بازشناسی برخط حروف مجزای دست‌نویس فارسی بر اساس تشخیص گروه بدنهٔ اصلی با استفاده از ماشین بُردار پشتیبان

محمّدامین مهرعلیان[1] و کاظم فولادی[2]
[1] دانشکدهٔ مهندسی کامپیوتر و فناوری اطّلاعات، دانشگاه صنعتی امیرکبیر، تهران
[2] دانشکدهٔ مهندسی برق و کامپیوتر، پردیس دانشکده‌های فنّی، دانشگاه تهران، تهران

چکیده

در این مقاله روشی جدید برای بازشناسی برخط حروف مجزای فارسی ارائه شده است که با استخراج چند ویژگی ساده از دنبالهٔ نمونه‌برداری شده از حروف و استفاده از دسته‌بندی‌کنندهٔ ماشین بردار پشتیبان(SVM) نتایج قابل قبولی را ارئه می‌دهد. الگوریتم پیش‌پردازش استفاده شده در این کار امکان یکسان سازی ابعاد ویژگی‌ها به ازای حروف متعدد را فراهم می‌کند تا در مرحلهٔ بعدی به‌منظور بازشناسی به دسته‌بندی کننده ارسال شود. فرآیند بازشناسی در دو مرحله صورت می‌گیرد: در مرحلهٔ اوّل بدنهٔ اصلی حرف ورودی (اوّلین حرکت قلم) پس از استخراج ویژگی با استفاده از دسته‌بندی‌کننده در قالب یکی از هجده گروه بدنهٔ اصلی حروف، طبقه‌بندی می‌شود و سپس در مرحلهٔ دوم، موقعیت، تعداد و شکل سایر حرکت‌ها مانند نقطه و سرکش (ریزحرکت‌ها)، نوع حرف نهایی را تعیین می‌کند. به‌عنوان نمونه برای تشخیص حرف «ت» ابتدا گروه بدنهٔ «ب، پ، ت، ث» تشخیص داده می‌شود و سپس وجود ریزحرکت «دونقطه» در بالای آن منجر به انتخاب «ت» از این گروه می‌شود. درنهایت در فرآیند پس‌پردازش با استفاده از تطبیق اطّلاعات مربوط به بدنهٔ اصلی و ریزحرکات سامانه به تصحیح خطاهای احتمالی موجود در مراحل قبلی پرداخته و دقّت بازشناسی را افزایش می‌دهد به‌عنوان مثال اگر در مرحله دسته‌بندی بدنهٔ حرف «ل» تشخیص داده شود ولی یک نقطه در بالای آن قرار داشته باشد آنگاه سامانه تشخیص خود را به حرف «ن» تغییر خواهد داد. نتایج تجربی این کار پژوهشی که بر اساس مجموعه دادهٔ Online-TMU صورت گرفته است، متوسط نرخ بازشناسی بدنهٔ اصلی را ۹۴٪ نشان می‌دهد و با درنظرگرفتن پس‌پردازش‌ها بر اساس ریزحرکت‌ها این نرخ به حدود ۹۸٪ می‌رسد.

واژگان کلیدی: بازشناسی برخط حروف فارسی، بازشناسی دست‌نوشته/ دست‌نویس، ریزحرکت، ماشین بردار پشتیبان (SVM).

۱- مقدمه

امروزه ابزارهای تجاری چون Tablet PC و PDA که از صفحات حساس به جای صفحه کلید برای ارتباط با کاربر استفاده می‌کنند توسعهٔ فراوانی یافته است، از این‌رو تقاضا برای سامانه‌های بازشناسی برخط (online) افزایش‌یافته است. اصلی‌ترین تفاوت میان دو روی‌کرد برخط و برون خط (offline) دسترسی به دنبالهٔ نقاط در هر حرکت دست و همچنین ترتیب و موقعیت زیرحرکات است.

در این مقاله قصد داریم به ارائهٔ روشی جدید برای بازشناسی برخط حروف مجزای فارسی بپردازیم. در روش ارائه‌شده سعی شده است کم‌ترین تعداد قید ممکن بر روی نحوهٔ نگارش حروف قرارداده شود. اصلی‌ترین قیدی که وجود دارد، این است که کاربر بایستی ابتدا «بدنهٔ» حروف را نوشته و سپس نقطه‌گذاری آنها را انجام دهد.

در مقایسه با زبان‌های لاتین، برای بازشناسی برخط حروف فارسی و عربی، تاکنون کارهای کم‌تری انجام شده است. مرور مختصری برروی مهم‌ترین کارهای مرتبط انجام‌شده در این زمینه، در بخش دو ارائه می‌شود.

این مقاله در هشت بخش سازمان‌دهی شده است. در بخش دو مروری بر کارهای انجام شده در زمینهٔ بازشناسی حروف دست‌نویس برخط فارسی/ عربی ارائه می‌شود. بخش سه مدلی برای حروف دست‌نویس فارسی ارائه و خصوصیات آن را بیان می‌کند. الگوریتم بازشناسی حروف و جزئیات آن در بخش‌های چهار، پنج و شش تحت عنوان‌های پیش‌پردازش، استخراج ویژگی و طبقه‌بندی و پس‌پردازش

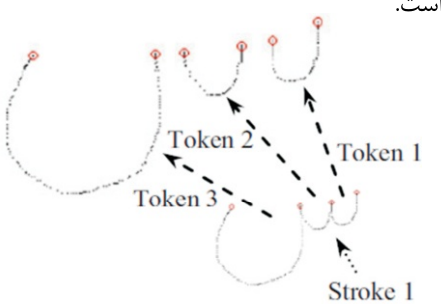
ارائه می‌شود. در بخش هفت به مسایل مربوط به پیاده‌سازی و نتایج تجربی پرداخته می‌شود. سرانجام بخش هشت به خلاصهٔ مطالب، بحث و نتیجه‌گیری اختصاص داده شده است.

