بسم الله الرحمن الرحيم

دانشگاه کاشان

دانشکده مهندسی

گروه مهندسی کامپیوتر

پایان نامه

جهت اخذ درجه کارشناسی ارشد

در رشته مهندسی کامپیوتر

عنوان:

تشخیص برون خط حروف مقطع دستنویس فارسی

استاد راهنما:

دکتر حسین ابراهیم پور کومله

استاد مشاور:

دکتر سید مرتضی بابامیر

به وسیله:

میلاد اسدی

بهمن ماه ۱۳۹۰

تقدیم به:

پدر و ماردم که در تمام مراحل زندگی راهنمای من بوده‌اند.

تشکر و قدر دانی

حمد و سپاس خدای را که توفیق کسب دانش و معرفت را به ما عطا فرمود. در این‌جا بر خود لازم می‌دانم از تمامی اساتید بزرگوار به ویژه اساتید دوره کارشناسی ارشد که در طول سالیان گذشته مرا در تحصیل علم و معرفت و فضایل اخلاقی یاری نموده‌اند تقدیر و تشکر نمایم.

از استاد گرامی و بزرگوار جناب آقای دکتر ابراهیم‌پور که راهنمایی اینجانب را در انجام تحقیق، پژوهش و نگارش این پایان‌نامه تقبل نموده‌اند نهایت تشکر و سپاسگذاری را دارم.

همچنین از تشریک مساعی آقای دکتر فرجی بعنوان داور داخل دانشگاه و آقای دکتر منجمی بعنوان استاد داور مدعو خارج از دانشگاه که این پایان‌نامه را مورد مطالعه قرار داده و در جلسه دفاعیه شرکت نموده‌اند تشکر می‌نمایم.

در پایان از ناظر محترم تحصیلات تکمیلی دانشگاه، جناب آقای دکتر زنگنه و استاد مشاور محترم، جناب آقای دکتر بابامیر سپاسگذاری می‌نمایم.

چکیده

هدف اصلی این پایان نامه ایجاد سیستمی کاربردی برای خواندن و شناسایی حروف دستنویس فارسی توسط رایانه است. بدیهی است در چنین سیستمی بهبود صحت شناسایی بسیار ضروری می‌باشد. بنابراین در این مستند ابتدا به بررسی روش‌های مختلف پرداخته و صحت تشخیص آن‌ها مورد بررسی قرار گرفته است. سپس جهت افزایش صحت به ترکیب این روش‌ها پرداخته شده است.

قسمت اصلی این پایان نامه به بررسی روش‌های مختلف استخراج ویژگی برای مسئله شناسایی حروف می‌پردازد. در هر روش پارامترها و پیش‌پردازش‌های مختلف و همچنین سه کلاسه‌بند شبکه عصبی، نزدیک‌ترین همسایگی و ماشین بردار پشتیبانی مورد استفاده قرار گرفته و مقایسه می شوند و در پایان بهترین حالت هر روش انتخاب می‌شود. پس از این مرحله با استفاده از روش‌های رای گیری به ترکیب سیستم‌های موجود پرداخته می‌شود. جهت سرعت بخشیدن به سیستم مورد نظر در هر روش، بخش استخراج ویژگی به صورت موازی و با استفاده از توان پردازنده‌های چند هسته‌ای پیاده‌سازی شده است.

برای ارزیابی این روش‌ها از بزرگترین پایگاه داده حروف مقطع فارسی، حدف (حروف دستنویس فارسی) استفاده شد. نتایج تجربی نشان می‌دهد که بالاترین درصد بازشناسی پیش از ترکیب متعلق به عملگر کرش (Kirsch) با صحت ۹۷.۴۴٪ و پس از ترکیب این روش‌ها با استفاده از رای گیری ۹۷.۷۳٪ بوده است.

کلمات کلیدی: شناسایی حروف دستنویس، استخراج ویژگی، کلاسه‌بندی، شبکه عصبی، نزدیک‌ترین همسایگی و ماشین بردار پشتیبانی

