

بازشناسی کلمات دست‌نوشته با ویژگی‌های نوین و کاهش فرهنگ لغت

سمیه برومند^۱ و مجید ایرانپور مبارکه^۲

چکیده

بازشناسی کلمات دست‌نوشته و تبدیل آن به متن تایپی معادل می‌تواند در تفسیر دست‌نوشته و جستجو در اسناد بسیار حائز اهمیت باشد. در این مقاله سیستمی به منظور تشخیص برون‌خط دست‌نوشته فارسی در یک فرهنگ لغت محدود معرفی شده است. به منظور استخراج ویژگی، بعد از بلوک‌بندی تصویر ورودی و استخراج مرکز هر بلوک توسط مرکز ثقل، میانگین مرکز اجزای متصل از الگوریتم سیفت متراکم استفاده شده است. از روش آنالیز تفکیک‌کننده خطی برای کاهش تعداد ویژگی‌ها استفاده شده است. در این مقاله، در مرحله نخست کلمات موجود در فرهنگ لغت بر اساس شباهت با استفاده از الگوریتم ISOCLUSE به همراه الگوریتم سلسله‌مراتبی خوشه‌بندی شده است. میانگین هر خوشه در فضای ویژگی به عنوان نماینده آن خوشه و مدخل مشترک اعضای آن خوشه در فرهنگ لغت تصویری، در نظر گرفته شده است. در مرحله دوم به منظور بازشناسی کلمه‌ی جدید از کلمات کاندید، از روش ماشین‌بردار پشتیبان به صورت چندکلاسه و دوکلاسه استفاده شده است. نتایج بدست آمده نشان دهنده عملکرد بهتر روش ماشین‌بردار پشتیبان دو کلاسه از لحاظ سرعت و دقت بر روی مجموعه داده ایران‌شهر می‌باشد. در مرحله بازشناسی با انتخاب ۵ خوشه نزدیک به کلمه دست‌نوشته مورد آزمون با دقت ۹۳٫۳۷٪ حدود ۷۶٫۶۵٪ از کلمات مورد بررسی کاهش خواهد یافت.

کلیدواژه‌ها

تشخیص کلمه دست‌نوشته، آنالیز تفکیک‌کننده خطی، سیفت متراکم، کاهش فرهنگ لغت، الگوریتم خوشه‌بندی ISOCLUSE، کلاس بند ماشین‌بردار پشتیبان دوکلاسه.

محققین در زبان‌های مختلف قرار گرفته است. بازشناسی یا پردازش اسناد، خواندن فرم‌های مخصوص، خواندن و مرتب‌کردن خودکار مرسولات پستی و چک‌های بانکی، بازیابی و مرتب‌کردن تصاویر متون دست‌نوشته تاریخی اسکن شده با کیفیت پایین از کاربردهای مهم بازشناسی دست‌نوشته هستند. تشخیص دست‌نوشته می‌تواند بصورت کل‌نگر یعنی در سطح کلمه و یا از بالا به پایین در سطح حروف انجام شود. در حالت بالا به پایین متون دست‌نوشته را به کلمه و کلمات را به حروف تشکیل دهنده‌اش تقسیم می‌کند و سپس عمل تشخیص انجام می‌گیرد. سیستم‌های تشخیص دست‌نوشته موجود به خصوص برای زبان‌های لاتین و چینی قادرند با دقتی در حدود ۹۹٪ - ۹۰٪ البته با یک سری محدودیت‌های آزمایشی دست‌نوشته‌ها را شناسایی کنند.

۱ مقدمه

بازشناسی دست‌نوشته شامل تبدیل خودکار تصویر اسکن شده متون دست‌نوشته به متون ماشینی است و یکی از کاربردهای جالب بازشناسی الگو در تصویر می‌باشد که مورد توجه بسیاری از

این مقاله در خردادماه سال ۱۳۹۵ دریافت، در آذرماه بازنگری و پذیرفته شد.

^۱ دانشجوی کارشناسی ارشد مهندسی کامپیوتر، گرایش علوم تصمیم و مهندسی دانش، دانشگاه آزاد اسلامی واحد مبارکه.

رایانامه: somaye.boroumand@gmail.com

^۲ دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه پیام‌نور.

رایانامه: iranpour@pnu.ac.ir

هستند. این ویژگی‌ها به نحوه نوشتن وابسته نیستند و نسبت به نویز حساس و استخراج آن‌ها بسیار مشکل است.

بازشناسی متون لاتین بر اساس شکل کلی کلمات با تحقیق [۱] شروع شد. در این تحقیق از روش تولید فرض برای تعیین کلمات با شکل کلی یکسان در یک دیکشنری کلمات استفاده شده است. در تحقیق [۲] از ویژگی‌های ناحیه‌ای شکل کلمه برای بازشناسی کلمات دست‌نویس لاتین استفاده شده است. در این مقاله از اطلاعات وجود یا عدم وجود بالارونده‌ها و پایین‌رونده‌ها، مکان تقریبی آن‌ها و ترتیب قرارگیری آن‌ها برای توصیف شکل کلی کلمات دست‌نویس لاتین استفاده شده است. بالارونده‌ها و پایین‌رونده‌ها با استفاده از اطلاعات ناحیه‌ای کلمه بدست می‌آیند، که با دو روش هیستوگرام افقی نقاط سیاه تصویر کلمه و بیشینه محلی تعداد نقاط سیاه تصویر کلمه استخراج می‌شود. در تحقیق [۳] سیستمی برای بازشناسی کلمات چاپی عربی با شبکه عصبی پیشنهاد شده است. بازشناسی کلمات در سیستم پیشنهادی به سه مرحله تقسیم می‌شود. در مرحله پیش‌پردازش، تصویر با یک پوشگر نوری با درجه تفکیک ۳۰۰ نقطه در اینچ، تصویربرداری شده و اجزاء متصل آن مشخص می‌شوند. سپس ویژگی‌های عمومی مانند تعداد اجزاء متصل، تعداد بیشینه‌ها در هر جزء متصل، تعداد و محل نقاط و علائم و ویژگی‌های هندسی دیگری از شکل کلی کلمه استخراج می‌شود.

روش‌های مبتنی بر ویژگی‌های آماری نسبت به نویز مقاوم هستند. ضرایب تبدیل فوریه، گشتاورها و پروفایل‌ها از این نوع ویژگی‌ها هستند. در تحقیق [۴] یک روشی برای تشخیص متون دست‌نوشته تاریخی لاتین اسکن شده با کیفیت پایین ارائه شده است. در تشخیص متون تاریخی بیشتر از روش تطبیق کلمه استفاده می‌شود. البته این روش نیز به دلیل پایین بودن کیفیت تصاویر با مشکلات زیادی روبه‌رو می‌شود. در تحقیق [۵] از ویژگی‌های شبه‌زیرینکی برای تشخیص متون دست‌نویس و چاپی با شبکه عصبی احتمالاتی استفاده شده است. روش ارائه شده در این تحقیق بر روش مبتنی بر شکل زیرکلمات استوار است. از شبکه عصبی برای بازشناسی استفاده شده است. در تحقیق [۶]، از ضرایب ویولت استخراج شده از پروفایل تصویر هموار شده در چهار جهت استفاده شده است. دلیل انتخاب این ویژگی ساختار ویژه زبان فارسی بیان شده است. در زبان فارسی، شکل حروف بسته به موقعیتش در کلمه پروفایل‌های مختلفی تولید می‌کنند. بنابراین ویژگی پروفایل تقریباً تمام ویژگی یک کلمه در زبان فارسی را بیان می‌کند. در مرحله تشخیص از شبکه عصبی تابع پایه‌ای شعاعی^۳ که برای معماری بهینه آن الگوریتم ژنتیک بکار رفته، استفاده شده است. در تحقیق [۷] یک ویژگی برای تشخیص متون دست‌نوشته معرفی شده است که نسبت به چرخش و اندازه تغییرناپذیر و مبتنی بر شکل کلی کلمات است. این تحقیق از m باند بسته‌ی تبدیل

ولی متاسفانه سیستمی با دقت مناسب که بتواند برای تشخیص دست‌نوشته فارسی عملکرد مناسبی از خود نشان دهند هنوز وجود ندارد و محققان بسیار در سال‌های اخیر به این موضوع پرداخته‌اند. یکی از دلایل این امر ماهیت ساختار نگارشی زبان فارسی هست. در همین راستا، این ضرورت احساس می‌شود که تحقیقات و پروژه‌های علمی متعددی در این زمینه انجام شود.

در این تحقیق روشی برای بازشناسی کلمه‌های دست‌نوشته فارسی ارائه شده است. این روش شامل دو مرحله اساسی می‌باشد. در مرحله اول برای تسریع تشخیص و کاهش محاسبات، روشی برای کاهش کلمات مورد بررسی در تشخیص کلمه مورد آزمون ارائه شده است. برای این منظور کلمات در فرهنگ لغت توسط الگوریتم‌های خوشه‌بندی، گروه‌بندی می‌شوند. برای هر خوشه تشکیل شده نماینده‌ای به عنوان مدخل اعضای آن خوشه در نظر گرفته شده است.

در مرحله دوم کلمه دست‌نوشته مورد آزمون از بین کلمات کاندید تشخیص داده می‌شود. تعدادی از خوشه‌های نزدیک به کلمه دست‌نوشته ورودی توسط الگوریتم انطباق زمانی پویا^۱ انتخاب می‌شود. کلمات موجود در این خوشه در لیست کاندید قرار می‌گیرند. در این مرحله از روش دسته‌بند ماشین‌بردار پشتیبان به صورت دو کلاسه و چند کلاسه استفاده شده است.

در هر دو مرحله ویژگی استفاده شده شامل موقعیت هر قسمت از کلمه در بلاک‌های جدا و روش سیفت متراکم می‌باشد. در ادامه به منظور کاهش ویژگی‌های استخراج شده توسط این روش‌ها، از الگوریتم آنالیز تفکیک‌کننده خطی^۲ استفاده شده است.

ساختار این مقاله به صورتی خواهد بود که در قسمت دوم به ادبیات تحقیق پرداخته شده و چالش‌های موجود در این زمینه مورد بررسی قرار گرفته شده است. در ادامه در قسمت سوم، کارهای انجام شده در زمینه کلمه دست‌نوشته برون-خط مبتنی بر شکل کلی کلمه مرور می‌شود. معرفی روش پیشنهادی در قسمت چهارم بوده و در ادامه توضیحات مربوط به چگونگی پیاده‌سازی و ارزیابی روش پیشنهادی مطرح شده است.

