

کاهش ویژگی توسط توده ذرات دودویی برای بازشناسی ارقام دستنویس فارسی

محسن صدیقی ناو^۱، علی سلیمانی ایوری^۲، حسین خسروی^۳

۱- دانشجوی کارشناسی ارشد برق، دانشکده مهندسی برق و رباتیک- دانشگاه صنعتی شاهرود- شاهرود- ایران

Mohsen.sedighi_nav@yahoo.com

۲- عضو هیأت علمی، دانشکده مهندسی برق و رباتیک- دانشگاه صنعتی شاهرود- شاهرود- ایران

solimani_ali@shahroodut.ac.ir

۳- عضو هیأت علمی، دانشکده مهندسی برق و رباتیک- دانشگاه صنعتی شاهرود- شاهرود- ایران

hosseinkhosravi@gmail.com

چکیده: بازشناسی ارقام دستنویس یکی از مسائل مهم در حوزه شناسی الگو است. در این مقاله با ترکیب روشهای هیستوگرام گرادیان و مکان مشخصه توسعه یافته، ویژگیهای تصاویر ارقام دستنویس فارسی استخراج شده است. توسط الگوریتم بهینه سازی توده ذرات دودویی بهبود یافته جدید (INBPSO) و ارائه تابع برازندگی مناسب، ویژگیها با اهمیت بیشتر انتخاب و ارقام توسط طبقه بند ماشین بردار پشتیبان (SVM) شناسایی شده‌اند. ابتدا در مرحله آموزش پس از استخراج ویژگی بر روی داده‌های آموزش با استفاده از روش پیشنهادی، الگوی مناسبی برای انتخاب ویژگی بدست می‌آید سپس در مرحله آزمون پس از استخراج ویژگی داده‌های آزمون بردار ویژگی توسط الگوی بدست آمده کاهش داده شده و عمل طبقه بندی نهایی صورت می‌گیرد. با اعمال روش ذکر شده روی پایگاه داده تصاویر ارقام دستنویس فارسی هدی، دقت بازشناسی ۹۹٫۴۰٪ بدون کاهش ویژگی و دقت بازشناسی ۹۹٫۲۸٪ با کاهش ویژگی بدست آمده است. مقایسه نتایج با کار محققان دیگر حاکی از آن است روش ارائه شده در استخراج ویژگی و انتخاب ویژگی از کارایی مناسبی برخوردار است.

واژه‌های کلیدی: بازشناسی ارقام دستنویس، هیستوگرام گرادیان، مکان مشخصه توسعه یافته، الگوریتم بهینه سازی توده ذرات دودویی بهبود یافته جدید، ماشین بردار پشتیبان

۱- مقدمه

دارند و همچنین، تفاوت در شیوه رسم آنها، ایجاد یک سیستم بازشناسی با دقت قابل قبول برای استفاده عملی با مشکلاتی مواجه است. از این رو، توسعه روشها برای بالا بردن نرخ بازشناسی در آنها ضروری است. روشهای مختلفی برای شناسایی ارقام دستنویس، به ویژه در زبان فارسی ارائه شده است که از جمله آنها می‌توان به روشهای آماری و ساختاری و روشهای مبتنی بر تبدیلات اشاره نمود. انتخاب روش استخراج ویژگی، مهمترین عامل در بازشناسی الگو مطرح است [۲]. برای بازشناسی ارقام و حروف از ویژگی‌های ناحیه‌ای [۳]، گشتاورهای هندسی [۴]، گشتاورهای زرنیکی [۵]، توصیفگرهای فوریه [۶]، پایاهای گشتاوری [۷]، هیستوگرام‌نما و ویژگی‌های مکان مشخصه

بازشناسی ارقام دستنویس کاربردهای زیادی در شناسایی کدهای درج شده بر فرم‌های مختلف و چک‌های بانکی و کدپستی دارد [۱]. استفاده از یک سیستم بازشناسی ارقام دستنویس، در عمل با چالش‌هایی مواجه است که مهمترین آنها ضرورت بالا بودن نرخ بازشناسی است. در حوزه زبان فارسی به علت شباهت زیادی که ارقام به هم

۱ تاریخ ارسال مقاله : ۱۳۹۱/۰۵/۰۷

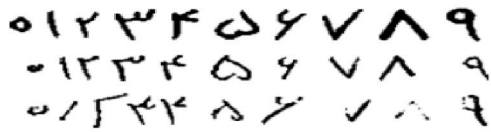
تاریخ پذیرش مقاله : ۱۳۹۲/۱۱/۰۵

نام نویسنده مسؤول : محسن صدیقی ناو

نشانی نویسنده مسؤول : ایران - شاهرود - دانشگاه صنعتی

شاهرود - دانشکده برق و رباتیک

را نشان می‌دهد.



شکل (۱): نمونه ای از اعداد پایگاه داده هدی

در این مقاله، ابتدا سیستم پیشنهادی برای شناسایی ارقام دستنویس فارسی معرفی شده است. در بخش ۳ دو روش استخراج ویژگی معرفی شده است، در بخش ۴ بهینه سازی توده ذرات معرفی شده است و در انتها در بخش ۵ نتایج به دست آمده با استفاده از روش پیشنهادی قرار داده شده است.

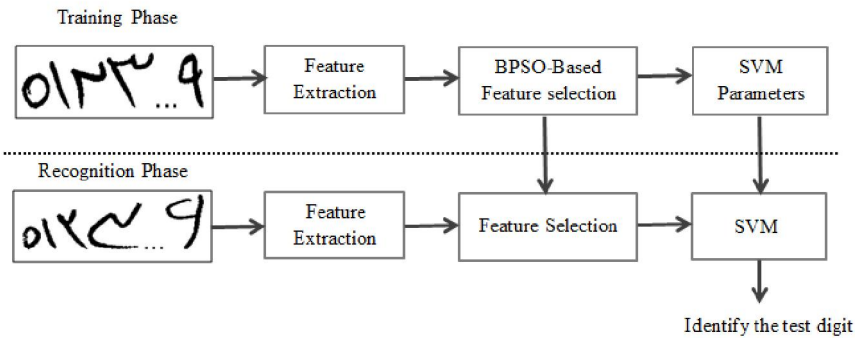
۲- معرفی سیستم پیشنهادی

در سیستم پیشنهادی برای بازشناسی ارقام دستنویس فارسی مطابق شکل (۲)، مجموعه داده‌های آزمون و آموزش به طور جداگانه در دو مرحله بررسی می‌شوند. در مرحله آموزش، ابتدا ویژگی‌های ارقام دستنویس فارسی با استفاده از ترکیب الگوریتم‌های هیستوگرام گرادیان و مکان مشخصه توسعه یافته استخراج می‌شود. سپس با استفاده از الگوریتم PSO دودویی بهبود یافته، الگویی از نحوه انتخاب ویژگی‌ها برای داشتن بهترین دقت و کمترین تعداد ویژگی به دست آمده و نیز پارامترهای مربوط به طبقه بند SVM که برای محاسبه تابع برازندگی استفاده شده، ذخیره می‌شود. برای مجموعه آزمون در مرحله آزمون، ابتدا ویژگی‌های ارقام با استفاده از ترکیب الگوریتم‌های هیستوگرام گرادیان و مکان مشخصه توسعه یافته استخراج شده، سپس با استفاده از الگویی به دست آمده در مرحله آموزش، ویژگی‌ها برای هر رقم انتخاب می‌شود. در نهایت، با استفاده از پارامترهای مربوط به طبقه بند SVM که در مرحله آموزش به دست آمده، بازشناسی انجام گرفته و درصد تشخیص برای مجموعه آزمون به دست می‌آید.

