

## Ensemble Recognition of Persian Typed Sub-word in limited Space Using Smart weighted voting

Ali Miri<sup>1</sup>, Seyyed-Mohammad Razavi<sup>2\*</sup> and Ismail Miri<sup>3</sup>

1- Department of Industrial Engineering, Sharif University of Technology, Tehran, Iran.

2\*- Faculty of Electrical and Computer Engineering, University of Birjand, Birjand, Iran.

3- Faculty of Electrical and Computer Engineering, University of Birjand, Birjand, Iran

<sup>1</sup>Ali.miri@ie.sharif.edu, <sup>2\*</sup> Smrazavi@birjand.ac.ir and <sup>3</sup> Miri.esmail@birjand.ac.ir

Corresponding author's address: Seyyed-Mohammad Razavi, Faculty of Electrical and Computer Engineering, University of Birjand, Birjand, Iran.

**Abstract-** In this paper, an ensemble method for recognition of Persian typed sub-words is proposed. First, the search space is limited to a very small number of sub-words using a few simple features. Then, by combining the six basic classifications with the weighted voting method, the sub-word is recognized. One of basic classifiers is the same as the search space limiter. Four of the basic classifiers use the nearest neighbor method with each feature of the loci, zoning, the number of the vertical cross between text and background and DCT, respectively. In another classifier, using the product of the normalized image of the input sub-word and the images of the reduced training sub-words, A degree of similarity is obtained for each training sub-words And with its help, the sub-word is recognized. The final sub-word is selected from the options obtained in a weighted voting process whose optimal weights are obtained by an intelligent algorithm. This method has been tested for lotus font and 98.34% recognition rate has been gained for this data.

**Keywords-** Ensemble, Recognition, Persian, sub-words, Typed, Voted.

## بازشناسی شورایی زیرکلمه تایپی فارسی در فضای محدود شده با روش ترکیب رای گیری وزن دار هوشمند

علی میری<sup>۱</sup>، سید محمد رضوی<sup>۲\*</sup>، اسماعیل میری<sup>۳</sup>

۱- دانشکده مهندسی صنایع، دانشگاه صنعتی شریف، تهران، ایران.

۲\* - دانشکده مهندسی برق و کامپیوتر، دانشگاه بیرجند، بیرجند، ایران.

۳- دانشکده مهندسی برق و کامپیوتر، دانشگاه بیرجند، بیرجند، ایران.

<sup>1</sup>Ali.miri@ie.sharif.edu, <sup>2\*</sup> Smrazavi@birjand.ac.ir, <sup>3</sup> Miri.esmail@birjand.ac.ir

\* نشانی نویسنده مسئول: سید محمد رضوی، بیرجند، شوکت آباد، دانشگاه بیرجند، دانشکده مهندسی برق و کامپیوتر.

چکیده- در این مقاله روشی شورایی برای بازشناسی زیرکلمات تایپی فارسی ارائه می‌شود. ابتدا فضای جستجو با استفاده از چند ویژگی ساده به تعداد خیلی کم از زیرکلمات محدود می‌شود. سپس با ترکیب شش طبقه‌بند پایه با روش رای گیری وزن دار زیرکلمه بازشناسی می‌شود. یک طبقه‌بند پایه همان محدود کننده فضای جستجو است. چهار طبقه‌بند پایه از روش نزدیکترین همسایگی و به ترتیب با ویژگی‌های مکان مشخصه، ناحیه‌بندی، تعداد تقاطع عمودی متن و زمینه و DCT استفاده می‌کنند. در یک طبقه‌بند دیگر با استفاده از حاصل ضرب تصاویر نرمالیزه زیرکلمه ورودی و زیرکلمات محدود شده آموزشی یک میزان شباهت برای هر زیرکلمه آموزشی به دست می‌آید و با آن بازشناسی را انجام می‌دهد. سپس زیرکلمه نهایی در یک فرایند رای گیری وزن دار که وزن‌های بهینه آن توسط الگوریتم هوشمند به دست می‌آیند از بین این گزینه‌ها انتخاب می‌گردد. این روش برای قلم لوگوس آزمایش شده و نرخ بازشناسی ۹۸/۳۴٪ برای داده‌های آزمون به دست آمده است.

واژه‌های کلیدی: بازشناسی، تایپی، شورایی، زیرکلمات، رای گیری، فارسی.

### ۱- مقدمه

به طور پیوسته از راست به چپ حرکت نمی‌کند بلکه به صورت گسسته از یک نقطه تمرکز به نقطه تمرکز دیگر جهش می‌کند. این تحقیقات نشان داده زمان لازم برای بازشناسی یک کلمه چهار حرفی برابر با زمان لازم برای بازشناسی یک حرف مجزا است. تحقیقات انجام شده بر اهمیت شکل کلی کلمه در فرآیند بازشناسی تاکید کرده و با توجه به این موضوع برای بکارگیری ویژگی‌های تصویری در سطح کلمه، روش‌هایی پیشنهاد شده است [۱] تا [۳].

چالش اصلی روش‌های مبتنی بر شکل کلی، دامنه وسیع جستجوی این روش‌ها با عنایت به حجم انبوه کلاس‌های در دسترس است. تعداد کلاس‌هایی که باید جستجو گردد بر اساس [۴] حداقل ۱۲۷۰۰ زیرکلمه پرتکرار استفاده شده در زبان

در زمینه بازشناسی کلمات فارسی، عربی و دستنویس انگلیسی سه رویکرد جداسازی حروف، شکل کلی کلمه (زیرکلمه) و ترکیبی از این دو استفاده شده است. اکثر روش‌های بازشناسی نوری نویسه با استفاده از رویکرد جداسازی، با شکستن کلمه به حروف تشکیل دهنده آن و بازشناسی این حروف، کلمه را بازشناسی می‌کنند. این رویکرد به دلیل مشکلات جداسازی حروف و بازشناسی دقیق آن‌ها در متون با کیفیت تصویری پایین با مشکلاتی مواجه است. بنابراین رویکرد مبتنی بر بازشناسی بدون جداسازی در چنین مواردی کارساز خواهد بود. از طرفی تحقیقات روانشناسی زیادی در مورد نحوه قرائت انسان انجام شده است که یکی از نتایج این تحقیقات این است که چشم انسان در زمان مطالعه یک خط متن،

## سه‌منابنقصا تخفیفی هفتگی یک‌کند لفتیم کلیل لمعد ینصو تکیننی لتتیجه صلا چینه بیس بگنجا لیز با کینه سفلد تیا مهذ چیف

شکل ۱: نمونه تصاویر زیرکلمات پایگاه داده

ساختار کلی سامانه پیشنهادی ارائه می‌شود. ارزیابی روش پیشنهادی و نتایج در بخش ۴ آمده و در نهایت در بخش ۵ جمع‌بندی و نتیجه‌گیری ارائه گردیده است.

### ۲- مروری بر کارهای گذشته

پیوستگی حروف در برخی خطها و در بعضی از شیوه‌های نگارش، قطعه‌بندی شکل کلمه را پیچیده‌تر می‌سازد و موجب می‌شود که محققان گرایش بیشتری به سمت روش‌های مبتنی بر شکل کلی داشته باشند. خطوط فارسی، عربی و همچنین دست‌نویس انگلیسی از این دسته هستند. تحقیقات متعددی درباره بازشناسی خطوط پیوسته انجام شده است [۱۶].

در [۱۷] روش‌های مختلف ارائه شده برای بازشناسی کلمات دست‌نویس با رویکرد مبتنی بر شکل کلی بررسی شده است. در روش‌های مبتنی بر شکل کلی، عموماً پس از استخراج ویژگی‌های زیرکلمه، این ویژگی‌ها در واژه‌نامه تصویری تشکیل شده در مرحله آموزش جستجو می‌گردد. با توجه به زیاد بودن تعداد کلاس‌ها انتخاب ویژگی‌های مناسب که امکان تفکیک به این تعداد کلاس را فراهم کند یکی از گلوگاه‌های این روش است. مقالات مختلف برای استخراج ویژگی از روش‌های متفاوتی استفاده کرده‌اند. استخراج ویژگی‌های سراسری از شکل کلمات با استفاده از انواع توصیف-گرهای ساختاری و آماری مانند افکنش‌های افقی و عمودی، تعداد و موقعیت نقاط، علائم، بالارونده‌ها، پایین‌رونده‌ها، حفره‌ها، گودی‌ها، نمایه‌های بالا و پایین شکل، تراکم پیکسل‌ها و تبدیل‌های سراسری در مقالات بررسی شده است.

در مرجع [۱۸] به مرور روش‌های بازشناسی کلمات عربی با هر دو رویکرد مبتنی بر قطعه‌بندی و مبتنی بر شکل کلی پرداخته است.

در [۱۹] روشی برای بازشناسی کلمات دست‌نویس انگلیسی بر اساس ماشین بردار پشتیبان ( $SVM^1$ ) ارائه شده است. این پژوهش از ویژگی‌های توزیع زاویه‌ای استورک که با استفاده از

فارسی خواهد بود. لذا ارائه روش‌هایی برای محدود کردن دامنه این جستجو از فعالیت‌های اصلی موجود در روش‌های مبتنی بر شکل کلی است. برای حل این مشکل معمولاً برای کاهش دامنه جستجو یک سامانه سلسله‌مراتبی به کار می‌رود [۴] تا [۶]. علاوه بر سامانه‌های بازشناسی، استفاده از اطلاعات شکل کلی کلمه در بازیابی کلمه میان مجموعه محدود کلمات، حجم پردازش را به شکل قابل ملاحظه‌ای کاهش می‌دهد. همچنین، توصیف شکل کلی کلمه، روشی کارآمد برای نشان کردن کلمات پرس و جو در تصاویر اسناد است [۷] تا [۱۳]. بدیهی است که کاهش فضای جستجو، علاوه بر کاهش حجم محاسبات مورد نیاز در مراحل بعد، دقت نهایی سامانه بازشناسی را نیز افزایش خواهد داد.