فهرست مطالب

فصل اول: دیباچه

۱-۱- مقدمه ۱

۱-۲- کاربردها ۲

۱-۳- انگیزش ۳

۱-۴- نوشتار فارسی ۴

۱-۵- اهداف تحقیق ۶

۱-۶- چشم انداز پایان نامه ۷

فصل دوم: مروری بر پژوهش‌های مرتبط

۲-۱- مقدمه ۸

۲-۲- تاریخچه ۸

۲-۳- پایگاه داده‌های موجود ۹

۲-۴- پژوهش‌های انجام شده در زبان‌های فارسی و عربی ۱۱

۲-۵- رقابت‌ها ۱۶

فصل سوم: روش تشخیص

۳-۱- مقدمه ۱۸

۳-۲- پیش پردازش ۱۹

۳-۲-۱- باینری کردن ۱۹

۳-۲-۲- حذف موارد مشابه پس زمینه ۱۹

۳-۲-۳- مرکز آوردن ۲۰

۳-۲-۴- مرکز آوردن با مرکز ثقل ۲۱

۳-۲-۵- تصحیح کجی ۲۱

۳-۲-۶- نرمال‌سازی سایز ۲۲

۳-۳- نحوه نمایش ۲۳

۳-۴- بخشبندی ۲۴

۳-۵- استخراج ویژگی ۲۵

۳-۵-۱- ویژگی‌های ساختاری ۲۶

۳-۵-۲- ویژگی‌های آماری ۲۶
۳-۶- کلاسه‌بندی ۴۰
۳-۶-۱- نزدیکترین همسایگی (KNN) ۴۰
۳-۶-۲- شبکه های عصبی (NN) ۴۱
۳-۶-۳- ماشین بردار پشتیبانی (SVM) ۴۲
۳-۷- روش‌های رای گیری ۴۵

فصل چهارم:پیاده‌سازی

۴-۱- مقدمه ۴۹
۴-۲- موازی‌سازی در Matlab ۵۱
۴-۳- سیستم‌های غیر ترکیبی بر مبنای ویژگی‌های مختلف ۵۳
۴-۳-۱- ویژگی بلوک‌بندی ۵۳
۴-۳-۲- ویژگی تعداد خطوط تقاطع ۵۶
۴-۳-۳- ویژگی DPP ۶۰
۴-۳-۴- ویژگی مکان مشخصه ۶۳
۴-۳-۵- ویژگی گرادیان روشنایی روبرتز ۶۷
۴-۳-۶- ویژگی گرادیان روشنایی سوبل ۷۱
۴-۳-۷- ویژگی گرادیان روشنایی کرش ۷۵
۴-۳-۸- ویژگی گرادیان روشنایی کنی ۷۹
۴-۳-۹- ویژگی توصیف‌گر SIFT ۸۲
۴-۳-۱۰- ویژگی تبدیل فوریه ۸۴
۴-۳-۱۱- ویژگی تبدیل کسینوسی ۸۸
۴-۳-۱۲ ویژگی هیستوگرام کد زنجیر ۹۰
۴-۴- ترکیب روش‌ها ۹۲

فصل پنجم: نتیجه‌گیری و پیشنهادات

۵-۱- جمع بندی روش‌های استخراج ویژگی ۹۶
۵-۲- مقایسه ۹۸
۵-۳- نتیجه ۱۰۰
منابع ۱۰۱

فهرست شکل‌ها

شکل ۱-۱حروف پیوندی ۵
شکل ۲-۱ نمونه فرم جمع آوری اطلاعات ۱۰
شکل ۲-۲ نتایج تشخیص در مسابقات ICDAR فارسی ۱۷
شکل ۳-۱ روال تشخیص در یک سیستم OCR ۱۸
شکل ۳-۲ حذف پس زمینه ۲۰
شکل ۳-۳ تبدیل حذف پس‌زمینه ۲۰
شکل ۳-۴ متمرکز کردن ۲۱
شکل ۳-۵ نمونه تصاویر پس از اصلاح کجی ۲۲
شکل ۳-۶ نرمالسازی سایز ۲۲
شکل ۳-۷ نمونه کد زنجیر ۲۳
شکل ۳-۸ نحوه نمایش ۲٤
شکل ۳-۹ نمونه ابهامات اسکلت ۲٤
شکل ۳-۱۰ نمایش حصار ۲٤
شکل ۳-۱۱ ترتیب نامناسب حروف ۲۵
شکل ۳-۱۲ ویژگی‌های ساختاری ۲٦
شکل ۳-۱۳ بلوکبندی حرف شین ۲۷
شکل ۳-۱۴ ویژگی خطوط تقاطع ۲۷
شکل ۳-۱۵ ویژگی مکان مشخصه ۲۷
شکل ۳-۱۶ گشتاور مرکزی پس از انتقال و جابجایی ۲۹
شکل ۳-۱۷ گشتاورهای ثابت ۳۰
شکل ۳-۱۸ نمونه پراجکشن پروفایل ۳۱
شکل ۳-۱۹ نمونه پروفایل مرزی ۳۲
شکل ۳-۲۰ فیلتر سوبل ۳۳
شکل ۳-۲۱ فیلتر روبرتز ۳۳
شکل ۳-۲۲ فیلترهای کرش ۳۳
شکل ۳-۲۳ تصویر گرادیان ۳٤
شکل ۳-۲۴ هشت جهت اصلی ۳۵
شکل ۳-۲۵ تجزیه بردار گرادیان ۳۵
شکل ۳-۲۶ نحوه به دست آوردن ویژگی توصیف‌گر SIFT ۳٦

