وزارت علوم، تحقیقات و فناوری
مؤسسه آموزش عالی غیردولتی-غیرانتفاعی سجاد

پایان نامه دوره کارشناسی ارشد برق

گرایش مهندسی مخابرات

تشخیص حروف دست‌نویس فارسی با استفاده از شبکه تابع پایه شعاعی

سید محمدرضا حافظ

استاد راهنما :

دکتر وحید اسدپور

اساتید دفاع :

دکتر ایمان احدی اخلاقی

دکتر مسعود امینیان مدرس

شهریور ۱۳۹۱

بسم الله الرحمن الرحيم

تقدیر و تشکر

سپاس خداوندی را که به انسان موهبت تفکر کردن عطا کرد و در این مسیر، امتیاز استفاده از تجربیات دیگران را ارزانی داد،

و تقدیر فراوان از اساتید گرامی، که با نهایت دلسوزی تجربیات خود را در اختیار رهجویان علم قرار می‌دهند،

و تشکر بسیار از پدران و مادران که بدون هیچ چشم داشت، دارای سهم به سزایی در به عمل نشستن علوم فرا گرفته شده در فرزندان خود هستند.

باشد که در قبال این همه لطف، در راستای انجام وظیفه کوشا بوده، تا بلکه بتوانیم گوشه ای هر چند ناچیز از این محبت‌ها را پاسخگو باشیم.

چکیده

در تحقیق ارائه شده ، مساله بازشناسی حروف دستنویس فارسی به کمک شبکه‌های عصبی مورد بررسی و مطالعه قرار گرفته است. جهت انجام این پروژه، مراحلی تعیین و تعریف می‌شود که شامل انتخاب و تعیین خصوصیت منحصر به هر حرف بوده، به گونه‌ای که در مقابل مواردی از قبیل چرخش، تغییر اندازه و جابجایی مقاوم باشد. تبدیل فرمت ورودی به شکل قابل ارائه به شبکه و آنگاه آموزش شبکه عصبی از دیگر مراحل اجرای پروژه است. در این پروژه از یک خصوصیت مهم به نام Shape Context با توجه به استفاده از الگوریتم Otsu جهت انجام عملیات باینری کردن تصاویر نمونه و چندین خصوصیت تکمیلی دیگر در کنار آن به صورت ترکیبی استفاده شده است. استفاده از این خصوصیت مهم به ما در دستیابی به اهداف ذکر شده از قبیل قابلیت انعطاف بالای ویژگی‌های استخراج شده در برابر چرخش و همچنین محل قرار گرفتن حرفی که مورد پردازش قرار گرفته، کمک می‌کند. این خصوصیت به شیوه‌ی خاصی با محیط کردن یک دایره به دور حرف مورد نظر با عدم توجه به اینکه تصویر حرف مورد نظر در کدام قسمت از کادر قرار گرفته، اقدام به قسمت بندی داخل دایره مورد نظر در راستای شعاعی و همچنین در راستای قطر دایره کرده و با شمارش تعداد پیکسل‌های موجود در هرکدام از قسمت‌ها ماترس ویژگی حاصل از خصوصیت Shape Context را در خروجی نمایش می‌دهد. ماتریس حاصل شاکله اصلی ماتریس ویژگی ما را تشکیل خواهد داد که در کنار آن ویژگی‌هایی نظیر تعداد نقاط موجود در هر شکل و ویژگی‌هایی از این دست قرار داده می‌شود تا تشخیص به صورت دقیق‌تر صورت پذیرد. در انتهای عملیات، آموزش بر روی نه کلاس داده که از تصاویر استخراج شده تشکیل شده‌اند صورت گرفته و بر روی داده آزمون که داده‌ای است که در مجموعه داده‌های مورد آموزش توسط شبکه نمی‌باشد، امتحان می‌شود، که در ۹۰/۲۳٪ موارد، خروجی به درستی تشخیص داده می‌شود. در قسمت پایانی آزمایش ها سعی شد تا خروجی شبکه‌ای که توسط تمام داده‌ها آموزش دیده است مورد بررسی قرار گیرد، که با دادن کلماتی که شامل حروف مقطع از داخل خود مجموعه بودند در اکثریت کلمات خروجی صحیح را داشته باشیم.

