مهندسی صنایع و مدیریت شریف، (تابستان ۱۴۰۲)
دوره‌ی ۳۹، شماره‌ی ۱، صص. ۲۳-۳۴، (پژوهشی)

Original Article

توسعه‌ی الگوریتم بازشناسی اعداد دست‌نویس فارسی، بر پایه‌ی الگوریتم‌های طبقه‌بندی شبکه‌ی عصبی چندلایه و احتمالاتی، به کمک مراکز خوشه

علی میری (دانشجوی کارشناسی ارشد)
مجید خدمتی* (دانشیار)
دانشکده‌ی مهندسی صنایع، دانشگاه صنعتی شریف

در این پژوهش تلاش شده است تا با ارائه‌ی الگوریتمی بهبودیافته و مبتنی بر خوشه‌بندی، بازشناسی اعداد دست‌نویس فارسی با دقت قابل توجهی صورت پذیرد. بر این اساس، آموزش و بازشناسی الگوها به کمک شبکه‌ی عصبی احتمالاتی و چندلایه‌ی پرسپترون میسر شده است، به این صورت که پس از استخراج دو دسته ویژگی مکان مشخصه و ناحیه‌یی از داده‌های آموزشی، داده‌های هریک از کلاس‌های دهگانه بر اساس هر ویژگی با استفاده از روش‌های پیوند کامل، PAM و FCM خوشه‌بندی شده و کلاس‌های دهگانه‌ی جدید حاصل از خوشه‌بندی، توسط یکی از دو الگوریتم طبقه‌بندی کننده آموزش می‌بینند. تعداد بهینه خوشه‌های هر کلاس با استفاده از الگوریتم بهینه‌سازی جست‌وجوی ممنوعه با تابع برازندگی نرخ بازشناسی تعیین می‌شود. میزان دقت الگوریتم در نهایت با استفاده از داده‌های آزمایش مورد سنجش قرار می‌گیرد و با توجه به نتایج ملاحظه می‌شود که الگوریتم پیشنهادی، بازشناسی اعداد دست‌نویس فارسی را با دقت بالایی انجام می‌دهد.

واژگان کلیدی: خوشه‌بندی، شبکه‌ی عصبی چندلایه، شبکه‌ی عصبی احتمالاتی، بازشناسی، جست‌وجوی ممنوعه.

ali.miri@ie.sharif.edu
khedmati@sharif.edu

۱. مقدمه

بازشناسی حروف و اعداد دست‌نویس شاخه‌یی از بازشناسی الگوست که تحقیقات گسترده‌یی را به خود اختصاص داده و تا به امروز نیز در حال تحول است. تحقیقات در این زمینه از حدود ۶۰ سال پیش آغاز شده و هم اکنون در کاربردهای مختلفی همچون بازشناسی آدرس‌های پستی، خوانش خودکار چک‌ها و فرم‌های بانکی، ثبت اسناد و مدارک و ... از آن استفاده می‌شود. روش‌های متفاوتی برای بازشناسی اعداد و ارقام در پژوهش‌ها مورد استفاده قرار گرفته است که عمدتاً در یکی از دسته‌های روش‌های آماری، ساختاری، شبکه‌ی عصبی و فازی قرار می‌گیرند.[۱]

اعداد و ارقام در تمامی زبان‌ها رسم‌الخطی اختصاصی دارند و با وجود تفاوت‌های اعداد دهگانه، شباهت‌هایی نیز میان آنها قابل مشاهده است که این امر موجب خطای سیستمی و انسانی در خوانش و بازشناسی اعداد می‌شود. بنابراین، برای کاهش خطاهای مذکور، حرکت به سوی اتوماسیون و تشخیص دقیق اعداد مورد توجه پژوهشگران قرارگرفته و روش‌های مختلفی برای بازشناسی اعداد دست‌نویس فارسی پیشنهاد شده است.

در بازشناسی اعداد دست‌نویس فارسی یکی از چالش‌های اصلی تخصیص مشاهدات به کلاس‌های دهگانه است چرا که با توجه به شباهت‌های بین کلاسی و تنوع درون کلاسی در رسم‌الخط‌های مختلف، توسعه‌ی الگوریتم‌های طبقه‌بندی و بازشناسی که دقت قابل قبول و کارایی زمانی مناسبی داشته باشند، برای محققین دشوار است. بر این اساس اغلب روش‌های ارائه شده قادر نیستند دقت بالا در بازشناسی الگو و کاهش زمان جست‌وجو را به صورت همزمان ارائه دهند.

در این پژوهش تلاش شده است تا با پرداختن به هر دو چالش مطرح شده، الگوریتمی با کارایی زمانی بالا و دقت مناسب برای بازشناسی اعداد دست‌نویس فارسی ارائه شود. در این راستا پس از استخراج دو دسته ویژگی متفاوت از داده‌های هر کلاس، به کمک روش‌های خوشه‌بندی و طبقه‌بندی کننده‌های شبکه‌ی عصبی

* نویسنده مسئول
تاریخ: دریافت ۱۴۰۰/۶/۱، اصلاحیه ۱۴۰۱/۴/۱۲، پذیرش ۱۴۰۱/۴/۲۱.
DOI:10.24200/J65.2022.58031.2220

استناد به این مقاله:
میری، علی و خدمتی، مجید، ۱۴۰۲. توسعه‌ی الگوریتم بازشناسی اعداد دست‌نویس فارسی، بر پایه‌ی الگوریتم‌های طبقه‌بندی شبکه‌ی عصبی چندلایه و احتمالاتی، به کمک مراکز خوشه. مهندسی صنایع و مدیریت شریف، ۳۹(۱)، صص. ۲۳-۳۴.

چندلایه و احتمالاتی، مراکز خوشه‌های بهینه در هرکلاس تعیین می‌شود. پس از آن، از مراکز خوشه و یک طبقه‌بندی کننده مبتنی بر شبکه‌های عصبی، برای آموزش الگوریتم بازشناسی استفاده می‌شود و الگوریتم مورد ارزیابی قرار می‌گیرد. از مزایای روش پیشنهادی می‌توان به موارد زیر اشاره کرد:

- الگوریتم‌های طبقه‌بندی کننده‌ی شبکه‌ی عصبی، به خصوص شبکه‌ی عصبی احتمالاتی و چندلایه، قدرت تشخیص بالایی دارند و از قدرت‌مندترین روش‌های آموزش نظارت شده محسوب می‌شوند که قادرند بیشترین تطابق‌سنجی میان داده‌های آموزش و آزمایش را ایجاد کنند؛
- شبکه‌های عصبی با وجود توانمندی بالای تشخیصی، در مواجهه با حجم زیاد داده‌ها با چالش‌هایی مواجه می‌شوند. لذا استفاده از مراکز خوشه به جای اصل داده‌ها، راه حل مناسبی برای مقابله با این مشکلات را فراهم می‌کند؛
- مراکز هر خوشه به عنوان نماینده و میانگین ویژگی‌های هر خوشه محسوب می‌شوند که می‌تواند آنها را به جایگزین‌های مناسبی برای داده‌ها تبدیل کند؛
- استفاده از الگوریتم فرااِبتکاری قوی جست‌وجوی ممنوعه در کنار شبکه‌های عصبی کمک می‌کند تا خوشه‌بندی داده‌های هرکلاس با توجه به چالش‌های آن دسته، در حالت بهینه صورت پذیرد.

۲. مرور ادبیات

در زمینه‌ی بازشناسی اعداد دست‌نویس فارسی تلاش‌ها و پژوهش‌های بسیاری صورت گرفته است. عمده‌ی این تلاش‌ها را می‌توان در دو دسته: ۱. تلاش برای بهبود طبقه‌بندها؛ ۲. شناسایی و استخراج ویژگی‌ها از داده‌های در دسترس، دسته‌بندی کرد. پژوهشگران در تلاش‌اند تا ویژگی‌هایی را از داده‌ها استخراج کنند که بیشترین سازگاری و همخوانی را با الگوریتم طبقه‌بندی کننده مورد استفاده داشته باشد.

در این راستا جوهری و همکاران[۲] بازشناسی اعداد دست‌نویس فارسی را به کمک روش‌های فازی انجام داده‌اند. خلیلی و امیری[۳] از طبقه‌بندی شبکه‌ی عصبی چندلایه برای بازشناسی اعداد دست‌نویس استفاده کرده‌اند. درویش و همکاران[۴] الگوریتم تطابق شکل و طبقه‌بندی نزدیک‌ترین فاصله از نماینده یا نماینده‌های هر کلاس را برای بازشناسی اعداد دست‌نویس فارسی به کار برده‌اند. همچنین نحوی و همکاران[۵] با استفاده از شبکه‌ی عصبی و ترکیب ده الگوریتم طبقه‌بندی دوکلاسی، راهکاری برای بهبود نرخ بازشناسی اعداد دست‌نویس فارسی ارائه کرده‌اند. خسروی و همکاران[۶] نیز با تمرکز بر شبکه‌های عصبی و ادغام ویژگی‌ها، از طبقه‌بندی‌کننده‌ی شبکه‌ی عصبی و ترکیب سه ویژگی مکان مشخصه، گرادیان بهبودیافته و کریش، الگوریتمی برای بهبود نرخ بازشناسی اعداد توسعه دادند.

در راستای بهبود نرخ بازشناسی اعداد دست‌نویس فارسی و عربی، بوچنه و همکاران[۷] به کمک نسخه‌ی بهبودیافته از هیستوگرام گرادیان جهت‌یافته به عنوان توصیف‌گر ویژگی و راهکار کاهش ابعاد مبتنی بر تحلیل مؤلفه‌ی اصلی تلاش کردند تا دقت بازشناسی اعداد را بهبود دهند. خراشادی‌زاده و همکاران[۸] با استفاده از یک مجموعه ویژگی ثابت و کارآمد با ترکیب چهار هیستوگرام کد زنجیره‌یی جهت‌دار و هیستوگرام گرادیان جهت‌دار به کمک ماشین بردار پشتیبان (SVM) [۱] بهبود قابل توجهی در نرخ بازشناسی کسب کردند.