۲ کارهای انجام شده

به طور کلی می‌توان روش‌ها و تکنیک‌های موجود برای بازشناسی دست‌نوشته را بر اساس ویژگی‌های استفاده شده برای توصیف شکل کلی کلمه، به سه دسته: روش‌های مبتنی بر ویژگی‌های ساختاری یا توپولوژیکی، مبتنی بر ویژگی‌های آماری و مبتنی بر ویژگی‌های محلی یا ناحیه‌ای تقسیم کرد. البته بعضی از روش‌ها از ترکیب این سه دسته بوجود آمده‌اند.

روش‌های مبتنی بر ویژگی‌های ساختاری، ساختار کلی و توپولوژی هر کلمه را توصیف می‌کنند. تعداد عناصر متصل، نقاط، حفره‌ها، بالارونده‌ها، پایین‌رونده‌ها و ... مثال‌هایی از ویژگی‌های ساختاری

^۱ Dynamic Time Warping (DTW)

^۲ Linear Discriminant Analysis (LDA)

^۳ Radial Basis Function (RBF)

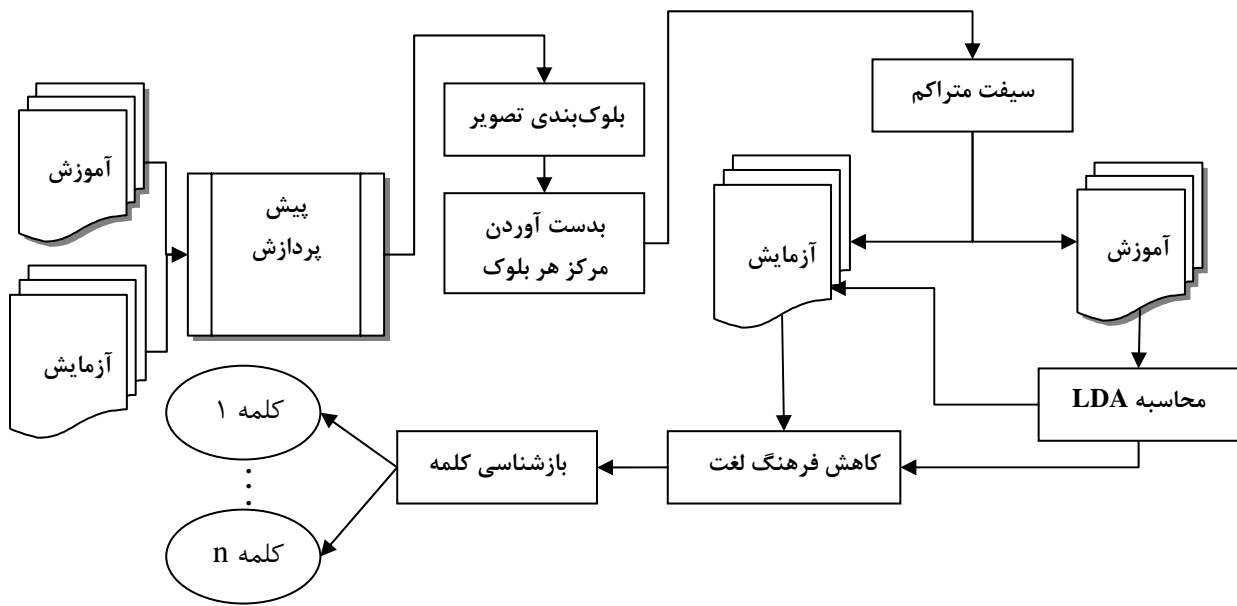
می‌گیرد: یافتن خوشه‌های نزدیک به ورودی و یافتن نزدیک‌ترین زیرکلمه از بین خوشه‌های نزدیک و روش دوم علاوه بر این مراحل شامل یک مرحله‌ی اضافی برای یافتن زیرکلمه نهایی بر اساس الگوی نقاط می‌باشد. ویژگی که در مرحله آموزش برای خوشه‌بندی استفاده شده، میانگین بلوکی است. برای بدست آوردن این ویژگی، تصویر کلمه به بلوک‌هایی تقسیم می‌شود و از هر بلوک میانگین تعداد پیکسل‌های سیاه محاسبه می‌شود. این ویژگی کلیات شکل را بیان می‌کند. در مرحله تشخیص با استفاده از فاصله بلوک شهری تعدادی از خوشه‌های نزدیک به کلمه ورودی انتخاب می‌شود. سپس با استفاده از ترکیب ویژگی‌های هیستوگرام‌گردادیان و تبدیل فوریه کانتور، نزدیک‌ترین زیرکلمات انتخاب می‌شود. هر دو روش نتایج قابل قبولی روی تصاویر تمیز ارائه می‌دهند. اما در برخورد با تصاویر کم کیفیت و نویزی دچار افت دقت می‌شوند که این کاهش در برخی موارد بسیار شدید است. در تحقیق [۱۳] تاثیر طبقه‌بند K نزدیکترین همسایه^۲ و نزدیکترین همسایه در تشخیص کلمات دست‌نویس عربی بررسی شده است. الگوریتم پیشنهادی شامل سه مرحله پیش‌پردازش، استخراج ویژگی و کلاس بندی می‌باشد. در تحقیق [۱۴] از ترکیب سه مدل پنهان مارکوف یکسان که با ویژگی‌های متفاوت آموزش داده شده‌اند، برای تشخیص کلمات دست‌نویس پایگاه داده IFN/ENIT استفاده شده است. در تحقیق [۱۵] روشی برای تشخیص زیرکلمات چاپی فارسی ارائه شده است. در این تحقیق از ویژگی‌های کلی و جزئی برای توصیف شکل کلی زیرکلمات در یک روش سلسله-مراتبی استفاده شده است. تصاویر کامل زیرکلمات با استفاده از ویژگی مکان مشخصه که اطلاعات کلی شکل زیرکلمات را می‌دهد برای کاهش دامنه جستجو خوشه‌بندی استفاده شده است. تعداد مناسب خوشه‌ها برای تصاویر کامل زیرکلمات، با دو شاخص اعتبارسنجی خوشه‌بندی محاسبه شده است. میانگین هر خوشه به عنوان نماینده آن خوشه در نظر گرفته می‌شود. برای ارزیابی خوشه‌ها از طبقه‌بند حداقل میانگین فاصله بر روی ۵۰۰۰ زیرکلمه با ۴ نوع قلم در ۳ اندازه مختلف استفاده شده است. دقت تشخیص در انتخاب نزدیکترین خوشه ۷۸/۷۱٪، در پنج خوشه اول ۹۹/۱٪ و در ده خوشه اول ۱۰۰٪ گزارش شده است. تحقیق [۱۶] را می‌توان اولین تحقیق در تشخیص متون دست‌نویسه فارسی دانست که دارای مرحله کاهش فرهنگ لغت است. مرحله کاهش فرهنگ لغت توسط کاهش کلمات نامشابه با کلمه ورودی سبب افزایش سرعت و دقت سیستم در مرحله تشخیص می‌شود. معمولاً هنگامی که فرهنگ لغت کوچک باشد دقت تشخیص مهم‌تر از زمان تشخیص است اما در فرهنگ لغت بزرگ سرعت تشخیص یک معیار مهم تلقی می‌شود. در این تحقیق از تعداد و موقعیت نقاط در کلمه برای کاهش تعداد کلماتی که در تشخیص کلمه آزمون ورودی بکار می‌رود، استفاده شده است.

موجک تصویر قطبی کلمه برای استخراج بردار ویژگی استفاده کرده است. فاصله ماهالانویس برای اندازه‌گیری فاصله بین بردار ویژگی استخراج شده از کلمه ورودی و کلمات فرهنگ لغت استفاده شده است.

روش‌های مبتنی بر ویژگی‌های محلی تصاویر را به بلوک‌های کوچکتر تقسیم می‌کنند و از هر یک از این بلوک‌ها، ویژگی‌های لازم استخراج می‌شود. ویژگی‌های محلی می‌تواند جزئیات شکل را بهتر از ویژگی‌های کلی توصیف کنند. در تحقیق [۸] از مدل مخفی مارکف برای بازشناسی اسامی دست‌نویس شهرهای ایران استفاده شده است. در این کار، از اطلاعات کانتور کلمات استفاده شده است. تصویر کلمات در راستای افقی به پنج قسمت مساوی و در راستای عمودی به قسمت‌هایی که ۵۰٪ باهم همپوشانی دارند تقسیم شده است. در هر یک از این پنجره‌ها، برداری شامل هیستوگرام‌های جهتی کدهای زنجیره‌ای بعنوان مدل کلمه انتخاب می‌شود. برای کم کردن تعداد مشاهده‌هایی که به مدل مخفی مارکف گسسته اعمال می‌شود، فضای ویژگی با استفاده از یک شبکه عصبی خودسامان کوهن، کوانتیزه شده است. برای هر اسم شهر یک مدل پنهان مارکوف^۱ گسسته با الگوریتم بام-ولش بطور مجزا آموزش داده می‌شود. برای هر تصویر ورودی، اسامی شهرها بر حسب میزان شباهت مدل آن‌ها به کلمه ورودی مرتب می‌شوند. در تحقیق [۹] از ویژگی‌های بر پایه گردادیان برای دسته‌بندی کلمات دست‌نویسه مجموعه داده‌ی ایران‌شهر استفاده شده است. در این تحقیق دسته‌بندی توسط دو شبکه عصبی پرسپترون و یک شبکه تابع بر پایه شعاعی انجام شده است. نتایج بدست آمده در این مقاله برای مجموعه داده ایران‌شهر ۸۹،۱۲ گزارش شده است. از تحقیقات دیگر انجام شده بر روی مجموعه داده ایران‌شهر می‌توان به تحقیق انجام شده در مرجع [۱۰ و ۱۱] اشاره نمود که در مرجع [۱۰] به مانند مرجع [۹] از ویژگی مبتنی بر گردادیان استفاده شده است و در ادامه برای دسته‌بندی دست‌نویسه از چندین شبکه عصبی پرسپترون به همراه روش الگو تصمیم استفاده شده است. در واقع خروجی شبکه پرسپترون توسط روش الگوهای تصمیم ترکیب شده‌اند. در این تحقیق دقت ۹۱،۵۵ برای شناسایی دست‌نویسه ایران‌شهر گزارش شده است. آقای ابراهیمی و همکاران در مرجع [۱۱] از ترکیب خبره‌ها برای دسته‌بندی استفاده کرده‌اند. در این مقاله از شبکه عصبی پرسپترون با ترم‌های ممنوع استفاده شده است و نتیجه ۹۱،۱۱ برای دسته‌بندی ایران‌شهر گزارش شده است. در تحقیق [۱۲] دو روش برای بازشناسی متون تایپی فارسی یا استفاده از شکل کلی تصویر ارائه شده است. در هر دو روش مرحله‌ی آموزش، زیرکلمات مجموعه‌ی آموزش، خوشه‌بندی می‌شوند. برای خوشه‌یابی از الگوریتم ISODATA استفاده شده و مراکز اولیه‌ی خوشه‌ها توسط یک الگوریتم خوشه‌یابی سلسله‌مراتبی تهیه شده‌اند. در روش اول تشخیص طی دو مرحله صورت

