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

محمدامین مهرعلیان[۱] و کاظم فولادی[۲]
[۱] دانشکده‌ی مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی امیرکبیر، تهران (mehralian@aut.ac.ir)
[۲] دانشکده‌ی مهندسی برق و کامپیوتر، پردیس دانشکده‌های فنی، دانشگاه تهران، تهران (kfouladi@ut.ac.ir)

چکیده

در این مقاله روشی جدید برای بازشناسی برخط حروف مجزای فارسی ارائه شده است که با استفاده از دسته‌بندی کننده‌ی ماشین بردار پشتیبان یا SVM عمل می‌کند. این عملیات در دو مرحله صورت می‌گیرد. در مرحله‌ی اول حرف ورودی با استفاده از دسته‌بندی کننده در قالب یکی از هجده گروه بدنه‌ی اصلی حروف، طبقه‌بندی می‌شود و سپس در مرحله‌ی دوم، موقعیت، تعداد و شکل سایر حرکت‌ها مانند نقطه و سرکش (ریزحرکت‌ها)، حرف نهایی را تعیین می‌کند. به عنوان نمونه برای تشخیص حرف «ت» ابتدا گروه بدنه‌ی «ب، پ، ت، ث» تشخیص داده می‌شود و سپس وجود ریزحرکت «دونقطه» در بالای آن منجر به انتخاب «ت» از این گروه می‌شود. نتایج تجربی این کار پژوهشی که بر اساس مجموعه داده‌ی Online-TMU صورت گرفته است، متوسط نرخ بازشناسی بدنه‌ی اصلی را ۹۴% نشان می‌دهد و با در نظر گرفتن پس‌پردازش‌ها بر اساس ریزحرکت‌ها این نرخ به حدود ۹۸% می‌رسد.

کلمات 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 توسعه‌ی فراوانی یافته است، از اینرو تقاضا برای سیستم‌های بازشناسی برخط (online) افزایش یافته است. اصلی‌ترین تفاوت میان دو رویکرد برخط و برون خط (offline) دسترسی به دنباله‌ی نقاط در هر حرکت دست و همچنین ترتیب و موقعیت زیر حرکات است.

در این مقاله قصد داریم به ارائه‌ی روشی جدید برای بازشناسی برخط حروف مجزای فارسی بپردازیم. در روش ارائه شده سعی شده است کمترین تعداد قید ممکن بر روی نحوه‌ی نگارش حروف قرار داده شود. اصلی‌ترین قیدی که وجود دارد، این است که کاربر بایستی ابتدا «بدنه»ی حروف را بنویسد 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 بر روی مهم‌ترین کارهای مرتبط انجام شده در این زمینه در بخش ۲ ارائه می‌شود.

این مقاله در هشت بخش سازمان‌دهی شده است. در بخش۲ مروری بر کارهای انجام شده در زمینه‌ی بازشناسی حروف دست‌نویس برخط فارسی/ عربی ارائه می‌شود. بخش ۳ مدلی را برای حروف دست‌نویس فارسی ارائه می‌کند و خصوصیات آن را بیان می‌کند. الگوریتم بازشناسی حروف و جزئیات آن در بخش‌های ۴، ۵ و ۶ تحت عنوان‌های پیش‌پردازش، استخراج ویژگی و طبقه‌بندی و پس‌پردازش ارائه می‌شود. در بخش ۷ به مسایل مربوط به پیاده‌سازی و نتایج تجربی پرداخته می‌شود. سرانجام بخش ۸ به خلاصه‌ی مطالب، بحث و نتیجه‌گیری اختصاص داده شده است.

۲- مروری بر کارهای انجام شده توسط دیگران

در زمینه‌ی بازشناسی دست‌نوشته‌ی برخط، برخی کارها به بازشناسی ارقام و برخی دیگر به بازشناسی حروف یا کلمات پرداخته‌اند. از آنجا شرایط آزمایش و معیارهای ارزیابی این کارها با یکدیگر متفاوت است، مقایسه‌ی عددی آنها را بیان نمی‌کنیم و فقط روش‌های مورد استفاده در آنها را مورد توجه قرار می‌دهیم. کارهای قدیمی‌تر توسط مؤلفین همین مقاله در [۱] بررسی شده است و در اینجا تنها به مرور چند کار مرتبط جدیدتر می‌پردازیم.