هدف از انجام این تحقیق ارائه یک روش ساده و کارا جهت بازشناسی زیرکلمات تایپی فارسی با قلم لوتوس در چند مرحله است. در اولین مرحله با استفاده از روش کاهش فضای جستجوی معرفی شده در [۱۴] فضای جستجو برای هر یک از زیرکلمات به متوسط ۸۵ زیرکلمه محدود می‌شود. در مرحله بازشناسی زیرکلمه نهایی از بین زیرکلمات پیشنهادی فضای جستجوی محدود شده، با توجه به شباهت زیاد این زیرکلمات، فرآیندی ترکیبی طراحی شده است. بدین منظور ابتدا نزدیکترین گزینه به زیرکلمه ورودی بر اساس هر یک از ویژگی‌های مکان مشخصه، ناحیه‌بندی، تعداد تقاطع عمودی متن و زمینه، ضرب، DCT و همچنین نزدیکترین گزینه بر اساس فضای محدود شده قبلی انتخاب شده سپس زیرکلمه نهایی در یک فرآیند رای‌گیری از بین این گزینه‌ها انتخاب می‌گردد.

پایگاه داده استفاده شده در این مقاله شامل مجموعه تصویر ۱۲۷۰۰ زیرکلمه رایج زبان فارسی است که با قلم لوتوس با اندازه قلم ۱۴ نگارش و چاپ شده و با درجه تفکیک ۴۰۰ نقطه در اینچ روبش شده‌اند [۱۵]. برای ایجاد نمونه‌های آزمایشی، ۱۰۰۰ زیرکلمه از میان زیرکلمات پایگاه داده به تصادف انتخاب شده است. این زیرکلمات با قلم لوتوس و در سه اندازه قلم ۱۰، ۱۲، ۱۴ چاپ شده و برای اندازه قلم‌های ۱۴ و ۱۲ در درجات تفکیک ۲۰۰ و ۳۰۰ نقطه بر اینچ و برای اندازه قلم ۱۰ در درجه تفکیک ۳۰۰ نقطه بر اینچ روبش شده و در مجموع ۵۰۰۰ نمونه آزمون را ایجاد کرده‌اند. دلیل استفاده از این پایگاه داده جامع بودن آن و امکان مقایسه با سایر پژوهش‌های مرتبط و ارزیابی منصفانه نتایج است. نمونه‌هایی از تصاویر این پایگاه داده در شکل ۱ دیده می‌شود.

در ادامه در بخش ۲ به مرور روش‌های موجود پرداخته و در بخش ۳

شده و در یک واژه‌نامه تصویری ذخیره می‌شوند. الگوریتم اول استخراج ویژگی‌های زیرکلمه بر اساس ضرایب تقریب سطح سوم تبدیل موجک هار است که بدون جداسازی زیرکلمه به حروف تشکیل دهنده آن، ویژگی‌های زیرکلمه استخراج می‌شود و یک کد رقمی برای هر زیرکلمه ساخته می‌شود. این کدها برای همه زیرکلمات محاسبه شده و در واژه‌نامه ذخیره می‌شود.

در الگوریتم دوم از تبدیل بسته موجک برای استخراج ویژگی‌های زیرکلمات استفاده شده است. در این الگوریتم ضرایب زیرباند سطح ۲۰ تبدیل بسته موجک با تابع موجک سیملت ۸ به عنوان ویژگی‌های زیرکلمه بدون جداسازی آن به حروف سازنده زیرکلمه استخراج شده‌اند. ویژگی‌های زیرکلمات به صورت یک کد رقمی برای همه زیرکلمات محاسبه شده و در واژه‌نامه ذخیره می‌شوند.

برای بازشناسی زیرکلمات چاپی فارسی و بازشناسی متون چاپی با استفاده از این ویژگی‌ها بعد از استخراج ویژگی‌های زیرکلمه ناشناس ورودی و مقایسه آن با ویژگی‌ها زیرکلمات واژه‌نامه، یک رتبه‌بندی از زیرکلمات با ویژگی‌های مشابه ویژگی‌های زیرکلمه ناشناس بدست می‌آید. در مرحله آخر برای بازشناسی نهایی زیرکلمه ورودی، کد نقاط آن محاسبه شده و با کد نقاط ۱۰ زیرکلمه آخر مرحله قبل مقایسه می‌شود. اولین زیرکلمه با کد نقطه مشابه زیرکلمه ناشناس، به عنوان زیرکلمه صحیح بازشناسی می‌شود. بهترین نرخ بازشناسی حاصل شده این تحقیق ۹۷/۶ گزارش شده است.

محققان در مرجع [۲۲] با رویکرد حل مشکل با طراحی ویژگی‌های با دقت بالا، الگوریتمی برای بازشناسی کلمات عربی بدون تقطیع آن‌ها طراحی کرده‌اند که در آن ابتدا هر کلمه به یک تصویر قطبی نرمالیزه تبدیل می‌شود و سپس یک تبدیل فوریه دو بعدی به تصویر قطبی اعمال می‌شود. طیف منتج شده نسبت به دوران، تغییر اندازه و تغییر مکان کلمات، غیر حساس است. مرحله بازشناسی با به کارگیری تابع فاصله اقلیدسی نرمالیزه شده انجام می‌گیرد. جهت آزمایش این الگوریتم، بیش از ۱۷۰۰ کلمه در ۴ قلم مختلف مورد استفاده قرار می‌گیرد. کلمات در جهت‌های مختلف (۰ تا ۱۸۰ درجه) و اندازه‌های متفاوت (۱۸ تا ۴۸pt) در نظر گرفته می‌شوند. نرخ خطا برای این سیستم در بهترین حالت

تبدیل آرنولد تخمین زده می‌شوند و از تبدیل هاف تبعیت می‌کنند، بهره می‌برد. در کنار این ویژگی‌ها برخی ویژگی‌های جهتی اولیه به منظور ایجاد بردار ویژگی به کار گرفته شده است. در مرجع [۲۰] در یک رویکرد ابتکاری به منظور ایجاد یک سیستم بازشناسی با نرخ بالا برای اجتناب از مشکلات جداسازی، زیرکلمات به عنوان پایه سیستم بازشناسی در نظر گرفته شده‌اند. در این تحقیق ابتدا روش‌های جدیدی در خصوص تشخیص کجی و جداسازی در سطح صفحات که دو مرحله مهم پیش‌پردازش می‌باشند معرفی شده، سپس متن به زیرکلمات تجزیه شده است. برای استخراج ویژگی یک طرح ترکیبی بین سه روش استفاده شده متداول استخراج ویژگی (تبدیل فوریه گسسته، گشتاورهای زرنیکی و گشتاورهای Hu) مورد استفاده قرار گرفته و سرانجام برای طبقه‌بندی از طبقه‌بندی‌کننده k نزدیکترین همسایگی استفاده شده است. با توجه به مشاهدات و تجربیات، قوانین زیر در روش ترکیبی مورد استفاده قرار گرفته است.

- وقتی تبدیل فوریه گسسته نتیجه غلط می‌دهد دو روش دیگر نتیجه درست می‌دهند.
- وقتی نتیجه گشتاورهای Hu و زرنیکی متفاوت است نتایج تبدیل فوریه گسسته صحیح است.

با توجه به شرایط بالا نتیجه روش ترکیبی به این ترتیب تعریف شده است که اگر نتیجه گشتاورهای Hu و زرنیکی مشابه هستند نتیجه نهایی همین است و در غیر این صورت نتیجه تبدیل فوریه گسسته به عنوان نتیجه نهایی پذیرفته می‌شود. جالب توجه است که نرخ بازشناسی تبدیل فوریه گسسته، گشتاورهای Hu و زرنیکی به تنهایی به ترتیب عبارت است از: ۷۹/۷، ۹۲/۸ و ۸۶ درصد در حالی که روش ترکیبی به نرخ ۹۷ درصد دست یافته است. در بررسی این روش کم بودن سرعت سیستم بازشناسی با عنایت به اینکه علاوه بر زمانبر بودن استخراج ویژگی‌های معرفی شده، عملاً برای هر زیرکلمه سه بار عملیات بازشناسی باید انجام گیرد (برای هر ویژگی یک بار) ایراد اصلی روش به نظر می‌رسد.