² K- Nearest Neighbor (KNN)

¹ Hidden Markov Model (HMM)



شکل ۱ نمای کلی از روش پیشنهادی

گرایان و کلاس‌بند ماشین بردار پشتیبان^۱ پیشنهاد شده است. با توجه به ماهیت پیوسته بودن متون دست‌نوشته اردو، در این تحقیق از ویژگی ساختاری و گرایان که مبتنی بر شکل کلی کلمه هستند و نیازی به جداسازی کلمات به اجزای سازنده‌اش نیست، استفاده شده است. بعد از انجام پردازش‌های اولیه دو نوع ویژگی استخراج می‌شود: ویژگی گرایان که یک ویژگی جهتی است و توسط گرایان تصویر خاکستری شده تولید می‌شود و ویژگی ساختاری شامل ویژگی‌های پروفایل، ویژگی‌های توپولوژیکی و ویژگی افکنش و... است. که در این تحقیق فقط از ویژگی پروفایل بالایی که ساختار بیرونی کلمه را نشان می‌دهد، استفاده شده است. برای کلاس‌بندی ۵۷ کلاس از یک پایگاه داده شامل ۵۷ کلمه از کلاس‌بند ماشین بردار پشتیبان استفاده شده است و نرخ تشخیص ۹۷٪ گزارش شده است.

۳ روش پیشنهادی

در این قسمت به بررسی فازهای مختلف روش پیشنهادی پرداخته شده است. فرایند کلی استفاده شده در این مقاله در شکل (۱) نشان داده شده است.

۳-۱ پیش‌پردازش

گام نخست در بازشناسی دست‌نوشته فاز مربوط به پیش‌پردازش بر روی تصاویر ورودی می‌باشد. در مراحل پیش‌پردازش به‌منظور بهبود تصویر ورودی باینری کردن، هموارسازی و حذف نویز، آنالیز مولفه‌های متصل، اصلاح کجی تصویر و نازک‌سازی و نرمال‌سازی اندازه استفاده شده است.

۳-۱-۱ هموارسازی و حذف نویز

هموارسازی تصاویر دو سطحی برای کاهش نویز و درست کردن لبه‌ها، می‌باشد. این مرحله برای حذف نویزهایی مثل نمک و فلفل

نقاط، کوچکترین واحد در یک متن فارسی و عربی است. تعدادی از حروف در زبان فارسی و عربی دارای شکل یکسان می‌باشند و وجه تمایز بین آن‌ها تعداد و موقعیت نقاط می‌باشد. شکل نقاط در دست‌نوشته‌ها بسیار متغیر است. قبل از استخراج نقاط خط اصلی کلمه محاسبه می‌شود. خط اصلی محل اتصال حروف به یکدیگر می‌باشد. خط اصلی در این تحقیق با استفاده از نقاط مینیمم محلی کانتور کلمه و استفاده از رگرسیون خطی بر روی این نقاط تقریب زده می‌شود. در گام بعدی با استفاده از یک سری قواعد نقاط استخراج می‌شود و یک عدد و یک حرف به هر نقطه تخصیص داده می‌شود. در واقع برای هر کلمه یک کد توصیفگر تولید می‌شود. با استفاده از روش تطبیق رشته دمرا- لون اشتین کد توصیفگر کلمه‌ی ورودی با تمام کدهای توصیفگر کلمات موجود در فرهنگ لغت مقایسه می‌شود و k تا از بهترین کلمات کاندید انتخاب می‌شود. مقدار k وابسته به نرخ دقت تشخیص و درجه کاهش فرهنگ لغت است. برای هر کلمه موجود در فرهنگ لغت یک مدل پنهان مارکوف گسسته با استفاده از ویژگی انتقال ساخته می‌شود. ویژگی انتقال، مکان و تعداد انتقال‌ها از پیکسل سیاه به سفید را در هر ستون، از بالا به پایین و از پایین به بالا، محاسبه می‌کند. در مرحله تشخیص کلمه آزمون ورودی تنها مدل پنهان مارکوف کلماتی شرکت می‌کند که در مرحله کاهش فرهنگ لغت انتخاب شده‌اند. این الگوریتم بر روی یک پایگاه داده شامل ۱۷۰۰۰ تصویر از اسامی ۲۰۰ شهر ایران آزمایش شده است. نرخ تشخیص در انتخاب اول بدون کاهش فرهنگ لغت ۶۰/۲٪ و با کاهش فرهنگ لغت ۷۳/۶۱٪ گزارش شده است.

در تحقیق [۱۷] آزمایش‌هایی بر روی زبان اردو انجام شده است. زبان اردو همانند زبان فارسی و عربی است اما تعداد حروفش بیشتر می‌باشد. در این تحقیق الگوریتمی با استفاده از ویژگی

¹ Support Vector Machine (SVM)

در بعضی از دست‌نوشته‌ها فاصله بین زیرکلمات بسیار زیاد است. در این مرحله برای یکسان کردن فاصله بین زیرکلمات، فضای خالی بینشان حذف می‌شود. در بعضی از دست‌نوشته‌ها، بعضی از حروف یا زیرکلمات بدون اینکه به هم متصل شوند در زیرحروف یا زیرکلمات دیگر قرار می‌گیرند و سبب ایجاد خطا در استخراج بعضی از ویژگی‌ها می‌شود، برای حل این مشکل از الگوریتم کشیدن استفاده شده است [۲۰].

۳-۱-۴ نازک سازی

یکی از مهم‌ترین مراحل فاز پیش‌پردازش مرحله نازک‌سازی یا اسکلت‌بندی است. نازک‌سازی یعنی تبدیل شکل به منحنی‌ها و خطوطی با ضخامت یک پیکسل می‌باشد به‌گونه‌ای که ویژگی‌های ساختاری شکل حفظ شود. نازک‌سازی با یکسان کردن ضخامت کلمه در تصاویر کلمات، سبب بهبود دقت سیستم‌های تشخیص کلمه می‌شود.

۳-۱-۵ نرمال‌سازی

نرمال‌سازی اندازه تصاویر نقش مهمی در بهبود دقت شناسایی کلمات را ایفا می‌کند. در این مرحله اندازه همه تصاویر برای داشتن ابعاد ویژگی ثابت، یکسان می‌شود. در این مقاله از روش مقاله [۲۱] استفاده شده است.

۳-۱-۶ بلوک‌بندی

برای اینکه در گام بعدی بتوان از هر قسمت از تصویر تفسیر جداگانه‌ای را بیان نمود تصویر را بلوک‌بندی می‌نمایند. بلوک‌بندی در واقع تقسیم تصویر به زیرتصویر کوچکتر می‌باشد که هر بلوک می‌تواند اطلاعات مفیدی از کلمه موجود در تصویر باشد. در واقع در این روش محل قرار گرفتن قسمت‌های مختلف از کلمات جدا می‌شود و با توجه به محل قرار گرفتن (بلوک مربوط به هر بخش-های مختلف از کلمات) ویژگی استخراج می‌شود. اندازه بلوک‌ها در تمام تصاویر می‌تواند ثابت یا متغیر باشد. تعداد بلوک‌ها در تمام تصاویر ثابت و به توافق بین دقت و پیچیدگی وابسته است. یعنی تعداد بلوک‌ها به‌گونه‌ای انتخاب می‌شود که پیچیدگی کم و در عین حال دقت بالا باشد. در این مقاله به منظور بالا بردن سرعت از بلوک‌بندی ثابت استفاده شده است در واقع تصویر به ۶ قسمت جدا تقسیم می‌شود و استخراج ویژگی در هر بلوک به‌صورت جداگانه انجام می‌شود. در این مقاله به‌مانند شکل (۲) تصویر به دو سطر و سه ستون مساوی تقسیم شده است. در تصویر زیر ابتدا با استفاده از فیلتر سوبل، لبه‌های تصویر استخراج می‌شود و سپس عمل بلوک‌بندی صورت می‌گیرد. در بلوک‌بندی کلمات باید دقت کرد که نباید بلوک‌ها زیاد بزرگ و یا زیاد کوچک باشند در صورت بزرگ بودن بلوک‌ها حروف زیادی در یک بلوک قرار گرفته و در واقع تأثیر بلوک‌بندی را کاهش می‌دهد و در صورت کوچک بودن بلوک‌ها این امکان وجود دارد که یک حرف در چند بلوک مجزا قرار گیرد و عملکرد سیستم پیشنهادی را کاهش دهد. به همین دلیل در این مقاله از یک نسبت منطقی دو سطر و سه ستون استفاده شده

و پر کردن شکاف‌ها و دره‌های کوچک و حذف دست‌اندازهای بوجود آمده در لبه‌ها مناسب می‌باشد. برای حذف نویزها و هموارسازی تصویر از فیلترهای میانگین و میانه استفاده شده است. گاهی اوقات نویزهای بوجود آمده در تصویر بزرگتر از یک پیکسل هستند و با به کار بردن یک بار فیلترهای میانه یا میانگین بر روی تصویر این نویزها از بین نمی‌روند. اگر چندین بار بر روی تصویر پیاده شود ممکن است بر روی لبه‌های حروف تأثیر بگذارد و این باعث کاهش دقت خواهد شد. در این مقاله از الگوریتم [۱۸] برای کاهش نویز استفاده شده است. در این روش ابتدا، تصویر ورودی توسط فیلتر میانگین هموار می‌شوند در گام بعدی، اندازه گرادیان تصویر هموار شده توسط روش سوبل محاسبه می‌شود و هم‌پیکسل‌های منفرد از پیکسل دوسطحی حذف می‌شوند. برای هر مولفه متصل در تصویر دوسطحی، میانگین اندازه گرادیان پیکسل‌های لبه محاسبه می‌شود و در نهایت مولفه‌هایی که میانگین اندازه گرادیان پیکسل‌های لبه‌اش کمتر از یک حد آستانه باشد حذف می‌شوند.