گلزاری و همکاران[۹] نیز با استفاده از پیش‌پردازش تصاویر، استفاده از ویژگی‌های سطح بالا و پایین و در نهایت ماشین بردار پشتیبان روشی جدید در راستای بهبود نرخ بازشناسی ارقام دست‌نویس ارائه دادند.

در این مسیر استفاده از روش‌های مبتنی بر هوش مصنوعی همچون یادگیری عمیق نیز مورد توجه پژوهشگران قرار گرفته است. فرخی و همکاران[۱۰] با استفاده از ماشین بولتزمن عمیق و دولایه‌ی رمزگذار خودکار که در هر لایه از ۲۰۰ عدد نورون استفاده شده، الگوریتمی ارائه دادند که با دقت ۹۳٪ اعداد دست‌نویس را بازشناسی می‌کند. اما چنان که بیان شد روش‌ها و الگوریتم‌های یاد شده عمدتاً بر ویژگی‌ها و طبقه‌بندی کننده‌های متفاوت و ترکیب این دو متمرکز بوده‌اند. دو مورد از ایرادات اصلی روش‌های ارائه شده این است که اولاً با وجود دقت بالای بازشناسی در برخی الگوریتم‌ها، زمان اجرای آنها بیشتر است، دوماً الگوریتم‌هایی که با استفاده از روش‌هایی همچون کاهش داده‌ها و تغییر ویژگی زمان اجرای الگوریتم را کاهش داده‌اند، بهبود مناسبی در دقت بازشناسی کسب نکرده‌اند. لذا دقت بازشناسی و زمان اجرای الگوریتم در این روش‌ها با یکدیگر تناسب ندارند.

در این راستا برخی از پژوهشگران تلاش کردند تا در راستای حل این چالش به کمک پیشرفته‌ترین طبقه‌بندها بهبودهایی حاصل کنند. دهقانیان و قدس[۱۱] با توسعه‌ی شبکه‌ی عصبی کانولوشن (CNN) [۲] علی‌رغم کاهش ۵۰٪ داده‌های ورودی و کاهش زمان آموزش و آزمایش، فقط ۱٫۵٪ نرخ بازشناسی را بهبود دادند. حقیقی و همکاران[۱۲] نیز تلاش کردند در پژوهش خود با ارائه‌ی الگوریتمی مبتنی بر شبکه‌ی عصبی کانولوشن و حافظه‌ی بلندمدت دوجهته، ضمن استفاده از مفهوم بردار احتمال کلاس تصاویر، دقت مدل یادگیری عمیق خود را بهبود بخشند.

خوشه‌بندی به عنوان یک ابزار کاربردی دیگر داده‌کاوی، از اواخر قرن بیستم وارد ادبیات موضوع بازشناسی الگو شد. این مفهوم در ابتدا برای یادگیری‌های نظارت نشده در مقایسه با روش‌های طبقه‌بندی و نظارت شده مورد استفاده قرار گرفت اما مشخص شد که می‌توان با کنار هم قرار دادن این دو سیستم یادگیری، بهبودهای قابل توجهی مشاهده کرد که نمونه‌های استفاده از آن در مطالعات موجود[۱۳-۱۵] قابل بررسی‌اند. کاربرد روش‌های خوشه‌بندی در بازشناسی اعداد و کاراکترهای فارسی در پژوهش ابراهیمی و همکاران[۱۶] برای بازشناسی زیرکلمات چاپی مورد استفاده قرار گرفت. همچنین در حوزه‌ی بازشناسی ارقام فارسی، تقی‌پور و همکاران[۱۷] روشی ترکیبی برای بهبود بازشناسی ارقام دست‌نویس ارائه دادند. میری و همکاران[۱۸] نیز در پژوهش خود الگوریتمی مبتنی بر شبکه‌های عصبی و خوشه‌بندی ارائه دادند که نتایج حاصله بیان‌گر بهبود نتایج در الگوریتم‌های ترکیبی است. اما چنان که بیان شد در روش‌های توسعه‌یافته چالش‌هایی وجود دارد که در این پژوهش تلاش می‌شود با ارائه‌ی الگوریتمی نوآورانه در راستای حل و بهبود آنها حرکت شود.

۳. روش‌شناسی پژوهش

با توجه به مشاهده‌ی بهبود اولیه در نرخ بازشناسی اعداد دست‌نویس فارسی به کمک خوشه‌بندی در پژوهش میری و خدمتی[۱۸] و با هدف اثبات بهبود در نرخ بازشناسی در صورت استفاده از خوشه‌بندی، الگوریتمی در دوازده حالت با استفاده از دو روش طبقه‌بندی شبکه‌ی عصبی احتمالاتی و چندلایه و سه روش خوشه‌بندی کاملاً پیوسته پیوند کامل،[۳] PAM [۴] و FCM،[۵] مبتنی بر ویژگی‌های مکان مشخص شده و ناحیه‌یی در این مطالعه توسعه داده شد تا با مقایسه‌ی نرخ‌های بازشناسی، در کنار تعیین دقیق‌ترین الگوریتم شناساگر، میزان بهبود ایجاد شده توسط الگوریتم‌ها نسبت به حالت بدون خوشه‌بندی مورد ارزیابی قرار گیرد.

پایگاه داده‌ی مورد استفاده در این تحقیق پایگاه داده‌ی هدی[۱۹] است که با جمع‌آوری نمونه‌های استاندارد پیش‌پردازش شده از اعداد دست‌نویس فارسی، مورد

استفاده‌ی بسیاری از پژوهشگران قرار گرفته است. تعداد کل تصاویر این پایگاه داده ۱۰۲۳۵۲ رقم بوده که ۸۰۰۰۰ تصویر در دو بخش آموزشی و آزمایشی تفکیک شده است. در این پژوهش داده‌های آموزشی شامل ۶۰۰۰۰ داده‌یی که برای آموزش در نظر گرفته شده است و ۲۰۰۰۰ داده نیز برای تصدیق الگوریتم و تعیین نرخ بازشناسی (داده‌های آزمایش) استفاده می‌شود.

در ادامه، در بخش ۴ مبانی نظری و در بخش ۵، الگوریتم کلی راهکارهای توسعه یافته بر اساس سه روش خوشه‌بندی و دو روش طبقه‌بندی تشریح می‌شود. در بخش ۶، نرخ‌های بازشناسی استخراج شده است. نهایتاً در بخش ۷، نتایج الگوریتم مورد بررسی قرار می‌گیرد و با سایر پژوهش‌ها مقایسه می‌شود؛ پیشنهادهای پژوهش‌های آتی نیز مورد بحث قرار می‌گیرد.

۴. مبانی نظری

در این بخش برخی از اصطلاحات رایج در ادبیات موضوع تشریح شده و مفاهیم مرتبط با پژوهش جاری تبیین می‌شود.

۱.۴. خوشه‌بندی

خوشه‌بندی[6] یکی از شاخه‌های یادگیری بدون نظارت[7] است که در یک فرایند تکراری، داده‌ها به دسته‌هایی که اعضای آن مشابه یکدیگرند تقسیم می‌شوند. به هریک از این دسته‌ها به صورت کلی خوشه[8] گفته می‌شود لذا هر خوشه مجموعه‌یی از اشیاء بوده که در آن هر شی با هم‌گروهی خود مشابهت بالایی داشته و با اشیاء موجود در خوشه‌های دیگر کم‌ترین مشابهت را دارد. برای مشابه بودن می‌توان معیارهای مختلفی را در نظر گرفت از جمله معیار فاصله که خود با روش‌های متفاوتی از جمله اقلیدسی، منهتن، مجذور اقلیدسی و ... اندازه‌گیری می‌شود.

تجزیه و تحلیل خوشه‌یی خود یک الگوریتم خاص نیست، بلکه روند کلی است و می‌تواند توسط الگوریتم‌های مختلفی به دست آید که هر روش درک متفاوتی از ماهیت خوشه ارائه می‌کند.[20] روش‌های خوشه‌بندی از جنبه‌های مختلفی دسته‌بندی شده و هر روش طبیعتاً در چند دسته می‌تواند قرار گیرد. از جمله‌ی این دسته‌ها، خوشه‌بندی مبتنی بر توزیع داده‌ها، اتصال داده‌ها، تراکم داده‌ها یا قطعیت عضویت در خوشه، عضویت احتمالی، فازی و ... است.

در این پژوهش روش‌های خوشه‌بندی پیوند کامل و PAM به عنوان روش‌های خوشه‌بندی قطعی و روش FCM به عنوان روش خوشه‌بندی غیرقطعی فازی مورد بررسی قرار می‌گیرند. این سه روش با توجه به ماهیت پژوهش که بررسی اثرگذاری خوشه‌بندی در بهبود نرخ بازشناسی است به علت عضویت در چندین دسته متفاوت انتخاب شده و سایر روش‌ها نیز می‌تواند در پژوهش‌های آتی مورد بررسی قرار گیرد.

۱.۱.۴. خوشه‌بندی پیوند کامل

یکی از مدل‌های قطعی خوشه‌بندی، روش‌های سلسله‌مراتبی هستند. این روش بر مبنای ایده‌ی اصلی اشیائی است که بیشتر مربوط به اشیاء نزدیک، نسبت به اشیاء دورتر است. این الگوریتم‌ها «اشیاء» را برای ایجاد «خوشه‌ها»، بر اساس فاصله‌ی آنها به یکدیگر متصل می‌کنند. خوشه را می‌توان به طور کلی، با بیشترین فاصله مورد نیاز برای اتصال اشیاء یک خوشه توصیف کرد. در فاصله‌های مختلف، خوشه‌های متفاوتی شکل می‌گیرد که می‌توان آنها را به کمک یک دندروگرام نشان داد. این نمودار با توجه به ساختار بالا به پایینی که دارد، توصیفی دقیق برای مفهوم خوشه‌بندی سلسله‌مراتبی ارائه می‌کند.