شکل ۳-۲۷ نمونه‌هایی از بازسازی تصویر پس از تبدیل کسینوسی ۳۹
شکل ۳-۲۸ کلاسه‌بند نزدیک‌ترین همسایگی ۴۰
شکل ۳-۲۹ نمونه شبکه عصبی چند لایه ۴۱
شکل ۳-۳۰ مثالی برای کلاسه‌بندی دو کلاس تفکیک پذیر خطی ۴۲
شکل ۳-۳۱ مسئله کلاسه‌بندی XOR ۴۳
شکل ۳-۳۲ نمونه‌ای از ماشین بردار پشتیبانی غیر خطی با استفاده از تابع کرنل RBF ۴۴
شکل ۴-۱ استفاده از حلقه‌های موازی در MATLAB ۵۲
شکل ۵-۱ نمونه‌های بد ۱۰۰

فهرست جدول‌ها

جدول ۱-۱ حالت‌های مختلف حروف ٤
جدول ۲-۱ نتایج مسابقات ICDAR 2009 ..... ١٦
جدول ۴-۱ نحوه توزیع حروف در بخش‌های پایگاه داده ٥٠
جدول ۴-۲ آزمون پیش‌پردازش‌های مختلف بر روی ویژگی بلوک‌بندی ٥٣
جدول ۴-۳ آزمون تغییرات سایز پنجره برای ویژگی بلوک‌بندی ٥٣
جدول ۴-۴ آزمون تعداد مختلف همسایگی KNN برای ویژگی بلوک‌بندی ٥٤
جدول ۴-۵ آزمون تعداد مختلف نرون‌های لایه میانی MLP برای ویژگی بلوک‌بندی ٥٤
جدول ۴-۶ ماتریس کارایی ویژگی بلوک‌بندی ٥٥
جدول ۴-۷ آزمون تسریع در ویژگی بلوک‌بندی ٥٦
جدول ۴-۸ آزمایش پیش‌پردازش‌های مختلف برای ویژگی تعداد خطوط تقاطع ٥٦
جدول ۴-۹ آزمون فاصله‌های مختلف برای ویژگی خطوط تقاطع ٥٧
جدول ۴-۱۰ آزمون تعداد مختلف همسایگی KNN ٥٧
جدول ۴-۱۱ آزمون مقادیر مختلف برای تعداد نرون‌های میانی در ویژگی تعداد خطوط تقاطع ٥٨
جدول ۴-۱۲ آزمون تسریع ویژگی تعداد خطوط تقاطع ٥٨
جدول ۴-۱۳ ماتریس کارایی برای ویژگی تعداد خطوط تقاطع ٥٩
جدول ۴-۱۴ آزمون پیش‌پردازش‌های مختلف روی ویژگی DPP ٦٠
جدول ۴-۱۵ آزمون تعداد همسایگی مختلف KNN در ویژگی DPP ٦٠
جدول ۴-۱۶ آزمون تعداد مختلف نرون لایه میانی برای ویژگی DPP ٦١
جدول ۴-۱۷ آزمون تسریع روی ویژگی DPP ٦١
جدول ۴-۱۸ ماتریس کارایی برای ویژگی DPP ٦٢
جدول ۴-۱۹ نتایج اولیه به دست آمده از ویژگی مکان مشخصه با پیش پردازش‌های مختلف ٦٣
جدول ۴-۲۰ آزمون حاشیه‌های مختلف برای ویژگی مکان مشخصه ٦٤
جدول ۴-۲۱ آزمون تعداد همسایگی‌های مختلف KNN در ویژگی مکان مشخصه ٦٤
جدول ۴-۲۲ آزمون تعداد محتلف نرون میانی برای ویژگی مکان مشخصه ٦٤
جدول ۴-۲۳ ماتریس کارآیی برای ویژگی مکان مشخصه ٦٥
جدول ۴-۲۴ تسریع ویژگی مکان مشخصه ٦٦
جدول ۴-۲۵ آزمون پیش‌پردازش‌های مختلف روی ویژگی گرادیان روشنایی روبرتز ٦٧
جدول ۴-۲۶ آزمون سایزهای مختلف پنجره برای ویژگی هیستوگرام روشنایی روبرتز ٦٨
جدول ۴-۲۷ آزمون تعداد مختلف همسایگی کلاسه‌بند KNN برای ویژگی گرادیان روشنایی روبرتز ٦٨
جدول ۴-۲۸ آزمایش تعداد نرون‌های میانی مختلف برای ویژگی گرادیان روشنایی روبرتز ٦٩

