

بازشناسی کلمات دست‌نویس فارسی به کمک تبدیل قطبی - لگاریتمی و مدل مخفی مارکوف

غلامرضا نادعلی نیا چاری، خشایار یغمایی، حامد فضل‌اللهی آقاملک و سیدمحمد رضوی

دست‌نویس فارسی و عربی می‌پردازیم. در [۳] از مدل مخفی مارکوف نیمه‌پیوسته یک‌بعدی برای بازشناسی کلمات دست‌نویس عربی استفاده شده است. در این شیوه پس از مرحله پیش‌پردازش طول و عرض تصویر کلمه ورودی نرمالیزه شده، سپس کجی موجود در تصویر کلمه اصلاح می‌شود و بعد به کمک پنجره لغزان عمودی استخراج ویژگی صورت می‌گیرد. برای ارزیابی این سیستم از پایگاه داده IFN/ENIT استفاده شده است. در [۶] برای بازشناسی کلمات دست‌نویس نام ۱۰۰ شهر مختلف، از مدل مخفی مارکوف استفاده شده است. در این سیستم بازشناسی ابتدا تصویر کلمه دست‌نویس (نام شهر) به ابعاد مناسب نرمالیزه می‌شود، سپس از تصویر نرمالیزه‌شده ویژگی‌هایی چون طول نرمالیزه‌شده، پیمایش، شیب، نوع همسایگی، نقاط انتهایی، تعداد قطعه‌ها و لوپ‌ها استخراج می‌شود و این ویژگی‌های استخراج‌شده به صورت یک بردار توالی مشاهده مرتب می‌شوند. در [۷] یک روش جدید برای بازشناسی کلمات فارسی و عربی ارائه شده است. برای تشکیل بردار ویژگی ابتدا تبدیل قطبی تصویر کلمات محاسبه می‌شود، سپس از این تصویر به دست آمده تبدیل موجک گرفته می‌شود. به این ترتیب برای هر کلمه یک بردار ویژگی تشکیل می‌شود که پایگاه داده کلی را تشکیل می‌دهد. در انتها به کمک یک طبقه‌بند، کلمات پیشنهادی سیستم بازشناسی بر اساس فاصله مورد نظر لیست می‌شوند. نتایج تجربی به دست آمده از ارزیابی‌های مختلف کلمات با سایز و چرخش‌های مختلف نشان‌دهنده عملکرد بهتر این روش نسبت به روش‌های تبدیل فوریه، تبدیل موجک و گشتاورهای زرنیکی می‌باشد.

از مراحل مهم در یک سیستم بازشناسی کلمات، مرحله استخراج ویژگی است. با توجه به این که کلمات دست‌نویس دارای تغییرات نگارشی مختلفی هستند بنابراین باید از یک طرف، یک ویژگی از تصویر کلمات استخراج شود که در مقابل این تغییرات نگارشی مقاوم بوده و از طرف دیگر بتواند نماینده منحصر به فرد آن در یک کلاس خاص باشد. بر خلاف بیشتر سیستم‌های بازشناسی بر پایه HMM [۸] تا [۱۱] که از پنجره لغزان در مرحله استخراج ویژگی به صورت یکنواخت نمونه‌گیری می‌کنند، در این مقاله روشی ارائه می‌شود که استخراج ویژگی به صورت غیر یکنواخت و متمرکز صورت می‌گیرد. در واقع به کمک تبدیل قطبی - لگاریتمی^۱ که در ادامه مورد بحث قرار می‌گیرد می‌توان از تصویر کلمه به گونه‌ای نمونه‌برداری انجام داد که بیشترین نمونه‌ها در یک ناحیه خاص متمرکز باشد و هر چه از آن ناحیه به سمت حاشیه و پیرامون تصویر کلمه پیش می‌رویم نقاط نمونه‌برداری شده کمتر شود. مزیت این روش این است که می‌توان تعداد نقاط نمونه‌برداری از بخشی از تصویر کلمه که دارای اطلاعات (پیکسل‌های سیاه) بیشتری است را افزایش داد. همچنین از آنجا که چرخش و تغییر مقیاس در تصویر کلمه به صورت شیب در

چکیده: در این مقاله یک سیستم بازشناسی کلمات فارسی معرفی می‌شود که از خودهمبستگی محلی مرتبه بالای تصویر قطبی - لگاریتمی برای استخراج ویژگی از زیرکلمات فارسی استفاده می‌کند. این شیوه استخراج ویژگی باعث می‌شود سیستم در مقابل تغییرات نگارشی مانند تغییر مقیاس‌های خطی و چرخش مقاوم شود. همچنین به کمک تبدیل قطبی - لگاریتمی، نمونه‌برداری از تصویر زیرکلمه به صورتی انجام شده که بیشترین نمونه‌ها در یک ناحیه خاص متمرکز باشد. در روش ارائه‌شده از مدل مخفی مارکوف گسسته به عنوان طبقه‌بند و همچنین برای افزایش امنیت و دقت خروجی سیستم بازشناسی از یک شبکه واژه‌نامه استفاده شده و برای ارزیابی سیستم از پایگاه داده ایران‌شهر استفاده شده بود. مقایسه نتایج حاصل از روش پیشنهادی با نتایج سایر روش‌های استخراج ویژگی مؤید این است که سیستم بازشناسی پیشنهادشده در این مقاله از حساسیت کمتری نسبت به تغییرات نگارشی برخوردار است.

کلید واژه: تبدیل قطبی - لگاریتمی، مدل مخفی مارکوف، بازشناسی کلمات دست‌نویس فارسی.

۱- مقدمه

تحقیقات گسترده‌ای در زمینه بازشناسی نوشتار انجام شده و می‌شود. بازشناسی دست‌نویس^۱ به دلیل راحت‌تر بودن نوشتن از تایپ کردن، سرعت نوشتن تا تایپ کردن، عدم امکان تایپ در بعضی موارد و عدم وجود یک صفحه کلید کامل روی کامپیوترهای کوچک، مورد توجه خاصی قرار گرفته است. روش‌های بازشناسی کلمات دست‌نویس شناخته شده‌اند و در وسعت زیادی برای کلمات در زبان‌های مختلف استفاده می‌شوند به خصوص این که روش‌های زیادی مبتنی بر مدل مخفی مارکوف^۲ (HMM) به طور موفقیت‌آمیزی برای بازشناسی کلمات دست‌نویس پیوسته و بدون محدودیت پیاده‌سازی شده است [۱] تا [۵]. اما مشکل اساسی این است که تفاوت وسیعی بین شکل‌ها و منحنی‌های دست‌نویس زبان‌های فارسی و لاتین وجود دارد که کار بازشناسی را برای زبان فارسی پیچیده‌تر می‌کند. در زیر به بررسی چند مقاله در زمینه بازشناسی کلمات

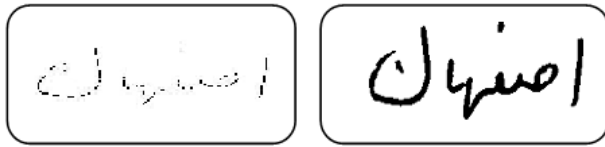
این مقاله در تاریخ ۲ تیر ماه ۱۳۹۱ دریافت و در تاریخ ۱۱ آبان ماه ۱۳۹۲ بازنگری شد.

غلامرضا نادعلی نیا چاری، دانشکده مهندسی برق و کامپیوتر، دانشگاه سمنان، سمنان، (email: Nadaliniareza@gmail.com).
خشایار یغمایی، دانشکده مهندسی برق و کامپیوتر، دانشگاه سمنان، سمنان، (email: Khyaghmaie@semnan.ac.ir).
حامد فضل‌اللهی آقاملک، دانشکده مهندسی برق و کامپیوتر، دانشگاه بیرجند، بیرجند، (email: Fazlollahi.hamed@yahoo.com).
سیدمحمد رضوی، دانشکده مهندسی برق و کامپیوتر، دانشگاه بیرجند، بیرجند، (email: smrazavi@birjand.ac.ir).