فهرست

فصل اول – مقدمه و مروری بر مقالات

۱-۱- مقدمه ۲

۱-۲- تاریخچه تشخیص کاراکتر ۳

۱-۳- تشخیص دست‌خط ۵

۱-۳-۱- تشخیص دست خط عربی و فارسی ۷

۱-۳-۲- روش‌های تشخیص دست‌خط ۹

۱-۳-۳- مقایسه دو روش تشخیص هم زمان و غیر هم زمان ۱۰

۱-٤- مروری بر مقالات ۱۱

فصل دوم : مروری بر عملکرد شبکه‌های عصبی و بررسی شبکه‌های تابع پایه شعاعی

۲-۱- مقدمه ۱۸

۲-۲- تابع انتقال نرون‌های شبکه عصبی ۲۰

۲-۳-آموزش شبکه عصبی ۲۳

۲-۳-۱- تقسیم بندی قوانین آموزش براساس نحوه اعمال آموزش به شبکه ۲٤

۲-۳-۲- آموزش با نظارت ۲۵

۲-۳-۳ – آموزش بدون نظارت ۲۵

۲-۳-٤- تشخیص الگو ۲۸

۲-۳-۵- روش‌های مختلف تشخیص الگو ۲۹

۲-۳-۵-۱- روش‌های ریاضی ۲۹

۲-۳-۵-۲- استفاده از شبکه‌های عصبی ۲۹

۲-٤ – شبکه‌های عصبی چندلایه به عنوان دسته بندی کننده ۳۰

۲-۵ – روش‌های کاهش احتمال مینیمم‌های محلی ۳۱

۲-۵-۱ – کاهش تدریجی نرخ یادگیری ۳۱

۲-۵-۲ – افزایش نرون‌های لایه پنهان ۳۱

۲-٦- بررسی شبکه های تابع پایه شعاعی ۳۲

۲-٦-۱- معماری شبکه RBF ۳۲

۲-۷ – آموزش یک شبکه RBF ۳٤

۲-۷-۱ – آموزش لایه پنهان ۳٤

۲-۷-۲ – آموزش لایه خروجی ۳٦

۲-۸- مزایای شبکه RBF ۳٦

۲-۹- توقف آموزش ۳۶

فصل سوم : بررسی فعالیت‌های پیشینیان

۳-۱ – بررسی فعالیت‌های پیشینیان..... ٤٠

فصل چهارم : روش پیشنهادی

٤-۱ – مقدمه ٤٥

٤-۲ – مجموعه داده‌ها..... ٤٥

٤-۳ – پیش پردازش..... ٤٧

٤-۳-۱ – باینری کردن..... ٤٧

٤-۳-۲ – اسکلت بندی ٥٠

٤-٤ – استخراج ویژگی ٥١

٤-٤-۱ – چرا استخراج ویژگی؟ ٥١

٤-٤-۳ – تعداد اشیای موجود در شکل ٥٧

٤-٤-٤ – تعداد نقطه‌ها ٥٩

٤-٤-٥- عدد اویلر ٦٠

٤-٤-٦ – پراکندگی پیکسل‌ها..... ٦١

٤-٤-۸ – هیستوگرام عمودی و افقی..... ٦١

٤-٤-٩ – موقعیت نقطه‌ها نسبت به قطعه اصلی شکل..... ٦٣

٤-٥ – تشکیل بردار ویژگی ٦٤

٤-٦ – طراحی شبکه جهت دسته‌بندی بردارهای ویژگی ٦٧

فصل پنجم : آزمایش‌های انجام شده وبررسی نتایج

٥-١ – مقدمه ٦٩

٥-٢- نتایج آزمایش بر روی داده‌های آزمون ٧٠

٥-٣– نمودار میانگین خطاها ٧٨

٥-٤ – نتیجه آموزش شبکه بر روی کلیه داده‌ها ٧٩

فصل ششم : جمع‌بندی و پیشنهادات

٦-١ – جمع‌بندی ٨٢

٦-٢ – پیشنهادات ٨٣

مراجع ٨٥

فهرست اشکال

فصل اول

شکل ۱-۱- ابهامات در تشخیص کلمات دست نوشته a .یک تکه پیکسل به تنهایی باعث مشاهده یک کاراکتر به تنهایی و یا دو کاراکتر شده، b. یک قسمت از پیکسل ها باعث آشکار سازی دو نوع کاراکتر شده ٦

