دسته بندی پاکتهای نامه با استفاده از بینایی ماشین در خواندن و پردازش آدرسهای دستنویس

سید علیرضا سیدین°، نرگس طباطبائی مشهدی و سید محمد مهدی سیدین

°. دانشگاه فردوسی مشهد، گروه برق

seyedin@ferdowsi.um.ac.ir, ntabatabaey@gmail.com, mahdi.seyedin@gmail.com

چکیده

دستگاه های طبقه بندی نامه، سامانه هایی هستند که برای دسته بندی نامه ها، بر اساس آدرس مقصد، در ادارات پست مورد استفاده قرار می گیرند. در این پژوهش تلاش کردیم سامانه ای طراحی و ساخته شود که نامه های ورودی به آن با توجه به نام شهر در آدرس دستنویسِ گیرنده، طبقه بندی شوند. این سامانه نمونه ای اولیه از یک دستگاه کاربردی دسته بندی پاکتهای نامه است که برای اداره پست مرکزی استان خراسان رضوی طراحی شده است. داده های تصویری این پژوهش کاملا اصیل بوده و طی چند مرحله تصویر برداری از پاکتهای نامه موجود در مرکز پست، تهیه شده اند. دسته بندی داده ها با استفاده از ماشینهای بردار تکیه گاه (*SVM*) انجام شد که نتایج خوبی را به ارمغان آورد. داده های ورودی به *SVM* تصاویر دستنویس فارسی مربوط به نامِ شهرِ مقصد بودند. برای اجرای طرح پیش فرضهایی در رابطه با رعایت اصول اولیه در آدرس نویسی، الصاق تمبر، و عدم درج آگهی بر روی پاکت نامه لحاظ شده است. در این مقاله، جداسازی آدرس از تصویر پاکت نامه و تشخیص کلمات دستنویس توضیح داده می شود. این پژوهش مبتنی بر امکانات موجود در داخل کشور، سیستم کارآمدی را برای دسته بندی خودکار نامه ها ارائه می دهد.

کلمات کلیدی- دستگاه طبقه بندی نامه (*LSM*) ، ماشینهای بردار تکیه گاه (*SVM*)، تشخیص نوری نویسه ها (*OCR*)، تشخیص آدرس دستنویس.

۱- مقدمه

مراکز پست برای تفکیک نامه ها پس ازجمع آوری، غربالگری، مرتب کردن و ابطال تمبر، آنها را بر اساس آدرس مقصد، طبقه بندی کرده و به قفسه های مربوطه منتقل می کنند. این فرآیند، روندی طولانی، زمانبر و خسته کننده دارد که در عین حال با توجه به حجم بالای مرسولات، این فرآیند نیازمند سرعت ودقت زیاد است. اگر تمام فرآیند طبقه بندی نامه ها به صورت سنتی و توسط نیروی انسانی صورت گیرد باعث می شود که دقت و سرعت این فرآیند تابعی از وضعیت جسمی و روحی نیروی انسانی بوده و در نهایت باعث پایین آمدن کیفیت فرآیند و استفاده ی غیر بهینه از نیروی انسانی شود. در بسیاری از کشورهای جهان، این فرآیند بصورت خودکار و مکانیزه انجام می شود.

دستگاه های طبقه بندی نامه(Letter Sorting Machines) سامانه هایی خودکار هستند که برای طبقه بندی نامه ها در ادارات پست مورد استفاده قرار می گیرند. این دستگاه ها در

حالت کلی از بخشهای مختلفی تشکلیل شده اند. این بخشها شامل جمع آوری، ابطال تمبر، تشخیص مقصد، کُد گذار، کُد خوان، بخش هدایتی، و بخش های مکانیکی می باشند که نهایتا نامه را، بطور خودکار، در صندوق مقصدِ مربوطه جای می دهند. معمولا این دستگاه ها بر اساس بودجه، موقعیت مکانی، نیازِ سرعتی، حجم مرسولات و بطور کلی نیاز مشتری ساخته و تهیه می شوند. دستگاه هایی نیز موجود می باشند که یک یا بخشی از فعالیتهای فوق را انجام می دهند و بقیه آن عملیات توسط نیروی انسانی انجام می شود. در این مقاله ما به بخش کلیدی، یعنی تشخیص مقصد، با جداسازی آدرس از تصویر پاکت نامه و تشخیص کلمات دستنویس، خواهیم پرداخت.

عملیات خودکار سازی روند پردازش نامه ها، خود پیچیدگیهای خاصی دارد. رعایت نشدن قواعد مربوط به فضاهای اختصاص داده شده روی پاکت نامه ها برای جای تمبر، آدرس گیرنده و فرستنده؛ همچنین ناخوانا بودن برخی دست خطها و یا بطور کلی دست خطهای به هم پیوسته در زبان انگلیسی، و بغرنج تر از آن دستخطهای تو در توی فارسی و انواع نوشتارهای نستعلیق و شکسته و غیره، همه از جمله مواردی هستند که عملیات پردازش خودکار را پیچیده می کنند.