1. Handwritten
2. Hidden Markov Model



شکل ۲: (الف) تصویر کلمه اصفهان و (ب) تصویر زیرکلمات جداسازی شده به روش رنگ‌آمیزی.



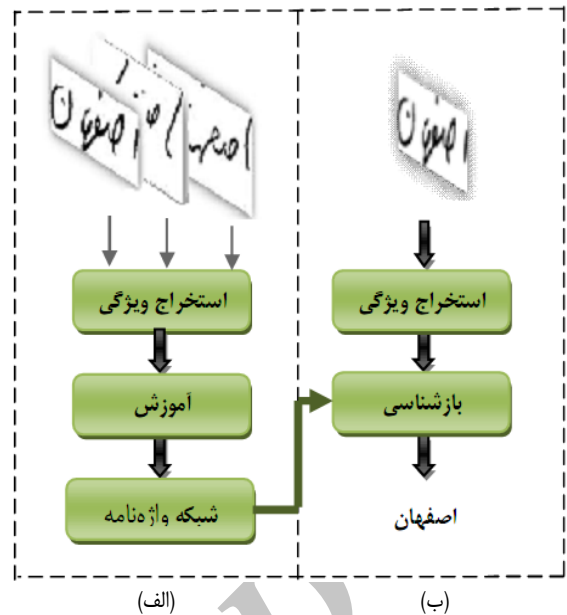
شکل ۳: (الف) تصویر کلمه اصفهان و (ب) تصویر کلمه اصفهان بعد از مرحله نازک‌سازی.

از اخذ به کمک فیلترهای میانگین‌گیری^۲ [۱۲] فیلتر می‌شود تا نقاط گسستگی تصویر محو یا کم‌رنگ شود و سپس تصویر کلمه به کمک آستانه‌گیری به یک تصویر باینری تبدیل می‌شود. در مرحله بعد کلیه نقطه‌ها از تصویر کلمه برداشته شده و در تصویر مجزای دیگری ذخیره می‌شوند، به عبارتی تصویر نقطه‌ها جداگانه و تصویر زیرکلمات نیز به صورت جداگانه ذخیره می‌شوند. اطلاعات مربوط به نقاط در نهایت برای افزایش دقت بازشناسی به کار گرفته شده و پس از آن مشابه [۱۳]، به کمک روش رنگ‌آمیزی^۳ زیرکلمه‌ها از هم جدا می‌شوند و قطعه‌بندی به صورت زیرکلمات انجام می‌شود (شکل ۲).

در مرحله بعد برای حذف برخی از تغییرات نگارشی، برای هر زیرکلمه نازک‌سازی اعمال می‌شود [۱۴] (شکل ۳) و در انتها زیرکلمات به دست آمده از مرحله قطعه‌بندی به صورت جداگانه به مرحله استخراج ویژگی فرستاده شده تا بردار ویژگی مربوط به هر زیرکلمه محاسبه شود.

۲-۲ استخراج ویژگی

یکی از مراحل مهم در یک سیستم بازشناسی، مرحله استخراج ویژگی است زیرا این مرحله به طور مستقیم روی نرخ بازشناسی سیستم طراحی شده اثر می‌گذارد. در این مقاله از خودهمبستگی محلی مرتبه بالای تصویر قطبی- لگاریتمی برای استخراج ویژگی از زیرکلمات فارسی استفاده می‌شود. تغییر مقیاس‌های خطی و چرخش به صورت شیفت در تصویر log-polar ظاهر می‌شود [۱۵] تا [۱۸] و در نتیجه می‌توان به کمک تابع خودهمبستگی محلی مرتبه بالا که یک تابع غیر حساس به شیفت است، ویژگی خاصی را از تصویر کلمه یا زیرکلمه استخراج کرد که به تغییراتی چون تغییر مقیاس و چرخش غیر حساس باشد [۱۹]. از آنجا که چرخش و تغییر مقیاس از تغییرات نگارشی رایج در دست‌نوشته‌های افراد مختلف است به کمک این شیوه استخراج ویژگی می‌توان در مقابل این تغییرات (چرخش و تغییر مقیاس) مقاومت کرد. همچنین به کمک تبدیل log-polar می‌توان از تصویر کلمه به گونه‌ای نمونه‌برداری مجدد انجام گیرد که در یک ناحیه خاص (مثلاً اطراف خط مبنا) دارای بیشترین نمونه باشیم. به عبارت دیگر به کمک این روش قادر به نمونه‌برداری غیر یکنواخت و متمرکزتر خواهیم بود [۱۸]. این روش به روش‌های رایج بر اساس طبقه‌بندی با استفاده از HMM که از تصویر کلمه به صورت



شکل ۱: ساختار سیستم بازشناسی، (الف) بخش آموزش و (ب) بخش تست.

تصویر log-polar ظاهر می‌شود، می‌توان به کمک تابع خودهمبستگی^۱ که یک تابع غیر حساس به شیفت است، سیستم بازشناسی را در مقابله با تغییراتی همچون تغییر مقیاس‌های خطی و چرخش مقاوم کرد. در این سیستم بازشناسی کلمات فارسی از مدل مخفی مارکوف به عنوان طبقه‌بند استفاده شده است، به این صورت که برای هر زیرکلمه یک مدل تعریف می‌شود. مدل‌های زیرکلمات با یک شبکه واژه‌نامه خاص به یکدیگر متصل شده و استفاده از شبکه واژه‌نامه امنیت و دقت سیستم بازشناسی را افزایش می‌دهد. انتهای هر مسیر در شبکه واژه‌نامه به یک کلمه منحصر به فرد ختم می‌شود و مجموع این مسیرها در شبکه، پایگاه داده ما را تشکیل می‌دهند. برای ارزیابی سیستم از پایگاه داده "ایران‌شهر" استفاده شده است.

در این مقاله ابتدا راجع به معماری سیستم بازشناسی ارائه‌شده بحث می‌شود و در بخش بعد نحوه استخراج ویژگی توضیح داده می‌شود. در قسمت‌های بعدی درباره ارائه مدل برای هر زیرکلمه و آموزش مدل‌ها و شبکه واژه‌نامه صحبت می‌شود. بخش‌های آخر مربوط به ارزیابی سیستم بازشناسی و نتیجه می‌باشد.

۲- ساختار سیستم بازشناسی

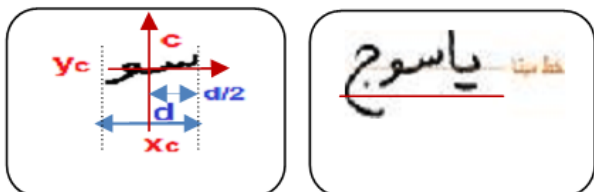
سیستم بازشناسی کلمات از دو بخش اصلی آموزش و تست تشکیل شده است (شکل ۱). در بخش آموزش پس از استخراج ویژگی از تمامی الگوهای ورودی، پایگاه داده جامعی تشکیل می‌شود. در بخش تست، پایگاه داده به دست آمده به عنوان یک پایگاه داده (شبکه واژه‌نامه) مورد استفاده قرار می‌گیرد یا به عبارتی تصویر کلمه ورودی پس از استخراج ویژگی با تمامی الگوهای موجود در پایگاه داده شبکه واژه‌نامه مقایسه می‌شود تا نزدیک‌ترین کلمه به تصویر کلمه ورودی انتخاب شود. در ادامه قسمت‌های مختلف ساختار سیستم پیشنهادی به طور کامل شرح داده می‌شود.

