دانشگاه آزاد اسلامی

واحد علوم و تحقیقات

دانشکده برق الکترونیک

پایان نامه برای دریافت درجه کارشناسی ارشد در رشته مهندسی برق

گرایش الکترونیک

عنوان:

تشخیص نوع احساس از روی گفتار با استفاده از ویژگی های زمان – فرکانسی

استاد راهنما:

جناب آقای دکتر حسین مروی

استاد مشاور:

جناب آقای دکتر امیدرضا معروضی

نگارش:

زینب اسماعیلیان

تابستان ۹۲

بسم الله الرحمن الرحيم

معاونت پژوهش و فن آوری

به نام خدا

منشور اخلاق پژوهش

با یاری از خداوند سبحان و اعتقاد به این که عالم محضر خداست و همواره ناظر بر اعمال انسان و به منظور پاس داشت مقام بلند دانش و پژوهش و نظر به اهمیت جایگاه دانشگاه در اعتلای فرهنگ و تمدن بشری، ما دانشجویان و اعضاء هیات علمی واحدهای دانشگاه آزاد اسلامی متعهد می گردیم اصول زیر را در انجام فعالیت های پژوهشی مد نظر قرار داده و از آن تخطی نکنیم:

۱- اصل حقیقت جویی: تلاش در راستای پی جویی حقیقت و وفاداری به آن و دوری از هرگونه پنهان سازی حقیقت.

۲- اصل رعایت حقوق: التزام به رعایت کامل حقوق پژوهشگران و پژوهیدگان (انسان، حیوان و نبات) و سایر صاحبان حق.

۳- اصل مالکیت مادی و معنوی: تعهد به رعایت کامل حقوق مادی و معنوی دانشگاه و کلیه همکاران پژوهش.

۴- اصل منافع ملی: تعهد به رعایت مصالح ملی و در نظر داشتن پیشبرد و توسعه کشور در کلیه مراحل پژوهش.

۵- اصل رعایت انصاف و امانت: تعهد به اجتناب از هرگونه جانب داری غیر علمی و حفاظت از اموال، تجهیزات و منابع در اختیار.

۶- اصل رازداری: تعهد به صیانت از اسرار و اطلاعات محرمانه افراد، سازمان ها و کشور و کلیه افراد و نهادهای مرتبط با تحقیق.

۷- اصل احترام: تعهد به رعایت حریم ها و حرمت ها در انجام تحقیقات و رعایت جانب نقد و خودداری از هرگونه حرمت شکنی.

۸- اصل ترویج: تعهد به رواج دانش و اشاعه نتایج تحقیقات و انتقال آن به همکاران علمی و دانشجویان به غیر از مواردی که منع قانونی دارد.

۹- اصل برائت: التزام به برائت جویی از هرگونه رفتار غیر حرفه ای و اعلام موضع نسبت به کسانی که حوزه علم و پژوهش را به شائبه های غیر علمی می آلایند.

دانشگاه آزاد اسلامی

پردیس تحصیلات تکمیلی علوم و تحقیقات شاهرود

تعهد نامه اصالت رساله یا پایان نامه

اینجانب زینب اسماعیلیان دانش آموخته مقطع کارشناسی ارشد ناپیوسته در رشته مهندسی برق که در تاریخ ۱۳۹۲/۰٤/۳۰ از پایان نامه خود تحت عنوان تشخیص نوع احساس از روی گفتار با استفاده از ویژگی های زمان – فرکانسی **با کسب نمره ۱۹/۱٦ و درجه عالی دفاع نموده ام بدینوسیله متعهد می‌شوم :**

۱) **این پایان نامه حاصل تحقیق و پژوهش انجام شده توسط اینجانب بوده و در مواردی که از دستاوردهای علمی و پژوهشی دیگران (اعم از پایان نامه، کتاب، مقاله و ....) استفاده نموده‌ام، مطابق ضوابط و رویه موجود، نام منبع مورد استفاده و سایر مشخصات آنرا در فهرست مربوطه ذکر و درج کرده‌ام.**