این الگوریتم‌ها یک ساختار سلسله‌مراتبی از خوشه‌هایی که در فاصله‌های معینی با یکدیگر ادغام می‌شوند، ارائه می‌دهند. در یک دندروگرام، محور y نشان‌دهنده‌ی فاصله‌یی است که خوشه‌ها ادغام می‌شوند، در حالی که اشیاء در امتداد محور x قرار می‌گیرند به طوری که خوشه‌ها با هم مخلوط نمی‌شوند. خوشه‌بندی سلسله مراتبی شامل دو نوع خوشه‌بندی است:

۱. پیوند منفرد:[9] در این روش که به پایین به بالا[10] نیز معروف است، ابتدا هر داده به عنوان یک خوشه در نظر گرفته می‌شود. در ادامه با به کارگیری یک الگوریتم، هر بار خوشه‌های دارای ویژگی‌های نزدیک به هم با یکدیگر ادغام می‌شوند و این کار تا تشکیل چند خوشه‌ی مجزا ادامه می‌یابد. این روش به نویز و مصرف زیاد حافظه حساس است؛

۲. پیوند کامل: در این روش که به روش بالا به پایین[11] نیز معروف است،ابتدا تمام داده‌ها به عنوان یک خوشه در نظر گرفته و با به کارگیری یک الگوریتم تکرار شونده، هر بار داده‌یی که کم‌ترین شباهت را با داده‌های دیگر دارد به خوشه‌های مجزا تقسیم می‌شود. این کار تا تشکیل یک یا چند خوشه یک عضوی ادامه می‌یابد و مشکلات روش قبل را ندارد.[21]

با توجه به این موارد، داده‌های مرتبط با هر دسته‌ی عددی به کمک الگوریتم خوشه‌بندی لینک‌کامل، خوشه‌بندی می‌شود. برای محاسبه‌ی شباهت بین دو خوشه‌ی A و B، از معیار رابطه‌ی ۱ استفاده می‌شود:

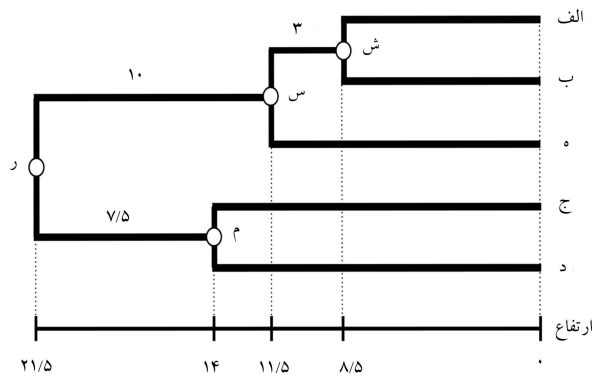
$$d_{AB} = Max_{i\in A, j\in b} d_{ij} \qquad (۱)$$

در شکل ۱ نمونه‌یی از دندروگرام خوشه‌بندی با روش پیوند کامل نشان داده شده است.[22]

۲.۱.۴. خوشه‌بندی PAM

این روش خوشه‌بندی با نام k-medoids (نمایندگان خوشه‌ها) نیز شناخته شده و از روش‌های نرم و مسطح است که نسخه‌ی ارتقاء یافته‌ی الگوریتم k-means نیز محسوب می‌شود. در الگوریتم PAM، به جای استفاده از میانگین اعضای خوشه، از خود نمونه‌ها برای مرکز ثقل و نمایندگی خوشه‌ها استفاده می‌شود. با انتخاب نمونه‌های واقعی برای نمایش یک خوشه، حساسیت روش نسبت به نمونه‌های نویز و خارج از محدوده کاهش می‌یابد.

در این الگوریتم فاصله‌ی هریک از اعضای خوشه نسبت به مرکز خوشه، که خود یکی از این اعضاست، سنجیده می‌شود و داده‌یی به عنوان مرکزیت خوشه



شکل ۱. دندروگرام خوشه‌بندی با روش لینک کامل.[22]

انتخاب می‌شود که کم‌ترین فاصله را با سایر داده‌های موجود در خوشه داشته باشد. تابع هدف این روش در رابطه‌ی ۲ ذکر شده است:

$$j = \sum_{j=1}^{k} \sum_{i=1}^{n} \left\| x_i^{(j)} - o_j \right\|^2 \quad (۲)$$

|| || معیار فاصله‌ی بین نقاط و o_j مرکز خوشه‌ی jام است. الگوریتم با طی کردن مراحل زیر، داده‌ها را خوشه‌بندی می‌کند:

۱. در ابتدا k نقطه به عنوان نقاط مراکز خوشه‌ها انتخاب می‌شود؛

۲. هر نمونه داده، به خوشه‌یی که مرکز آن خوشه کم‌ترین فاصله تا آن داده را دارد، نسبت داده می‌شود؛

۳. پس از تعلق تمام داده‌ها به یکی از خوشه‌ها، برای هر خوشه یکی دیگر از داده‌ها به عنوان مرکز خوشه در نظر گرفته شده و تابع هدف برای آن محاسبه می‌شود. مراحل ۲ و ۳ تکرار می‌شوند تا زمانی که دیگر هیچ تغییری در مراکز خوشه‌ها حاصل نشود.

۳.۱.۴. خوشه‌بندی فازی (FCM)

در خوشه‌بندی فازی به جای تعیین وقوع یا عدم وقوع رویداد (مانند مورد با احتمال)، فازی‌سازی درجه‌یی وقوع یک واقعه، اندازه‌گیری می‌شود. هدف خوشه‌بندی سنتی اختصاص هر نقطه داده به یک و فقط یک خوشه است در حالی که خوشه‌بندی فازی، درجه‌های مختلف عضویت را برای هر نقطه اختصاص می‌دهد. عضویت یک نقطه مبتنی بر مقدار اشتراکی روی خوشه‌های مختلف است.

در حالتی که داده‌ها ماهیت پارامتری داشته باشند، این الگوریتم برای کمینه‌سازی یک تابع هزینه تلاش می‌کند. تابع هزینه موجود در رابطه‌ی ۳، نسبت به u (یک افراز k تایی فازی از مجموعه داده) و c (یک مجموعه از k نمونه (مراکز خوشه‌ها)) کمینه شده است:

$$J_q(u,c) = \sum_{j=1}^{m} \sum_{i=1}^{k} u_{ij} d^2(x_j, c_i); K \leq N \quad (۳)$$

که در این رابطه:

q: یک عدد حقیقی بزرگ‌تر از ۱؛

x_j: j امین بردار ویژگی n بعدی؛

x_i: مرکز خوشه‌ی iام؛

u_{ij}: درجه عضویت x_j به iامین خوشه؛

$d^2(x_j, c_i)$: فاصله بین x_j و c_i؛

m: تعداد نقاط داده؛

k: تعداد خوشه‌ها است.

پارامتر q وزن تعریف کننده‌ی u_{ij} و کنترل درجه فازی بودن نتایج خوشه‌بندی است؛ الگوریتم روش خوشه‌بندی فازی عبارت است از:

۱. انتخاب مراکز اولیه؛

۲. محاسبه‌ی درجه‌ی عضویت همه‌ی بردارهای ویژگی در همه‌ی خوشه‌ها با استفاده از رابطه‌ی ۴؛

$$u_{ij} = \frac{\frac{1}{d^2(x_j, c_j)}^{\frac{1}{q-1}}}{\sum_{k=1}^{K} \frac{1}{d^2(x_j, c_j)}^{\frac{1}{q-1}}} \quad (۴)$$

۳. محاسبه‌ی مراکز جدید و به‌روزرسانی درجه‌های عضویت با استفاده از رابطه‌ی ۵:

$$\widehat{c}_i = \frac{\sum_{j=1}^{M} u_{ij}{}^q x_j}{\sum_{i=1}^{M} u_{ij}{}^q} \quad (۵)$$

۴. اگر $Max \|u_{ij} - \widehat{u_{ij}}\| \leq tol$ توقف و در غیر این‌صورت بازگشت به گام ۳.

در رابطه‌ی فوق، tol حد نهایی هم‌گرایی درجه‌ی عضویت خوشه‌هاست که رسیدن به آن، معیار پایان الگوریتم است. ذکر این نکته نیز حائز اهمیت است که محاسبه‌ی درجه‌ی عضویت به تعریف اندازه‌ی فاصله بستگی دارد.[۲۳]

۲.۴. استخراج ویژگی

انتخاب روش استخراج ویژگی، یکی از مراحل اثرگذار در یک سیستم بازشناسی الگو است. برای بازشناسی اعداد دست‌نویس فارسی، از ویژگی‌های متفاوتی همچون ویژگی‌های ناحیه‌یی،[۱۲] گشتاورهای هندسی، گشتاورهای زرنیکی، ویژگی‌های مکان‌مشخصه،[۱۳] کریش، پروفایل مرزی و ... می‌توان استفاده کرد. در این پژوهش، ویژگی‌های ناحیه‌یی و مکان مشخصه از تصاویر استخراج شده و برای بازشناسی الگو مورد استفاده قرار می‌گیرند.