۳-۱-۲ اصلاح کجی

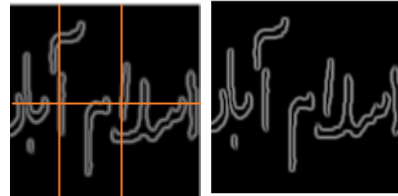
کجی، انحراف میانگین پاره‌خط‌های نزدیک به عمودی نسبت به جهت عمودی می‌باشد. این مرحله در فاز پیش‌پردازش برای کاهش بازه اختلافات در متون دست‌نوشته و تایپی لازم است. کجی در متون دست‌نوشته به سبب نحوه نوشتن خاص یک فرد و در متون تایپی به دلیل خصوصیات ذاتی برخی فونت‌ها بوجود می‌آید. گاهی اوقات کجی‌های موجود باعث کاهش نرخ تشخیص می‌شود. کجی در یک کلمه می‌تواند به دو صورت باشد: کجی یکنواخت هنگامی که پاره‌خط‌های نزدیک به خط عمودی دارای زاویه کجی یکسان باشند. کجی غیر یکنواخت هنگامی است که زاویه کجی در پاره‌خط‌های نزدیک به خط عمودی متفاوت باشند که بیشتر در متون دست‌نوشته دیده می‌شود. در همه تکنیک‌های اصلاح کجی یکنواخت، ابتدا زاویه کجی تخمین می‌شود سپس توسط تبدیل shear نسبت به خط افقی تصویر اصلاح می‌شود. در این مقاله از روش ارائه شده در [۱۹] استفاده شده است. این الگوریتم بر این ایده استوار است که کلمه نوشته شده بدون کجی دارای پروفایل افکنش عمودی با قله‌های بلند و دره‌هایی عمیق است. بر اساس این ایده، هر تصویر با زاویه‌های مختلف کجی چرخانده می‌شود و پروفایل افکنش عمودی آن کشیده می‌شود، هر تصویری که قله‌های بلندتر و دره‌های عمیق‌تر داشته باشد به عنوان تصویر اصلاح شده و زاویه کجی آن به عنوان زاویه اصلاح کجی در نظر گرفته می‌شود.

۳-۱-۳ آنالیز زیرکلمات

در متون دست‌نوشته تنوع در شیوه نگارش بسیار زیاد است و این باعث پیچیده شدن فرآیند تشخیص کلمات می‌شود. فاز پیش-پردازش با انجام یک سری عملیات بر روی تصاویر این تغییرات و تنوع را کاهش می‌دهد. یکی از این عملیات آنالیز زیرکلمات است.

است. البته باید به این نکته اشاره نمود که می‌توان اندازه بلوک‌ها را وابسته به اندازه حروف نیز در نظر گرفت. با توجه به نتایج بدست آمده در جدول (۱) بر روی مجموعه داده اعتبارسنجی، وابسته بودن بلوک‌بندی به اندازه تصویر (کلمات) عملکرد بهتری داشته است. در این آزمایشات فقط از ماشین بردار پشتیبان دو کلاس با در نظر گرفتن دو نوع بلوک‌بندی متفاوت استفاده شده است.

است. البته باید به این نکته اشاره نمود که می‌توان اندازه بلوک‌ها را وابسته به اندازه حروف نیز در نظر گرفت. با توجه به نتایج بدست آمده در جدول (۱) بر روی مجموعه داده اعتبارسنجی، وابسته بودن بلوک‌بندی به اندازه تصویر (کلمات) عملکرد بهتری داشته است. در این آزمایشات فقط از ماشین بردار پشتیبان دو کلاس با در نظر گرفتن دو نوع بلوک‌بندی متفاوت استفاده شده است.



شکل ۲ بلوک‌بندی تصویر، تصویر سمت تصویر اصلی، تصویر سمت راست تصویر بلوک‌بندی شده

جدول ۱ کلاس‌بندی داده‌های اعتبارسنجی با بلوک‌بندی‌های متفاوت

دقت	SVM دو کلاس با بلوک‌بندی وابسته به اندازه حرف	SVM دو کلاس با بلوک‌بندی وابسته به اندازه تصویر (کلمات)
۶۹,۵۱	۷۵,۳۲	

۲-۳ استخراج ویژگی

یکی از مهم‌ترین مراحل در سیستم‌های تشخیص الگو، مرحله استخراج ویژگی می‌باشد. هدف از استخراج ویژگی تبدیل تصویر به یک مجموعه‌ای از بردارهای ویژگی است که بتواند تصویر را به خوبی توصیف کند و در مرحله کلاس‌بندی و تشخیص به کار رود.

۱-۲-۳ موقعیت هر بلاک

یکی از ویژگی‌هایی که در این تحقیق استفاده شده موقعیت بلاک-ها می‌باشد. به همین منظور بعد از اعمال بلاک بندی، مرکز تصویر را به عنوان ابتدای مبدأ مختصات در نظر گرفته و موقعیت هر بلاک با توجه به این مبدأ مختصات به عنوان ویژگی در نظر گرفته شده است. برای محاسبه مرکز هر بلاک از مرکز ثقل استفاده شده است. در واقع ابتدا مرکز ثقل هر جز متصل از کلمات موجود محاسبه شده و در نهایت میانگین مراکز ثقل موجود به عنوان مرکز هر بلاک محاسبه می‌شود. مرکز ثقل یک جسم همیشه روی مرکز هندسی آن نیست و نقطه‌ی دیگری می‌تواند مرکز ثقل جسم باشد. مرکز ثقل برای هر جسم در واقع، محل اثر برآیند نیروهای استاتیکی (مانند وزن) بر جسم می‌باشد. برای بدست آوردن مرکز ثقل هر شی از رابطه (۱) که در مرجع [۲۲] معرفی شده استفاده شده است.

$$X_m = \frac{\sum_{i=1}^n \sum_{j=1}^m ib(i, j)}{A}, Y_m = \frac{\sum_{i=1}^n \sum_{j=1}^m jb(i, j)}{A} \quad (1)$$

۲-۲-۳ ویژگی سیفت^۱ متراکم (Dense Sift) توصیفگر سیفت یک روش تناظریابی مبتنی بر ویژگی است که در سال ۱۹۹۱ ارائه شده است و اولین بار در سال ۲۰۰۴ جهت شناسایی الگو در تصاویر بصری استفاده شد [۲۳]. این توصیفگر از ۴ مرحله تشکیل شده است. اولین مرحله، ایجاد فضای مقیاس، نمایش ساختارهای تصویر در مقیاس‌های مختلف است که از مجموعه‌ای از تصاویر گوسی و تفاوت گاوسی^۲ در ابعاد مختلف تشکیل شده است. این تصاویر در لایه‌های منظمی تحت عنوان اکتاو گروه‌بندی و مرتب شده‌اند. تصویر ورودی در روند تکراری با کرنل گوسی در مقیاس‌های مختلف کانولوشن می‌شود و در نهایت تصاویر گوسی فضای مقیاس در هر اکتاو ایجاد می‌شود. هر یک از پیکسل‌های تصویر در تفاوت گاوسی با ۸ پیکسل همسایه از همان مقیاس و ۹ پیکسل همسایه از مقیاس بالا و پایین مقایسه شده و چنانچه مینیمم یا ماکزیمم این پیکسل‌ها باشد، به عنوان نقطه اکستریم انتخاب خواهد شد.

در مرحله دوم نقاط استخراج شده بیشتر مورد بررسی قرار می‌گیرند تا نقاط پایدارتر انتخاب شوند. پایداری برای هر نقطه به صورت $|D(x, y, \sigma)|$ مشخص شده و با استفاده از حد آستانه، نقاط با وضوح پایین و ناپایدار در لبه‌ها حذف خواهند شد. در مرحله سوم به منظور مقاوم بودن الگوریتم در برابر چرخش، برای تمام نقاط استخراج شده پارامتر اندازه‌گرادیان و جهت تعیین می‌گردد. در مرحله چهارم متناظر با هر نقطه، یک توصیفگر به عنوان ابزاری برای تناظریابی میان نقاط ایجاد می‌شود. توصیفگر، هیستوگرامی سه‌بعدی از شیب و جهت است. شیب یک شبکه 4×4 و جهت نیز در ۸ قسمت تفکیک شده است. ویژگی‌های نهایی هر یک از نقاط استخراج شده، در یک بردار یکه با ۱۲۸ مولفه $(4 \times 4 \times 8)$ قرار می‌گیرند.

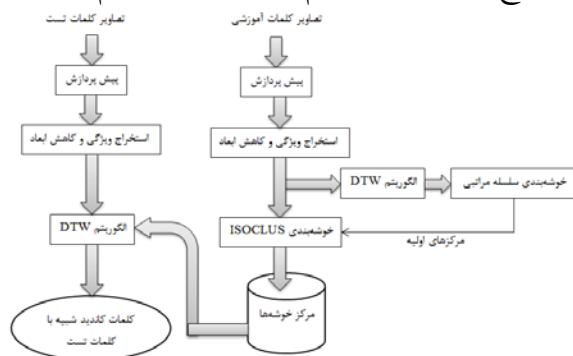
در سال‌های اخیر روشی به منظور بهبود روش سیفت با نام سیفت متراکم ارائه شده است [۲۴]. در این مقاله نقاط خروجی این روش به عنوان ویژگی برای شناسایی استفاده می‌شود. الگوریتم سیفت شامل دو مرحله شناسایی و توصیفگر می‌باشد که در الگوریتم سیفت متراکم فقط شامل مرحله توصیفگر می‌باشد. این روش به دلیل وابسته نبودن به اندازه و همچنین زاویه و جهت حرف می‌تواند بسیار حائز اهمیت باشد. در سال‌های اخیر از الگوریتم سیفت برای دسته‌بندی تصاویر کلمات فارسی استفاده شده است که از جمله آن می‌توان به مرجع [۲۵] اشاره نمود که از الگوریتم سیفت برای تشخیص فونت متون فارسی و عربی استفاده شده و عملکرد خوبی را داشته است. ولی تاکنون از سیفت متراکم