۲) **این پایان نامه قبلا برای دریافت هیچ مدرک تحصیلی (هم سطح، پایین تر یا بالاتر) در سایر دانشگاهها و موسسات آموزش عالی ارائه نشده است.**

۳) **چنانچه بعد از فراغت از تحصیل، قصد استفاده و هرگونه بهره برداری اعم از چاپ کتاب، ثبت اختراع و ... از این پایان نامه یا رساله را داشته باشم، از حوزه معاونت پژوهشی واحد، مجوزهای مربوطه را اخذ نمایم.**

٤) **چنانچه در هر مقطع زمانی خلاف موارد فوق ثابت شود، عواقب ناشی از آن را می‌پذیرم و دانشگاه آزاد اسلامی واحد علوم و تحقیقات مجاز است با اینجانب مطابق ضوابط و مقررات رفتار نموده و در صورت ابطال مدرک تحصیلی‌ام هیچگونه ادعایی نخواهم داشت.**

نام و نام خانوادگی: زینب اسماعیلیان

تاریخ و امضاء:

دانشگاه آزاد اسلامی

پردیس علوم و تحقیقات شاهرود

پایان نامه کارشناسی ارشد رشته برق(M.Sc)

گرایش: الکترونیک

عنوان :

تشخیص نوع احساس از روی گفتار با استفاده از ویژگی های زمان-فرکانسی

نگارنده

زینب اسماعیلیان

سال تحصیلی ۱۳۹۲ - ۱۳۹۱

هیأت داوران :

۱. دکتر حسین مروی ـ استاد راهنما

۲. دکتر امیدرضا معروضی ـ استاد مشاور

۳. دکترعلیرضا احمدی فرد ـ استاد داور

سپاسگزاری

به مصداق « من لم یشکر المخلوق لم یشکر الخالق» بسی شایسته است از استاد فرهیخته و فرزانه جناب آقای دکتر حسین مروی که با راهنمایی های کارساز و سازنده خود، مرا در انجام هرچه بهتر این پایان نامه یاری نمودند، تقدیر و تشکر نمایم.

همچنین از استاد گرانقدر جناب آقای دکتر امیدرضا معروضی به خاطر قبول مشاوره این پایان نامه کمال تشکر را دارم.

واز پدر و مادرم ، این دو معلم بزرگوارم که همواره بر کوتاهی و درشتی من، قلم عفو کشیده و کریمانه از کنار غفلت هایم گذشته اند و در تمام عرصه های زندگی یار و یاوری بی چشم داشت برای من بوده اند، تشکر و قدردانی می نمایم.

تقدیم به همسرم:

که سایه مهربانیش سایه سار زندگیم می باشد، او که اسوه صبر و تحمل بوده و مشکلات مسیر را برایم تسهیل نمود.

فهرست مطالب

عنوان	شماره صفحه
چکیده	۱
فصل اول: احساس و تأثیر آن بر روی گفتار	
۱-۱- مقدمه	۳
۱-۲- تئوری احساس	۵
۱-۳- مدل احساس	٦
۱-۳-۱- مدل گسسته	٦
۱-۳-۲- مدل پیوسته	٦
۱-٤- ارتباط احساس و مغز	۸
۱-۵- وابستگی احساس با جنسیت	۹
فصل دوم: مروری بر پیشینه ی تشخیص احساس از روی گفتار	
۲-۱- مقدمه	۱۲
۲-۲- سیستم تشخیص احساس از روی گفتار	۱۲
۲-۲-۱- استخراج ویژگی	۱۲
۲-۲-۱-۱- ویژگی های عروضی	۱۳
۲-۲-۱-۱-۱- فرکانس گام	۱۳
۲-۲-۱-۱-۲- انرژی	۱٤
۲-۲-۱-۲- ویژگی های طیفی	۱۵
۲-۲-۱-۲-۱- فرمنت ها	۱۵
۲-۲-۱-۲-۲-ضرایب مل کپستروم (MFCC)	۱٦
۲-۲-۱-۲-۳- ضرایب پیشگویی خطی (PLP)	۱۷