در مرجع [۲۱] دو روش برای استخراج ویژگی‌های زیرکلمات چاپی فارسی مبتنی بر شکل کلی آن‌ها ارائه داده است. در این دو روش ویژگی‌های ثابت اندازه و قلم برای همه زیرکلمات محاسبه

زیرکلمه استخراج می‌شود. هر ویژگی، ورودی یک طبقه‌بندی‌کننده با خروجی امتیازدار است. یعنی هر طبقه‌بندی‌کننده به هر کلاس یک امتیاز با توجه به زیرکلمه ورودی می‌دهد. خروجی ۱۰ طبقه‌بندی‌کننده با هم ترکیب شده و برای هر کلاس یک امتیاز نهایی حاصل می‌شود. کلاس‌ها بر اساس امتیازشان به صورت نزولی مرتب شده و تعدادی از کلاس‌ها با بیشترین امتیاز انتخاب می‌شوند. ساختار این روش در شکل ۲ دیده می‌شود. همچنین شکل ۳ نمونه‌هایی از فضای محدود شده پیشنهادی برای چند زیرکلمه نمونه ناشی از اجرای الگوریتم کاهش فضای جستجو فوق را نشان می‌دهد. (فضای محدود شده مرتبط با زیرکلمات قشنگ در کادر آبی، تعمیق در کادر قرمز، شصتمین در کادر سبز و بیفتند در کادر بنفش)

همانطور که ملاحظه می‌گردد با توجه به شباهت زیاد زیرکلمات پیشنهادی ناشی از الگوریتم محدود کننده قبلی، در مرحله بازشناسی باید از ویژگی‌هایی که قدرت تفکیک‌پذیری بالایی داشته باشند استفاده کرد. لذا ویژگی‌های مختلف مورد آزمایش قرار گرفت که استفاده از هر یک از ویژگی‌ها به تنهایی همانطور که در بخش نتایج به آن اشاره می‌گردد نتیجه مطلوبی به دست نمی‌دهد. لذا ایده استفاده از سیستم شورایی با ترکیبی از چند طبقه‌بند (در جهت ایجاد یک مرز تصمیم دقیق با توجه به شباهت زیاد زیرکلمات محدود شده) مورد توجه قرار گرفت که با توجه به نتایج، پیاده‌سازی این سیستم بهبود قابل ملاحظه‌ای در نرخ بازشناسی را نشان می‌دهد. روش‌های طراحی و قواعد ترکیب شورایی و چگونگی ایجاد مرزهای پیچیده مورد نیاز در مرجع [۲۵] مورد بحث قرار گرفته است.

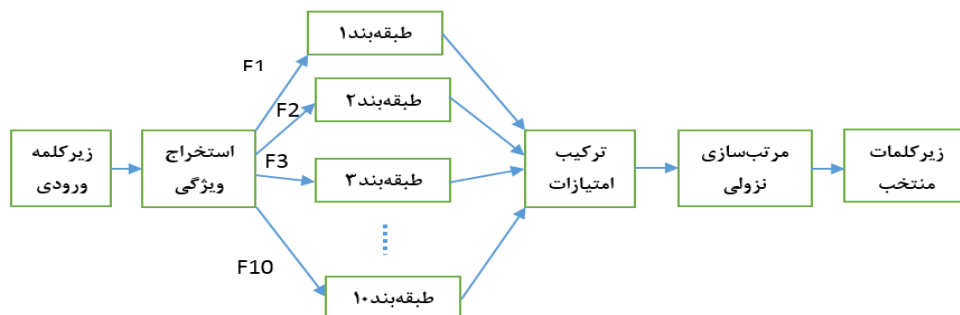
(با بکارگیری ۱۶ باند فرکانس پایینی) در حدود ۶ درصد برآورد شده است.

در مرجع [۲۳] نیز با ترکیب رویکرد سلسله مراتبی و ویژگی‌های دقیق، ابتدا ویژگی‌های مکان مشخصه شکل زیر کلمه استخراج شده و پس از کاهش بعد با استفاده از PCA با مراکز خوشه‌های سیصدگانه‌ای که قبلاً با خوشه‌بندی پایگاه داده به دست آمده‌اند مقایسه شده و ۱۰ خوشه نزدیک به آن به عنوان گزینه‌های احتمالی انتخاب می‌شوند سپس از توصیفگرهای فوریه منحنی پیرامونی زیرکلمات برای طبقه‌بندی در داخل خوشه‌ها استفاده شده است.

در مرجع [۲۴] پس از کاهش فضای ویژگی در چند مرحله، در مرحله بازشناسی ابتدا با استفاده از ویژگی‌های مکان مشخصه و معیار کمترین فاصله، دو گزینه نزدیکتر به زیرکلمه ورودی انتخاب می‌شود. همین مرحله با ویژگی‌های ناحیه‌بندی نیز تکرار می‌گردد. در مرحله نهایی بازشناسی، زیرکلمه منتخب از حداکثر چهار زیرکلمه منتخب مرحله قبل با استفاده از ویژگی‌های تعداد تقاطع متن با زمینه در جهت عمودی و طبقه‌بندی‌کننده نزدیکترین همسایگی صورت می‌گیرد. در این روش، بازشناسی زیرکلمات با دقت بیش از ۹۸٪ برای داده‌های آزمون صورت پذیرفته است.

### ۳- ساختار روش پیشنهادی

هدف در نظر گرفته شده در این تحقیق بازشناسی زیرکلمه نهایی از بین مجموعه زیرکلمات محدود شده به روش ارائه شده در [۱۴] است. در الگوریتم معرفی شده در این مرجع به منظور بالابردن سرعت و دقت بازشناسی سعی شده طی یک فرایند سلسله مراتبی فضای جستجو کاهش یابد. در این تحقیق ۱۰ ویژگی ساده از هر



شکل ۲: فرایند سلسله مراتبی کاهش فضای جستجو [۱۴]

بجنگند	نشیمین	تخمینی	تجز	خیز	خیش	شبق	شبق	ضبق	نچش	تمیز	تمیز	تفنگ	
پیگمنتو	مستقیمش	سنخیتی	شینر	تیمز	حشیش	حضبض	حنیش	حنین	خمین	خیشو	خینی	فتقتا	
بیختن	متخصصینی	ششمین	منیژ	منیز	ستیز	سنجش	شیشو	شبخو	شمیز	شیتو	شیخو	منتنگ	
بیفتند	متحصنین	تحقیقی	تحسین	تحصین	تحمیض	شینو	صفین	ضعین	قیخو	مضیق	مفیز	قتنگ	
چیفتن	شصتمین	تضمینی	حشیشی	حضبضی	حقیققر	حقیقی	نضی	نچشو	نچفی	نچمن	نیشو	فتنشا	
پیشگفتا	متخصصین	محققین	سخیفی	شبستر	شیشو	ضعیتی	ضعیفی	تحمیق	تعمیق	تعمیق	تیسفو	منتفذ	
بیفکنند	شخصیتی	تضمن	نسیتو	نسیتو	نمیتو	نمیخو	نمیشو	نیستر	خمیتر	حنیفی	خمینی	فتقتله	
	تصنیفی	تصنیفی	نیستو	نیستو	نینهو	تحسینی	مفیسو	منچستر	قیمتم	نصیحتی	ضعیمی	خفتند	
	تنسیق	تنسیق	نحیفی	نحیفی	نجمنی	ستینو	حسنین	تجتر	منجق	شیخی	ستیز	قشنگشا	
د	ج	ب	الف										

شکل ۳: فضای محدود شده برای چند زیر کلمه منتخب

انجام می‌گیرد.

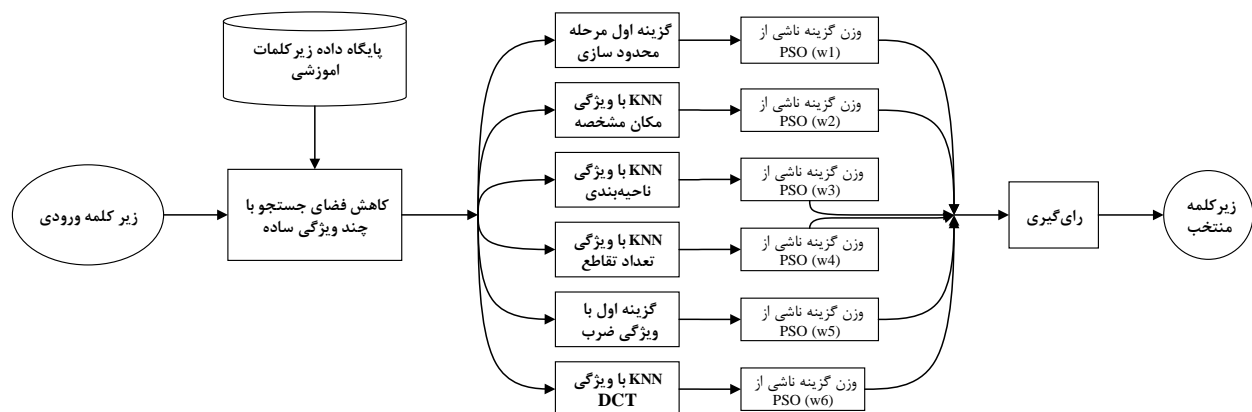
پس از آن زیر کلمه نهایی در یک فرایند رای گیری از بین این گزینه‌ها انتخاب می‌گردد. جهت بهبود بیشتر نتایج و با توجه به کیفیت متفاوت نتایج ناشی از هر کدام از ویژگی‌ها، حضور وزن دار گزینه‌ها در رای گیری مد نظر قرار گرفت. جهت تعیین وزن‌های ایده‌ال از الگوریتم PSO<sup>۲</sup> کمک گرفته شد. نتایج به دست آمده از این تحقیق در قسمت نتایج تشریح شده است. در ادامه نحوه استفاده از ویژگی‌های مورد اشاره در روش پیشنهادی تشریح شده است.

