

روشی دومرحله‌ای برای بازشناسی کلمات دستنوشته فارسی به کمک بلوک‌بندی تطبیقی گرادیان تصویر

الهام بایسته تاشک، علیرضا احمدی فرد و حسین خسروی

دانشگاه شهرورد، شهرورد، ایران

چکیده

در این مقاله یک روش دومرحله‌ای برای بازشناسی کلمات دستنوشته فارسی ارائه شده است. در مرحله اول برای افزایش دقت و کاهش بار محاسباتی، الگوریتمی برای کاهش کلمات فرهنگ لغت قابل مقایسه با کلمه مورد آزمون ارائه شده است. برای این منظور کلمات موجود در فرهنگ لغت توسط الگوریتم‌های خوشبندی، دسته‌بندی می‌شوند. خوشبندی در این مرحله بر اساس ویژگی‌هایی که شکل کلی کلمه را توصیف می‌کنند، می‌باشد. در مرحله دوم یک روش جدید جهت استخراج ویژگی هیستوگرام گرادیان تصویر کلمه پیشنهاد شده است که این روش پیشنهادی تناظر بین نمونه‌های مختلف تصاویر یک کلمه دسته‌بندی را بهتر نشان می‌دهد. با مقایسه بردار ویژگی استخراج شده از کلمه ورودی و بردار ویژگی کلمات نامزد (به دست آمده از مرحله اول) در یک طبقه‌بند K‌نزدیک‌ترین همسایه بهترین نامزد برای کلمه ورودی شناسایی می‌شود.

نتایج پیاده‌سازی روش پیشنهادی بر روی پایگاه داده "ایران شهر" نشان می‌دهد که مرحله کاهش فرهنگ لغت و روش جدید جهت استخراج ویژگی هیستوگرام گرادیان، دقت و سرعت سامانه را به طور قابل ملاحظه‌ای هم از لحاظ دقت و تا حدودی از لحاظ سرعت بهبود می‌بخشد.

واژگان کلیدی: بازشناسی کلمه دستنوشته، الگوریتم خوشبندی ISOCLUS، DTW، ویژگی نمایه، ویژگی هیستوگرام گرادیان.

۱- مقدمه

یکی از مسائل پرکاربرد در حوزه شناسایی الگو تشخیص خودکار متون دستنوشته است. از کاربردهای مهم این مسئله می‌توان به پردازش اسناد، خواندن فرم‌های ثبت‌نام، پرسشنامه‌ها، اسناد و اوراق بهادر، خواندن و مرتب‌کردن خودکار مرسولات پستی، چک‌های بانکی و بازیابی و مرتب‌کردن تصاویر متون دستنوشته تاریخی پویش شده اشاره کرد (چن، ۱۹۹۵) و (گولوج، ۱۹۹۸) و (لوریگو، ۲۰۰۶) به طور کلی بازشناسی کلمات دستنوشته را می‌توان به دو گروه تقسیم کرد: بازشناسی کلمه دستنوشته به دو حوزه متفاوت هستند. در سامانه‌های گروه اول تصویر پویش شده‌ای از تصویر کلمه دستنوشته در دست است. در این خصوص اطلاعات مربوط به حرکت قلم موجود

نیست؛ اما در گروه دوم علاوه‌بر اطلاعات مکانی، اطلاعات زمانی و ایستای حرکتی فرآیند نوشتمن که از طریق قلم به سامانه منتقل می‌شود، نیز وجود دارد. این اطلاعات اضافی باعث شده تحت شرایط آزمایشی مشابه نرخ شناسایی در سامانه‌های برخط بالاتر از سامانه‌های برون خط باشد (پلاماندن، ۲۰۰۰). همچنین سامانه‌های برخط می‌توانند در فرهنگ لغت بزرگ‌تری کلمات را شناسایی کنند.

سامانه‌های بازشناسی کلمه دستنوشته برون خط موجود، به خصوص برای زبان‌های لاتین و چینی قادرند با دقیقی در حدود ۹۹٪-۹۰٪، البته با یک سری محدودیت‌های آزمایشی، دستنوشته‌ها را شناسایی کنند. یکی از رایج‌ترین محدودیت‌های سامانه‌های موجود این است که بیشتر این سامانه‌ها دارای یک فرهنگ لغت

از روی شکل کلی آن شناسایی می‌کنیم؛ لذا در دسته دوم سامانه‌های بازشناسی دستنوشته برونو خط جای می‌گیرد. در ادامه برخی کارهای گزارش شده در زمینه بازشناسی متون فارسی و عربی مبتنی بر شکل کلی کلمه را مرور می‌کنیم.

در پژوهش (برومندنیا، ۲۰۰۸) یک ویژگی آماری برای بازشناسی متون دستنوشته معرفی شده است که نسبت به چرخش و اندازه، تغییرناپذیر و مبتنی بر شکل کلی کلمات است. این پژوهش از m باند تبدیل موجک گسسته تصویر قطبی کلمه برای استخراج بردار ویژگی برای بازشناسی یکصد کلمه فارسی استفاده کرده است. برای تغییرناپذیربودن ویژگی‌ها نسبت به مقیاس، تصویر کلمه در دستگاه قطبی هنجارسازی می‌شود. از فاصله ماهالانوبیس برای اندازه‌گیری شباهت بین بردار ویژگی استخراج شده از کلمه آزمون ورودی و کلمات فرهنگ لغت استفاده شده است. دقت بازشناسی بروی یکصد کلمه برای این ویژگی ۹۲٪ گزارش شده که نسبت به ویژگی‌های فوریه و گشتاورهای زرنیکی متغیر با چرخش ۱۲٪ افزایش پیدا کرده است.

در مرجع (دهقان، ۲۰۰۱) از مدل مخفی مارکف و اطلاعات محلی کانتور برای بازشناسی اسمای دستنویس شهرهای ایران استفاده شده است. تصویر کلمات در راستای افقی به پنج قسمت مساوی و در راستای عمودی به قسمت‌هایی که ۵۰٪ با هم همپوشانی دارند تقسیم می‌شود. برداری شامل هیستوگرام‌های جهتی کدهای زنجیره‌ای در هر یک از این پنجره‌ها به عنوان مدل کلمه انتخاب می‌شود. برای کم کردن تعداد ویژگی‌ها، فضای ویژگی با استفاده از یک شبکه عصبی خودسامان کوهن، کمی شده است. برای هر اسم شهر یک HMM گسسته با الگوریتم بام-ولش به طور مجزا آموزش داده می‌شود. برای هر تصویر کلمه ورودی مورد آزمون، اسمای شهرها بر حسب میزان شباهت مدل آنها به کلمه ورودی مرتباً شوند. نتایج به دست آمده بر روی یک مجموعه کلمات دستنویس شامل اسمای ۱۹۸ شهر ایران، در انتخاب اول ۶۵٪ و در بیست انتخاب اول ۹۵٪ گزارش شده است.

در پژوهش (مصطفی، ۲۰۰۸) از کد توصیف‌گر تعداد و موقعیت نقاط در کلمه برای کاهش تعداد کلماتی که در بازشناسی کلمه ورودی مورد آزمون به کار می‌رود، استفاده شده است. مرحله کاهش فرهنگ لغت برای سامانه‌های با اندازه فرهنگ لغت بزرگ برای کاهش کلمات نامتشابه با

محدود هستند و تنها قادرند کلماتی را شناسایی کنند که در این فرهنگ لغت وجود داشته باشند. این فرهنگ لغت^۱ می‌تواند شامل سی تا هزار کلمه باشد (کورریج، ۲۰۰۳).

هرچه تعداد کلمات موجود در فرهنگ لغت افزایش یابد، احتمال وجود کلمات مشابه در فرهنگ لغت بالا

می‌رود و درنتیجه کار بازشناسی مشکل‌تر می‌شود. پیچیدگی محاسباتی سامانه بازشناسی نیز به اندازه فرهنگ لغت وابسته است و هر چه اندازه فرهنگ لغت افزایش یابد، با همان نسبت پیچیدگی محاسباتی بالا می‌رود. در بعضی از سامانه‌های شناسایی کلمه که در ارتباط با یک فرهنگ لغت بزرگ است برای کاهش پیچیدگی محاسباتی و افزایش دقت و سرعت، تعداد کلمات مورد بررسی در فرآیند بازشناسی کلمه ورودی مورد آزمون را کاهش می‌دهند که به این روش کاهش فرهنگ لغت^۲ یا هرس کردن^۳ گفته می‌شود (مصطفی، ۲۰۰۸) و (وشاه، ۲۰۱۰).

سامانه‌های بازشناسی دستنوشته برونو خط را می‌توان به دو دسته تقسیم کرد (چریت، ۲۰۰۷) و (مدھونت، ۲۰۰۱): ۱- سامانه‌های مبتنی بر قطعه‌بندی ۲- سامانه‌های مستقل از قطعه‌بندی.

روش‌های مبتنی بر قطعه‌بندی که به طور معمول روش‌های تجزیه‌ای گفته می‌شوند، روش‌هایی هستند که کلمه را به اجزای تشکیل‌دهنده‌اش تجزیه می‌کنند و هر جزء را به طور مجزا، با استفاده از ویژگی‌های استخراج شده از آن بازشناسی می‌کنند. در روش‌های بازشناسی مستقل از قطعه‌بندی که به طور معمول روش‌های کلی نامیده می‌شوند ویژگی‌ها به طور مستقیم از تصویر کلمه دستنوشته بدون هیچ گونه قطعه‌بندی استخراج می‌شوند و با استفاده از یک فرهنگ لغت محدود، کلمه ورودی شناسایی می‌شود. روش‌های دسته دوم به دلیل اینکه از مشکلات ناشی از قطعه‌بندی رها شده‌اند، ساده‌تر و دقیق‌تر هستند (چریت، ۲۰۰۷).