۲-۲-۱-۲-٤- آنالیز پیشگویی خطی (LPC) ۱۹
۲-۲-۱-۳- ویژگی های کیفیت صدا ۱۹
۲-۲-۱-۳-۱- Jitter ۲۰
۲-۲-۱-۳-۲- Shimmer ۲۰
۲-۲-۱-۳-۳- نسبت هارمونیک به نویز(HNR) ۲۱
۲-۲-۱-٤- سایر ویژگی ها ۲۱
۲-۲-۱-٤-۱- اپراتور انرژی تیگر(TEO) ۲۱
۲-۲-۱-٤-۲- نرخ عبور از صفر ۲۲
۲-۲-۱-۵- ویژگی های محلی و سراسری ۲۳
۲-۲-۲- کاهش ویژگی ۲۳
۲-۲-۲-۱- فیلترها ۲۵
۲-۲-۲-۱-۱- رتبه بندی با معیار فیشر (FDR) ۲۵
۲-۲-۲-۲- Wrapper ۲۵
۲-۲-۲-۳- جاسازها ۲۷
۲-۲-۲-۳-۱- آنالیز مؤلفه های اصلی (PCA) ۲۷
۲-۲-۲-۳-۲- الگوریتم جداساز خطی (LDA) ۲۷
۲-۲-۳- کلاسه بندی ۲۸
۲-۲-۳-۱- الگوریتم ماشین بردار پشتیبان (SVM) ۲۸
۲-۲-۳-۲- الگوریتم کلاسه بند جداساز خطی (LDA) ۳۱
۲-۳- پایگاه داده احساسی گفتار ۳۲
۲-۳-۱- معرفی چند پایگاه داده احساسی متداول ۳۳
۲-۳-۱-۱- پایگاه داده احساسی برلین ۳۵
۲-۳-۱-۲- پایگاه داده احساسی VAM ۳۷

۲-۳-۲- مشکلات موجود در پایگاه های داده احساسی ۳۷
۲-٤- جمع بندی مطالعات ۳۸
فصل سوم: پایگاه داده پیشنهادی و سیستم طراحی شده برای تشخیص احساس از روی گفتار
۳-۱- مقدمه ٤۳
۳-۲- پایگاه داده پیشنهادی ٤۳
۳-۳- سیستم تشخیص احساس طراحی شده ٤٥
۳-۳-۱- پیش پردازش ٤۷
۳-۳-۲- استخراج ویژگی ٤۷
۳-۳-۲-۱- ویژگی های مبتنی بر وریوگرام ٤۸
۳-۳-۲-۲- ویژگی های مبتنی بر مل کپستروم دو بعدی ٥۹
فصل چهارم: نتایج
٤-۱- مقدمه ٦٤
٤-۲- آزمایش اول ٦٤
٤-۲-۱- استخراج ویژگی ٦٤
٤-۲-۲- کاهش ویژگی ٦٥
٤-۲-۳- کلاسه بندی ٦٥
٤-۲-٤- نتایج آزمایش اول ٦٦
٤-۳- آزمایش دوم ۷۳
٤-۳-۱- استخراج ویژگی ۷۳
٤-۳-۲- کاهش ویژگی ۷٤
٤-۳-۳- کلاسه بندی ۷٤
٤-۳-٤- نتایج آزمایش دوم ۷٤

٤-٤- آزمایش سوم ٨٤
٤-٤-١- استخراج ویژگی ٨٤
٤-٤-٢- کاهش ویژگی ٨٤
٤-٤-٣- کلاسه بندی ٨٥
٤-٤-٤- نتایج آزمایش سوم ٨٥

فصل پنجم: نتیجه گیری و پیشنهادات

٥-١- مقدمه ٩٢
٥-٢- آزمایش اول ٩٢
٥-٣- آزمایش دوم ٩٣
٥-٤- آزمایش سوم ٩٣
٥-٥- جمع بندی ٩٣
٥-٦- پیشنهادات ٩٤

منابع و مآخذ

فهرست منابع فارسی ٩٦
فهرست منابع انگلیسی ٩٧
فهرست مقالات مستخرج شده از پایان نامه ١٠٦
چکیده انگلیسی ١٠٧