### ۳-۱- ویژگی‌های تبدیل کسینوسی گسسته

یکی از ویژگی‌های استفاده شده در این تحقیق، تبدیل کسینوسی گسسته (DCT) است که یک تبدیل یکتااست. از مهمترین خواص DCT این است که بیشترین اطلاعات سیگنال را در ضرایب معدودی که در فرکانس‌های پایین قرار دارند ذخیره می‌کند.

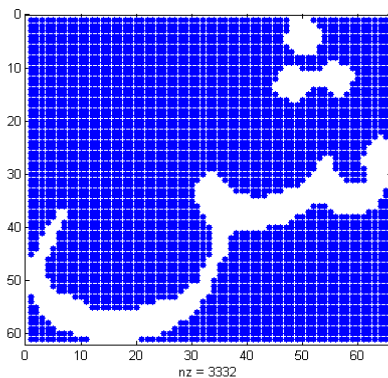
شکل ۴ ساختار روش پیشنهادی را نشان می‌دهد. همانطور که در این شکل دیده می‌شود، در روش پیشنهادی، ابتدا فضای جستجو با استفاده از چند ویژگی ساده به تعداد خیلی کم از زیر کلمات محدود می‌شود [۱۴]. سپس با ترکیب نتایج شش طبقه‌بند پایه با روش رای گیری وزن دار، زیر کلمه بازشناسی می‌شود.

یک طبقه‌بند پایه محدود کننده فضای جستجو است که کلاسی که بالاترین امتیاز در این مرحله را کسب کرده است به عنوان نتیجه این طبقه‌بند لحاظ می‌گردد. چهار طبقه‌بند پایه بعدی از روش نزدیکترین همسایگی (KNN<sup>۲</sup>) و به ترتیب با ویژگی‌های مکان مشخصه، ناحیه‌بندی، تعداد تقاطع عمودی متن و زمینه و DCT<sup>۳</sup> استفاده می‌کنند. در یک طبقه‌بند پایه دیگر با استفاده از حاصل ضرب تصاویر نرمالیزه زیر کلمه ورودی و زیر کلمات محدود شده آموزشی یک میزان شباهت برای هر زیر کلمه آموزشی به دست می‌آید و با استفاده از آن بازشناسی



شکل ۴: ساختار روش پیشنهادی

این ترتیب صورت گرفته است که ابتدا تبدیل DCT دو بعدی شکل زیرکلمه که یک ماتریس با ابعاد زیرکلمه خواهد بود محاسبه می‌گردد. سپس ماتریس  $7 \times 7$  از گوشه بالا-چپ یعنی ماتریس حاوی ستون‌های ۲:۸ و سطرهای ۲:۸ ماتریس DCT جداسازی می‌گردد، از کنار هم قرار دادن ستون‌های این ماتریس و نرمال‌سازی آن‌ها بر اساس بزرگترین مقدار ماتریس DCT که همان مقدار DC است یک بردار ویژگی ۴۹ عضوی به دست می‌آید.



شکل ۵: زیرکلمه نمونه

همینطور در مورد تصویر هم بیشترین اطلاعات تصویر در ماتریس گوشه بالا-سمت چپ ماتریس خروجی ذخیره می‌شود. می‌توان از همین اطلاعات به عنوان ویژگی استفاده کرد.

این که چند ضریب از ماتریس خروجی به عنوان ویژگی در نظر گرفته شود، سوالی است که غالباً مطرح می‌شود. در کاربردهای مختلف، معمولاً این کار به صورت شهودی انجام می‌گیرد، یعنی به ازای تعداد ضرایب مختلف کیفیت تصویر بازسازی شده را بررسی می‌کنند و در مورد تعداد ضرایب تصمیم می‌گیرند.

شکل ۵ یک زیرکلمه نمونه و جدول ۱ قسمتی از ماتریس DCT آن (ابعاد ماتریس معادل ابعاد زیرکلمه خواهد بود) و قسمتی از ماتریس ویژگی‌های DCT (ناحیه زرد رنگ) که با توجه به آزمایشات متعدد به عنوان ویژگی انتخاب شده‌اند و در نهایت جدول ۲ ویژگی‌های استخراجی از زیرکلمه که از کنار هم قرار دادن ستون‌های ماتریس منتخب نرمال شده به مقدار DC (عدد گوشه بالا-چپ (۵۲/۹) در این نمونه)) را نشان می‌دهد. همانطور که در این جداول دیده می‌شود در این تحقیق با آزمایش روش‌های مختلف در نهایت استخراج ویژگی مناسب DCT به

جدول ۱: قسمتی از ویژگی‌های DCT زیرکلمه نمونه

۵۲/۹	۱/۹	-۰/۶	-۱/۵	-۰/۳	-۱/۴	۱/۲	-۰/۲	-۱/۹	۱/۳	۰/۷	-۰/۵	-۰/۲	۰/۸	۰/۴	۰/۵
۳/۸	۶/۳	-۱/۰	-۱/۱	۳/۰	-۱/۴	۰/۵	۰/۷	-۰/۸	۰/۸	-۰/۸	-۰/۴	۰/۷	-۰/۶	-۰/۹	۰/۶
۰/۷	-۶/۳	-۰/۲	۱/۱	۳/۶	-۲/۴	-۰/۳	۱/۵	۰/۵	-۰/۳	-۰/۵	-۰/۱	۱/۴	-۱/۳	-۰/۴	۰/۷
-۰/۹	۳/۴	۱/۸	-۴/۵	-۱/۸	-۰/۴	۰/۷	-۱/۶	-۱/۴	۰/۴	۰/۸	-۱/۳	۰/۴	-۰/۱	۱/۱	-۰/۴
-۳/۸	۱/۶	-۰/۶	۲/۵	۱/۷	۱/۴	۰/۷	۰/۹	۰/۲	۱/۷	-۰/۶	۰/۹	۰/۳	-۰/۴	۰/۷	-۰/۵
۴/۴	-۲/۰	-۱/۴	۱/۱	-۳/۱	-۱/۲	-۱/۲	۱/۰	-۱/۸	۰/۰	-۱/۸	۱/۱	-۱/۰	-۱/۲	۰/۴	-۰/۶
۲/۱	-۳/۳	۱/۷	-۰/۱	۱/۳	۲/۹	-۱/۵	-۰/۱	۰/۷	۰/۸	-۰/۱	-۰/۷	۱/۲	۰/۳	۰/۶	-۰/۴
-۲/۳	۱/۰	۱/۶	-۲/۱	-۱/۰	-۱/۰	۰/۳	۰/۷	۰/۴	-۰/۶	۱/۲	-۰/۳	۱/۲	-۱/۱	۰/۴	۰/۵
۰/۵	۱/۴	-۱/۲	۱/۲	۱/۳	-۰/۵	۱/۱	-۰/۹	-۲/۰	۰/۲	-۰/۴	-۰/۸	-۱/۳	-۰/۳	۰/۱	-۰/۶
-۰/۲	-۱/۳	-۰/۶	۰/۲	-۰/۲	-۰/۹	-۰/۷	۱/۵	۱/۰	۰/۹	-۰/۴	۱/۱	۰/۸	۱/۲	-۰/۵	۰/۶
-۰/۷	۲/۴	۰/۴	-۱/۸	۱/۱	۱/۰	-۰/۷	-۰/۵	-۰/۶	-۱/۲	۰/۶	-۰/۹	-۰/۲	-۰/۴	-۰/۷	۰/۷
-۲/۵	۰/۸	-۰/۵	۰/۶	۰/۷	-۱/۵	۱/۱	۰/۲	۰/۷	۰/۲	۰/۷	-۰/۸	۰/۵	۰/۳	-۰/۴	۰/۱

جدول ۲: ویژگی‌های DCT استخراجی از زیرکلمه نمونه

۰/۱۱۹۸	-۰/۱۱۹۵	۰/۰۶۳۷	۰/۰۳۰۶	-۰/۰۳۷۶	-۰/۰۶۱۵	۰/۰۱۹۴	-۰/۰۱۸۳	-۰/۰۰۴۵	۰/۰۳۳۵
-۰/۰۱۰۴	-۰/۰۲۷۳	۰/۰۳۱۶	۰/۰۳۰۲	-۰/۰۲۱۲	۰/۰۲۱۶	-۰/۰۸۵۲	۰/۰۴۷۳	۰/۰۲۱۱	-۰/۰۰۲۳
-۰/۰۴۰۲	۰/۰۵۷۳	۰/۰۶۸۸	-۰/۰۳۴۲	۰/۰۳۲۷	-۰/۰۵۸۲	۰/۰۲۴۲	-۰/۰۱۸۳	-۰/۰۲۵۹	-۰/۰۴۵۶
-۰/۰۰۷۵	۰/۰۲۶۲	-۰/۰۲۳۲	۰/۰۵۵۱	-۰/۰۱۸۷	۰/۰۰۹۷	-۰/۰۰۵۱	۰/۰۱۲۷	۰/۰۱۲۹	-۰/۰۲۲۳
-۰/۰۲۷۴	۰/۰۰۰۶	۰/۰۱۳۲	۰/۰۲۷۵	-۰/۰۲۹۶	۰/۰۱۷۱	۰/۰۱۹۵	-۰/۰۰۰۲	۰/۰۱۳۷	

