



طراحی و تهیه دادگان تلفنی اعداد متصل زبان فارسی و استفاده از آن در یک سیستم بازشناسی خودکار گفتار مبتنی بر کلمه

محمد بحرانی

حسین ثامتی

امین فاضل

دانشکده مهندسی کامپیوتر - دانشگاه صنعتی شریف

bahrani@ce.sharif.edu

sameti@sharif.edu

afazel@mehr.sharif.edu

بگیرد. هدف اصلی از تهیه این دادگان فراهم کردن داده‌های آموزشی برای یک سیستم بازشناسی گفتار مبتنی بر واحد کلمه برای بازشناسی اعداد از طریق خط تلفن است و هدف نهایی، استفاده از آن در محدوده وسیعی از طرح‌های پژوهشی و تجاری می‌باشد. برای زبان فارسی دادگان تلفنی معتبر زیادی در دسترس نیست. دو دادگان تلفنی OGI multi-language [۱] و Call Friend Farsi [۲] برای زبان فارسی وجود دارند که جهت بازشناسی خودکار زبان^۱ مورد استفاده قرار می‌گیرند و تنها دادگان تلفنی آماده شده جهت بازشناسی گفتار دادگان TFarsdat [۳] می‌باشد. این دادگان در سطح واج تقطیع شده است و معمولاً برای سیستم‌های بازشناسی گفتار که واحد آوایی آن‌ها واج می‌باشد مورد استفاده قرار می‌گیرد. از آنجایی که در این دادگان فرکانس تکرار هر کلمه، به خصوص اعداد، کم می‌باشد از آن نمی‌توان برای سیستم‌های بازشناسی مبتنی بر کلمه استفاده کرد. البته دادگان تلفنی دیگری که در دانشگاه‌های کشور تهیه شده‌اند نیز وجود دارند. این دادگان‌ها معمولاً در سطح آزمایشگاهی و برای پروژه‌های خاص تهیه شده‌اند. از آن جمله می‌توان به دادگان اعداد FarsNumbers1 که در دانشکده مهندسی کامپیوتر دانشگاه امیرکبیر تهیه شده است اشاره کرد. این دادگان دارای برچسب‌دهی در سطح واج می‌باشد. دادگان

چکیده: این مقاله تحقیقات و کارهای انجام شده در یک پروژه تهیه دادگان تلفنی اعداد زبان فارسی برای سیستم‌های بازشناسی گفتار را شرح می‌دهد. این دادگان در حال حاضر شامل داده‌های ضبط شده تلفنی از اعداد فارسی به صورت متصل است و مشتمل بر ۱۱۰ گوینده می‌باشد که هر یک تقریباً ۷۰ رشته عددی را بیان کرده‌اند. هر رشته عددی جداگانه برچسب‌دهی شده است. آزمایش‌های متنوعی که به وسیله شنوندگان مختلف بر روی این دادگان صورت گرفته، کیفیت دادگان از نظر قابل فهم بودن برای استفاده‌کنندگان انسانی و همچنین صحت برچسب‌گذاری دادگان را تضمین کرده است. از این دادگان در یک سیستم بازشناسی گفتار پیوسته استفاده شده است. با به‌کارگیری این دادگان، دقت بازشناسی اعداد متصل تلفنی بیش از ۹۷٪ بوده است.

واژه‌های کلیدی: دادگان اعداد تلفنی، بازشناسی گفتار تلفنی، بازشناسی خودکار گفتار

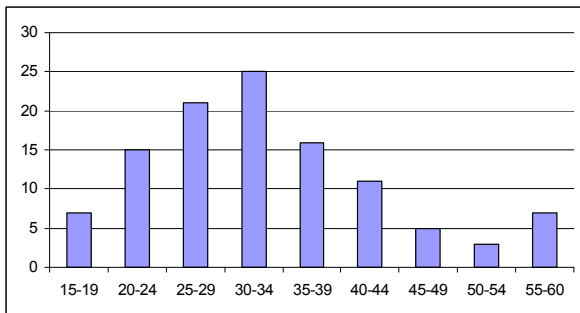
۱- مقدمه

برای آموزش سیستم‌های بازشناسی گفتار به یک دادگان معتبر نیاز است به طوری که سیستم بتواند از روی دادگان آموزشی تمامی واحدهای آوایی و گذارهای بین آن‌ها را فرا

