

بازشناسی ارقام دستنویس فارسی با استفاده از ماشین بردار پشتیبان

سلمان کریم زاده بجزستانی^۱، علیرضا محمدی عنبران^۲

^۱ دانشجوی کارشناسی ارشد مخابرات، دانشگاه آزاد اسلامی واحد گناباد

^۲ عضو هیئت علمی الکترونیک، دانشگاه آزاد اسلامی واحد گناباد

نام و نشانی ایمیل نویسنده مسئول:

سلمان کریم زاده بجزستانی

Salman.karimzade@yahoo.com

چکیده

تشخیص و بازشناسی ارقام و حروف فارسی نوری (OCR)، امروزه یکی از مهمترین و جذاب ترین حوزه ها در بازشناسی الگو و ماشین بینایی است. هدف اصلی از ارائه این مقاله، بازشناسی ارقام دستنویس فارسی می باشد تا بتوان توسط آن اعداد دستنویس افراد مختلف را شناسایی و تشخیص داد. در این مقاله، جهت بازشناسی ارقام دستنویس فارسی از پایگاه داده هدی استفاده شده است که شامل ۱۰ کلاس مختلف برای ارقام ۰ تا ۹ است. جهت استخراج ویژگی ارقام دستنویس فارسی از تبدیل موجک دو بعدی به همراه الگوریتم PCA استفاده شده است و جهت طبقه بندی ارقام دستنویس فارسی دقیق اعداد از ماشین بردار پشتیبان (SVM) استفاده شده است. دقت تشخیص اعداد دستنویس فارسی ۹۱٫۷۵ درصد است.

واژگان کلیدی: ارقام و حروف نوری (OCR)، بازشناسی الگو، ماشین بردار پشتیبان (SVM)، تبدیل موجک، الگوریتم PCA

مقدمه

رایانه تنها یک ابزار است و ما انسانها باید شیوه‌ی کارکرد را به آن بیاموزیم، و امکانات کار را برای آن فراهم کنیم. یکی از این نرم افزارها، نرم افزارهای بازشناسی حروف فارسی^۱ می باشد. استفاده از الگوریتم های هوشمند در بازشناسی حروف و اعداد، یکی از مباحث جذاب در چند دهه اخیر بوده است. در چند دهه گذشته بازشناسی الگوهای نوشتاری شامل حروف، ارقام و دیگر نمادهای متداول در اسناد نوشته شده به زبان های مختلف، توسط گروه های مختلفی از محققین مورد مطالعه و بررسی قرار گرفته است. نتیجه این تحقیقات منجر به پیدایش مجموع های از روشهای سریع و تا حد زیادی مطمئن موسوم به OCR به منظور وارد نمودن اطلاعات موجود در اسناد، مدارک، کتا بها و سایر مکتوبات تایپی و حتی دست نوشت به داخل رایانه شده است. در ساختار این مقاله در بخش دوم به بررسی سیستم بازشناسی حروف و ارقام نوری فارسی پرداخته می شود. در بخش سوم الگوریتم تبدیل موجک جهت استخراج ویژگی ارقام دستنویس فارسی مورد بررسی قرار می گیرد. در بخش چهارم، از روش ماشین بردار پشتیبان طبقه بندی داده های ورودی استفاده می شود و در نهایت، در بخش پنجم نتایج حاصل ارائه خواهند شد.

۱- سیستم بازشناسی ارقام دستنویس

نویسه کوچک ترین جز نوشتاری است که شامل ارقام، حروف و علائم است. بازشناسی نویسه ها به معنی شناسایی و تشخیص ارقام از روی شکل ظاهری تصویر آنها می باشد. به طور کلی بخش های مختلف یک سیستم بازشناسی ارقام دستنویس به صورت شکل (۱) است.