خصوصیات خاصی نظیر پیوستگی برخی حروف در کلمات، تغییر شکل حروف در موقعیت‌های مختلف کلمه، وجود نقاط، علائم و سرکش در حروفی مانند "ك" ، "گ" و "آ" در زبان‌های فارسی و عربی باعث شده که بازشناسی متون فارسی و عربی مشکل‌تر و پیچیده‌تر از متون لاتین باشد. در روش پیشنهادشده در این مقاله، کلمات فارسی را

¹ lexicon

² Lexicon reduction
³ pruning

می‌کند تا تصاویر یک کلمه با دستخط‌های مختلف را بتوان به درستی منطبق کرد.
برای ارزیابی الگوریتم پیشنهادی از پایگاه داده "ایران شهر"^۱ شامل ۱۶۰۰۰ تصویر کلمه دست‌نوشتۀ از ۵۰۳ نام شهرهای ایران استفاده شده است. در ادامه، در بخش دو پیش‌پردازش‌های انجام‌شده برای کاهش اختلافات در تصاویر کلمات دست‌نوشتۀ جهت کاهش فضای خوشبندی کلمات دست‌نوشتۀ جهت کاهش فضای جستجو در بخش سه آمده است. بخش چهار به بازشناسی کلمه دست‌نوشتۀ ورودی می‌پردازد. بخش پنجم نتایج ارزیابی الگوریتم پیشنهادی بر روی پایگاه داده "ایران شهر" را بیان می‌کند و بخش شش نتیجه‌گیری ارائه شده است.

۲- پیش‌پردازش

در مرحله پیش‌پردازش، یک دنباله‌ای از روش‌ها برای بهبود کیفیت و کاهش اختلافات و پیچیدگی‌های موجود در تصاویر، بر روی تصاویر کلمات دست‌نوشتۀ انجام می‌شود. این روش‌ها می‌توانند بر روی دقت نهایی سامانه‌های بازشناسی متون دست‌نوشتۀ تأثیر بسزایی داشته باشند. روش‌های پیش‌پردازش شامل:

- دودویی کردن: ممکن است متون دست‌نوشتۀ توسط قلم‌های رنگی و یا بر روی کاغذهای رنگی نوشته شده باشند؛ اما پردازش متون دست‌نوشتۀ اغلب بر روی تصاویر دوستخی انجام می‌شود. بنابراین اولین عمل در مرحله پیش‌پردازش دودویی کردن تصاویر رنگی پویش شده می‌باشد. در این مقاله برای دودویی کردن تصاویر کلمات دست‌نوشتۀ از الگوریتم آستانه‌گیری otsu استفاده شده است (اوتسو، ۱۹۷۹).

- هموارسازی و حذف نوحفه: مرحله بعدی در مرحله پیش‌پردازش، هموارسازی تصاویر دوستخی به دست آمده از مرحله دودویی برای کاهش نوحفه و اصلاح لبه‌های در این مرحله برای حذف نوحفه‌ای نمک و فلفل، پرکردن شکاف‌ها، دره‌های کوچک و حذف دست‌اندازهای به وجود آمده در لبه‌ها از فیلتر میانگین و عملیات ریخت‌شناسی بستن و باز کردن^۲ استفاده شده است.

^۱ این پایگاه داده از دانشگاه تربیت مدرس دانشکده فنی و مهندسی گروه الکترونیک گرفته شده است

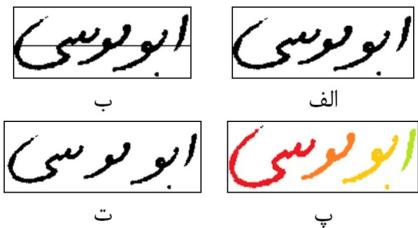
^۲ Closing and Opening Morphology

کلمه ورودی در مرحله بازشناسی و افزایش سرعت و دقت سامانه، مناسب است. برای هر کلمه موجود در فرهنگ لغت یک مدل HMM گستته 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 بازشناسی کلمه آزمون ورودی، تنها مدل HMM کلماتی شرکت می‌کنند که در مرحله کاهش فرهنگ لغت انتخاب شده‌اند. این الگوریتم بر روی یک پایگاه داده شامل ۱۷۰۰۰ تصویر از اسامی دویست شهر ایران آزمایش شده است. نرخ بازشناسی در انتخاب اول بدون کاهش فرهنگ لغت ۶۰٪ و با کاهش فرهنگ لغت ۷۳٪ گزارش شده است.

در این مقاله یک روش دو مرحله‌ای برای بازشناسی کلمه دست‌نوشتۀ فارسی ارائه شده است. در مرحله اول برای تسریع بازشناسی و کاهش محاسبات، روشی برای کاهش کلمات مورد بررسی در بازشناسی کلمه مورد آزمون ارائه شده است. برای این منظور ابتدا کلمات در فرهنگ لغت خوشبندی می‌شوند. برای هر خوشة تشکیل شده نماینده‌ای به عنوان مدخل اضافی آن خوشه در نظر گرفته می‌شود. در این مرحله از ویژگی‌هایی که شکل کلی کلمه را توصیف می‌کنند، استفاده شده است. در این رابطه از ویژگی‌های مبتنی بر نمایه تصویر کلمه بهره می‌بریم. در مرحله شناسایی، تعدادی از خوشه‌ها که توصیف‌گرها مشابه با کلمه دست‌نوشتۀ ورودی دارند، انتخاب می‌شوند. کلمات موجود در این خوشه‌ها در فهرست نامزدها برای شناسایی کلمه ورودی قرار می‌گیرند. از کلمه ورودی و نامزدهای به دست آمده ویژگی‌هایی که جزئیات شکل کلمه را بهتر بیان می‌کنند، استخراج می‌شود. با مقایسه بین ویژگی‌های استخراج شده از نامزدهای به دست آمده در مرحله خوشبندی و کلمه مورد آزمون به کمک یک طبقه‌بند K نزدیک‌ترین همسایه، هویت کلمه ورودی شناسایی می‌شود.

استفاده از ویژگی‌های مبتنی بر نمایه تصویر کلمه جهت کاهش فرهنگ لغت کار جدیدی است که در این مقاله پیشنهاد شده است. از طرفی در این مقاله در مرحله شناسایی کلمه، از ویژگی‌های گردایان در بلوک‌های تطبیقی استفاده شده است. پیشنهاد تقسیم‌کردن تصویر کلمه به بلوک‌هایی که با شکل کلمه تطبیق داده می‌شوند جدید است. همان‌طوری که نتایج نشان می‌دهد، استفاده از ویژگی‌های استخراج شده از بلوک‌های تطبیق‌یافته کمک

میانگین تعداد پیکسل های مؤلفه های متصل آن کلمه است، به عنوان مؤلفه های متصل اصلی کلمه در نظر گرفته می شوند. مؤلفه های متصل غیر اصلی (نقاط، علامت و مکمل ها) متعلق به مؤلفه اصلی است که مرکز ثقلش کمترین فاصله را با مرکز ثقل آن مؤلفه داشته باشد. مرحله بعدی گذاشتن فاصله بین مؤلفه های متصل اصلی است. شکل (۱) مراحل اجرای الگوریتم کشیدن را نشان می دهد.



(شکل - ۱): مراحل اجرای الگوریتم کشیدن الف- تصویر نمونه ب- مرحله اول پ- مرحله دوم ت- مرحله سوم

- نازک سازی: یکی دیگر از مراحل پیش پردازش، مرحله نازک سازی یا اسکلت بنده است. نازک سازی یعنی تبدیل شکل به منحنی ها و خطوطی با ضخامت یک پیکسل است؛ به گونه ای که ویژگی های ساختاری شکل حفظ شود. نازک سازی با یکسان کردن ضخامت پیکسل های کلمه در تصاویر، در دقت سامانه های بازناسی کلمه تأثیر دارد. برای نازک سازی تصاویر کلمات دستنوشته از الگوریتم Hilditch استفاده شده است (هیلدیچ، ۱۹۸۳).

- نممال سازی بعد از تصاویر: ننممال سازی بعد از تصاویر یکی از مهم ترین مراحل پیش پردازش است. در این مرحله ابعاد همه تصاویر برای داشتن ابعاد ویژگی ثابت، یکسان می شود. بر اساس فن ننممال سازی تطبیقی نسبت ابعاد (چریت، ۲۰۰۷)، نزممال سازی ابعاد تصاویر به دو دسته تقسیم می شود: نزممال سازی خطی و نزممال سازی غیر خطی. با توجه به نتایج پیاده سازی روش های مختلف نزممال سازی، روش گشتاور با تبدیل نسبت ابعاد سینوسی بهترین جواب را داده است.

۳- خوشبندی کلمات دستنوشته

در سامانه های بازناسی مبنی بر فرهنگ لغت، هر چه تعداد کلمات موجود در فرهنگ لغت بیشتر باشد، به همان اندازه دقت و سرعت سامانه های بازناسی کلمه کاهش می یابد.