جدول ۴-۲۹ آزمون تسریع برای ویژگی گرادیان روشنایی روبرتز ٦٩
جدول ۴-۳۰ ماتریس کارایی برای ویژگی گرادیان روشنایی روبرتز ۷۰
جدول ۴-۳۱ آزمایش پیش‌پردازش‌های مختلف روی ویژگی گرادیان روشنایی سوبل ۷۱
جدول ۴-۳۲ آزمایش سایزهای مختلف پنجره برای ویژگی گرادیان روشنایی سوبل ۷۲
جدول ۴-۳۳ آزمون تعداد همسایگی‌های مختلف KNN برای ویژگی گرادیان سوبل ۷۲
جدول ۴-۳۴ آزمایش تعداد لایه‌های میانی مختلف MLP در ویژگی گرادیان روشنایی سوبل ۷۳
جدول ۴-۳۵ آزمون تسریع برای ویژگی هیستوگرام روشنایی سوبل ۷۳
جدول ۴-۳۶ ماتریس کارایی ویژگی گرادیان سوبل ۷٤
جدول ۴-۳۷ آزمایش پیش‌پردازش‌های مختلف بر روی ویژگی گرادیان روشنایی کرش ۷۵
جدول ۴-۳۸ آزمون سایزهای مختلف پنجره در ویژگی کرش ۷٦
جدول ۴-۳۹ آزمون تعداد مختلف لایه میانی برای MLP در ویژگی کرش ۷٦
جدول ۴-۴۰ ماتریس کارایی برای ویژگی کرش ۷۷
جدول ۴-۴۱ آزمون تعداد مختلف لایه میانی در ویژگی کرش ۷۸
جدول ۴-۴۲ تسریع ویژگی کرش ۷۸
جدول ۴-۴۳ آزمایش پیپش‌پردازش‌های مختلف در ویژگی گرادیان روشنایی کنی ۷۹
جدول ۴-۴۴ آزمون سایزهای مختلف پنجره نمونه‌برداری در ویژگی گرادیان روشنایی کنی ۷۹
جدول ۴-۴۵ آزمون تعداد همسایگی‌های مختلف برای کلاسه‌بند KNN در ویژگی گرادیان روشنایی کنی ۸۰
جدول ۴-۴۶ تغییرات تعداد نرون‌های لایه میانی در ویژگی گرادیان روشنایی کنی ۸۰
جدول ۴-۴۷ آزمون میزان تسریع در ویژگی گرادیان روشنایی کنی ۸۰
جدول ۴-۴۸ ماتریس کارایی ویژگی گرادیان روشنایی کنی ۸۱
جدول ۴-۴۹ آزمون تغییرات سایز پنجره در ویژگی توصیف‌گر SIFT ۸۲
جدول ۴-۵۰ آزمون تعداد مختلف همسایگی KNN در ویژگی توصیف‌گر SIFT ۸۲
جدول ۴-۵۱ ماتریس کارایی برای ویژگی توصیف‌گر SIFT ۸۳
جدول ۴-۵۲ آزمون میزان تسریع برای ویژگی توصیف‌گر SIFT ۸٤
جدول ۴-۵۳ آزمون پیش‌پردازش‌های مختلف برای ویژگی تبدیل فوریه ۸٤
جدول ۴-۵۴ آزمون اندازه‌های مختلف ویژگی FFT ۸۵
جدول ۴-۵۵ آزمون تعداد مختلف همسایگی برای ویژگی تبدیل فوریه ۸٦
جدول ۴-۵۶ آزمون تعداد مختلف نرون میانی برای ویژگی تبدیل فوریه ۸٦
جدول ۴-۵۷ محاسبه میزان تسریع ویژگی تبدیل فوریه ۸٦
جدول ۴-۵۸ ماتریس کارایی ویژگی تبدیل فوریه ۸۷
جدول ۴-۵۹ آزمون سایزهای مختلف ویژگی تبدیل کسینوسی ۸۸
جدول ۴-۶۰ آزمون تعداد همسایگی‌های مختلف KNN برای ویژگی تبدیل کسینوسی ۸۸
جدول ۴-۶۱ ماتریس کارایی ویژگی تبدیل کسینوسی ۸۹

جدول ۴-۶۲ آزمایش تعداد نرون‌های میانی مختلف کلاسه‌بند MLP برای ویژگی تبدیل کسینوسی..... ۹۰
جدول ۴-۶۳ آزمون تسریع در ویژگی تبدیل کسینوسی..... ۹۰
جدول ۴-۶۴ سایزهای مختلف پنجره برای ویژگی هیستوگرام کد زنجیر ۹۱
جدول ۴-۶۵ آزمایش سایزهای مختلف برای ویژگی هیستوگرام کد زنجیر ۹۱
جدول ۴-۶۶ آزمایش تعداد نرون‌های مختلف MLP برای ویژگی هیستوگرام کد زنجیر ۹۲
جدول ۴-۶۷ آزمون تسریع برای ویژگی هیستوگرام کد زنجیر ۹۲
جدول ۴-۶۸ آزمون روش‌های مختلف رای‌گیری ۹٤
جدول ۴-۶۹ ماتریس کارایی رای‌گیری روش اکثریت ۹۵
جدول ۵-۱ چند سیستم مشابه جهت مقایسه..... ۹۹