تا به حال تلاشهای متعددی در زمینه تشخیص متون فارسی و عربی انجام شده است که عمده موفقیتها در زمینه تشخیص متون تایپی و چاپی با فونتهای مختلف بوده است و نتیجه موثر آنها ایجاد چند نرم افزار OCR فارسی و عربی می باشد. بطور کلی در مقایسه با سیستمهای تشخیص متون تایپی، سیستمهای تشخیص دستخطهای پیوسته -به خصوص نوع غیرهمزمان آن(offline)- بسیار پیچیده و با کارآمدیهای پایینتری می باشند. سیستمهای تشخیصی برای دستخطهای زبانهایی مانند عربی و فارسی نیز بنا به شیوه نوشتاری این زبانها، جزء دسته مسائل پیچیده و سخت قرار دارند[۱]. در زمینه تشخیص آدرس و کدپستی دستنویس به زبان فارسی نیز تلاشهایی انجام شده است. برخی پژوهشها، بر اساس تشخیص اعداد فارسی، سیستمهایی برای تشخیص کدپستی طراحی کرده اند.

پیچیدگی و تنوع دستخطهای فارسی باعث شده است که تا کنون نتوان یک سیستم عام در مورد تشخیص دست نوشته ها ارائه داد و تلاشهای انجام شده با نتایج مناسب، تا کنون عمدتا روی حوزه های محدود انجام شده است. معمولا در این نوع پژوهشها تلاش بر ساده کردن مساله است و بر این اساس معولا به دنبال استخراج ویژگیهای موثر نویسه ها برای دسته بندی و تشخیص هستیم؛ در این زمینه تا به حال ویژگیهای مختلفی استخراج شده و بااستفاده از دسته بندی کننده های متفاوتی دسته بندی شده اند[۱]. برخی از روشها تشخیص را بر اساس تشخیص حروف و برخی بر اساس تشخیص کل کلمه انجام می دهند که معمولا وقتی مجموعه کلمات ورودی سیستم محدود باشد (مثلا در اینجا نام شهرهای مقصد)، روش تشخیص کل کلمه، عملی و اجرایی خواهد بود. در برخی پژوهشها از ویولت برای استخراج ویژگی کلمه استفاده می شود[۲]. در[۳] برای خواندن نام شهرهای آدرسهای پستی از ویژگیهایی مانند منحنی محیط کلمه و هستوگرام مربوط به فریمهای خاصِ تعریف شده، استفاده شده است و سپس با استفاده از سیستم های فازی و مدل مارکوف دسته بندی انجام گرفته است. در این مقاله ما پیش پردازشها را کاهش داده و ویژگی اصلی و ممیزی را همان شکل ظاهری کلمه در نظر گرفته ایم و با استفاده از ماشینهای بردار تکیه گاه دسته بندی را انجام می دهیم.

در میان این تلاشهای جهانی، ما نیز در این طرح علی رغم پیچیدگی های ساختاری، در سدد ایجاد یک نمونه اولیه از سیستم پردازش و طبقه بندی نامه ها بر اساس داده ها و امکانات موجود در داخل کشور هستیم. این سامانه اولیه برای مرکز پست خراسان رضوی و برای دسته بندی نامه های این استان طراحی شده است. برای اجرای این طرح، پیش فرضهایی در رابطه با رعایت اصول اولیه در آدرس نویسی، الصاق تمبر، و عدم درج آگهی بر روی پاکت نامه لحاظ شده است. در این پژوهش تلاش می کنیم تا با استفاده از پردازش تصویر، آدرس، و سپس نام شهر مقصد را از پاکت نامه استخراج کرده و با استفاده از ماشینهای بردار تکیه گاه تصویرِ نام شهر را تشخیص داده و سپس دسته بندی کنیم.

در ادامه به بخش های تهیه و تنظیم داده های تصویری، جداسازی کلمات آدرس، تشخیص کلمات دستنویس و در پایان نتیجه گیری خواهیم پرداخت.

۲- تهیه و تنظیم داده های تصویری

از آنجایی که این سامانه برای تشخیص شهرهای استان خراسان رضوی طراحی می شود، تصاویر تهیه شده مربوط به شانزده شهرِ این استان می باشند (به جز شهر مشهد). بنا بر این تصاویر پاکتهای نامه مربوط به این شهر ها، با استفاده از دستگاه آزمایشی تصویر برداری، که به این منظور تهیه شده بود در اداره پست مرکزی استان خراسان رضوی تصویر برداری شد. این عملیات طی چند روز پی یا پی انجام شد و در نتیجه حجم زیادی از داده های تصویری بدست آمد.