فهرست جداول

عنوان	شماره صفحه
جدول(۲-۱): معرفی چند پایگاه داده متداول.	۳٤
جدول(۲-۲): مروری بر کارهای گذشته در زمینه تشخیص احساس از روی گفتار.	۳۹
جدول(۳-۱): اطلاعات پایگاه داده درام (پیشنهادی).	٤٤
جدول(٤-۱): لیست ویژگی های استخراج شده در آزمایش اول.	٦٥
جدول(٤-۲): نرخ تشخیص برای ویژگی های مختلف برای پایگاه داده درام(پیشنهادی) و پایگاه داده برلین در آزمایش اول.	٦۷
جدول(٤-۳): ماتریس تداخل برای تشخیص ٥ نوع احساس برای گویندگان زن با ویژگی های عروضی برای پایگاه داده درام.	٦۸
جدول(٤-٤): ماتریس تداخل برای تشخیص ٥ نوع احساس برای گویندگان زن با ویژگی های عروضی برای پایگاه داده برلین.	٦۹
جدول(٤-٥): ماتریس تداخل برای تشخیص ٥ نوع احساس برای گویندگان مرد با ویژگی های عروضی برای پایگاه داده درام.	٦۹
جدول(٤-٦): ماتریس تداخل برای تشخیص ٥ نوع احساس برای گویندگان مرد با ویژگی های عروضی برای پایگاه داده برلین.	٦۹
جدول(٤-۷): ماتریس تداخل برای تشخیص ٥ نوع احساس برای گویندگان زن با ویژگی های طیفی برای پایگاه داده درام.	۷۰
جدول(٤-۸): ماتریس تداخل برای تشخیص ٥ نوع احساس برای گویندگان زن با ویژگی های طیفی برای پایگاه داده برلین.	۷۰
جدول(٤-۹): ماتریس تداخل برای تشخیص ٥ نوع احساس برای گویندگان مرد با ویژگی های طیفی برای پایگاه داده درام.	۷۱
جدول(٤-۱۰): ماتریس تداخل برای تشخیص ٥ نوع احساس برای گویندگان مرد با ویژگی های طیفی برای پایگاه داده برلین.	۷۱
جدول(٤-۱۱): ماتریس تداخل برای تشخیص ٥ نوع احساس برای گویندگان زن با همه ی ویژگی ها برای پایگاه داده درام.	۷۲

جدول(٤-١٢): ماتریس تداخل برای تشخیص ۵ نوع احساس برای گویندگان زن با همه ی ویژگی ها برای پایگاه داده برلین. ۷۲
جدول(٤-١٣): ماتریس تداخل برای تشخیص ۵ نوع احساس برای گویندگان مرد با همه ی ویژگی ها برای پایگاه داده درام. ۷۲
جدول(٤-١٤): ماتریس تداخل برای تشخیص ۵ نوع احساس برای گویندگان مرد با همه ی ویژگی ها برای پایگاه داده برلین. ۷۳
جدول(٤-١٥): لیست ویژگی های استخراج شده در آزمایش دوم. ۷٤
جدول(٤-١٦): نرخ تشخیص ویژگی های مختلف برای پایگاه داده درام(پیشنهادی) و برلین در آزمایش دوم. ۷۵
جدول(٤-١٧): ماتریس تداخل برای تشخیص ۵ نوع احساس برای گویندگان زن با ویژگی های عروضی برای پایگاه داده درام. ۷۷
جدول(٤-١٨): ماتریس تداخل برای تشخیص ۵ نوع احساس برای گویندگان زن با ویژگی های عروضی برای پایگاه داده برلین. ۷۷
جدول(٤-١٩): ماتریس تداخل برای تشخیص ۵ نوع احساس برای گویندگان مرد با ویژگی های عروضی برای پایگاه داده درام. ۷۷
جدول(٤-٢٠): ماتریس تداخل برای تشخیص ۵ نوع احساس برای گویندگان مرد با ویژگی های عروضی برای پایگاه داده برلین. ۷۸
جدول(٤-٢١): ماتریس تداخل برای تشخیص ۵ نوع احساس برای گویندگان زن با ویژگی های طیفی برای پایگاه داده درام. ۷۸
جدول(٤-٢٢): ماتریس تداخل برای تشخیص ۵ نوع احساس برای گویندگان زن با ویژگی های طیفی برای پایگاه داده برلین. ۷۹
جدول(٤-٢٣): ماتریس تداخل برای تشخیص ۵ نوع احساس برای گویندگان مرد با ویژگی های طیفی برای پایگاه داده درام. ۷۹
جدول(٤-٢٤): ماتریس تداخل برای تشخیص ۵ نوع احساس برای گویندگان مرد با ویژگی های طیفی برای پایگاه داده برلین. ۷۹
جدول(٤-٢۵): ماتریس تداخل برای تشخیص ۵ نوع احساس برای گویندگان زن با ویژگی ۸۰

