بسم الله الرحمن الرحيم

دانشگاه یزد
دانشکده مهندسی برق و کامپیوتر
گروه مهندسی کامپیوتر

پایان‌نامه
جهت دریافت درجه کارشناسی ارشد
مهندسی کامپیوتر - هوش مصنوعی

تشخیص زبان اشاره فارسی با استفاده از کینکت

استاد راهنما:
دکتر مهدی رضائیان

استاد مشاور:
دکتر ولی درهمی

پژوهش‌گر:
عباس دهقان حسام پور

اسفند ۱۳۹۲

کلیه‌ی حقوق مادی و معنوی مترتب بر نتایج مطالعات، ابتکارات و نوآوری‌های ناشی از تحقیق موضوع این پایان‌نامه متعلق به دانشگاه یزد است و هرگونه استفاده از نتایج علمی و عملی از این پایان‌نامه برای تولید دانش فنی، ثبت اختراع، ثبت اثر بدیع هنری، همچنین چاپ و تکثیر، نسخه‌برداری، ترجمه و اقتباس و ارائه مقاله در سمینارها و مجلات علمی از این پایان‌نامه منوط به موافقت کتبی دانشگاه یزد است.

تقدیم به

روح آرام و بلند ابرمرد تاریخ معاصر ایران و انقلاب، **خمینی کبیر** (قدس) که با انفاس مسیحایی‌اش صفحه‌ی جدیدی از تاریخ کشورم را فراروی من و هم‌وطنانم گشود تا ما توفیق تنفس در فضای جدیدی که معنوی و اسلامی است را داشته باشیم. به **رهبر عزیزم** که علی‌وار پرچم آن سفرکرده را برافراشته نگه‌داشته است. به **شهدای زنده‌ای** که با بذل جانشان طعم سرافرازی و عزت را به ما چشاندند. به **امام زمان (عج)** ، که بی‌صبرانه منتظر قدوم مبارکش هستیم و به **پدر و مادر عزیزم** که همواره با حمایت‌های خود من را یاری نموده‌اند.

سپاس‌گزاری

سپاس خداوند یکتای عزتمندی که رحمت و دانش او در سراسر گیتی گسترده شده، آسمان‌ها و زمین همه از آن اوست و علم و دانش حقیقی را بر هر که بخواهد موهبت می‌فرماید. رحمت و لطف او را بی‌نهایت سپاس می‌گویم چرا که فهم و درک مطالب این پژوهش را بر من ارزانی داشت و مرا به این اصل رساند که علم و ایمان دو بال یک پروازند. توفیق تلاش به من داد و هر بار که خطا کردم فرصتی دوباره، تا با امید، تلاشی تازه را آغاز کنم و به خواست او به نتیجه‌ی مطلوب نائل آیم. به‌راستی که همه چیز از آن اوست و همه‌چیز به خواست اوست.

برخود لازم می دانم که از استاد عزیز و گرامی جناب آقای دکتر رضائیان که همواره با صبر و بردباری پاسخگوی اینجانب بوده‌اند، کمال تشکر و سپاس‌گزاری را داشته باشم. همچنین از زحمات آقای دکتر درهمی و دیگر اساتید بخش مهندسی کامپیوتر تقدیر و تشکر می‌نمایم. از همکاری مسئولان آموزش پرورش استثنایی شیراز آقایان دانشور، لطفی، کریمی و همچنین آقای عباس باباجعفری معاون دبیرستان ناشنوایان بهاران شیراز، آقای همایونی مدیر مجتمع آموزشی خاتم النبیا یزد که در گردآوری داده‌های مورد نیاز این پژوهش همکاری کردند تشکر می‌نمایم.