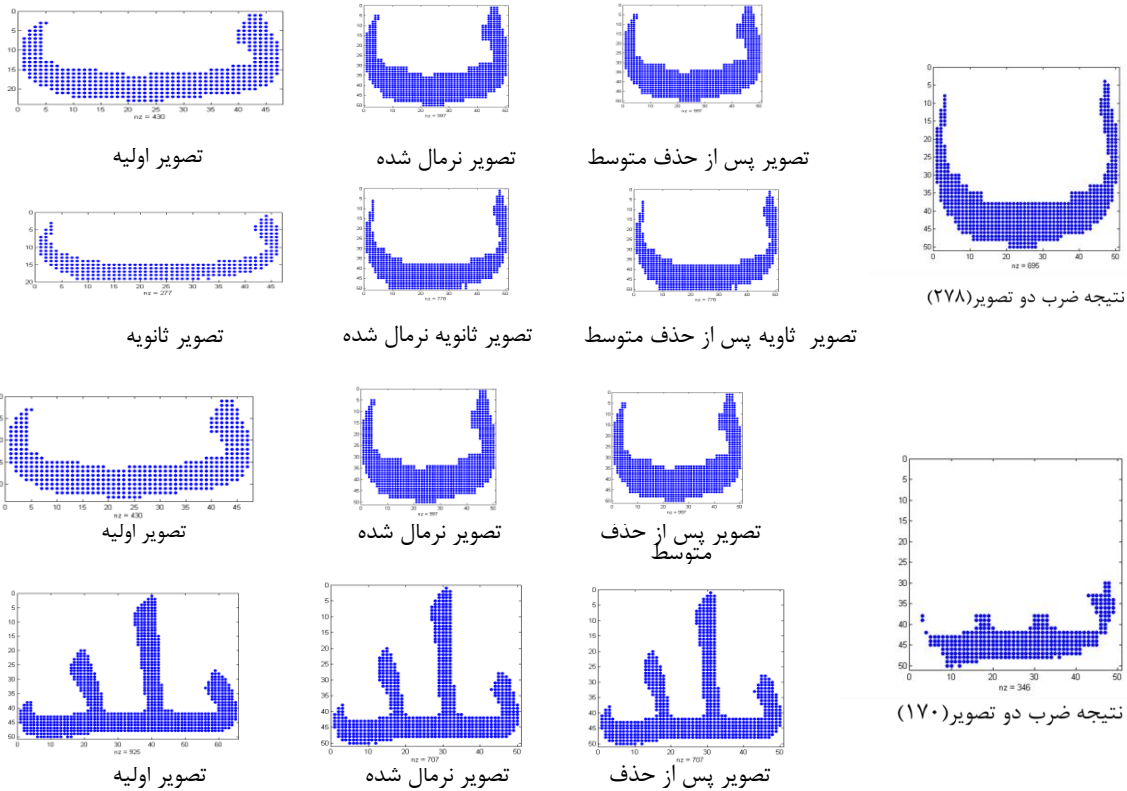
• انتخاب تصویر با حاصل بیشتر به عنوان پاسخ

شکل ۶ روال یاد شده برای یک زیرکلمه نمونه که یک بار با زیرکلمه مشابه خود با اندازه قلم و درجه تفکیک متفاوت و بار دیگر با یک زیرکلمه متفاوت مقایسه شده است را نشان می‌دهد. همانطور که در این شکل دیده می‌شود با صفر در نظر گرفتن زمینه و یک در نظر گرفتن بدنه شکل زیرکلمه (تصویر باینری شکل) در صورت نزدیک بودن دو زیرکلمه که در هم ضرب می‌شوند با شمارش تعداد یک‌ها در تصویر نتیجه به عدد بزرگتری خواهیم رسید یعنی هر چه این عدد بزرگتر باشد نشانه نزدیکتر بودن دو زیرکلمه مورد مقایسه است.

۳-۲- ضرب داخلی تصویر زیرکلمه ورودی با زیرکلمات فضای محدود شده

روال تعریف شده به منظور یافتن نزدیکترین گزینه عبارت است از:

- استانداردسازی ابعاد تصویر
- کم کردن تصویر از متوسط آن
- ضرب تصویر پیش‌پردازش شده ورودی در تصاویر پیش‌پردازش شده فضای محدود شده
- جمع مقادیر حاصل

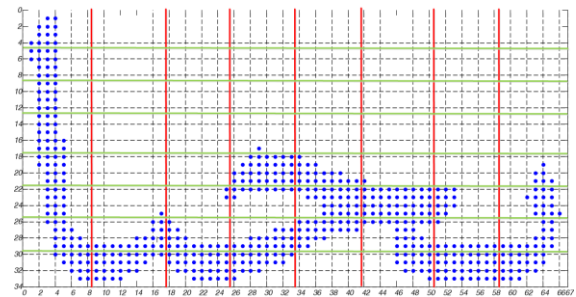


شکل ۶: فرایند ضرب دو زیرکلمه نمونه



### ۳-۳- استخراج ویژگی ناحیه‌بندی

شکل ۷ نحوه استخراج ویژگی ناحیه‌بندی در این تحقیق برای یک زیرکلمه نمونه را نشان می‌دهد. همانطور که در این شکل دیده می‌شود به منظور استخراج ویژگی‌های ناحیه‌بندی در هر جهت شکل به هشت قسمت در حد امکان مساوی تقسیم می‌گردد (رند شدن حاصل تقسیم‌ها)، در هر بلوک تعداد یک‌ها (خانه‌های اشغال شده توسط زیرکلمه) شمرده می‌شود و جهت نرمال‌سازی بر کل تعداد خانه‌های بلوک تقسیم می‌گردد. بدین ترتیب ۶۴ ویژگی به ازای هر زیرکلمه به دست می‌آید.



شکل ۷: استخراج ویژگی‌های ناحیه‌بندی

### ۳-۴- ویژگی‌های مکان مشخصه

ویژگی‌های مکان مشخصه معمولاً در راستاهای عمودی، افقی و یا زوایای ۴۵ و ۱۳۵ درجه تعریف می‌شوند. بردار ویژگی مکان مشخصه برای هر تصویر به این ترتیب محاسبه می‌گردد که به ازای هر نقطه از زمینه تصویر، یک عدد نسبت داده می‌شود. این عدد با توجه به اینکه خطوط عمودی و افقی رسم شده از آن نقطه در جهت‌های چهارگانه بالا، پایین، راست و چپ، بدنه زیرکلمه را در چند نقطه قطع کنند، محاسبه می‌شود.

اگر تعداد تقاطع‌ها با بدنه رقم در هر جهت به ۳ محدود شود، یک عدد چهار رقمی در مبنای ۴ برای نمایش نقاط مختلف نواحی زمینه به دست می‌آید. برای نمایش مکان‌های مشخصه از

معادل مبنای ۱۰ این اعداد استفاده می‌شود. لذا بردار ویژگی مکان‌های مشخصه در این حالت دارای ۲۵۶ مؤلفه است که هر کدام فراوانی نسبی عدد مربوط به خود در تصویر ورودی را نشان می‌دهند. برای آن که عناصر بردار ویژگی مقادیری بین صفر تا یک داشته باشند، مقادیر بردار ویژگی بر تعداد نقاط سفید زمینه تقسیم می‌شوند. (نرمالیزه می‌شود)

شکل ۸ نحوه محاسبه بردار ویژگی مکان‌های مشخصه را برای یک نقطه از تصاویر دو زیرکلمه نشان می‌دهد.

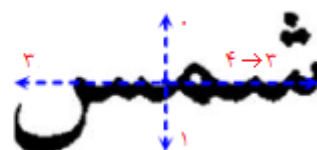
### ۳-۵- استخراج ویژگی تعداد تقاطع عمودی متن با زمینه

شکل ۹ نحوه استخراج این ویژگی مربوط به یک زیرکلمه را نشان می‌دهد. در این شکل محل تقاطع متن با زمینه با نقاط سفید مشخص شده که با شمارش این نقاط در هر ستون ویژگی‌ها به دست می‌آیند با توجه به اینکه پهنای همه زیرکلمات یکسان نیست بردار ویژگی به دست آمده باید از نظر تعداد عناصر نرمالیزه شود. برای یکسان کردن تعداد ویژگی در این ویژگی، در صورتی که تعداد ویژگی N بوده اما نیاز به M ویژگی باشد، بدین صورت عمل می‌شود که ابتدا M عدد حقیقی تولید می‌گردد که اولین آن‌ها برابر ۱ و آخرین آن‌ها برابر N است. سپس بین ۱ تا N به M-۱ بازه مساوی تقسیم شده و M-۲ عدد حقیقی دیگر به دست می‌آید. اعداد حاصله را گرد کرده تا شماره عناصری را که باید از N عنصر اولیه انتخاب گردد، به دست آید.