وریوگرام برای پایگاه داده درام.
جدول(٤-٢٦): ماتریس تداخل برای تشخیص ٥ نوع احساس برای گویندگان زن با ویژگی وریوگرام برای پایگاه داده برلین. ٨٠
جدول(٤-٢٧): ماتریس تداخل برای تشخیص ٥ نوع احساس برای گویندگان مرد با ویژگی وریوگرام برای پایگاه داده درام. ٨٠
جدول(٤-٢٨): ماتریس تداخل برای تشخیص ٥ نوع احساس برای گویندگان مرد با ویژگی وریوگرام برای پایگاه داده برلین. ٨١
جدول(٤-٢٩): ماتریس تداخل برای تشخیص ٥ نوع احساس برای گویندگان زن با ویژگی های عروضی و طیفی برای پایگاه داده درام. ٨١
جدول(٤-٣٠): ماتریس تداخل برای تشخیص ٥ نوع احساس برای گویندگان زن با ویژگی های عروضی و طیفی برای پایگاه داده برلین. ٨١
جدول(٤-٣١): ماتریس تداخل برای تشخیص ٥ نوع احساس برای گویندگان مرد با ویژگی های عروضی و طیفی برای پایگاه داده درام. ٨٢
جدول(٤-٣٢): ماتریس تداخل برای تشخیص ٥ نوع احساس برای گویندگان مرد با ویژگی های عروضی و طیفی برای پایگاه داده برلین. ٨٢
جدول(٤-٣٣): ماتریس تداخل برای تشخیص ٥ نوع احساس برای گویندگان زن با همه ی ویژگی ها برای پایگاه داده درام. ٨٢
جدول(٤-٣٤): ماتریس تداخل برای تشخیص ٥ نوع احساس برای گویندگان زن با همه ی ویژگی ها برای پایگاه داده برلین. ٨٣
جدول(٤-٣٥): ماتریس تداخل برای تشخیص ٥ نوع احساس برای گویندگان مرد با همه ی ویژگی ها برای پایگاه داده درام. ٨٣
جدول(٤-٣٦): ماتریس تداخل برای تشخیص ٥ نوع احساس برای گویندگان مرد با همه ی ویژگی ها برای پایگاه داده برلین. ٨٣
جدول(٤-٣٧): لیست ویژگی های استخراج شده در آزمایش سوم. ٨٤
جدول(٤-٣٨): بهترین نرخ تشخیص هر یک از ویژگی ها به همراه تعداد ویژگی برتر. ٨٨