۱- دیباچه

۱-۱- مقدمه

پیش از ورود به بحث تشخیص نوری حروف[1] لازم است اشاره مختصری به شناسایی الگو شود. شناسایی الگو[2] یکی از زیرمجموعه‌های هوش مصنوعی است که به طبقه‌بندی و توصیف مشاهدات می‌پردازد. شناسایی الگو با استفاده از یک سری قوانین علمی کمک می‌کند تا داده‌ها با تکیه بر دانش قبلی یا اطلاعات استخراج‌شده از الگوها طبقه‌بندی[3] شوند[۱]. تشخیص نوری حروف یکی از قدیمی‌ترین زیر شاخه‌های شناسایی الگو می‌باشد. تحقیقات زیادی که در فصول آینده ارائه شده موید این مطلب است که تشخیص نوری حروف به عنوان یک فعالیت کاربردی و تکنیکی جهت تصحیح و در اختیار گرفتن مستندات یکی از مهم‌ترین و فعال‌ترین زمینه‌های تحقیقاتی شناسایی الگو می‌باشد.

به طور کلی تشخیص نوری حروف عبارت است از فرایند تشخیص حروف و کلمات در یک تصویر و تبدیل آن به صورتی که برای یک سیستم کامپیوتری قابل فهم باشد (مانند کد اسکی[4]) که این فرایند می‌تواند الکترونیکی یا مکانیکی باشد [۲]. البته پسوند نوری[5] این عبارت در مقابل مرکب مغناطیسی قرار داده شده تا آن را از روش‌های قدیمی‌تر متمایز کند.

سیستم‌های تشخیص نوری حروف به دو دسته تقسیم می‌شوند، برون خط[6] و برخط[7]. در نوع برون خط یک تصویر ثابت به سیستم داده می‌شود تا حروف آن تشخیص داده شود. اما در نوع برخط در همان لحظه نوشتن حروف اطلاعات به شکل مجموعه‌ای از نقاط وارد

---

[1] Optical character recognition (OCR)
[2] Pattern recognition
[3] Classification
[4] ASCII code
[5] Optical
[6] Offline
[7] Online

سیستم شده و پردازش می‌شود. بنابراین سیستم برخط می‌تواند برای بهبود کارایی از داده‌های زمانی مانند ترتیب نوشتار، سرعت و مسیر خطوط نیز استفاده کند. گوشی‌های تلفن همراه لمسی و کامپیوترهای لوحه‌ای[1] نمونه‌های سیستم برخط می‌باشند. یک سیستم تشخیص حروف برخط از سه جز تشکیل شده است. اول یک قلم یا وسیله‌ای مانند آن برای نوشتن، دوم یک سطح حساس برای نوشتن و سوم یک نرم افزار که بتواند این حرکات را تشخیص بدهد. سیستم برخط به اسامی دیگری نیز معروف است مانند ICR[2] و سیستم تشخیص بلادرنگ حروف.

سیستم‌های برون خط به دو دسته تقسیم می‌شوند، چاپ شده و دستنویس[3]. همان طور که از اسم این دو پیداست در نوع دستنویس متنی که به سیستم داده می‌شود توسط انسان نوشته و در نوع چاپ شده تصویری که قبلا توسط چاپگر یا ماشین واژه پرداز تهیه شده است به سیستم داده می‌شود.

۱-۲- کاربردها

برای ورود اطلاعات یک فرم دستنویس به کامپیوتر اولین راهکاری که به ذهن می‌رسد این است که یک تایپیست آن را وارد سیستم کند. اما این روش هزینه مالی داشته و زمان بر است. این مساله موقع بکار بردن این روش در مقیاس وسیع‌تر، مثلاً در ورود فرم‌های آماری، بیشتر بروز می‌کند و لذا استفاده از روش‌های سنتی دیگر مناسب نخواهد بود.

کاربردهای OCR محدوده وسیعی را شامل می‌شود. از مهمترین کاربردهای آن خواندن فرم‌هایی است که با دست پر شده‌اند. کاربردهای دیگری نیز برای آن وجود دارد مانند مرتب سازی پاکت‌های نامه برای ارسال، خواندن اطلاعات از چک‌های بانکی و مدیریت مستنداتی که قبلا کاغذی بوده‌اند [۳]. تبدیل یک سند سنتی به یک سند دیجیتال مزایای بسیاری از جمله کم کردن حجم اطلاعات، قابلیت ویرایش و جستجوی سریع به دنبال دارد که در مستندات دستی و سنتی این گونه نمی‌باشند.