دستگاه آزمایشی تصویر برداری، شامل چهار دوربین است که تصویر پشت و روی هر پاکت نامه را مشابه نمونه شکل ۱ تصویر برداری می کند. همان طور که مشاهده می شود هر طرف نامه توسط دو دور بین عکس برداری شده که تصویر این دو دوربین نیز هم پوشانی دارند. البته بنا به ایده اولیه این نوع تصویر برداری، اصولا آدرس گیرنده باید در پایین سمت چپ پشت پاکت

نامه نوشته شود؛ که در این وضعیت (فرای نحوه قرار گیری پاکت نامه در جایگاه عکسبرداری که وارونه باشد یا نه) آدرس بطور کامل توسط یکی از چهار دوربین قابل رویت خواهد بود.

شکل ۱: تصویر تهیه شده از پاکت نامه توسط دستگاه تصویربرداری آزمایشی (مقیاس ۴۰٪)

به این ترتیب تصاویر پاکتهای نامه مربوط به شهرهای بردسکن، چناران، درگز، فریمان، فرودس، قوچان، گناباد، کاشمر، خواف، نیشابور، سبزوار، سرخس، تایباد، تربت جام، تربت حیدریه و کلات تهیه و ذخیره سازی شدند.

۳- جدا سازی آدرس از تصویر پاکت نامه

در این مرحله ابتدا تصاویر را ضمن آشکار سازی لبه، باینری می نماییم. به دلیل باریک بودن نوشته ها لبه های آشکار شده به نوعی خود دست نوشته ها هستند، که این نتیجه به ما در تسریع پردازش ها کمک می نماید. سپس با انجام عملیات حذف نویز، تصویر برای انجام پردازش های بعدی آماده می شود.

در ادامه این بخش مباحث تشخیص جهت نامه و موقعیت آدرس گیرنده، جداسازی کلمات آدرس و جداسازی نام شهر مقصد از بین بقیه کلمات، شرح داده خواهند شد.

۳-۱- تشخیص جهت نامه و موقعیت آدرس

با توجه به اینکه در نوشتار فارسی تعداد زوایای ۹۰ درجه بیش از ۲۷۰ درجه می باشد، با اعمال دو ماسکِ به نمایش در آمده در شکل۲، بر تصویر باینریِ کامل نامه، می توان تعداد زاویه های ۹۰ و ۲۷۰ درجه را در خطوط دست نوشته محاسبه کرد و بسته به این که تعداد زوایای ۹۰ درجه بیشتر است و یا ۲۷۰ درجه می توان به ترتیب نتیجه گرفت که نامه در جهت صحیح قرار گرفته و یا نیاز به اصلاح دارد. در صورت نیاز به اصلاح، تصویر را ۱۸۰ درجه می چرخانیم.

پس از آن، با محاسبه چگالی پیکسل های مقدار دار در نیمه ی بالا و پایین تصویر، روی نامه تشخیص داده می شود. با فرض رعایت شدن استاندارد های مربوط به نامه نگاری، آدرس گیرنده در نیمه راست آن بخش (روی نامه) قرار خواهد داشت (ربع پایین سمت راست و یا بالا سمت راست نامه ی اصلاح شده).

۱	۰	۰
۱	۰	۰
۱	۱	۱

(الف)

۱	۱	۱
۱	۰	۰
۱	۰	۰

(ب)

شکل ۲: الف)برای زوایای ۹۰ درجه. ب)برای زوایای۲۷۰ درجه.

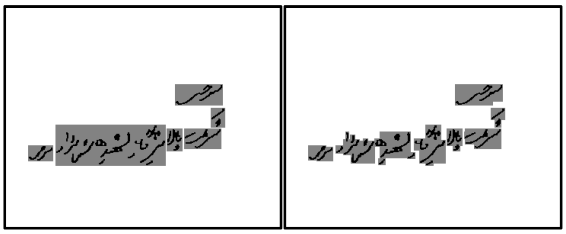
۳-۲- جداسازی کلمات آدرس

پس از تشخیص موقعیت آدرس گیرنده، در این مرحله، با توجه به توضیحاتی که در ابتدای این بخش داده شد، پیکسل های مقدارداری که در مجاورت هم قرار دارند تشخیص داده شده و برای هر یک از این مجموعه پیکسل ها ناحیه ای مستطیل شکل تعریف می شود (کوچکترین ناحیه ی ممکن که به طور کامل هر یک از مجموعه ها را دربر بگیرد). سپس هر دو ناحیه ای که دارای همپوشانی باشند یکی می شوند. این عمل تا زمانی ادامه می یابد که دیگر هیچ دو مجموعه ی دارای همپوشانی باقی نماند. در نهایت هر یک از مجموعه های ایجاد شده به عنوان یک کلمه در نظر گرفته می شود.

۳-۳- جداسازی نام شهر مقصد از بقیه کلمات

در این مرحله بر اساس استانداردهایی که انتظار می رود رعایت شده باشند نام شهر مقصد را تشخیص می دهیم. در اینجا فرض شده است که نام شهر مقصد اولین کلمه ی آدرس باشد.

نمونه ای از مراحل توضیح داده شده در شکل ۳ به نمایش در آمده است.