۱-۲ پیش‌پردازش و قطعه‌بندی

در سیستم بازشناسی ارائه‌شده در این مقاله تصویر کلمه ورودی پس

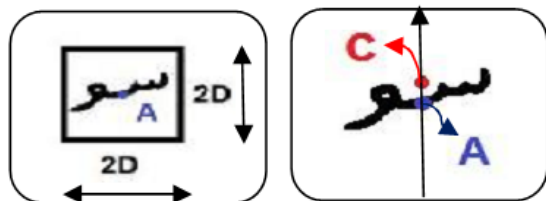
2. Average Filters
3. Coloring

1. Autocorrelation



(الف) (ب)

شکل ۶: (الف) تصویر کلمه یاسوج با خط مبنای محاسبه شده و (ب) محاسبه نقطه مناسب نمونه برداری برای زیر کلمه "سو".



(الف) (ب)

شکل ۷: (الف) تصویر زیر کلمه "سو" با نقطه C بهترین نقطه برای نمونه گیری و نقطه A مرکز چرخش و (ب) انتقال دادن نقطه A به مرکز تصویر.

محور θ ظاهر می شود [۱۵] تا [۱۹]. از این ویژگی برای حذف تغییرات نگارشی مانند تغییر مقیاس و چرخش استفاده می شود. شکل ۴ تبدیل log-polar تصویر کلمه "آبیک" را نشان می دهد که این تبدیل در دو نمودار قطبی (شکل ۴-ب) و کارترین (شکل ۴-ج) نمایش داده شده است. در این تحقیق از تصویر log-polar که در نمودار کارترین نمایش داده شده، برای استخراج ویژگی استفاده شده است. از مزایای نمایش تصویر log-polar در نمودار کارترین این است که هر گونه تغییر مقیاس خطی و چرخش در تصویر اصلی به صورت شیفت در این تصویر ظاهر می شود [۱۵] تا [۱۹].

نقطه $C = (x_c, y_c)$ در (۱) مرکز ناحیه ای می باشد که بیشترین نمونه برداری روی آن صورت می گیرد به طوری که هر چه از نقطه C دور شویم، تعداد نمونه ها یا به عبارتی دقت نمونه برداری کاهش می یابد. در واقع با تغییر x_c و y_c در (۱) می توان مرکز ناحیه ای را که در آن باید نمونه برداری با دقت بیشتری انجام شود مشخص کرد. شکل ۵ تصویر log-polar کلمه "آبیک" را با وضوح 40×40 به ازای مقادیر C های مختلف نشان می دهد.

آزمایش های متعدد نشان داد که با در نظر گرفتن عرض خط مبنا و نصف طول تصویر زیر کلمه برای x_c و y_c بالاترین نرخ بازشناسی حاصل می شود. همچنین نقطه $C = (x_c, y_c)$ در تصویر زیر کلمه باید به نحوی معرف مرکز چرخش زیر کلمه نیز باشد تا نسبت به چرخش های حول این نقطه، غیر حساس شود. به عبارت دیگر نقطه C در عین حال که مرکز نمونه برداری را تعیین می کند، باید معرف مرکز ثقل زیر کلمه به منظور غلبه بر چرخش و تغییر مقیاس نیز باشد. برای برقراری هم زمان نقطه C در دو شرط اشاره شده در بالا، کافی است قبل از گرفتن تبدیل log-polar از تصویر، مراحل زیر دنبال شود:

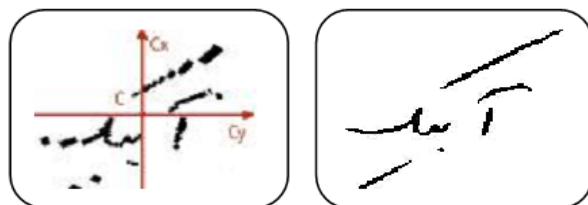
(۱) در ابتدا نقطه C با مختصات x_c و y_c به ترتیب از روی عرض خط مبنا و نصف طول تصویر زیر کلمه محاسبه می شود (شکل ۶-الف).

(۲) سپس روی خط $x = x_c$ نزدیک ترین پیکسل سیاه تصویر (نقطه $A = (x_a, y_b)$) در زیر کلمه به خط $y = y_c$ مشخص می شود (شکل ۶-ب و ۷-الف). دقت شود طول نقطه به دست آمده A برابر طول نقطه C می باشد یا به عبارتی $x_a = x_c$ است.

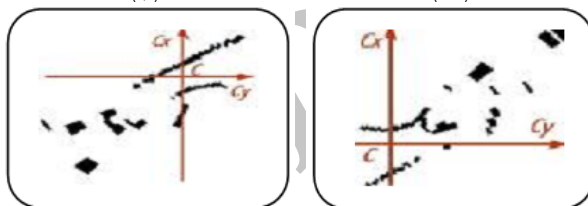


(الف) (ب) (ج)

شکل ۴: (الف) تصویر کلمه آبیک، (ب) تبدیل قطبی- لگاریتمی در نمودار قطبی و (ج) تبدیل قطبی- لگاریتمی در نمودار کارترین.



(الف) (ب)



(د) (ج)

شکل ۵: (الف) تصویر کلمه آبیک، (ب)، (ج) و (د) تصویر قطبی- لگاریتمی آن به ازای مقادیر مختلف C.

یکنواخت نمونه برداری می کند [۸] تا [۱۱] ارجحیت دارد زیرا در مورد کلمات دستنویس اکثر اطلاعات نوشتاری در اطراف خط مبنا متمرکز می باشد. در قسمت بعد نحوه گرفتن تبدیل log-polar از یک تصویر کلمه و انتقال تبدیل تصویر به مختصات کارترین شرح داده می شود.

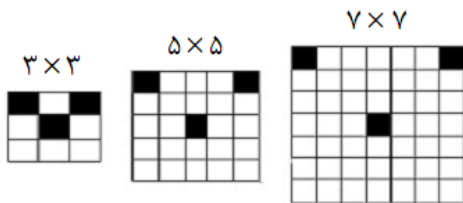
۲-۲-۱ تبدیل قطبی- لگاریتمی

در روش های معمول پردازش متن معمولاً تصویر کلمات ورودی در مختصات کارترین ارائه می شود [۹] تا [۱۱]. تبدیل log-polar چنان که ملاحظه خواهد شد، مزایای متعددی در پردازش متن دارد. برای محاسبه تبدیل log-polar ابتدا نقاط (x, y) در مختصات کارترین با استفاده از (۱) و (۲) به نقاط (ρ, θ) در مختصات قطبی تبدیل می شوند

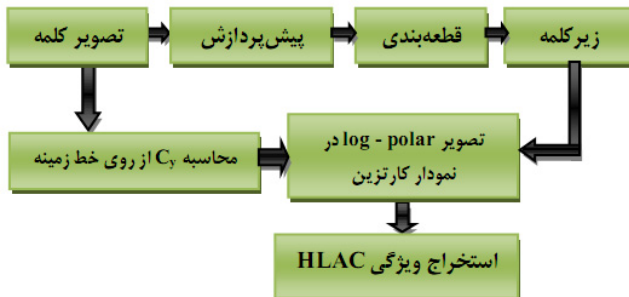
$$\rho = \sqrt{(x-x_c)^2 + (y-y_c)^2} \quad (۱)$$

$$\theta = \arctan \frac{y-y_c}{x-x_c} \quad (۲)$$

که در آن θ زاویه و ρ فاصله شعاع از مرکز (x_c, y_c) می باشد. سپس با محاسبه لگاریتم از شعاع ρ ، نقاط از روی مختصات قطبی به نقاط $(z = \log(\rho), \theta)$ در مختصات log-polar در می آیند. در این مقاله مشابه [۱۹] برای به دست آوردن تصویر log-polar از نمونه برداری مجدد و انتقال معکوس استفاده شده، به این صورت که برای به دست آوردن مقادیر پیکسل در نقطه (z_i, θ_j) روی تصویر log-polar، نقاط به صورت معکوس به نقاط $(\exp(z_i) \cos(\theta_j), \exp(z_i) \sin(\theta_j))$ در نمودار کارترین انتقال داده می شوند. سپس مقادیر نقاط (z_i, θ_j) به صورت میانگین مقادیر چگالی همسایگی نقاط انتقال داده شده در آن نقطه تخمین زده می شود. یک ویژگی مهم تصویر log-polar این است که تغییر مقیاس های خطی در تصویر اصلی به صورت شیفت در امتداد محور $z = \log(\rho)$ و چرخش در تصویر اصلی به صورت شیفت در امتداد



شکل ۹: ماسک با الگوی گسترش یافته.