۱.۲.۴. ناحیه‌بندی (زونینگ یا بلوک‌بندی)

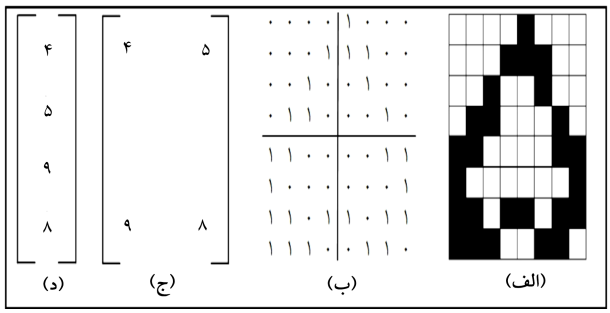
در تصاویر باینری اعداد یا حروف، صفرهای زیادی (یعنی نقاط سفید که مربوط به زمینه هستند و حاوی هیچ اطلاعاتی نیستند) در چهار طرف تصویر وجود دارند. برای استخراج ویژگی ناحیه‌بندی، در ابتدا چهار وجه که حاوی اطلاعات زائد هستند حذف می‌شوند و هریک از اعداد محصور به یک چهار وجهی می‌شود. در این تحقیق، اعداد از نظر اندازه به صورت ۳۲ × ۳۲ نرمالیزه شدند. منظور از نرمال‌سازی، تبدیل تصویر ورودی با هر اندازه‌یی به تصویری با ابعاد از پیش تعیین شده $m \times m$ است.

روش ساده و مرسوم برای نرمال‌سازی تصویر به این صورت است که تصویر از هر دو جهت طولی و عرضی به اندازه مورد نظر تغییر مقیاس داده شود. برای کاهش تعداد ویژگی‌ها، تصویر ۳۲ × ۳۲ به بلوک‌های کوچک‌تری تقسیم می‌شود و نقاط داخل هر بلوک شمرده و به عنوان ویژگی در نظر گرفته می‌شوند. به طور مثال اگر بلوک‌ها به صورت ۴ × ۴ انتخاب شوند، تصویر دارای ۶۴ بلوک ۱۶ نقطه‌یی است. تعداد نقاط سیاه موجود در هر بلوک ۴ × ۴ به عنوان ویژگی‌های سیستم بازشناسی در نظر گرفته می‌شوند. در این حالت هر رقم ۶۴ زیر ویژگی دارد که مقادیر این ویژگی‌ها می‌تواند بین صفر تا ۱۶ تغییر کند.[۲]

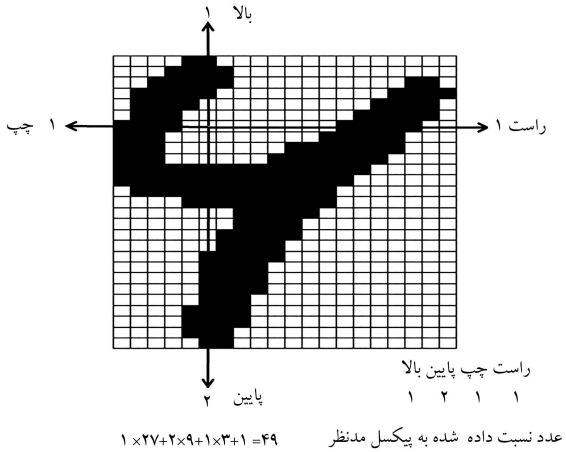
برای مثال در شکل ۲ الف، تصویر نرمالیزه شده رقم ۵ با اندازه ۸ × ۸ قابل مشاهده است. این تصویر به ۴ بلوک ۴ × ۴ تقسیم شده که با شمردن تعداد نقاط سیاه در هر بلوک یا یک‌های متناظر آنها در شکل ۲ ب، مقادیر ویژگی‌ها به صورت یک ماتریس ۲ × ۲ بدست می‌آید که در شکل ۲ ج قابل ملاحظه است. این ماتریس را می‌توان به صورت یک بردار ۴ عنصری نمایش داد و از آن به عنوان بردار ویژگی (شکل ۲ د) استفاده کرد.

۲.۲.۴. ویژگی مکان مشخصه

ویژگی‌های مکان مشخصه معمولاً در راستاهای عمودی، افقی و یا زوایای ۴۵ و ۱۳۵ درجه تعریف می‌شود. بردار ویژگی مکان‌های مشخصه برای هر تصویر به این صورت محاسبه می‌شوند که به هر نقطه از زمینه تصویر، یک عدد نسبت داده که این



شکل ۲. استخراج ویژگی ناحیه‌بندی.[۲]



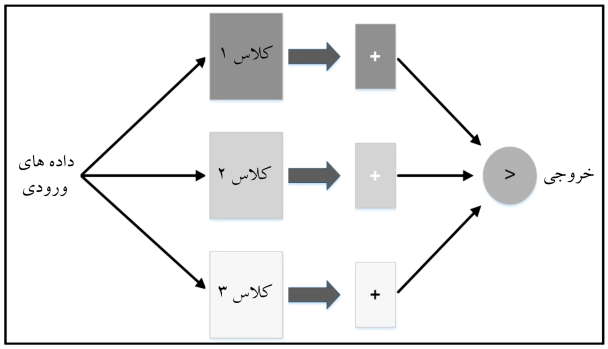
شکل ۳. محاسبه‌ی ویژگی مکان مشخصه برای یک نقطه از تصویر.

عدد با توجه به این‌که خطوط عمودی و افقی رسم شده از آن نقطه در جهت‌های چهارگانه بالا، پایین، راست و چپ، بدنه عدد را در چند نقطه قطع کنند به دست می‌آید. اگر تعداد تقاطع‌ها با بدنه رقم به ۲ محدود شود، می‌توان از یک عدد چهار رقمی در مبنای ۳ برای نمایش نقاط مختلف نواحی زمینه استفاده کرد. بردار ویژگی مکان‌های مشخصه در این حالت دارای ۸۱ مؤلفه است که هرکدام فراوانی نسبی عدد مربوط به خود در تصویر ورودی را نشان می‌دهد. برای آن که عناصر بردار ویژگی مقادیری بین صفر تا یک داشته باشند، مقادیر بردار ویژگی بر تعداد نقاط سفید زمینه تقسیم می‌شوند (نرمالیزه کردن).[۲۴] شکل ۳ نحوه‌ی محاسبه‌ی بردار ویژگی مکان‌های مشخصه را برای یک نقطه از تصاویر رقم ۶ نشان می‌دهد.

۳.۴. طبقه‌بندی‌کننده‌های شبکه‌ی عصبی

طبقه‌بندی و روش‌های آن، به عنوان یکی دیگر از شیوه‌های مهم داده‌کاوی شناخته می‌شود. در این روش، برخلاف خوشه‌بندی، کلاس‌های داده‌ها برای دسته‌بندی آنها مبنای کار بوده و دسته‌بندی با ناظر صورت می‌پذیرد. روش‌های متنوع ابتکاری و فراابتکاری برای طبقه‌بندی داده‌ها وجود دارد. مبنای هریک از این روش‌ها، پدیده‌یی طبیعی (شبکه‌ی عصبی، ژنتیک و ...)، یا نظریه‌ی ریاضیاتی است.

یکی از روش‌های پرکاربرد در طبقه‌بندی داده‌ها، روش شبکه‌های عصبی است که زیر مجموعه‌های متنوعی دارد. این روش با مبنا قرار دادن شبکه‌ی عصبی انسان، الگوریتم طبقه‌بندی را بر پایه مهم‌ترین عنصر پردازش در سیستم‌های زیستی یعنی نرون طراحی می‌کند. در مغز انسان بیش از ۱۰۰ میلیارد نرون وجود دارد که هریک با حدود ده هزار نرون دیگر مرتبط بوده و یک ساختار متراکم را ایجاد کرده است. بر اساس ترکیب‌های مختلف نرون‌ها و قوانین مختلف یادگیری، شبکه‌های



شکل ۴. ساختار نوعی از شبکه‌ی عصبی احتمالاتی.[۲۵]

عصبی مصنوعی مختلفی تشکیل می‌شود. شبکه‌های عصبی مصنوعی مدل‌های محاسباتی هستند که به دلیل یادگیری ساده‌ی الگوهای تازه، برای مسائل بازشناسی الگو مناسب‌اند. از جمله‌ی این روش‌ها می‌توان به پرسپترون چندلایه (MLP)،[۱۴] ماشین بردار پشتیبان (SVM)، شبکه‌های خودسازمانده (SOM)،[۱۵] شبکه‌های احتمالاتی (PNN) [۱۶] و انواع دیگر اشاره کرد.

در این پژوهش به کمک هریک از دو روش MLP و PNN الگوریتم‌های بازشناسی اعداد دست‌نویس توسعه می‌یابند.

۱.۳.۴. شبکه‌ی عصبی احتمالاتی

شبکه‌ی عصبی احتمالاتی یا PNN یکی از شبکه‌های عصبی مصنوعی است که بر پایه‌ی قوانین و مفاهیم بیزی بنا شده است و به عنوان شبکه‌های بیزی نیز شناخته می‌شوند. یکی از کاربردهای این شبکه‌ها، استفاده به عنوان یک روش طبقه‌بندی در مسائل داده‌کاوی است. چنان که در شکل ۴ نشان داده شده است، این شبکه از توپولوژی جلورونده[۱۷] با دو لایه‌ی مخفی و یک لایه‌ی خروجی استفاده می‌کند. لایه‌های مخفی شامل لایه‌ی الگو و لایه‌ی جمع است.[۲۵]

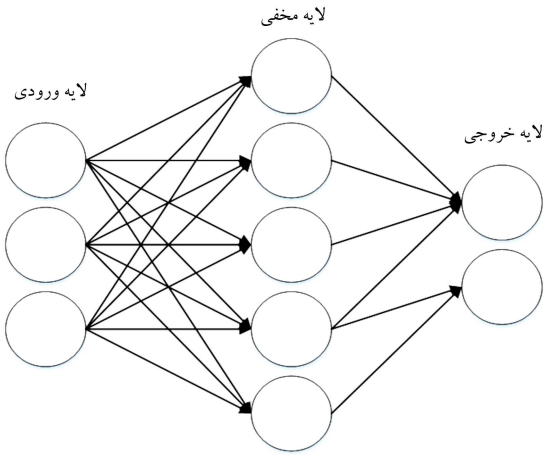
در این نوع از شبکه‌های عصبی، هر نرون با یک داده‌ی ورودی متناظر است که حاوی اطلاعات خاصی از داده ورودی (ویژگی‌های تصاویر) است. فرض می‌شود این اطلاعات توزیع گوسی دارند که در رابطه‌ی ۶ قابل مشاهده است.