(۳-الف) (۳-ب)

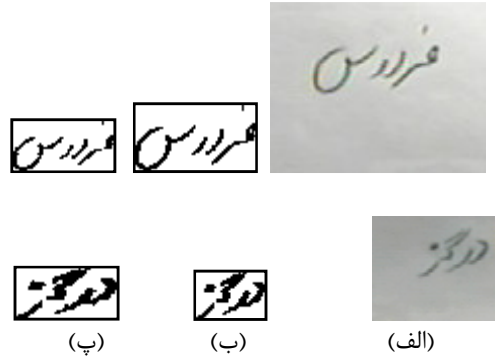
(۳-پ)

شکل ۳: مراحل جداسازی نام شهر مقصد. در تصویر (الف) شناسایی خطوط پیوسته و تشکیل نواحی مستطیل شکل و در تصویر (ب) ناحیه های دارای همپوشانی یکی می شوند و در نهایت در قسمت (پ) بالاترین ناحیه به عنوان نام شهر مقصد در نظر گرفته خواهد شد. (مربوط به شهر سرخس)

۴- تشخیص کلمات دستنویس

برای آموزش سیستم اولیه، از تعداد محدودی از تصاویر تهیه شده، بطور دستی، کلمه شهرِ مقصد جداسازی و ذخیره شد (شکل۴، تصویر اصلی). برای آماده سازی داده ها برای پردازش، ابتدا تصاویر ذخیره شده بازبینی شد تا مطمئن شویم که داخل هر تصویر فقط اسم شهر نوشته شده و خطوط اضافی دیگر (به جز خط افق نوشتاری در برخی پاکتهای نامه استاندارد)، حذف شده باشند. در این مرحله برخی خطوط اضافی در تصویر که مربوط به مُهرِ روی پاکت می شدند و یا به علت تو در تو نوشتن آدرس، خطوط اضافی که مربوط به کلمات قبل یا بعد از کلمه شهر مقصد می شدند نیز به شکل دستی حذف شدند. به این ترتیب مجموعه تصاویر برای انجام پیش پردازشها آماده شد. کل تصاویری که از تمام ۱۶ شهر به این شکل تهیه و در دسته بندی استفاده شدند، برابر ۷۷۵ تصویر می باشد.

در اولین مرحله پردازش، تصویر دستخط، به دو سطح سیاه و سفید تقلیل داده شد. به عبارت دیگر تصاویر باینری شدند. برای باینری سازی از روش استاندارد "اوتسا" استفاده شد[۴]. در این روش آستانه باینری طوری انتخاب می شود که واریانسِ درونی هر دسته سیاه و سفید، کمینه شود. در نتیجه ی این عمل پس زمینه سفید، و دستخط با پیکسلهای سیاه نمایش داده می شود. پس از این مرحله، تصویر به ابعاد دستخط مماس می شود بطوریکه ابعاد مستطیلِ تصویر همان ابعاد طول و عرض نوشتارِ شهر را نشان می دهد. قدم بعدی نرمال سازی اندازه های تصویر می باشد. این بدان معنی است که تمام تصاویر براساس درون یابی استاندارد bicubic به ابعاد ۲۵×۵۰ برگردانده می شوند تا به این ترتیب تصاویر نسبت به یکدیگر قابل قیاس شوند. برای نمونه، تصویر دو دستخط در مراحل مختلفِ ذکر شده، در شکل۴ نشان داده شده است.



(الف) (ب) (پ)

شکل ۴: مراحل پردازش تصویر ورودی به SVM، الف: تصویر اصلی، ب: تصویر باینری با مرز مماس پ: تصویر نرمال شده

آنچه مسلم است استخراج ویژگی از دستنوشته ها بطوریکه مستقل از نوع نوشتار افراد، تنها کلمه نوشته شده را استخراج کند، با توجه به تنوع دستخطها و خصوصیات ویژه مربوط به خط فارسی، کاری بسیار دشوار است. این پیچیدگی برای یافتن یک خصوصیات عام در تمام دستخطهای مختلف طوری است که می توان گفت بهترین ویژگی همان شکل کلی کلمه است که پارسی زبانان خود نیز در بسیاری از موارد دستخطهای ناخوانا را بر اساس آن خوانده و تشخیص می دهند.

با توجه به آنکه در این مقاله از دسته بندی کننده SVM (ماشینهای بردار تکیه گاه) برای دسته بندی الگوها استفاده می کنیم، محدودیت زیادی روی اندازه بعد بردار ویژگی نداریم؛ چراکه SVM با ورودیهای بُعد بالا نیز به راحتی آموزش می بیند و مشکل خاصی برای عملکرد آن ایجاد نمی شود. از این رو در این پژوهش از عناصر خود ماتریس تصویر برای ساخت بردار ویژگی استفاده شده است. در پژوهشهای پیشین نیز نشان داده شده است که استفاده از ماتریس تصویر در دسته بندی تصاویر پیچیده کروموزمی و اجسام سه بعدی با استفاده از ماشینهای بردار تکیه گاه نتایج بسیار خوبی به همراه داشته است[۵, ۶]. به این ترتیب، با استفاده از تمام ویژگیهای در دسترس-یعنی پیکسلهای تصویر که کل اطلاعات را شامل می شدند- بردار ویژگی از کنار هم قرار دادن سطرهای ماتریس تصویر، ساخته شد.