شکل ۱۰: بلوک دیاگرام استخراج ویژگی.

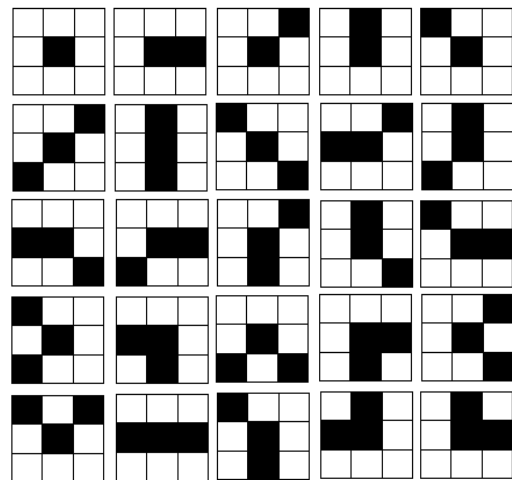
برای به دست آوردن مقادیر ویژگی HLAC، ابتدا تصویر log-polar هر زیرکلمه با ماسک ۳×۳ محلی پوش می‌شود. سپس برای هر ماسک، مجموع حاصل ضرب مقادیر پیکسل‌های متناظر با مقدار یک (یا پیکسل‌های سیاه ماسک) محاسبه می‌شود [۲۱] که ویژگی به دست آمده از این روش به ویژگی^۱ HLAC معروف است. هر مقدار ویژگی، توان طیف الگوی ماسک مربوط به خودش را نشان می‌دهد [۲۱] که مشابه تحلیل تابع در حوزه فرکانس می‌باشد، زیرا سایز ماسک‌ها در روش HLAC را می‌توان مشابه با اجزای فرکانسی در حوزه فرکانس و پراکندگی و تغییر مکان‌ها در روش HLAC را می‌توان معادل اجزای چپتی در حوزه فرکانس در نظر گرفت با این تفاوت که روش HLAC به خاطر بهره مناسب‌تر از اطلاعات دوبعدی جهت‌ها، دارای آنالیز بهتری نسبت به تبدیل فوریه می‌باشد [۲۱]. علاوه بر این در این تحقیق برای به دست آوردن ویژگی‌هایی با فرکانس‌های پایین‌تر، از ماسک‌های بزرگ‌تری با ابعاد ۵×۵ و ۷×۷ نیز استفاده شده است.

همان‌طور که در شکل ۹ نشان داده شده، ماسک‌های گسترش یافته دارای الگوهای مشابه ماسک‌های شکل ۸ هستند با این تفاوت که به دلیل ابعاد بزرگ‌ترشان شامل اطلاعاتی با فرکانس‌های پایین‌تر می‌باشند. به این ترتیب با استفاده از تبدیل log-polar و ویژگی HLAC تابعی حاصل شد که در مقابل چرخش و تغییر مقیاس مقاوم است. نحوه استخراج ویژگی از تصویر زیرکلمه در بلوک دیاگرام شکل ۱۰ نمایش داده شده است.

۳- ساختمان بردار ویژگی

در سیستم بازشناسی پیشنهاد شده، هنگامی که تعداد الگوهای آموزشی افزایش زیادی یابد، قدرت تفکیک‌پذیری ویژگی HLAC کاهش یافته و موجب تداخل کلاس‌ها می‌شود. در این تحقیق برای افزایش فاصله و قدرت تفکیک‌پذیری بین کلاس‌های مختلف، از آنالیز تفکیک خطی^۲ (LDA) استفاده شده است [۲۲]. برای این منظور ویژگی به دست آمده از روش HLAC که به صورت $x = (x_1, \dots, x_r)$ نشان داده می‌شود به کمک ترکیب‌های خطی x در غالب تابع $z = W^T x$ به بردار ویژگی

1. Higher-Order Local Autocorrelation
2. Linear Discriminant Analysis



شکل ۸: ۲۵ ماسک با الگوهای متفاوت برای استخراج ویژگی HLAC.

(۳) در این مرحله فاصله اقلیدسی بین تمام نقاط زیرکلمه و نقطه A محاسبه شده و سپس بیشترین فاصله به دست آمده را به عنوان طول D معرفی می‌کنیم.

(۴) در انتها یک تصویر زمینه سفید با طول و عرض برابر $2 \times D$ ایجاد می‌کنیم. سپس تصویر زیرکلمه را به گونه‌ای در این تصویر قرار می‌دهیم که نقطه A دقیقاً در مرکز تصویر تمام سفید قرار گیرد (شکل ۷-ب). در انتها تصویر به دست آمده با ابعاد 50×50 نرمالیزه می‌شود.

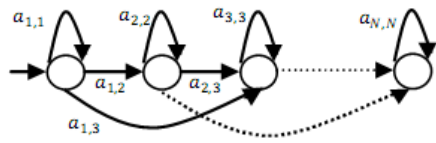
به این ترتیب به کمک مراحل ۱ تا ۴ تصویری به دست می‌آید که در آن مختصات مرکز نمونه‌برداری، همان مختصات مرکز تصویر است. در این مقاله از تصویر اصلاح‌شده به شیوه گفته‌شده در مراحل بالا، برای گرفتن تبدیل log-polar در مرحله استخراج ویژگی استفاده می‌شود. همان‌طور که در بالا به آن اشاره شد تغییراتی همچون چرخش و تغییر مقیاس به صورت شیفت در تصویر log-polar نمایان می‌شود. در نتیجه برای دستیابی به ویژگی که به چرخش و تغییر مقیاس حساس نباشد باید تابعی را بیابیم تا نسبت به شیفت به وجود آمده در تصویر log-polar غیر حساس باشد. از جمله معروف‌ترین تابع‌های غیر حساس به شیفت می‌توان به خودهمبستگی اشاره کرد [۱۹] و [۲۰] که در زیر به معرفی نوع خاصی از آن می‌پردازیم.

۲-۲-۲ خودهمبستگی محلی مرتبه بالا (HLAC)

تابع خودهمبستگی به صورت یک تابع غیر حساس به شیفت شناخته شده است [۱۹] و [۲۰]. N امین مرتبه تابع خودهمبستگی با N تغییر مکان (a_1, a_2, \dots, a_N) از نقطه مرجع r به صورت (۳) تعریف می‌شود

$$x^N(a_1, a_2, a_3, \dots, a_N) = \int I(r)I(r+a_1) \dots I(r+a_N) dr \quad (3)$$

که تابع $I(r)$ همان تصویر log-polar است که $r = (z = \log(\rho), \theta)$ است. تعداد تابع خودهمبستگی که از ترکیب‌های مختلف تغییر مکان روی تصویر مد نظر حاصل می‌شود، خیلی زیاد است. در این تحقیق مقدار N به ۲ محدود شده ($N = 0, 1, 2$) و همچنین نرخ تغییر مکان و جابه‌جا شدن در یک پنجره موضعی 3×3 محدود شده است [۱۹] چرا که میزان خودهمبستگی در یک ناحیه کوچک در تصویر زیرکلمه خیلی بیشتر از میزان خودهمبستگی بین نقاط دورتر می‌باشد. با حذف تغییر مکان‌هایی که به خاطر شیفت یکسان معادل هم بودند، تعداد الگوهای تغییر مکان به ۲۵ الگو کاهش می‌یابد. شکل ۸ الگوهای تغییر مکان محلی به کار رفته در این تحقیق را نشان می‌دهد.