¹ Scale-Invariant Feature Transform (SIFT)

² Difference of Gaussian (Dog)

۳-۱ خوشه‌بندی

هدف از تحلیل خوشه، شناسایی گروه‌هایی از اشیاء مشابه و آشکار کردن توزیع الگوها و همبستگی‌های موجود در مجموعه داده‌های بزرگ می‌باشد. خوشه‌یابی زمینه تحقیقاتی گسترده‌ای در کاربردهای مهندسی، اقتصاد و علوم اجتماعی داشته و دارد. در سال‌های اخیر به دلیل فراهم بودن مجموعه داده‌های تجربی بزرگ و نیاز فزاینده به کاوش این داده‌ها، وجود الگوریتم‌های کارآمد و قابل اعمال در حوزه‌های مختلف ضروری به نظر می‌رسد. هدف اصلی فرآیند خوشه‌بندی، سازمان‌دهی داده‌ها به گروه‌های معنادار است به نحوی که شباهت‌ها و تضادها قابل مشاهده باشند و بتوان نتیجه‌گیری مناسبی در مورد آن‌ها انجام داد. روش‌های خوشه‌بندی متعددی در مقالات ارائه شده‌اند. الگوریتم‌های خوشه‌بندی را می‌توان به پنج دسته کلی، الگوریتم‌های ترتیبی، الگوریتم‌های سلسله



شکل ۳ روند کلی برای کاهش فرهنگ لغت

مراتبی، الگوریتم‌های خوشه‌بندی افزایی، الگوریتم‌های بر اساس شبکه، الگوریتم‌های بر اساس تراکم، دسته‌بندی نمود. در این مقاله از دو الگوریتم سلسله مراتبی و افزایی استفاده شده است.

● الگوریتم‌های سلسله مراتبی

در روش خوشه‌بندی سلسله مراتبی، به خوشه‌های نهایی بر اساس میزان عمومیت آن‌ها ساختاری سلسله مراتبی، معمولاً به صورت درختی نسبت داده می‌شود. به این درخت سلسله مراتبی دندوگرام (dendrogram) می‌گویند. روش کار تکنیک‌های خوشه‌بندی سلسله‌مراتبی معمولاً بر اساس الگوریتم‌های حریصانه (Greedy Algorithms) و بهینگی مرحله‌ای (stepwise-optimal) است.

درخت دندروگرام ارتباط بین خوشه را نشان می‌دهد. با قطع دندروگرام در یک گره دلخواه، یک افراز از مجموعه داده به خوشه‌های مجزا بدست می‌آید. در مواردی که یک خوشه‌بندی مناسب برای یک مجموعه داده‌ی مشخص لازم است باید بررسی شود که کدام یک از افرازهای تولید شده توسط الگوریتم سلسله مراتبی برای مجموعه داده بهتر است و همان‌جا الگوریتم پایان داده شود. الگوریتم‌های سلسله مراتبی با پیچیدگی محاسباتی $O(N^3)$ (N تعداد نمونه‌ها) برای مجموعه داده‌های بزرگ مناسب نیستند.

● الگوریتم‌های خوشه‌بندی افزای

در این روش‌ها، مجموعه داده مستقیماً به مجموعه‌ای از خوشه‌های جدا از هم تجزیه می‌شود. به عبارت بهتر، در این روش‌ها

برای شناسایی کلمات دست‌نوشته و تایپی استفاده نشده است که در این مقاله از این روش استفاده شده است.

۳-۲-۳ استفاده از آنالیز تفکیک‌پذیری خطی به منظور

ارزش‌دهی و کاهش ویژگی

بعد از مرحله استخراج ویژگی، نوبت به کاهش ویژگی و تفسیر ویژگی می‌رسد. در این راستا از روش تجزیه و تحلیل مولفه‌های اصلی برای کاهش و ارزش‌دهی به ویژگی‌ها استفاده شده است. این روش بهترین روش برای کاهش ابعاد داده به صورت خطی می‌باشد.

جدول ۲ کلاس‌بندی داده‌های اعتبارسنجی با روش PCA و LDA

	PCA دو کلاسه SVM	LDA دو کلاسه SVM
دقت	۷۵,۸۷	۸۳,۶۵

یعنی اطلاعات از دست رفته با حذف ضرایب کم‌اهمیت نسبت به روش‌های دیگر کمتر است [۲۶]. در این روش جداسازی بر اساس کلاس صورت می‌گیرد و با افزایش پراکندگی بین کلاس‌ها و کاهش پراکندگی درون کلاسی جداسازی بین کلاس‌ها بهبود می‌یابد. این روش عملکرد مناسبی در کاهش ویژگی از خود نشان می‌دهد. در این الگوریتم، N نمونه با ابعاد D به فضایی با ابعاد M ($M < D$)، که در آن واریانس بین داده‌های جدید حداکثر می‌باشد نگاشت داده می‌شوند. یعنی در واقع برای جداسازی بهتر داده‌ها، پراکندگی بین داده‌ها در کلاس‌های متفاوت بیشتر می‌شود. البته باید به این نکته اشاره نمود روش تجزیه و تحلیل مولفه‌های اصلی (PCA) نیز بر روی مجموعه داده اعتبارسنجی مورد ارزیابی قرار گرفت که نسبت به روش LDA عملکرد بهتری را از خود نشان نداده است. جدول (۲) نتایج بدست آمده از مقایسه بین نتایج بین PCA و LDA بر روی ویژگی‌های استخراج شده از سیف متراکم (بدون بلوک‌بندی و مرکز ثقل) را نشان می‌دهد. همان‌طور که مشخص می‌باشد روش LDA عملکرد بهتری داشته است.

۳-۳ کاهش فرهنگ لغت

هرچه تعداد کلمات موجود در فرهنگ لغت بیشتر باشد به همان اندازه دقت و سرعت سیستم‌های تشخیص کلمه کاهش می‌یابد. بعضی از سیستم‌ها برای کاهش مشکلات ناشی از فرهنگ لغت بزرگ، در ابتدا تعداد کلمات مورد بررسی برای تشخیص کلمه آزمون ورودی را کاهش می‌دهند. این کار نه تنها سرعت سیستم را بالا می‌برد بلکه دقت را نیز افزایش می‌دهد. در این تحقیق، برای کاهش فرهنگ لغت از خوشه‌بندی تصاویر کلمات دست‌نوشته استفاده شده است. ابتدا ویژگی‌های معرفی شده از تصویر استخراج شده و سپس از توالی سلسله مراتبی و ISOCCLUS برای خوشه‌بندی تصاویر کلمات دست‌نوشته استفاده می‌شود. مراحل کلی استفاده از این روش‌ها در شکل (۳) نشان داده شده است.

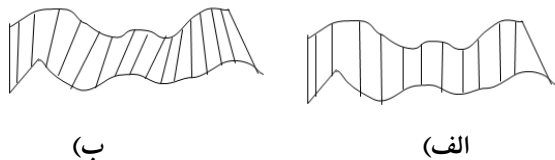
تعیین تعداد اولیه خوشه‌ها استفاده شده است. با این کار هم می‌توان تعداد خوشه‌های مناسب را تخمین زد و هم مراکز اولیه برای الگوریتم ISOCLUS را آماده کرد.

۳-۳-۲ معیار فاصله

معیار فاصله برای محاسبه فاصله بین دو بردار استفاده خواهد شد. روش‌های زیادی برای محاسبه معیار فاصله وجود دارد که از جمله می‌توان به فاصله مینکوسکی، زاویه‌ای، انطباق زمانی پویا اشاره نمود که در این تحقیق از معیار انطباق زمانی پویا استفاده شده است.

• الگوریتم انطباق زمانی پویا (DTW)

یکی از الگوریتم‌هایی که برای مقایسه و اندازه‌گیری میزان شباهت کلمات دست‌نوشته می‌توان به کاربرد، الگوریتم انطباق زمانی پویا است. این الگوریتم برای تطبیق سیگنال‌های یک‌بعدی به کار می‌



شکل ۴ محاسبه فاصله انطباق با دو مجموعه زمانی، (الف) انطباق یا

روش ساده، (ب) انطباق زمانی پویا

رود. برای به کار بردن این الگوریتم در تطبیق تصاویر، باید تصویر به سیگنال‌های یک‌بعدی تبدیل شوند. الگوریتم انطباق زمانی پویا برای محاسبه فاصله بین دو مجموعه زمانی استفاده می‌شود. یک روش ساده برای محاسبه فاصله انطباق بین دو مجموعه زمانی، مقایسه نمونه به نمونه است. اما این روش نتیجه شهودی و قابل درکی نمی‌دهد چون ممکن است نمونه‌هایی که باهم مقایسه می‌شوند مشابه هم نباشند (شکل ۴-الف). الگوریتم انطباق زمانی پویا با تنظیم بهینه نمونه‌ها در دو مجموعه زمانی، اختلاف بین نمونه‌های متناظر و محاسبه را کاهش می‌دهد. این تنظیم بهینه با مینیمم کردن یک مجموعه‌ای از فاصله‌های محلی نمونه‌ها بدست می‌آید. (شکل ۴-ب)

فاصله انطباق زمانی پویا بین دو مجموعه زمانی x_1, x_2, \dots, x_M و y_1, y_2, \dots, y_N ، با استفاده از روش برنامه نویسی پویا به صورت زیر محاسبه می‌شود.

$$D(i, j) = \begin{cases} D(i, j-1) \\ D(i-1, j) \\ D(i-1, j-1) \end{cases} + d(x_i, y_j) \quad (2)$$

انتخاب تساوی بازگشتی و تابع فاصله محلی $d(.,.)$ به نوع کاربرد بستگی دارد. با استفاده از $D(i, j-1)$ ، $D(i-1, j)$ ، $D(i-1, j-1)$ یک قید پیوستگی در محاسبه $D(i, j)$ ایجاد می‌شود که باعث هموار شدن انطباق زمانی می‌شود. مسیر انطباق زمانی پویا، با دنبال کردن جفت شاخص‌های هزینه مینیمم شده، (i, j) بدست

تعداد مناسب افزایها با بهینه‌کردن يك تابع هزینه تعیین می‌شوند. تابع هزینه ممکن است بر ساختار کلی یا محلی داده‌ها تاکید داشته باشد و فرآیند بهینه‌سازی آن يك فرآیند تکراری است. یکی از الگوریتم‌های معرف خوشه‌بندی الگوریتم ISOCLUS است که در بسیاری از کاربردها به صورت گسترده از آن استفاده می‌شود. این الگوریتم بر مبنای الگوریتم ISODATA با کمی اصلاحات می‌باشد و به طور کلی شبیه به الگوریتم k-means است با این تفاوت که در الگوریتم k-means تعداد خوشه‌ها ثابت است اما در الگوریتم ISOCLUS با حذف، تقسیم یا ادغام خوشه‌ها در هر مرحله، ممکن است تعداد خوشه‌های نهایی با تعداد خوشه‌های اولیه متفاوت باشد. هدف این الگوریتم پیدا کردن بهترین مرکز خوشه‌ها با یک روش تکراری می‌باشد. ورودی‌های این الگوریتم شامل: مجموعه داده آموزش، تعداد اولیه خوشه‌ها و یک سری پارامترهای از پیش تعیین شده برای حذف، تقسیم یا ادغام خوشه‌ها می‌باشد. این پارامترها توسط کاربر تعیین می‌شود و شامل، مینیمم تعداد نمونه‌ها در هر خوشه (SampRn)، ماکزیمم تعداد تکرار (MaxIter)، ماکزیمم انحراف استاندارد برای هر خوشه (stdDev)، ماکزیمم تعداد جفت خوشه‌هایی که در هر تکرار می‌توانند ادغام شوند (MaxPair)، ماکزیمم تعداد خوشه‌ها (MaxDis) و مینیمم فاصله بین خوشه‌ها (MinDis) می‌باشد.