- اصلاح کجی: کجی، انحراف میانگین پاره خط های نزدیک به خط عمودی نسبت به خط عمود است. کجی در متون دستنوشته به سبب نحوه نوشتن خاص یک فرد و در متون تایبی به دلیل خصوصیات ذاتی برخی قلم ها به وجود می آید. برای اصلاح کجی تصویر، ابتدا زاویه کجی تخمین زده می شود؛ سپس با تبدیل shear افقی، تصویر نسبت به خط عمودی اصلاح می شود. برای تخمین زاویه کجی از روش نمایه بالایی تصویر کلمه (کاوالیرتو، ۲۰۰۰) استفاده شد.

- تحلیل زیر کلمات: در متون دستنوشته تنوع در شیوه نوشتن بسیار زیاد است و این باعث پیچیده شدن فرآیند بازناسی کلمات می شود. در این بخش با انجام یک سری عملیات بر روی تصاویر تاحدودی تغییرات و تنوع را کاهش می دهیم. در بعضی از دستنوشته ها فاصله بین زیر کلمات بسیار زیاد است. در این مرحله برای یکسان کردن فاصله بین زیر کلمات، در صورت وجود فضای خالی در بین زیر کلمات، فضای خالی حذف می شود.

در متون دستنوشته، بعضی از حروف یا زیر کلمات در زیر حروف یا زیر کلمات دیگر قرار می گیرند بدون اینکه به هم متصل شوند و این ممکن است در استخراج بعضی ویژگی ها خطای ایجاد کند، برای حل این مشکل از الگوریتم کشیدن^۱ (اگبری، ۲۰۰۹) استفاده شده است.

- الگوریتم کشیدن: هدف از اجرای این الگوریتم حذف هم پوشانی های ایجاد شده در بعضی از زیر کلمات تصویر دستنوشته است. اولین مرحله این الگوریتم پیدا کردن خط کرسی کلمه دستنوشته است. خط کرسی یک کلمه، خطی فرضی است که بر روی آن حروف تشکیل دهنده کلمه به یکدیگر متصل می شوند. یکی از رایج ترین روش ها برای پیدا کردن خط کرسی یک کلمه بر اساس نمایه افکشن افقی است. به این صورت که خطی که بیشترین مقدار را در نمایه افکشن افقی داشته باشد، به عنوان خط کرسی شناخته می شود. بیشترین مقدار در نمایه افکشن افقی، ردیفی با بیشترین تعداد پیکسل را نشان می دهد. مرحله بعدی در این الگوریتم تعیین مؤلفه های متصل اصلی کلمه که همان زیر کلماتند، است. برای این کار بر روی خط کرسی حرکت کرده و مؤلفه های متصلی که با خط کرسی یا خطوط نزدیک به خط کرسی برخورد و تعداد پیکسل هایشان بزرگ تر از

¹ stretching

دو ویژگی بیان شده ساختار بیرونی شکل کلمه را توصیف می‌کند. برای بیان ساختار داخلی کلمه از ویژگی پروفایل انتقال سیاه به سفید استفاده شده است. این ویژگی در هر ستون تعداد انتقال‌ها از سیاه به سفید را محاسبه می‌کند.

بنابراین از هر ستون تصویر کلمه دست‌نوشته چهار ویژگی استخراج می‌شود. تعداد کل ویژگی‌ها برابر با $3 \times m$ است که m تعداد ستون‌های تصویر کلمه است. در مرحله پیش‌پردازش، در مرحله نرمال‌کردن ابعاد تصاویر، ابعاد همه تصاویر برابر با 200×200 شد. بنابراین در این مرحله تعداد کل ویژگی‌ها $= 600 = 3 \times 200$ می‌شود. با توجه به بعد بالای بُردار ویژگی، برای کاهش تعداد ویژگی‌ها از تبدیل موجک یک‌بعدی بر روی هر یک از ویژگی‌ها استفاده شده است. به کاربردن تبدیل موجک بر روی بُردارهای ویژگی نه تنها ابعاد ویژگی‌ها را کاهش می‌دهد، بلکه لرزش‌های موجود در لبه‌های تصویر را که در اثر نوشتمندی پویش کردن یا دودویی کردن به وجود آمده‌اند نیز کاهش می‌دهد.

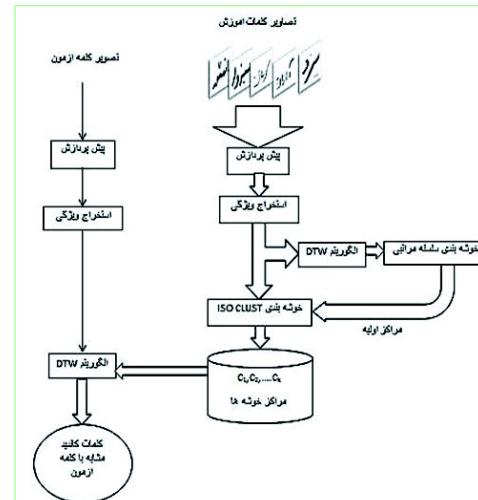
شکل (۳) ویژگی نمایه بالا را به کاربردن تبدیل موجک بر روی آن نشان می‌دهد. همان‌طور که شکل نشان می‌دهد تبدیل موجک بدون تغییر در شکل کلی نمایه ابعاد ویژگی را کاهش می‌دهد.

بعد از استخراج ویژگی‌ها از تمام تصویر، مرحله بعدی، مرحله گروه‌بندی تصاویر کلمه دست‌نوشته بر مبنای شکل ظاهری‌شان که با ویژگی‌های بیان شده در بالا توصیف می‌شوند، است. در این مرحله از الگوریتم خوش‌بندی سلسله‌مراتبی ISOCLUS با کمی استفاده می‌شود.

الگوریتم ISOCLUS یکی از الگوریتم‌های معروف خوش‌بندی است که در بسیاری از کاربردها به طور گسترده از آن استفاده می‌شود. این الگوریتم بر مبنای الگوریتم ISODATA با کمی اصلاحات است (معمار صادقی، ۲۰۰۳). این الگوریتم‌ها به طور کلی شبیه به الگوریتم k-میانگین است با این تفاوت که در الگوریتم k-میانگین تعداد خوش‌های ثابت است؛ اما در الگوریتم ISOCLUS با حذف یا تقسیم یا ادغام خوش‌های در هر مرحله، ممکن است تعداد خوش‌های نهایی با تعداد خوش‌های اولیه متفاوت باشد.

بعضی از سامانه‌ها برای کاهش مشکلات ناشی از فرهنگ لغت بزرگ، تعداد کلمات مورد بررسی برای بازشناسی کلمه آزمون ورودی را کاهش می‌دهند. این کار نه تنها سرعت سامانه را بالا می‌برد، بلکه ممکن است دقت را نیز افزایش دهد. در این تحقیق برای این منظور از خوش‌بندی تصاویر کلمات دست‌نوشته برای کاهش دامنه جستجوی بازشناسی کلمه آزمون ورودی استفاده شده است. در این مرحله ویژگی‌هایی استفاده شده است که بتواند شکل کلی کلمه را به خوبی توصیف کنند همچنین سریع استخراج شوند؛ سپس از الگوریتم‌های سلسله‌مراتبی و ISOCLUS (معمار صادقی، ۲۰۰۳) برای خوش‌بندی تصاویر کلمات دست‌نوشته استفاده می‌شود.

شکل (۲) مراحل خوش‌بندی تصاویر کلمات دست‌نوشته را نشان می‌دهد. همان‌طور که در شکل آمده است، این مراحل شامل پیش‌پردازش، استخراج ویژگی و خوش‌بندی سلسله‌مراتبی و ISOCLUS است.



(شکل -۲): مراحل خوش‌بندی تصاویر کلمات دست‌نوشته

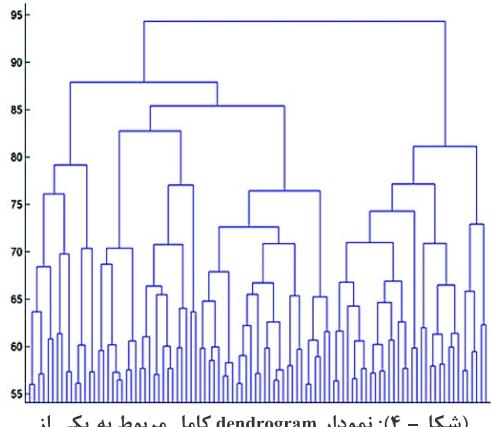
در این مرحله، از ویژگی‌های نمایه بالایی، پایینی و نمایه انتقال سیاه به سفید استفاده شده است. این ویژگی‌ها می‌توانند شکل کلمه را به طور کلی توصیف کنند (اغربی، ۲۰۰۹) و (رات، ۲۰۰۳).

ویژگی‌های نمایه بالا (و پایین)، ساختار بیرونی شکل کلمه را از بالا (و پایین) بیان می‌کنند. این ویژگی‌ها با اندازه‌گیری فاصله اولین پیکسل سیاه تا مرز بالایی (و پایینی) تصویر در هر ستون به دست می‌آیند. تعداد ویژگی‌ها برابر با تعداد ستون‌های هر تصویر است.