شکل ۱۲: ارائه مدل HMM برای هر زیرکلمه.

به راست گسسته [۲۳] با یک حالت پریدن به جلو برای پیاده‌سازی هر مدل زیرکلمه استفاده شده است.

۴-۱ آموزش مدل‌ها

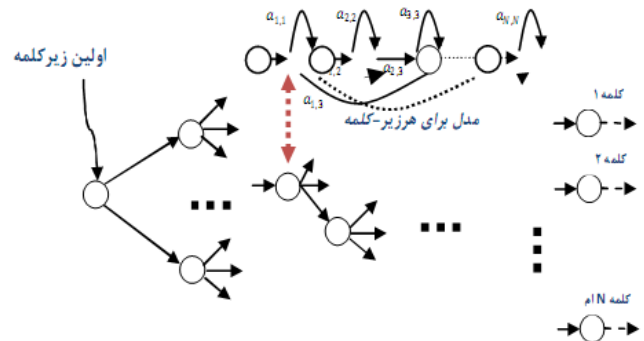
برای آموزش پارامترهای HMM از الگوریتم Baum-Welch استفاده شد [۲۳] به این معنا که برای نمایش هر مدل زیرکلمه از $\lambda = (A, B, \pi)$ استفاده شده است. توزیع حالت اولیه برای مدل‌ها به صورت $\pi = \{\pi_i\}$ می‌باشد که در آن $\pi_1 = 1$ و برای $1 < i \leq N$ ، $\pi_i = 0$ است و N تعداد حالت‌ها در مدل می‌باشد. ماتریس احتمال گذر $A = \{a_{i,j}\}$ به صورت اولیه برای بازه $1 \leq i < N$ به صورت $a_{i,j} = a_{i,i+1} = a_{i,i+2} = \dots = a_{i,i+N}$ و برای جایی که $j \neq i+1$ و $j \neq i+2$ و $j \neq i+3$ به صورت $a_{i,j} = 0$ تعریف شده و همچنین $a_{N,N}$ برابر با یک در نظر گرفته می‌شود. ماتریس مشاهده B در ابتدا طوری تعریف می‌شود تا توزیع یکنواختی را انعکاس دهد. تعداد حالت‌های انتخابی برای هر مدل بر اساس پیچیدگی هندسی شکل ظاهری آن زیرکلمه در نظر گرفته شده و تعداد حالت‌ها از ۵ تا ۱۰ متغیر است.

۵- ارائه واژه‌نامه برای مدل‌های آموزشی

برای افزایش دقت و امنیت خروجی سیستم بازشناسی و محدود کردن فضای جستجو، یک واژه‌نامه از کلمات معتبر و ممکن همانند [۲۴] مورد استفاده قرار گرفته است. واژه‌نامه کلمات فارسی به مجموعه‌ای از زیرواژه‌نامه‌های $\{D_1, D_2, \dots, D_N\}$ بر اساس تعداد زیرکلمات در کلمه‌ها تقسیم می‌شود و زیرواژه‌نامه D_k شامل همه کلماتی است که شامل K زیرکلمه می‌باشد، به طور مثال $\{ر، ضا، مسعود، مهدی\}$ D_2 سپس هر زیرواژه‌نامه را بر اساس بخش‌های زیرکلمه با نام $SWD_{k,i}$ اندیس‌گذاری می‌کنیم، مثلاً $SWD_{2,1}$ شامل $\{ر، مسعود، مهدی\}$ می‌شود. بنابراین باید تا $SWD_{k,i=k}$ زیرواژه‌نامه داشته باشیم.

زیرکلمات در یک شبکه واژه‌نامه با نام $WN_{k,i}$ که کلمات را نمایش می‌دهد تعبیه شده است که در آن مدل‌های زیرکلمه به کمک مسیرهای خاصی به هم متصل می‌شوند و هر یک از این مسیرها به یک کلمه معتبر ختم می‌شود. این شبکه با گروه کردن همه زیرکلمات مشابه بهینه شده است (شکل ۱۱). هر گره در این شبکه یک زیرکلمه را نشان می‌دهد و هر مسیر که از گره آغازین شروع و به یک گره انتهایی ختم می‌شود، متناظر با یک کلمه منحصر به فرد در D_k می‌باشد.

نحوه عملکرد سیستم بازشناسی برای جستجوی یک کلمه روی شبکه واژه‌نامه به این صورت می‌باشد [۲۴] که در ابتدا به شیوه‌ای که در مرحله استخراج ویژگی شرح داده شد، توالی مشاهده $O_s = [o_1, o_2, \dots, o_k]$ از یک کلمه دست‌نویس فارسی را محاسبه کرده که $O_i = [o_1, o_2, \dots, o_N]$ یک توالی مشاهده ساخته شده از زیرکلمه i ام دست‌نویس است. در مرحله بعد به کمک الگوریتم ویتربی [۲۳] و شبکه $WN_{k,i}$ که $1 \leq i \leq K$ می‌باشد، برای O_s داده شده، کلمه $\{SW_1, SW_2, \dots, SW_k\}$ که SW_i یک زیرکلمه i ام در W در یک زیرواژه‌نامه D_k است را با بیشترین احتمال شباهت به دست می‌آوریم (رابطه (۸))



شکل ۱۱: شبکه واژه‌نامه ارائه شده برای N کلمه.

جدیدی به نام $z = (z_1, \dots, z_i)$ تبدیل می‌شود [۲۲]. برای به دست آوردن بردار ویژگی جدید z ، ابتدا بردارهای آموزشی به صورت $\tau_{XG} = \{(x_1, g_1), \dots, (x_i, g_i)\}$ مرتب می‌شوند که در این رابطه x_i همان بردارهای آموزشی هستند که از روش HLAC حاصل شده‌اند و $g \in y = \{1, 2, \dots, c\}$ شماره کلاس (یا زیرکلمه) مربوط به بردارهای ویژگی x_i می‌باشد. سپس ماتریس پراکندگی درون کلاسی S_w و ماتریس پراکندگی میان کلاسی S_B به صورت (۴) و (۵) تعریف می‌شود

$$S_w = \sum_{g \in y} S_g, \quad S_g = \sum_{i \in \tau_g} (x_i - \mu_i)(x_i - \mu_i)^T \quad (4)$$

$$S_B = \sum_{g \in y} \tau_g |(\mu_g - \mu)(\mu_g - \mu)^T \quad (5)$$

که در این رابطه μ بردار متوسط کل و μ_g بردار متوسط کلاس‌ها از (۶) به دست می‌آید

$$\mu_g = \frac{1}{|\tau_g|} \sum_{i \in \tau_g} x_i, g \in y, \quad \mu = \frac{1}{l} \sum_{i=1}^l x_i \quad (6)$$

در انتها به کمک (۷) بهترین W برای تابع $z = W^T x$ به صورتی که معیار تفکیک‌پذیری کلاس‌ها بیشینه شود، به دست می‌آید

$$F(W) = \frac{\det(S_B)}{\det(S_W)} \quad (7)$$

وقتی که از مدل مخفی مارکوف گسسته برای طبقه‌بندی استفاده می‌شود، مقادیر ورودی باید به شکل گسسته باشند. بردار ویژگی استخراج شده به شیوه‌ای که در قسمت‌های قبل توضیح داده شد، به شکل پیوسته می‌باشد و بنابراین باید به نحوی این بردار به دست آمده به شکل گسسته تبدیل شود. برای این منظور از عمل کوانتیزه کردن استفاده می‌کنیم. در سیستم بازشناسی پیاده‌سازی شده برای هر زیرکلمه، یک بردار ویژگی (توالی مشاهده O_i) با مقدار عددی صحیح بین ۰ و ۱۲۷ در نظر گرفته شده است.