مراحل اجرای این الگوریتم به صورت زیر می‌باشد [۲۷]:

مرحله اول: وارد کردن مقدار پارامترها به الگوریتم (مقدار پارامترها با مجموعه اعتبارسنجی بدست می‌آید)

مرحله دوم: فاصله هر داده تا مرکز خوشه‌ها محاسبه می‌شود و هر داده به نزدیک‌ترین خوشه بر چسب می‌خورند.

مرحله سوم: خوشه‌هایی که تعداد نمونه‌هایش کمتر از SampRn باشد حذف می‌شوند.

مرحله چهارم: مرحله دوم و سوم دوباره تکرار می‌شود تا اینکه هیچ خوشه‌ای در مرحله سوم حذف نشود.

مرحله پنجم: میانگین فاصله نقاط هر خوشه تا مرکز آن خوشه و میانگین کلی فاصله محاسبه می‌شود.

مرحله ششم: برای هر خوشه، انحراف معیار هر بعد و ماکزیمم انحراف محاسبه می‌شود با توجه به پارامترهای stdDev و MaxNum، خوشه‌های بزرگ از بعدی که بالاترین انحراف معیار را دارد نصف می‌شود.

مرحله هفتم: اگر در مرحله ششم خوشه‌ای نصف شد به مرحله دوم می‌رود در غیر این صورت از پارامترهای MinDis و MaxPair برای ادغام خوشه‌های نزدیک استفاده می‌شود.

مرحله هشتم: با توجه به مقدار پارامترهای MaxNum و MaxIter تعیین می‌شود که به مرحله دو باز گردد یا الگوریتم تمام شود.

تعداد و مراکز اولیه خوشه‌ها در نتیجه نهایی تاثیر دارد و انتخاب اشتباه باعث ایجاد نتایج اشتباه و غیر معتبر خواهد شد. به همین دلیل از الگوریتم سلسله مراتبی که مستقل از پارامتر است، برای

می‌دهند. این کار نه تنها سرعت سیستم را بالا می‌برد بلکه دقت را نیز افزایش می‌دهد. برای کاهش فرهنگ لغت از خوشه‌بندی تصاویر کلمات دست‌نوشته استفاده شده است. ابتدا پیش‌پردازشی با استفاده از روش‌های معرفی شده در قسمت‌های قبل برای بهبود کلمات و یکسان‌سازی آن‌ها صورت می‌گیرد. در گام بعدی بلوک-بندی بر روی تصاویر انجام می‌شود در ادامه ویژگی‌های معرفی شده در روش پیشنهادی، موقعیت مرکز هر بلوک با استفاده از مرکز ثقل و ویژگی سیفت متراکم استخراج می‌شود. در این مرحله تعداد ویژگی‌های استخراج شده برابر با ۷۸۰ نمونه می‌باشد که در هر بلاک ۱۲۸ ویژگی سیفت متراکم و ۲ ویژگی مربوط به موقعیت مرکز بلاک می‌باشد. روش کاهش ویژگی LDA بر ویژگی‌های استخراج شده اعمال می‌شود با توجه به ماهیت این روش تعداد ویژگی‌ها به ۵۰۲ کاهش می‌یابد (یکی کمتر از تعداد کلاس‌ها که ۵۰۳ بوده است). بعد از استخراج ویژگی از توالی سلسله مراتبی و ISOCLUS برای خوشه‌بندی تصاویر دست‌نوشته استفاده می‌شود. مراحل کلی این روش در شکل (۱) نشان داده شده است. همانطور که در شکل (۱) نشان داده شده مراحل الگوریتم شامل، پیش‌پردازش، استخراج ویژگی و خوشه‌بندی توسط الگوریتم سلسله‌مراتبی و ISOCLUSE می‌باشد. پارامترهای مربوط به این روش در جدول (۳) بیان شده است.

جدول ۳ پارامترهای استفاده شده برای روش ISOCLUSE

پارامتر	Samp Rn	Max Iter	Std Dev	Max Pair	Max Dis	Min Dis
مقادیر	۱۵	۳۰	۰.۳۵	۶	۱۰۰	۱۰

در این الگوریتم به منظور استخراج تعداد کلاس اولیه از الگوریتم خوشه‌بندی استفاده شده است. نمودار دندروگرام بدست آمده توسط این روش در شکل (۵) نشان داده شده است. با نگاه به این نمودار و دقت در ارتفاع شاخه‌های حاصله می‌توان تشخیص داد که توقف الگوریتم در کدام سطح فاصله خوشه‌های مناسبی را ایجاد خواهد کرد. برای بدست آوردن نقطه توقف از مجموعه داده اعتبار سنجی نیز می‌توان استفاده نمود. با توجه به نمودار بدست آمده بعد از توقف الگوریتم در سطح ۵۵ خوشه‌هایی بدست می‌آید که فاصله بین نمونه‌های آن‌ها به طور شهودی، معقول بوده و اندازه‌ی خوشه‌ها نیز مناسب است. از الگوریتم انطباق زمانی پویا برای اندازه‌گیری فاصله بین نمونه‌های بدست آمده از الگوریتم سلسله‌مراتبی استفاده شده است. دلیل استفاده از این الگوریتم وجود طول‌های متفاوت حروف در کلمات مشابه در نوشته‌های افراد مختلف است. الگوریتم انطباق زمانی پویا سعی می‌کند تا حدودی با استفاده از فاصله نقاط مجاور این تغییرات را جبران کند. با اجرای الگوریتم انطباق زمانی پویا بر روی مجموعه زمانی $X(I)$ و $Y(J)$ استخراج شده از دو تصویر I و J ، تابع فاصله محلی

می‌آید. برای اینکه مسیر انطباق نزدیک به قطر ماتریس $D(i,j)$ باشد، از محدودیت باند ساکو-چیبا استفاده می‌شود.

۳-۴ شناسایی کلمات

روش‌های زیادی برای شناسایی و دسته‌بندی کلمات وجود دارد ولی در این تحقیق از روش ماشین بردار پشتیبان به دو صورت چند کلاسه و دو کلاسه استفاده شده است. ماشین بردار پشتیبانی یکی از روش‌های یادگیری بانظارت است که از آن برای طبقه‌بندی و رگرسیون استفاده می‌کنند. در ماشین بردار پشتیبان دو کلاسه، برای هر جفت کلاس از کلاس‌ها، یک ماشین بردار پشتیبان ساخته می‌شود. بنابراین تعداد ماشین بردار پشتیبان برابر با $M(M-1)/2$ است. برای ساختن ماشین بردار پشتیبان دو کلاسه (k,m) ، تنها داده‌های آموزش این دو کلاس استفاده می‌شود. برای تعیین کلاس یک نمونه آزمون، معمولاً از استراتژی ماکزیمم آرا استفاده می‌شود. اگر مقدار تابع مشخص‌کننده ماشین بردار پشتیبان جفت کلاس (k,m) و $f_{km}(x)$ مثبت باشد، یک رای به کلاس k در غیر این صورت یک رای به کلاس m داده می‌شود. به همین ترتیب از همه ماشین‌های بردار پشتیبان برای دادن رای به کلاس‌ها استفاده می‌شود. در نهایت نمونه آزمون متعلق به کلاسی است که بیشترین رای را داشته باشد.

۴ پیاده‌سازی و ارزیابی روش پیشنهادی

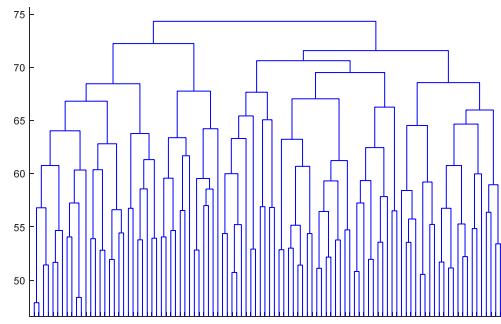
در این قسمت به تحلیل پیاده‌سازی و نتایج الگوریتم پیشنهادی در مقاله پرداخته شده است. این الگوریتم بر روی پایگاه داده ایران‌شهر مورد ارزیابی قرار گرفته است. این پایگاه داده شامل ۱۶۰۰۰ تصویر دست‌نوشته از ۵۰۳ نام شهرهای ایران می‌باشد. برای هر کلمه حداقل ۲۵ نمونه از دست‌نوشته افراد مختلف وجود دارد. برای هر کلمه از این پایگاه داده ۶۰٪ از نمونه‌ها برای مجموعه آموزش، ۲۰٪ برای مجموعه اعتبار سنجی و ۲۰٪ آن برای آزمون به صورت تصادفی انتخاب شده است. الگوریتم پیشنهادی در این تحقیق شامل دو مرحله اصلی می‌باشد. در مرحله اول شهرهای نامشابه با کلمه مورد آزمون، از مجموعه کلمات فرهنگ لغت حذف می‌شوند. برای این منظور در مرحله آموزش کلمه‌ها خوشه‌بندی شده تا کلمات متشابه و غیرمتشابه تفکیک شوند. مرحله دوم تشخیص کلمه آزمون ورودی از زیرمجموعه کلمات کاندید (بدست آمده از مرحله اول) است.