$$f(x_i;\mu,\delta^2)=\frac{1}{\sqrt{2\pi\delta^2}}e^{\frac{-(x-\mu)^2}{2\delta^2}} \qquad (۶)$$

در این رابطه x_i ویژگی‌های داده‌ی ورودی است و δ یک پارامتر هموارسازی یا گسترش که همان مفهوم واریانس را دارد. لذا در لایه‌ی مخفی دوم، توزیع‌های کلاس iام بر اساس معادله‌ی ۷ با هم جمع می‌شوند. توزیع گوسی مربوطه در این معادله قابل مشاهده است:

$$f_{c_i}=\sum_{j=1}^{k} f_j \qquad (۷)$$

که در آن f_j توزیع گوسی متناظر با داده jام از کلاس c_i است. در لایه‌ی خروجی، PNN از قواعد تصمیم‌گیری بیزی برای مقایسه‌ی خروجی‌های لایه‌های قبلی و تعیین کلاس درست برای داده‌ی ورودی استفاده می‌کند. روش کار شبکه به این ترتیب است که زمانی که یک بردار ورودی به شبکه داده می‌شود، لایه‌ی اول فاصله‌ی بردار ورودی تا بردارهای ورودی آموزشی را محاسبه کرده و برداری تولید می‌کند که عناصر آن نشان‌دهنده‌ی این است که چگونه ورودی با یک ورودی آموزشی تطبیق پیدا می‌کند. لایه‌ی دوم، این مقادیر را برای هر کلاس از ورودی‌ها جمع می‌کند و خروجی‌اش را که یک بردار احتمالی است تولید می‌کند. در نهایت یک تابع انتقال



شکل ۵. یک نمونه شبکه‌ی عصبی ۲ × ۵ × ۳.[۲۰]

روی خروجی لایه‌ی دوم، بیشینه‌ی این احتمالات را برمی‌گزیند و مقدار ۱ برای آن کلاس و مقدار صفر برای کلاس‌های دیگر تولید می‌کند.

PNN مانند هر روش طبقه‌بندی‌کننده‌ی دیگری معایب و مزایایی دارد. مزایای آن عبارت‌اند از:

۱. تمامی فرایندها در این نوع شبکه به طور موازی انجام می‌شود؛

۲. بار محاسباتی زیادی در فرایند آموزش مورد نیاز نیست (برخلاف شبکه‌های ترویج به عقب[۱۸] و گرادیان). بنابراین PNN دارای یکی از سریع‌ترین آموزش‌هاست.

از جمله معایب PNN عبارتند از:

۱. به حافظه زیادی نیاز دارد؛

۲. به مجموعه‌های آموزشی نیازمند است؛

۳. اغلب اوقات فقط برای طبقه‌بندی کاربرد دارد.

۲.۳.۴. شبکه‌ی عصبی چندلایه

یکی از کاربردی‌ترین شبکه‌های عصبی مصنوعی در علوم مهندسی برای نگاشت غیرخطی در مورد مسائل طبقه‌بندی، شبکه‌های عصبی پرسپترون چندلایه با روش یادگیری پسا انتشار خطا هستند. شبکه‌ی عصبی پرسپترون در ۱۹۵۸ توسط روزنبلات[۱۹] معرفی شد.[۲۶] یک نمونه شبکه‌ی عصبی پرسپترون چندلایه در شکل ۵ دیده می‌شود.

عملکرد بازشناسی شبکه‌های عصبی پسا انتشار خطا به میزان زیادی به ساختار شبکه بستگی دارد. تعداد گره‌ها در لایه‌های خروجی، مخفی و ورودی، ساختار شبکه را تعیین می‌کنند. در این شبکه‌ی عصبی، نرون‌های یک لایه به همه نرون‌های لایه‌های قبل و بعد متصل‌اند. تعداد گره‌های مخفی تأثیر قابل توجهی روی کارایی شبکه‌ی عصبی دارد. اگر تعداد گره‌های مخفی کم‌تر از تعداد مورد نیاز انتخاب شود، کم برازش[۲۰] اتفاق می‌افتد. در این حالت شبکه نمی‌تواند اعداد را درست بازشناسی کند و این موضوع از آن جهت است که پارامترهای قابل تنظیم کافی برای مدل کردن ارتباط بین خروجی و ورودی وجود ندارد. اگر تعداد گره‌های مخفی بیشتر از تعداد مورد نیاز انتخاب شود، بیش برازش[۲۱] اتفاق می‌افتد که در این حالت شبکه قابلیت تعمیم ندارد.

برای تعیین تعداد بهینه‌ی نرون‌ها در لایه‌های مخفی، پیشرفت‌هایی صورت گرفته است اما هنوز راه حل دقیقی برای آن ارائه نشده است. در اینجا بررسی مختصری از نحوه‌ی انتخاب تعداد گره‌های لایه‌های میانی یا مخفی صورت می‌گیرد. در حالت کلی تعداد گره‌های لایه میانی به چندین مورد وابسته است:

۱. ساختار شبکه؛

۲. تعداد گره‌های لایه ورودی و خروجی؛

۳. تعداد نمونه‌های آموزش؛

۴. میزان پیچیدگی نمونه‌هایی که باید یاد گرفته شوند؛

۵. نوع توابع فعال‌سازی در لایه میانی؛

۶. الگوریتم آموزش.

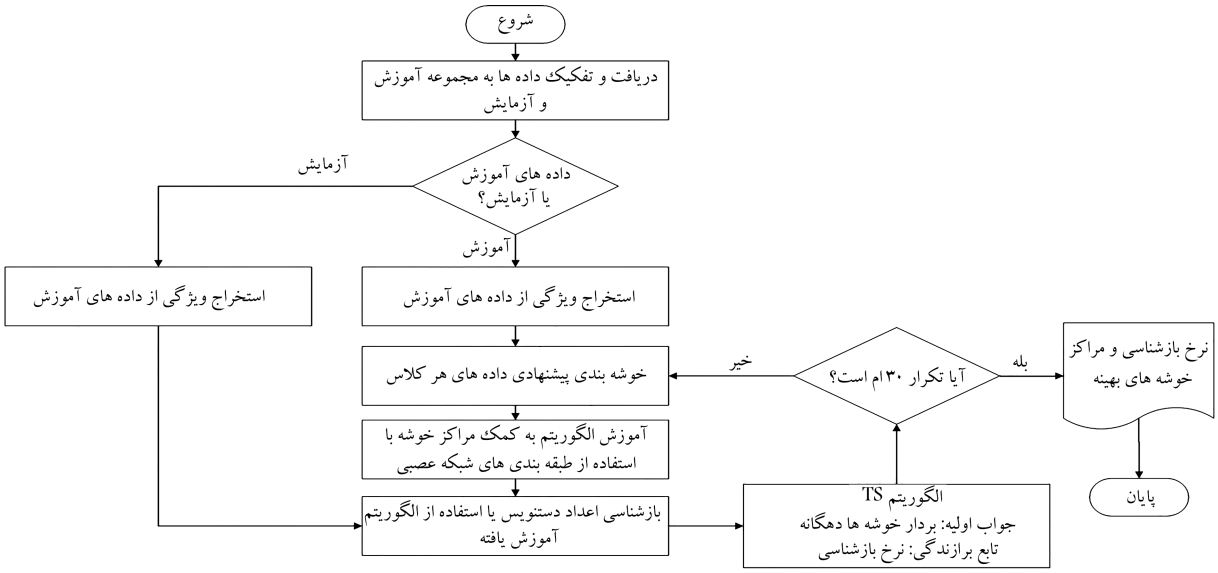
روش‌های متفاوتی برای پیدا کردن وزن‌های بهینه شبکه‌ی عصبی وجود دارد. در این تحقیق از روش پسا انتشار خطا برای این منظور استفاده شده است. در این روش، ماتریس‌های وزن شبکه به گونه‌یی تغییر می‌یابند که خطای شبکه کمینه شود. مقدار مربعات خطا در این روش به صورت لایه به لایه به عقب برگشت داده می‌شود و در هر لایه اصلاحات لازم روی وزن‌ها انجام می‌گیرد. این روند تا جایی ادامه می‌یابد که خطای خروجی کل کمینه شود و یا تعداد دفعات تکرار به تعداد معینی برسد. منظور از خطا، اختلاف بین برچسب‌های پیشنهادی شبکه‌ی عصبی و برچسب‌های واقعی است. اگر t برچسب واقعی مورد نظر و a برچسبی باشد که شبکه‌ی عصبی اختصاص می‌دهد، آنگاه خطا طبق رابطه‌ی ۸ تعریف می‌شود.

$$e = MSE(t - a) \qquad (۸)$$

۴.۴. الگوریتم جست‌وجوی ممنوعه

الگوریتم جست‌وجوی ممنوعه (TS) [۲۲] یک الگوریتم بهینه‌سازی فراابتکاری است که برای اولین بار در سال ۱۹۸۶ توسط گلوور معرفی شد.[۲۷] این الگوریتم بر مبنای مغز انسان پایه‌ریزی شده است و دارای مفاهیم مشابهی همچون حافظه کوتاه مدت، یادگیری، فراموشی، استنتاج و ... است. این الگوریتم در صورت پیاده‌سازی درست، بهترین الگوریتم فراابتکاری محسوب می‌شود.