¹ Automatic Language Recognition

جدول ۱: نحوه توزیع گویندگان دادگان با توجه به محدوده سنی آنها

محدوده سنی	مرد	زن	تعداد کل	درصد
۱۹-۱۵	۳	۴	۷	۶.۳۶٪
۲۴-۲۰	۸	۷	۱۵	۱۳.۴۶٪
۲۹-۲۵	۱۲	۹	۲۱	۱۹.۰۹٪
۳۴-۳۰	۱۶	۹	۲۵	۲۲.۷۳٪
۳۹-۳۵	۱۱	۵	۱۶	۱۴.۵۵٪
۴۴-۴۰	۶	۵	۱۱	۱۰٪
۴۹-۴۵	۲	۳	۵	۴.۵۵٪
۵۴-۵۰	۳	۰	۳	۲.۷۳٪
۶۰-۵۵	۶	۱	۷	۶.۳۶٪
مجموع	۶۷	۴۳	۱۱۰	۱۰۰٪



شکل ۱: نحوه توزیع گویندگان دادگان با توجه به محدوده سنی آنها

۲-۲- واژگان

گفتاری که از گویندگان در این دادگان جمع‌آوری شده است، رشته‌های عددی می‌باشد. اعدادی که گویندگان بیان کرده‌اند عبارتند از: "صفر"، "یک"، "دو"، "سه"، "چهار"، "پنج"، "شش"، "هفت"، "هشت" و "نه". هر گوینده ۷۰ رشته از این اعداد را به طور متصل^۱ به شکل زیر بیان کرده است.

۲۰ رشته یک رقمی (هر کدام از ارقام "صفر" تا "نه" دو بار ادا شده اند).

۱۰ رشته دو رقمی

۱۰ رشته سه رقمی

۱۰ رشته چهار رقمی

۱۰ رشته پنج رقمی

۱۰ رشته هفت رقمی

¹ connected

اعداد دیگری نیز در مرکز تحقیقات فناوری اطلاعات دانشگاه علم و صنعت تهیه شده است.

این مقاله کارهای انجام شده برای رسیدن به هدف فوق را بیان می‌کند. در بخش بعد مشخصات کلی دادگان بیان می‌شود، بخش ۳ کارهای انجام شده برای بیان اعتبار دادگان را معرفی می‌کند و سپس در بخش ۴ استفاده از این دادگان در یک سیستم بازشناسی گفتار آمده است. در بخش انتهایی نیز نتیجه‌گیری و کارهای آینده مشخص شده است.

۲- مشخصات کلی دادگان

این دادگان که FCDigits نامیده شده، به منظور استفاده در سیستم‌های بازشناسی گفتار، در پژوهشکده پردازش هوشمند علائم ضبط گردیده است. کار مطالعه و طراحی این دادگان در حدود ۶ ماه به طول انجامیده است. در زیر مشخصات کلی این دادگان آمده است.

۲-۱- توصیف گویندگان

گویندگان با توجه به سن، جنس و میزان تحصیلات انتخاب شده‌اند. تعداد کل گویندگانی که در جمع‌آوری این دادگان همکاری کرده‌اند برابر با ۱۱۰ نفر می‌باشد. این دادگان فقط در بردارنده گویندگانی با لهجه تهرانی است که از نقاط مختلف شهر تهران از طریق شبکه تلفن با مرکز جمع‌آوری دادگان تماس گرفته‌اند و رشته‌های عددی از پیش تعیین شده را بیان کرده‌اند. نسبت تعداد گویندگان مرد به زن در این دادگان تقریباً برابر با ۳ به ۲ می‌باشد که بیشترین آنها در محدوده سنی بین ۲۵ تا ۴۰ سال انتخاب شده‌اند. تعداد زنان و مردان در هر محدوده سنی و تعداد کل آنها در جدول ۱ آمده است. نمودار نشان داده شده در شکل ۱ نیز بیان کننده توزیع مناسب افراد در محدوده‌های سنی مختلف می‌باشد.