۲- مروری بر کارهای انجام‌شده توسط دیگران

در زمینهٔ بازشناسی دست‌نوشتهٔ برخط، برخی کارها به بازشناسی ارقام و برخی دیگر به بازشناسی حروف یا کلمات پرداخته‌اند. از آنجاکه شرایط آزمایش و معیارهای ارزیابی این کارها با یکدیگر متفاوت است، مقایسهٔ عددی آنها را بیان نمی‌کنیم و فقط روش‌های مورد استفاده در آنها را مورد توجّه قرار می‌دهیم. کارهای قدیمی‌تر توسط مؤلّفان همین مقاله در *(مهرعلیان، فولادی، ۱۳۸۸)* بررسی شده است و در اینجا تنها به مرور چند کار مرتبط جدیدتر می‌پردازیم.

در کار قبلی مؤلّفان این مقاله، (مهرعلیان، فولادی، ۱۳۸۸)، بازشناسی حروف مجزای دست‌نویس فارسی، بر اساس تشخیص بدنهٔ اصلی به‌کمک مدل مخفی مارکوف (HMM) انجام شده است. پس از تشخیص بدنهٔ اصلی، ریزحرکت‌ها (مانند نقاط، سرکش و مد) مورد توجّه قرار می‌گیرند و از ترکیب اطّلاعات آنها با بدنهٔ اصلی تشخیص‌داده‌شده، حرف نوشته‌شده بازشناسی می‌شود. تفاوت اصلی این کار با کار قبلی تغیر نوع دسته‌بندی کننده از نوع تعمیمی (Generative) به تمایزی (discriminative) است که در بخش ۰ برتری آن به تفصیل توضیح داده خواهد شد.

در (*Harouni, Mohamad, 2010*) نویسنده با تعریف نقاط بحرانی در شکل حروف (نقاط کمینه و بیشنه محلی) هر یک از آنها را به بخش‌هایی تقسیم‌بندی کرده است ((شکل ۱)، سپس با استفاده از یک مجموعه ویژگی‌های هندسی و با به‌کارگیری شبکهٔ عصبی به دسته‌بندی حروف پرداخته است.



(شکل ۱): تقطیع حروف با استفاده از نقاط بحرانی در *(Harouni, Mohamad, 2010)*

مشکل اصلی این مقاله درواقع استفاده از همین روش تقطیع است زیرا اوّل این‌که محل تقطیع همیشه با استفاده از نقاط کمینه و بیشینهٔ محلی، قابل شناسایی نیست و دوم این‌که این تقطیع تنها در شرایطی یکسان خواهد بود که کاربر محدود به نوشتن حروف در غالب استاندارد آنها (رعایت دندانه‌ها و ...) باشد. به‌عنوان نمونه در ((شکل۲) حرف «ن» دارای یک نقطهٔ بحرانی پیش از انتهای حرف است و دندانه‌ها حرف «س» نوشته نشده است

(شکل۲) : دو نمونه از نحوهٔ نگارش حروف با حالت‌های غیر استاندارد.

علاوه‌بر این، نتایج مقاله با استفاده از یک پایگاه جمع‌آوری‌شده توسط نویسنده ارائه شد است و مقایسه‌ای بین روش ذکر شده و سایر کارها ارائه نشده است.

در (*omer, 2010*) برای باشناسی حروف از درخت تصمیم استفاده شده است که در سطح اوّل از آن، نمونه‌ها بر اساس تعداد نقاط دسته‌بندی می‌شوند. سپس در سطح بعدی موقعیت نقاط مورد بررسی قرار گرفته که برای این منظور خط کرسی نیز استخراج می‌شود. در آخرین مرحله نیز با استفاده از یک الگوریتم تطبیق الگو، بدنهٔ اصلی حروف با بدنه اصلی حروف معیاری که از هر حرف در پایگاه ذخیره شده است، مقایسه می‌شود. اگرچه نتایج نهایی ارائه‌شده برای پنج کاربر مطلوب به‌نظر می‌رسد امّا وجود ویژگی تعداد نقاط در اولین سطح از درخت تصمیم محدودیت‌هایی را برای کاربران سامانه ایجاد می‌کند.

همچنین در این مقاله نیز از پایگاه جمع‌آوری شده توسط نویسنده از تنها ۹ کاربر استفاده شده است که چهار نفر از آنها به‌منظور آموزش و پنج نفر دیگر برای آزمایش سامانه استفاده شده‌اند.

در (*Izadi, Suen, 2009*) از دو دسته‌بندی‌کنندهٔ مختلف استفاده شده است که هر یک، ویژگی‌های متفاوتی را مورد استفاده قرار می‌دهد. در دسته‌بندی‌کنندهٔ اوّل، یک ویژگی با عنوان «هیستوگرام همبستگی» به شبکه عصبی پرسپترون داده می‌شود و در دسته‌بندی‌کنندهٔ دوم، ویژگی محتوای همبستگی[1] با دسته‌بندی‌کنندهٔ SVM مورد استفاده قرار داده شده، مشکل اصلی این روش، ابعاد بالای بردارهای ویژگی مورد استفاده است؛ به‌عنوان مثال در دسته‌بندی‌کنندهٔ SVM به‌ازای N نمونه ورودی برای

[1] relational context(RC)

هرحرف، N(N - 1) ویژگی استخراج می‌شود. هر چند این مقاله به‌لحاظ استفاده از دسته‌بندی‌کننده با این کار دارای شباهت است؛ امّا تفاوت اصلی این دو کار در انتخاب نوع ویژگی و اصولاً ابعاد بالای ویژگی‌هاست که باتوجّه به کاربرد این سامانه‌ها در محیط‌های بلادرنگ مانند نگارش در محیط تلفن‌های همراه ممکن است باعث کاهش سرعت سامانه در شناسایی حروف شود.