[۸] استفاده شده است. یکی دیگر از کارهای مؤثر در بالا بردن نرخ بازشناسی، تفکیک ویژگی‌هایی است که علاوه بر اینکه در بالا بردن نرخ بازشناسی نقشی ندارند، به کاهش نرخ بازشناسی نیز منجر می‌شوند. با انتخاب زیر مجموعه مناسب از ویژگی‌های استخراجی می‌توان در بالا نگه داشتن نرخ بازشناسی گام مناسبی برداشت.

الگوریتم‌های زیادی برای استخراج ویژگی بررسی و معرفی و سپس استفاده می‌شوند و حتی هم اکنون نیز در راه بهبود این الگوریتم‌ها پژوهش‌ها و مقالات زیادی در دست اجراست و هر یک به نوبه خود در جایی که استفاده می‌شوند، مزیت‌ها و محدودیت‌هایی دارند. این الگوریتم‌ها داده‌ها را در ابعاد مختلف تولید می‌نمایند. همچنین، پردازش سریع اطلاعات به عنوان چالشی برای پژوهشگران محسوب می‌شود و برای پردازش سریع و بلادرنگ به داده‌ها با ابعاد ویژگی کم نیاز است. ویژگی‌هایی نیز در بین این داده‌های استخراجی وجود دارد که اثر سوء بر نرخ بازشناسی دارند. در همین راستا، بهتر است علاوه بر پژوهش و استفاده از انواع روشهای استخراج ویژگی، روی مسأله انتخاب ویژگی نیز تمرکز کرده، با کاهش ابعاد داده سرعت و دقت شناسایی را بالا نگه داریم.

در پژوهش‌هایی که در زمینه بازشناسی ارقام دستنویس فارسی صورت گرفته، پایگاه داده استاندارد ارائه شده است. این پایگاه از ۱۰۲۳۵۲ تصاویر ارقام دودویی که از حدود ۱۱۹۴۲ از دو نوع فرم ثبت نام آزمون سراسری، که توسط دانشجویان مقاطع کارشناسی و کارشناسی ارشد پرشده، استخراج شده است. از این فرم‌ها دونوع پایگاه داده ارقام و حروف تهیه شده است که در این مقاله از پایگاه داده ارقام آن استفاده شده است. تعداد کل ارقام این پایگاه داده ۱۰۲۳۵۲ رقم است که ۶۰۰۰۰ رقم آن به عنوان داده‌های آموزش و ۲۰۰۰۰ رقم آن به عنوان نمونه‌های آزمایش استفاده شده و ۲۲۳۵۲ رقم آن هم به عنوان ارقام باقی مانده در نظر گرفته شده است [۹]. شکل (۱) تعدادی از این ارقام



شکل (۲): سیستم پیشنهادی برای بازشناسی ارقام دستنویس

۳- استخراج ویژگی

ویژگی‌های هیستوگرام گرادیان اولین بار توسط آقای خسروی و همکارش برای شناسایی ارقام دستنویس فارسی استفاده شده است. نتایج تجربی نشان از دقت و سرعت بهتر این الگوریتم نسبت به فیلتر گابور دارد [۱۰]. در این مقاله، از ترکیب الگوریتم‌های هیستوگرام گرادیان و مکان مشخصه توسعه یافته برای استخراج ویژگی از ارقام پایگاه داده استاندارد ذکر شده، استفاده شده است.

۳-۱- الگوریتم هیستوگرام گرادیان

۱- ابتدا تصویر ارقام به اندازه 64×64 نرمال شده و سپس به 16×16 بلوک غیر همپوشان با ابعاد 16×16 تقسیم شده است.
۲- برای هر پیکسل از هر بلوک، گرادیان با استفاده از عملگر سوبل که در جدول (۱) نمایش داده شده، محاسبه شده است. روش کار براساس معادله‌های (۱) تا (۳) است.

$$g(x, y) = [g_x, g_y]^T \quad \text{where} \quad (1)$$

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

$$g_y(x, y) = f(x-1, y+1) + 2f(x, y+1) + f(x+1, y+1) - f(x-1, y-1) - 2f(x, y-1) - f(x+1, y-1) \quad (3)$$

عبارت $f(x, y)$ مقدار پیکسل در مکان (x, y) است.

۳- برای هر پیکسل فاز و اندازه گرادیان با استفاده از

معادله‌های (۴) و (۵) محاسبه می‌شود.

$$P(x, y) = \tan^{-1}(g_x, g_y) \quad (4)$$

$$A(x, y) = \sqrt{(g_x^2 + g_y^2)} \quad (5)$$

۴- فاز به 16 زاویه از صفر تا $2\pi \frac{15}{16}$ کوانتیزه شده سپس برای هر بلوک 16 ویژگی مطابق با 16 فاز با استفاده از معادله (۶) استخراج شده است.

$$F_\theta = \sum_{x_\theta, y_\theta} A(x, y) \quad (6)$$

عبارت F_θ ویژگی متحد با فاز θ است. عبارت‌های x_θ و y_θ تمام مختصات‌هایی هستند که فازشان برابر θ است. در نتیجه، برای 16 چرخش به تعداد $16 \times 4 \times 4 = 256$ ویژگی با استفاده از عملگر سوبل به دست آمده و نرمالیزه شده‌اند. می‌توان تعداد چرخش‌ها را به تعداد کمتری مانند 4 یا 8 جهت نیز کوانتیزه کرد. تعداد ویژگی‌ها برای چرخش‌های مختلف مطابق جدول (۲) است.