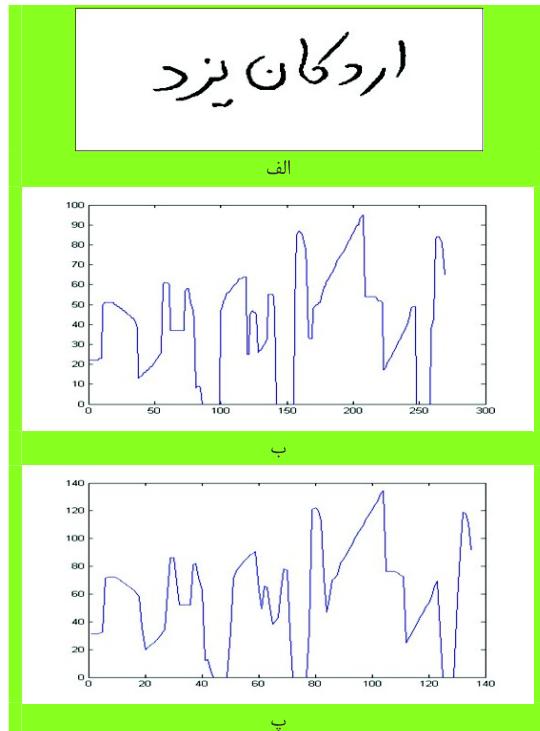
می توان تعداد خوشه های مناسب را تخمین زد و هم مراکز اولیه برای الگوریتم ISOCLUS آماده می شود.

در الگوریتم سلسله مراتبی پس از استخراج ویرگی ها از تصاویر آموزش، در مرحله اول فاصله بین تمام نمونه ها بر اساس معیار فاصله مورد نظر، محاسبه و در مرحله دوم از روی بردار فاصله ها، نمونه های نزدیک به هم پیدا شده و با یکدیگر ترکیب می شوند تا تشکیل یک خوشه جدید بدene. در دوره بعدی بردارهای میانگین خوشه های جدید به عنوان ورودی در نظر گرفته شده و مرحله اول و دوم تکرار می شود. باز هم خوشه های نزدیک به هم با یکدیگر ترکیب می شوند؛ این روال ادامه می یابد تا تمام نمونه ها در یک خوشه قرار بگیرند. در حین اجرای این فرایند نمودار تشکیل می شود که شاخه های آن بیان گر خوشه هایی است که با یکدیگر ترکیب می شود و ارتفاع شاخه ها بیان گر میزان فاصله بین خوشه های ترکیب شده است (خسروی، ۱۳۸۸).

اگر خوشه یابی سلسله مراتبی تا جایی که فقط یک خوشه باقی بماند، ادامه می یابد، نمودار درختی شبیه نمودار شکل (۴) حاصل می شود. در این نمودار خوشه های بسیار کوچک اولیه نشان داده نشده اند.



با نگاه به این نمودار و دقت در ارتفاع شاخه های حاصله می توان تشخیص داد که در کدام سطح فاصله اگر الگوریتم متوقف شود، خوشه های مناسبی به دست خواهد آمد. در نمودار بالایی اگر الگوریتم بعد از سطح هفتاد روی محور فاصله متوقف شود، خوشه هایی به وجود خواهد آمد که فاصله بین نمونه های آنها به طور شهودی معقول بوده و اندازه خوشه ها نیز مناسب است.



(شکل-۳): الف . یک نمونه از کلمه دستنوشته ب نمایه بالا پ. نمایه بالا پس از اعمال تبدیل موجک

هدف الگوریتم ISOCLUS پیدا کردن بهترین مراکز خوشه ها با یک روش تکراری است. ورودی های این الگوریتم شامل: مجموعه داده آموزش، تعداد اولیه خوشه ها و یک سری مشخصه های از پیش تعیین شده برای حذف یا تقسیم یا ادغام خوشه هاست. این مشخصه ها به صورت سعی و خطای توسعه کاربر تعیین می شود و شامل :

- کمینه تعداد نمونه ها در هر خوشه (SampRn)
- بیشینه تعداد تکرار (MaxIter)
- بیشینه انحراف استاندارد برای هر خوشه (StdDev)
- بیشینه تعداد جفت خوشه هایی که در هر تکرار می توانند ادغام شوند (MaxPair)
- بیشینه تعداد خوشه ها (MaxNum)
- کمینه فاصله بین مرکز خوشه ها (MinDis)
- مقدار و تعداد اولیه خوشه ها در نتیجه پایانی، تأثیر به سزایی دارد و انتخاب اشتباہ باعث ایجاد نتایج اشتباہ و غیر معتبر خواهد شد. به همین دلیل از الگوریتم سلسله مراتبی که مستقل از مشخصه است (خسروی، ۱۳۸۸)، برای تعیین تعداد اولیه خوشه ها استفاده شده است. با این کار هم



که $\alpha \in [0, 1]$ است. درجه و دقت کاهش نسبت به هم رابطه عکس دارند. این دو معیار باهم ترکیب می‌شوند و یک معیار کلی η را تولید می‌کنند. ضرب k تأکید دقت کاهش نسبت به درجه کاهش را بیان می‌کند.

دقت در فرهنگ لغت‌های کوچک، اهمیت بیشتری دارد؛ اما در فرهنگ لغت‌های متوسط و بزرگ علاوه‌بر دقت، سرعت نیز مهم است.

۴- بازشناسی کلمه دست نوشته و رویدی

در مرحله اول الگوریتم پیشنهادی با استفاده از گروه‌بندی تصاویر کلمات دست نوشته توسط ویژگی‌هایی نمایه، یک فهرستی از کلمات نامزد برای بازشناسی کلمه آزمون ورودی تولید می‌شود. این مرحله به عنوان مرحله‌ای برای کاهش دامنه جستجو است. در مرحله دوم با استفاده از ویژگی‌های جزئی و دقیق‌تر بهترین نامزد از فهرست به دست آمده از مرحله اول برای شناسایی کلمه ورودی انتخاب می‌شود.

در مرحله دوم شناسایی، از هیستوگرام گردایان استخراج شده از بلوک‌های تصویر کلمه استفاده می‌شود تا نزدیک‌ترین کلمه به کلمه ورودی از فهرست نامزد‌های مرحله قبل انتخاب شود. این کلمه هویت کلمه ورودی را تعیین می‌کند. در این بخش یک روش جدید جهت بلوك‌بندی تصویر ارائه می‌دهیم تا بتوان دست نوشته‌های مختلف یک کلمه را بهتر مقایسه کرد. برای انتخاب بهترین نامزد از طبقه‌بند k نزدیک‌ترین همسایه استفاده کرده‌ایم. در ادامه روش پیشنهادی با جزئیات بیشتر توضیح داده خواهد شد.

۴-۱- ویژگی هیستوگرام گردایان روشنایی

محلی

از هیستوگرام محلی گردایان روشنایی نتایج خوبی در بازشناسی ارقام و حروف گزارش شده است (خسروی، ۱۳۸۵). ویژگی گردایان با استفاده از جهت‌های محلی، به خوبی کلمات داخل یک طبقه را توصیف کرده و تغییرات بین طبقه‌ای را مدل می‌کند. این ویژگی از هیستوگرام محلی گردایان تصاویر خاکستری شده به دست می‌آید (ساغر، ۲۰۱۰).

مراحل استخراج ویژگی هیستوگرام محلی گردایان به صورت زیر است:

برای اندازه‌گیری فاصله بین نمونه‌ها از الگوریتم استفاده شده است (رات ۳۰۰۳). دلیل استفاده از این الگوریتم وجود طول‌های متفاوت حروف در کلمات مشابه در نوشته‌های افراد مختلف است شکل (۵). الگوریتم سعی می‌کند تا حدودی با استفاده از فاصله نقاط همسایه، این تغییرات را جبران کند.



(شکل-۵): تصاویر یک کلمه با دست نوشته‌های متفاوت

۱-۳ معیار عملکرد الگوریتم کاهش فرهنگ لغت

اگر یک فرهنگ لغت شامل n کلمه $\{l_1, l_2, \dots, l_n\} = L$ باشد. یک سامانه کاهش‌دهنده فرهنگ لغت، یک زیرمجموعه‌ای $L' \subset L$ از نامزدها برای بازشناسی کلمه آزمون ورودی تولید می‌کند. کلماتی که در این زیرمجموعه نسیتند ($L - L'$) در بازشناسی کلمه آزمون ورودی شرکت نمی‌کنند. در فهرست کلمات نامزد هیچ رتبه‌بندی وجود ندارد.

برای اندازه‌گیری عملکرد سامانه کاهش‌دهنده فرهنگ لغت در مراجع (مصطفیری، ۲۰۰۸) و (مدهونت، ۲۰۰۱) تعدادی معیار معرفی شده است.

اگر اتفاق درست بودن کلمه‌ای از فرهنگ لغت (l_i) در زیرمجموعه L' را یک متغیر تصادفی A در نظر بگیریم. تعریف می‌شود که :

$$l_i \in L' \quad \begin{cases} A = 1 \\ A = 0 \end{cases} \quad (1)$$

درجه کاهش سامانه با متغیر R تعریف می‌شود:

$$R = \frac{|L| - |L'|}{|L|} \quad (2)$$

سه معیار اندازه‌گیری عملکرد سامانه کاهش‌دهنده فرهنگ لغت :

- دقت کاهش : $\alpha = E(A)$

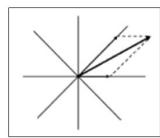
- درجه کاهش : $\rho = E(R)$

- درجه تأثیر کاهش : $\eta = \alpha k \cdot \rho$



(شکل-۷):الف- تصویر ورودی. ب- اندازه گرادیان. پ- جهت گرادیان

- برای استخراج ویژگی، جهت‌های گرادیان به ۳۲ جهت با فاصله‌های $16/\pi$ مقدار کمی می‌شوند. هر زاویه را می‌توان به دو روش مقدار کمی کرد. در روش اول هر جهت به دو جهت پایه تجزیه می‌شود. خطای مقدار کمی کردن در این روش کم است و هر جهت گرادیان به دو جهت پایه برچسب می‌خورد (لیو، ۲۰۰۷) شکل (۸).