برای رسیدن به جواب بهینه در یک مسئله بهینه‌سازی، الگوریتم جست‌وجوی ممنوعه ابتدا از یک جواب اولیه شروع به حرکت می‌کند. سپس الگوریتم بهترین جواب همسایه را از میان همسایه‌های جواب فعلی انتخاب می‌کند. در صورتی که این جواب در فهرست ممنوعه قرار نداشته باشد، الگوریتم به جواب همسایه حرکت می‌کند؛ در غیر این‌صورت الگوریتم معیاری به نام معیار تنفس را بررسی خواهد کرد. بر اساس معیار تنفس، اگر جواب همسایه از بهترین جواب یافت شده تاکنون بهتر باشد، الگوریتم به آن حرکت خواهد کرد، حتی اگر آن جواب در فهرست ممنوعه باشد. پس از حرکت الگوریتم به جواب همسایه، فهرست ممنوعه به‌روزرسانی می‌شود؛ به این معنا که با توجه به حرکت قبل، جواب همسایه در فهرست ممنوعه قرار داده می‌شود تا از بازگشت مجدد الگوریتم به آن جواب و ایجاد چرخه جلوگیری شود. در واقع فهرست ممنوعه ابزاری در الگوریتم جست‌وجوی ممنوعه است که توسط آن از قرار گرفتن الگوریتم در بهینه‌ی محلی جلوگیری می‌شود. پس از قرار دادن حرکت قبلی در فهرست ممنوعه، تعدادی از حرکت‌هایی که قبلاً در فهرست ممنوعه قرار گرفته بودند از فهرست خارج می‌شوند. مدت زمانی که حرکت‌ها در فهرست ممنوعه قرار می‌گیرند توسط یک پارامتر که «زمان ممنوعه» نام دارد تعیین می‌شود. حرکت از جواب فعلی به جواب همسایه تا جایی ادامه می‌یابد که شرط خاتمه دیده شود. شرط‌های خاتمه متفاوتی می‌توان برای الگوریتم در نظر گرفت. به طور مثال محدودیت تعداد حرکت به جواب همسایه می‌تواند یک شرط خاتمه باشد.



شکل ۶. فلوچارت‌های الگوریتم بازشناسی اعداد دست‌نویس مبتنی بر خوشه‌بندی.

دو بخش مهم در این الگوریتم، راهکارهای قوی‌سازی و متنوع‌سازی هستند. در راهکار قوی‌سازی، با تعدیل قوانین و فهرست ممنوعه، فضای جست‌وجو به جهت بررسی بیشتر جواب‌های همسایه افزایش می‌یابد اما در راهکار متنوع‌سازی، با ترکیب جواب‌های فعلی، جواب‌های جدید تولید شده و پس از مطابقت آن با لیست ممنوعه مورد بررسی قرار می‌گیرد.

۵. توسعه‌ی الگوریتم بازشناسی اعداد دست‌نویس فارسی

در این پژوهش تلاش شد تا با توسعه و بررسی الگوریتمی مبتنی بر شبکه‌ی عصبی و خوشه‌بندی، خطای حاصل از بازشناسی اعداد کاهش یافته و در نرخ بازشناسی بهبود قابل توجهی حاصل شود. به این منظور، با استفاده از طبقه‌بندی کننده‌های شبکه‌ی عصبی احتمالاتی و چندلایه، مراکز خوشه‌ها به جای مجموعه داده‌ها مورد آموزش قرار می‌گیرد. مراکز خوشه‌ها در این الگوریتم به کمک یکی از روش‌های پیوند کامل، PAM و FCM تعیین می‌شود. مراحل روش پیشنهادی عبارت‌اند از:

۱. تفکیک داده‌های ورودی به داده‌های آموزش و آزمایش؛

۲. استخراج یکی از دو دسته ویژگی از داده‌های آموزش و آزمایش؛

۳. با استفاده از برچسب‌های دهگانه اعداد آموزش و یکی از روش‌های خوشه‌بندی روی داده‌های هر عدد جداگانه انجام می‌پذیرد؛

۴. شبکه بر اساس مراکز خوشه‌های به دست آمده و هریک از روش‌های طبقه‌بندی شبکه‌ی عصبی احتمالاتی و چندلایه آموزش می‌بینند؛

۵. آزمایش شبکه با استفاده از داده‌های آزمایش و پس از استخراج ویژگی‌ها صورت پذیرفته و نرخ بازشناسی محاسبه می‌شود؛

۶. نحوه‌ی تعیین تعداد خوشه‌های بهینه با استفاده از الگوریتم بهینه‌سازی جست‌وجوی ممنوعه به این ترتیب است که بردار ۱ × ۱۰ حاوی تعداد خوشه‌های دهگانه به عنوان جواب اولیه، افزایش و کاهش و تغییر هم‌زمان به عنوان قاعده‌ی تعیین جواب همسایگی، و نرخ بازشناسی به عنوان تابع برازندگی لحاظ می‌شود. در هر مرحله از اجرای الگوریتم جست‌وجوی ممنوعه، تعداد خوشه‌های پیشنهادی به عملیات مرحله‌ی ۳ تزریق می‌شود و سپس نرخ بازشناسی مربوطه در عملیات مرحله‌ی ۵ به الگوریتم جست‌وجوی ممنوعه انعکاس می‌یابد و این چرخه تا پایان ۳۰ تکرار که شرط اختتام الگوریتم است، ادامه می‌یابد. در انتها تعداد خوشه‌های بهینه‌ی هرکلاس و نرخ بازشناسی به عنوان خروجی الگوریتم ارائه می‌شود.

فلوچارت الگوریتم طبقه‌بندی اعداد دست‌نویس فارسی در شکل ۶ نشان داده شده است. در این الگوریتم شبکه با مراکز خوشه‌ی اولیه آموزش می‌بینند و میزان دقت این شبکه‌ی آموزش یافته به کمک داده‌های آزمایش سنجیده می‌شود. الگوریتم فرابتکاری TS نیز تلاش می‌کند با تغییر مراکز خوشه الگوریتم را به سمت حالت بهینه بازشناسی سوق دهد.

بعد از چند مرحله اجرای کامل الگوریتم و اطمینان از نتایج، تعداد خوشه‌های بهینه در دسترس خواهد بود. الگوریتم مورد نظر در کل ۶ مرتبه با استفاده از طبقه‌بندی‌کننده‌ی PNN و شش مرتبه نیز با استفاده از طبقه‌بندی‌کننده‌ی MLP اجرا می‌شود. این شش تکرار بر اساس ترکیب استفاده از دو ویژگی ناحیه‌بندی و مکان مشخصه و سه روش خوشه‌بندی complete linkage، PAM و FCM است.

۶. نتایج اجرای الگوریتم‌ها

در این فصل نتایج حاصل از اجرای الگوریتم با استفاده از داده‌های پایگاه داده هدی توصیف می‌شود.

برای مقایسه‌ی عملکرد الگوریتم در هر بخش (آموزش با شبکه‌ی عصبی احتمالاتی و چندلایه) ابتدا نرخ بازشناسی بدون استفاده از خوشه‌بندی و تنها با استفاده از داده‌های اولیه و به کمک الگوریتم طبقه‌بندی کننده انجام پذیرفت.[۳و۶]

جدول ۱. تعداد خوشه‌های بهینه هر کلاس (ناحیه‌بندی).

کلاس	۰	۱	۲	۳	۴	۵	۶	۷	۸	۹
پیوند کامل										
تعداد خوشه‌ی بهینه‌ی عصبی احتمالاتی	۱۵۶	۱۳۰	۹۰	۱۶۰	۲۱۱	۱۲۳	۲۰۰	۱۸۹	۱۹۰	۲۰۵
تعداد خوشه‌ی بهینه‌ی عصبی چندلایه	۹	۱۰	۲۱	۸	۹	۸	۷	۱۰	۱۷	۱۱
PAM										
تعداد خوشه‌ی بهینه‌ی عصبی احتمالاتی	۱۵۱	۱۴۹	۱۸۰	۱۷۶	۲۲۰	۱۵۰	۱۶۰	۱۱۲	۱۳۰	۱۶۱
تعداد خوشه‌ی بهینه‌ی عصبی چندلایه	۱۲	۱۰	۸	۵	۲۰	۱۵	۱۰	۶	۹	۲۴
FCM										
تعداد خوشه‌ی بهینه‌ی عصبی احتمالاتی	۹۲	۸۱	۹۲	۱۰۱	۷۵	۶۲	۸۲	۸۵	۹۷	۴۳
تعداد خوشه‌ی بهینه‌ی عصبی چندلایه	۱۸	۳	۵	۴	۱۴	۶	۱۷	۱۱	۶	۷

جدول ۲. تعداد خوشه‌های بهینه هرکلاس (مکان مشخصه).

کلاس	۰	۱	۲	۳	۴	۵	۶	۷	۸	۹
پیوند کامل										
تعداد خوشه بهینه عصبی احتمالاتی	۲۰۹	۱۴۰	۲۲۰	۱۶۵	۱۲۰	۱۴۹	۱۴۴	۲۲۰	۱۴۵	۲۳۴
تعداد خوشه بهینه عصبی چند لایه	۹	۱۴	۱۶	۸	۹	۱۰	۱۱	۲	۸	۹
PAM										
تعداد خوشه بهینه عصبی احتمالاتی	۲۸۰	۲۹۳	۲۸۰	۱۸۹	۱۹۵	۱۶۳	۱۷۹	۱۳۲	۲۱۶	۱۹۶
تعداد خوشه بهینه عصبی چند لایه	۱۰	۹	۷	۶	۱۸	۱۱	۸	۵	۵	۳
FCM										
تعداد خوشه بهینه عصبی احتمالاتی	۱۰	۱	۲	۷	۱۰	۳	۴	۲	۴	۹
تعداد خوشه بهینه عصبی چند لایه	۱۵	۱۴	۱۰	۲	۸	۲	۱۱	۱۲	۹	۹

اما به علت محدودیت حافظه و نیاز الگوریتم به حافظه بالا، با حافظه‌های موجود (8G) تنها آموزش ۱۰۰۰۰ رقم در حالت بدون خوشه بندی میسر شد.