۲- مدل حروف فارسی

حروف مجزا در نگارش فارسی در دو بخش اصلی نوشته می‌شوند. قسمت اصلی حروف «بدنه» نامیده می‌شود و علایم اضافی، مانند نقطه، سرکش یا کلاه را «ریزحرکت» می‌گوییم (سلیمانی، باقری، ۱۳۸۵). ویژگی مهمی که در بازشناسی حروف فارسی قابل توجّه است این است که برخی از حروف بدنهٔ یکسان دارند و تنها تفاوت آنها در ریزحرکت‌های آنهاست. با توجه به بدنه‌های مشابه، حروف فارسی را می‌توان در قالب هیجده گروه دسته‌بندی کرد که این گروه‌ها در (جدول ۱) قابل مشاهده هستند.

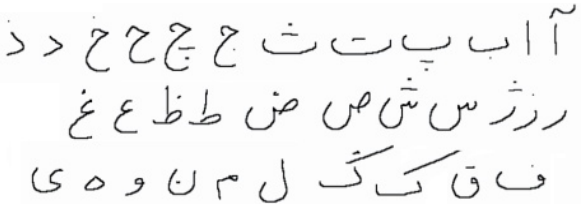
(جدول ۱): گروه‌بندی انجام شده براساس بدنهٔ حروف

گروه ۱	آ ، ا	گروه ۷	ص، ض	گروه ۱۳	ل
گروه ۲	ب، پ، ت، ث	گروه ۸	ط، ظ	گروه ۱۴	م
گروه ۳	ج، چ، ح، خ	گروه ۹	ع، غ	گروه ۱۵	ن
گروه ۴	د، ذ	گروه ۱۰	ف	گروه ۱۶	و
گروه ۵	ر، ز، ژ	گروه ۱۱	ق	گروه ۱۷	ه
گروه ۶	س، ش	گروه ۱۲	ک، گ	گروه ۱۸	ی

استفاده از روی‌کرد برخط این امکان را به ما می‌دهد که اطّلاعات ترتیب زمانی در هنگام نوشتن حروف و یا تعداد حرکات قلم را در اختیار داشته باشیم. در این کار فرض می‌کنیم که بدنهٔ اصلی حروف در ابتدا نوشته و سایر حرکات مانند نقاط، سرکش‌ها و... در مراحل بعدی گذاشته می‌شود. به‌عنوان نمونه حرف «چ» در (شکل۳) با سه حرکت قلم نوشته می‌شود، بدین‌ترتیب که ابتدا بدنهٔ اصلی «ح» و در مراحل بعدی نقاط قرار داده می‌شوند. به‌نظر می‌رسد که ذهن انسان نیز در فرآیند تشخیص حروف، ابتدا بدنهٔ اصلی را بازشناسی می‌کند و سپس موقعیت، تعداد و شکل ریزحرکت‌ها هستند که مشخّص می‌کند حرف مورد نظر چیست.

(شکل۳): ترتیب نگارش بدنهٔ اصلی و ریزحرکت‌ها

تنها محدودیتی که در نحوهٔ نوشتن ریزحرکت‌ها وجود دارد، این است که «سه نقطه» می‌بایستی از ترکیب «دو نقطه» و یک «تک نقطه» ساخته شده باشد. همچنین دو سرکش «گ» باید در دو حرکت مجزای قلم نوشته شود. به‌علاوهٔ دستهٔ «ط» و «ظ» و سرکش «ک» بایستی در حرکتی به‌جز حرکت بدنهٔ اصلی نوشته شود. به‌عبارت دیگر حروف «ط» و «ک» باید در دو حرکت قلم و حروف «گ» و «ظ» باید در سه حرکت قلم نگاشته شوند. ((شکل ۴)، صورت صحیح نگارش ۳۳ حالت مختلف حروف فارسی (شامل ۳۲ حرف و آ) را نشان می‌دهد که در آن هر قسمت با یک حرکت قلم نوشته شده است *(رضوی، کبیر، ۱۳۸۳)*.



(شکل ۴): نگارش صحیح برای ۳۳ حالت مختلف از حروف

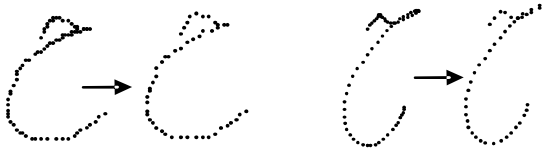
۳- پیش‌پردازش

از آنجاکه داده‌های برخط توسط قلم بر روی صفحهٔ حساس نوشته می‌شوند، مختصات نقاط نمونه‌برداری شده، تعداد و فاصلهٔ آنها و همچنین ابعاد منحنی‌های ایجاد شده از یک حرف به حرف دیگر و از یک نویسنده به نویسندهٔ دیگر می‌تواند به‌شدّت تغییر کند. برای اینکه بتوان تأثیر این تغییرات را در فرآیند بازشناسی به کمینه رساند، باید نوعی هنجارسازی روی مجموعه نقاط نمونه‌برداری‌شده از حرکات قلم انجام شود. در این مقاله هنجارسازی فاصلهٔ بین نقاط نمونه برداری، یکسان‌کردن تعداد نقاط نمونه برداری و یکسان کردن ابعاد حرکت‌ها به‌عنوان مهم‌ترین موارد پیش‌پردازش انجام می‌شود.

۳-۱- هنجارسازی فواصل میان نمونه‌ها

برای هنجارسازی فواصل میان نمونه‌ها، ابتدا میانگین فاصلهٔ نقاط متوالی در مجموعه‌نقاط مربوط به یک حرکت را اندازه‌گیری و نقاطی را که در فاصله‌ای کم‌تر از دو برابر این میانگین باشند، حذف می‌کنیم. با حذف این نقاط اضافی،

لرزش‌های کم‌تری نیز در مسیر حرکت دیده و از این رو لرزش‌های ناخواسته در نوشتن حرف نیز تا حدّی فیلتر می‌شود. (شکل ۵) دو نمونه از حروف و نتیجهٔ هنجارسازی فواصل میان نقاط آنها دیده می‌شود.