شکل ۱ - سیستم بازشناسی ارقام فارسی دستنویس

همانطور که ملاحظه می شود بخش های اصلی یک سیستم بازشناسی نویسه بصورت زیر است:

- ورودی دیتا
- پیش پردازش
- بخش بندی
- استخراج ویژگی
- طبقه بندی

در شکل (۱)، پس از اعمال دیتای ورودی، ویژگی هر یک از تصاویر با استفاده از الگوریتم تبدیل موجک دوبعدی استخراج می شود. سپس جهت کاهش ویژگی از شبکه عصبی PCA استفاده شده است. در نهایت جهت طبقه بندی دیتای ورودی از ماشین بردار پشتیبان (SVM) استفاده شده است. در ادامه به بررسی هر یک از این قسمت ها پرداخته می شود.

۲- استخراج ویژگی

در هر سیستم بازشناسی الگو، بخش استخراج ویژگی یکی از بخش های اصلی است. چرا که نتایج حاصل از این مرحله، مستقیماً بر روی کیفیت مرحله بازشناسی اثر می گذارد. در این مرحله، به هر الگوی ورودی، یک کد یا بردار ویژگی نسبت داده می شود که معرف آن الگو در فضای ویژگی ها است و آن را از دیگر الگوها متمایز می سازد. حال در ادامه، به بررسی الگوریتم تبدیل موجک می پردازیم.

۱-۲- تبدیل موجک

اخیراً تبدیلات زیر باند و موجک گسسته^۲ (DWT) در عرصه کاربردهای کدینگ، شاخص گذاری تصاویر^۳ و بازشناسی حروف با مقبولیت فراوانی مواجه شده اند. در تبدیل فوریه که نوعی از تبدیل با تابع پایه متعامد (ارتوگونال^۴) است، از موج های سینوسی استفاده می شود که در

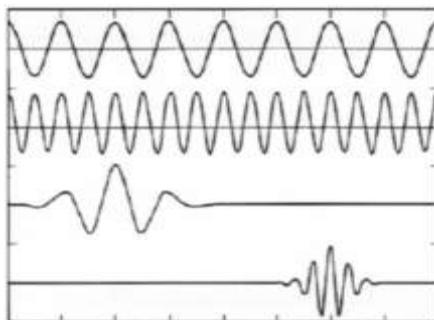
¹ Optical Character Reader

² Discrete Wavelet Transform

³ Image Indexing

⁴ Orthogonal

هر دو جهت دارای دوره نامحدود هستند. به طور مشابه در حالت گسسته، بردارهای پایه تبدیل DFT نیز دارای دوره نامحدود هستند و بنابراین این تبدیل دارای ویژگی (دوره-محدود) نیست. اما در عمل به سیگنال‌های گذرای برمی‌خوریم که تنها در یک دوره کوتاه زمانی غیر صفر هستند. مثلاً بسیاری از خصوصیات مهم تصاویر مانند لبه، در یک مکان مشخص قرار دارند. این ترکیبات به هیچ یک از توابع پایه تبدیل فوریه شبیه نیستند و در نتیجه توابع مذکور ابزار بهینه‌ای برای آنالیز سیگنال‌ها و تصاویری حاوی مؤلفه‌های گذرا یا متمرکز در یک زمان یا مکان، نیستند. البته می‌دانیم که تبدیل فوریه هر تابع تحلیلی را به صورت مجموعی از سینوس‌ها نمایش می‌دهد. اما ضرایب بدست آمده از تبدیل، معیار فشرده‌ای از اطلاعات تصویر نخواهد بود. برای رفع این نقیصه ریاضیدانان و مهندسان به کوشش‌هایی دست زده‌اند که در آنها توابع پایه با دوره محدود بکار برده می‌شوند. به این توابع پایه "موجک" و به تبدیلاتی که بر اساس آنها هستند، "تبدیلات موجک" می‌گویند. شکل (۲) تفاوت بین موج و موجک را نشان می‌دهد. دو منحنی بالایی، امواج سینوسی با فرکانس‌های مختلف و دو منحنی پایینی، دو موجک با فرکانس و موقعیت متفاوت هستند.