چکیده

ناشنوایان برای ارتباط با دیگران از زبان اشاره استفاده می‌کنند. فهم این زبان برای دیگران امری مشکل است و معمولاً نیاز به مترجم دارند. هدف این تحقیق طراحی و ساخت مترجم زبان اشاره فارسی است. داده عمق کینکت به نور محیط وابسته نیست، بنابراین ویژگی‌های استخراج شده از آن در محیط‌های مختلف تغییر ناپذیر می‌باشد. برای به دست آوردن یک سیستم پایدار در این تحقیق تنها از داده‌های عمق برای استخراج ویژگی استفاده گردیده است. در این پژوهش با استفاده از کینکت بیست و شش علامت از الفبای زبان اشاره فارسی تشخیص داده شده است. ابتدا با استفاده از داده‌های کینکت اسکلت بدن به دست آمده است. سپس زاویه بازوی دست راست نسبت به خط افق محاسبه شده و بر اساس این زاویه الفبای زبان اشاره به سه گروه تقسیم گردیده‌اند. آن‌گاه برای هر گروه یک پنجره مناسب که بتواند دست را از پس‌زمینه جدا کند در نظر گرفته شده و دست از سایر قسمت‌های بدن و پس زمینه جدا گردیده است. سپس در قسمت جدا شده نزدیک‌ترین نقطه به دوربین را محاسبه کرده، و نقاطی که در فاصله دورتر از یک مقدار مشخص از آن نقطه قرار دارند، حذف گردیده‌اند. مقدار آستانه برای حذف عمق برای هر دسته به صورت متفاوت تعیین شده است تا بهترین جداسازی انجام گیرد. سپس برای حذف اغتشاش، بزرگ‌ترین قطعه متصل در تصویر پیدا شده و مابقی نقاط حذف گردیده‌اند. پس از جدا سازی دست از پس زمینه، فیلتر گابور در ۶ جهت و ۵ مقیاس بر روی تصویر دست اعمال شده است. بدین ترتیب ۳۰ تصویر به دست می‌آید که انرژی و میانگین هر کدام محاسبه می‌گردد. پس از این مرحله الگوریتم کاهش بعد LDA را بر روی این ویژگی‌ها اعمال می‌کنیم. برای دسته بندی از ماشین‌های بردار پشتیبان و دسته بند بیز استفاده شده است. برای آزمایش روش پیشنهادی پایگاه داده‌ای به کمک ۲۰ نفر تهیه شده است. دقت به دست آمده برای هر گروه بیش از ۹۱ درصد می‌باشد. برای مقایسه بهتر دقت روش پیشنهادی برای استخراج ویژگی، این روش بر روی پایگاه داده‌ای از زبان اشاره آمریکایی آزمایش گردیده است. دقت به دست آمده بر روی این پایگاه داده بیش از ۸۲ درصد بوده است که از نتایجی که تاکنون بر روی این پایگاه داده گزارش شده، بهتر می‌باشد.

کلمات کلیدی: کینکت، داده عمق، تشخیص ژست، تشخیص زبان اشاره فارسی

فهرست مطالب

۱ مقدمه **۱**

۱.۱ پیش گفتار ۲

۲.۱ زبان اشاره ۲

۱.۲.۱ روش‌های اشاره دستی ۳

۲.۲.۱ الفبای گویای باغچه بان ۴

۲ آشنایی با کینکت **۷**

۱.۲ مقدمه ۸

۲.۲ کینکت چیست؟ ۸

۳.۲ اجزای کینکت ۹

۱.۳.۲ میکروفن‌های کینکت ۱۰

۲.۳.۲ حس‌گر عمق ۱۰

۴.۲ نحوه استخراج اطلاعات عمق در کینکت ۱۱

۵.۲ راه‌اندازهای کینکت ۱۳

۶.۲ تشخیص و دنبال کردن اسکلت بدن با استفاده از داده عمق ۱۳

۷.۲ تشخیص صورت در کینکت ۱۴

۸.۲ محدودیت‌های کینکت ۱۵

۹.۲ موارد استفاده از کینکت ۱۷

۳ مروری بر کارهای انجام شده **۱۹**

۱.۳ مقدمه ۲۰
۲.۳ سنسورهای پوشیدنی ۲۱
۳.۳ دوربین‌های رنگی ۲۲
۱.۳.۳ تشخیص مکان دست ۲۳
۲.۳.۳ تشخیص مکان دست به کمک ویژگی‌های شکلی دست ۳۱
۳.۳.۳ استخراج ویژگی در دوربین‌های رنگی ۳۱
۴.۳ دوربین‌های عمق ۳۴
۱.۴.۳ جداسازی دست از پس زمینه در دوربین‌های عمق ۳۶
۵.۳ مطالعات انجام شده با استفاده از کینکت ۳۸