(شکل ۵) : دو نمونه از هنجارسازی فواصل میان نقاط

۳-۲- یکسان‌سازی تعداد نقاط نمونه‌برداری

یکسان‌سازی تعداد نقاط نمونه‌برداری برای سهولت استفاده در SVM و افزایش کارایی آن، ضروری است. درواقع با این کار همهٔ حروف با یک تعداد مشخّص از برش‌های زمانی نمونه‌برداری مجدد می‌شوند و به داده‌هایی با ابعاد یکسان تبدیل می‌شوند. برای این منظور، ابتدا با استفاده از اسپیلاین‌های درجهٔ سوم یک درون‌یابی بین هر چهار نقطهٔ متوالی از منحنی حرف انجام می‌شود و بدین‌ترتیب تخمینی از شکل کلی حرف صورت می‌گیرد. سپس طول منحنی به بیست قسمت مساوی تقسیم می‌شود، به‌طوری‌که طول منحنی بین هر دو نقطهٔ متوالی، اندازه‌ای برابر دارد و درنتیجه ۲۱ نقطه با فواصل مساوی از روی منحنی حاصل نمونه‌برداری می‌شود. البته تعداد این نقاط برای ریزحرکت‌ها به ۱۰ نقطه کاهش می‌یابد. (شکل ۶) دو مثال از انجام یکسان‌سازی تعداد نقاط نمونه‌برداری را نشان می‌دهد.

(شکل ۶): دو نمونه از یکسان‌سازی تعداد نقاط نمونه‌برداری

۳-۳- یکسان‌سازی اندازه و مختصات نقطهٔ شروع

از آنجا که اندازهٔ حرف نوشته‌شده توسط کاربران می‌تواند متفاوت باشد، بنابراین لازم است پیش از بازشناسی آن، یک همسان‌سازی در ابعاد صورت گیرد. برای این منظور حرف نوشته شده در چارچوب مختصاتی به ابعاد ۱۰۰×۱۰۰ قرار می‌گیرد. به‌علاوه از آنجا که نقطهٔ شروع نگارش حرف می‌تواند هر نقطه‌ای روی صفحه باشد، نقطهٔ شروع برای همهٔ حروف به نقطهٔ (۰،۰) منتقل می‌شود.

۴- استخراج ویژگی‌ها و طبقه‌بندی

در مرحلهٔ پیش‌پردازش تمامی حروف با نمونه‌برداری‌های متفاوت یکسان‌سازی شده و برای استخراج ویژگی آماده شدند. برای انجام بازشناسی بدنهٔ حروف و زیرحرکت‌ها با استفاده از SVM لازم است که ویژگی‌هایی با قابلیت تمایزی بالا را جهت بازنمایی هر یک از آنها انتخاب کنیم.

۴-۱- استخراج ویژگی‌ها

برای انتخاب ویژگی‌ها، زیرمجموعه‌ای از ویژگی‌های معرّفی شده در (*Biem, 2006*) را مورد استفاده قرار می‌دهیم. برای هر نقطهٔ نمونه مجموعهٔ چهار ویژگی زیر را محاسبه می‌کنیم:

$$f_1(x_i, y_i) = x_i \quad (۱)$$

$$f_2(x_i, y_i) = y_i \quad (۲)$$

$$f_3(x_i, y_i) = x_i - x_{i-1} (= \Delta x_i) \quad (۳)$$

$$f_4(x_i, y_i) = y_i - y_{i-1} (= \Delta y_i) \quad (۴)$$

یکی از بهبودهای این روش در مقایسه با روش قبلی ارائه شده در (مهرعلیان، فولادی، ۱۳۸۸) کاهش ابعاد بردار ویژگی برای هر نقطه از ۶ به ۴، بدون اُفت دقّت در نتایج نهایی است. دو ویژگی کنار گذاشته‌شده، معیاری از زاویهٔ نقطهٔ فعلی را نسبت به نقطهٔ قبلی بازنمایی می‌کردند.

ویژگی (۱) و (۲) همان مختصات دکارتی نقطه است که در مرحلهٔ پیش‌پردازش مقادیر هنجارشدهٔ آن محاسبه شده است.

ویژگی (۳) و (۴) میزان تغییر مختصات در راستای افقی و عمودی را نشان می‌دهد. از آنجا که در مرحلهٔ یکسان‌سازی تعداد نقاط نمونه درپیش‌پردازش همگی حروف با تعداد ثابتی نقطه بازنمایی شده‌اند، این ویژگی‌ها نیز که متأثّر از شکل و طول حروف هستند، می‌توانند یکی از پارامترهای ایجاد تمایز بین بدنهٔ حروف در نظر گرفته شوند.

مجموعهٔ این چهار ویژگی برای هر دو نقطهٔ متوالی استخراج می‌شود و کل آنها به‌عنوان بردار ویژگی به SVM داده می‌شود.

برای تعیین وضعیت ریزحرکت‌ها، موقعیت هر ریزحرکت با توجّه به بدنهٔ اصلی حرف و در سه دستهٔ بالا، پایین و وسط تقسیم‌بندی می‌شود. این دسته‌بندی براساس مقایسهٔ نقطهٔ مرکزی ارتفاع ریزحرکت با کمینه و بیشینهٔ ارتفاع بدنه صورت می‌گیرد. درحالتی‌که نقطهٔ مرکزی ارتفاع

ریزحرکت بین کمینه و بیشینهٔ ارتفاع بدنهٔ اصلی قرار داشته باشد (دستهٔ وسط)، بسته به گروه ممکن است به یکی از دو حالت بالا و پایین تبدیل شود. برای مثال برای حرف «ج» وسط به معنی پایین و در حرف «ن» وسط به معنی بالا خواهد بود.