شکل ۲- تفاوت بین موج و موجک. دو منحنی بالایی، امواج سینوسی با فرکانس‌های مختلف و دو منحنی پایینی، دو موجک با فرکانس و موقعیت متفاوت

تبدیل موجک انواع مختلفی دارد که از توابع پایه متفاوتی برخوردارند، نظیر: هار^۵، مایر^۶، کلاه^۷، مکزیکی^۸ و سیملت^۹. از آنجائیکه توابع پایه موجک هار شکل گسسته دارند، تصور می‌شود که این موجک برای کار با تصاویر باینری کارا کترها، نسبت به سایر موجک‌ها مناسب‌تر باشد. تبدیل موجک، در واقع تکنیک پنجره با اندازه متغیر است. یا به عبارتی با شیفت و تغییر دادن مقیاس انواع موجک‌ها و ضرب در سیگنال اصلی و جمع کردن آنها حاصل می‌شود. این گفته با رابطه (۱) زیر نشان داده شده است.

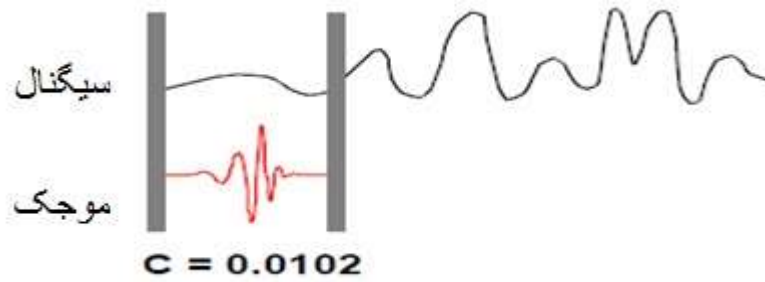
$$C(\text{scale, position}) = \int_{-\infty}^{\infty} f(t)\psi(\text{scale, position, } t)dt \quad (1)$$

از فرمول بالا ضرائب موجک (C) زیادی به دست می‌آید، که تابعی از مقیاس^{۱۰} و موقعیت^{۱۱} هستند. برای بدست آوردن تبدیل موجک ۵ مرحله زیر انجام می‌شود.

انتخاب یک موجک و مقایسه آن با یک بخش در شروع سیگنال

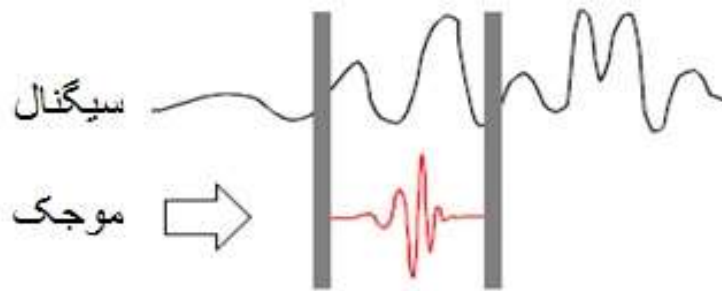
محاسبه ضریب C، که مقدار همبستگی موجک را با آن قسمت سیگنال نشان می‌دهد.

- 5 Haar
- 6 Meyer
- 7 Hat
- 8 Mexican
- 9 Symlets
- 10 Scale
- 11 Position



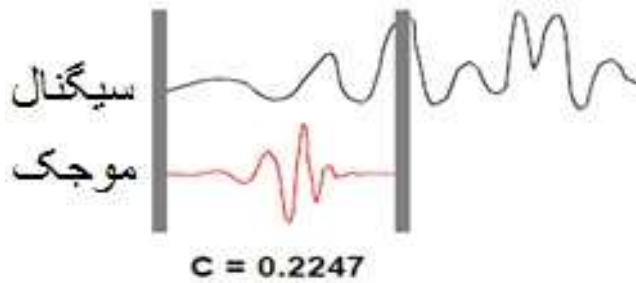
شکل ۳: محاسبه مقدار همبستگی موجک با یک قسمت سیگنال

موجک را به سمت راست شیف‌ت می‌دهیم و مراحل ۱ و ۲ را تکرار می‌کنیم تا کل سیگنال را بپوشانند.