۴ روش پیشنهادی جهت تشخیص زبان اشاره فارسی ۵۳

۱.۴ جمع آوری داده ۵۴
۲.۴ دسته بندی اشارات بر اساس مکان دست ۵۵
۳.۴ جدا کردن دست از سایر قسمت‌ها ۵۶
۴.۴ استخراج ویژگی ۶۱
۵.۴ کاهش ابعاد ۶۱
۶.۴ بررسی و تحلیل نتایج ۶۱
۷.۴ بررسی نتایج مربوط به دسته اول ۶۳
۸.۴ بررسی نتایج مربوط به دسته دوم ۶۳
۹.۴ بررسی نتایج مربوط به دسته سوم ۶۴
۱۰.۴آزمایش روش پیشنهادی بر روی زبان اشاره آمریکایی ۶۴

۵ نتیجه گیری و کارهای آینده ۶۷

۱.۵ نتیجه گیری ۶۸
۲.۵ کارهای آینده ۶۸
۱.۲.۵ اضافه کردن ویژگی شکل لب ۶۹

۲.۲.۵ تشخیص کلمات و جملات با استفاده از هجی انگشتی ۶۹
۳.۲.۵ تشخیص کلمات و حروف پویا ۶۹

آ تاریخچه **۷۱**

آ.۴.۰ مسابقه برای هک کردن کینکت ۷۲

ب ماشین‌های بردار پشتیبان **۷۷**

ب.۵.۰ تئوری ماشین‌های بردار پشتیبان ۷۷
ب.۶.۰ ماشین بردار پشتیبان چند کلاسی ۷۸
ب.۷.۰ ماشین‌های بردار پشتیبان غیرخطی ۷۹

مراجع **۷۹**

فهرست شکل‌ها

شکل ۱.۱: نمونه ای از علائم پویا در زبان اشاره فارسی [۹۴] ۳

شکل ۲.۱: نمونه ای از علائم ایستا در زبان اشاره فارسی [۹۴] ۳

شکل ۳.۱: جبار باغچه‌بان ۴

شکل ۱.۲: وسایل مشابه کینکت مایکروسافت ۸

شکل ۲.۲: نمونه اسکلت تشخیص داده شده روی بدن انسان ۹

شکل ۳.۲: ارتباطات داخلی کینکت [۲۳] ۱۰

شکل ۴.۲: تصویر بالا: تصویر رنگی یک مبل - تصویر پایین: تصویر همان مبل به صورتی که دوربین مادون قرمز آن را می‌بیند. [۳۳] ۱۱