۴-۱ مرحله اول: کاهش فرهنگ لغات

همانطور که در قسمت‌های قبل بیان شد هر چه تعداد کلمات موجود در فرهنگ لغت بیشتر باشد به همان اندازه دقت و سرعت سیستم‌های تشخیص کلمه کاهش می‌یابد. بعضی از سیستم‌ها برای کاهش مشکلات ناشی از فرهنگ لغت بزرگ، در ابتدا تعداد کلمات مورد بررسی برای تشخیص کلمه آزمون ورودی را کاهش

است یک شهر (یک کلمه در فرهنگ لغت)، متعلق به بیش از یک خوشه باشد.

تعداد خوشه‌های بدست آمده توسط الگوریتم خوشه‌بندی سلسله مراتبی برابر با ۴۰ و تعداد خوشه‌های نهایی برابر با ۵۷ خوشه است. برای ارزیابی عملکرد الگوریتم کاهش‌دهنده فرهنگ لغت از دو معیار دقت و درجه کاهش استفاده می‌شود. که این دو معیار با هم ترکیب می‌شوند و معیار درجه تاثیر را تولید می‌کند. بنابراین معیار درجه تاثیر برای مقایسه عملکرد الگوریتم‌های کاهش‌دهنده فرهنگ لغت مناسب است.



شکل ۵ نمودار دندروگرام بر روی مجموعه داده آموزش

در ارزیابی عملکرد کاهش فرهنگ لغت پیشنهادی، فاصله کلمه آزمون ورودی با میانگین خوشه‌ها توسط الگوریتم انطباق زمانی پویا محاسبه می‌شود. تعدادی از نزدیک‌ترین خوشه‌ها به کلمه آزمون ورودی انتخاب می‌شود (n). دقت و درجه کاهش الگوریتم وابسته به این تعداد است. اگر n زیاد باشد، دقت افزایش و درجه کاهش کم می‌شود و اگر n کم باشد، دقت کاهش می‌یابد اما درجه کاهش فرهنگ لغت، زیاد می‌شود. برای تعیین مقدار n مناسب، استفاده از معیار درجه تاثیر مناسب‌تر است. جدول (۴) میزان دقت، درجه کاهش و درجه تاثیر را برای مقادیر مختلفی از n را نشان می‌دهد در این آزمایش‌ها ۲۵۰۰ تصویر برای آزمون در نظر گرفته شده است. با توجه به جدول، با انتخاب ۵ خوشه نزدیک به تصویر کلمه مورد آزمون ورودی، احتمال وجود کلمه آزمون در این خوشه ۹۶٫۲۷٪ و میزان کاهش کلمات کاندید مورد بررسی برای تشخیص تصویر ورودی ۷۹٫۶۳٪ و درجه تاثیر الگوریتم ۷۶٫۶۵٪ خواهد شد. یعنی در این مرحله به‌طور میانگین تعداد کلمات مورد بررسی برای تشخیص تصویر مورد آزمون ورودی از ۵۰۳ به ۹۷ کلمه کاهش می‌یابد. جدول (۵) عملکرد الگوریتم کاهش‌دهنده فرهنگ لغت پیشنهادی را با ویژگی پیشنهادی بدون LDA و با LDA را نشان می‌دهد.

جدول ۵ عملکرد الگوریتم کاهش فرهنگ لغت با ویژگی پیشنهادی

ویژگی	درجه کاهش	دقت کاهش	درجه تاثیر
ویژگی‌های پیشنهادی با LDA	۷۹٫۶۳	۹۶٫۲۷	۷۶٫۶۵
ویژگی‌های پیشنهادی بدون LDA	۷۹٫۸۳	۹۵٫۱۱	۷۵٫۹۲

مورد نیاز برای محاسبه فاصله انطباق زمانی پویا به صورت زیر تعریف می‌شود.

$$d(x_i, y_j) = \sum_{k=1}^4 (f_k(I, i) - f_k(J, j))^2 \quad (3)$$

۲-۴ معیار عملکردی الگوریتم کاهش فرهنگ لغت

اگر یک فرهنگ لغت شامل n کلمه $L = \{l_1, l_2, \dots, l_n\}$ باشد. یک سیستم کاهش‌دهنده فرهنگ لغت، یک زیرمجموعه‌ای $L' \subset L$ از کاندیدها برای تشخیص کلمه آزمون ورودی تولید می‌کند. کلماتی که در این زیرمجموعه نیستند ($L - L'$)، در تشخیص کلمه آزمون ورودی شرکت نمی‌کنند. این سیستم برخلاف یک طبقه‌بند، نیازی به رتبه‌بندی کلمات موجود در زیرمجموعه بدست آمده ندارد. برای اندازه‌گیری عملکرد سیستم کاهش‌دهنده فرهنگ لغت از معیارهای معرفی شده در مرجع [۲۸] استفاده شده است. اگر اتفاق درست قرار گرفتن کلمه‌ای (l_i) از فرهنگ لغت در زیرمجموعه L' را یک متغیر تصادفی A در نظر بگیریم داریم:

$$\begin{cases} A=1 & \text{if } l_i \in L' \\ A=0 & \text{if otherwise} \end{cases} \quad (4)$$

مقدار کاهش سیستم با متغیر R تعریف می‌شود:

$$R = \frac{(|L| - |L'|)}{|L|} \quad (5)$$

سه معیار اندازه‌گیری عملکرد سیستم کاهش‌دهنده فرهنگ لغت:

● دقت کاهش: $\alpha = E(A)$

● درجه کاهش: $\rho = E(R)$

● درجه تاثیر کاهش: $\eta = \alpha k \cdot \rho$

که $\alpha, \rho, \eta \in [0, 1]$ است. درجه و دقت کاهش معمولاً نسبت عکس نسبت به هم دارند. این دو معیار با هم ترکیب می‌شوند و یک معیار کلی η را تولید می‌کنند. ضریب k تاکید دقت کاهش نسبت به درجه کاهش را بیان می‌کند. در گزارش‌های علمی مقدار k حدود ۰٫۱، در نظر گرفته شده است [۶ و ۲۹]. در فرهنگ لغت کوچک، دقت اهمیت بیشتری دارد اما در فرهنگ لغات متوسط و بزرگ علاوه بر دقت، سرعت نیز مهم است.

۳-۴ نتایج مرحله کاهش فرهنگ لغت

خوشه‌بندی تصاویر کلمات دست‌نوشته توسط الگوریتم ISOCLUS و الگوریتم انطباق زمانی پویا انجام می‌شود. بدین ترتیب کلمات با شکل کلی مشابه در خوشه‌های یکسان قرار گرفته و یک فرهنگ لغت تصویری تولید می‌شود. میانگین خوشه‌ها به عنوان مدخل اعضای آن خوشه در فرهنگ لغت تصویری بدست آمده در نظر گرفته می‌شود. به دلیل تنوع و تغییرات زیاد در بین دست‌نوشته‌ها، نمی‌توان برای هر کلمه (کلاس) نماینده‌ای پیدا کرد که بتواند تمام ویژگی‌های نمونه‌های آن کلاس را بیان کند. بنابراین خوشه‌بندی کلمات دست‌نوشته برای افزایش دقت از تمام نمونه‌های آموزش هر کلمه استفاده شده است. در این حالت ممکن

جدول ۴ دقت درجه کاهش و درجه تاثیر برای مقادیر مختلف n

خوشه‌های نزدیک (n)	درجه کاهش	دقت کاهش	درجه تاثیر
۱	۹۲,۸۳	۷۸,۲۸	۷۲,۶۶
۲	۸۳,۵	۸۷,۲۷	۷۲,۸۷
۵	۷۹,۶۳	۹۶,۲۷	۷۶,۶۵
۷	۷۱,۱۲	۹۶,۳۸	۶۸,۵۱
۱۰	۶۳,۴۷	۹۷,۸۷	۶۲,۱۱
۱۲	۵۸,۳۴	۹۸,۱	۵۷,۶۸
۱۵	۵۲,۳۱	۹۸,۱۵	۵۱,۳۴
۲۰	۴۱,۲۲	۹۸,۹۲	۴۰,۷۷

جدول ۶ نتایج الگوریتم پیشنهادی برای بازشناسی کلمات

روش - کاهش فرهنگ لغت	روش دسته‌بندی	دقت تشخیص	سرعت تشخیص (ثانیه بر نمونه)
بازشناسی با کاهش فرهنگ لغت	SVM چند کلاسه	۹۱,۲۳	۱,۳۶
	SVM دو کلاسه	۹۳,۳۷	۰,۵۶
بازشناسی بدون کاهش فرهنگ لغت	SVM چند کلاسه	۷۸,۲۳	۱,۵۴
	SVM دو کلاسه	۸۴,۲۷	۰,۶۲
Gradient base [۹]	-	۸۹,۱۲	-
Decision Templates [۱۰]	-	۹۱,۵۵	-
Mixture of Experts [۱۱]	-	۹۱,۱۱	-

فارسی نسبت به ماشین بردار پشتیبان چندکلاسه از خود نشان می‌دهد. در این جدول نشان داده شده که روش پیشنهادی نسبت به روش‌های معرفی شده در سال‌های اخیر بر روی مجموعه داده ایران‌شهر عملکرد بهتری را از خود نشان داده است.

۵ نتیجه‌گیری

در این مقاله، روشی نوین بر پایه‌ی کاهش فرهنگ لغت و ویژگی‌های جدید به منظور بهبود عملکرد سیستم‌های بازشناسی کلمات دست‌نوشته فارسی ارائه شده است. از روش سیفت متراکم و مرکز ثقل اجزای متصل در هر بلاک به عنوان ویژگی استفاده شده است و در گام بعدی با استفاده از روش‌های خوشه‌بندی با نگرش نوین به کاهش فرهنگ لغت پرداخته شده است. در نهایت نتایج بدست آمده از روش پیشنهادی توسط ماشین بردار پشتیبان مورد ارزیابی قرار گرفته است. نتایج بدست آمده نشان دهنده بهبود عملکرد این نوع از سیستم‌ها از لحاظ زمان و دقت نسبت به کارهای دیگر می‌باشد. در کارهای آینده در راستای شناسایی کلمات در دست‌نوشته‌ها می‌توان بر روی ویژگی‌ها ترکیبی بر پایه ویژگی‌های ساختاری، آماری و محلی تمرکز نمود و همچنین از چند دسته‌بند جدید برای هر ویژگی استفاده نمود و در نهایت از ترکیب این طبقه‌بندها بهره‌گرفت.

