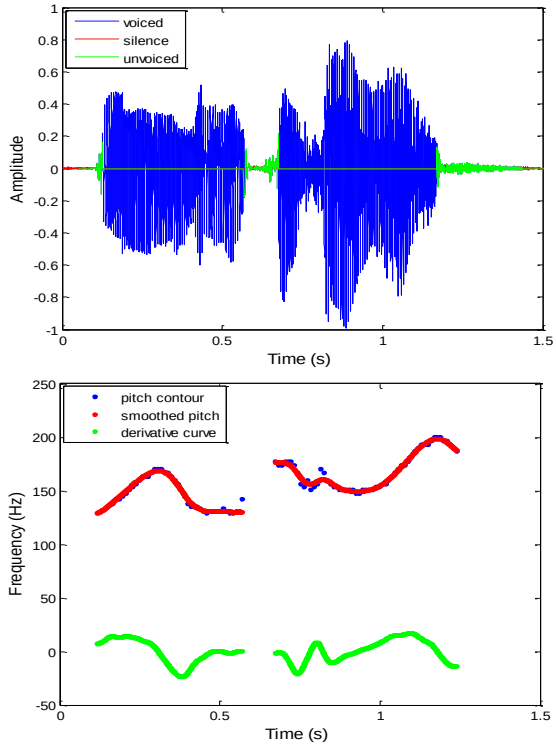
برخی از مشخصات هر لهجه با استفاده از زیر و بمی گفتار مشخص می‌شود. کانتور فرکانس پایه مدل پارامتری زیر و بمی گفتار است. در این مقاله، ما بر استخراج ویژگی از کانتور فرکانس پایه متمرکز شده‌ایم. مراحل کار به صورت بلوک دیاگرام در شکل (۱) نمایش داده شده است که در ادامه تشریح خواهند شد.



شکل (۱): فرایند استخراج ویژگی

از آنجائیکه فرکانس پایه تنها برای بخش‌های واکدار[۸] گفتار تعریف می‌شود، بنابراین ابتدا باید این نواحی مشخص شوند. برای این منظور پس از فریم‌بندی گفتار، به طول ۲۵ و همپوشانی ۱۵ میلی‌ثانیه، با استخراج نرخ عبور از صفر و انرژی هر فریم، فریم ها را به سه دسته سکوت، واکدار و بی‌واک تقسیم‌بندی می‌کنیم. فریمی به عنوان فریم واکدار در نظر گرفته می‌شود که نرخ عبور از صفر و انرژی سه فریم متوالی از یک حد آستانه بیشتر باشند. سپس با استفاده از تابع خودهمبستگی، مقادیر فرکانس پایه تعیین می‌شود. در این مدل برای نواحی سکوت و بی‌واک، مقدار صفر برای کانتور در نظر گرفته می‌شود.

کانتور به دست آمده را با استفاده از روش Cubic Spline هموار کرده و مشتق منحنی هموار شده محاسبه می‌شود. در شکل (۲) نمونه‌ای از کانتور استخراج شده، کانتور هموار شده و مشتق آن، برای یک نمونه گفتار، نشان داده شده است. منحنی مشتق، به منظور نمایش بهتر بزرگنمایی شده است.



شکل (۲):کانتور فرکانس پایه و مشتق آن

ویژگی‌های آماری که از این دو منحنی استخراج می‌شود عبارتند از:

ویژگی‌های استخراج شده از کانتور هموار شده، شامل میانگین، کمینه، بیشینه، واریانس، میانه، دامنه میان چارکی و محدوده (تفاضل بیشینه و کمینه) منحنی، نرخ صعود اولیه و نرخ نزول انتهایی منحنی و درصد وجود فرکانس پایه.

ویژگی‌های استخراج شده از منحنی مشتق، شامل میانگین، کمینه، بیشینه، واریانس، میانه و دامنه میان چارکی.

ویژگی‌های استخراج شده از علامت منحنی مشتق، شامل میانگین، کمینه، بیشینه، واریانس، میانه و دامنه میان چارکی طول قسمت‌های مثبت و منفی. تعداد تغییر علامت منحنی و نسبت طول نواحی مثبت و طول نواحی منفی به مجموع نواحی مثبت و منفی.

در پایان این مراحل، هر نمونه گفتار با یک بردار ویژگی با ابعاد ۳۱، در فضای ویژگی نمایش داده می‌شود.

۴- انتخاب موثرترین ویژگی‌ها

انتخاب ویژگی‌های برتر، تکنیکی است برای انتخاب زیرمجموعه‌ای از ویژگی‌هایی که قدرت تمایز بیشتری دارند. بدین منظور در این مقاله از الگوریتم انتخاب متوالی پیشرو استفاده شده است. در این الگوریتم، ابتدا مجموعه‌ای تهی در نظر گرفته می‌شود و در هر مرحله، ویژگی‌ای که به همراه این مجموعه از امتیاز بالاتری برخوردار باشد، به مجموعه افزوده می‌شود [10]. در این مقاله، نرخ طبقه‌بندی به عنوان معیاری برای امتیازدهی به ویژگی در نظر گرفته شده است. طبقه‌بندهای مورد استفاده در این مقاله KNN و PNN هستند که در ادامه مختصرا توضیح داده می‌شوند.