شکل ۲-۱- ابهام ایجاد شده توسط قطعه بندی. a. کلمه مورد نظر b. حتی در حالتی که کاراکترهای معقول در تصویر مشاهده می شوند، قطعه بندی نادرست منجر به تطبیق اشتباه شده ۷

شکل ۳-۱- دو دست نوشته متفاوت از یک کلمه که کاملا روی هم افتادگی حروف را در هر دو نشان میدهد. ۸

فصل دوم

شکل ٤-۲- یک ساختار کلی از یک شبکه عصبی ۱۸

شکل ٥-۲- مدل ساده از یک پردازنده عصبی ۲۱

شکل ٦-۲- تابع غیرخطی محدودکننده سخت ۲۱

شکل ۷-۲- تابع انتقال خطی ۲۲

شکل ۸-۲- تابع انتقال غیرخطی سیگموئید ۲۳

شکل ۹-۲- یک شبکه سه لایه ۲٦

شکل ۱۰ - ۲- یک شبکه عصبی چند لایه با دولایه پنهان ۳۰

شکل ۱۱-۲- شبکه تابع پایه شعاعی(RBF) ۳۲

شکل ۱۲-۲- یک منحنی گوسی به مرکز صفر و انحراف معیار ۰/٥ ..... ۳۳

شکل ۱۳-۲- خطای مرتبط با مجموعه های آموزش و آزمون در خلال آموزش ۳۸

فصل چهارم

شکل ١٤-٤- مجموعه دادهها (کلاس‌ها) ٤٦

شکل ۱۵-٤- نمونه ای از یک کلاس ٤٦
شکل ۱٦-٤- نمونه حروف با پیکسل‌هایی به ارزش ۰ تا ۲۵۵ ..... ٤٧
شکل ۱۷-٤- تصاویر باینری شدن یک حرف با اعمال آستانه‌های متفاوت ٤۹
شکل ۱۸-٤- نمونه‌ای از یک تصویر باینری شده به روش OTSU ٤۹
شکل ۱۹-٤- نمونه حروف با ضخامت‌های متفاوت ۵۰
شکل ۲۰-٤- نمونه تصاویر باینری و اسکلت بندی شده آنها ۵۱
شکل ۲۱-٤- نمونهای از یک حرف با ابعاد، زوایا و محل‌های متفاوت ۵۲
شکل ۲۲-٤- تعداد اشیای موجود در تصویر هر حرف(a) که میتواند برای حروف یکسان، متفاوت باشد ۵۳
شکل ۲۳-٤- SHAPE CONTEXT ۵٤
شکل ٢٤-٤ الف- مرکز ثقل ب- تعداد اشیای یک شکل ۵٦
شکل ۲۵-٤- حروف با تعداد اشیاء تشکیل دهنده متفاوت ۵۷
شکل ۲٦-٤- خروجی‌های متفاوت ناشی از ویژگی «چند شیئی بودن»، در چند حرف نمونه ۵۸
شکل ۲۷-٤- انواع حالت‌های نمایش دونقطه ۵۹
شکل ۲۸-٤- انواع حالت‌های نمایش سهنقطه ۵۹
شکل ۲۹-٤- نمونه حرفی که در آن نقطه‌ها از بدنه اصلی حروف بزرگترند ٦۰
شکل ۳۰-٤- عدد اویلر برای چند نمونه حرف ٦۰
شکل ۳۱-٤- هیستوگرام عمودی و افقی تصویر نمونه ٦۲
شکل ۳۲-٤- مستطیل محاط شده به حرف نمونه ٦٦

فصل پنجم

شکل ۳۳-۵- میانگین خطا بر روی هر کدام از داده های آزمون ۷۸
شکل ٣٤-۵- الف- تصویر کلمه همراه با چرخش ۲۰ درجه ب- تصویر کلمه بون چرخش ۷۹