در کار قبلی مولفین این مقاله، [۱]، بازشناسی حروف مجزای دست‌نویس فارسی، بر اساس تشخیص بدنه‌ی اصلی به کمک مدل مخفی مارکوف (HMM) انجام شده است. پس از تشخیص بدنه‌ی اصلی، ریزحرکت‌ها (مانند نقاط، سرکش و مد) مورد توجه قرار می‌گیرند و از ترکیب اطلاعات آنها با بدنه‌ی اصلی

تشخیص داده شده، حرف نوشته شده بازشناسی می‌شود.

در [5] نویسنده با تعریف نقاط بحرانی در شکل حروف (نقاط کمینه و بیشنه محلی) هر یک از آنها را به بخش‌هایی تقسیم‌بندی کرده است، سپس با استفاده از یک مجموعه ویژگی‌های هندسی و با بکارگیری شبکه‌ی عصبی به دسته‌بندی حروف پرداخته است.

در [6] برای باشناسی حروف از درخت تصمیم استفاده شده است که در سطح اول از آن، نمونه‌ها بر اساس تعداد نقاط دسته‌بندی می‌شوند. سپس در سطح بعدی موقعیت نقاط مورد بررسی قرار می‌گیرد که برای این منظور خط کرسی نیز استخراج می‌گردد. در آخرین مرحله نیز با استفاده از یک الگوریتم تطبیق الگو، بدنه اصلی حروف با بدنه اصلی حروف معیاری که از هر حرف در پایگاه ذخیره شده است مقایسه می‌گردد. اگرچه نتایج نهایی ارائه شده مطلوب به نظر می‌رسد اما وجود ویژگی تعداد نقاط در اولین سطح از درخت تصمیم محدودیت‌هایی را برای کاربران سیستم ایجاد می‌کند.

در [7] از دو دسته‌بندی کننده مختلف استفاده شده است که هر یک، ویژگی‌های متفاوتی را مورد استفاده قرار می‌دهد. در دسته‌بندی‌کننده‌ی اول، یک ویژگی با عنوان «هیستوگرام همبستگی» به شبکه عصبی پرسپترون داده می‌شود و در دسته‌بندی‌کننده دوم، ویژگی محتوی همبستگی با دسته‌بندی کننده‌ی SVM مورد استفاده قرار داده شده، مشکل اصلی این روش، ابعاد بالای بردارهای ویژگی مورد استفاده است؛ مثلاً در دسته‌بندی کننده SVM به ازای N نمونه ورودی برای هرحرف، $N(N-1)$ ویژگی استخراج می‌شود.

۳- مدل حروف فارسی

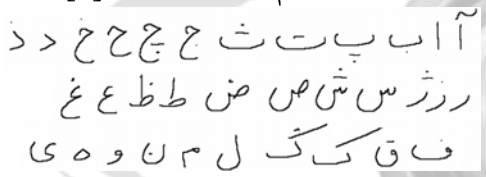
حروف مجزا در نگارش فارسی در دو بخش اصلی نوشته می‌شوند. قسمت اصلی حروف «بدنه» نامیده می‌شود و علایم اضافی، مانند نقطه، سرکش یا کلاه را «ریز حرکت» می‌گوییم [۲]. ویژگی مهمی که در بازشناسی حروف فارسی قابل توجه است این است که برخی از حروف بدنه‌ی یکسان دارند و تنها تفاوت آنها در ریزحرکت‌های آنهاست. با توجه به بدنه‌های مشابه، حروف فارسی را می‌توان در قالب ۱۸ گروه دسته‌بندی کرد که این گروه‌ها در جدول (۱) قابل مشاهده هستند.