۲-۳- جمع‌آوری دادگان

برای جمع‌آوری دادگان ابتدا لیست رشته‌های عددی مربوط به هر گوینده به آنها داده شد؛ سپس از طریق تماس تلفنی که از پژوهشکده پردازش هوشمند علائم با آنها گرفته شد، از آنها خواسته شد که رشته‌های عددی را بیان کنند. قابل ذکر است که گویندگان رشته‌های عددی را از طریق تلفن ثابت بیان کردند. از گویندگان خواسته شده است که تا حد امکان در اتاقی بدون نویز رشته‌های عددی را بیان کنند و از تلفن‌های بی‌سیم تیز استفاده نکنند. همچنین اعداد تا حد امکان به صورت متصل و بدون مکث ادا شده‌اند. صدای هر گوینده از طریق رابطی که از گوشی تلفن به ورودی کارت صدای کامپیوتر متصل بوده ضبط گردیده است. این رابط یک مدار یک طرف کننده صدا و دارای خروجی برای کارت صدای کامپیوتر می‌باشد. هنگامی که این رابط در ورودی کارت صدای کامپیوتر قرار می‌گیرد تنها صدای کسی که با مرکز تماس گرفته است به ورودی کارت صدا وارد می‌شود. نرخ نمونه برداری برابر با ۱۱,۰۲۵ کیلوهرتز و هر نمونه، شانزده بیتی می‌باشد. برای رقمی کردن نمونه‌ها از کارت صدای Creative PC64 استفاده شده است. این کارت صدای PCI ضبط صدا را با A/D ۱۶ بیتی انجام می‌دهد و دارای سیگنال به نویز حدود ۹۵ دسی‌بل می‌باشد.

برای هر گوینده یک فایل برچسب^۱ وجود دارد که برچسب رشته‌های عددی را به ترتیبی که گوینده ادا کرده، در خود جای داده است. برچسب هر رشته عددی شامل دنباله ارقام موجود در آن رشته می‌باشد.

۳- تصدیق صحت دادگان و آمایش آن

برای اطمینان از انطباق دادگان جمع‌آوری شده با لیست‌های آماده شده جهت هر گوینده، از گروهی از شنوندگان استفاده گردید. آنان هر یک از فایل‌های صوتی دادگان را با لیست‌های از پیش آماده شده مقایسه می‌کردند و در صورت مشاهده عدم تطابق بین لیست‌ها و گفتار ضبط شده، اشکال موجود گزارش می‌شد تا آن قسمت از دادگان دوباره ضبط گردد.

بنابراین هر گوینده ۲۳۰ عدد و ۱۶۰ حالت گذر بین عددی را بیان کرده است. از الگوریتم زیر برای تولید لیست یکتایی از رشته‌های عددی، که هر گوینده بیان می‌کند، استفاده شده است.

مرحله ۱: یک آرایه هفتاد عنصری تولید کن. این آرایه در بر دارنده طول رشته‌های عددی است. به بیست عنصر این آرایه، که به‌طور تصادفی انتخاب شده‌اند، مقدار ۱ و به هر ده عنصر آرایه، که آنها نیز به‌طور تصادفی انتخاب شده‌اند، مقادیر ۲، ۳، ۴، ۵ و ۷ نسبت بده. این آرایه بیان کننده مکان هر رشته عددی یک، دو، سه، چهار، پنج و هفت رقمی در لیستی که هر گوینده بیان می‌کند، می‌باشد.

مرحله ۲: برای رشته اعداد یک رقمی، از یک لیست بیست عنصری حاوی ارقام ۰ تا ۹، که هر یک دو بار تکرار شده‌اند، یک رقم را به‌طور تصادفی و بدون جایگزینی انتخاب کن. این ارقام به ترتیبی که آرایه هفتاد عنصری برای مکان رشته‌های عددی یک رقمی تعیین کرده است در لیست نهایی مربوط به گوینده قرار می‌گیرند.