فهرست نمودارها

عنوان	شماره صفحه
نمودار(۲-۱): هیستوگرام طول گفته در عرض چند ثانیه برای پایگاه داده برلین.	۳۵
نمودار(۲-۲): تعداد جملات پایگاه داده برلین در هر کلاس.	۳۶
نمودار(۲-۳): توزیع هندسی احساس ها در پایگاه داده VAM.	۳۷
نمودار(۳-۱): نسبت جملات با احساس های مختلف برای پایگاه داده درام، (الف) گویندگان زن و (ب) گویندگان مرد.	٤٥
نمودار(٤-۱): میانگین نرخ تشخیص برای ۵ احساس با استفاده از همه ی ویژگی ها برای پایگاه داده درام(پیشنهادی) و برلین در آزمایش اول.	٦٨
نمودار (٤-۲): میانگین نرخ تشخیص برای ۵ احساس با استفاده از همه ی ویژگی ها و ترکیب ویژگی های عروضی و طیفی برای پایگاه داده درام(پیشنهادی) و پایگاه داده برلین، (الف) گویندگان زن و(ب) گویندگان مرد در آزمایش دوم.	٧٦
نمودار(٤-۳): منحنی دقت برای پایگاه داده درام(پیشنهادی)، (الف) گویندگان زن و (ب) گویندگان مرد.	٨٦
نمودار(٤-٤): منحنی دقت برای پایگاه داده برلین، (الف) گویندگان زن و (ب) گویندگان مرد.	٨٧
نمودار(٤-۵): بهترین نرخ تشخیص هر یک از ویژگی ها، (الف) پایگاه داده درام (پیشنهادی) و (ب) پایگاه داده برلین.	٨٩

فهرست شکل ها

عنوان **شماره صفحه**

شکل(۱-۱): سیستم تشخیص احساس از روی گفتار. ٤

شکل(۱-۲): تأثیر احساس های مختلف بر حالت چهره. ٥

شکل(۱-۳): مدل سه بعدی احساس و (ب) مدل سه بعدی با مانکن ها. ۷

شکل(۱-٤): مدل دو بعدی احساس. ۸

شکل(۱-٥): الف) شماتیک اجزای تولید گفتار. ب) اثرات فیزیولوژیکی احساسات بر اعضای بدن و سیگنال گفتار. ۱۰

شکل(۲-۱): مدل عمومی سیستم تشخیص احساس از روی گفتار. ۱۲

شکل(۲-۲): تغییرات گام و انرژی گفتار هنگام ادای کلمه "ONE" با احساس های مختلف. ۱۳

شکل(۲-۳): موقعیت فرمنت ها برای دو حرف "i" و "e" و تصاویر اسپکتروگرام این دو سیگنال. ۱٥

شکل(۲-٤): بلوک دیاگرام محاسبه ضرایب مل کپستروم. ۱٦

شکل(۲-٥): بلوک دیاگرام استخراج ویژگی از PLP، RASTA PLP، GPLP. ۱۸

شکل(۲-٦): تأثیر استفاده توأم از ویژگی هایی که به تنهایی کاربردی ندارند در یک مسئله ی دو کلاسه با دو ویژگی هم بسته. ۲٤