۴-۲- طبقه‌بندی

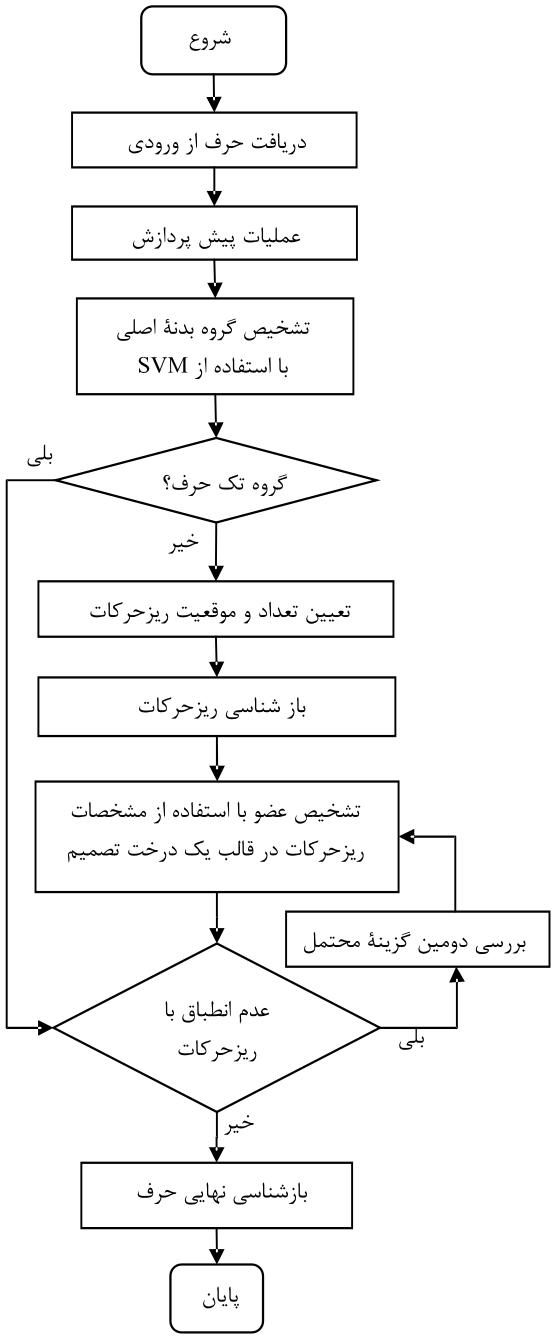
در این مقاله از «ماشین بردار پشتیبان» یا SVM به‌عنوان دسته‌بندی‌کنندهٔ اصلی در بازشناسی حروف استفاده شده است. بررسی‌های صورت گرفته نشان می‌دهد استفاده از SVM به‌طور مستقیم در بازشناسی برخط چندان مورد توجّه قرار نگرفته است و علّت این امر نیز ناشی از متغیربودن طول ویژگی‌های استخراج شده، بسته به طول منحنی حروف است (*Bahlmann, Haasdonk, 2002*) . به‌عنوان نمونه در (*Ahmad, Khalia, 2004*) از SVM تنها برای جداسازی (segmentation) حروف استفاده شده است و برای دسته‌بندی نهایی HMM به‌کارگرفته شده است. امّا فرآیند پیش‌پردازش استفاده شده در این مقاله علاوه‌بر ایجاد یک استانداردسازی اولیهٔ حروف، قابلیت استفاده از ویژگی‌ها را برای دسته‌بندی‌کنندهٔ SVM فراهم می‌کند.

دسته‌بندی‌کنندهٔ HMM مورد استفاده در (مهرعلیان، فولادی، ۱۳۸۸) و روش مورد استفاده در این کار هریک دارای ویژگی‌هایی هستند. در HMM از آنجا که ویژگی‌های مربوط به نقاط، در توالی‌های زمانی قرار گرفته‌اند. بنابراین جابه‌جایی زمانی ظاهرشدن ویژگی‌ها در آن چندان اثر گذار نیست. اما از آنجا که در SVM توالی نمونه‌ها در ابعاد بردار ویژگی ظاهر می‌شود (نقاطی که در زمان t نمونه‌برداری شده‌اند در بعد tأم ظاهر می‌شوند) جابه‌جاشدن این تغییرات می‌تواند بر نتیجهٔ دسته‌بندی اثرگذار باشد. در مقابل از آنجا که SVM به‌عنوان یک دسته‌بندی‌کنندهٔ تمایزی[1] برای تعیین مرز دسته‌بندی، علاوه‌بر نمونه‌های کلاس مورد آموزش، سایر کلاس‌ها را نیز در نظر می‌گیرد یک دسته‌بندی‌کننده با تفکیک‌پذیری بالا خواهیم داشت که با ویژگی‌های کم‌تر به نرخ بازشناسی مطلوبی می‌رسد، امّا در HMM به‌عنوان یک دسته‌بندی‌کنندهٔ تعمیمی[2] مدل هر طبقه به‌صورت مجزا و بدون در نظر گرفتن سایر طبقه‌ها آموزش داده می‌شود، بنابراین با افزایش تعداد طبقه‌ها دقّت آن کاهش یافته و گاهی برای تفکیک بهتر میان طبقه‌ها نیاز به ویژگی‌های بیشتری خواهد بود.