مرحله ۳: در هر یک از ۵۰ مکان باقی‌مانده (مربوط به رشته‌های عددی با طول دو رقم و بیشتر) برای تعیین اولین رقم رشته‌های عددی، از یک لیست ده عنصری شامل ارقام ۰ تا ۹، یک رقم را به‌صورت تصادفی و بدون جایگزینی انتخاب کن. اگر عدد انتخابی اول به عنوان مثال ۳ باشد یعنی این که رقم اول اولین رشته‌های عددی دو، سه، چهار، پنج و هفت رقمی برابر ۳ خواهد بود.

مرحله ۴: برای تعیین ارقام بعدی در رشته‌های عددی با طول دو رقم و بیشتر، به‌صورت تصادفی و بدون جایگزینی از لیست‌های گذار مربوط به ارقام قبلی یک رقم را انتخاب کن. ده لیست گذار، مربوط به هر رقم، وجود دارد که شامل ارقام ۰ تا ۹ است و بیان کننده رقم بعدی هر رقم می‌باشد. پس از انتخاب یک رقم از لیست گذار، رقم مذکور از این لیست حذف می‌گردد (این کار تضمین می‌کند که در رشته‌های با طول مختلف، گذارهای بین عددی مشابه وجود نداشته باشد). مرحله ۵: به ازای هر گوینده، مراحل ۱ تا ۴ را تکرار کن.

¹ label

همچنین آزمایش‌هایی توسط گروهی از شنوندگان خبره برای تعیین کیفیت دادگان و همچنین قابل فهم بودن آن صورت پذیرفت تا در صورت وجود اشکال، کار تهیه آن قسمت از دادگان دوباره صورت گیرد. در انتها و پس از جمع‌آوری کامل دادگان نیز فایل‌های صوتی دوباره مورد بازبینی قرار گرفتند. در این مرحله نیز آن دسته از فایل‌های صوتی که دارای اشکال بودند از دادگان حذف گردیدند. در جدول ۲ مشخصات کلی این دادگان همراه با مقایسه آن با سایر دادگان‌های فارسی موجود آمده است. همانطور که بیان شد دادگان‌های تهیه شده در دانشگاه‌ها از آن جهت که اطلاعات آن‌ها در اختیار عموم قرار نگرفته‌اند و موارد استفاده خاص در یک یا چند پروژه تحقیقاتی داشته‌اند در این جدول برای مقایسه نیامده‌اند.

جدول ۲: مشخصات کلی دادگان تلفنی اعداد متصل زبان فارسی و مقایسه آن با سایر دادگان‌های تلفنی موجود زبان فارسی

نام دادگان	کاربرد	تعداد گویندگان	اندازه کل دادگان (ساعت)	اندازه دادگان برچسب دهی شده (ساعت)
TFarsdat	بازشناسی گفتار	۶۴	۸	۸
OGI Multilanguage	بازشناسی زبان (همه زبان‌ها)	۲۰۵۲	۳۸.۵	۸
Call Friend FARSI	بازشناسی زبان	۶۰	۳۰	-
FCDigits	بازشناسی گفتار	۱۱۰	۷	۷
FarsNumbers 1	بازشناسی گفتار	۵۷	۱:۴۰	-

آوایی را در این سیستم می‌توان جداگانه تعیین کرد. تعیین توپولوژی شامل تعیین تعداد حالت^۱، تعداد تلفیق^۲ در هر حالت و تعیین پرش‌های مجاز بین حالت‌ها می‌باشد. در این آزمایش‌ها تعداد حالت‌ها برای هر مدل مخفی مارکف برابر ۱۶ حالت و در هر حالت ۳ تلفیق می‌باشد و گام پرش مجاز بین حالت‌های هر مدل برابر ۱ در نظر گرفته شده است. با استفاده از این مدل‌ها سیستم مذکور تطبیق الگو را برای بازشناسی سیگنال گفتار ورودی انجام می‌دهد و برای این منظور از جستجوی ویتربی شعاعی برای یافتن بهترین دنباله واحدهای آوایی متناظر با دنباله آکوستیکی ورودی بهره می‌گیرد. واحد آوایی به کار برده شده در این سیستم، کلمه می‌باشد یعنی به ازای هر کلمه در دادگان آموزشی، یک مدل مخفی مارکوف آموزش داده می‌شود؛ بنابراین ورودی سیستم یک سیگنال گفتار و خروجی سیستم، بهترین دنباله کلمات یافت شده متناظر با دنباله آکوستیک ورودی است. بازنمایی‌های مورد استفاده در این سیستم MFCC بوده که شامل ۱۲ ضریب استاندارد مل-کپستروم و مشتقات اول و دوم آن‌ها، همراه با یک ضریب لگاریتم انرژی و مشتقات اول و دوم آن می‌باشد. اندازه طول هر قاب ۲۵ میلی‌ثانیه و میزان هم‌پوشانی ۱۰ میلی‌ثانیه در نظر گرفته شده است. همچنین برای حذف DC احتمالی موجود در فایل صوتی ورودی، تابعی به شکل زیر در نظر گرفته شده است که تا حد امکان بتواند آن را حذف کند:

$$y[n] = x[n] - x[n-1] + 0.999 * y[n-1] \quad (1)$$

که $x[n]$ ورودی این تابع و $y[n]$ خروجی آن می‌باشد. جهت استفاده از دادگان FCDigits در سیستم بازشناسی گفتار مذکور، این دادگان به دو مجموعه آموزش و آزمایش تقسیم شد. از بین ۱۱۰ گوینده موجود در دادگان ۱۳ گوینده شامل ۸ گوینده مرد و ۵ گوینده زن به عنوان گویندگان مجموعه آزمایش به طور کاملاً تصادفی انتخاب گردیدند. مجموعه آموزش شامل ۷۲۱۶ رشته عددی می‌باشد که دارای ۲۴۰۹۶ کلمه (عدد) است. در مجموعه آزمایش نیز ۹۷۵ رشته عددی شامل ۳۲۶۳ کلمه (عدد) وجود دارد. از آنجایی که در

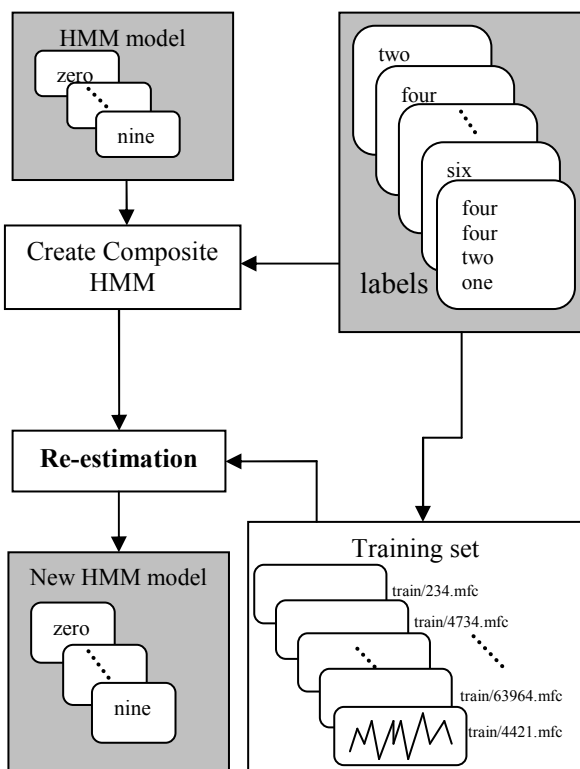
۴- آزمایش‌های انجام شده بر روی دادگان به وسیله سیستم بازشناسی گفتار مبتنی بر کلمه

پس از آماده‌سازی دادگان فوق، آزمایش‌هایی به کمک یک سیستم بازشناسی گفتار پیوسته [۴] انجام شد. سیستم بازشناسی مذکور، مدل‌سازی واحدهای آوایی را با استفاده از مدل مخفی مارکوف انجام می‌دهد. مدل‌های مخفی مارکوف به کار برده شده از نوع چگالی پیوسته و با تلفیق‌های گوسی می‌باشند و پرش بین حالت‌های آن فقط به صورت چپ به راست مجاز می‌باشد. توپولوژی مدل‌های مربوط به هر واحد

¹ state

² mixture

مدل‌های اولیه محاسبه می‌شوند. مراحل به‌دست آوردن تخمین جدید در شکل ۳ آمده است.