شکل ۴: محاسبه مقدار همبستگی موجک با یک قسمت سیگنال

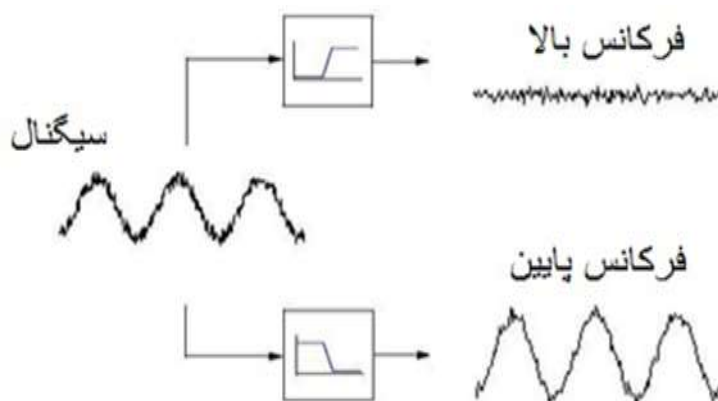
موجک را می‌کشیم^{۱۲} (باز کردن سیگنال) و مراحل ۱ تا ۳ را تکرار می‌کنیم.



شکل ۵: باز کردن تصویر برای تغییر مقیاس و تکرار مراحل ۱ تا ۳

مراحل ۱ تا ۴ را برای همه مقیاس‌ها تکرار می‌کنیم.

در واقع در تبدیل موجک تصویر به طور پیاپی از مجموعه‌ای از فیلترهای پایین‌گذر و بالاگذر عبور داده می‌شود. خروجی فیلتر بالاگذر، اطلاعات و جزئیات ریز تصویر را به ما می‌دهد و فیلتر پایین‌گذر اطلاعات کلی تصویر را به ما می‌دهد (شکل ۶). از خروجی‌های فیلتر پایین‌گذر به عنوان ضرائب استفاده می‌شود.



شکل ۶: عبور تصویر به طور پیاپی از مجموعه‌ای از فیلترهای پایین‌گذر و بالاگذر برای استخراج ضرائب موجک

۳- طبقه بندی

آخرین بخش یک سیستم بازشناسی نویسی، طبقه بندی داده‌ها می‌باشد. فرایند طبقه بندی، در واقع نوعی یادگیری با ناظر^{۱۳} می‌باشد که در طی دو مرحله انجام می‌گردد. در مرحله اول مجموعه‌ای از داده‌ها که در آن هر داده شامل تعدادی خصوصیت دارای مقدار و یک خصوصیت بنام خصوصیت کلاس می‌باشد، برای ایجاد یک مدل داده بکار می‌روند که این مدل داده در واقع توصیف کننده مفهوم و خصوصیات مجموعه داده‌هایی است که این مدل از روی آنها ایجاد شده است. مرحله دوم فرایند طبقه بندی اعمال یا بکارگیری مدل داده ایجاد شده بر روی داده‌هایی است که شامل تمام خصوصیات داده‌هایی که برای ایجاد مدل داده بکار گرفته شده‌اند، می‌باشد، به جز خصوصیت کلاس این مقادیر که هدف از عمل طبقه بندی نیز تخمین مقدار این خصوصیت می‌باشد. روش‌های زیادی جهت طبقه بندی وجود دارد که ما در ادامه به بررسی روش پیشنهادی خود می‌پردازیم.