K-nearest neighbor روشی برای طبقه‌بندی است که در آن نمونه‌های تست، بر اساس برچسب نزدیکترین نمونه‌های آموزشی در فضای ویژگی، طبقه‌بندی می‌شوند.

شبکه عصبی احتمالاتی (PNN)، از نوع شبکه‌های عصبی RBF می‌باشد و بر پایه تابع چگالی احتمال نمایی و قانون تصمیم‌گیری Bayes عمل می‌کند. این شبکه، یک شبکه‌ی پیش-رونده سه لایه بوده و از یادگیری با نظارت استفاده می‌کند.

۵- پیاده سازی و نتایج

در این مرحله الگوریتم تشریح شده را بر پایگاه داده معرفی شده اجرا می‌کنیم. این مجموعه از گفتار ۴۰ نفر در ۵ لهجه مختلف (در هر لهجه ۸ نفر) تشکیل شده است.

در تمامی آزمایشات از الگوریتم leave-one-out برای انتخاب داده‌های آموزش و تست استفاده شده است. بدین ترتیب از هر کلاس (لهجه)، یک نمونه به صورت تصادفی به عنوان داده تست در نظر گرفته می‌شود و سایر نمونه‌ها، برای آموزش شبکه استفاده می‌شوند. با توجه به تصادفی بودن انتخاب داده‌ها، برای بیان نتایج مورد اعتماد، عمل طبقه‌بندی را چند بار اجرا کرده و میانگین نتایج اعلام می‌شود.

در اولین آزمایش، پس از مرحله استخراج ویژگی، عمل طبقه‌بندی با استفاده از تمام ویژگی‌های معرفی شده و با بکارگیری طبقه‌بندهای KNN و PNN انجام گرفته است. نتایج این آزمایش در جدول (۲) نشان داده شده است.

جدول (۲): نتایج طبقه‌بندی با استفاده از تمام ویژگی‌ها (%)

	PNN	KNN
نرخ طبقه‌بندی	۴۵	۴۷,۴

در ادامه به منظور کاهش فضای ویژگی و تعیین ویژگی‌های موثرتر، الگوریتم انتخاب متوالی پیشرو اجرا می‌شود. با توجه به جدول (۲)، استفاده از KNN در طبقه‌بندی نتایج بهتری به همراه دارد. بنابراین برای اجرای این الگوریتم، نرخ طبقه‌بندی با استفاده از KNN، به عنوان معیار امتیازدهی در نظر گرفته شده است. با اجرای الگوریتم، هشت ویژگی به عنوان ویژگی‌های برتر تعیین می‌شوند. انتخاب این هشت ویژگی، به این دلیل است که با اضافه کردن ویژگی‌های جدید به مجموعه، افزایش محسوسی در نرخ‌طبقه‌بندی حاصل نمی‌شود. این هشت ویژگی تعیین شده بترتیب عبارتند از:

نرخ نزول انتهایی کانتور، میانگین طول ناحیه مثبت مشتق، میانه مشتق، میانه، واریانس و دامنه میان چارکی طول ناحیه مثبت مشتق، کمینه طول ناحیه منفی مشتق، میانگین مشتق.

با بررسی این ویژگی‌ها مشخص می‌شود که تفاوت در لهجه‌ها، بیشتر در مشتق کانتور فرکانس پایه ظاهر می‌شود.

در آزمایش بعدی تنها از هشت ویژگی برتر معرفی شده، برای طبقه‌بندی استفاده می‌شود. جدول (۳) نتایج این آزمایش را نشان می‌دهد.

جدول (۳): نتایج طبقه‌بندی با استفاده هشت ویژگی برتر (%)

	PNN	KNN
نرخ طبقه‌بندی	۶۳,۷	۷۰,۴

مقایسه نتایج این دو آزمایش، مشخص می‌کند که با استفاده از ویژگی‌های تعیین شده، علاوه بر کاهش ابعاد بردار ویژگی، نرخ طبقه‌بندی بهتری حاصل می‌شود.

جدول (۴) ، ماتریس سردرگمی نتایج این آزمایش را با استفاده از طبقه‌بند KNN نشان می‌دهد. با توجه به نتایج ملاحظه می‌شود که لهجه کردی بالاترین و لهجه اصفهانی پایین-ترین نرخ طبقه‌بندی را دارد. همچنین تداخل بین لهجه‌های تهرانی و اصفهانی بیشتر از سایر لهجه‌ها است.