۴-۴ مرحله دوم: تشخیص کلمه آزمون ورودی

در مرحله اول روش پیشنهادی با استفاده از خوشه‌بندی تصاویر کلمات دست‌نوشته با ویژگی‌های معرفی شده یک لیست از کلمات کاندید برای تشخیص کلمه آزمون ورودی تولید می‌شود. این مرحله به‌عنوان مرحله‌ای برای کاهش دامنه جستجو استفاده می‌شود. در مرحله دوم بهترین کاندید از لیست بدست آمده از مرحله اول برای شناسایی کلمه ورودی انتخاب می‌شود.

۴-۴-۱ نتایج مرحله بازشناسایی

برای دسته‌بندی نمونه‌های آزمون از روش دسته‌بندی ماشین بردار پشتیبان به‌صورت چندکلاسه و دوکلاسه استفاده می‌شود. آزمایش‌ها به‌گونه‌ای است که استفاده از روش خوشه‌بندی و بدون استفاده از خوشه‌بندی نمایش داده شده است. در این قسمت از مقاله، معیار دقت یعنی تعداد کلاس‌هایی که شناسایی شده، استفاده شده است. علاوه بر آن مدت زمان در نظر گرفته شده برای مجموعه آزمون نیز مورد ارزیابی قرار گرفته است. نتایج بدست آمده از آزمایشات بر روی مجموعه داده ایران‌شهر به‌صورت جدول (۶) است.

همان‌طور که از جدول (۶) مشخص است استفاده از روش کاهش فرهنگ لغت در بهبود سرعت و دقت شناسایی بسیار مؤثر بوده است. همچنین با توجه به نتایج بدست آمده روش ماشین بردار پشتیبان دو کلاسه نتایج بهتری را در دسته‌بندی کلمات دست‌نوشته

۶ مراجع

- [13] AlKhateeb, J. H., Jiang, J., Ren, J., Khelifi, F., and Ipson, S. S., "Multiclass Classification of Unconstrained Handwritten Arabic Words Using Machine Learning Approaches "The Open Signal Processing Journal, 2, pp.21-28, 2009.
- [14] Mohamad, R. A. H., Likforman-Sulem, L., and Mokbel, C., "Combining Slanted-Frame Classifiers for Improved HMM-Based Arabic Handwriting Recognition" IEEE transactions on pattern analysis and machine intelligence, Vol. 31, No. 7, pp.1165-1177, July 2009.
- [۱۵] ابراهیمی، ا.، کبیر، ا.، "استفاده از شکل کلی زیرکلمات چابی در بازیابی تصویر مستندات و بازشناسی متون فارسی" رساله دکترای مهندسی برق، گرایش الکترونیک، دانشگاه تربیت مدرس، تابستان ۸۴.
- [16] Mozaffari, S., Faez, K., Märgner, V., and El-Abed, H., "Lexicon reduction using dots for off-line Farsi/Arabic handwritten word recognition" Pattern Recognition Letters, Vol. 29, No. 7, pp. 724-734, 2008.
- [17] Sagheer, M. W., He, C. L., Nobile, N., and Suen, C. Y., "Holistic Urdu Handwritten Word Recognition Using Support Vector Machine" In Pattern Recognition (ICPR), 2010 20th International Conference on, pp. 1900-1903. IEEE, 2010..
- [18] Hennig, A., Sherkat, N., "Cursive script recognition using wildcards and multiple, " Pattern Analysis and Applications, Vol. 4, No. 1, pp. 51-60, 2001.
- [19] Haji. M. M., Farsi Handwritten Word Recognition Using Continous Hidden Markov Model and Structural Feature, Iran: MSC, Computer Engineering Shiraz University Shiraz (2005).
- [20] Al Aghbari, Z., Brook, S., "HAH manuscripts: Aholistic paradigm for lassifying and retrieving historical Arabic handwritten documents "Expert Systems with Applications 36, No. 8, pp. 10942-10951, 2009.
- [21] Cheriet, M., Kharma, N., Liu, C. L., and Suen, C., "character recognition system, A Guide for Students and Practioners", John Wiley and Sons, 2007.
- [22] Nasrollahi, K., Moeslund. T. B., "Face Quality Assessment System in Video Sequences," In European Workshop on Biometrics and Identity Management, pp. 10-18. Springer Berlin Heidelberg, 2008.
- [23] Lowe, D. G., "Distinctive Image Features from Scale-Invariant Keypoints", International Journal of Computer Vision, Vol. 60, No. 2, pp 91-110, 2004.
- [24] Wang, J. G., Li, J., Lee, C. Y., and Yau, W. Y., "Dense SIFT and Gabor descriptors-based face representation with applications to gender recognition," In Control Automation Robotics & Vision (ICARCV), 2010 11th International Conference on, pp. 1860-1864. IEEE, 2010.
- [25] Zahedi, M., Eslami, S., "Farsi/Arabic Optical Font Recognition Using SIFT Features," Procedia Computer Science, Vol. 3, pp.1055-1059, 2011.
- [1] Hull, J. J., "Word Shape Analysis in a Knowledge-based System for Reading Text", In The Second IEEE Conference on Artificial Intelligence Applications, pp. 114-119, 1985.
- [2] Powalka, R. K., Sherkat, N. and Whitrow, R. J., "The Use of Word Shape Information for Cursive Script Recognition", In Fourth International Workshop on Frontiers of Handwriting Recognition, pp. 67-76. 1994.
- [3] Amin, A. and Mansoon, W., "Recognition of Arabic Printed Text using Neural Networks", Document Analysis and Recognition, 1997., Proceedings of the Fourth International Conference on. Vol. 2. IEEE, 1997.
- [4] Toni, M. R., and Manmatha, R., "Word Image Matching Using Dynamic Time Warping" In Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on, vol. 2, pp. II-521. IEEE, 2003.
- [۵] شیرعلی شهرضا، م.، فائز، ک.، "تشخیص کلمات و ارقام دستنویس فارسی بوسیله شبکه‌های عصبی (خط‌نسخ)"، رساله دکترای مهندسی برق - کامپیوتر، دانشگاه صنعتی امیرکبیر، ۱۳۷۴.
- [6] Guillevic, D, Suen, C. Y., "HMM-KNN word recognition engine for bank cheque processing" In Pattern Recognition, 1998. Proceedings. Fourteenth International Conference on, vol. 2, pp. 1526-1529. IEEE, 1998.
- [7] Kim, J. H., Kim, K. K., and Suen. C. Y., "Hybrid schemes of homogeneous and heterogeneous classifiers for cursive word recognition". In Proc 7th International Workshop on Frontiers in Handwriting Recognition, Amsterdam. Netherlands, pp. 433-442. 2000.
- [8] Dehghan, M., Faez. K., Ahmadi M., and Shridhar M., "Handwritten Farsi (Arabic) word recognition: a holistic approach using discrete HMM", Pattern Recognition, Vol. 34, No. 5, pp. 1057-1065, 2001.
- [9] Ebrahimpour, R., Sarhangi, S., and Sharifzadeh, F., "Mixture of Experts for Persian handwritten word recognition." Iranian Journal of Electrical & Electronic Engineering, Vol. 7, No. 4, pp. 217-224, 2011.
- [10] Ebrahimpour, R., Amini, M. Shams, A. V., "A New Combination Method Based on Different Representation of Data," International Journal of Hybrid Information Technology, Vol. 4 No. 3, pp.51-60, July, 2011.
- [11] Ebrahimpour, R., Vahid, R. D., Nezhad, B. M., "Decision Templates with Gradient based Features for Farsi Handwritten Word Recognition," International Journal of Hybrid Information Technology, Vol. 4, No. 1, pp.1-12, 2011.
- [۱۲] خسروی، ح.، کبیر، ا.، "معرفی دو ویژگی سریع و کارآمد برای تشخیص ارقام دست‌نویس فارسی" چهارمین کنفرانس بینایی‌ماشین و پردازش‌تصویر ایران، ۲۵-۲۶ بهمن ۱۳۸۵.

- [26] Duda, R. O., Hart, P. E., Stork, D. G., "pattern classification," Wiley-Interscience Publication, newyork, 2001.
- [27] Memarsadeghi, N., Mount, D. M., Netanyahu, N. S., and Le Moigne, J., "A Fast Impemention of the ISOCLUS Alghorithm" In international geoscience and remote sensing symposium, Vol. 3, pp. III-2057., 2003.
- [28] Mozaffari, S., Faez, K., Märgner, V., and El-Abed, H., "Lexicon reduction using dots for off-line Farsi/Arabic handwritten word recognition" Pattern Recognition Letters, Vol. 29, No. 6, pp. 724-734, 2008.
- [29] Koerich. A. L., Sabourin. R., Suen. C. Y., "Large vocabulary off-line handwriting recognition: A survey", Pattern Analysis & Applications, Vol. 6, No. 2, pp. 97-121, 2003.



سمیه برومند در حال حاضر دانشجوی کارشناسی ارشد علوم تصمیم و مهندسی دانش در دانشگاه آزاد اسلامی واحد مبارکه است. ایشان مدرک کارشناسی را در رشته مهندسی کامپیوتر گرایش نرم افزار سال ۱۳۸۳ از این دانشگاه اخذ نمودند و بعد از آن در شرکت تامین ماشین آلات ساختمانی و نصب - تامکو در سمت های برنامه نویس و تجزیه و تحلیل گر سیستم های اطلاعاتی، مدیر IT، مدیر تضمین کیفیت همکاری کرده است. زمینه تحقیقاتی مورد علاقه ایشان شناسایی الگو و پردازش تصویر می باشد.



مجید ایرانپور مبارکه در حال حاضر استادیار دانشکده مهندسی کامپیوتر و فناوری اطلاعات دانشگاه پیام نور می باشند. ایشان مدرک دکتری و کارشناسی ارشد خود را در رشته مهندسی کامپیوتر-هوش مصنوعی به ترتیب از دانشگاه صنعتی شاهرود در سال ۱۳۹۵ و دانشگاه علم و صنعت ایران در سال ۱۳۸۷ اخذ نموده اند. زمینه های تحقیقاتی مورد علاقه ایشان بینایی ماشین، پردازش اسناد تصویری، شناسایی الگو و داده کاوی است.