۴- ارائه مدل برای زیرکلمه‌ها

در این مقاله از HMM گسسته برای ارائه هر زیرکلمه استفاده شده است. برای افزایش دقت و امنیت سیستم بازشناسی، مدل‌های ساخته شده برای زیرکلمات در یک شبکه واژه‌نامه که کلمات معتبر را ارائه می‌دهند، جاسازی شده‌اند (شکل ۱۱). در این مقاله هر زیرکلمه به صورت یک مدل جداگانه در نظر گرفته شده است (شکل ۱۲). از مدل مخفی مارکوف چپ

1. Within-Class
2. Between-Class

جدول ۱: نرخ بازشناسی برای تصویر LOG - POLAR با وضوح‌های مختلف.

وضوح تصویر Log - Polar	۲۰×۲۰	۳۰×۳۰	۴۰×۴۰	۶۰×۶۰
نرخ بازشناسی	%۹۴٫۵	%۹۶٫۲	%۸۸٫۹	%۵۶٫۷

جدول ۲: نرخ بازشناسی روش پیشنهادی با سایر روش‌ها.

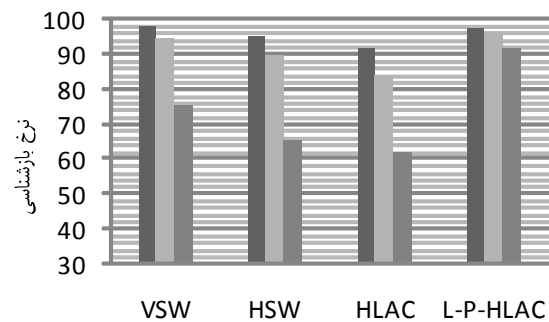
روش بازشناسی	[۲۵]	[۲۶]	[۲۷]	روش پیشنهادی
نرخ بازشناسی (%)	۹۱٫۵۵	۹۱٫۱۱	۹۲٫۲۱	۹۶٫۲

دست آمده از روش پیشنهادی، بدون استفاده از تصویر log - polar نیز در ستون سوم شکل ۱۳ با عنوان HLAC آمده است.

برای نشان دادن مقاومت سیستم بازشناسی در مقابل چرخش‌های موضعی، آزمایشی صورت گرفت که در آن زیرکلمات در تصویر کلمه به صورت تصادفی برای مقادیر $|\theta| < 30^\circ$ چرخانده می‌شوند که نتیجه این آزمایش در نمودار شکل ۱۳ رسم شده است. در این نمودار بیشترین نرخ بازشناسی برای داده‌های آموزشی، مربوط به استخراج ویژگی به روش پنجره لغزان عمودی است که دلیل آن عدم احتیاج این شیوه به مرحله قطعه‌بندی می‌باشد. همچنین بیشترین نرخ بازشناسی برای داده‌های تست مربوط به استخراج ویژگی به روش پیشنهادی می‌باشد که مؤید حساسیت کم سیستم بازشناسی نسبت به تغییرات نگارشی است. نکته دیگری که در این نمودار بارز است اختلاف نرخ بازشناسی برای داده‌های تست و آموزشی (به ویژه داده تست با چرخش موضعی) می‌باشد که برای روش پیشنهادی (L - P - HLAC) این اختلاف حداقل است. این اختلاف کم بیانگر آن است که سیستم بازشناسی به روش پیشنهادی، وابستگی کمتری نسبت به یک نویسنده خاص دارد.

۶-۱ مقایسه با سایر روش‌ها

در این مقاله، نتایج روش پیشنهادی با نتایج [۲۵] تا [۲۷] مقایسه شد و این مقایسه به دو دلیل انجام شده است. دلیل اول به علت استفاده مشترک همه این مراجع از پایگاه داده ایران‌شهر برای ارزیابی سیستم بازشناسی و دلیل دوم به روز بودن این مراجع می‌باشد. در [۲۵] برای استخراج ویژگی از کلمات فارسی از روش گرادیان استفاده شده که در برابر تغییرات مقیاس غیر حساس است و برای طبقه‌بندی کلاس‌ها از شبکه عصبی پرسپترون و کلیشه تصمیم^۳ استفاده شد. در [۲۶] یک روش طبقه‌بندی جدید برای بازشناسی کلمات فارسی بر اساس گرادیان و پایه گرادیان ارائه شده است. بردار ویژگی استخراج شده به کمک ترکیب دو شبکه عصبی چندلایه پرسپترون^۴ و تابع پایه شعاعی^۵ طبقه‌بندی می‌شود. در [۲۷] یک سیستم بازشناسی بر اساس ترکیبی از طبقه‌بندی با روش توده تصمیم‌یافته^۶ ارائه شده است. توده تصمیم‌یافته یک روش عمومی برای ترکیب سطح پایین طبقه‌بندی می‌باشد که به وسیله آن می‌توان به سطح بالاتری از طبقه‌بندی برای افزایش نرخ بازشناسی رسید. ویژگی‌های استخراج شده در این مقاله گرادیان و ناحیه‌بندی بوده است. نرخ بازشناسی روش پیشنهادی به همراه نرخ بازشناسی [۲۵] تا [۲۷] در جدول ۲ آمده است. این نتایج بیانگر آن می‌باشد که روش



شکل ۱۳: نمودار نرخ بازشناسی برای روش‌های مختلف استخراج ویژگی.

$$P(W|O_s) = \prod_{i=1}^k P(SW_i|O_i) \quad (8)$$

که $P(SW_i|O_i) = P(O_i|SW_i)P(SW_i)/P(O_i)$ است [۲۳] و [۲۴]. همچنین برای سادگی فرض می‌کنیم همه قسمت‌های کلمه در زیرواژه‌نامه با احتمال برابر اتفاق می‌افتد، بنابراین $P(O_i)$ برای همه قسمت‌های کلمه یکی است. پس مسئله به بیشینه‌نمودن $P(O_i|SW_i)$ تقلیل می‌یابد که این مقدار را هم می‌توان به طور مؤثر برای شبکه $WN_{k,i}$ با الگوریتم ویتربی محاسبه کرد [۲۳]. در نهایت برای جستجوی کلمه W در D_k به صورت (۹) اقدام می‌کنیم

$$W = \arg \max \prod_{i=1}^k P(O_i|SW_i, WN_{k,i}) \quad (9)$$

$$W = \{SW_1, SW_2, \dots, SW_k\} \in D_k$$

که W همان متن کلمه بازشناسی شده است.

۶-۲ ارزیابی سیستم

برای ارزیابی سیستم از پایگاه داده ایران‌شهر استفاده شده است [۲۵] تا [۲۷]. این پایگاه داده شامل نام ۳۰ شهر از شهرهای ایران می‌باشد که توسط ۲۶ نویسنده مختلف نوشته شده و در مجموع شامل ۷۸۰ نمونه کلمه دست‌نویس می‌باشد که از ۶۰ نمونه آن برای آموزش و ۱۸۰ نمونه آن برای تست استفاده شده است. نرخ بازشناسی به دست آمده برای وضوح‌های مختلف تصویر log - polar در مرحله استخراج ویژگی در جدول ۱ نشان داده شده است.