جدول (۴): ماتریس سردرگمی نتایج با استفاده از ویژگی‌های برتر و طبقه‌بند KNN (%)

	آذری	مازنی	تهرانی	کردی	اصفهانی
آذری	۶۶	۲	۴	۱۸	۱۰
مازنی	۰	۸۰	۱۲	۸	۰
تهرانی	۰	۲	۶۰	۸	۳۰
کردی	۱۰	۰	۰	۸۸	۲
اصفهانی	۰	۰	۳۶	۶	۵۸

متوسط نرخ طبقه‌بندی با استفاده از ویژگی‌های برتر تعیین شده و طبقه‌بند KNN، ۷۰,۴% است که نسبت به بالاترین نرخ گزارش شده در [۹] (۶۸%) بهبود ۲,۴% را نشان می‌دهد. همچنین این نتیجه اگر چه با مقدار ایده‌آل فاصله دارد اما در مقایسه با متوسط نرخ طبقه‌بندی توسط انسان (۷۶,۷% در جدول ۲) قابل قبول به نظر می‌رسد. البته، الگوریتم پیشنهادی در تشخیص برخی از لهجه‌ها (لهجه‌ی کردی در جدول ۴) بهتر عمل می‌کند.

۶- نتیجه‌گیری و کارهای آینده

این مقاله، به طبقه‌بندی لهجه‌های مختلف زبان فارسی پرداخته است. در راستای این پژوهش مجموعه دادگانی، تقریبا جامع از لهجه‌های مختلف زبان فارسی شامل تهرانی، آذری، کردی، اصفهانی و مازندرانی گردآوری شده است.

به منظور طبقه‌بندی، برخی از ویژگی‌های آماری مربوط به کانتور فرکانس پایه و مشتق آن استخراج می‌شود. در ادامه، برای انتخاب ویژگی‌های متمایز کننده، از الگوریتم انتخاب متوالی پیشرو استفاده می‌شود. نتایج پیاده‌سازی این الگوریتم نشان می‌دهد که ویژگی‌های مربوط به مشتق کانتور فرکانس پایه قدرت تمایز بیشتری برای طبقه‌بندی لهجه‌ها دارند. همچنین با بکارگیری این ویژگی‌ها نرخ طبقه‌بندی قابل مقایسه‌ای در برابر قدرت تشخیص انسان حاصل شده است.

این مقاله تنها به بررسی یکی از زیرمجموعه‌های ویژگی نوایی گفتار پرداخته است. در مطالعات آینده می‌توان از تمام ویژگی‌های نوایی گفتار، شامل الگوی استرس، ضرب‌آهنگ و زیر و بمی گفتار، برای بهبود نتایج استفاده کرد. همچنین با گسترش مجموعه دادگان و افزودن گفتار خودانگیز[9] به آن، می‌توان نتایج قابل اعتمادتری کسب کرد.

سپاسگزاری

از جناب آقای امیر پیله‌ور، جهت گردآوری مجموعه دادگان مورد استفاده در این تحقیق، بسیار سپاسگزاریم.

[1] C. Huang, T. Chen and E. Chang, *"Accent Issues in Large Vocabulary Continuous Speech Recognition"*, *International Journal of Speech Technology*, Vol. 7, No. 2/3, pp. 141-153, 2004.

[2] Q. Yan and S. Vaseghi, *"Modeling and synthesis of English regional accents with pitch and duration correlates"*, *Computer Speech and Language*, Vol. 24 No. 4, 2010.

[3] L. M. Arsalan and J. H. L. Hansen, *"Language accent classification in American English"*, *Speech Communication*, Vol. 18, No. 4, 1996.

[4] C. Pedersen and J. Diederich, *"Accent Classification Using Support Vector Machines"*, *6th IEEE. Conf. on Computer and Information Science*, pp. 444 – 449, 2007.

[5] A. Rabiee and S. Setayeshi, *"Persian Accents Identification Using an Adaptive Neural Network"*, *2th .Int. Conf. on Education Technology and Computer Science*, pp. 7-11, 2010.

[6] P. Mareuil and B. Vieru, *"The contribution of prosody to the perception of foreign accent"*, Phonetica, vol 63, No. 4, pp. 247–267, 2006.

[7] M. Jilka, *"The Contribution of Intonation to the Perception of Foreign Accent"*, Ph.D. Thesis, University of Stuttgart, Germany, 2000.

[8] B. Vieru, P. Boula and M. A. Decker, *"Characterisation and identification of non-native French accents"*, *speech communication*, Vol. 53, No 3, pp. 292-310, 2011.

[۹] علی قلی‌پور، محمد حسین صداقی و موسی شمسی، " طبقه بندی برخی از لهجه‌های زبان فارسی با استفاده از شبکه عصبی احتمالاتی"، بیستمین کنفرانس مهندسی برق ایران، تهران، ۱۳۹۱.

[10] M. H. Sedaaghi, C. Kotropoulos, D. Ververidis, *"Using Adaptive Genetic Algorithms to Improve Speech Emotion Recognition"*, 9th IEEE. Workshop. On MMSP, 2007.

زیرنویس‌ها

[1] Prosody
[2] Rhythm
[3] Intonation
[4] Mel Frequency Cepstral Coefficient
[5] Formant
[6] Pitch or Fundamental Frequency Contour
[7] Sequential Forward Selection
[8] Voiced Region
[9] Spontaneous Speech