۳-۱- ماشین بردار پشتیبان

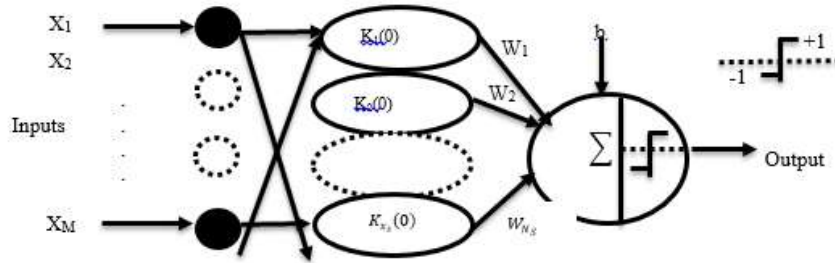
ماشین بردار پشتیبان یا در واقع یک طبقه بندی کننده دودویی است که دو کلاس را با استفاده از یک مرز خطی از هم جدا می‌کند. الگوریتم SVM اولیه در ۱۹۶۳ توسط ولادیمیر وپنیک ابداع شد و در سال ۱۹۹۵ برای حالت غیرخطی تعمیم داده شد. ماشین بردار پشتیبان یکی از روش‌های یادگیری بانظارت است که از آن برای طبقه بندی و رگرسیون استفاده می‌کنند. این روش از جمله روش‌های نسبتاً جدیدی است که در سال‌های اخیر کارایی خوبی نسبت به روش‌های قدیمی‌تر برای طبقه بندی از جمله شبکه‌های عصبی پرسپترون نشان داده است. شاید به گونه‌ای بتوان محبوبیت کنونی روش ماشین بردار پشتیبان را با محبوبیت شبکه‌های عصبی در دهه گذشته مقایسه کرد. علت این قضیه نیز قابلیت استفاده این روش در حل مسائل گوناگون می‌باشد، در حالیکه روش‌هایی مانند درخت تصمیم گیری را نمی‌توان به راحتی در مسائل مختلف به کار برد. ماشین بردار پشتیبان در انواع دسته بندی‌ها همچون تشخیص ارقام دستنویس، تشخیص شیء، شناسایی صورت، دسته بندی انواع صداها و مانند آن مورد استفاده قرار گرفته است که درمقایسه با تکنیک‌های دیگر از کارایی قابل ملاحظه‌ای برخوردار است.

مبنای کاری دسته بندی کننده SVM دسته بندی خطی داده‌ها است و در تقسیم خطی داده‌ها سعی می‌شود خطی انتخاب شود که حاشیه اطمینان بیشتری داشته باشد. حل معادله پیدا کردن خط بهینه برای داده‌ها به وسیله روش‌های QP^{14} که روش‌های شناخته شده‌ای در حل مسائل محدودیت‌دار هستند صورت می‌گیرد. SVM از یک تکنیک که kernel trick نامیده می‌شود، برای تبدیل داده‌ها استفاده می‌کند و سپس بر اساس این تبدیل، مرز بهینه بین خروجی‌های ممکن را پیدا می‌کند. به عبارت ساده تبدیلات بسیار پیچیده را انجام می‌دهد، سپس مشخص می‌کند چگونه داده‌ها را بر اساس برچسب‌ها یا خروجی‌ها جدا کند. از مزایای SVM جداسازی کلاس‌ها با توجه به پراکندگی آنها می‌باشد. این تکنیک به عنوان یک جداساز بهینه کارایی خوبی را در طبقه بندی داده‌ها در کاربردهای مختلف از خود نشان داده است. در مرحله آموزش، برای یک تعداد بردار ویژگی با کلاس معلوم، مرز بهینه بین دو کلاس تعیین می‌شود. در SVM با استفاده از تمامی بردارهای آموزشی و یک الگوریتم بهینه‌سازی، تعدادی از نمونه‌های آموزشی که مرزهای کلاس‌ها را ایجاد می‌کنند بدست

¹³ Supervised Learning

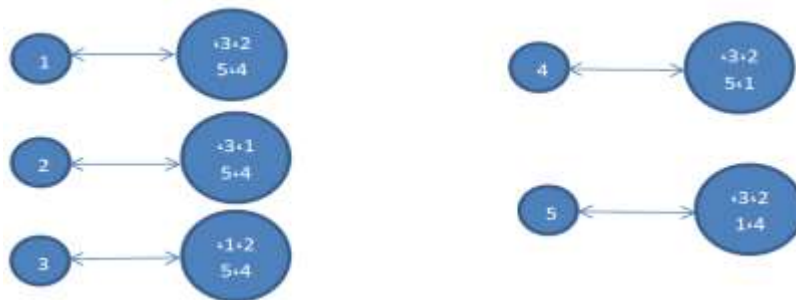
¹⁴ Quadratic Programming

می‌آید که این نمونه‌های آموزشی را بردارهای پشتیبان می‌نامند. نمونه‌های آموزشی که به عنوان بردار پشتیبان در نظر گرفته می‌شوند کمترین فاصله تا مرز جداسازی دو کلاس را دارند و با استفاده از آنها یک مرز جداسازی خطی بهینه برای جداسازی دو کلاس بدست می‌آید.