فهرست جداول

فصل اول

جدول ۱-۱- موقعیت هر حرف به تنهایی درون کلمه ۹

فصل پنجم

جدول ۲-۵- محاسبه میزان خطای شبکه بر روی داده آزمون بدون وزن دادن به ویژگی‌ها ۷۰

جدول ۳-۵- محاسبه میزان خطای شبکه با وزن‌دهی ویژگی‌ها ۷۱

جدول ٤-۵- محاسبه میزان خطای شبکه با وزن‌دهی ویژگی‌ها ۷۲

جدول ۵-۵- محاسبه میزان خطای شبکه با استخراج ویژگی‌ها از روی تصویر اصلی ۷٤

جدول ٦-۵- محاسبه میزان خطای شبکه با کوچک کردن ابعاد Shape Context ۷۵

جدول ۷-۵- محاسبه میزان خطای شبکه با توجه به افزایش وزن عدد اویلر ۷٦

جدول ۸-۵- محاسبه میزان خطای شبکه با افزایش ابعاد Shape Context ۷۷

فصل اول

مقدمه و مروری بر مقالات

۱-۱- مقدمه

در بخش ابتدایی فصل اول که به عنوان مقدمه‌ای از معرفی موضوع سمینار از آن نام برده می‌شود، به صورت خلاصه به بررسی علل و عوامل برگزیده شدن موضوع روش‌های تشخیص دست خط به عنوان موضوع ارائه شده در سمینار توجه کرده و در ادامه این بحث به بررسی انگیزه بوجود آمده در راستای انتخاب این موضوع پرداخته خواهد شد و تعدادی کاربرد های کلی بر گرفته شده از این موضوع معرفی می‌شوند .

موضوع روش‌های تشخیص دست خط موضوعی است که در معرفی اولیه و دیده شدن عنوان آن، هر فردی، حتی اگر این فرد تحصیلات ویژه ای جهت شناخت اولیه این موضوع هم نداشته باشد، میتواند یک تداعی ذهنی و یک پیش زمینه قبل از ارائه آن در ذهن خود در رابطه با این عنوان و موضوع پیدا کرده و در شرایطی که این خصوصیت را به همراه دارد، موضوعی پر محتوا که جای بحث و پیش برد در همه بخش ها و زیر شاخه ها در آن موجود می‌باشد را در خود حس می‌کند.

موضوع تشخیص دستخط زیر مجموعه ای از بحث OCR[۱] می باشد که به زعم مقالات متعدد پر محتواترین و گسترده ترین زیر مجموعه این مقوله محسوب می شود، به گونه ای که از موضوع تشخیص دست خط به عنوان یک مساله باز[۲] یاد شده، که این خود نشان دهنده بی انتها بودن فعالیت در تمام شاخه ها و گرایش های مربوط به این موضوع می‌باشد. هدف اصلی این موضوع سمینار، رسیدن به یک جمع بندی کلی و یک معرفی کلی از تمام فعالیت های متنوعی که در مسیر تحول انجام شده در راستای این موضوع صورت پذیرفته می‌باشد. امید است با این توضیحات و بررسی های مناسب بتوان مکمل خوبی برای اطلاعات از گذشته موجود در ذهن شما در رابطه با این موضوع بوجود آورد.

تعدادی از کاربرد های کلی تشخیص خودکار یک متن از تصویر برگرفته از روی متن:

۱. جستجو برای پیدا کردن عبارتی خاص در یک مجموعه نوشته بزرگ
۲. تنظیم و طبقه‌بندی نامه های پستی

---

۱ - Optical Character Recognition
۲ - open problem

۳. پردازش های دلخواه بر روی متن تشخیص داده شده و چاپ متن دست نوشته با معیارهای دلخواه

۱-۲- تاریخچه تشخیص کاراکتر

مبحث باز شناسی متون، همانند مبحث تشخیص صحبت از اوایل ظهور مبحث شناسی الگو مطرح بود. در ابتدا شناسایی آنها آسان به نظر می‌رسید ولی در عمل بر خلاف انتظار بعد از یکسری پیشرفت کوچک، این مساله با مشکلات بزرگی روبرو شد و توجه محققین به سمت سایر زمینه های شناخت الگو معطوف گشت. با پیدایش وسایل ارتباط جمعی الکترونیکی، نیاز شدید به سیستمهایی ایجاد می شد که بتوانند با دقت و سرعت، اطلاعات نوشتاری موجود را خوانده و ذخیره نمایند.

در سال ۱۹۲۹Taushech در آلمان و در سال ۱۹۳۳ Handle در آمریکا ابداعاتی در زمینه OCR ثبت نمودند. اینها اولین ایده های شناسایی حروف بودند. این روش ها، تطبیق الگو نام داشتند و بدین صورت کار می‌کردند که، به هر حرف نور تابیده می شود و نور باز تابیده شده از حروف ، از قالب های مکانیکی عبور داده می شود. هر گاه که نوری از قالب عبور نکرد، حروف تشخیص داده می‌شود.