در صورتی که N بزرگتر یا مساوی M باشد شماره‌های عناصر تکراری نخواهد بود ولی اگر N کوچکتر یا مساوی M باشد شماره‌های عناصر تکراری نیز در اختیار بوده که نمونه‌ها چند بار انتخاب می‌شود. با توجه نتایج تحقیق [۲۴] تعداد ویژگی تقاطع متن با زمینه در این تحقیق ۱۱۸ ویژگی لحاظ شده است.



$$(2001)_4 = 1*1 + 0*4 + 0*4^2 + 2*4^3 = 129$$

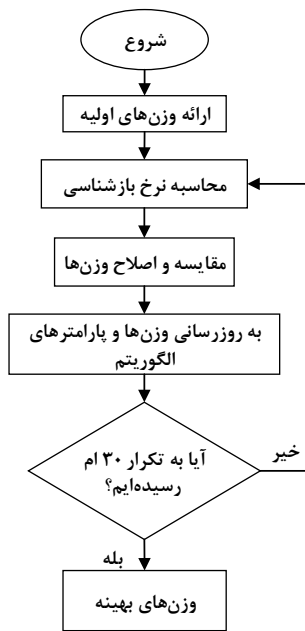


$$(1303)_4 = 3*1 + 0*4 + 3*4^2 + 1*4^3 = 115$$

شکل ۸: استخراج ویژگی‌های مکان مشخصه



شکل ۹: استخراج ویژگی تعداد تقاطع عمودی متن با زمینه



شکل ۱۱: دیاگرام نحوه عملکرد الگوریتم PSO

#### ۴- آزمایش‌ها و تحلیل نتایج

پایگاه داده استفاده شده در این تحقیق مجموعه ۱۲۷۰۰ زیرکلمه رایج زبان فارسی است که با قلم لوتوس با اندازه قلم ۱۴ نگارش و چاپ شده و با درجه تفکیک ۴۰۰ نقطه در اینچ رویش شده‌اند و برای ایجاد نمونه‌های آزمایشی، ۱۰۰۰ زیرکلمه از میان زیرکلمات پایگاه داده به تصادف انتخاب شده است. این زیرکلمات با قلم لوتوس و در سه اندازه قلم ۱۰، ۱۲، ۱۴ چاپ و با درجات تفکیک ۲۰۰ و ۳۰۰ نقطه بر اینچ رویش شده و ۵۰۰۰ نمونه را شامل شده‌اند. از این ۵۰۰۰ نمونه جهت آزمون استفاده شده است. همانطور که تشریح شد، ابتدا فضای جستجو با استفاده از چند ویژگی ساده به تعداد خیلی کم از زیرکلمات محدود می‌شود. در مرحله بازشناسی ابتدا با استفاده از هر یک از ویژگی‌های مکان مشخصه، ناحیه‌بندی، تعداد تقاطع عمودی متن و زمینه، DCT

#### ۳-۶ الگوریتم بهینه‌سازی جمعیت ذرات

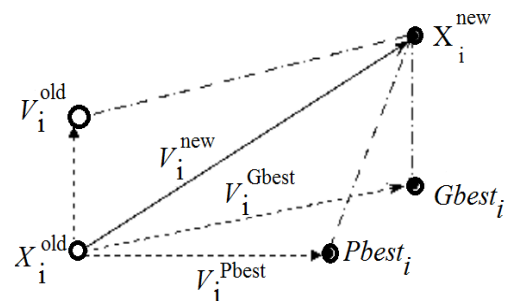
روش بهینه‌سازی جمعیت ذرات (PSO) در سال ۱۹۹۵ توسط کندی و ابرهات مطرح شده است [۲۶] روابط (۱) و (۲) به ترتیب تغییرات سرعت و مکان ذرات در هر تکرار را نشان می‌دهند.

در اینجا،  $V$  سرعت ذره،  $X_{old}$  مکان (موقعیت) کنونی یا به عبارتی قبلی (قبل از بروزسانی) هر ذره،  $P_{Gbest}$  و  $P_{Pbest}$  به ترتیب مکان‌های بهترین تجربه شخصی و بهترین تجربه کلی، ضریب  $r$  یک عدد تصادفی بین ۰ تا ۱ و  $c_1$  و  $c_2$  به فاکتورهای آموزش معروفند. همچنین در رابطه (۳) وزن اینرسی،  $W_{max}$  و  $W_{min}$  به ترتیب مقادیر حداکثر و حداقل وزن اینرسی،  $Max\_It$  حداکثر تکرار الگوریتم و  $t$  بیانگر تعداد تکرارهای حال حاضر الگوریتم می‌باشند. یک جابه‌جایی از ذرات به صورت نمونه در شکل ۱۰ نشان داده شده است.

$$V_{new} = w \times V_{old} + c_1 \times r \times (P_{pbest} - X_{old}) + c_2 \times r \times (P_{Gbest} - X_{old}) \quad (1)$$

$$X_{new} = X_{old} + V_{new} \quad (2)$$

$$w = w_{max} - \left( \frac{w_{max} - w_{min}}{Max\_It} \right) \times t \quad (3)$$



شکل ۱۰: نمونه‌ای از جابجایی ذره در الگوریتم PSO [۲۶]

شکل ۱۱ دیاگرام نحوه عملکرد الگوریتم PSO برای یافتن وزن‌های بهینه در این تحقیق را نشان می‌دهد.

در مرحله بعد، با عنایت به دقت متفاوت هر کدام از طبقه‌بندها رای‌گیری وزن‌دار این گزینه‌ها مورد آزمایش قرار گرفت. فرایند (یافتن وزن‌های بهینه) توسط الگوریتم PSO با در نظر گرفتن  $C1=C2=2$ ،  $W=0.8$  و جمعیت اولیه ۲۰ و ۳۰ تکرار انجام شد. هدف در نظر گرفته شده در این مساله کمینه کردن تعداد خطاهای بازشناسی در ۲۵۰۰ نمونه اول است که الگوریتم PSO این تعداد را کمینه می‌کند. الگوریتم بهینه‌سازی برای حالت سه گزینه‌ای بالا وزن‌های بهینه (۲ ۳ ۳) را پیشنهاد می‌کند که با این وزن‌ها تعداد زیرکلماتی که به درستی بازشناسی شده‌اند از ۲۴۵۰ به ۲۴۵۴ زیرکلمه افزایش می‌یابد. با اجرای الگوریتم برای کلیه گزینه‌ها یعنی رای‌گیری بین شش گزینه ۱- مکان مشخصه ۲- ناحیه‌بندی ۳- تعداد تقاطع ۴- ضرب ۵- DCT ۶- اولین گزینه پیشنهادی مرحله کاهش فضای جستجو، وزن‌های [۱۰ ۴ ۱۷ ۱۶ ۱۵] به عنوان یکی از وزن‌های بهینه پیشنهادی به دست آمده است که در نتیجه تعداد بازشناسی صحیح از ۲۴۴۵ زیرکلمه در حالت بدون وزن به ۲۵۶۲ زیرکلمه با وزن‌های بهینه افزایش یافته است. همین وزن‌ها تعداد زیرکلمه بازشناسی صحیح در ۲۵۰۰ زیرکلمه دوم را که اصلا در فرایند بهینه‌سازی شرکت نداشته‌اند را از ۲۴۲۲ زیرکلمه به ۲۴۵۶ زیرکلمه افزایش می‌دهد. شکل ۱۱ روند کمینه‌سازی تعداد خطاها با تغییر وزن ویژگی‌ها با استفاده از الگوریتم PSO در چهار اجرای نمونه را نشان می‌دهد و جدول ۵ مجموعه نتایج قید شده در بالا را نشان می‌دهد. همانطور که در نتایج نهایی دیده می‌شود روش پیشنهادی در عین سادگی نتایج نسبتاً مناسبی نیز در بر دارد.

جدول ۴: دقت بازشناسی با رای‌گیری از ترکیب‌های مختلف

نوع ترکیب	T3_1,2	T3_1,3	T3_1,4	T3_1,5	T3_2,3	T3_2,4	T3_2,5	T3_3,4
بازشناسی صحیح	۹۷٪/۴۸	۹۶٪/۴	۹۷٪/۴	۹۶٪/۴	۹۷٪/۱۱۶	۹۸٪	۹۷٪/۴۴	۹۷٪/۳۲
نوع ترکیب	T3_3,5	T3_4,5	T4_1	T4_2	T4_3	T4_4	T4_5	T5
بازشناسی صحیح	۹۶٪/۶۴	۹۷٪/۳۶	۹۷٪/۵۲	۹۷٪/۱۸۸	۹۷٪/۳۶	۹۸٪	۹۷٪/۴۴	۹۸٪/۱۰۸

جدول ۵: نتایج بازشناسی با رای‌گیری بین گزینه اول چند ویژگی

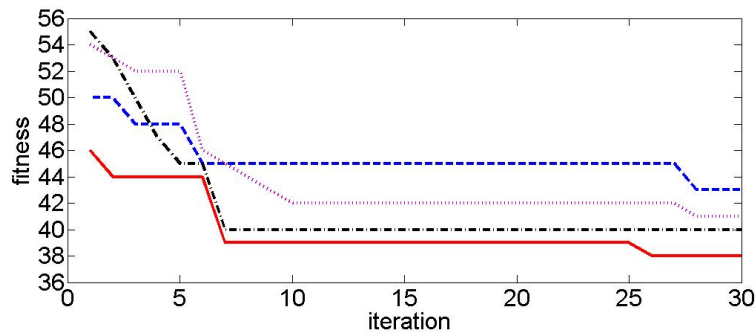
نوع داده	۲۵۰۰ داده اولیه	۲۵۰۰ داده آزمون	مجموع ۵۰۰۰ داده
نتیجه بازشناسی در حالت بدون وزن	۲۴۴۵	۲۴۲۲	۴۸۶۷
دقت بازشناسی در حالت بدون وزن	۹۷٪/۸	۹۶٪/۱۸۸	۹۷٪/۳۴
نتیجه بازشناسی در حالت وزن‌دار	۲۴۶۲	۲۴۵۵	۴۹۱۷
دقت بازشناسی در حالت وزن‌دار	۹۸٪/۴۸	۹۸٪/۲	۹۸٪/۳۴

و نتیجه ضرب تصویر ورودی در تصاویر فضای محدود شده مراحل قبل و معیار کمترین فاصله، نزدیکترین گزینه به زیرکلمه ورودی انتخاب می‌شود. جدول ۳ تعداد گزینه صحیح به دست آمده برای هر یک از این ویژگی‌ها را برای ۲۵۰۰ نمونه آزمون اول نشان می‌دهد. برای ۵۸ زیرکلمه، فضای محدود شده تنها شامل یک زیرکلمه است که برای سنجش دقت بازشناسی به نتایج حاصل از همه این ویژگی‌ها می‌بایست اضافه گردد.