جدول (۱) : گروه بندی انجام شده بر اساس بدنه 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 ٤	د، ذ	گروه ۱۰	ف	گروه ۱٦	و
گروه ٥	ر، ز، ژ	گروه ۱۱	ق	گروه ۱۷	ه
گروه ٦	س، ش	گروه ۱۲	ک، گ	گروه ۱۸	ی

استفاده از رویکرد برخط این امکان را به ما می‌دهد که اطلاعات ترتیب زمانی در هنگام نوشتن حروف و یا تعداد حرکات قلم را در اختیار داشته باشیم. در این کار فرض می‌کنیم که بدنه‌ی اصلی حروف در ابتدا نوشته می‌شود و سایر حرکات مانند نقاط، سرکش‌ها و... در مراحل بعدی گذاشته می‌شود. به عنوان نمونه حرف «چ» با سه حرکت قلم نوشته می‌شود، به این ترتیب که ابتدا بدنه‌ی اصلی «ح» و در مراحل بعدی نقاط قرار داده می‌شوند.

تنها محدودیتی که در نحوه‌ی 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. بعلاوه‌ی دسته‌ی «ط» و «ظ» و سرکش «ک» بایستی در حرکتی بجز حرکت بدنه‌ی اصلی نوشته شود. به عبارت دیگر حروف «ط» و «ک» باید در دو حرکت قلم و حروف «گ» و «ظ» باید در سه حرکت قلم نگاشته شوند. شکل (۱)، 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 [۳].



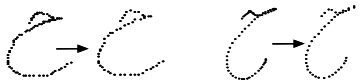
شکل (۱) : نگارش صحیح برای ۳۳ حالت مختلف از حروف

۴- پیش‌پردازش

از آنجا که داده‌های برخط توسط قلم بر روی صفحه‌ی حساس نوشته می‌شوند، مختصات نقاط نمونه‌برداری شده، تعداد و فاصله‌ی آنها و همچنین ابعاد منحنی‌های ایجاد شده از یک حرف به حرف دیگر و از یک نویسنده به نویسنده‌ی دیگر می‌تواند بشدت تغییر کند. برای اینکه بتوان تاثیر این تغییرات را در فرآیند بازشناسی به حداقل رساند، باید نوعی نرمال‌سازی روی مجموعه نقاط نمونه‌برداری شده از حرکات قلم انجام شود.

۴-۱- نرمال‌سازی فواصل میان نمونه‌ها

برای نرمال‌سازی فواصل میان نمونه‌ها، ابتدا میانگین فاصله‌ی نقاط متوالی در مجموعه نقاط مربوط به یک حرکت را اندازه‌گیری می‌کنیم 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 می‌شود و از این رو لرزش‌های ناخواسته در نوشتن حرف نیز تا حدی فیلتر می‌شود. شکل (۲) دو نمونه از حروف و نتیجه‌ی نرمال‌سازی فواصل میان نقاط آنها دیده می‌شود.



شکل (۲) : دو نمونه از نرمال‌سازی فواصل میان نقاط

۴-۲- یکسان‌سازی تعداد نقاط نمونه‌برداری

یکسان‌سازی تعداد نقاط نمونه‌برداری برای سهولت استفاده در SVM و افزایش کارایی آن، ضروری است. در واقع با این کار همه‌ی حروف با یک تعداد مشخص از برش‌های زمانی نمونه‌برداری مجدد می‌شوند و به داده‌هایی با ابعاد یکسان تبدیل می‌شوند.

برای این منظور، ابتدا با استفاده از اسپیلاین‌های درجه‌ی سوم یک درون‌یابی بین هر چهار نقطه‌ی متوالی از منحنی حرف انجام می‌شود و به این ترتیب تخمینی از شکل کلی حرف صورت می‌گیرد. سپس طول منحنی به ۲۰ قسمت مساوی تقسیم می‌شود به طوری که طول منحنی بین هر دو نقطه‌ی متوالی، اندازه‌ای برابر دارد و در نتیجه ۲۱ نقطه با فواصل مساوی از روی منحنی حاصل نمونه‌برداری می‌شود. البته تعداد این نقاط برای ریزحرکت‌ها به ۱۰ نقطه کاهش می‌یابد. شکل (۳) دو مثال از انجام یکسان‌سازی تعداد نقاط نمونه‌برداری را نشان می‌دهد.

شکل (۳) : دو نمونه از یکسان‌سازی تعداد نقاط نمونه‌برداری

۴-۳- یکسان‌سازی اندازه و مختصات نقطه‌ی شروع

از آنجا که اندازه حرف نوشته شده 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. بعلاوه از آنجا که نقطه‌ی شروع نگارش حرف می‌تواند هر نقطه‌ای روی صفحه باشد، نقطه‌ی شروع برای همه‌ی حروف به نقطه‌ی (۰,۰) منتقل می‌شود.