جدول (۱): (الف) و (ب) عملگرهای سوبل

۱	۰	-۱	۱	۲	۱
۲	۰	-۲	۰	۰	۰
۱	۰	-۱	-۱	-۲	-۱

ب

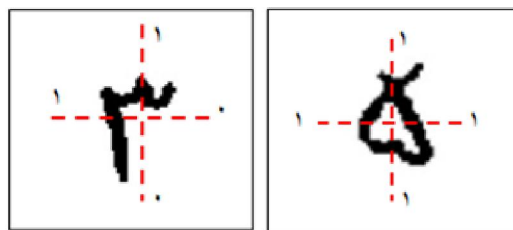
الف

جدول (۲): تعداد ویژگی‌ها برای چرخش‌های مختلف

تعداد چرخش	تعداد ویژگی‌ها
۱۶	$16 \times 4 \times 4 = 256$
۸	$8 \times 4 \times 4 = 128$
۴	$4 \times 4 \times 4 = 64$

۳-۲- الگوریتم مکان مشخصه توسعه یافته

مکان مشخصه توسعه یافته، روش خوبی برای ترکیب با سایر روش‌های استخراج ویژگی است، چون اطلاعاتی می‌دهد که سایر روش‌های استخراج ویژگی چنین اطلاعاتی در اختیار قرار نمی‌دهد. مراحل کار در این روش این گونه است که به ازای هر پیکسل در چهار جهت بالا، پایین، چپ و راست تعداد گذرها شمارش می‌شود. گاهی اوقات برای غلبه بر نویز تعداد گذرها محدود می‌گردد. همچنین، با این عمل تعداد ویژگی‌ها نیز محدود می‌شوند. سپس اعداد به دست آمده در جهت‌های مختلف را به ترتیب به صورت (راست، بالا، چپ، پایین) چیده و این رشته به عددی بر مبنای یک واحد بیشتر از ماکزیمم عدد به دست آمده در رشته بالا برده می‌شود. شکل (۳) جزئیات این روش را نمایش داده است.



شکل (۳): جزئیات روش مکان مشخصه توسعه یافته

به دست آمده به مبنای ذکر شده در دو جهت عمودی و افقی بر اساس معادله (۷) منتقل می‌شود. طبق جدول (۳) کمترین و بیشترین مقدار طول بردار ویژگی از مجموع انتقال یافته اعداد به پایه‌های مربوطه از رشته (۰۰۰۰) تا (۲۳۲۳) برابر [۰ تا ۱۴۳] است. حال طبق توضیحات داده شده باید رشته اعداد فوق به مبنای ۳ و ۴ برده شود. جدول (۳) این انتقال را نمایش داده است.

در نتیجه، طول بردار ویژگی که در این روش به دست می‌آید، برابر ۱۴۴ ویژگی است. نهایتاً طول بردار ویژگی اصلی از مجموع طول بردار ویژگی ۲ روش فوق برابر ۴۰۰ است.

جدول (۳): انتقال رشته اعداد به مبنای ۳ و ۴

رشته اعداد	تعداد ویژگی‌ها
۰۰۰۰	$0 \times (4 \times 3 \times 4 \times 1) + 0 \times (3 \times 4 \times 1)$ $+ 0 \times (4 \times 1) + 0 \times (1) = 0$
۲۳۲۳	$2 \times (4 \times 3 \times 4 \times 1) + 3 \times (3 \times 4 \times 1)$ $+ 2 \times (4 \times 1) + 3 \times (1) = 143$

۴- بهینه سازی توده ذرات

PSO یک الگوریتم محاسبه‌ای تکاملی مبتنی بر قوانین احتمال، الهام گرفته از طبیعت و بر اساس تکرار است که برای اولین بار توسط Kennedy و Eberhart در سال ۱۹۹۵ مطرح شد و یکی از الگوریتم‌های بسیار پرکاربرد در زمینه بهینه‌سازی استاتیکی و دینامیک است [۱۱]. در این سیستم، عامل‌ها به طور محلی با هم همکاری می‌نمایند و رفتار جمعیت باعث همگرایی در نقطه‌ای نزدیک به جواب بهینه سراسری می‌شود. نقطه قوت این الگوریتم بی‌نیازبودن از یک کنترل سراسری است. هرذره در این الگوریتم خودمختاری نسبی دارد که می‌تواند در سراسر فضای جواب حرکت کند و باید با سایر ذرات همکاری داشته باشد. تفاوت عمده‌ای که این روش با الگوریتم‌های دیگر چون الگوریتم ژنتیک دارد، این است که اعضای جامعه از وضعیت سایر اعضا و یا بهترین عضو جامعه باخبر هستند و نتیجه به دست آمده توسط آنها را در تصمیم‌گیری‌های خود لحاظ می‌کنند. همچنین، اعضا بهترین نتیجه خود، در طی

اگر رشته به دست آمده به صورت $(N_3 N_2 N_1 N_0)$ نشان داده شود، عدد خروجی از انتقال رشته فوق بر مبنای ۳ و ۴ به صورت معادله (۷) به دست می‌آید.

$$(N_3 N_2 N_1 N_0) = N_3(4 \times 3 \times 4 \times 1) + N_2(3 \times 4 \times 1) + N_1(4 \times 1) + N_0 \quad (7)$$

در ادامه کار باید نمودار فراوانی اعداد به دست آمده از انتقال رشته اعداد به مبنای مشخص شده را به دست آورده و به عنوان بردار ویژگی استفاده کرد. همان گونه که بیان شد در این مقاله گذرهای افقی به مبنای ۴ و گذرهای عمودی به مبنای ۳ منتقل شده است. این بدین معنی است که در جهت‌های افقی باید تعداد گذرها به ۳ و در جهت عمودی تعداد گذرها به مقدار ۲ محدود گردد. بنابراین، بیشترین تعداد بردار ویژگی زمانی رخ می‌دهد که با شمارش تعداد گذرها، رشته (۲۳۲۳) تولید شود. در مرحله بعد، این رشته

در بازه [۰,۱] هستند. همگرایی الگوریتم شدیداً به ضریب اینرسی وابسته است و به صورت دینامیک در بازه [۰,۲,۰,۸] انتخاب می‌شود که به صورت خطی در طی روند تکامل جمعیت کاهش می‌یابد و در ابتدا بزرگ است تا امکان یافتن جواب‌های خوب در همان مراحل اولیه فراهم شود و در مراحل پایانی کوچک بودن w همگرایی بهتری را سبب می‌شود. در این مقاله، این مقدار بر اساس تعداد تکرار با معادله (۹) تعیین شده است که مقدار $W_{initial}$ برابر با ۰,۹ است.

$$W(t) = W_{initial} - \frac{0.5 \times iteration}{\max iteration} \quad (9)$$

سرعت ذره‌ها در هر بعد، برای جلوگیری از واگرایی الگوریتم محدود به ماکزیمم سرعت $V_{ik} \in [-V_{max}, V_{max}]$ است که چگونگی حرکت ذره‌ها در فضای جستجو را مشخص می‌کند و اگر V_{max} خیلی کوچک باشد، ممکن است ذره‌ها بخوبی در فضاهای مناسب محلی کاوش نکنند و در بهینه محلی به دام بیفتند. از سویی دیگر، اگر این مقدار بیش از حد زیاد باشد، ممکن است ذرات در فضاهای دور از جواب حرکت کنند معادله (۱۰).