جدول ۳: دقت بازشناسی با گزینه اول هریک ویژگی‌ها

نوع ویژگی	مکان مشخصه	ناحیه‌بندی	تعداد تقاطع	DCT	ضرب	انتخاب	اولین مرحله قبل
بازشناسی صحیح	۹۱٪/۱۶	۹۱٪/۴۸	۹۰٪/۱۸	۹۴٪/۴۸	۹۰٪/۷۶	۹۰٪/۵۶	۶۹٪/۵۶

در مرحله بعد، انتخاب گزینه نهایی بازشناسی با رای‌گیری از ترکیب‌های مختلف گزینه‌های بالا مورد آزمایش قرار گرفت. جدول ۴ نتیجه بازشناسی صحیح برای این ترکیب‌ها را نشان می‌دهد. در این جدول ترتیب گزینه‌ها به ترتیب ۱- مکان مشخصه ۲- ناحیه‌بندی ۳- تعداد تقاطع ۴- ضرب ۵- DCT لحاظ شده و T4-1 به مفهوم رای‌گیری از چهار تا از گزینه‌ها است که گزینه حذف شده از گزینه‌های پنج‌گانه گزینه ۱ است. بر اساس اطلاعات این جدول ترکیب T3\_2,4 که به مفهوم رای‌گیری بین گزینه‌های منتخب مکان مشخصه، تعداد تقاطع و DCT است، یکی از بهترین نتیجه‌ها (دقت ۹۸ درصد) را به دست آورده است.



شکل ۱۱: روند کمینه‌سازی تعداد خطاها با تغییر وزن ویژگی‌ها با استفاده از الگوریتم PSO

قبلا توضیح داده شد روش پیشنهادی نسبت به روش پژوهش [۲۴] که هر دو بر روی یک پایگاه داده اجرا شده‌اند، دارای کارایی زمانی بهتر و نرخ بازشناسی بالاتر است. روش ارائه شده در پژوهش [۲۰] علی‌رغم استفاده از پایگاه داده ساده‌تر دقت بازشناسی پایین‌تری دارد و روش [۲۱] که از پایگاه داده مشابه استفاده کرده به علت سرعت کم ناشی از یک مرحله‌ای بودن و دقت پایین مرحله نهایی (کد نقاط) نرخ بازشناسی پایین‌تری نسبت به روش پیشنهادی این پژوهش دارد.

جدول ۶: مقایسه نتایج با سایر پژوهش‌ها

نرخ بازشناسی نهایی	پژوهش‌ها
۹۸٪/۱۸	روش سلسله مراتبی مرجع [۲۴]
۹۷٪	روش ابتکاری مرجع [۲۰]
۹۷٪/۶	روش مرجع ویژگی‌های موجک [۲۱]
۹۸٪/۳۴	روش تحقیق جاری

جدول ۷ مجموعه زیرکلماتی که با استفاده از الگوریتم به کار رفته در تحقیق در مجموعه ۵۰۰۰ زیرکلمه داده‌های آزمون به درستی بازشناسی نشده‌اند به همراه کد زیرکلمه در پایگاه داده آزمایش را نشان می‌دهد. زیرکلمات دارای کد ۱ تا ۱۰۰۰ زیرکلماتی‌اند که با اندازه قلم ۱۰ نوشته شده و با قدرت تفکیک ۲۰۰ روبش شده‌اند، زیرکلمات دارای کد ۱۰۰۱ تا ۲۰۰۰ زیرکلماتی‌اند که با اندازه قلم ۱۰ نوشته شده و با قدرت تفکیک ۳۰۰ روبش شده‌اند، زیرکلمات دارای کد ۲۰۰۱ تا ۳۰۰۰ زیرکلماتی‌اند که با اندازه قلم ۱۲ نوشته شده و با قدرت تفکیک ۲۰۰ روبش شده‌اند، زیرکلمات دارای کد ۳۰۰۱ تا ۴۰۰۰ زیرکلماتی‌اند که با اندازه قلم ۱۲ نوشته شده و با قدرت تفکیک ۳۰۰ روبش شده‌اند، زیرکلمات دارای کد ۴۰۰۱ تا ۵۰۰۰ زیرکلماتی‌اند که با اندازه قلم ۱۴ نوشته شده و با قدرت تفکیک ۳۰۰ روبش شده‌اند، از مجموعه زیرکلماتی که به درستی بازشناسی نشده‌اند ۱۷ زیرکلمه نشان داده شده در جدول ۸ خطای انتشار یافته از مرحله قبل است.

این الگوریتم با استفاده از نرم افزار متلب (در رایانه‌ای با مشخصات (RAM:8GB , CPU:Core i7-3520M) کد نویسی شده که متوسط زمان صرف شده برای بازشناسی یک زیرکلمه از فضای محدود شده قبلی (کاهش فضای جستجو برای بازشناسی زیرکلمات با استفاده از ویژگی‌های ساده و ترکیب طبقه‌بندها) تنها در حد ۳ میلی ثانیه است. به عبارت دیگر مجموعه روش، شامل محدودسازی فضای جستجو و بازشناسی زیرکلمه نهایی به طور متوسط برای هر زیرکلمه ۱۲ میلی ثانیه زمان خواهد برد که به این عدد زمان لازم برای استخراج ویژگی‌های مورد نیاز اضافه می‌گردد.

کارایی زمانی این روش برای بازشناسی زیرکلمات (۳ میلی ثانیه برای هر زیرکلمه) در مقایسه با روش سلسله مراتبی مرجع [۲۴] (مقایسه منصفانه به دلیل پایگاه داده و سیستم سخت‌افزاری یکسان در هر دو پژوهش) که این فرایند را به طور متوسط در مدت زمان ۸ میلی ثانیه برای هر زیرکلمه انجام می‌دهد بهبود قابل ملاحظه‌ای را نشان می‌دهد و این در حالی است که نرخ بازشناسی نیز بهبود یافته است.

متوسط زمان استخراج ویژگی‌های استفاده شده به ترتیب برای ویژگی‌های مکان مشخصه، ناحیه‌بندی، تقاطع متن و زمینه، DCT، ضرب و نمایه عبارت است از: ۳۰۰، ۲/۱، ۰/۱۲، ۰/۵، ۱/۵ و ۳ میلی ثانیه است.

تعامل بین دقت و زمان بازشناسی، مشخص کننده روش انتخابی با توجه به نوع مسئله خواهد بود. به عنوان مثال، در صورت نیاز به دقت بالا، استفاده از ویژگی مکان مشخصه در ترکیب با سایر ویژگی‌های توصیف شده در فرایند رای‌گیری مفید خواهد بود؛ در حالی که اگر سرعت بازشناسی مدنظر باشد، با توجه به زمانبر بودن استخراج این ویژگی، از ترکیب سایر ویژگی‌ها باید استفاده شود.

جدول ۶ مقایسه بین دقت بازشناسی این پژوهش و چند روش ارائه شده در مرور کارهای گذشته را نشان می‌دهد. همانطور که

جدول ۷: زیرکلماتی که به درستی بازشناسی نشده‌اند

زیرکلمه	کد	زیرکلمه	کد	زیرکلمه	کد	زیرکلمه	کد	زیرکلمه	کد	زیرکلمه	کد		
'و'	۲	'تمس'	۸۸	'قشم'	۱۹۰	'پپسی'	۱۲۸۶	'جنیه'	۲۴۰۴	'منچو'	۲۶۲۰	'تلفیق'	۳۷۶۵
'پم'	۴	'تمس'	۱۰۸۸	'قشم'	۱۱۹۰	'پپسی'	۳۲۸۶	'سپشل'	۴۳۲	'مهنا'	۲۶۲۵	'ستخفا'	۷۸۶
'پم'	۱۰۰۴	'جیه'	۲۱۰۳	'قشم'	۳۱۹۰	'چنگو'	۳۰۱	'گلبن'	۳۰۸	'یستی'	۶۹۴	'سینین'	۸۰۷
'نا'	۱۰۱۴	'حسو'	۱۰۷	'قمر'	۱۹۲	'ئینگ'	۳۱۹	'نتخا'	۱۳۱۷	'یشین'	۷۰۰	'سینین'	۳۸۰۷
'حظ'	۲۰۱۶	'ختش'	۱۱۶	'قیز'	۱۹۴	'بخیر'	۳۳۲۹	'نتخا'	۳۳۱۷	'ستطا'	۱۴۳۹	'سینین'	۴۸۰۷
'فح'	۴۰۲۷	'سا'	۳۱۲۶	'قیز'	۱۱۹۴	'بلم'	۳۴۱	'طیسی'	۲۴۸۲	'طیسی'	۴۸۲	'عملیه'	۱۸۲۵
'ننس'	۳۰۶۱	'سمت'	۲۱۳۴	'کچی'	۱۱۹۵	'تعمد'	۴۳۸۳	'فتها'	۱۵۱۳	'طیسی'	۱۴۸۲	'میثمی'	۸۹۷
'ئیک'	۱۰۶۲	'سمت'	۴۱۳۴	'نتد'	۲۵۵	'تمیل'	۱۳۸۶	'فتها'	۳۵۱۳	'یقها'	۳۷۰۳	'نخجیر'	۴۹۰۳
'ئیک'	۳۰۶۲	'فقی'	۲۱۸۱	'نقه'	۳۲۶۲	'تمیل'	۴۳۸۶	'فیتی'	۵۲۴	'یقها'	۴۷۰۳	'نستیم'	۲۹۰۵
'بفو'	۷۳	'فنی'	۳۱۸۳	'نقه'	۴۲۶۲	'تیکن'	۳۹۵	'فیتی'	۱۵۲۴	'بنیه'	۷۱۶	'نستیم'	۳۹۰۵
'تیس'	۳۰۸۰	'فیض'	۳۱۸۵	'نیک'	۲۶۷	'تیکن'	۱۳۹۵	'لینی'	۵۸۶	'بینه'	۱۷۱۸	'شبختیم'	۱۱۹۵۱
'نته'	۱۰۸۱	'قسا'	۴۱۸۹	'نیک'	۳۲۶۷	'تیمی'	۳۹۶	'لینی'	۲۵۸۶	'بینه'	۳۷۱۸		