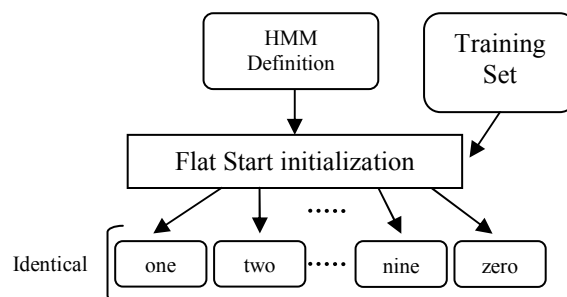


شکل ۳: بلوک دیاگرام مراحل به‌دست آوردن تخمین جدید

از آن‌جا که ممکن است گوینده در فاصله‌ی ادای هر رقم کمی مکث کرده باشد، یک مدل سکوت نیز آموزش داده شده تا مکث‌های گوینده در حین صحبت را پوشش دهد. نتایج حاصل از بازشناسی در جدول ۳ آمده است. همان‌طور که در این جدول مشاهده می‌شود، سیستم بازشناسی پس از آموزش با این دادگان توانسته است بر روی دادگان آزمایش به دقتی در حدود ۹۷٪ دست یابد. قابل ذکر است که آزمایش سیستم فوق در محیط عملی و از پشت خط تلفن نیز گویای نتایج زیر است و سیستم در اکثر موارد توانایی بازشناسی صدای گوینده را دارا می‌باشد.

فرکانس‌های بالاتر از ۸ کیلوهرتز عملاً هیچگونه داده‌ای وجود ندارد، فایل‌های این دادگان جهت انجام آزمایش‌ها به ۸ کیلوهرتز زیرنمونه برداری شده‌اند.

از آنجایی که این دادگان دارای تقطیع در سطح کلمه نمی‌باشد و تنها هر رشته‌ی عددی یک تا هفت کلمه‌ای دارای برجسب می‌باشد، برای آموزش سیستم، در قسمت مقداردهی اولیه از روش Flat Start استفاده شده است. در این روش به حالت‌های همه‌ی مدل‌ها یک میانگین و واریانس برابر، به عنوان مقدار اولیه، نسبت داده می‌شود. بلوک دیاگرام این روش در شکل ۲ آمده است.



شکل ۲: بلوک دیاگرام مقداردهی اولیه

برای تخمین دوباره‌ی مدل‌های به‌دست آمده، با استفاده از همه‌ی داده‌های آموزشی، به طور همزمان همه‌ی مدل‌ها به‌روز می‌شوند. در ابتدا مدل‌های مخفی مارکف اولیه که به روش فوق به‌دست آمده‌اند، خوانده می‌شوند. همان‌طور که بیان شد هر یک از فایل‌های صوتی رشته‌های عددی دارای برجسبی معادل با آن چیزی که گوینده بیان کرده است، می‌باشد. با استفاده از این برجسب‌ها، مدل مخفی مارکف مرکبی ساخته می‌شود که در بر گیرنده کل رشته عددی بیان شده توسط گوینده می‌باشد. این مدل مرکب از پشت سرهم قرار دادن مدل مخفی مارکف هر یک از کلمات موجود در رشته عددی ساخته می‌شود. سپس الگوریتم جلورو_عقب‌رو^۱ [۵] بر روی هر یک از این مدل‌های مرکب اجرا می‌شود و مقادیر متغیرهای جلورو و عقب‌رو محاسبه می‌گردند. هنگامی که همه‌ی فایل‌های موجود در مجموعه آموزش پردازش شدند، پارامترهای تخمین جدید با روش Baum-Welch [۵]، از روی متغیرهای جلورو و عقب‌رو و

^۱ forward-backward

جدول ۳: نتایج بازشناسی حاصل از آموزش سیستم با دادگان FCDigits

دقت بازشناسی رشته‌های عددی	کلمه		مجموعه مورد آزمایش
	صحت	دقت	
%۹۱.۴۹	%۹۷.۳۰	%۹۷.۱۲	مجموعه آزمون
%۹۰.۳۷	%۹۶.۹۷	%۹۶.۶۵	مجموعه آموزش