۵- استخراج ویژگی‌ها و طبقه‌بندی

برای انجام بازشناسی بدنه‌ی حروف و زیرحرکت‌ها با استفاده از SVM لازم است که ویژگی‌های مناسبی را جهت بازنمایی هر یک از آنها انتخاب کنیم.

۵-۱- استخراج ویژگی‌ها

برای انتخاب ویژگی‌ها، زیرمجموعه‌ای از ویژگی‌های معرفی شده در [8] را مورد استفاده قرار می‌دهیم. برای هر نقطه‌ی نمونه مجموعه‌ی 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 [۱] کاهش ابعاد بردار ویژگی برای هر نقطه از ۶ به ۴، بدون افت دقت در نتایج نهایی است. دو ویژگی کنار گذاشته شده، معیاری از زاویه‌ی نقطه‌ی فعلی را نسبت به نقطه‌ی قبلی بازنمایی می‌کردند.

ویژگی (۱) و (۲) همان مختصات دکارتی نقطه است که در مرحله‌ی پیش‌پردازش مقادیر نرمال‌شده‌ی آن محاسبه شده است. ویژگی (۳) و (۴) میزان تغییر مختصات در راستای افقی و عمودی را نشان می‌دهد. از آنجا که در مرحله‌ی یکسان‌سازی تعداد نقاط نمونه درپیش‌پردازش همگی حروف با تعداد ثابتی نقطه بازنمایی شده‌اند، این ویژگی‌ها نیز که متاثر از شکل و طول حروف می‌باشند، می‌توانند یکی از پارامترهای ایجاد تمایز بین بدنه‌ی حروف در نظر گرفته شوند.

مجموعه‌ی این چهار ویژگی برای هر دو نقطه‌ی متوالی استخراج می‌شود و کل آنها به عنوان بردار ویژگی به SVM داده می‌شود.

برای تعیین وضعیت ریزحرکت‌ها، موقعیت هر ریزحرکت با توجه به بدنه‌ی اصلی حرف و در سه دسته‌ی بالا، پایین و وسط تقسیم‌بندی می‌شود. این دسته‌بندی بر اساس مقایسه‌ی نقطه‌ی مرکزی ارتفاع ریزحرکت با حداقل و حداکثر ارتفاع بدنه صورت می‌گیرد. در حالتی که نقطه‌ی مرکزی ارتفاع ریزحرکت بین حداقل و حداکثر ارتفاع بدنه‌ی اصلی قرار داشته باشد (دسته‌ی 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.

۵-۲- طبقه‌بندی

در این مقاله از «ماشین بردار پشتیبان» یا SVM به عنوان دسته‌بندی‌کننده اصلی در بازشناسی حروف استفاده شده است. بررسی‌های صورت گرفته نشان می‌دهد استفاده از SVM به طور مستقیم در بازشناسی برخط چندان مورد توجه قرار نگرفته است و علت این امر نیز ناشی از متغیر بودن طول ویژگی‌های استخراج شده، بسته به طول منحنی حروف است [9]. به عنوان نمونه در [10] از SVM تنها برای جداسازی (segmentation) حروف استفاده شده است و برای دسته‌بندی نهایی HMM بکارگرفته شده است. اما فرآیند پیش‌پردازش استفاده شده در این مقاله علاوه بر ایجاد یک استانداردسازی اولیه‌ی حروف، قابلیت استفاده از ویژگی‌ها را برای دسته‌بندی کننده‌ی SVM فراهم می‌کند.

۶- پس‌پردازش

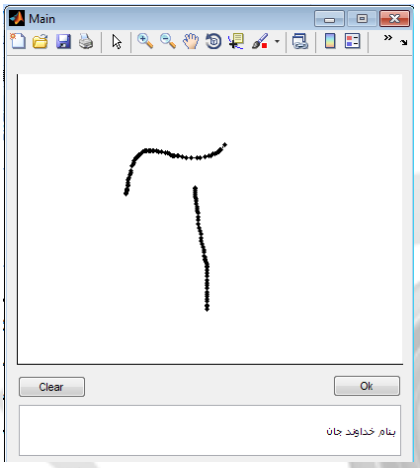
پس از تشخیص و بازشناسی گروه بدنه‌ی اصلی، نتایج همراه با اطلاعات مربوط به ریزحرکت‌ها، به یک درخت تصمیم داده می‌شود تا با ترکیب اطلاعات این دو بخش بررسی کند که حرف ناشناخته، متعلق به کدام‌یک از اعضای گروه‌های هجده‌گانه است و بدین ترتیب فرآیند بازشناسی کامل حرف کامل شود [۱].