$$V_i^{t+1} = \begin{cases} V_i^{t+1} & \text{if } V_i^{t+1} \in (V_{min}, V_{max}) \\ \max(\min(V_{max}, V_i^{t+1}), V_{min}) & \text{else} \end{cases} \quad (10)$$

در نهایت، ذرات به سوی موقعیت جدید، بنا به رابطه مکانی معادله (۱۱) حرکت می‌نمایند.

$$X_i^{t+1} = X_i^t + V_i^{t+1} \quad (11)$$

مدلی که طی رابطه ۸ و ۱۱ ارائه شد، مدل کلی نام دارد. برای این الگوریتم مدل دیگری به نام مدل محلی نیز ارائه شده است که در مدل محلی برای هر ذره یک همسایگی تعریف شده و بهترین موقعیت دیده شده در تمام همسایگان ذره محاسبه شده و با $(l_{i1}^t, l_{i2}^t, \dots, l_{ik}^t)$ نمایش داده می‌شود. در این شرایط به جای آنکه تمام ذرات در به روز رسانی از G_{best} استفاده کنند، هر ذره از l_i مربوط به خود استفاده می‌کند. آزمایش‌ها نشان می‌دهد که مدل محلی در

اجرای الگوریتم را به یادداشته، همواره سعی می‌کنند تا آن را نیز در تصمیمات خود دخالت دهند. به همین علت، در صورتی که اشتباهی رخ دهد و تصمیم بدی بگیرند، به زودی آن را جبران می‌کنند. به این ترتیب، اعضای جامعه می‌توانند به راحتی محدوده اطراف خود را جستجو کنند، بدون آنکه نگران خرابتر شدن نتیجه باشند. امروزه در بسیاری از زمینه‌ها این الگوریتم کاربرد دارد. انتخاب ویژگی یکی از این کاربردها بوده که در این مقاله به این منظور از آن استفاده شده است [۱۲ و ۱۳].

۴-۱- الگوریتم PSO

بهینه ساز توده ذرات با یک جمعیت از جواب‌های تصادفی (ذره‌ها) شروع به کار می‌کند. هر ذره i ام در جمعیت با دو بردار موقعیت $X_i^t = (X_{i1}^t, X_{i2}^t, \dots, X_{ik}^t)$ و سرعت $V_i^t = (V_{i1}^t, V_{i2}^t, \dots, V_{ik}^t)$ مدل می‌شود که k بیانگر بعد هر ذره است. اجزا در هر تکرار، بنابر جدیدترین بردار سرعت، از موقعیتی به موقعیت‌های دیگر می‌روند. بردار سرعت هر ذره با استفاده از دو مقدار بهینه بروز رسانی می‌شود. اولین مقدار بهینه، بهترین مقدار شخصی هر ذره P_{best} است که بهترین موقعیتی است که تاکنون ذره موفق به رسیدن به آن شده است و دومین مقدار بهینه، بهترین مقدار سراسری که با G_{best} نشان داده می‌شود، بهترین موقعیتی است که تاکنون توسط جمعیت ذرات به دست آمده است. پس از هر تکرار t الگوریتم PSO بردار سرعت جدید را به صورت معادله (۸) بروز رسانی می‌کند.

$$V_i^{t+1} = W V_i^t + c_1 r_1 (P_i^t - X_i^t) + c_2 r_2 (G_i^t - X_i^t) \quad (8)$$

که در آن w ضریب اینرسی، c_1 فاکتور بهترین جواب محلی و c_2 فاکتور بهترین جواب عمومی است. ضرایب یادگیری c_1 و c_2 در بازه [۰,۲] و معمولاً $c_1 = c_2$ اختیار می‌شود. r_1 و r_2 اعداد تصادفی حاصل از توزیع یکنواخت

در موقعیت بعدی ذره، نسبت به موقعیت فعلی آن است. مقادیر بزرگ V نشان دهنده فاصله زیاد ذره تا رسیدن به مکان بهینه بوده، بنابراین جابه‌جایی بیشتری برای ذره لازم است و بالعکس، مقادیر کوچک نشان دهنده نزدیک شدن مکان ذره به جواب مسأله به گونه‌ای که اگر موقعیت ذره با موقعیت بهینه یکی شود، مقدار سرعت ذره برابر صفر می‌شود و این نشان دهنده آن است که دیگر جابه‌جایی ذره لازم نیست.

در نسخه دودویی الگوریتم، تعاریف متفاوتی از X و V ارائه شده است. به طوری که سرعت ذره به جای آنکه ذره را به سمت جواب‌های بهینه مسائل هدایت کند، به صورت تابع احتمالات، یک یا صفر شدن X را مطرح می‌کند؛ یعنی مقدار V_{ik} احتمال اینکه مقدار X_{ik} یک یا صفر باشد را تعیین می‌کند. از آنجایی که مقدار احتمال باید در بازه صفر و یک باشد، V_{ik} از تابع محدود کننده سیگموئید می‌گذرد.

همان‌طور که مشاهده می‌شود، در نسخه دودویی، بزرگ بودن مقدار V_{ik} نشان دهنده تغییرات زیاد X_{ik} نیست، بلکه متناسب با بزرگ بودن سرعت در جهت مقادیر مثبت احتمال یک بودن X_{ik} ، افزایش می‌یابد و به همین نحو متناسب با بزرگ بودن سرعت در جهت منفی احتمال صفر بودن X_{ik} زیاد می‌شود. از سویی دیگر، اگر سرعت برابر صفر شود، مقدار X_{ik} ثابت نمانده و با احتمال ۵۰٪ صفر و با همین احتمال یک خواهد شد. با توجه به آنچه بحث شد می‌توان ایرادات زیر را به نسخه دودویی متداول وارد دانست [۱۵ و ۱۶].

الف: ایراد نخست از تابع سیگموئید رابطه (۱۲) است. در حالی که بزرگ شدن سرعت در جهت مثبت و منفی از نظر مفهومی با یکدیگر فرق ندارد و نشان دهنده این است که موقعیت ذره برای این بعد خاص باید تغییر کند، در الگوریتم دودویی متداول بین این دو حالت فرق قائل شده است؛ به گونه‌ای که بزرگ شدن آن در جهت مثبت باعث افزایش احتمال یک شدن موقعیت ذره و در جهت منفی باعث افزایش احتمال صفر شدن آن می‌شود. همچنین، وقتی

حل مسائل پیوسته به مراتب بهتر از مدل کلی است. در این مقاله، از مدل کلی این الگوریتم استفاده شده است. شرط خاتمه الگوریتم همگرایی و یا توقف بعد از تعداد معینی تکرار است.