[3] A. Amin, "Recognition Of Printed Arabic Text Based On Global Features And Decision Tree Learning Techniques," Pattern Recognition, vol. 33, no. 8, pp. 1309-1323, 2000.

[۴] افشین ابراهیمی، "استفاده از شکل کلی زیرکلمات چاپی در بازیابی تصویر مستندات و بازشناسی متون فارسی"، رساله دکتری مهندسی برق- الکترونیک، دانشگاه تربیت مدرس، تهران، ۱۳۸۴.

[۵] حسین خسروی، احسان الله کبیر، "ارزیابی روش‌های بازشناسی متون فارسی بر مبنای شکل کلی زیرکلمات"، نشریه مهندسی برق و کامپیوتر ایران، جلد ۷، شماره ۴، صص. ۲۸۰-۲۶۷، ۱۳۸۸.

[6] S. Madhvanath, G. Kim, and V. Govindaraju, "Chain Code Contour Processing For Handwritten Word Recognition," IEEE Transactions on Pattern Recognition and Machine Intelligence, vol. 21, no. 9, pp. 928-932, Sep. 1999.

[7] K. Zagoris, K. Ergina, and N. Papamarkos, "A Document Image Retrieval System," Engineering Application of Artificial Intelligence, vol. 23, no. 6, pp. 872-879, 2010.

[8] S. Bai, L. Li, and C. L. Tan, "Keyword Spotting In Document Images Through Word Shape Coding," in Proc. 10th International Conference on Document Analysis and Recognition, ICDAR'09, pp. 331-335, 26-29 Jul. 2009.

[9] L. Li, S. Lu, and C. L. Tan, "A Fast Keyword-Spotting Technique," in Proc. 9th Int. Conference on Document Analysis and Recognition, ICDAR'07, pp. 68-72, 23-26 Sep. 2007.

[10] S. Lu and C. L. Tan, "Document Image Retrieval Through Word Shape Coding," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 30, no. 11, pp. 1913-1918, Nov. 2008.

[11] J. A. Rodriguez-Serrano and F. Perronnin, "Handwritten Word Spotting Using Hidden Markov Models And Vocabularies", Pattern Recognition, vol. 42, no. 9, pp. 2106-2116, Sep. 2009.

[12] T. M. Rath and R. Manmatha, "Word Spotting For Historical Documents," International Journal on Document Analysis and Recognition, Vol. 9, no. 2-4, pp. 139-152, Apr. 2007.

[13] Y. Lu and C. L. Tan, "Information Retrieval In Document Image Databases," IEEE Transactions on Knowledge and Data Engineering, Vol. 16, no. 11, pp. 1398-1410, Nov. 2004.

جدول ۸: خطاهای انتشار یافته از مرحله قبل

'شبختیم'	'بلم'	'مهنا'	'نستیم'	'سینین'	'تمس'
'نخجیر'	'میثمی'	'نستیم'	'فح'	'سینین'	'تمس'
'تعمد'	'عملیه'	'نتخا'	'قسا'	'سینین'	

## ۵- جمع بندی

در این تحقیق جهت بازشناسی زیرکلمه نهایی از فضای محدود شده، ابتدا نزدیکترین گزینه به زیرکلمه ورودی بر اساس هر یک از ویژگی‌های مکان مشخصه، ناحیه‌بندی، تعداد تقاطع عمودی متن و زمینه، ضرب، DCT و همچنین نزدیکترین گزینه بر اساس فضای محدود شده قبلی انتخاب شده سپس زیرکلمه نهایی در یک فرایند رای‌گیری از بین این گزینه‌ها انتخاب می‌گردد. با توجه به دقت متفاوتی که برای هر یک از ویژگی می‌توان متصور بود، هر یک از گزینه‌ها با وزن متفاوتی در رای‌گیری شرکت می‌کنند و وزن‌های بهینه مدنظر توسط الگوریتم بهینه‌یابی PSO پیشنهاد می‌گردند. ویژگی عمده روش استفاده شده سادگی و در عین حال کارا بودن این روش است به طوری که نتیجه نهایی دقت بازشناسی بیش از ۹۸٪ را نشان می‌دهد.

## مراجع

[1] T. Adamek, N. E. Connor, and A. F. Smeaton, "Word matching using single closed contours for indexing Handwritten Historical Documents," International Journal of Document Analysis and Recognition, vol. 9, no. 2-4, pp. 153-165, 2007.

[2] J. R. Pinales, R. J. Rivas, and M. J. C. Bleda, "Holistic Cursive Word Recognition Based On Perceptual Features," Pattern Recognition Letters, vol. 28, no. 13, pp. 1600-1609, 1 Oct. 2007.

- [۱۴] اسماعیل میری، سید محمد رضوی، ناصر مهرشاد، "کاهش فضای جستجو برای بازشناسی زیرکلمات تایپی فارسی با استفاده از ویژگی-های ساده، کوانتیزاسیون ویژگی و ترکیب طبقه‌بندها"، مجله علمی پژوهشی رایانش نرم و فناوری اطلاعات دانشگاه صنعتی نوشیروانی بابل، جلد ۹، شماره ۲، صفحه ۷۳-۶۱، ۱۳۹۹.
- [15] A. Ebrahimi and E. Kabir, "A Pictorial Dictionary For Printed Farsi Sub Words," *Pattern Recognition Letters*, Vol. 29, no. 5, pp. 656-663, 2008.
- [16] A. Rehman and T. Saba, "Off - Line Cursive Script Recognition: Current Advances, Comparisons And Remaining Problems," *Artificial Intelligence Review*, vol. 37, no. 4, pp. 261-288, 2012.
- [17] S. G. Madhvanath and V. Govindaraju, "The Role Of Holistic Paradigms In Handwritten Word Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 2, pp. 149-164, Feb. 2001.
- [18] L. M. Lorigo and V. Govindaraju, "Off - Line Arabic Handwriting Recognition: A Survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 5, pp. 712-724, May 2008.
- [19] Jija Dasgupta, Kallol Bhattacharya, Bhabatosh Chanda, "A holistic approach for Off-line handwritten cursive word recognition using directional feature based on Arnold transform", *Pattern Recognition Letters*, Vol 79, pp. 73-79, Aug. 2016
- [20] M. Shafii, "Optical Character Recognition of Printed Persian/Arabic Documents", Ph.D. dissertation, Windsor Univ., Ontario, Canada, 2014.
- [21] S. Nasrollah and A. Ebrahimi, "Printed persian subword Recognition Using Wavelet Packet Descriptors", *Journal of Engineering (Hindawi Publishing Corporation)*, 2013
- [22] M. S. Khorsheed and W. F. Clocksin, "Multi-Font Arabic Word Recognition Using Spectral Features", *Proc. 16th Pattern Recognition Int. Conf.*, 2000.
- [۲۳] افشین ابراهیمی، احسان الله کبیر، "یک روش دو مرحله‌ای برای بازشناسی زیرکلمات چاپی"، نشریه مهندسی برق و مهندسی کامپیوتر ایران، سال ۲، شماره ۲، ۱۳۸۳.
- [24] E. Miri, S.M. Razavi, N. Mehrshad, "Recognition Of The Persian Typed Sub-Words With A Hierarchical Manner," *Journal of Engineering and Applied Sciences* 12 (8): 2009-2017, 2017.
- [۲۵] محمدعلی باقری، غلامعلی منتظر، احسان الله کبیر، "سیستم‌های دسته‌بند چندگانه: روش‌های طراحی و قواعد ترکیب شورا"، دو فصلنامه پردازش علائم و داده‌ها، شماره ۲، پیاپی ۱۶، سال ۱۳۹۰
- [26] J. Kennedy and R. Eberhart, "Particle Swarm Optimization", *4th IEEE Neural Networks Int. Conf.*, 1995, pp. 1942-1948. doi:10.1109/ICNN.1995.488968.

## باورقی‌ها:

- 
- Support Vector Machine 1
  - K- Nearest Neighbors 2
  - Discrete Cosine Transform (DCT) 3
  - PSO: Particle Swarm Optimization ۴