با توجه به نرخ بازشناسی به دست آمده در جدول ۱، وضوح بالای تصویر log - polar موجب کاهش شدید نرخ بازشناسی می‌شود چرا که با افزایش وضوح در تصویر log - polar، نمونه‌برداری (به خصوص در مرکز زیرکلمه) به صورت نمایی افزایش می‌یابد به صورتی که کمترین اختلاف بین دو زیرکلمه هم‌کلاس، به یک اختلاف خیلی بزرگ تبدیل می‌شود.

همچنین برای مقایسه، نتایج دیگر روش‌های استخراج ویژگی مانند پنجره لغزان عمودی^۱ (VSW) و افقی^۲ (HSW) نیز در نمودار شکل ۱۳ رسم شده است. برای این منظور یک پنجره لغزان عمودی (یا افقی) در طول (یا عرض) تصویر کلمه لغزانده می‌شود تا توالی مشاهده تولید شود [۸] تا [۱۱]. نرخ بازشناسی به دست آمده با روش پیشنهادی در نمودار شکل ۱۳ با نام L - P - HLAC نشان داده شده است. برای نمایش نقش بهتر تصویر log - polar در استخراج ویژگی به روش پیشنهادی، نتایج به

3. Decision Template
4. Multi Layer Perceptron
5. Radial Basis Function
6. Stack Generalization

1. Vertical Slider Window
2. Horizontal Slider Window

جدول ۳: نرخ بازشناسی برای سایز قلم‌های مختلف.

سایز قلم	۲۰	۲۶	۳۶	۴۸
نرخ بازشناسی	%۶۱	%۶۲	%۶۹	%۷۲

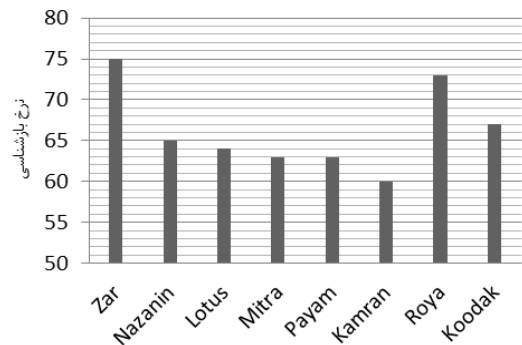
این مقاله باعث می‌شود سیستم بازشناسی بدون استفاده از روش‌های تصحیح کجی در مرحله پیش‌پردازش، در مقابل تغییرات نگارشی مانند تغییر مقیاس‌های خطی و چرخش مقاوم باشد. به دلیل شیوه استخراج ویژگی، روش ارائه‌شده نسبت به چرخش‌های موضعی زیرکلمه نیز غیر حساس است. علاوه بر این، روش به کار گرفته شده در نمونه‌گیری متمرکز و غیر یکنواخت در اطراف خط مبنا نیز به افزایش نرخ بازشناسی منجر شد.

۸- سپاس‌گزاری

با درود فراوان به روح پر فتوح استاد عزیزم دکتر خشایار یغمایی که با نکته‌های دلایز و گفته‌های بلند، صحیفه‌های سخن را علم‌پرور نمود و همواره راهنما و راهگشای اینجانب بوده است.

مراجع

- [1] S. Almaadeed, C. Higgins, and D. Elliman, "Recognition of off-line handwritten arabic words using hidden markov model approach," in *Proc. 16th Int. Conf. Pattern Recognition*, vol. 3, pp. 481-484, Feb. 2002.
- [2] S. Almaadeed, C. Higgins, and D. Elliman, "Off-line recognition of handwritten arabic words using multiple hidden markov models," *Knowledge-Based Systems*, vol. 17, no. 4, pp. 75-79, May 2004.
- [3] M. Pechwitz and V. Margner, "HMM based approach for handwritten arabic word recognition using the IFN/ENIT-database," in *ICDAR03 Proc. of the 7th Int. Conf. on Document Analysis and Recognition*, IEEE Computer Society, vol. 2, pp. 890-894, Edinburgh, Scotland, Aug. 2003.
- [4] R. Safabakhsh and P. Adibi, "Nastaaligh handwritten word recognition using a continuous - density variable-duration HMM," *The Arabian J. Science and Eng.*, vol. 30, no. 1B, pp. 95-118, Apr. 2005.
- [5] M. Mohamed and P. Gader, "Handwritten word recognition using segmentation-free hidden markov modeling and segmentation-based dynamic programming techniques," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 18, no. 5, pp. 548-554, May 1996.
- [6] M. M. Haji, Farsi Handwritten Word Recognition Using Continuous Hidden Markov Models and Structural Features, M.Sc. Thesis, Shiraz University, Shiraz, Iran, Jan. 2005.
- [7] A. Broumandnia, J. Shanbehzadeh, and M. Nourani, "Handwritten Farsi/Arabic word recognition," in *Proc. Int. Conf. on Computer Systems and Applications*, pp. 767-771, 13-16 May 2007.
- [8] S. Al-Qahtani and M. Khorshood, "An omni-font HTK-based Arabic recognition system," in *Proc. 8th IASTED Int'l Conf. Artificial Intelligence and Soft Computing*, vol. 3, pp. 317-321, Sep. 2004.
- [9] S. Al-Qahtani and M. Khorshood, "A HTK-based system to recognise arabic script," in *Proc. 4th IASTED Int. Conf. Visualization, Imaging, and Image Processing*, vol. 4, pp. 133-140, Sep. 2004.
- [10] H. Bunke, S. Bengio, and A. Vinciarelli, "Off-line recognition of unconstrained handwritten texts using HMMS and statistical language models," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 26, no.6, pp. 709-720, Jun. 2.
- [11] M. Khorshood, "Recognising handwritten arabic manuscripts using a single hidden markov model," *Pattern Recognition Letters*, vol. 24, no. 14, pp. 2235-2242, Oct. 2003.
- [12] C. G. Rafael and R. E. Woods, *Digital Image Processing*, Addison-Wesley, 1993.
- [13] I. A. Jannoud, "Automatic arabic handwritten text recognition system," *American J. of Applied Sciences*, vol. 4, no. 11, pp. 857-864, Mar. 2007.
- [14] L. Lam, S. W. Lee, and C. Y. Suan, "Thinning methodologies - a comprehensive survey," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 14, no. 9, pp. 869-885, Sep. 1992.



شکل ۱۴: نمودار نرخ بازشناسی برای کلمات تایپی با فونت‌های مختلف.

پیشنهادی ما نسبت به سایر روش‌های ذکرشده از نرخ بازشناسی بالاتری برخوردار است.