کارهای زیر برای ادامه کار وجود دارند تا این که بتوان به یک دادگان غنی تلفنی اعداد در زبان فارسی دست یافت. تهیه دادگان تلفنی برای اعداد پیوسته زبان فارسی و همچنین افزایش تعداد گویندگان دادگان فعلی از مهمترین کارهایی است که باید انجام گردد. دادگان حاضر فقط در بر دارنده لهجه تهرانی است. برای داشتن یک دادگان کامل جهت سیستم‌های بازشناسی مستقل از گوینده نیاز به لهجه‌های متنوع زبان فارسی می‌باشد. همچنین تقطیع در سطح کلمه برای دادگان اعداد متصل زبان فارسی می‌تواند تاثیر مثبت زیادی در بهبود کارایی سیستم‌های بازشناسی گفتار داشته باشد.

برای افزایش دقت بازشناسی نیز، کارهای دیگری مانند تطبیق مدل‌ها با شرایط تلفنی مختلف و همچنین استفاده از بازنمایی‌های مقاوم جهت استفاده از آن‌ها در شرایط نویزی متفاوت در حال انجام است.

۶- سپاسگزاری

از پژوهشکده پردازش هوشمند علائم و به خصوص آقای دکتر شیخ‌زادگان و خانم حسینی که در تهیه این دادگان ما را یاری دادند، متشکریم.

مراجع

- [1] Muthusamy, Y. K. et. al., "The OGI Multi-Language Telephone Speech Corpus," *Proc. of the ICSLP*. USA, pp. 895-898, 1992.
- [2] LDC Homepage:
<http://WWW ldc.upenn.edu/catalog/catalogEntry.jsp?catalogId=LDC 96S50>.
- [3] Bijankhan M., et. al., "TFarsdat - the Telephone Farsi Speech Database," *Proc. Eurospeech*. Scandinavia, pp. 1525-1528, 2003.
- [4] B. Babaali, H. Sameti, "The Sharif Speaker-Independent Large Vocabulary Speech Recognition System", The 2nd Workshop on Information Technology & Its Disciplines (WITID 2004), Kish Island, Iran, pp. 85-88 2004.
- [5] L. R. Rabiner and B. H. Juang, *Fundamentals of Speech Recognition*, Englewood Cliffs, New Jersey: Prentice Hall, 1993.

در آزمایش دیگر، مدل‌های آموزش داده شده با این دادگان، با مدل‌های آموزش داده شده به وسیله دادگان TFarsdat، از نظر دقت بازشناسی مقایسه گردیدند. برای این که مقایسه معنی‌داری صورت بگیرد، باید مجموعه آزمون از خارج از دو دادگان انتخاب شده باشد. بدین منظور ۱۰۲ رشته عددی متصل (در کل شامل ۴۰۶ عدد) از طریق خط تلفن ضبط گردید. نتایج حاصل از بازشناسی این رشته‌های عددی در جدول ۴ آمده است. همان‌طور که مشاهده می‌شود، با استفاده از دادگان طراحی شده، حدود ۸۲ درصد کاهش در نرخ خطای کلمه نسبت به دادگان TFarsdat حاصل شده است.

جدول ۴: مقایسه دقت بازشناسی حاصل از آموزش سیستم با دادگان‌های

TFarsdat و FCDigits

دقت بازشناسی رشته‌های عددی	کلمه		مجموعه مورد استفاده جهت آزمون
	صحت	دقت	
%۹۴.۱۲	%۹۹.۵۱	%۹۸.۵۲	FCDigits
%۷۰.۵۹	%۹۲.۶۱	%۹۱.۸۷	TFarsdat

۵- بحث و نتیجه گیری

در این مقاله پروژه در حال انجامی برای تهیه دادگان تلفنی اعداد فارسی برای سیستم‌های بازشناسی خودکار گفتار معرفی شد که تا کنون قسمت اعداد متصل آن آماده شده است و قسمت مربوط به اعداد پیوسته^۱ در حال آماده‌سازی می‌باشد. نتایج چشمگیر حاصل از استفاده از این دادگان در سیستم بازشناسی گفتار برای بازشناسی اعداد متصل تلفنی، بیانگر کیفیت خوب طراحی این دادگان می‌باشد.

¹ continuous