[1] Tablet pc
[2] Intelligent character recognition
[3] Handwritten

۱-۳- انگیزش

مستندات کاغذی که در زمان‌های گذشته برای ایجاد ارتباط بین افراد و نگهداری اطلاعات به وجود آمده‌اند، با وجود پیشرفت الکترونیک امروزی هنوز هم مهمترین رسانه انتقال و نگهداری اطلاعات در دنیا می‌باشند. این مستندات به دلیل اطمینان، قابلیت حمل و سهل الوصول بودن حتی در آینده نیز نقش مهمی در زندگی روزمره بازی خواهند نمود. لذا سوالی که به ذهن خطور می‌کند این است که «با وجود استفاده از مستندات کاغذی آیا تشخیص ماشینی متون ضرورتی خواهد داشت؟» . پاسخ این سوال واضح است. اگرچه نسخه پویش[1] شده مستندات درون کامپیوتر می‌تواند روی صفحه نمایش نشان داده شود؛ اما این تصویر قابل خواندن و فهم توسط ماشین نمی‌باشد و همچنین از مزایای دیجیتال شدن مستندات (از قبیل ویرایش) نیز نمی‌توان استفاده کرد.

طبق تحقیقات انجام شده در سال ۲۰۰۹ توسط شرکت Dlib تشخیص حروف چاپی روزنامه توسط نرم افزارهای تجاری لاتین درستی بین ۷۱ تا ۹۸ درصد داشته اند [۴]، بنابراین حتی در تشخیص متون لاتین هم هنوز جای کار وجود دارد. همچنین به دلیل آن که نوشتار هر فرد متفاوت از دیگری است مسئله بازشناسی حروف دستنویس پیچیده‌تر می‌شود و لذا مسئله بهبود فرایند هنوز حل نشده است.

در حالی که بیش از نیم میلیارد نفر در جهان با زبان‌های فارسی، عربی، کردی و اردو تکلم می‌کنند (که از مجموعه حروف تقریبا یکسانی استفاده می‌کنند) تحقیقات کمی در زمینه تشخیص حروف این زبان‌ها شده است و در مقابل تحقیقات بسیاری درخصوص زبان‌های لاتین و چینی صورت گرفته است [۵]. مهمترین دلیل کمی تحقیقات در این زبان‌ها کمبود پشتیبانی ابزاری مانند پایگاه داده و فرهنگ لغت و همچنین ذات پیوسته و پیچیده نوشتار آن‌ها می‌باشد [۶]. بنابراین توانایی تشخیص حروف فارسی و عربی می‌تواند مزایای زیادی داشته باشد. از جهت دیگر حروف فارسی/عربی تغییر چندانی در طول تاریخ نکرده است و می‌توان از چنین سیستمی‌برای تشخیص متون تاریخی نیز استفاده کرد [۷] مانند کاری که در [۸] توسط آقای خورشید برای شناسایی متون کهن استفاده شد. تا چندین سال پیش حتی یک سیستم تجاری قابل اعتماد برای بازشناسی متن پیوسته فارسی وجود نداشت [۹]. با توجه به دلایل ذکر شده ضرورت وجود یک سیستم OCR با دقت بالا احساس می‌شود.

[1] Scan

از طرفی با توجه به روند همه‌گیر شدن سیستم‌های چندهسته‌ای لزوم استفاده از الگوریتم‌های موازی برای افزایش سرعت احساس می‌شود که می‌تواند موجب بالا نگه داشتن صحت پاسخ‌ها مانند روش ترتیبی باشد. با توجه به جدید بودن این تکنیک هنوز در سیستم‌های تشخیص حروف چندان از الگوریتم‌های موازی استفاده نشده است.

۱-۴- نوشتار فارسی

برای تمرکز بیشتر بر حروف فارسی، در این قسمت اندکی به نحوه نوشتار متون فارسی و عربی پرداخته خواهد شد. نوشتار عربی و فارسی بر خلاف انگلیسی بصورت پیوسته و راست به چپ می‌باشد و همچنین حروف کوچک و بزرگ ندارند. نوشتار فارسی متشکل از ۳۲ حرف بوده و نوشتار عربی ۲۸ حرف از زیر مجموعه آن است که هر حرف ۲ یا ۴ شکل دارد و بسته به جای آن در زیرکلمه و حروف اطراف آن یکی از این ۴ حالت استفاده می‌شود. این ۴ حالت، ابتدای زیرکلمه، میان زیرکلمه، پایان زیرکلمه و جداگانه هستند که در جدول ۱-۱ می‌توان آن‌ها را مشاهده نمود. حروفی نیز هستند که نمی‌توانند به صورت متصل به کار روند و حالت وسط آن‌ها همان حالت جدای آن‌ها و حالت اول همان حالت مجزای آن‌ها می‌باشد.