۴-۲- الگوریتم دودویی PSO

برای حل مسائل گسسته روش دودویی این الگوریتم توسط کندی در سال ۱۹۹۷ ارائه شد و به دودویی PSO شهرت یافت [۱۴]. استفاده از الگوریتم دودویی PSO همچنین زمان محاسبات را کاهش می‌دهد. تفاوت عمده این روش با الگوریتم اصلی در روابط بروز رسانی موقعیت و سرعت است که ابتدا باید مقدار بردار سرعت توسط تابع سیگموئید به بازه $[0,1]$ منتقل و سپس موقعیت جدید ذره محاسبه شود. در نسخه دودویی موقعیت هر ذره در هر بعد با دو مقدار صفر و یک مشخص می‌شود و وضعیت انتخاب ویژگی را نمایش می‌دهد. در الگوریتم نسخه دودویی، مفهوم سرعت به مفهوم احتمال، تغییر یافته و V_{ik} احتمال یک بودن X_{ik} را بیان می‌کند. هر ذره بر اساس معادله‌های (۱۲) و (۱۳) به روزرسانی می‌شود که سرعت بر اساس معادله ۹ به دست می‌آید.

$$S(V_i^{t+1}) = \frac{1}{1 + e^{-V_i^{t+1}}} \quad (12)$$

$$X_i^{t+1} = \begin{cases} 1 & \text{if } Rand < S(V_i^{t+1}) \\ 0 & \text{else} \end{cases} \quad (13)$$

در اینجا $Rand$ یک عدد تصادفی در بازه $[0,1]$ است. همان‌گونه که در الگوریتم پیوسته مقدار V_{ik} به یک مقداری محدود می‌شد، در این روش نیز مقدار سرعت محدود به یک مقدار بیشینه V_{max} است. شایان ذکر است که الگوریتم دودویی نیز همانند نوع پیوسته با استفاده از مدل‌های محلی و کلی قابل پیاده سازی است. همان گونه که قبلاً بیان شد، ذرات در الگوریتم اجتماع ذرات پیوسته با دو کمیت X و V تعریف می‌شوند که X موقعیت ذره و V معرف سرعت برای هر ذره است و نشان دهنده میزان تغییر

الگوریتم استفاده شده و با استفاده از رابطه (۱۷) به دست می‌آید [۱۵].

$$A = k(1 - e^{-\frac{F}{T}}) \quad (17)$$

در این رابطه k ، یک ضریب ثابت، T ثابت زمانی که با توجه به نوع و بعد مساله انتخاب می‌شود و F یک شمارنده است و دفعاتی را که الگوریتم دچار شکست می‌شود، می‌شمارد. شکست این گونه تعریف می‌شود که الگوریتم نتواند موقعیت بهتری را به دست آورد یا به عبارتی دیگر $G_{best}(t) = G_{best}(t-1)$ ؛ بنابراین، با هر دفعه شکست به مقدار F یکی افزوده می‌شود. شایان ذکر است که F در لحظه شروع برابر صفر و پس از هر بار موفقیت، الگوریتم این شمارنده صفر خواهد شد. هنگامی که مقدار F برابر صفر باشد، مقدار A نیز صفر خواهد بود و بدان معنی است که الگوریتم در مسیر درستی قرار دارد و هیچ گونه جهشی صورت نمی‌گیرد. الگوریتم بهبود یافته فوق در محیط گسسته را INBPSO می‌نامیم [۱۵].

تابع برازندگی مورد استفاده در این مقاله که توسط الگوریتم BPSO بهینه می‌شود، به دو مقدار تعداد ویژگی-های انتخاب شده و دقت بازشناسی وابسته است. برای این منظور، از طبقه‌بند ماشین بردار پشتیبان SVM کتابخانه LIBSVM استفاده شده است. با ترکیب دو مقدار تعداد ویژگی و دقت بازشناسی تابع برازندگی با معادله (۱۸) تعیین می‌گردد.

$$Fitness(i) = \frac{K_1}{|FS|} + K_2 \times Accuracy \quad (18)$$

که در آن $|FS|$ برابر تعداد ویژگی‌های انتخابی و K_1 و K_2 اعداد ثابت هستند.

۴-۳- انتخاب ویژگی

مسأله انتخاب ویژگی، یکی از مسائلی است که در مبحث یادگیری ماشین و در بررسی مسائل شناخت الگو از دیدگاه آمار مطرح است. با انتخاب ویژگی مناسب در نرخ

سرعت ذره برای یک بعد مشخص به صفر نزدیک می‌شود، به معنی آن است که ذره در آن بعد موقعیت مناسبی دارد و باید موقعیت ذره تغییر نکند. این در حالی است که با تابع احتمال سیگموئید، احتمال صفر یا یک شدن موقعیت ذره در آن بعد برابر خواهد شد.

ب: ایراد دوم از رابطه به روزرسانی موقعیت ذره رابطه (۱۲) است. در این رابطه موقعیت قبلی ذره برای محاسبه موقعیت بعدی آن در نظر گرفته نمی‌شود.

این دو ایراد باعث می‌شود الگوریتم همگرا نشده و ذرات از مکان بهینه دور شوند. برای رفع ایراد اول از تابع احتمال رابطه (۱۴) و در ادامه سرعت با استفاده از رابطه (۱۵) به روز می‌شود.

$$S(V_{ik}^{t+1}) = |\tanh(\alpha V_{ik})| \quad (14)$$

$$(15)$$

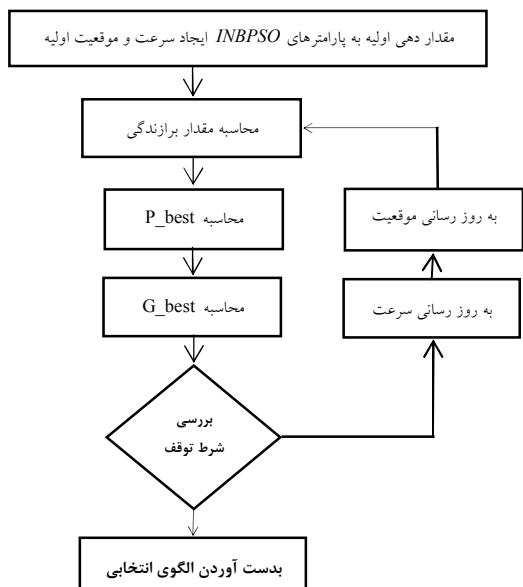
$$if \quad rand < S(V_{ik}(t+1)) \begin{cases} then \quad X_{ik}(t+1) = complement(X_{ik}(t)) \\ else \quad X_{ik}(t+1) = X_{ik}(t) \end{cases}$$

در رابطه (۱۴) مقدار α یک ضریب ثابت است که تقریباً یک در نظر گرفته می‌شود.