در عمل تا زمان به عرصه آمدن کامپیوتر یعنی سال ۱۹۵۰ سیستم مطلوبی ساخته نشد. تاریخچه تشخیص کاراکتر به عنوان یک مساله پردازشی کلاسیک به زمان پیدایش پردازش ها توسط شبکه عصبی[۱] باز می‌گردد که این ایده استفاده از شبکه عصبی در این رابطه، ابتدا در سال ۱۹۵۰ توسط Frank Rosenblatt مطرح می شود. این اولین تابع پردازشی توسط یک پردازنده شبکه عصبی متشکل از نرون ها بود که وظیفه تشخیص کاراکتر های با شکل ثابت و تنظیم شده در محلی مشخص را بر عهده داشت.در سال ۱۹۵۱ بود که ایده OCR به عنوان یک پدیده قابل پیاده سازی پذیرفته شد.از این زمان به بعد تحقیقات گسترده ای بر روی این تکنولوژی انجام گرفته است و در نتیجه سیستم های تجاری عرضه شدند که دارای قابلیت های خوبی هستند. سیستم های تجاری عرضه شده را می توان به سه نسل تقسیم بندی نمود.

سیستم های نسل اول در اوایل دهه ۱۹۶۰ به بازار آمدند که نتیجه تحقیقات دهه ۱۹۵۰ بودند. آنها فقط قابلیت تشخیص کاراکتر های خاص با اندازه و فونت مشخصی را داشتند. در آنها از روش هایی استفاده شده بود که نسبت به

---

۱ - Neural Network

تغییر جای کاراکترها، اندازه و دوران آنها فوق العاده حساس بودند. شاخص ترین سیستم این نسل، NCR 420 بود که می توانست اعداد و پنج نماد دیگر را تشخیص دهد.

سیستم های نسل دوم در اواسط دهه ۱۹۶۰ به بازار آمدند و می توانستند بعضی از کاراکتر های دست نویس مانند اعداد دست نویسی شده را تشخیص دهند.

سیستم های نسل سوم مربوط به اواخر دهه ۱۹۶۰ هستند که با هدف تشخیص کاراکتر هایی با کیفیت چاپ پایین و همچنین متون دست نویس لاتین ساخته شدند. این نسل از سیستم های تشخیص کاراکتر تا به امروز هم ادامه دارد و فعالیت های گسترده ای در این زمینه در حال انجام است. امروزه سیستم های OCR قادر به تشخیص دقیق کاراکتر های تایپی لاتین با انواع فونت ها و در اندازه های مختلف هستند(تا ٪۹۹ میزان صحت). ولی هنوز در مورد تشخیص متون دست نویس لاتین و یا فونت هایی که در آنها از خطوط خمیده استفاده می‌شود(مثل زبان های فارسی و عربی) مشکلات متعددی وجود دارد. با توجه به کاربرد های هوش مصنوعی تا سال ۱۹۸۰ به مقوله تشخیص دست خط به عنوان یک امر بسیار دشوار و غی قابل دستیابی نگاه می شد، اما با پیدایش سرعت در پردازش ها و توانایی همگرا شدن تعداد زیادی از تکنولوژی ها بر روی موضوع تشخیص کاراکتر، این امکان از سال ۱۹۹۰ به بعد به محققان داده شد تا جهت دستیابی به سیستم های معتبر و قابل اطمینان اقدام کنند[۸].

روش های زیادی تا به امروز برای شناسایی متون فارسی و عربی به کار رفته‌اند. در سال ۱۳۵۹ به دنبال ابداع سیستمی برای شناسایی دست نوشته های عربی موسوم به IRAC توجهات به این زمینه نوین در قلمرو باز شناسی الگو معطوف شد. به دنبال آن سیستم دیگری برای تشخیص کلمات تایپ شده عربی پیشنهاد شد که در آن ابتدا کلمات به حروف تفکیک شده و سپس توصیف کننده های فوریه دوره خارجی حروف را مبنای شناسایی این حروف قرار می-دهند. به موازات آن دو روش ساختاری مطرح شدند:

- یکی مبتنی بر ایده دنبال کردن کانتور برای شناسایی حروف مجزای عربی
- دیگری بر اساس استفاده از محل های تقاطع و انشعاب پاره‌خط ها برای شکستن کلمات به زیر حرف ها و دسته بندی آنها به کمک ویژگی های هندسی و توپولوژیکی.