(شکل-۸): مقدار کمی کردن با تجزیه بردار جهت گرادیان

- در روش دیگر هر جهت گرادیان به نزدیکترین جهت پایه برچسب می‌خورد. این روش نسبت به روش اول خطای مقدار کمی کردن بیشتر ولی ساده‌تر است.
- برای محاسبه هیستوگرام محلی جهت گرادیان، تصاویر گرادیان را به بلوک‌هایی تقسیم می‌کنیم. ابعاد 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 به گونه‌ای انتخاب می‌شود که پیچیدگی کم و در عین حال دقت بالا باشد.
- در هر بلوک اندازه گرادیان‌ها را در هر ۳۲ جهت با هم جمع می‌شود. بنابراین از هر بلوک یک بردار ۳۲ تایی استخراج می‌شود.

- اگر تعداد بلوک‌ها در جهت افقی n و در جهت عمودی m باشد، ابعاد ویژگی برابر با $m \times n \times 32$ خواهد شد. برای کاهش ابعاد ویژگی می‌توان تعداد بلوک‌ها را با فیلتر گوسی و تعداد جهت‌های پایه را با کاهش نمونه‌ها با بردار وزن [۱ ۴ ۶ ۴ ۱] کاهش داد (پال، ۲۰۰۹).

در تصاویر دستنوشتۀ فارسی، بلوک‌بندی منظم تصاویر، روش مناسبی برای استخراج ویژگی و مقایسه دو تصویر نمی‌باشد. چون بهدلیل تنوع در نحوه نوشتن این خط، در مثال‌های مختلف نگارش یک کلمه، ممکن است

- ابتدا تصاویر کلمات دستنوشته به تصاویر سطح خاکستری تبدیل می‌شوند. برای این منظور از فیلتر میانگین 3×3 استفاده می‌کنیم. در فیلتر میانگین مقدار هر پیکسل با مقدار میانگین شدت همسایگی 3×3 به مرکز آن پیکسل جایگزین می‌شود.

- برای محاسبه گرادیان روشناجی در هر پیکسل تصویر یکی از عملگرهای سوبل، روبرتز یا کریش استفاده می‌شود. عملگر کریش از این نظر که مؤلفه گرادیان را در چهار راستای افقی، عمودی، قطر اصلی و قطر فرعی می‌دهد، از بقیه پیچیده‌تر است. عملگرهای روبرتز و سوبل تنها دو مؤلفه گرادیان در اختیار می‌گذارند. در اینجا از عملگر سوبل استفاده شده است. این عملگر نسبت به عملگر کریش محاسبات کمتری دارد و با توجه به این که نسبت به عملگر روبرتز، همسایگی بزرگتری برای هر پیکسل در نظر می‌گیرد نتایج بهتری نسبت به عملگر روبرتز به دست می‌دهد. فیلتر عملگرهای سوبل در شکل (۶) آمده‌اند (لیو، ۲۰۰۷).

-1	0	1
-2	0	2
-1	0	1

1	2	1
0	0	0
-1	-2	-1

(شکل-۶): عملگرهای سوبل

- مؤلفه‌های گرادیان عمودی g_y و افقی g_x تصویر ورودی f در موقعیت (x, y) به کمک عملگر سوبل به صورت زیر محاسبه می‌شود :

$$g_x(x, y) = f(x+1, y-1) + 2f(x+1, y) + f(x+1, y+1) - f(x-1, y-1) - 2(x-1, y) - f(x-1, y+1)$$

$$g_y(x, y) = f(x-1, y+1) + 2f(x, y+1) + f(x+1, y+1) - f(x-1, y-1) - 2f(x, y-1) - f(x+1, y-1)$$

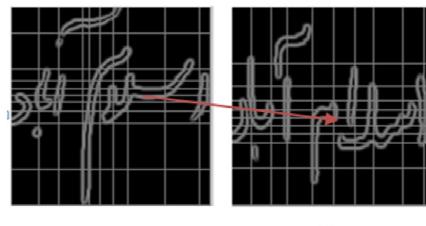
- جهت و اندازه گرادیان هر پیکسل به صورت زیر محاسبه می‌شود :

$$\text{اندازه} \quad s(x, y) = \sqrt{g_x^2 + g_y^2} \quad (۳)$$

$$\text{جهت} \quad \theta(x, y) = \tan^{-1}(g_y / g_x) \quad (۴)$$

در شکل (۷) تصاویر گرادیان یک نمونه از تصویر کلمه دستنوشته را نشان می‌دهد.





ب

الف

(شکل-۱۰): تصویر گرادیان بلوک‌بندی شده با
بلوک‌های تطبیقی روش اول

۲-۱-۴ روش دوم: بلوک‌بندی تطبیقی زیر کلمات
هر کلمه دست نوشته را می‌توان به مؤلفه‌های اصلی آن که زیر کلمه نیز نامیده می‌شوند تجزیه کرد. هر زیر کلمه به لحاظ تصویری یک جزء مستقل است. به عنوان مثال زیر کلمات تصویر دست نوشته "ایوانکی" در شکل (۱۱) شامل "ا"، "یو"، "ا"، "نکی" می‌باشد.

نظر به اینکه مقایسه زیر کلمات دو تصویر برای شناسایی کلمه، منطقی به نظر می‌رسد، در روش دوم پیشنهادی ابتدا کلمه را به مؤلفه‌های اصلی آن تجزیه کرده، سپس تصویر هر مؤلفه توسط روش پیشنهادی اول به بلوک‌های تطبیقی تقسیم شده و ویژگی‌های هیستوگرام گرادیان از هر مؤلفه استخراج می‌شود. ویژگی‌های استخراج شده از مؤلفه‌های اصلی تصویر کلمه در کنار هم قرار داده می‌شوند تا بردار ویژگی کلمه ساخته شود.

برای تجزیه کلمه به مؤلفه‌های اصلی، ابتدا خط کرسی کلمه بر اساس نمایه افکشن افقی بیدار می‌شود. در مرحله بعد مؤلفه‌های متصلی که از خط کرسی یا خطوط نزدیک به خط کرسی عبور می‌کنند و تعداد پیکسل هایشان از یک آستانه از پیش تعیین شده بیشتر است، به عنوان مؤلفه‌های اصلی کلمه انتخاب می‌شوند (اغبری، ۲۰۰۹).

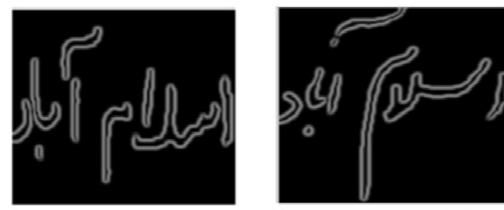
در بررسی کلمات موجود در پایگاه داده مورده استفاده ("ایران شهر")، بیشینه تعداد مؤلفه‌های اصلی کلمات برابر با هفت به دست آمد. برای قابل مقایسه بودن بردارهای ویژگی کلمات، تعداد مؤلفه‌های اصلی تمام کلمات را هفت در نظر گرفتیم. بر این اساس ابعاد بردار ویژگی استخراج شده از هر تصویر برای $32 \times n \times m$ است شکل (۱۱). در صورتی که تعداد مؤلفه‌های اصلی یک کلمه کمتر از هفت باشد، ویژگی‌های متناظر با آن زیر کلمه صفر در نظر گرفته می‌شود.

بلوک‌های متناظر تصاویر محتوی بخش‌های متناظر از حروف کلمه نباشند؛ بدین دلیل ویژگی‌های هیستوگرام جهت گرادیان محلی با بلوک‌بندی منظم، مناسب برای مقایسه کلمات دست نوشته فارسی نیستند. شکل (۹) دو مثال از یک کلمه در بلوک‌بندی منظم را نشان می‌دهد.

در این پژوهش برای کاهش عدم تناقض بلوک‌ها، دو روش بلوک‌بندی تطبیقی پیشنهاد شده است.