جدول ۱-۱ حالت‌های مختلف حروف

نام	ابتدایی	میانه	پایانی	جدا
الف	ـ	ـا	ـا	ا
ب	بـ	ـبـ	ـب	ب
پ	پـ	ـپـ	ـپ	پ
ت	تـ	ـتـ	ـت	ت
ث	ثـ	ـثـ	ـث	ث
جیم	جـ	ـجـ	ـج	ج
چ	چـ	ـچـ	ـچ	چ
ح	حـ	ـحـ	ـح	ح
خ	خـ	ـخـ	ـخ	خ
دال	ـ	ـد	ـد	د
ذال	ـ	ـذ	ـذ	ذ
ر	ـ	ـر	ـر	ر
ز	ـ	ـز	ـز	ز
ژ	ـ	ـژ	ـژ	ژ
سین	سـ	ـسـ	ـس	س
شین	شـ	ـشـ	ـش	ش

نام	ابتدایی	میانه	پایانی	جدا
صاد	صـ	ـصـ	ـص	ص
ضاد	ضـ	ـضـ	ـض	ض
طا	طـ	ـطـ	ـط	ط
ظا	ظـ	ـظـ	ـظ	ظ
عین	عـ	ـعـ	ـع	ع
غین	غـ	ـغـ	ـغ	غ
ف	فـ	ـفـ	ـف	ف
قاف	قـ	ـقـ	ـق	ق
کاف	کـ	ـکـ	ـک	ک
گاف	گـ	ـگـ	ـگ	گ
لام	لـ	ـلـ	ـل	ل
میم	مـ	ـمـ	ـم	م
نون	نـ	ـنـ	ـن	ن
واو	ـ	ـو	ـو	و
ه	هـ	ـهـ	ـه	ه
یا	یـ	ـیـ	ـی	ی

حروف میان کلمه ای به طور معمول به همدیگر متصل می‌شوند. چسبیدن حروف به یکدیگر به مکان آن در کلمه و حروف اطراف آن بستگی دارد. حروف هنگامی که به هم می‌چسبند در یک ارتفاع و روی خط پایه وصل می‌شوند(مانند خط پایه در انگلیسی). خط پایه در متون دستنویس یک خط فرضی ایده‌آل است که اتصالات در نزدیکی آن رخ می‌دهند.

در فارسی هیچ دو کلمه‌ای به هم اتصال ندارند بنابراین همواره بین هر دو کلمه فاصله وجود دارد. زمانی که حروف بدون حالت وسط در میان کلمه قرار می‌گیرند کلمه به چند زیرکلمه تبدیل می‌شود. این زیر کلمات نیز با یکدیگر فاصله دارند اما فاصله آن‌ها کمتر از فاصله بین کلمات است.

در نوشتار عربی علائم مشخصه ای مانند فتحه و کسره وجود دارد که بیشتر برای قرائت صداهاست. این علائم در خط الرسم استاندارد استفاده نمی‌شوند.

در نوشتار فارسی تعدادی حروف پیوندی[1] وجود دارد، به این معنی که وقتی دو یا چند حرف خاص کنار همدیگر می‌آیند به نحوی دیگر نوشته می‌شوند و استثناهایی برای قواعد اتصال هستند مانند اتصال یا و میم و اتصال لام و الف که در متون چاپی به شکل "لا" و در دستنویس به شکل ۱-۱ نمایش داده می‌شود.

شکل ۱-۱حروف پیوندی

در زیر برخی از مشکلات تشخیص نوری آورده شده است.

مشکلات عمومی:

- وجود خطوط و اشیای غیر بخش اصلی حروف مانند نقطه، همزه و مد
- وجود نویز در تصویر پویش شده

مشکلات تشخیص حروف دستنویس:

- تنوع در ابزار نوشتاری که ممکن است باعث نازکی، کلفتی ، کیفیت و رنگ متفاوت آن شود. کلفتی ممکن است باعث این موارد شود: متصل شدن حروف به هم و پر شدن حفره های میان حروف
- نویسندگان مختلف و یا حتی یک نویسنده در حالات مختلف ممکن است یک حرف را به شکل مختلفی بنویسد.
- نوشته نشدن حروف در یک موقعیت نسبت به هم

[1] Ligature

- وجود سمبل های مشابه مثل "۱" و "l"

مشکلات تشخیص حروف فارسی/عربی:

- ذات پیوسته آن
- اشکال مختلف حروف بنابه مکان آن در کلمه
- وجود حروفی که شکل آن ها بسیار مشابه می‌باشند و تفاوت آن‌ها ممکن است فقط در یک نقطه باشد [۱۰].

۱-۵- اهداف تحقیق

هدف ابتدایی این تحقیق تولید یک سیستم کارای خودکار برای تشخیص حروف مقطع دستنویس فارسی می‌باشد. این سیستم تصویر حرف مورد نظر را به صورت دیجیتال به عنوان ورودی دریافت کرده و کلاس آن حرف که عددی بین ۱ تا ۳۲ می‌باشد را به عنوان خروجی تولید می‌نماید.