هنگامی که الگوریتم اجتماع ذرات دچار همگرایی می‌شود، ذرات به سمت یک بهینه محلی حرکت کرده، در آن گرفتار می‌شوند. در این حالت از آنجایی که ذرات به ذره با موقعیت G_{best} نزدیک شده‌اند، سرعت آنها به صفر نزدیک می‌شود. در چنین حالتی احتمال تغییر مکان ذره‌ها در الگوریتم دودویی روابط (۱۴) و (۱۵) کم شده، به سمت صفر میل می‌کند و در بهینه محلی گرفتار می‌شود. برای حل این مشکل و گریز از این مشکل، رابطه (۱۴) به صورت رابطه ۱۶ تغییر می‌کند.

$$S(V_i^{t+1}) = A + (1 - A) \times |\tanh(\alpha V_{ik})| \quad (16)$$

در این رابطه ضریب A برای جلوگیری از رکود



شکل (۴): روندنما به دست آوردن الگوی انتخابی

در ادامه در جدول (۵) نتیجه به دست آمده با کارهایی که قبلاً بر روی این دیتابیس صورت گرفته، بررسی شده است. سپس تعداد ویژگی‌ها با استفاده از الگوریتم بهینه سازی توده ذرات دودویی بر اساس انتخاب ویژگی کاهش داده شده است؛ به طوری که دقت بازشناسی بالا نگه داشته شود. برای این کار در تابع برازندگی معادله (۱۸) مقدار K_2 بزرگتر از K_1 انتخاب شده است.

پارامترهای انتخاب شده برای الگوریتم PSO دودویی بهبود یافته به صورت جدول (۶) است.

مقدار $W_{initial}$ برابر با ۰٫۹ در نظر گرفته شده که پس از هر تکرار با استفاده از معادله (۹) بروز رسانی شده است. نتایج به دست آمده از ترکیب ویژگی‌های استخراج شده با استفاده از دو روش بالا و کاهش ویژگی در جدول (۷) نشان داده شده است.

بازشناسی و کنار گذاشتن ویژگی‌هایی که در طبقه بندی تأثیر منفی دارند، می‌توان هم نرخ بازشناسی را بهبود بخشید و همچنین، هزینه محاسباتی را کاهش و سرعت پردازش را افزایش داد که این امر در سیستم‌های بلا درنگ امر حیاتی است. عموماً از الگوریتم‌های متفاوتی می‌توان برای استخراج ویژگی استفاده نمود که در سال‌های اخیر کارهای فراوانی روی آن صورت گرفته است، اما استفاده از کل این ویژگی‌ها که ممکن است ویژگی‌هایی نیز در آن باشد که اثر منفی در نرخ بازشناسی دارد، مقرون به صرفه نبوده و نیز طول بردار ویژگی بزرگ سرعت و کارایی سیستم بازشناسی را کاهش داده، ما را در یافتن پاسخ بهینه و سریع باز می‌دارد [۱۷]. در این مقاله پس از استخراج ویژگی با استفاده از ترکیب الگوریتم‌های هیستوگرام گرادیان و مکان مشخصه توسعه یافته ویژگی‌های مناسب با استفاده از الگوریتم دودویی PSO بهبود یافته انتخاب شده است. روندنما برای به دست آوردن الگوی انتخابی به صورت شکل (۴) است.

۵- نتایج

در این مقاله از پایگاه داده ارقام دستنویس فارسی هدی استفاده شده است [۹]. با استفاده از ترکیب الگوریتم‌های هیستوگرام گرادیان و مکان مشخصه توسعه یافته ویژگی‌ها برای ۶۰۰۰۰ داده آموزش و ۲۰۰۰۰ داده آزمون استخراج شد و با کنار هم قرار دادن بردارهای ویژگی استخراجی از دو روش ذکر شده، بردار ویژگی اصلی با طول ۴۰۰ به دست آمد. دقت بازشناسی برای ۲۰۰۰۰ داده آزمون با ۱۶ چرخش در روش گرادیان با استفاده از طبقه‌بند SVM بدون کاهش ویژگی در تعداد نمونه‌های آموزش مختلف به صورت جدول (۴) است. در بین داده‌ها بازشناسی ارقام ۲ و ۳ نسبت به سایر ارقام بازشناسی پایینی دارند و ارقام ۱ و ۸ بیشترین بازشناسی را به خود اختصاص داده‌اند.

جدول (۴): دقت بازشناسی برای ۲۰۰۰۰ نمونه آزمون و تعداد نمونه‌های آموزش مختلف بدون کاهش ویژگی با استفاده از طبقه‌بند

SVM

تعداد نمونه‌های آموزش	۱۰۰۰	۵۰۰۰	۱۰۰۰۰	۲۰۰۰۰	۳۰۰۰۰	۴۰۰۰۰	۵۰۰۰۰	۶۰۰۰۰
دقت بازشناسی	%۹۶,۶۶	%۹۸,۴۶	%۹۸,۸۴	%۹۹,۱۱	%۹۹,۲۲	%۹۹,۲۹	%۹۹,۳۳	%۹۹,۴۰

جدول (۵): مقایسه کارهای قبلی با روش پیشنهادی بر روی دیتابیس هدی

Algorithms	Dataset Size		Recognition Rate %	
	Train	Test	Train	Test
Harifi., Aghagolzadeh [18]	230	500	-	97.60
Hosseini, Bouzardum [19]	480	480	-	92
Mowlaei et al. [20]	2240	1600	99.29	91.88
Mozaffari et al. [21]	2240	1600	98.00	91.37
Mozaffari et al. [22]	2240	1600	100	94.44
Mowlaei, Faez [23]	2240	1600	100	92.44
Shirali-Shahreza et al.[24]	2600	1300	-	97.80
Soltanzadeh, Rahmati [25]	4979	3939	-	99.57
Dehghan, Faez [26]	6000	4000	-	97.01
Ziaratban et al. [27]	6000	4000	100	97.65
Sadri et al. [28]	7390	3035	-	94.14
A. Alaei et al. [29]	60000	20000	99.99	98.71
A. Alaei et al. [30]	60000	20000	99.99	99.02
H. Parvin et al. [31]	60000	20000	-	98.27
Parvin et al. H [32]	60000	20000	-	98.89
نجوی، کیانی، ابراهیم پور [۳۳]	60000	20000	-	99.15
Proposed Algorithm	60000	20000	100	99.40

جدول (۶): پارامترهای انتخاب شده برای الگوریتم INBPSO

تعداد ذرات	۳۰
ضریب C_1 و C_2	۲
ضریب K_1	۴
ضریب K_2	۶
تعداد تکرار	۱۵
T	۱۰۰۰
A	۱
α	۰,۹۸