شکل ۵.۲: گوشه‌ای از مبل شکل ۴.۲ ، الگوی نقطه‌ای مبل [۳۳] ۱۲

شکل ۶.۲: نحوه تشخیص عمق توسط کینکت [۳۹] ۱۲

شکل ۷.۲: تقسیم تصویر عمق بدن به قسمت های مختلف و تشخیص اسکلت بدن [۷۱] ۱۴

شکل ۸.۲: اتصالات تشخیص داده شده توسط Microsoft SDK 1.6 . . . . . ۱۴

شکل ۹.۲: شناسایی اتصالات برای حالت ایستاده و نشسته با استفاده از کینکت ۱۵

شکل ۱۰.۲: نقاط صورت تشخیص داده شده در Microsoft SDK 1.6 . . . . . ۱۶

شکل ۱۱.۲: مدل سه بعدی ایجاد شده با Kinect Fusion (اتاق دکتر رضائیان) ۱۷

شکل ۱.۳: دستکش داده‌ها ۲۱

شکل ۲.۳: دستکش طراحی شده در [۵۸] ۲۲

شکل ۳.۳: مراحل یک سیستم تشخیص ژست [۷۴] ۲۳

شکل ۴.۳: مکعب رنگی RGB [۹۶] ۲۵

شکل ۵.۳: تبدیل فضای رنگی RGB به فضای رنگی YCbCr ۲۵
شکل ۶.۳: تشخیص دست و حذف پس زمینه در [۱۸] ۲۷
شکل ۷.۳: حذف پس زمینه در فضای YCbCr ۲۹
شکل ۸.۳: حذف پس زمینه در فضای HSV ۲۹
شکل ۹.۳: پیدا کردن قسمت‌هایی که رنگ پوست دارند (دو دست و صورت) ۳۰
شکل ۱۰.۳: از چپ: ۱- دایره‌ها و بیضوی‌هایی که در تصویر پیدا شده‌اند. ۲- ویژگی‌های انتخاب شده بر اساس کف دست ، انگشتان و نوک انگشتان ۳- ترکیبی از هسته‌های گاوسی بر اساس حباب‌ها [۱۳] ۳۲
شکل ۱۱.۳: بانک فیلتر تبدیل موجک ۳۳
شکل ۱۲.۳: تبدیل موجک دو بعدی [۳۸] ۳۳
شکل ۱۳.۳: اعمال بانک فیلتر بر روی تصویر دست [۳۸] ۳۴
شکل ۱۴.۳: پیش پردازش تصویر دست در [۳۸] ۳۴
شکل ۱۵.۳: بانک فیتر تبدیل موجک اعمال شده بر روی تصویر در [۳۸] ۳۵
شکل ۱۶.۳: نمونه ای از دوربین‌های TOF (swisranger) ۳۶
شکل ۱۷.۳: جداسازی دست به کمک یگ مچبند رنگی ۳۷
شکل ۱۸.۳: علائم تشخیص داده شده در [۴۷، ۴۸] ۳۸
شکل ۱۹.۳: پوسته محدب نشان حرف آ ۳۹
شکل ۲۰.۳: نقاط پیدا شده حاصل از تلاقی کانتور و پوسته محدب ۳۹
شکل ۲۱.۳: نقاط باقی مانده پس از خوشه بندی ۴۰
شکل ۲۲.۳: نقاط باقی مانده پس از اعمال روش تطبیق سه گوشه ۴۱
شکل ۲۳.۳: شناسایی نام انگشتان در [۴۸، ۴۷] ۴۱
شکل ۲۴.۳: تکنیک تطبیق گوشه ۴۲
شکل ۲۵.۳: مراحل تشخیص زبان اشاره آلمانی در [۴] ۴۲
شکل ۲۶.۳: کانتورهای خطی متعامد ۴۳

شکل ۲۷.۳: علائم تشخیص داده شده در [۲، ۱۰] ۴۴
شکل ۲۸.۳: نتایج به دست آمده در [۲، ۱۰] برای دو هسته خطی و گاوسی در ماشین‌های
بردار پشتیبان ۴۵
شکل ۲۹.۳: نمایش گرافیکی فیلتر گابور در ۶ جهت و چهار مقیاس [۹۷] ۴۷
شکل ۳۰.۳: مراحل تشخیص زبان اشاره در [۶۲] ۴۷
شکل ۳۱.۳: نمایش ویژگی فشردگی برای سه شکل متفاوت [۹۵] ۵۱
شکل ۳۲.۳: پوسته محدب برای حرف y (دایره‌های توخالی قزمر: چاله ها، دایره‌های زرد
رنگ: نقاط پایانی) [۹۵] ۵۱
شکل ۳۳.۳: نتایج به دست آمده شده در [۹۵] برای ۱ تا ۵ نفر ۵۲

شکل ۱.۴: برنامه استفاده شده جهت جمع آوری داده ۵۵
شکل ۲.۴: تقسیم بندی زبان اشاره فارسی بر اساس مکان دست ۵۵
شکل ۳.۴: زاویه‌های ایجاد شده بین خط افق و بازو ۵۶
شکل ۴.۴: گروه اول: زاویه دست نسبت به خط افق ۹۰ درجه می‌باشد ۵۶
شکل ۵.۴: گروه دوم: زاویه دست نسبت به خط افق ۱۸۰ درجه می‌باشد ۵۷
شکل ۶.۴: گروه سوم: زاویه دست نسبت به خط افق ۱۳۵ درجه می‌باشد ۵۸
شکل ۷.۴: پنجره انتخاب شده برای گروه اول ۵۹
شکل ۸.۴: پنجره انتخاب شده برای گروه دوم ۵۹
شکل ۹.۴: پنجره انتخاب شده برای گروه سوم ۶۰
شکل ۱۰.۴: اعمال فیلتر گابور در ۶ جهت و ۵ مقیاس برای حرف آ ۶۲
شکل ۱۱.۴: تصاویر موجود در پایگاه داده [۶۳] ۶۵

شکل آ.۱: جدا سازی نقاط مربوط به بدن فرد از پس زمینه و قسمت بندی آن ۷۲
شکل آ.۲: اطلاعات اسکلت بدن انسان تشخیص داده شده توسط اطلاعات کینکت . . . ۷۳

شکل ب.۱: ماشین‌های بردار پشتیبان ۷۹