از آنجا که بازشناسی و تفکیک ریزحرکت‌ها مانند نقطه‌ها، سرکش‌ها و... نسبتاً دشوار است و اغلب به لحاظ ظاهری نیز با هم تداخل دارند، سعی ما بر این بوده است که از این اطلاعات در آخرین مراحل تصمیم‌گیری استفاده شود. از این رو در ساختار درخت تصمیم، اولویت تصمیم‌گیری با تعداد و محل ریزحرکت‌هاست، نه شکل واقعی آنها. به عنوان نمونه زمانی که بدنه‌ی اصلی حرفی در قالب گروه ۱، «ا» تشخیص داده می‌شود، تنها وجود یک زیرحرکت در بالای آن کافی است تا سیستم آن را «آ» تشخیص دهد و دیگر نیازی به تشخیص شکل علامت «مد» نمی‌باشد.

همچنین گاهی به خاطر عدم تطابق بین ریزحرکت‌ها و بدنه‌ی اصلی، سیستم قادر به تشخیص خطا در بازشناسی بدنه خواهد بود و می‌تواند بدنه‌ی تشخیص داده شده را تصحیح کند و آن را به بدنه‌ی محتمل دیگری تبدیل نماید. برای مثال ممکن است سیستم بدنه‌ی «ل» را به همراه یک ریزحرکت در بالای آن تشخیص دهد. در چنین حالتی سیستم با توجه عدم تطابق بین آنچه تشخیص داده شده و آنچه برای حرف «ل» مورد انتظار بوده است، تشخیص بدنه‌ی خود را به «ن» تغییر می‌دهد. به عبارت دیگر بازنگری نهایی بر اساس مشخصات بسیار ساده‌ی ریزحرکت‌ها صورت می‌گیرد.

۷- پیاده‌سازی و نتایج تجربی

برای پیاده‌سازی برنامه‌ی بازشناسی، از نرم‌افزار MATLAB استفاده شده است. برای استفاده از SVM نیز کتابخانه LIBSVM، به کار گرفته شده است. شکل (۴) نمایی از این نرم‌افزار را نشان می‌دهد.



شکل (٤) : نمایی از نرم‌افزار بازشناسی برخط حروف فارسی.

جدول(۲) : 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) : مقایسه نتایج ماتریس پراکنش با روش ارائه شده در [۱]

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

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.

مجموعه داده‌ی Online-TMU که توسط بخش مهندسی برق دانشگاه تربیت مدرس تهیه شده است [۴]، به عنوان مجموعه‌ی داده‌ی آموزشی و آزمایشی مورد استفاده قرار گرفته است. از ۷۰ درصد نمونه‌ها در مرحله آموزش و از ۳۰ درصد مابقی برای آزمایش استفاده شده است. نتیجه‌ی مرحله‌ی 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 دسته‌بندی کننده SVM را نسبت به روش ارائه شده در [۱] (با استفاده از HMM) برای هر گروه به صورت جداگانه نشان می‌دهد (اعداد نشان داده شده نماینده تعداد نمونه‌هایی است که درست دسته‌بندی شده‌اند).

نرخ بازشناسی کلی سیستم برای حروف پس انجام پس‌پردازش به ۹۸% درصد می‌رسد. به عنوان نمونه در مورد دو گروه ۲ و ۱۲ (بدنه‌های «ک» و «ب») که تداخل زیادی دیده می‌شود تصمیم‌گیری نهایی با دخیل کردن اطلاعات ریزحرکات صورت می‌گیرد.