بر اساس الگوریتم توصیف شده در این پژوهش، در راستای تعیین تعداد خوشه‌های بهینه هر کلاس با استفاده از ۶۰۰۰۰ داده‌آموزشی، ابتدا ویژگی‌های داده‌های هر کلاس (مکان مشخصه یا ناحیه‌بندی) استخراج شده و پس از آن به کمک هریک از روش‌های خوشه‌بندی پیوند کامل، PAM و FCM و با استفاده از الگوریتم بهینه‌سازی جست‌وجوی ممنوعه برای هریک از دو طبقه‌بندی کننده شبکه‌ی عصبی احتمالاتی (واریانس ۴ برای ناحیه‌بندی و ۰٫۵ برای مکان مشخصه) و شبکه‌ی عصبی چندلایه ($c = ۷۰$، $m = ۹۰$ و $\eta = ۰٫۳$) تعداد بهینه خوشه‌ها در هر کلاس تعیین می‌شود.

نتایج نهایی خوشه‌بندی در جدول ۱ متناسب با ویژگی‌های ناحیه‌بندی و در جدول ۲ متناسب با ویژگی‌های مکان مشخصه تعداد خوشه‌های بهینه هر کلاس تعیین شده‌است. برای مثال در جدول ۱ تعداد خوشه‌های بهینه‌ی کلاس صفر با استفاده از روش خوشه‌بندی PAM و طبقه‌بند شبکه‌ی عصبی احتمالاتی ۱۵۱ خوشه است.

پس از تعیین تعداد خوشه‌های بهینه، الگوریتم بازشناسی به کمک مراکز خوشه‌ها آموزش می‌بینند. در راستای بررسی دقت بازشناسی الگوریتم شبکه‌ی عصبی توسعه‌یافته در هر حالت (ترکیب شبکه‌ی عصبی، روش خوشه‌بندی و دسته ویژگی)، نرخ بازشناسی به کمک داده‌های آزمایش محاسبه می‌شود. در جدول ۳ نرخ‌های بازشناسی شبکه‌ی عصبی احتمالاتی با توجه به روش تعیین مراکز خوشه از هریک از دو دسته ویژگی استخراج شده قابل مشاهده است و به طور مشابه این مقادیر برای شبکه‌ی عصبی چندلایه در جدول ۴ مشخص شده است.

۷. بحث و نتیجه‌گیری

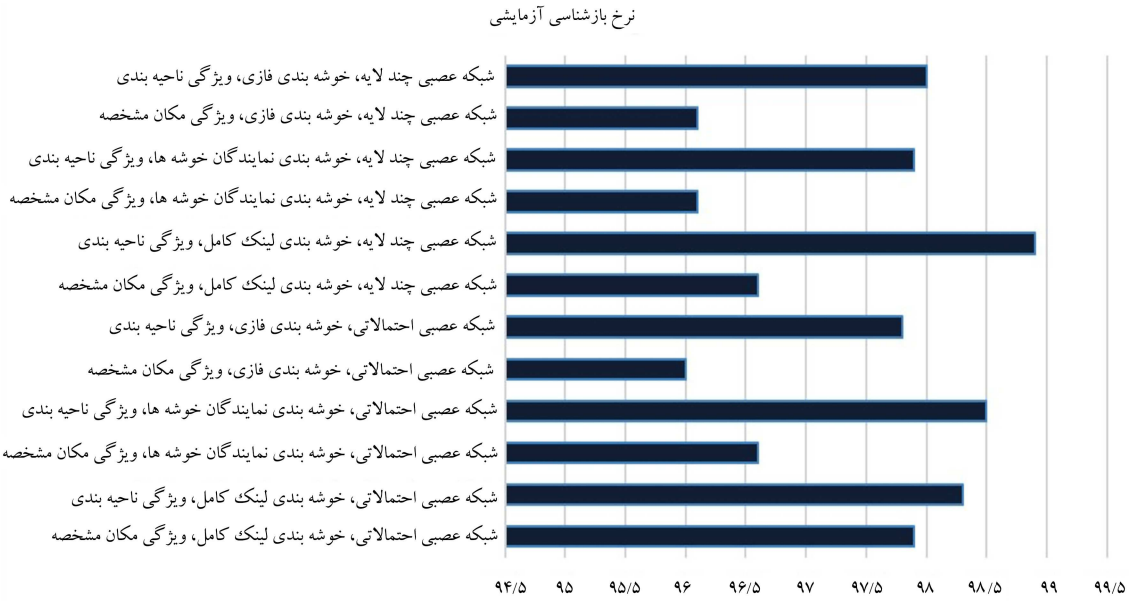
در این پژوهش، الگوریتم‌هایی برای بازشناسی اعداد دست‌نویس فارسی با استفاده از دو نوع روش طبقه‌بندی، در نظر گرفتن سه روش خوشه‌بندی و دو دسته ویژگی استخراج شده مورد بررسی قرار گرفتند. الگوریتم‌های توسعه‌یافته به کمک مراکز خوشه‌های تعیین شده، اعداد دست‌نویس فارسی را در بهترین حالت با نرخ‌های

جدول ۳. نرخ بازشناسی الگوریتم‌ها با استفاده از داده‌های آزمایش به کمک شبکه‌ی عصبی احتمالاتی.

نوع الگوریتم	نوع ویژگی	نرخ بازشناسی آزمایشی
طبقه‌بندی کننده PNN بدون خوشه بندی (۱۰۰۰۰ داده)	مکان مشخصه	۹۲/۱
طبقه‌بندی کننده PNN بدون خوشه بندی (۵۰۰۰ داده)	ناحیه‌بندی	۹۴/۳
طبقه‌بندی کننده PNN با خوشه‌بندی پیوند کامل (۶۰۰۰۰ داده)	مکان مشخصه	۹۷/۹
طبقه‌بندی کننده PNN با خوشه‌بندی پیوند کامل (۶۰۰۰۰ داده)	ناحیه‌بندی	۹۸/۳
طبقه‌بندی کننده PNN با خوشه‌بندی PAM (۶۰۰۰۰ داده)	مکان مشخصه	۹۶/۶
طبقه‌بندی کننده PNN با خوشه‌بندی PAM (۶۰۰۰۰ داده)	ناحیه‌بندی	۹۸/۵
طبقه‌بندی کننده PNN با خوشه‌بندی FCM (۶۰۰۰۰ داده)	مکان مشخصه	۹۶
طبقه‌بندی کننده PNN با خوشه‌بندی FCM (۶۰۰۰۰ داده)	ناحیه‌بندی	۹۷/۸

جدول ۴. نرخ بازشناسی الگوریتم‌ها با استفاده از داده‌های آزمایش به کمک شبکه‌ی عصبی چندلایه.

نوع الگوریتم	نوع ویژگی	نرخ بازشناسی آزمایشی
طبقه‌بندی کننده MLP بدون خوشه بندی (۱۰۰۰۰ داده)	مکان مشخصه	۸۹/۶
طبقه‌بندی کننده MLP بدون خوشه بندی (۵۰۰۰ داده)	ناحیه‌بندی	۹۵
طبقه‌بندی کننده MLP با خوشه‌بندی پیوند کامل (۶۰۰۰۰ داده)	مکان مشخصه	۹۶/۶
طبقه‌بندی کننده MLP با خوشه‌بندی پیوند کامل (۶۰۰۰۰ داده)	ناحیه‌بندی	۹۸/۹
طبقه‌بندی کننده MLP با خوشه‌بندی PAM (۶۰۰۰۰ داده)	مکان مشخصه	۹۶/۱
طبقه‌بندی کننده MLP با خوشه‌بندی PAM (۶۰۰۰۰ داده)	ناحیه‌بندی	۹۷/۹
طبقه‌بندی کننده MLP با خوشه‌بندی FCM (۶۰۰۰۰ داده)	مکان مشخصه	۹۶/۱
طبقه‌بندی کننده MLP با خوشه‌بندی FCM (۶۰۰۰۰ داده)	ناحیه‌بندی	۹۸



شکل ۷. مقایسه نرخ‌های بازشناسی داده‌های آزمایش.

۹۸/۹٪ (طبقه‌بندی کننده MLP با خوشه‌بندی complete linkage) و ۹۸/۵٪ (طبقه‌بندی کننده PNN با خوشه‌بندی PAM) بازشناسی کردند. نتایج حاصل که در جداول ۳ و ۴ و همچنین شکل ۷ قابل مشاهده است، نشان می‌دهد که نرخ بازشناسی داده‌های آزمایشی اعداد دست‌نویس فارسی در حالت خوشه‌بندی نسبت به حالت بدون خوشه‌بندی بهبود قابل ملاحظه‌یی داشته است.

اجرای الگوریتم در هر دو حالت، در سیستمی با قدرت پردازش (intel core i۵ – ۱۰۵۲۰U(۳/۱GHz و ۸GB RAM صورت پذیرفته است. نتایج فوق حاکی از آن است که بیشترین نرخ بازشناسی متعلق به الگوریتم طبقه‌بندی شبکه‌ی عصبی چندلایه به کمک خوشه‌بندی linkage و ویژگی‌های ناحیه‌بندی است.

تحلیل نتایج حاصل از اجرای الگوریتم مشخص می‌شود خروجی حاصل از بازشناسی اعداد به کمک ویژگی ناحیه‌بندی نسبت به ویژگی مکان مشخصه در

جدول ۵. مقایسه‌ی نتایج پژوهش‌ها.