۴-۱-۱-۱- روش اول : بلوک‌بندی تطبیقی کل کلمه

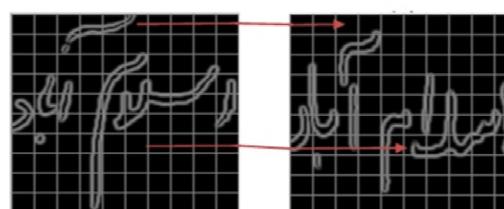
در این روش اندازه بلوک‌ها در یک تصویر ثابت نیست و بر اساس توزیع پیکسل‌های قلم در سطرها و ستون‌های تصویر تغییر می‌کند. برای بلوک‌بندی تصویر به $n \times m$ بلوک ابتدا تصویر به n نوار افقی تقسیم می‌شود؛ به طوری که تعداد پیکسل‌های سیاه (اثر قلم) در همه نوارها با هم برابر باشند؛ سپس تصویر به m نوار عمودی تقسیم می‌شود؛ به طوری که تعداد پیکسل‌های سیاه در همه نوارهای عمودی نیز با هم برابر باشد. به این ترتیب اندازه بلوک‌ها در یک تصویر کلمه با توجه به توزیع پیکسل‌های اثر قلم تغییر می‌کند. چنانچه در هر بلوک از ۳۲ جهت گرادیان در هیستوگرام استفاده شود، ابعاد بردار ویژگی یک تصویر برابر با $m \times n \times 32$ است. شکل (۱۰)



ب



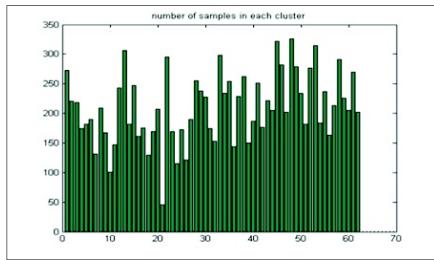
الف



ت

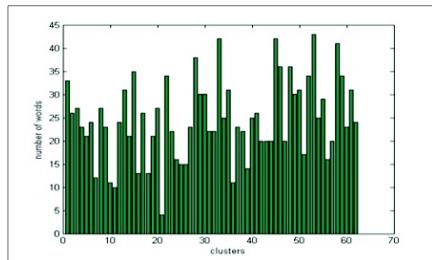
پ

(شکل-۹): (الف) (ب) تصاویر اندازه گرادیان نمونه‌های یک طبقه و (پ) (ت) تصاویر بلوک‌بندی شده با اندازه بلوک‌های ثابت



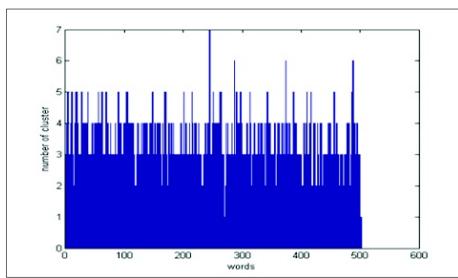
(شکل-۱۲): نمودار تعداد نمونه‌ها در هر خوشه

برای تعیین تعداد کلمه‌ها (طبقه‌ها) در هر خوشه، هر کلمه‌ای که حداقل سه نمونه‌اش در آن خوشه قرار بگیرد، متعلق به آن خوشه تلقی می‌شود. نمودار تعداد کلمه‌ها در هر خوشه در شکل (۱۳) نشان داده است. به طور میانگین تعداد کلمه‌ها در هر خوشه، ۲۳ است.

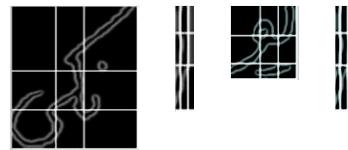


(شکل-۱۳): نمودار تعداد طبقه‌ها در هر خوشه

همان‌طور که گفته شده نمونه‌های هر کلمه ممکن است در چند خوشه قرار بگیرد. شکل (۱۴) نمودار تعداد خوشه‌هایی که نمونه‌های هر کلمه در آن قرار گرفته است، نشان می‌دهد. (تعداد کلمه‌ها در پایگاه داده "ایران شهر" برابر با ۵۰۳ کلمه است) میانگین تعداد خوشه برای نمونه‌های هر کلمه سه است؛ یعنی به طور میانگین نمونه‌های هر کلمه در سه خوشه قرار می‌گیرد.



(شکل-۱۴): نمودار تعداد خوشه‌هایی که هر کلمه در آن قرار گرفته است



(شکل-۱۱): تصویر گرادیان اجزای اصلی بلوک‌بندی شده کلمه دست‌نوشته

۵- نتایج

در این بخش ابتدا عملکرد مرحله کاوش فرهنگ لغت ارزیابی می‌شود، سپس نتایج مرحله بازشناسی کلمه دست‌نوشته ورودی از فهرست نامزد به دست آمده از مرحله اول ارائه خواهد شد. برای ارزیابی الگوریتم پیشنهادی از پایگاه داده "ایران شهر" شامل ۱۶۰۰۰ تصویر کلمه دست‌نوشته از ۵۰۳ نام شهرهای ایران استفاده شده است. برای هر شهر حداقل ۲۵ تصویر دست‌نوشته وجود دارد که توسط افراد مختلف نوشته شده است. تصویر برای مرحله آموزش الگوریتم و ۳۰۰۰ تصویر برای مرحله آزمون به صورت تصادفی انتخاب شده است.

۱-۵- نتایج مرحله کاوش فرهنگ لغت

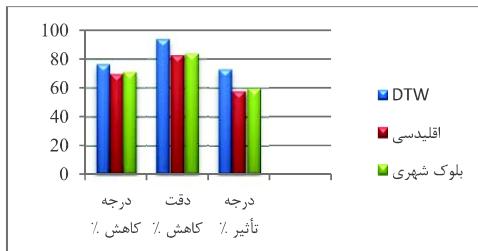
خوشه‌بندی تصاویر کلمات دست‌نوشته توسط الگوریتم خوشه‌بندی سلسله‌مراتبی و الگوریتم ISOCLUS با معیار اندازه‌گیری فاصله DTW انجام شد. بدین ترتیب کلمات با شکل کلی مشابه در خوشه‌های یکسان قرار گرفته‌اند. میانگین هر خوشه به عنوان مدخل اعضا آن خوشه در نظر گرفته شده است. برای خوشه‌بندی کلمات از تمام نمونه‌های آموزشی هر کلمه استفاده شده است.

تعداد اولیه خوشه‌های به دست آمده توسط الگوریتم خوشه‌بندی سلسله‌مراتبی برابر با ۴۵ و تعداد خوشه‌های نهایی در الگوریتم ISOCLUS برابر با ۶۲ خوشه است. نمودار تعداد نمونه‌های قرار گرفته در هر خوشه بر حسب خوشه‌ها در شکل (۱۲) نشان داده شده است. به طور میانگین در هر خوشه ۲۰۹ نمونه وجود دارد.

فصل نهم



از الگوریتم DTW برای اندازه‌گیری فاصله بین نمونه‌ها و فاصله بین کلمه آزمون ورودی و میانگین خوش‌ها استفاده شده است. در نمودار شکل (۱۵) نتایج بدست آمده با الگوریتم DTW را با نتایج حاصل از سایر معیارهای فاصله مانند فاصله اقلیدسی و بلوک شهری مقایسه می‌کنیم.



(شکل-۱۵): مقایسه عملکرد الگوریتم DTW با معیار فاصله اقلیدسی و بلوک شهری

همانطور که نمودار نشان می‌دهد نتایج بدست آمده با الگوریتم DTW بهتر از معیار بلوک شهری و نتایج بدست آمده از معیار بلوک شهری بهتر از فاصله اقلیدسی است.

۲-۵ نتایج مرحله بازشناسی

در مرحله اول روش پیشنهادی با استفاده از خوشبندی تصاویر کلمات دستنوشته توسط ویژگی‌های کلی کلمه فهرستی از کلمات نامزد برای بازشناسی کلمه آزمون ورودی تولید می‌شود. در مرحله دوم با استفاده از بردار ویژگی هیستوگرام گردیدن محلی بهترین نامزد از فهرست بدست آمده در مرحله اول برای شناسایی کلمه ورودی انتخاب می‌شود. برای طبقه‌بندی در این بخش از طبقه‌بندی K نزدیک‌ترین همسایه استفاده شده است.

همان‌طوری که اشاره شد در این تحقیق روشی جهت بلوک‌بندی تطبیقی تصویر کلمه ارائه شد. از هر بلوک به دست آمده بردار ویژگی هیستوگرام گردیدن محلی استخراج شد که با در کنار هم گذاشتن این ویژگی‌ها یک بردار ویژگی برای هر تصویر کلمه ساخته می‌شود. جدول (۲) نتایج بازشناسی کلمات دستنوشته (بدون مرحله کاهش فرهنگ لغت) به ازای تعداد بلوک‌های مختلف در تصویر نشان داده شده است.

با توجه به جدول، در تعداد بلوک‌های بیشتر، ویژگی هیستوگرام اصلاح شده به روش اول نسبت به دو روش دیگر دقیق بازشناسی بالاتری دارد. البته این به قیمت افزایش ابعاد بردار ویژگی و درنتیجه افزایش بار محاسباتی سامانه

برای ارزیابی عملکرد الگوریتم کاهش‌دهنده فرهنگ لغت از دو معیار دقت و درجه کاهش استفاده می‌شود که این دو معیار با هم ترکیب شده و معیار درجه تأثیر را تولید می‌کند. بنابراین معیار درجه تأثیر برای ارزیابی عملکرد الگوریتم کاهش‌دهنده فرهنگ لغت مناسب است.

در ارزیابی عملکرد الگوریتم کاهش فرهنگ لغت پیشنهادی، فاصله بردار ویژگی استخراج شده از کلمه آزمون ورودی با میانگین بردار ویژگی اعضای تشکیل‌دهنده هر خوش توسط الگوریتم DTW محاسبه می‌شود. تعدادی (n) از نزدیک‌ترین خوش‌ها به کلمه آزمون ورودی انتخاب می‌شود. دقت و درجه کاهش الگوریتم وابسته به این تعداد (n) است. اگر n زیاد باشد، دقت افزایش، ولی درجه کاهش فرهنگ لغت، کم می‌شود و اگر n کم باشد، دقت کاهش می‌یابد؛ اما درجه کاهش فرهنگ لغت، زیاد می‌شود. برای تعیین مقدار n مطلوب، استفاده از معیار درجه تأثیر مناسب‌تر است. جدول (۱) میزان دقت، درجه کاهش و درجه تأثیر را برای مقادیر مختلفی از n برای سه هزار تصویر مجموعه آزمون نشان می‌دهد.