در سال ۱۳۶۰ روشی برای شناسایی فونت های بزرگ فارسی ابداع شد که از ویژگی هایی نظیر وجود حفره و دهنه برای شناسایی حروف بهره می‌جست. در سال ۱۳۷۲ دو سیستم جهت شناسایی حروف در متن فارسی تایپی ارائه

گردید که یکی با روش شکل شناسی و دیگری بر اساس توصیف کننده های فوریه عمل می نمودند. امروزه با کمک شبکه های عصبی که قادرند ارتباط پیچیده بین ورودی و خروجی را برقرار کنند، می‌توان حروف فارسی را شناسایی نمود.

۱-۳- تشخیص دست‌خط

بیشترین مسائل دشوار در زمینه تشخیص کاراکتر ها (OCR) مربوط به تشخیص دست‌خط شکسته بدون قید می-باشد. ابزار های کنونی هنوز به آن میزان از شایستگی و قابلیت اطمینان نرسیده اند تا توانایی مدل کردن تشخیص تنوع زیادی از دست خط افراد مختلف را داشته باشند. تشابه اشکال کاراکتر های مجزا، روی هم افتادگی‌ها، اتصالات داخلی در کاراکتر های مجاور یکدیگر باعث پیچیده شدن این موضوع شده است. به اضافه، مشاهده شده در هنگام جداسازی، کاراکتر ها اغلب دچار ابهام شده و جهت کاهش خطاهای دسته بندی نیاز به اطلاعات متنی داریم.

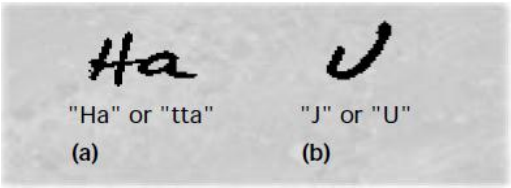
شناسایی دست خط نیازمند ابزارها و تکنیک های متعدد می باشد تا در شناسایی عبارات مجتمع شده از کاراکترها و همچنین موارد ابهام بوجود آمده در آنها از نظر برداشت های ادراک شده عادی راجع به اشکال و نمایش عمومی کاراکترها، کلمات، لغات و عبارات به درستی عمل کند.شبکه های عصبی و منطق فازی[۱] ابزار مکمل جهت حل پارهای از این مشکلات می‌باشند.شبکه های عصبی به صورت شبکه های غیر خطی و به هم متصل شده هستند، که جهت پردازش و تحلیل اطلاعات مبهم و مواردی که ابهام بر انگیز می‌باشد، به منظور تصمیم گیری‌های مرزی حساس، به کار می‌روند. روش های تنظیم فازی و مدل های مخفی مارکوف[۲] (HMMs) توانایی ارائه درجه درستی و یا میزان صحت و یا درجه وابستگی حروف به یکدیگر را دارند. منطق فازی یکی از چندین روش تنظیمات فازی می‌باشد، که به صورت رمز در آوردن معلومات و موارد مبهم و مشکل دار، و حفظ کردن فرضیه های گوناگون به صورت طبیعی، که نتیجه ای از تردید ها و موارد نامعلوم در مسئله ها و مشکلات حقیقی می‌باشند را بر عهده دارد. با ترکیب کردن رشته های مکمل از روش های عصبی و فازی به سمت سیستم های ترکیبی، ما می‌توانیم توانایی شناسایی بلند مرتبه ای را برای حل مساله و مشکل دست خط بدست آوریم.