شکل(۲-۷): الگوریتم ماشین بردار پشتیبان. ۲۹

شکل(۲-۸): تفکیک دو کلاس با اعمال یک صفحه در الگوریتم ماشین بردار پشتیبان. ۲۹

شکل(۲-۹): پروسه ی آموزش الگوریتم ماشین بردار پشتیبان. ۳۰

شکل(۲-۱۰): استفاده از کرنل در الگوریتم ماشین پشتیبان. ۳۱

شکل(۲-۱۱): تفسیر هندسی از امتیازات ممیزی. ۳۲

شکل(۲-۱۲): نرخ تشخیص هر احساس در پایگاه داده برلین توسط شنوندگان. ۳٦

شکل(۳-۱): فضای کار در نرم افزار Adobe Audition 3.0. ٤٤

شکل(۳-۲): بلوک دیاگرام سیستم تشخیص احساس. ٤٦

شکل(۳-۳): تصویر اسپکتروگرام و منحنی وریوگرام ۰ و ۹۰ درجه برای حالت عصبانیت ٤۹

(الف) گوینده زن و (ب)گوینده مرد از پایگاه داده برلین.
شکل(۳-٤): تصویر اسپکتروگرام و منحنی وریوگرام ۰ و ۹۰ درجه برای حالت ترس (الف) گوینده زن و (ب)گوینده مرد از پایگاه داده برلین. ۵۰
شکل(۳-۵): تصویر اسپکتروگرام و منحنی وریوگرام ۰ و ۹۰ درجه برای حالت خوشحالی (الف)گوینده زن و (ب)گوینده مرد از پایگاه داده برلین. ۵۱
شکل(۳-٦): تصویر اسپکتروگرام و منحنی وریوگرام ۰ و ۹۰ درجه برای حالت ناراحتی (الف) گوینده زن و (ب)گوینده مرد از پایگاه داده برلین. ۵۲
شکل(۳-۷): تصویر اسپکتروگرام و منحنی وریوگرام ۰ و ۹۰ درجه برای حالت طبیعی (الف) گوینده زن و (ب)گوینده مرد از پایگاه داده برلین. ۵۳
شکل(۳-۸): تصویر اسپکتروگرام و منحنی وریوگرام ۰ و ۹۰ درجه برای حالت عصبانیت (الف)گوینده زن و (ب)گوینده مرد از پایگاه داده درام ۵٤
شکل(۳-۹): تصویر اسپکتروگرام و منحنی وریوگرام ۰ و ۹۰ درجه برای حالت ترس (الف)گوینده زن و (ب)گوینده مرد از پایگاه داده درام. ۵۵
شکل(۳-۱۰): تصویر اسپکتروگرام و منحنی وریوگرام ۰ و ۹۰ درجه برای حالت خوشحالی (الف)گوینده زن و (ب)گوینده مرد از پایگاه داده درام. ۵٦
شکل(۳-۱۱): تصویر اسپکتروگرام و منحنی وریوگرام ۰ و ۹۰ درجه برای حالت ناراحتی (الف)گوینده زن و (ب)گوینده مرد از پایگاه داده درام. ۵۷
شکل(۳-۱۲): تصویر اسپکتروگرام و منحنی وریوگرام ۰ و ۹۰ درجه برای حالت طبیعی (الف)گوینده زن و (ب)گوینده مرد از پایگاه داده درام. ۵۸
شکل(۳-۱۳): بلوک دیاگرام استخراج ویژگی وریوگرام از سیگنال گفتار. ۵۹
شکل(۳-۱٤): (الف) یک نمونه از شبکه و (ب) نمونه وزن متناظر با آن . ٦۰
شکل(۳-۱۵): فرآیند اعمال یک شبکه به تبدیل فوریه دو بعدی یک تصویر. ٦۱

چکیده

تشخیص احساس از روی سیگنال گفتار یکی از موضوعات جدید و چالش برانگیز در پردازش گفتار می باشد که نقش مهمی در ارتباط انسان و کامپیوتر دارد. استفاده از یک پایگاه داده جامع در سیستم تشخیص احساس از اهمیت زیادی برخوردار است. تاکنون پایگاه های داده مختلفی به زبان های آلمانی، انگلیسی، دانمارکی و سایر زبان ها ارائه شده اند اما پایگاه داده فارسی جامعی تاکنون مشاهده نشده است از این رو در این تحقیق پایگاه داده فارسی احساسی درام برای تشخیص احساس از روی گفتار ارائه شده است. این پایگاه داده شامل ۷٤۸ جمله با ۸ احساس عصبانیت، خستگی، نفرت، ترس، طبیعی، ناراحتی، تعجب وخوشحالی می باشد. جملات توسط ۳۳ گوینده (۱۸ مرد و ۱۵ زن) بیان شده اند. به منظور ارزیابی و مقایسه ی پایگاه داده پیشنهادی و پایگاه داده معروف برلین، ویژگی های متداول عروضی و طیفی از جملات این دو پایگاه داده استخراج گردیدند. نوآوری های دیگر در این تحقیق، ارائه دو ویژگی زمان- فرکانسی جدید مبتنی بر وریوگرام و مل کپستروم دو بعدی به منظور تشخیص احساس از روی گفتار می باشد. درآزمایش های انجام شده، از الگوریتم انتخاب ویژگی مبتنی بر معیار *FDR* و الگوریتم *LDA* استفاده نموده ایم. به منظور کلاسه بندی احساس های مختلف ازالگوریتم های *LDA*، *SVM* خطی و *SVM* با کرنل *RBF* استفاده نموده ایم. بیشترین نرخ تشخیص از ترکیب ویژگی های وریوگرام (پیشنهادی) با ویژگی های عروضی و طیفی حاصل شده است که در پایگاه داده درام برای گویندگان مرد ٥٧/٣٤٪ و برای گویندگان زن ٦٣/١٨٪ و درپایگاه داده برلین برای گویندگان مرد ۹۰/٤۳٪ و برای گویندگان زن ۸٦/۸۲٪ محاسبه شد.