[1] Discriminative
[2] Generative

۵- پس‌پردازش

پس از تشخیص و بازشناسی گروه بدنهٔ اصلی، نتایج همراه با اطلاعات مربوط به ریزحرکت‌ها، به یک درخت تصمیم داده می‌شود تا با ترکیب اطلاعات این دو بخش بررسی کند که حرف ناشناخته، متعلق به کدام‌یک از اعضای گروه‌های هجده‌گانه است و بدین ترتیب فرآیند بازشناسی کامل حرف کامل شود (مهرعلیان، فولادی، ۱۳۸۸).



(شکل ۷) نمودار جعبه‌ای الگوریتم بازشناسی حرف

از آنجا که بازشناسی و تفکیک ریزحرکت‌ها مانند نقطه‌ها، سرکش‌ها و... تاحدودی دشوار است و اغلب به‌لحاظ ظاهری نیز با هم تداخل دارند، سعی ما بر این بوده است که

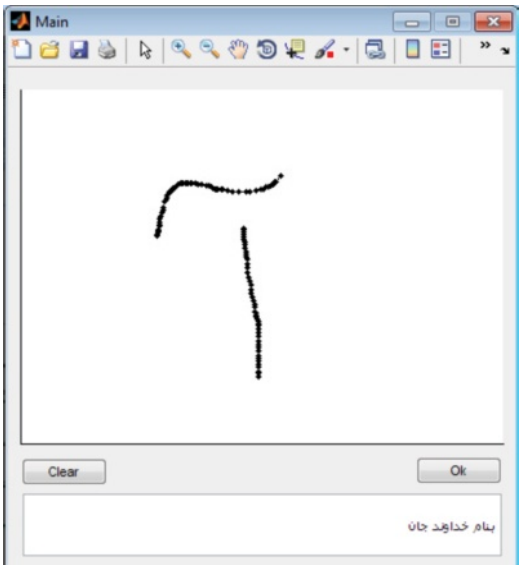
از این اطّلاعات در آخرین مراحل تصمیم‌گیری استفاده شود. از این‌رو در ساختار درخت تصمیم، اولویت تصمیم‌گیری با تعداد و محل ریزحرکت‌هاست، نه شکل واقعی آنها. به‌عنوان نمونهٔ زمانی که بدنهٔ اصلی حرفی در قالب گروه ۱، «ا» تشخیص داده می‌شود، تنها وجود یک زیرحرکت در بالای آن کافی است تا سامانه آن را «آ» تشخیص دهد و دیگر نیازی به تشخیص شکل علامت «مد» نیست.

همچنین گاهی به‌خاطر عدم تطابق بین تعداد ریزحرکت‌ها و بدنهٔ اصلی، سامانه قادر به تشخیص خطا در بازشناسی بدنه خواهد بود و می‌تواند بدنهٔ تشخیص داده شده را تصحیح کند و آن را به بدنهٔ محتمل دیگری تبدیل نماید. برای مثال ممکن است سامانه بدنهٔ «ل» را به همراه یک ریزحرکت در بالای آن تشخیص دهد. در چنین حالتی سامانه باتوجّه به عدم تطابق بین آنچه تشخیص داده شده و آنچه برای حرف «ل» مورد انتظار بوده است، تشخیص بدنهٔ خود را به «ن» تغییر می‌دهد. به‌عبارت دیگر بازنگری نهایی براساس مشخّصات بسیار ساده‌ی ریزحرکت‌ها صورت می‌گیرد.

(شکل ۷)، نمودار جعبه‌ای الگوریتم بازشناسی شامل کلیهٔ مراحل توضیح داده‌شده را در یک نما نشان می‌دهد.

۶- پیاده‌سازی و نتایج تجربی

برای پیاده‌سازی برنامهٔ بازشناسی، از نرم‌افزار MATLAB استفاده شده است. برای استفاده از SVM نیز کتابخانهٔ LIBSVM، به‌کار گرفته شده است. (شکل ۸) نمایی از این نرم‌افزار را نشان می‌دهد.



(شکل ۸): نمایی از نرم‌افزار بازشناسی برخط حروف فارسی

برای SVM از کرنل گاوسی استفاده شده که مقدار پارامتر σ در آن برای حروف و ریزحرکات به ترتیب ۰.۶ و ۱.۶ در نظر گرفته شده است.

یکی از مشکلات اصلی برای مقایسهٔ عملکرد این مقاله با مقالات جدیدتر، عدم دسترسی به پایگاه استفاده شده در این مقالات است. زیرا در برخی از این مقالات نویسنده به‌طور مستقیم پایگاه را جمع‌آوری کرده و تنوع پایگاه بسیار پایین است و یا در برخی پایگاه به‌صورت اختصاصی مورد استفاده قرار گرفته و اجازهٔ دسترسی عمومی به آن وجود ندارد. همچنین در بیشتر این کارها دست‌خط‌های استفاده شده از نگارش‌های عربی است که علاوه‌بر نبود برخی از حرف، نوع نگارش نیز تفاوت‌های زیادی با نگارش‌های فارسی دارد.

بنابراین مجموعه دادهٔ Online-TMU که توسط بخش مهندسی برق دانشگاه تربیت مدرس تهیه شده است (رضوی، کبیر، ۱۳۸۳)، به‌عنوان مجموعهٔ دادهٔ آموزشی و آزمایشی مورد استفاده قرار گرفته است.

از هفتاد درصد نمونه‌ها در مرحله آموزش و از سی درصد مابقی برای آزمایش استفاده شده است. نتیجهٔ مرحلهٔ آزمون در ماتریس پراکنش موجود در (جدول ۲) نشان داده شده است.

همانطور که مشاهده می‌شود، بازشناسی در خروجی SVM با میانگینی برابر با ۹۴.۳٪ درست صورت می‌گیرد. (جدول ۳) مقایسه‌ای بین ماتریس پراکنش حاصل از دسته‌بندی‌کنندهٔ SVM را نسبت به روش ارائه‌شده در (مهرعلیان، فولادی، ۱۳۸۸) با استفاده از HMM برای هر گروه به‌صورت جداگانه نشان می‌دهد (اعداد نشان داده شده نماینده تعداد نمونه‌هایی است که درست دسته‌بندی شده‌اند).