با توجه به جدول، با انتخاب پنج خوش نزدیک به تصویر کلمه آزمون ورودی، احتمال وجود کلمه آزمون در این پنج خوش ۹۴/۱۲٪ و میزان کاهش تعداد نمونه‌های مورد بررسی برای بازشناسی تصویر ورودی ۹۲٪ و درجه تأثیر الگوریتم ۸۶/۶۹٪ خواهد شد. یعنی در این مرحله با توجه به این که در هر خوش به‌طور متوسط ۲۰۹ نمونه وجود دارد، تعداد نمونه‌های مورد بررسی برای بازشناسی تصاویر آزمون ورودی از ۱۳۰۰۰ به ۱۰۵۰ کاهش می‌یابد.

(جدول-۱): دقت، درجه کاهش و درجه تأثیر برای مقادیر

n مختلفی از

خوشه‌های نزدیک (n)	درجه کاهش٪	دقت کاهش٪	درجه تأثیر٪
۱	۹۸/۳	۷۵/۲۸	۷۴
۲	۹۶/۷	۸۴/۲۷	۸۱.۴۸
۵	۹۲	۹۴/۱۴	۸۶/۶۹
۷	۸۸/۷	۹۴/۳۴	۸۳/۶۷
۱۰	۸۳/۹۲	۹۵/۰۱	۸۷/۴
۱۲	۸۰/۷	۹۵/۹	۷۷/۳۴
۱۵	۷۵/۸	۹۷/۰۷	۷۳/۵۷
۲۰	۶۷/۸	۹۸/۳۷	۶۶/۷

جدول (۳) دقت و سرعت بازشناسی را با استفاده از ویژگی گرادیان استخراج شده به روش اول پیشنهادی با اضافه کردن مرحله کاهش فرهنگ لغت نشان می‌دهد. در این آزمایش از ۲۵ بلوک (۵×۵)، در بلوک‌بندی تطبیقی استفاده شده است. همان‌طور که جدول (۳) نشان می‌دهد، استفاده از مرحله کاهش فرهنگ لغت به طور چشم‌گیری باعث افزایش دقت می‌شود. برخلاف اینکه انتظار می‌رود کاهش فرهنگ لغت زمان بازشناسی کلمه را کاهش دهد، اما به دلیل وقت‌گیری‌بودن الگوریتم DTW در کاهش فرهنگ لغت سرعت بازشناسی با کاهش فرهنگ لغت خیلی کاهش نمی‌یابد.

(جدول -۳): دقت و سرعت بازشناسی برای روش‌های مختلف

روش	دقت بازشناسی٪ (ثانیه بر هر نمونه)	سرعت بازشناسی٪ (ثانیه بر هر نمونه)
بازشناسی بدون مرحله کاهش فرهنگ لغت	۷۶/۶۲	۰/۶۶
بازشناسی با مرحله کاهش فرهنگ لغت	۸۱/۱۹	۰/۶۱

جهت مقایسه روش پیشنهادی با روش‌های دیگر، روش ارائه شده در این مقاله با روش پیشنهادشده در مرجع (مظفری، ۲۰۰۸)، که نخستین و تنهایترین روشی است که از مرحله کاهش فرهنگ لغت در بازشناسی کلمات دست‌نوشته فارسی استفاده کرده است، مقایسه می‌کنیم. پایگاه داده به کار رفته در این روش نسخه اولیه پایگاه داده ایران‌شهر که شامل ۲۰۰ نام شهر ایران است، می‌باشد. نتایج مقایسه در جدول (۴) آمده است.

(جدول -۴): مقایسه دقت بازشناسی روش‌های مختلف

روش‌ها	تعداد کلمات فرهنگ لغت	دقت بازشناسی٪
روش ارائه شده در (مظفری، ۲۰۰۸)	۲۰۰	۶۱/۷۲
روش پیشنهادی در این مقاله	۵۰۰	۸۱/۱۹

شکل (۱۷) تعدادی از تصاویر کلمات دست‌نوشته‌ای که در روش دو مرحله‌ای پیشنهادی به‌اشتباه بازشناسی شده‌اند، آورده شده است.

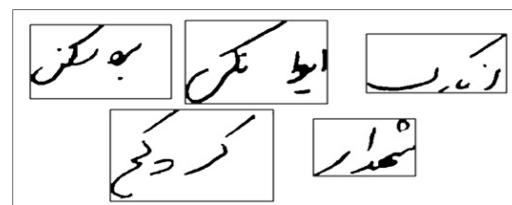
شناسایی است. با در نظر گرفتن پیچیدگی محاسباتی و دقت بازشناسی قابل قبول، ویژگی هیستوگرام گرادیان محلی اصلاح شده به روش اول با تعداد بلوک‌های ۵×۵ مناسب است.

دلیل کارایی پایین‌تر ویژگی گرادیان روشنایی استخراج شده به روش پیشنهادی دوم در بازشناسی تصاویر دست‌نوشته را می‌توان خطاهای ایجاد شده در جداسازی مؤلفه‌های اصلی کلمه دانست. همان‌طوری که اشاره شد در این روش ابتدا مؤلفه‌های اصلی کلمه جدا می‌شوند؛ سپس تصویر هر مؤلفه به کمک بلوک‌بندی تطبیقی تحلیل شده و بردار ویژگی هیستوگرام از آن استخراج می‌شود. جداسازی تا حدودی ناموفق مؤلفه‌های اصلی کلمه در این روش باعث خطا در شناسایی کلمات می‌شود.

دلایل خطاهای ایجاد شده در جداسازی مؤلفه‌های اصلی متفاوت است. در برخی موارد به خاطر پس‌بین مؤلفه‌های اصلی خطای ایجاد می‌شود. همچنین فاصله داشتن بعضی از مؤلفه‌های اصلی از خط کرسی باعث می‌شود به عنوان مؤلفه اصلی در نظر گرفته نشوند. جایه‌جايان مکمل‌هایی مثل سرکش "ک و گ" و کلاه "آ" از مؤلفه اصلی اش و یا بزرگ نوشتند بعضی از نقاط و علائم باعث می‌شود به عنوان یک مؤلفه اصلی در نظر گرفته شوند. در شکل (۱۶) برخی از کلمات که در جداسازی مؤلفه‌های اصلی خطای داشته‌اند، نشان داده شده است.

(جدول -۲): دقت بازشناسی برای ویژگی‌های مختلف

تعداد بلوک‌ها ویژگی	۳×۳	۵×۵	۷×۷	۱۰×۱۰
هیستوگرام گرادیان محلی	۴۹/۷۸	۶۵/۳۲	۶۷/۱۴	۶۹/۴۴
هیستوگرام گرادیان محلی اصلاح شده به روش اول	۵۵/۳۲	۷۶/۶۲	۷۸/۱۲	۷۹/۷۰
هیستوگرام گرادیان محلی اصلاح شده به روش دوم	۷۰/۱۰	-	-	-



(شکل -۱۶): برخی از کلماتی که در جداسازی مؤلفه‌های اصلی خطای داشته‌اند

فصل نهم



کاهش فرهنگ لغت نقش چشم‌گیری در افزایش دقت سامانه بازشناسی دارد و تا حدودی سرعت بازشناسی را افزایش می‌دهد.

در روش دوم برای استخراج ویژگی از تصویر کلمه، پیشنهاد شد تا کلمه به مؤلفه‌های اصلی تشکیل‌دهنده آن تجزیه شود؛ سپس از هر مؤلفه مشابه روش نخست بردار ویژگی مبتنی بر هیستوگرام گردیدن استخراج شود. نتایج نشان می‌دهد این روش نرخ بازشناسی پایین‌تری از روش نخست (استخراج ویژگی از تصویر کلی کلمه) دارد. دلیل این مسئله خطای در تجزیه کردن کلمه به مؤلفه‌های اصلی آن است. در صورتی که بتوان این خطای را کاهش داد، انتظار می‌رود نرخ بازشناسی روش دوم از روش پیشنهادی نخست بیشتر شود.

۶- مراجع

خسروی، ح، کبیر، ا، (۱۳۸۸)"ارزیابی روش‌های بازشناسی متون فارسی بر مبنای شکل کلی زیرکلمات"نشریه مهندسی برق و کامپیوتر زمستان، ۷(۴): ۲۸۰-۲۶۷.

خسروی، ح، کبیر، ا... (۱۳۸۵) "معرفی دو ویژگی سریع و کارآمد برای بازشناسی ارقام دستنویس فارسی" چهارمین کنفرانس ماشین‌بینایی و پردازش تصویر.

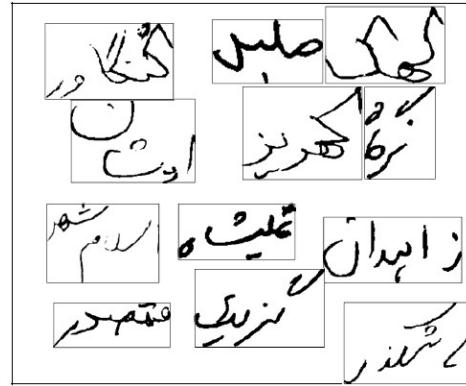
شیرعلی شهرضا، م، فائز، ک، (۱۳۷۴)، "تشخیص کلمات و ارقام دستنویس فارسی به‌وسیله شبکه‌های عصبی (خط نسخ)"، رساله دکترای مهندسی برق و کامپیوتر، دانشگاه صنعتی امیرکبیر.