---

۱ - Fuzzy Logic

۲ - Hidden Markov Models

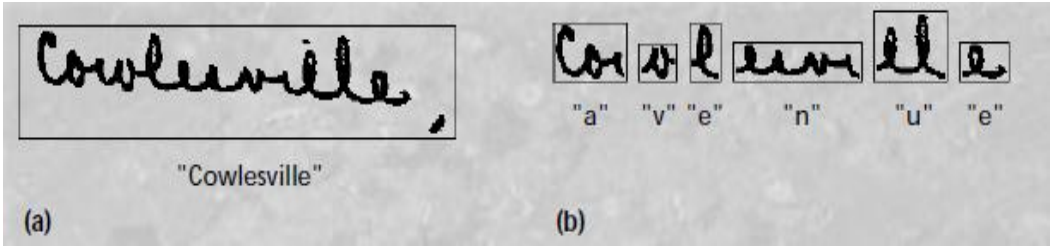
شناسایی حروف دست نوشته، در عملیات مقایسه یک تصویر دیجیتالی از یک کلمه دست نوشته با یک مجموعه کلمات شناخته شده از قبل خلاصه می‌شود. در یک تشخیص موفق، اگر چه به جداسازی معمولی حروف و تشخیص کاراکترهای مجزا و تکی اعتماد کرد، اما برای رسیدن به یک تعریف درست و به نوعی بهترین تعریف، تکنیک های تشخیص نیز باید اطلاعات ارزیابی شده ای در رابطه با لغات موجود در یک متن به ما بدهند.



شکل ۱-۱- ابهامات در تشخیص کلمات دست نوشته a .یک تکه پیکسل به تنهایی باعث مشاهده یک کاراکتر به تنهایی و یا دو کاراکتر شده، b. یک قسمت از پیکسل ها باعث آشکار سازی دو نوع کاراکتر شده

حتی آزمایش کاراکتر به وسیله کاراکتری دیگر در بحث تشخیص دست نوشته ابهام درست می‌کند،نمایش این موضوع در شکل ۱-۱ [۸] کاملا قابل مشاهده می باشد. روش های زیادی از مبحث تشخیص دست خط، یک تصویر را به تعدادی تصاویر پایه قبل از به کار بردن الگوریتم های شناسایی تقسیم می کنند.

شکل ۱-۲ [۸] نشان می دهد که قطعه بندی چگونه می تواند باعث اضافه شدن ابهامات شود. در اینجا کلمه "cowlesville" به صورت نادرست به تصاویر پایه تقسیم بندی شده، تعدادی از این کاراکتر ها هنگامی که به صورت جدا دیده شده اند محتمل وقابل قبول هستند. بنابر این یک سیستم تشخیص دست خط در حالتی که قطعه بندی به اشتباه انجام شده، بیشتر امتیاز را در شناسایی با مطابقت با مجموعه لغات از قبل فرا گرفته، به کلمه "Avenue" می دهد.



شکل ۲-۱- ابهام ایجاد شده توسط قطعه بندی. a. کلمه مورد نظر b. حتی در حالتی که کاراکترهای معقول در تصویر مشاهده می شوند، قطعه بندی نادرست منجر به تطبیق اشتباه شده

در ادامه با توجه به شکل ۱-۲ مشاهده می شود، تشخیص کاراکتر های جدا شده از یکدیگر با توجه به نادرستی قطعه بندی در قیاس کاراکتر های جدا شده با کلمات مجموعه لغات از قبل فرا گرفته شده، اطلاعات مناسبی را از کنار هم گذاشتن تصاویر کاراکتر ها به ما نمی دهد. حتی با وجود اینکه کلمات ششم و پنجم شبیه به e و u هستند از دید یک انسان کاملا قابل مشاهده می باشد که در درون کلمه دو حرف مورد نظر ترکیب مناسبی را رعایت نکرده اند، و یکی خیلی بزرگتر از دیگری است بنابراین برای داشتن یک تشخیص صحیح، احتیاج به یک چهارچوب معین و با قاعده، برای حروف قطعه بندی شده داریم.

۱-۳-۱- تشخیص دست خط عربی و فارسی

تشخیص دست خط عربی و فارسی چالش ها و مزیت های بی نظیری را ارائه می دهد و تحقیق راجع به رسیدن به این مهم، اخیرا بیشتر از تشخیص متن در دیگر متون با زبان متفاوت جلوه گر می‌شود.تشخیص دست خط عربی می تواند مطالعه خودکار متون دست نویس عربی قدیمی را فراهم سازد زیرا شیوه نوشته های عربی در طول زمان های زیاد به نسبت زبان های دیگر، تغییر اندکی داشته، و تکنیک های اینگونه جهت توسعه مطالعه متون دستی مختلف عربی کاربرد فراوان خواهد داشت. پردازش های خودکار می توانند رشد بزرگی را در مورد این مضمون فراهم آورند. زیرا دست نوشته در متون دستی، معمولا مرتب تر از دست نویس های دیگر زبان ها می باشد. اگر چه موانعی مانند، کیفیت پایین تصاویر، علامت گذاری های غیر منتظره، بعضی از روش های نوشتاری قبلی، چالش هایی را به وجود می‌آورند.