جدول (۷): دقت بازشناسی برای ۲۰۰۰۰ نمونه آزمون و تعداد نمونه‌های آموزش مختلف با کاهش ویژگی با استفاده از طبقه‌بند SVM

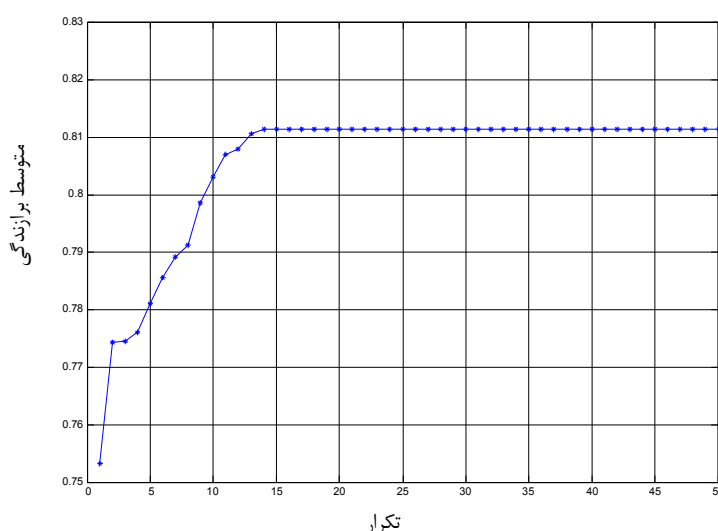
تعداد نمونه‌های آموزش	۱۰۰۰	۵۰۰۰	۱۰۰۰۰	۲۰۰۰۰	۳۰۰۰۰	۴۰۰۰۰	۵۰۰۰۰	۶۰۰۰۰
تعداد ویژگی	۴۹	۵۴	۵۴	۵۳	۵۶	۵۹	۵۸	۶۴
دقت بازشناسی	%۹۶,۱	%۹۸,۲۳	%۹۸,۴۸	%۹۸,۵۳	%۹۸,۸۰	%۹۸,۹۶	%۹۹,۰۹	%۹۹,۲۸

استفاده از الگوریتم INBPSO نشان می‌دهد. در شکل (۵) گراف مربوط به متوسط شایستگی بر اساس تعداد تکرار برای ۱۰۰۰ نمونه از داده‌های آموزش و ۲۰۰۰۰ نمونه آزمون رسم شده است. با توجه به شکل مشاهده می‌شود که متوسط شایستگی با افزایش تعداد تکرار الگوریتم رشد داشته، به مقدار خاصی همگرا می‌گردد.

با مقایسه جدول (۴) و جدول (۷) می‌توان دید که دقت‌های به دست آمده از روش بدون کاهش ویژگی و با کاهش ویژگی به وسیله الگوریتم بهینه سازی توده ذرات دودویی بهبود یافته، تفاوت چندانی در بازشناسی با هم ندارند و با کاهش ویژگی تا تعداد قابل قبولی هنوز دقت بازشناسی به مقدار قابل قبولی در حد بالا قرار دارد. جدول (۸) مقدار بازشناسی را برای تعداد ویژگی‌های مختلف با

جدول (۸): دقت بازشناسی برای ۲۰۰۰۰ نمونه آزمون و تعداد ویژگی‌های مختلف با استفاده از طبقه‌بند SVM

تعداد ویژگی	۶۴	۱۰۵	۱۵۱	۲۱۰	۲۷۰	۳۳۴	۴۰۰
دقت بازشناسی	%۹۹,۲۷	%۹۹,۱۹	%۹۹,۲۴	%۹۹,۳۱	%۹۹,۳۴	%۹۹,۳۷	%۹۹,۴۰



شکل (۵): همگرایی متوسط شایستگی، با افزایش تعداد تکرار در الگوریتم بهینه سازی توده ذرات دودویی بهبود یافته

بازشناسی در دو مرحله آموزش و آزمون صورت گرفت. در مرحله آموزش که برای ۶۰۰۰۰ نمونه صورت گرفت، پارامترهای طبقه‌بند SVM تعیین، و در مرحله آزمون استفاده شد.

مقایسه نتایج با کار محققان دیگر حاکی از آن است روش ارائه شده در استخراج ویژگی و انتخاب ویژگی از کارایی مناسبی برخوردار است.

پیشنهاد می‌شود برای کار در آینده قبل از استفاده از الگوریتم هیستوگرام گرادیان، ابتدا از تبدیل موجک روی تصویر اصلی استفاده شود و سپس از تصویر تقریب با

۶- نتیجه گیری

در این مقاله ویژگی‌های ارقام دستنویس فارسی با استفاده از ترکیب دو الگوریتم هیستوگرام گرادیان و مکان مشخصه توسعه یافته استخراج گردید و توسط طبقه‌بند SVM بازشناسی با ۴۰۰ ویژگی به دست آمده صورت گرفت و دقت %۹۹,۴۰ حاصل شد. توسط الگوریتم بهینه‌سازی توده ذرات دودویی و تعریف تابع برازندگی خاص، ۶۴ ویژگی انتخاب و دقت بازشناسی %۹۹,۲۸ حاصل گردید. در تابع برازندگی ارائه شده دقت بازشناسی و تعداد ویژگی‌های مورد استفاده لحاظ شده است.