Awaidah. S. M., Mahmoud. S.A. (2009) ;" A multiple feature/resolution scheme to Arabic (Indian) numerals recognition using hidden Markov models " ,Signal Processing 89 1176–1184.

Broumandnia. A, Shanbehzadeh .J, Rezakhah. M, faderani .V(2008)," Persian/arabic handwritten word recognition using M-band packet wavelet transform" , Image and Vision Computing 26, 829–842K.

Chen. M.Y, Kundu. A, Srihari .S.N.(1995)," Variable duration hidden Markov and morphological segmentation for handwritten word recognition" , IEEE Trans. Image Process. 4 (12) 1675–1688.

Cheriet. M, Kharma. N, liu .C ,ching Y.suen, (2007) , " character recognition system, A Guide for Students and Practitioners" wiley.



(شکل - ۱۷): تعدادی از تصاویر کلمات دستنوشته که به استیاه بازشناسی شده‌اند

۶- نتیجه‌گیری

در این مقاله یک روش دومرحله‌ای برای بازشناسی کلمات دستنوشته بیان شد. در مرحله نخست برای کاهش فضای جستجو، دامنه فرهنگ لغت نامزد برای کلمه مورد آزمون ورودی کاهش می‌یابد. برای این منظور از خوشبندی تصاویر توسط الگوریتم خوشبندی سلسه‌مراتبی و ISOCLUS استفاده کردیم. تعداد خوشبندی به دست آمده برای پایگاه داده "ایران شهر" با ۵۰۳ نام شهر ۶۲ است. میانگین بردار ویژگی نمونه‌های هر خوشبندی عنوان ورودی آن خوشبندی در نظر گرفته شده است. در خوشبندی بردار ویژگی مورد استفاده مبتنی بر نمایه‌های عمودی تصویر است. برای هر کلمه ورودی، پنج خوشبندی نزدیک به کلمه انتخاب می‌شوند و تصاویر نمونه‌های داخل این خوشبندی فهرست نامزدها برای بازشناسی کلمه قرار می‌گیرند. نتایج آزمایش‌ها نشان می‌دهد که در این مرحله با احتمال ۹۴٪ کلمه مورد بررسی در فهرست نامزد قرار می‌گیرد و تعداد نمونه‌های مورد بررسی ۹۲٪ کاهش می‌یابد.

در مرحله دوم که مربوط به بازشناسی کلمه است دو روش برای استخراج ویژگی از تصویر کلمه پیشنهاد می‌شود. در روش نخست تصویر کلمه به صورت تطبیقی بلوک‌بندی می‌شود؛ در این بلوک‌بندی چگالی لبه‌ها در بلوک‌ها تاحدودی یکنواخت می‌شود. از هر بلوک بردار ویژگی هیستوگرام گردیدن استخراج می‌شود. با کنار هم قرار گرفتن ویژگی‌های بالا بردار ویژگی تصویر کلمه تشکیل می‌شود. نتایج تجربی نشان می‌دهد بلوک‌بندی تطبیقی تصویر نسبت به بلوک‌بندی یکنواخت کارایی بسیار بالاتری در شناسایی کلمات دارد. همچنین نتایج نشان می‌دهد که

Pal. U, Wakabayashi .T, Kimura .F(2009)," Comparative Study of Devnagari Handwritten Character Recognition using Different Feature and Classifiers" , 10th International Conference on Document Analysis and Recognition.

Plamondon .R, Srihari S.N. (2000)," On-line and off-line handwriting recognition: A comprehensive survey". IEEE Trans Pattern Analysis and Machine Intelligence 11(1):68–89.

Ramy Al-Hajj .M, Laurence. L, and Chafic .M. (2009)," Combining Slanted-Frame Classifiers for Improved HMM-Based Arabic Handwriting Recognition" IEEE Transactions on Pattern Analysis and Machine Intelligence, VOL. 31, NO. 7.

Rath .T. M. and Manmatha .R. (2003)," Word Image Matching Using Dynamic Time Warping " 1063-6919/03 \$17.00 © IEEE.

Sagheer. M.W, Lei He. C, Nobile .N, Suen C. Y. (2010)," Holistic Urdu Handwritten Word Recognition Using Support Vector Machine" , 1051-4651/10 \$26.00 © IEEE.

Vaseghi. B, Alirezaee .S, Ahmadi .M, Amirkhattabi.R(2008)," Off-line Farsi / Arabic Handwritten Word Recognition Using Vector Quantization and Hidden Markov Model" , 978-1-4244-2824-3/08/\$25.00 © IEEE.

Wshah. S, Govindaraju. V, Cheng. Y ,and Li .H,(2010), "A Novel Lexicon Reduction Method For Arabic Handwriting Recognition", International Conference on Pattern Recognition. 1051-4651/10 \$26.00 © IEEE.



الهام بايسته تاشک مدرک کارشناسی
و کارشناسی ارشد خود را در رشته
الکترونیک در سال ۱۳۹۱ از دانشگاه
شهرورد دریافت کرده است . موضوعات
پژوهشی ایشان بارشناسی آماری الگو، پردازش تصویر و
صوت است.

نشانی رایانمۀ ایشان عبارت است از:

Bayesteh.el@gmail.com

Dehghan. M, Faez. K, Ahmadi. M and Shridhar .M(2001), "Handwritten Farsi (Arabic) word recognition: a holistic approach using discrete HM-M", Pattern Recognition ,vol. 34, no. 5, pp. 1057-1065.

Guillevic .D., Suen. C.Y. (1998), "HMM-KNN word recognition engine for bank cheque processing", Proceedings of International Conference on Pattern Recognition, Vol. 2, Brisbane, Australia, August, pp. 1526:1529.

Haji. M. M , Eghbali H. J, Katebi. S. D. (2005)," Farsi Handwritten Word Recognition Using Continuous Hidden Markov Models and Structural Features" MS These .

Hilditch .C. J. (1983), "Comparison of thinning algorithms on a parallel processor". Image Vision Computing, 1, 115–132.

Kavallieratou. E, Fakotakis .N, and Kokkinakis .G. (2000), "A Slant Removal Algorithm", Pattern Recognition., pp. 1261-1262.

Koerich A. L, Sabourin.R. C, Suen .Y (2003)," Large vocabulary off-line handwriting recognition: A survey" ,Pattern Anal Applica 6, 97–121.

Liu .C.L(2007)," Normalization-Cooperated Gradient Feature Extraction for Handwritten Character Recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence, VOL. 29, NO. 8.

Lorigo .L. M, and Govindaraju. V, (2006), " Offline Arabic Handwriting Recognition: A Survey" IEEE Transactions on Transactions on Pattern Analysis and Machine Intelligence . VOL. 28, NO. 5.

Madhvanath. S, Govindaraju. V,(2001)," The role of holistic paradigms in handwritten word recognition". IEEE Trans Pattern Analysis and Machine Intel-ligence, 23(2):149–164.

Memarsadeghi .N, Mount. D. M., Netanyahu. N. S., Moigne. J. L.(2003)," A Fast Implementation of the ISOCLUS Algorithm" ,0 – 7803 – 7929 .

Mozaffari .S, Faez .K, Ma'rgner .V, El-Abed .H , (2008)," Lexicon reduction using dots for off-line Farsi/Arabic handwritten word recognition "Pattern Recognition Letters 29 , 724–734.

Otsu. N.(1979), "A Threshold Selection Method from Gray Level Histograms", IEEE Trans. on Systems, Man and Cybernetics, vol. 9, pp. 62-66.



علیرضا احمدی‌فرد مدرک کارشناسی



را از دانشگاه صنعتی اصفهان در رشته
مهندسی الکترونیک و کارشناسی ارشد
را از دانشگاه صنعتی امیرکبیر در رشته
مهندسی مخابرات اخذ کردند. ایشان

مدرک دکترای تخصصی از مرکز CVSSP دانشگاه Surrey در سال ۲۰۰۲ اخذ کردند. ایشان از سال ۱۳۸۷ در دانشکده مهندسی برق و رباتیک دانشگاه شاهroud مشغول به فعالیت است. زمینه‌های پژوهشی مورد علاقه ایشان تشخیص و آشکارسازی اشیا در تصاویر و تشخیص هویت از روی تصاویر است.

نشانی رایانمۀ ایشان عبارت است از:

ahmadyfard@shahroodut.ac.ir

حسین خسروی در سال ۱۳۶۰ در



حضوری متولد شد. ایشان مدرک کارشناسی خود را در سال ۱۳۸۲ از دانشگاه صنعتی شریف و مدرک کارشناسی ارشد و دکترای خود را

به ترتیب در سال ۱۳۸۴ و ۱۳۸۷ از دانشگاه تربیت مدرس دریافت کرد. ایشان از سال ۱۳۸۷ در دانشکده مهندسی برق و رباتیک دانشگاه شاهroud مشغول به فعالیت است.

زمینه پژوهشی ایشان پردازش تصاویر، متون و صوت است.
نشانی رایانمۀ ایشان عبارت است از:

hosseinkhosravi@chmail.ir