جدول (٤) : مقایسه‌ی نرخ بازشناسی صحیح این سیستم با دو کار دیگر بر اساس مجموعه داده‌ی Online-TMU

سیستم ارائه شده در [۲]	%۹۰/۲۷
سیستم ارائه شده در [۳]	%۹۳/۳
سیستم پیشنهادی قبلی در [۱]	%۹۷
سیستم پیشنهادی جدید در این مقاله	%۹۸

جدول (۴) نرخ بازشناسی کلی این سیستم را با سه سیستم دیگر که آنها نیز برای ارزیابی از مجموعه داده‌ی Online-TMU استفاده کرده‌اند نشان می‌دهد. مشاهده می‌شود که در صورت یکسان بودن شرایط آزمایش‌ها، سیستم پیشنهادی ما نرخ بازشناسی صحیح بالاتری را ارائه می‌کند.

۸- 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 اول بدنه اصلی حروف تشخیص داده می‌شوند سپس با استفاده از موقعیت و شکل ریز حرکات نتیجه نهایی مشخص می‌شود.

آنچه بر مبنای نتایج تجربی به نظر می‌رسد، این است که تشخیص بدنه‌ی 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 به نظر نمی‌رسد. بنابراین استفاده از سایر اطلاعات مربوط به ریزحرکات مانند موقعیت و تعداد آنها در مرحله پس پردازش، می‌تواند تا حد زیادی مشکل ضعف در بازشناسی حروف را برطرف کند.

در مقایسه‌ی روش ارائه شده در این مقاله با روش قبلی که از HMM برای بازشناسی بدنه‌ی اصلی استفاده می‌کرد، چند بهبود ملاحظه می‌شود. اولاً همان طور که در بخش ۵-۱ اشاره شد، تعداد عناصر بردار ویژگی برای هر نقطه، از شش به چهار کاهش یافته است، بدون اینکه نرخ باشناسی اولیه کاهش یابد. ثانیاً ساختار SVM نسبت به HMM در هنگام اجرا از سرعت بیشتری برخوردار است که این باعث می‌شود پیاده‌سازی کارآمدتری در عمل ایجاد شود. علاوه بر این بهبود نرخ بازشناسی در حدود یک درصد نیز قابل توجه است.

۹- مراجع

[۱] مهرعلیان، م.ا.، فولادی، ک.، **"بازشناسی برخط حروف مجزای دست نویس فارسی بر اساس تشخیص گروه اصلی بدنه با استفاده از مدل مخفی مارکوف"** مجموعه مقالات پانزدهمین کنفرانس مهندسی کامپیوتر ایران، مرکز توسعه فناوری نیرو(متن)، تهران، ۱۳۸۸.

[۲] سلیمانی باغشاه، م.، باقری شورکی، س.، کسائی، ش.، **"بازشناسی و یادگیری حروف مجزای برخط فارسی به روش فازی"**، مجموعه مقالات چهاردهمین کنفرانس مهندسی برق ایران، دانشگاه صنعتی امیرکبیر، تهران، ۱۳۸۵.

[۳] رضوی، س. م.، کبیر، ا.، **"بازشناسی برخط حروف مجزای فارسی"**، مجموعه مقالات ششمین کنفرانس سیستم‌های هوشمند، دانشگاه شهید باهنر، کرمان، آذر ۱۳۸۳.

[۴] رضوی، س. م.، کبیر، ا.، **"یک پایگاه داده برای بازشناسی دست‌نوشته‌های برخط فارسی"**، مجموعه مقالات ششمین کنفرانس سیستم‌های هوشمند، دانشگاه شهید باهنر، کرمان، آذر ۱۳۸۳.

[5] 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)*, pp. 791-795, 2010.

[6] 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)*, pp. 259-262., 2010.

[7] Izadi, S., Suen, C.Y., "Integration of contextual information in online handwriting representation", *Lecture Notes in Computer Science*, LNCS 5716, pp. 132-142. 2009.

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

[9] Bahlmann, C., Haasdonk, B., Burkhardt, H., "Online handwriting recognition with support vector machines-a kernel approach," *Eighth International Workshop on Frontiers in Handwriting Recognition*, pp. 49-54, 2002.

[10] Ahmad, A.R., Khalia, M., Viard-Gaudin, C., Poisson, E., "Online handwriting recognition using support vector machine", *IEEE Region 10 Conference (TENCON)*, pp. 311-314, 2004.