در ادامه ماشینهای بردار تکیه گاه و نحوه استفاده از آن در این مقاله، همچنین آزمایشات انجام شده روی داده ها توضیح داده خواهند شد. در پایان این بخش نیز نتایج دسته بندی مورد بررسی قرار می گیرد.

۴-۱- ماشینهای بردار تکیه گاه

ماشینهای بردار تکیه گاه (Support Vector Machines)، یک روش تشخیص الگوست. این روش در قالب "حداقل سازی مخاطره ساختاری (Structural Risk Minimization)" و تئوری آماری فراگیری ارائه می شود. روشهای کلاسیک تشخیص الگو اکثرا بر اساس حداقل کردن مقدار قدر مطلق خطا روی داده های آموزشی، و یا حداقل کردن مجذور خطا روی این داده ها، تابع مناسب تفکیک گر را بدست می آوردند؛ ولی SVM تلاش می کند خطا را به شکل ساختاری حداقل کند. به این ترتیب خطا نه فقط روی الگوهای آموزشی بلکه روی الگوهای آزمایشی نیز حداقل خواهد شد. این همان علتی است که قابلیت تعمیم یافتگی خوب را برای SVM به ارمغان می آورد[۷, ۸].

پارامترهای انتخابی برای ساخت یک دسته بندی کننده SVM عبارتند از انتخاب یک کرنل(kernel) مناسب و تعیین پارامتر C (ضریب تاثیر خطا) . انتخاب کرنل در عملکرد SVM بسیار موثر است. با استفاده از تابع کرنل حجم محاسباتی که برای بدست آوردن ضرب داخلی در فضای ویژگی با ابعاد بزرگ، لازم

است، با حجم محاسباتی کمتر در فضای ورودی، جایگزین میشود. یکی از کرنلهای مشهور کرنل گوسی است. رابطه(۱) کرنل گوسی را نشان می دهد که مربوط به ضرب داخلی در فضای ویژگیِ دو بردار $\mathbf{u}$ و $\mathbf{v}$ می باشد (γ پارامتر کرنل گوسی است).

$$K(\mathbf{u},\mathbf{v}) = \exp(-\gamma(\|\mathbf{u}-\mathbf{v}\|^2) \quad (۱)$$

این کرنل در اغلب مواردی که استفاده شده است در مقابل کرنلهای مشهور دیگر بهترین نتیجه را ارائه داده است. در این مقاله نیز از این کرنل در SVM استفاده می کنیم. پارامتر کرنل در این پژوهش از اجرای محدودِ ژنتیک الگوریتم بدست آمده است ، و C=100 انتخاب شده است.

۴-۲- آزمایشات و نتایج آنها

برای دسته بندی تصاویر دستنویس، آزمایشات به دو صورت انجام شد. در روش اول تمام داده های آزمایشی به یکی از ۱۶ دسته موجود، مربوط به یکی از شهرها، اختصاص داده می شود؛ اما در آزمایش دوم نتیجه تفکیکگرها ابتدا بررسی شده و با یک آستانه مقایسه می شود، اگر شرایط آستانه برآورده شود آنگاه نتیجه دسته بندی معتبر بوده و تصویر ورودی به یکی از ۱۶ دسته اختصاص داده می شود، در غیر این صورت تصویر ورودی جزء دسته جدیدی به نام "نامعلوم" دسته بندی می شود. در عمل الگوهای این دسته از سیستم تشخیص خودکار خارج شده و بطور دستی تشخیص داده خواهند شد.

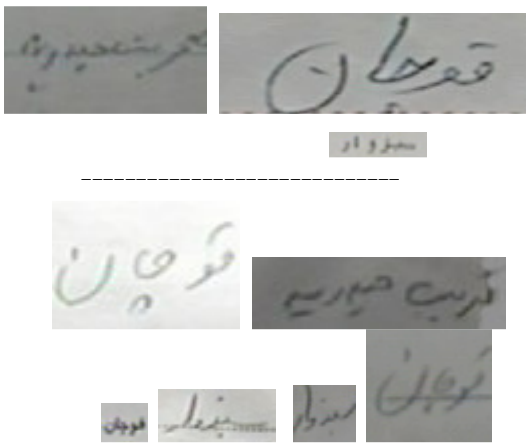
برای انجام آزمایشات یک چهارم داده ها از هر دسته، به شکل اتفاقی انتخاب شدند و به عنوان الگوهای آزمایشی مورد استفاده قرار گرفتند. سه چهارم، الگوهای باقی مانده از هر گروه به عنوان الگوهای آموزشی استفاده شدند.