- International Conference on Computational Cybernetics and Simulation, Vol. 5, pp.4104-4108, 1997.
- [15] H. Nezamabadipoor, M. Rostami, M. Maghfoori, "Particle Swarm Optimization, Challenges and New Solutions", Iranian Journal of Electrical and Computer Engineering, Vol.6, No. 1, pp. 21-32, 2008.
- [16] M. Rostami, H. Nezamabadipoor, "Introducing New Algorithm for Binary PSO", 14th Iranian Conference on Electrical Engineering, 2006.
- [17] R. M. Ramadan and R. F. Abdel-Kader, "Face Recognition Using Particle Swarm Optimization-Based Selected Features", International Journal of Signal Processing, Image Processing and Pattern Recognition, Vol. 2, No.2, pp. 51-66, June 2009.
- [18] A. Harifi and A. Aghagolzadeh., "A New Pattern for Handwritten Persian/ Arabic Digit Recognition", Journal of Information Technology, Vol. 3, pp. 249-252, 2004.
- [19] H. Mir Mohammad Hosseini and A. Bouzerdoum, "A Combined Method for Persian and Arabic Handwritten Digit Recognition", Australian New Zealand Conference on Intelligent Information System pp. 80 – 83, 1996.
- [20] A. Mowlaei, K. Faez & A. Haghghat, "Feature Extraction with Wavelet Transform for Recognition of Isolated Handwritten Farsi/Arabic Characters and Numerals", 14th International Conference on Digital Signal Processing, Vol. 2, pp. 923-926, 2002.
- [21] S. Mozaffari, K. Faez & H. RashidyKanan, "Recognition of Isolated Handwritten Farsi/Arabic Alphanumeric Using Fractal Codes", Image Analysis and Interpretation, 6th Southwest Symposium, pp. 104-108, 2004.
- [22] S. Mozaffari, K. Faez and M. Ziaratban, "Structural Decomposition and Statistical Description of Farsi/Arabic Handwritten Numeric Characters", Proceedings of the 8th Intl. Conference on Document Analysis and Recognition, Vol. 1, pp. 237- 241, 2005.
- [23] A. Mowlaei and K. Faez, "Recognition Of Isolated Handwritten Persian Arabic Characters And Numerals Using Support Vector Machines", Proceedings of XIII Workshop on Neural Networks for Signal Processing, pp. 547-554, 2003.
- [24] M. H. Shirali-Shahreza, K. Faez and A. Khotanzad, "Recognition of Hand-written Persian/Arabic Numerals by Shadow Coding and an Edited Probabilistic Neural Network", Proceedings of International Conference on Image Processing, Vol. 3, pp. 436-439, 1995.
- [25] H. Soltanzadeh and M. Rahmati, "Recognition of Persian Handwritten Digits Using Image Profiles of Multiple Orientations", Pattern Recognition Letters 25, pp. 1569–1576, 2004.
- [26] M. Dehghan and K. Faez, "Farsi Handwritten Character Recognition With Moment Invariants", Proceedings of 13th International Conference on Digital Signal Processing, Vol 2, pp. 507-510, 1997.
- [27] M. Ziaratban, K. Faez and F. Faradj, "Language-استفاده از الگوریتم هیستوگرام گرادیان ویژگی‌ها استخراج شود. این کار باعث افزایش سرعت برنامه خواهد شد.
- مراجع:**
- [1] HS. Impedovo, P.S. Wang, and H. Bunke, editors, "Automatic Bankcheck Processing", World Scientific Publ. Co, Singapore, 1997.
- [2] I.D. Trier, A.K. Jain, "Feature Extraction Methods for Character Recognition- A Survey", Pattern Recognition, Vol. 29, No. 4, pp. 641-662, 1996.
- [3] H. Takahashi, "A Neural Net OCR Using Geometrical and Zonal Pattern Features", in proceedings of the first international Conference on Document Analysis and Recognition, pp. 821-828, 1991.
- [4] L.O. Jimenez, A. Morales-Morell, "Classification of Hyperdimensional Data Based on Feature and Decision Fusion Approachs Using Projection Pursuit, Majority Voting, and Neural Networks", IEEE Trans. on Geoscience and Remote Sensing, Vol. 37, No.3, pp.1360,1366, May 1999.
- [5] M.H ShirAli, A.R. Khotanzad, "Shift and Scale Invariant Digit Recognition Using Zernik Moments and Neural Networks", 2nd Iranian Conference on Electrical Engineering, Vol. 5, pp. 417-424, 1994
- [6] R. Azmi, E. Kabir, K. Badi, "Printed Character Recognition Using Features of Peripheral Curvature", Iranian Journal of Electrical and Computer Engineering, Vol 1, Issue 1, pp. 29-37, 2003.
- [7] Y. Li., "Reforming the Theory of Invariant Moments for Pattern Recognition", Pattern Recognition Letters, Vol. 25, Issue 7, pp. 723-730, July1992.
- [8] H.A. Glucksman, "Multicategory of Patterns Represented by High-Order Vectors of Multilevel Measurement", IEEE Transaction Computer, pp. 1593- 1598, Dec. 1971.
- [9] H. Khosravi, E. Kabir, "Introducing a Very Large Dataset of Handwritten Farsi Digits and a Study on their Varieties", Pattern Recognition Lett, Vol. 28, Issue 10, pp.1133-1141.
- [10] H. Khosravi, E. Kabir, "Introducing Two Fast and Efficient Features for Recognition of Persian Handwritten Digits", 4th Conference on Machine Vision and Image Processing, pp.25-26, 2006.
- [11] J. Kennedy and R. C. Eberhart, "Particle Swarm Optimization", Proceedings of IEEE International Conference on Neural Networks, Vol. 4, pp.1942-1948, 1995.
- [12] M. Rostami, H. Nezamabadipoor, "Feature Selection in Content Based Image Retrieval Using PSO Algorithm", 11th Iranian CSI Computer Conference, pp.269-275, 2005.
- [13] H. A. Firip, and E. Goodman, "Swarmed Feature Selection", in IEEE Proceedings of the 33rd Applied Imagery Pattern Recognition Workshop, pp. 112-118, 2004.
- [14] J. Kennedy and R. C. Eberhart, "A Discrete Binary Version of the Particle Swarm Algorithm", IEEE

- Based Feature Extraction Using Template-Matching in Farsi/Arabic Handwritten Numeral Recognition", Proceedings of 9th International Conference on Document Analysis and Recognition, Vol.1, pp. 297-301, 2007.
- [28] J. Sadri, C. Y. Suen and T. D. Bui, "Application of Support Vector Machines for Recognition of Handwritten Arabic/Persian Digits", Proceedings of the 2nd Conference on Machine Vision and Image Processing & Applications, Vol. 1, pp. 300-307, 2003.
- [29] Alireza Alaei, Umapada Pal and P. Nagabhushan, "Using Modified Contour Features and SVM Based Classifier for the Recognition of Persian/Arabic Handwritten Numerals", Proceedings of 7th International Conference on Advances in Pattern Recognition, pp.391-394, 2009.
- [30] Alireza Alaei, P. Nagabhushan and Umapada Pal, "Fine Classification of Unconstrained Handwritten Persian/Arabic Numerals by Removing Confusion Amongst Similar Classes", Proceedings of 10th International Conference on Document Analysis and Recognition, pp. 601-605, 2009.
- [31] H. Parvin, H. Alizadeh and B. Minaei-Bidgoli, "A New Approach to Improve the Vote-Based Classifier Selection", Proceedings of Fourth International Conference on Networked Computing and Advanced Information Management, pp. 91-95, 2008.
- [32] H. Parvin, H. Alizadeh, B. Minaei-Bidgoli and M. Analoui, "A Scalable Method for Improving the Performance of Classifiers in Multiclass Applications by Pairwise Classifiers and GA", Proceedings of Fourth International Conference on Networked Computing and Advanced Information Management, pp.137-142, 2008.
- [33] M. Nahvi, K. Kiani, R. Ebrahimpoor, "Improving Gradient Feature Extraction Based on Discrete Cosine Transform Towards Recognition of Persian Handwritten Digits", 18th Iranian Conference on Electrical Engineering, pp.3067-3071, 2010.