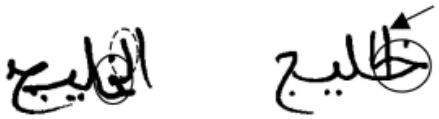
حروف زبان عربی و فارسی هر کدام بین دو تا چهار شکل متفاوت را دارا هستند و این شکل انتخابی جهت مورد استفاده شده، به موقعیت حرف در کلمه بستگی دارد،که می توانند در شروع کلمه، مابین کلمه، انتهای کلمه و یا

جدا از یکدیگر باشند.جدول شماره ۱-۱ [۸] اشکال مربوط به هر حرف و موقعیت های مختلف آن را نشان می دهد. حروف بدون شکل ابتدایی یا انتهایی، نمی توانند در حالت اتصال داده شده به ادامه دیگر حروف نمایش داده شوند. بنابراین اشکال اولیه آنها به سادگی اشکال جدا شده آنها هستند.

برای مبحث دست نویس در نظر گرفتن خط زمینه مناسب یک مفهوم ایده ال می باشد و باعث ساده سازی در یک نوشتار قانونمند می‌باشد. حروف به استثنای بعضی از آنها که حالت نزولی وصعودی دارند و بعضی از علامات خاص در زبان عربی(فتحه- کسره و...) بر روی خط زمینه قرار می‌گیرند.

اتصالات در زبان عربی جزیی از طبیعت این زبان می باشد، برخلاف زبان لاتین که به صورت طبیعی حروف جدا از یکدیگر می باشند و این موضوع خود مشکلات زیادی نظیر وجود کجی ها،انحناها و ... را به وجود می‌آورد. گاهی در نوشتار حروف در یک راستای عمودی قرار می گیرد و این موضوع امر تشخیص را سخت تر می‌کند. در شکل ۱-۳ [۸] می توانید این موضوع را به عینه مشاهده کنید.

با وجود این دلایل و موضوعاتی که مطرح شد، شناخت زبان فارسی و عربی به مراتب سخت تر از زبان لاتین می‌باشد، اما در صورت رعایت شدن بعضی از اصول می تواند نسبت به قبل ساده تر شود، نظیر: کم شدن تنوع در حالت های مختلف، میانگین کوتاه طول کلمات، تفکیک نقاط و دیگر علامات از کلمات، ایجاد تنوع قانونمند بر روی شکل حروف.



شکل ۳-۱- دو دست نوشته متفاوت از یک کلمه که کاملا روی هم افتادگی حروف را در هر دو نشان می‌دهد.

جدول۱ -۱- موقعیت هر حرف به تنهایی درون کلمه

Isolated	Initial	Medial	Final
ا	-		ـا
ب	بـ	ـبـ	ـب
ت	تـ	ـتـ	ـت
ث	ثـ	ـثـ	ـث
ج	جـ	ـجـ	ـج
ح	حـ	ـحـ	ـح
خ	خـ	ـخـ	ـخ
د	-		ـد
ذ	-		ـذ
ر	-		ـر
ز	-		ـز
س	سـ	ـسـ	ـس
ش	شـ	ـشـ	ـش
ص	صـ	ـصـ	ـص
ض	ضـ	ـضـ	ـض
ط	طـ	ـطـ	ـط
ظ	ظـ	ـظـ	ـظ
ع	عـ	ـعـ	ـع
غ	غـ	ـغـ	ـغ
ف	فـ	ـفـ	ـف
ق	قـ	ـقـ	ـق
ك	كـ	ـكـ	ـك
ل	لـ	ـلـ	ـل
م	مـ	ـمـ	ـم
ن	نـ	ـنـ	ـن
ه	هـ	ـهـ	ـه
و	-		ـو
ي	يـ	ـيـ	ـي

۱-۳-۲- روش‌های تشخیص دستخط

تشخیص دست خط عموما و به صورت مرسوم به دو روش

- تشخیص هم زمان[۱]
- و تشخیص غیر هم زمان[۲]

---

۱ - on line

۲ - off line