آزمایش اول: در این مدل، تعداد دسته های خروجی برابر تعداد کدِ شهرهای مختلف است. این به معنی است که شانزده خروجی مختلف می تواند در خروجی دسته بندی کننده ظاهر شود که هر یک نشانگر یک کد شهر می باشد. بنابراین هر کُدی که به ازای یک الگوی ورودی در خروجی ظاهر شود به این معنی است که الگوی ورودی به آن نسبت داده شده است. به این ترتیب در مرحله آزمایش، الگوی ورودی در تمام این شانزده تفکیک گر باینری امتحان می شود؛ و سپس مقدار خروجی هر یک بیشتر از بقیه بود، به این معنی است که الگوی ورودی به آن دسته شبیه تر بوده و بنابراین جزء آن گروه دسته بندی می شود. نتایج اجرای این روش در جدول(۱) ارائه شده است. حدود ۷۶٪ از داده های آزمایشی درست تشخیص داده شده است.

جدول ۱: نتایج آزمایش اول با لحاظ شانزده کد خروجی

دسته بندی کننده	درصد تشخیص درست	درصد خطا
SVM	76.3158	23.6842

نمونه ای از الگوهایی که نادرست تشخیص داده شده اند به همراه نمونه هایی از الگوهای درست تشخیص داده شده مربوط به دسته های مختلف، در شکل۵ نشان داده شده اند.



شکل ۵: نمونه های درست تشخیص داده شده در پایین خطچین و نمونه های غلط تشخیص داده شده در بالای خطچین مربوط به آزمایش اول

در این روش هیچ الگویی وجود ندارد که تشخیص داده نشود. به عبارت دیگر بالاخره ورودی –بنا به مقدار بیشینه- به یکی از دسته ها، شبیه تر تشخیص داده می شود هرچند که تفاوت معنی داری بین مقدار بیشینه خروجی SVM باینری با بقیه مقادیر نباشد. نهایتا الگوی ورودی یا درست تشخیص داده می شود و یا جزء خطاها خواهد بود.

آزمایش دوم: در این روش نیز عملیات آموزش و آزمایش مشابه مدل قبل انجام می شود، تنها تفاوت این روش با روش قبل در هنگام تصمیم گیری نهایی و اعلام نتیجه دسته بندی به خروجی است. در این روش نتیجه خروجی معتبر ارزیابی نمی شود مگر آنکه مقدار آستانه معینی را برآورده کند. بنابراین اگر مقدار آستانه برآورده شد، نتیجه دسته بندی اعلام شده و ورودی به یکی از ۱۶ دسته خروجی نسبت داده می شود؛ اما اگر این مقدار برآورده نشود آنگاه الگوی ورودی غیر قابل تشخیص اعلام شده و جزء داده های تشخیص داده نشده از سیستمِ خودکار خارج می شود.

پارامتر آستانه ای که در این فازِ آزمایش مورد بررسی قرار گرفت و سیستم روی آن آزمایش شد، به شکل رابطه(۲) تعریف می شود که در آن، خروجیِ تفکیکگرهایِ باینری، همان مقادیر خروجی از شانزده SVM باینری هستند.

(۲)

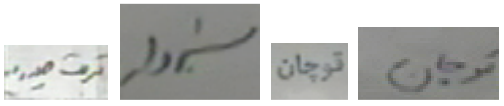
$$\frac{abs\left(max\left(\text{خروجیهای تفکیکگرهای باینری}\right)-mean\left(\text{خروجیهای تفکیکگرهای باینری}\right)\right)}{std\left(\text{خروجیهای تفکیکگرهای باینری}\right)}$$

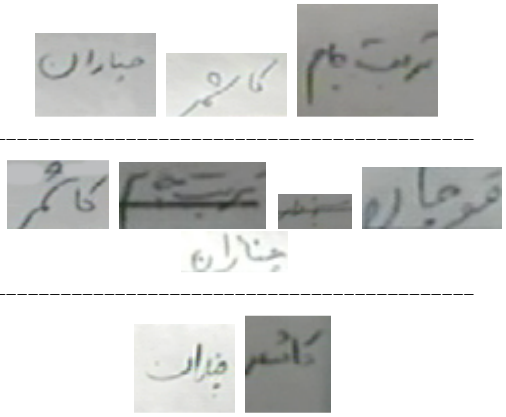
همان گونه که پیشتر بیان شد، مقدار ماکزیمم خروجیهای SVMهای باینری همان مقداری است که در شرایط کلی و بدون لحاظ آستانه، دسته خروجی بر اساس آن تعیین می شود. بر این اساس مقدار پارامتر آستانه برابر است با مقدار استاندارد شده همان خروجی. بنابراین، این مقدار هرچه از صفر دورتر باشد نشان دهنده آن خواهد بود که خروجی ماکزیمم، تفاوتِ محسوس تر و قابل توجه تری از بقیه خروجی ها داشته است و در نتیجه خروجی، معتبر تر است. همچنین اگر این مقدار نزدیک به صفر باشد به این معنی خواهد بود که تفاوت چندانی بین خروجیهای بدست آمده نبوده است و بنابراین دسته بندی چندان معتبر نیست (هنگامی که تعداد داده ها به سمت بی نهایت تعمیم داده شود، توزیع داده های استاندارد شده نیز به سمت نرمالِ استاندارد با میانگین صفر و واریانس یک تمایل خواهد شد). به این ترتیب پارامتر آستانه برای هر الگو محاسبه می شود و اگر این مقدار بزرگتر از حد مد نظر بود، آنگاه خروجی معتبر اعلام می شود و در غیر این صورت ورودی تشخیص داده نخواهد شد. با این تعریف، هنگامی که آستانه برابر ۲ در نظر گرفته شد، درصد الگو هایی که تشخیص داده می شدند حدود ۵۰٪ بود که از این مقدار حدود ۹۵٪ درست دسته بندی می شدند. حدود پنجاه درصد از الگوها نیز اصلا تشخیص داده نمی شدند. هنگامی که آستانه را به ۲.۵ تغییر دادیم درصد داده های تشخیص داده شده به حدود سی درصد کاهش یافت اما این بار از این مقدار، حدود ۹۸٪ درست تشخیص داده شدند. نتایج این آزمایشات در جدول(۲) نشان داده شده است. مشاهده می شود که با افزایش آستانه، درصد داده های تشخیص داده نشده زیاد شده است و البته همزمان درصد تشخیص درست نیز بالا رفته است.