با توجه به وجود کاربردهای واقعی مسئله، مانند خواندن اطلاعات چک، دستیابی به صحت بالای تشخیص از اهداف این سیستم می‌باشد. برای بالا بردن این صحت در این سیستم از ترکیب روش‌های موجود استفاده خواهد شد. میزان صحت یک سیستم توسط فرمول ۱-۱ محاسبه می‌شود.

$$\text{نرخ تشخیص} = \frac{\text{تعداد حروف درست تشخیص داده شده}}{\text{تعداد کل حروف مورد آزمون}} \qquad (۱-۱)$$

یک معیار دیگر کارایی سرعت انجام می‌باشد. از این رو در این سیستم بنا شده الگوریتم‌ها به صورت موازی روی هسته‌های پردازنده اجرا شوند تا زمان اجرای آن‌ها کاهش یابد. یک معیار مقایسه سرعت حالت موازی و ترتیبی تسریع می‌باشد که با فرمول ۱-۲ محاسبه می‌گردد.

$$\text{تسریع} = \frac{\text{زمان اجرای برنامه در حالت ترتیبی}}{\text{زمان اجرای برنامه در حالت موازی}} \qquad (۱-۲)$$

اگرچه هدف این پایان نامه تولید سیستمی برای تشخیص حروف مقطع دستنویس فارسی است اما از این سیستم با اندک تغییراتی می‌توان در شناسایی ارقام و حتی کلمات

چسبیده نیز استفاده کرد. از طرف دیگر الگوریتم‌های به کار رفته الگوریتم‌های زمانبری نبوده و زمان طولانی آن‌ها به دلیل تعداد زیاد داده می‌باشند، لذا از این الگوریتم‌ها می‌توان به صورت موازی و با پخش داده‌ها روی چند رایانه استفاده کرد.

۱-۶- چشم انداز پایان نامه

سازماندهی این پایان نامه به این شکل است. در فصل اول، دیباچه این پایان نامه به معرفی مسئله، اهداف پایان نامه و انگیزش آن پرداخته شده است. مروری بر تاریخچه، پایگاه داده‌های موجود و بررسی جامع تحقیقات در زمینه تشخیص حروف در فصل دوم صورت می‌گیرد. در فصل سوم معماری یک مدل پایه برای تشخیص حروف ارائه می‌شود. همچنین برخی از روش‌های تشخیص معمول و پایه در این فصل آورده شده است. فصل چهارم به ارائه جزئیات روش پیشنهادی برای یک سیستم تشخیص حروف دستنویس می‌پردازد و در پایان این فصل نتایج این سیستم‌ها به همراه مقایسه آن‌ها با روش‌های دیگر و کارهای آتی ارائه خواهند شد.

۲-مروری بر پژوهش‌های مرتبط

۲-۱- مقدمه

در این فصل ابتدا به تاریخچه و مراحل تکامل سیستم‌های شناسایی حروف پرداخته شده، سپس پایگاه داده‌های حروف فارسی مورد بررسی قرار می‌گیرد. پس از آن مروری بر کارهای انجام شده در زبان‌های فارسی و عربی پرداخته خواهد شد.

۲-۲- تاریخچه

با پیشرفت‌های روزافزون وسایل ارتباط جمعی الکترونیکی، نیاز به سیستم‌هایی احساس می‌شود که بتوانند با دقت و سرعت، اطلاعات مستندات نوشتاری موجود را خوانده و ذخیره نمایند. تحقیقات روی تشخیص نوری حروف از دهه ۱۹۵۰ میلادی آغاز شده است و امروزه نرم افزارهای تجاری مطمئنی برای متون چاپی بوجود آمده‌اند و تحقیقات روی بهبود این سیستم‌ها ادامه دارد.

سیستم‌های تشخیص نوری حروف در کل جهان در سه مرحله رشد تکاملی خود را داشته‌اند[۱۱].

- مرحله تکوین: در دهه‌های میانی قرن بیستم که سیستم‌ها کامپیوتری نبودند و فرایند به صورت مکانیکی بود، اقدامات اولیه در این زمینه صورت گرفت. اولین اختراع در این خصوص مربوط به سیستمی‌در دهه ۱۹۳۰ میلادی در آلمان است [۱۲]. این اختراع یک ماشین مکانیکی بود که حروف چاپی را با روش تطابق قالب[1] شناسایی می‌کرد. نحوه کار آن بدین صورت بود که به هر حرف نور تابیده می‌شد و نور بازتاب شده از قالب مکانیکی عبور داده می‌شد و هرگاه نور از قالب عبور نمی‌کرد حرف تشخیص داده می‌شد. در این

[1] Template matching