شکل ۷: شکل کلی ماشین بردار پشتیبان SVM

با توجه به اینکه تعداد کلاس‌ها در این نوع دیتا بسیار است، بنابراین باید از یکی از روش‌های مناسب ماشین بردار پشتیبان استفاده شود. یکی از بهترین روش‌های پیشنهادی، استفاده از روش یکی در برابر همه می‌باشد. در این روش به تعداد کلاسها SVM وجود دارد در هر SVM یک کلاس با کل کلاس‌های دیگر بعنوان SVM دو کلاسی در نظر گرفته می‌شود. از این روش در انجام آزمایشات این مقاله استفاده شده است.



شکل ۸: روش یکی در مقابل همه در SVM

۴- نتایج

۴-۱- ارزیابی روش پیشنهادی

برای این کار از داده های موجود در پایگاه داده دستنویس هدی استفاده شده است. داده های پایگاه داده شامل ۶۰۰۰۰ داده آموزش، ۲۰۰۰۰ داده تست و همچنین ۲۲۳۵۰ نمونه متفرقه است که در این مقاله از آنها استفاده شده است. توزیع ارقام در این پایگاه داده در جدول (۱) نشان داده شده است.

جدول ۱: توزیع ارقام در پایگاه داده هدی

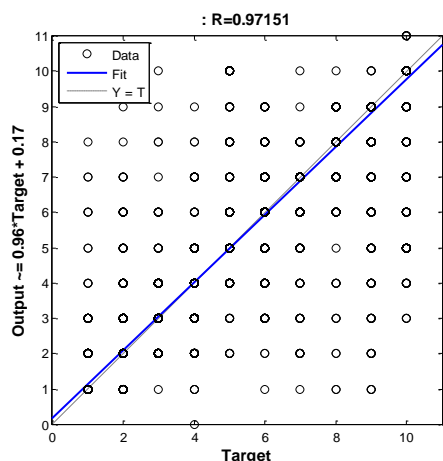
۰	۱	۲	۳	۴	۵	۶	۷	۸	۹
۱۰۰۷۰	۱۰۳۳۰	۹۹۲۳	۱۰۳۳۴	۱۰۳۳۳	۱۰۱۱۰	۱۰۲۵۴	۱۰۳۶۳	۱۰۲۶۴	۱۰۳۷۱

متاسفانه تصاویر موجود در پایگاه داده هدی دارای سایزهای مختلفی است و به همین دلیل، قبل از انجام هر کاری باید سایز تمام دیتاها را یکسان کرد. بنابراین ما برای راحتی سایز هر یک از اعداد را 28×28 در نظر می گیریم. در نتیجه تمام ارقام موجود در پایگاه داده هدی به مربع هایی با این اندازه تبدیل شدند. در تغییر اندازه مربع ها از روش نزدیک ترین همسایه استفاده شده است تا تصویر دقیقاً سیاه و سفید باقی بماند و خطای ناشی از تغییر اندازه به روش درون یابی وارد نشود. پس از این مرحله ما با استفاده از یک الگوریتم تبدیل موجک دو بعدی، ویژگی های هر یک از داده ها را استخراج کردیم. اما با توجه به تعداد ویژگی های زیاد استخراج، از الگوریتم تحلیل مولفه اساسی (PCA) جهت کاهش و انتخاب ویژگی استفاده شد. در نهایت جهت طبقه بندی از ماشین بردار پشتیبان (SVM) باینری استفاده شده است. نتایج حاصل از این الگوریتم طبقه بندی در جدول (۲) نشان داده شده است.

جدول ۲: دقت هر یک از کلاس ها با استفاده از ماشین بردار پشتیبان

class	0	1	2	3	4	5	6	7	8	9
A*	1856	1961	2019	2143	1910	2119	1973	2087	2042	1889
B**	%91.20	%95.70	%89.10	%91.50	%88.25	%95.20	%89.90	%93.30	%93.35	%89.75

در جدول فوق، A مقدار پیش بینی شده هر رقم در پایگاه داده است و B دقت هر کلاس می باشد. با توجه به جدول فوق، نمودار رگرسیون نتایج را می توان به صورت شکل (۸) نشان داد.