نرخ بازشناسی کلّی سامانه برای حروف پس انجام پس‌پردازش به ۹۸٪ درصد می‌رسد. به‌عنوان نمونه در مورد دو گروه ۲ و ۱۲ (بدنه‌های «ک» و «ب») که تداخل زیادی دیده می‌شود، تصمیم‌گیری نهایی با دخیل‌کردن اطّلاعات ریزحرکات صورت می‌گیرد.

(جدول ۴): مقایسهٔ نرخ بازشناسی صحیح این سیستم با دو کار دیگر براساس مجموعه دادهٔ Online-TMU

سامانهٔ ارائه شده در (سلیمانی، باقری، ۱۳۸۵)	٪۹۰/۲۷
سامانهٔ ارائه شده در (رضوی، کبیر، ۱۳۸۳)	٪۹۳/۳
سامانهٔ پیشنهادی قبلی در (مهرعلیان، فولادی، ۱۳۸۸)	٪۹۷
سامانهٔ پیشنهادی جدید در این مقاله	٪۹۸

(جدول ۴) نرخ بازشناسی کلّی این سامانه را با سه سامانهٔ دیگر که آنها نیز برای ارزیابی از مجموعه‌دادهٔ *Online-TMU* استفاده کرده‌اند نشان می‌دهد. مشاهده

(جدول۲): ماتریس پراکنش (Confusion) برای بازشناسی گروه‌ها

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	%
1	۷۱	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۱	۰	۰	۰	۰	0.99
2	۰	۱۳۶	۰	۰	۱	۰	۰	۰	۰	۰	۰	۸	۰	۰	۰	۰	۰	۰	0.94
3	۰	۰	۱۴۳	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	1.00
4	۰	۰	۰	۷۰	۱	۰	۰	۱	۰	۰	۰	۱	۰	۰	۰	۰	۰	۰	0.96
5	۰	۰	۰	۱	۱۰۹	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	0.99
6	۰	۰	۰	۰	۱	۶۸	۰	۰	۰	۰	۰	۱	۰	۱	۰	۱	۰	۱	0.93
7	۰	۰	۰	۰	۱	۱	۷۱	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	0.97
8	۰	۰	۰	۰	۰	۰	۰	۶۷	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	1.00
9	۰	۰	۰	۰	۰	۰	۰	۰	۷۲	۰	۰	۰	۰	۰	۰	۰	۰	۰	1.00
10	۰	۲	۰	۰	۰	۰	۰	۰	۰	۳۱	۴	۰	۰	۰	۰	۰	۰	۰	0.84
11	۰	۰	۰	۰	۰	۰	۰	۰	۰	۱	۳۵	۰	۰	۰	۱	۰	۰	۰	0.95
12	۰	۱۲	۰	۰	۱	۰	۰	۰	۰	۱	۱	۵۳	۰	۰	۱	۱	۰	۰	0.76
13	۰	۰	۰	۰	۱	۰	۰	۰	۰	۰	۱	۱	۳۲	۰	۱	۰	۱	۰	0.86
14	۰	۰	۰	۰	۱	۰	۰	۰	۰	۰	۰	۰	۰	۳۶	۰	۰	۰	۰	0.97
15	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۱	۲	۰	۰	۳۲	۰	۲	۰	0.86
16	۰	۰	۰	۰	۰	۰	۰	۲	۰	۰	۰	۰	۰	۰	۰	۳۶	۰	۰	0.95
17	۰	۰	۰	۰	۰	۰	۰	۱	۰	۰	۱	۰	۵	۰	۰	۰	۲۸	۰	0.80
18	۰	۰	۰	۰	۰	۱	۱	۰	۰	۰	۰	۰	۰	۰	۰	۰	۰	۳۶	0.95

درصد کل : ۹۴.۳

(جدول ۳): مقایسهٔ نتایج ماتریس پراکنش با روش ارائه‌شده در (مهرعلیان، فولادی، ۱۳۸۸)

گروه	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
SVM	۷۱	۱۳۶	۱۴۳	۷۰	۱۰۹	۶۸	۷۱	۶۷	۷۲	۳۱	۳۵	۵۳	۳۲	۳۶	۳۲	۳۶	۲۸	۳۶
HMM	۷۱	۱۲۵	۱۴۱	۶۹	۱۰۲	۷۱	۶۹	۶۴	۶۹	۳۱	۳۲	۵۱	۳۵	۳۶	۳۰	۳۵	۳۳	۳۶
اختلاف	۰	۱۱	۲	۱	۷	-۳	۲	۳	۳	۰	۳	۲	-۳	۰	۲	۱	-۵	۰

می‌شود که درصورت یکسان‌بودن شرایط آزمایش‌ها، سامانهٔ پیشنهادی ما نرخ بازشناسی صحیح بالاتری را ارائه می‌کند.

۷- خلاصه، بحث و نتیجه‌گیری

در این مقاله، روشی ساده برای بازشناسی برخط حروف فارسی ارائه شده است که مبنای آن، بازشناسی بدنه‌ی اصلی حروف با استفاده از SVM می‌باشد. بدین صورت که در گام اوّل بدنهٔ اصلی حروف تشخیص داده می‌شوند، سپس با استفاده از موقعیت و شکل ریزحرکات نتیجهٔ نهایی مشخّص می‌شود.