جدول ۲: نتایج آموزش سیستم با شانزده دسته خروجی و یک دسته نامعلوم

مقدار آستانه	2.0	2.5
درصدِ تشخیص داده شده	50.5263	30.5263
درصدِ تشخیصِ درست / درصدِ خطا	94.7917 / 5.2083	98.2759 / 1.7241
درصدِ تشخیص داده نشده	49.4737	69.4737

نمونه هایی از الگوهای تشخیص داده نشده، تشخیص داده شده به شکل صحیح، و تشخیص داده شده به شکل خطا با استفاده از این روش، در شکل۶ نشان داده شده است.





شکل ۶: نمونه هایی از الگوهای صحیح تشخیص داده شده(بالا)، تشخیص داده نشده(وسط)، و خطا(پایین)، با آستانهٔ ۲ به بالا.

۴-۳- بحث و بررسی نتایج دسته بندی

در آزمایش اول (جدول(۱)) درصد تشخیص ۷۶٪، بنا به حجم داده های استفاده شده و کیفیت آنها، خوب ارزیابی می شود. لازم به ذکر است که در برخی تصاویر به علت عدم نورپردازی مناسب در هنگام تصویر برداری، پس از باینری سازی تصویر، بخشی از سایه موجود در تصویر به عنوان نوشتار و با رنگ سیاه نشان داده شده است که البته برای لحاظ تمام این نقصها در سیستم، این داده ها حذف نشدند و به همان شکل مورد استفاده قرار گرفتند.

از آنجایی که تعداد داده های تهیه شده از هر دسته یکسان نبودند و از برخی گروه ها داده های بیشتری در آزمایش حضور داشتند، لازم به ذکر است که در بررسی ریز تر نتایج، به وضوح مشاهده می شد که دسته هایی که داده های بیشتری از آنها در آزمایش استفاده شده است درصد تشخیص بالاتری هم داشتند و اشتباهات برای آن دسته ها کمتر بود مثلا بیشترین داده مربوط به نیشابور و سبزوار بود که به ترتیب از ۲۵ و ۲۱ داده آزمایشی آنها تنها یک خطا ایجاد شده بود. بنابراین با کامل کردن پایگاه داده، بر اساس خروجی بخش اول، نتایج به شکل قابل توجهی بهبود خواهد یافت. به علاوه، کم تعداد بودن داده ها باعث کم بودن تنوع دستخطها شده و در مواردی مانند "کاشمر" باعث می شود که سیستم برای نحوه خاصی از نوشتار آموزش ببیند؛ مثلا کاشمر را با شین بدون دندانه و به شکل شکسته یاد بگیرد.

ایجاد یک دسته جدید به شکلی که در آزمایش دوم توضیح داده شد، باعث می شود که اصولا درصد تشخیص صحیح در الگوهای تشخیص داده شده، افزایش یابد و به این ترتیب قابلیت اطمینانِ سیستمِ خودکار نیز افزایش خواهد یافت. در این روش یک دسته جدید علاوه بر تعداد گروه های مربوط به دسته های مختلفِ داده، بطور مجزا ایجاد می شود. این دسته مربوط به الگوهایی است که به علت نویزی بودن شدید و یا نامفهوم بودن برای دسته بندی کننده، نمی توانند توسط آن با درصد اطمینان

بالایی دسته بندی شوند. احتمال آنکه اینگونه الگوها به شکل خطا دسته بندی شوند بالا است، بنابراین سیستم به جای آنکه ریسک کرده و آنها را دسته بندی کند، آنها را به عنوان الگوهای غیر قابل تشخیص از مسیر خودکار خارج می کند تا جداگانه به شکل دستی، دسته بندی شوند. با اعمال آستانه، تفاوت و پراکندگی الگوهایی که درست دسته بندی می شوند ظاهرا کمتر می شود و بسیاری از الگوهای نزدیک مرز که در روش بدون آستانه، درست دسته بندی می شدند و یا جزء خطاها بودند، اینجا جزء گروه غیر قابل تشخیص قرار می گیرند. این مفهوم با مقایسه شکلهای ۵ و ۶، تاحدی قابل مشاهده است.