۲-۶ ارزیابی دوم

برای ارزیابی بهتر سیستم بازشناسی آزمایش دومی انجام شد که هدف از انجام آن، نشان دادن نقش شبکه واژه‌نامه در افزایش نرخ بازشناسی سیستم پیشنهادشده می‌باشد. برای این منظور پایگاه داده دومی را تشکیل دادیم که همانند پایگاه داده آزمایش اول شامل همان اسامی ۳۰ شهر ایران می‌باشد با این تفاوت که این بار با ۸ فونت مختلف نوشته شده و به عبارت دیگر کلمات تایپ‌شده جایگزین کلمات دست‌نویس شده و همچنین برای هر فونت، ۴ سایز قلم متفاوت در نظر گرفته شده و سایز قلم‌های به کار گرفته شده در این آزمایش ۲۰، ۲۶، ۳۶ و ۴۸ بوده است. به این ترتیب کل کلمات موجود در پایگاه داده دوم شامل ۹۶۰ کلمه می‌باشد که از این ۹۶۰ کلمه برای نمونه‌های تست استفاده شده است. دقت شود در آزمایش دوم سیستم بازشناسی، آموزش جدیدی نمی‌بیند و به عبارتی با همان سیستم آموزش دیده شده در آزمایش اول، آزمایش دوم انجام می‌شود. نتیجه این آزمایش در نمودار شکل ۱۴ نشان داده شده که در این نمودار نرخ بازشناسی صحیح برای هر فونت رسم شده است. با توجه به این که سیستم بازشناسی در آزمایش دوم برای هیچ فونتی آموزش ندیده بود ولی میانگین نرخ بازشناسی به دست آمده بین ۶۰ تا ۷۰ درصد می‌باشد که مؤید نقش کلیدی شبکه واژه‌نامه در سیستم بازشناسی پیشنهادی است. بررسی خطاهای صورت‌گرفته نشان می‌دهد که بیشتر خطاها مربوط به مرحله پیش‌پردازش خصوصاً مرحله حذف نقطه‌ها بوده است چرا که سبک نوشتن نقطه در بعضی از فونت‌ها به شدت با سبک دست‌نویس آن متفاوت بود. همچنین در این آزمایش نرخ بازشناسی صحیح برای هر سایز قلم به طور جداگانه محاسبه شده که نتیجه آن در جدول ۳ نشان داده شده است. با توجه به نزدیک بودن مقادیر نرخ بازشناسی صحیح برای سایز قلم‌های مختلف می‌توان نتیجه گرفت که سیستم بازشناسی پیشنهادشده، تقریباً در مقابل تغییر مقیاس‌های خطی نیز مقاوم است.

۷- نتیجه‌گیری

در این مقاله یک سیستم بازشناسی کلمات فارسی و عربی که از خودهمبستگی محلی مرتبه بالای تصویر قطبی- لگاریتمی برای استخراج ویژگی از زیرکلمات استفاده می‌کند، معرفی شد. در این سیستم بازشناسی کلمات از مدل مخفی مارکوف به عنوان طبقه‌بند استفاده شده بود به این صورت که برای هر زیرکلمه یک مدل تعریف می‌شد. مدل‌های زیرکلمات با یک شبکه واژه‌نامه خاص به یکدیگر متصل شد تا امنیت و دقت سیستم بازشناسی را افزایش دهد. شیوه استخراج ویژگی پیشنهادشده در

[27] R. Ebrahimpour, M. Amini, and F. Sharifzadehi, "Farsi handwritten recognition using combining neural networks based on stacked generalization," *Int. J. on Electrical Engineering and Informatics*, vol. 3, no. 2, pp. 146-164, Jun. 2011.

غلامرضا نادعلی نیا چاری در سال ۱۳۸۷ مدرک کارشناسی مهندسی برق- الکترونیک خود را از دانشگاه شمال و در سال ۱۳۹۰ مدرک کارشناسی ارشد مهندسی برق- مخابرات خود را از دانشگاه سمنان دریافت نمود. زمینه‌های علمی مورد علاقه نام‌برده متنوع بوده و شامل موضوعاتی مانند بازشناسی الگو، پردازش تصویر و سیستم‌های هوشمند و شبکه عصبی است.

خشایار یغمایی در سال ۱۳۶۴ با درجه کارشناسی ارشد در رشته مهندسی مخابرات از دانشگاه تهران فارغ التحصیل شد و به عنوان عضو هیئت علمی دانشگاه سمنان مشغول به کار شد. وی در سال ۱۳۷۶ مدرک دکتری خود را در رشته مخابرات در حوزه پردازش صوت از دانشگاه Surrey انگلستان دریافت نمود و سپس فعالیت خود را در دانشگاه سمنان ادامه داد. وی تا کنون استاد راهنمای ده‌ها پایان‌نامه کارشناسی ارشد بوده است. زمینه‌های تحقیقاتی مورد علاقه وی پردازش گفتار، نهم‌نگاری، پردازش خط و زبان می‌باشد.

حامد فضل‌اللهی آقاملک در سال ۱۳۸۷ مدرک کارشناسی مهندسی برق- الکترونیک خود را از دانشگاه شمال و در سال ۱۳۹۱ مدرک کارشناسی ارشد مهندسی برق- الکترونیک خود را از دانشگاه بیرجند دریافت نمود. او اکنون به عنوان مدرس دانشگاه شمال مشغول به تدریس است. زمینه‌های تحقیقاتی مورد علاقه وی عبارتند از: ترکیب طبقه‌بندها، بازشناسی الگو، انتخاب ویژگی، ادغام ویژگی، پردازش تصویر.

سید محمد رضوی در سال ۱۳۷۳ مدرک کارشناسی خود را در رشته مهندسی برق- الکترونیک دانشگاه صنعتی امیرکبیر (واحد تفرش) اخذ نمود و سال‌های ۱۳۷۵ و ۱۳۸۵ در مقاطع کارشناسی ارشد و دکتری در دانشگاه تربیت مدرس دانش آموخته شد. وی از سال ۱۳۷۶ به عنوان عضو هیات علمی گروه مهندسی الکترونیک و مخابرات دانشگاه بیرجند مشغول به فعالیتهای آموزشی و پژوهشی است. زمینه‌های تحقیقاتی مورد علاقه وی بازشناسی الگو، پردازش تصویر و سیستم‌های هوشمند است..

- [15] V. J. Traver and F. Pla, "Log - polar mapping template design: from task - level requirements to geometry parameters," *Image Vision Computing*, vol. 26, no. 10, pp. 1354-1370, Nov. 2008.
- [16] W. K. Wong, C. W. Choo, C. K. Loo, and J. P. The, "FPGA implementation of log - polar mapping," in *Proc. Intl. Conf. on Mechatronics and Machine Vision in Practice*, vol. 4, pp. 45-50, Dec. 2008.
- [17] F. Yuan, H. Zhang, and R. Jia, "Digital image stabilization based on log-polar transform," in *Proc. Int. Conf. Image and Graphics*, pp. 769-773, Aug. 2007.
- [18] S. Zokai and G. Wolberg, "Image registration using log-polar mappings for recovery of large-scale similarity and projective transformations," *IEEE Trans. Image Processing*, vol. 14, no. 10, pp. 1422-1434, Oct. 2005.
- [19] T. Kurita, K. Hotta, and T. Mishima, "Scale and rotation invariant recognition method using higher - order local autocorrelation features of Log - Polar image," in *Proc. 3rd Asian Conf. on Computer Vision*, vol. 2, pp. 89-96, Apr. 1998.
- [20] N. Otsu and T. Kurita, "A new scheme for practical, flexible, and intelligent vision systems," in *Proc. IAPR Workshop Comput. Vis.*, pp. 431-435, Oct. 1988.
- [21] S. M. Lajevardi, Z. M. Hussain, "Facial expression recognition: gabor filters versus higher - order correlators," in *Proc. Int. Conf. on Communication, Computer and Power, ICCCP'09*, pp. 354-358, Muscat, Oman, Feb. 2009.
- [22] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, John Wiley & Sons, 2nd Edition, 2001.
- [23] L. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," *IEE Proc. on Readings in Speech Recognition*, vol. 77, no. 2, pp. 257-286, Feb. 1989.
- [24] F. Biadisy, J. El-Sana, and N. Habash, "Online arabic handwriting recognition using hidden markov models," in *Proc. the 10th Int. Workshop on Frontiers of Handwriting Recognition*, vol. 1, pp. 85-90, Oct. 2006.
- [25] R. Ebrahimpour, R. D. Vahid, and B. M. Nezhad, "Decision templates with gradient based features for farsi handwritten word recognition," *International J. of Hybrid Information Technology*, vol. 4, no. 1, pp. 1-12, Jan. 2011.
- [26] R. Ebrahimpour, M. Amini, and A. Vahidi Shams, "A new combination method based on different representation of data," *Int. J. of Hybrid Information Technology*, vol. 4, no. 3, pp. 51-60, Jul. 2011.

Archive