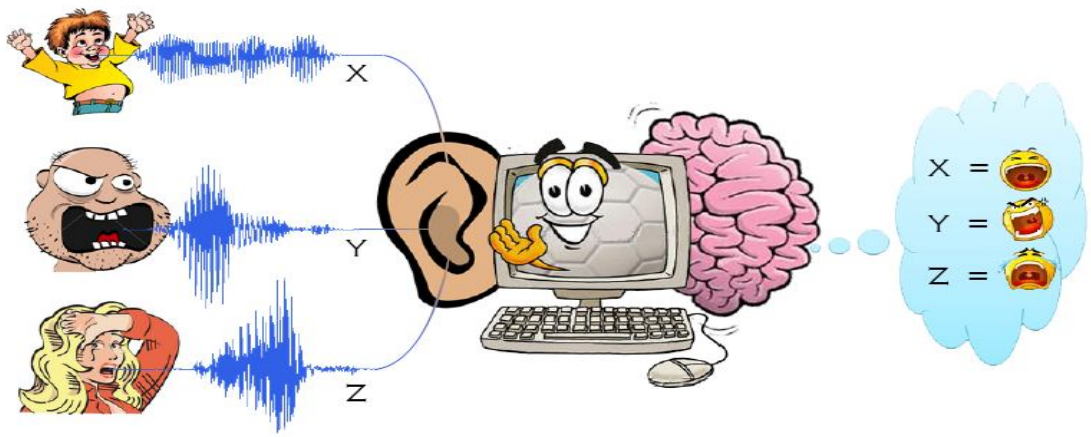
کلمات کلیدی- تشخیص احساس از روی گفتار، پایگاه داده ی درام، ویژگی های عروضی و طیفی، ویژگی های زمان- فرکانسی، وریوگرام، مل کپستروم دو بعدی.

فصل اول:

احساس و تأثیر آن برروی گفتار

۱-۱- مقدمه

سیگنال گفتار سریع ترین و طبیعی ترین روش ارتباط بین انسان ها می باشد. بر این اساس گفتار به عنوان یک روش سریع و کارآمد برای تعامل انسان و کامپیوتر بکار گرفته می شود. تا کنون تلاش های زیادی در زمینه تشخیص گفتار انجام شده است. با وجود پیشرفت های زیاد در این زمینه، فاصله ی زیادی بین تعامل طبیعی انسان و کامپیوتر وجود دارد. دلیل اصلی این موضوع عدم توانایی کامپیوتر در درک احساس کاربر می باشد. انسان به طور طبیعی، با استفاده از حواس پنجگانه تلاش خود را برای درک حداکثر پیام دریافتی می کند. انسان ها صدا را می شنوند و قادر به لب خوانی هستند. آنها قادر به درک طنز و طعنه موجود در گفتار می باشند و در مقابل واکنش مناسب را نشان می دهند. تعامل انسان و کامپیوتر در حال تکامل است. تلاش های بسیاری برای درک احساس انسان توسط کامپیوتر انجام شده است. هر چند که تشخیص احساس به عنوان یک امر طبیعی برای انسان می باشد، این امر برای کامپیوتر بسیار چالش برانگیز است. هدف نهایی سیستم های تشخیص احساس، به کار بردن احساسات به منظور ایجاد یک ارتباط بهتر بین انسان و کامپیوتر و احساس رضایت در کاربر می باشد. شکل (۱-۱) ساختار یک سیستم تشخیص احساس از روی گفتار را نشان می دهد. ورودی این سیستم سیگنال گفتار و خروجی آن احساس کاربر می باشد.



شکل(۱-۱): سیستم تشخیص احساس از روی گفتار (Chiriacescu 2009).