در بسیاری موارد، درصد تشخیص صحیح بالا برای الگوهایی که تشخیص داده می شوند آنچنان مهم و حیاتی می شود که آستانه را طوری تعیین می کنند که این درصد نزدیک به ۱۰۰ باشد؛ هرچند که مقدار الگوهایی که کلا تشخیص داده می شوند به ۳۰ تا ۴۰ درصد کاهش یابد. با این تفسیر نتیجه بدست آمده در آزمایش دوم که در آن ۳۰ درصد داده ها تا ۹۸٪ صحیح تشخیص داده شده اند، ارزشمند است هرچند که تلاش برای افزایش آنها و کارآمد تر شدن کل سیستم ادامه دارد.

۵- نتیجه‌گیری

آنچه در این مقاله ارائه شد، طراحی و پیاده سازی موفقیت آمیز الگوریتم تشخیص کلمات دستنویسِ نام شهر با استفاده از جداسازی کلمات نام شهر از تصویر اخذ شده توسط چهار دوربین از پاکت نامه و بکارگیری ماشین بردار تکیه گاه بعنوان دسته بندی کننده بود. با توجه به قابلیت روش، در صورت تکمیل بانک اطلاعاتی و استفاده از آن در آموزش سیستم، نتایج بهبود خواهند یافت. با فرض اینکه نوشتن آدرس در پاکت نامه ها بصورت استاندارد انجام شده باشد و به تبع آن نام شهر بطور کامل بعنوان ورودی به بخش تشخیص ارائه گردد، توانستیم به درصد تشخیص حدود ۷۶٪ برسیم که با اعمال آستانه، درصد تشخیص صحیح به ۹۸٪ نیز رسید.

همانگونه که پیشتر اشاره شد در راستای اجرایی شدن این پروژه مشکلاتی از قبیل عدم رعایت حداقل استانداردهای آدرس نویسی چه از نظر محل نوشتن آدرس و یا استفاده از پاکتهای نامه غیراستاندارد و نامناسب کار را پیچیده و گاه غیرعملی و غیراقتصادی می کند. همچنین، مشکل تو در تو نویسی در زبان فارسی باعث می شود که امکان تفکیک کلمات از هم، پیچیده شود و بنابراین اگر نام شهر در یک سطر مجزا نوشته نشده باشد باید از راههای دیگری برای استخراج آن استفاده کرد.

در ادامه این پژوهش، دو بخش جدا سازی آدرس و تشخیص با یک سیستم میکانیکی و کنترلی (در حال تکمیل است) با هم ادغام خواهند شد. همچنین بدست آوردن یک آستانه مناسب تر که علاوه بر بالا بردن درصد تشخیص صحیح، درصد الگوهای تشخیص نداده را کاهش دهد نیز یکی از افقهای مد نظر این پروژه است.

سپاسگزاری

بدینوسیله از اداره پست مرکزی و مرکز تحقیقات شرکت مخابرات استان خراسان رضوی برای پشتیبانی از این پروژه تشکر می گردد.

مراجع

[۱] L.M. Lorigo, V. Govindaraju, "Offline Arabic handwriting recognition: a survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 28, Issue: 5, pp.712-724, 2006.

[۲] A. Broumandniaa, J. Shanbehzadehb and M. Rezakhah Varnoosfaderanic, " Persian/Arabic handwritten word recognition using M-band packet wavelet transform," *Image and Vision Computing*, vol. 26, Issue:6, pp. 829-842, 2008.

[۳] M. Dehghan, K. Faez, M. Ahmadi, M. Shridhar, "Off-line unconstrained Farsi handwritten word recognition using fuzzy vector quantization and hidden Markov word models," Proceedings. 15th International Conference on Pattern Recognition, vol. 2, pp.351-354, 2000.

[۴] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, No. 1, pp. 62-66, 1979.

[۵] N. Tabatabaey-Mashadi and S. A. Seyedin, "Direct classification of human G-banded Chromosome images using support vector machines," presented at International Symposium on Signal Processing and its Applications (ISSPA), Sharjah, United Arab Emirates (U.A.E.), 2007.

[۶] M. Pontil and A. Verri, "Support vector machines for 3-d object recognition," *IEEE Transactions on Pattern Analysis and Machines Intelligence*, pp. 637-646, 1998.

[۷] C. Cortes and V. Vapnik, "Support vector network," *Machine Learning*, vol. 20, pp. 1-25, 1995.

[۸] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, pp. 121-167, June 1998.