شکل ۹: نمودار رگرسیون خطی در SVM

۲-۴- نتایج

در این مقاله از یک سیستم بازشناسی الگو جهت تشخیص ارقام دستنویس فارسی استفاده شد. در این مقاله از الگوریتم تبدیل موجک و تحلیل مولفه اساسی جهت استخراج و انتخاب ویژگی استفاده شد. با توجه به اینکه تعداد کلاسها در این پروژه ۱۰ کلاس است، بنابراین از یک ماشین بردار پشتیبان باینری و از نوع یکی در مقابل همه استفاده شده است. نتایج بررسی در این پژوهش نشان می دهد ماکزیمم دقت طبقه بندی ارقام دستنویس فارسی حدود ۹۱,۷۵ درصد می باشد. همچنین الگوریتم پیشنهادی با تعدادی از روشها مانند شبکه پرسپترون چند لایه (MLP) و همچنین الگوریتم KNN مقایسه شد. نتایج نشان داد میزان دقت طبقه بندی ارقام دستنویس فارسی به ازای شبکه عصبی پرسپترون چند لایه ۹۰,۲۵ درصد و به ازای الگوریتم KNN در حدود ۹۱,۰۰ درصد است.

منابع و مراجع

- [1] Parhami, B. and M. Taraghi, Automatic recognition of printed Farsi texts. *Pattern Recognition*, 14(1): p. 395-403, 1981.
- [2] G. A. Abandah, F. Jamour, “ Recognizing handwritten Arabic words using grapheme segmentation and recurrent neural networks”, *IJDAR manuscript No.2014*
- [3] K.Borna, V.Hashemi“*A Fast Algorithm for Persian Handwritten Number Recognition with Computational Geometry Techniques*” ".*ACSII Advances in Computer Science: an International Journal*, Vol. 3, Issue 3, And No.9 -ISSN: 2322-5157, May 2014.
- [4] J.Sadri, y.Akbari, M.Jalili, A.Farahi, M.Habibi, A New System for Recognition of Handwritten Persian Bank Checks,in: *International Conference on Document Analysis and Recognition*, 2011.
- [5] S.Zarezade, A.Akkasi, “Recognition of Farsi Letter using Hidden Markov Model”, *International Journal of Computer Applications (0975 – 8887) Volume 125 – No.1, September 2015*.
- [6] A.Ramzi, A. Zahary, “Online Arabic Handwritten Character Recognition using Online-Offline Feature Extraction and Back Propagation Neural Network” *1st International Conference on Advanced Technologies for Signal and Image Processing - ATSIP'2014 March 17-19, 2014*.
- [7] S.M. Saad., Application of fuzzy logic and genetic algorithm in biometric text-independent writer Identification, *IET Information Security.vol.5, no.1, pp.1-9, 2011*.
- [8] Shivram, A.; Ramaiah, C.; Porwal, U.; Govindaraju, V., "Modeling Writing Styles for Online Writer Identification: A Hierarchical Bayesian Approach," *Frontiers in Handwriting Recognition (ICFHR)*, 2012 *International Conference on* , vol., no., pp.387,392, 18-20 Sept. 2012.
- [9] Zhang Youzhi,WangJin,”Hidden Markov Model with parameter optimized K-Means clustering for handwriting recognition”, *Internet Computing & Information Services(ICICIS)*, pp. 435–438, Sept 2011.
- [10] Z. Imani, A.R. Ahmadyfard and A. Zohrevand, “Holistic Farsi handwritten word recognition using gradient features”, *Journal of AI and Data Mining* Published online: August 2015.
- [11] R. B. Yadava, N. K. Nishchala, A. K. Gupta, and V. K. Rastogi, “Retrieval and classification of shape-based objects using Fourier, generic Fourier, and wavelet-fourier descriptors technique: A comparative study,” *Optics and Lasers in Engineering*, vol. 45(6), pp. 695-708, 2007.