آنچه بر مبنای نتایج تجربی به‌نظر می‌رسد، این است که تشخیص بدنهٔ اصلی حروف با روش SVM نتایج مطلوبی را در بر دارد، اما بنا به دلایلی چون تعدد شکل ریزحرکت‌ها (مانند نوشتن سه نقطه در قالب یک، دو یا سه حرکت) و عدم انحصار در شکل (مانند شباهت مد و دو نقطه) استفاده از SVM برای بازشناسی آنها چندان کارامد به‌نظر نمی‌رسد. بنابراین استفاده از سایر اطّلاعات مربوط به ریزحرکات مانند موقعیت و تعداد آنها در مرحله پس‌پردازش، می‌تواند تا حدّ زیادی مشکل ضعف در بازشناسی حروف را برطرف کند.

در مقایسهٔ روش ارائه‌شده در این مقاله با روش قبلی که از HMM برای بازشناسی بدنهٔ اصلی استفاده می‌کرد، چند بهبود ملاحظه می‌شود. اوّل این‌که همان‌طور که در بخش ۵-۱ اشاره شد، تعداد عناصر بردار ویژگی برای هر نقطه، از شش به چهار کاهش‌یافته است؛ بدون اینکه نرخ باشناسی اولیه کاهش یابد. دوم این‌که ساختار SVM نسبت به HMM در هنگام اجرا از سرعت بیشتری برخوردار است که این باعث می‌شود پیاده‌سازی کارآمدتری در عمل ایجاد شود. هر چند استفاده از دسته‌بندی‌کنندهٔ تمایزی SVM باعث افزایش دقّت شده است، امّا نکتهٔ قابل توجّه در نتایج کاهش میزان تداخل میان طبقه‌ها یا همان حروف است.

۸- مراجع

مهرعلیان، محمدامین؛ فولادی، کاظم. "بازشناسی برخط حروف مجزای دست نویس فارسی بر اساس تشخیص گروه اصلی بدنه با استفاده از مدل مخفی مارکوف"، در مجموعه مقالات پانزدهمین کنفرانس مهندسی کامپیوتر ایران، تهران، مرکز توسعه فناوری نیرو(متن)، ۱۳۸۸.

سلیمانی باغشاه، مهدیه؛ باقری شورکی، سعید؛ کسائی،

شهره. "بازشناسی و یادگیری حروف مجزای برخط فارسی به روش فازی"، در مجموعه مقالات چهاردهمین کنفرانس مهندسی برق ایران، تهران، دانشگاه صنعتی امیرکبیر، ۱۳۸۵.

رضوی، سید محمد؛ کبیر، احسان‌اله."بازشناسی برخط حروف مجزای فارسی"، در مجموعه مقالات ششمین کنفرانس سامانه‌های هوشمند، کرمان، دانشگاه شهید باهنر، آذر ۱۳۸۳.

رضوی، سید محمد؛ کبیر، احسان‌اله. "یک پایگاه داده برای بازشناسی دست‌نوشته‌های برخط فارسی"، در مجموعه مقالات ششمین کنفرانس سامانه‌های هوشمند، دانشگاه شهید باهنر، کرمان، آذر ۱۳۸۳.

Harouni, M., Mohamad, D., Rasouli, A., "Deductive Method for Recognition of On-Line Handwritten Persian/Arabic Characters", The 2nd International Conference on Computer and Automation Engineering, (ICCAE), 2010, pp. 791-795.

Omer, M.A.H., Ma, S.L., "Online Arabic Handwriting Character Recognition Using Matching Algorithm", The 2nd International Conference on Computer and Automation Engineering, (ICCAE), 2010, pp. 259-262.

Izadi, S., Suen, C.Y., "Integration of Contextual Information in Online Handwriting Representation", Lecture Notes in Computer Science, LNCS 5716, 2009, pp. 132-142.

Biem, A., "Minimum classification error training for online handwriting recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, Vol. 28, No. 7.

Bahlmann, C., Haasdonk, B., Burkhardt, H., "Online Handwriting Recognition with Support Vector Machines-A Kernel Approach", Eighth International Workshop on Frontiers in Handwriting Recognition, 2002, pp. 49-54.

Ahmad, A.R., Khalia, M., Viard-Gaudin, C., Poisson, E., "Online Handwriting Recognition Using Support Vector Machine", IEEE Region 10 Conference (TENCON), 2004, pp. 311-314.

محمدامین مهرعلیان در سال ۱۳۸۸ مدرک کارشناسی خود را در رشتهٔ مهندسی کامپیوتر گرایش نرم‌افزار از دانشگاه اراک دریافت کرد. وی مدرک کارشناسی ارشد خود را در رشتهٔ مهندسی کامپیوتر گرایش هوش مصنوعی از دانشگاه صنعتی امیرکبیر دریافت نمود. یادگیری ماشین، شناسایی الگو و پردازش تصویر از زمینه‌های تحقیقاتی مورد علاقهٔ وی است.

نشانی رایانامهٔ ایشان عبارت است از:

mehralian@aut.ac.ir

کاظم فولادی، کارشناسی ارشد خود را در رشتهٔ مهندسی برق و کامپیوتر در گرایش هوش ماشین و روباتیک از دانشگاه تهران در سال ۱۳۸۵ دریافت کرد و هم‌اکنون در مرحلهٔ دفاع از رسالهٔ دکترای خود با موضوع بازشناسی دست‌نوشته‌های فارسی در همین رشته و دانشگاه است. وی به مدّت هفت سال تاکنون در دانشگاه تهران به تدریس دروس نظریهٔ زبان‌ها و ماشین‌ها، هوش مصنوعی و اصول طراحی کامپایلر پرداخته و در دانشگاه‌های اراک، شاهد در مقطع کارشناسی دروس متعددی را ارائه نموده است.

ایشان هم‌اکنون رییس مرکز مطالعات سیستم‌های ترنس‌فیزیکی است و زمینه‌های پژوهشی وی طیف خاصی از نظریهٔ سیستم‌ها و همچنین سیستم‌های بازشناسی الگوهای متنی است.

نشانی رایانامهٔ ایشان عبارت است از:

kazim@fouladi.ir