مرجع	ویژگی مستخرج	الگوریتم	پایگاه داده	نرخ بازشناسی داده‌های آزمایش (درصد)
خراشادی‌زاده و همکاران[۸]	هیستوگرام گرا[۱]	SVM	هدی	۹۹٫۳۱
فرخی و همکاران[۱۰]	فوریه، موجک	یادگیری عمیق	هدی	۹۳
دهقانیان و قدس[۱۱]	کرنل	CNN	هدی	۹۸٫۹۳
حقیقی و همکاران[۱۲]	بردار احتمال کلاس تصاویر	CNN	هدی	۹۹٫۳۹
میری و خدمتی[۱۴]	ناحیه‌بندی	PNN و خوشه‌بندی ناحیه‌بندی	هدی	۹۸٫۴
قنبری[۲۸]	ناحیه‌بندی	MLP و کاهش ابعاد بردارهای ویژگی به روش تهیه‌ی مؤلفه‌های اصلی	هدی	۹۶٫۸
عبدی و سلیمی[۲۹]	مکان مشخصه	RBF[۲]	هدی	۹۷٫۱
رضایی و کیمیایی[۳۰]	حفره، چگالی حفره و حجم پیکسل	KNN	هدی	۹۷٫۳
سرورامینی و همکاران[۳۱]	OCR[۳]	CNN	هدی	۹۷٫۷
پژوهش فعلی	ناحیه‌بندی و مکان مشخصه	MLP و خوشه بندی لینک کامل	هدی	۹۸٫۹

۱. Histogram of Oriented
۲. Radial Basis Function
۳. Optical Character Recognition

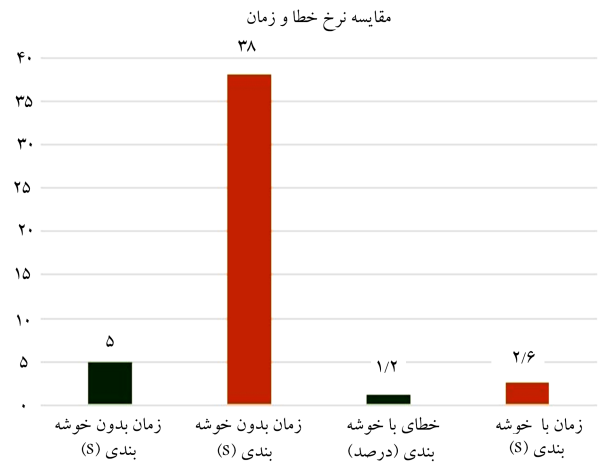
تمامی حالات از دقت بالاتری برخوردار بوده و نرخ بازشناسی بهتری دارد. لذا استفاده از ویژگی ناحیه‌بندی در کنار طبقه‌بندی‌کننده‌های شبکه‌ی عصبی MLP و PNN نسبت به مکان مشخصه نتایج بهتری ارائه می‌دهد.

با توجه به این که هدف پژوهش بررسی بهبود در نرخ بازشناسی به‌کمک خوشه‌بندی است، سایر روش‌های خوشه‌بندی و طبقه‌بندی کننده مورد بررسی قرار نگرفته است اما می‌توان برای پژوهش‌های آینده، میزان انطباق سایر ویژگی‌ها با طبقه‌بندهای موجود را مورد بررسی قرار داد. در جدول ۵ نتایج پژوهش‌های اخیر در حوزه‌ی بازشناسی اعداد ثبت شده است. نرخ‌های بازشناسی ذکر شده در جدول ۵ برای مقایسه‌ی بهتر، همگی براساس ۶۰۰۰۰ داده آموزشی و ۲۰۰۰۰ داده آزمایش پایگاه داده‌ی هدی هستند. با توجه به نتایج ارائه شده، الگوریتم توسعه‌یافته در این پژوهش که به بازشناسی اعداد دست‌نویس با استفاده از خوشه‌بندی می‌پردازد توانسته بهبود قابل توجهی در نرخ بازشناسی نسبت به الگوریتم‌های بدون خوشه‌بندی نشان دهد.

نرخ بازشناسی ۹۸٫۹٪ در روش ارائه شده این پژوهش، نسبت به روش‌های ارائه شده توسط حقیقی و همکاران،[۹] دهقانیان و همکاران[۱۲] و خراشادی‌زاده و همکاران،[۸] به ترتیب به میزان ۰٫۴۹٪، ۰٫۰۳٪ و ۰٫۴۱٪ خطای بیشتری داشته است اما به علت استفاده از خوشه‌بندی، الگوریتم ارائه شده در این مقاله در زمان بسیار کم‌تری تصاویر را بازشناسی کرده است که با توجه به میزان بهبود در زمان حل، خطای قابل اغماضی محسوب می‌شود.

بر این اساس در راستای مقایسه کارایی زمانی روش‌های ارائه شده، در شکل ۸ زمان و خطای بازشناسی در حالت خوشه‌بندی پیوند کامل (۶۰۰۰۰ داده) و بدون خوشه‌بندی (۵۰۰۰ داده) داده با کمک بردارهای ویژگی ناحیه‌بندی توسط الگوریتم MLP مقایسه شده است.

در برخی از پژوهش‌ها، اعتبارسنجی مدل با استفاده از داده‌های آموزشی صورت



شکل ۸. مقایسه‌ی خطا و زمان بازشناسی با و بدون خوشه‌بندی.

گرفته و نرخ‌های بازشناسی بر این اساس مشخص شده‌اند که بررسی میزان بهبود پیچیدگی زمانی و دقت الگوریتم در بازشناسی داده‌های آموزش می‌تواند موضوع پژوهش‌های آتی باشد.

در مطالعات آینده می‌توان با بررسی روش‌های طبقه‌بندی جدید، شبکه‌ی عصبی همچون CNN، یادگیری عمیق[۲۳] و سایر روش‌های طبقه‌بندی ابتکاری و نظری نیز مورد بررسی و پایش قرارگیرند. استفاده از الگوریتم فراابتکاری جست‌وجوی ممنوعه نیزکمک شایانی در جهت کاهش پیچیدگی زمانی الگوریتم تعیین تعداد خوشه‌های بهینه کرد اما در پژوهش‌های آتی، با در نظرگرفتن نوع طبقه‌بندی کننده و همچنین روش خوشه‌بندی، می‌توان سایر الگوریتم‌های فراابتکاری، ابتکاری و دقیق را با توجه به هدف پژوهشگر بررسی کرد.

پانوشت‌ها

1. Support Vector Machine
2. Convolutional Neural Network
3. Complete linkage
4. Partition Around Medoids
5. Fuzzy C-Means
6. Clustering
7. Unsupervised
8. Cluster
9. Single Linkage
10. Bottom-Up
11. Top-Down
12. Zoning
13. Loci
14. Multiple Layer Perceptron
15. Self-Organizing Maps
16. Probabilistic Neural Networks
17. Feed forward
18. Back forward
19. Rosenblatt
20. Under Fitting
21. Over Fitting
22. Tabu Search
23. Deep Learning

منابع (References)

1. Rajaee, A. and Moradgholi, F. "Persian handwritten digit recognition", *National Conference of Computer, IT and AI Applications*, (In Persian) (1397/2018).
2. Johari majd, V. and Razavi, S.M. "Fuzzy persian handwritten digit recognition", *The 1th Iranian International Conference on Machine Vision and Image Processing*, pp. 151-144 (In Persian) (1380/2001).
3. Khalili, F. and Fadavi amiri, M. "Persian handwritten digit recognition with MLP", *4th International Conference Modern Studies in Computer Science and Information Technology*, (In Persian) (1397/2018).
4. Darvish, A., Kabir, E. and Khosravi, H. "Application of figure matching in persian handwritten digit recognition", *Iranian Journal of Electrical and Computer Engineering*, **22**, pp. 37-48, (In Persian) (1385/2006).
5. Nahvi, M., Rafiee, M., Ebrahimpor, R. and et al. "Combining two class categories for persian handwritten digit recognition", *16th Iranian Electrical Engineering Conference*, (In Persian) (1387/2008).
6. Khosravi, H. "Persian handwritten digit and letters recognition in national exam registration forms", Master Thesis, Tarbiat Modares University, (In Persian) (1384/2005).
7. Bouchene, M.M. and Boukharouba, A. "Features extraction and reduction techniques with optimized SVM for Persian/Arabic handwritten digits recognition", *Iran Journal of Computer Science*, **5**, pp. 247-265 (2022).
8. Khorashadizadeh, S. and Latif, A. "Arabic/farsi handwritten digit recognition using histogram of oriented gradient and Chain code histogram", *The International Arab Journal of Information Technology*, **13**(4), pp. 367-374 (July 2016).
9. Golzari, S., Khalili, A. and Sabzi, R. "Combining convolutional neural networks with SVM classifier for recognizing persian and Arabic handwritten words", *Multimedia Tools and Applications*, **81**, pp. 33785-33799 (2022).
10. Farrokhi, A. "Recognize handwritten figures using deep learning", *International Conference on Nonlinear Systems and Electrical and Computer Engineering Optimization*, (In Persian) (1394/2015).
11. Dehghanian, A. and Ghods V. "Farsi handwriting digit recognition based on convolutional neural Networks", 6th International Symposium on Computational and Business Intelligence (2018).
12. Haghighi, f. and Omranpour, H. "Stacking ensemble model of deep learning and its application to Persian/Arabic handwritten digits recognition", *Knowledge-Based Systems*, **220**, 106940 (2021).
13. Yang, M. and Ahuja, N. "Gaussian mixture model for human skin color and its applications in image and video databases", *Proceedings of the SPIE*, 3656, pp. 458-466 (1998).
14. Rose, K., Gurewitz, E. and Fox, G.C. "Vector quantization by deterministic annealing", *IEEE Transactions On Information Theory*, **38**(4), pp. 1249-1257 (1998).
15. Fung, G. and Mangasarian, O.L, "Semi-supervised support vector machines for unlabeled data classification", *Optimization Methods and Software*, **15**(1), pp. 29-44 (2011).
16. Ebrahimi, A. and Kabir, E. "A two-step method for recognizing printed sub-words", *Iranian Journal of Electrical and Computer Engineering*, **2**(2), pp. 57-62 (2004).
17. Taghipor-Gorjikolaie, M., Miri, I., Razavi, S. and et al. "Persian hand written digit recognition using particle swarm probabilistic neural network", *Journal of Iranian Association of Electrical and Electronics Engineers*, **12**(3), pp. 101-110 (2015).
18. Miri, A. and Khedmati, M. "Improvement of persian handwritten digit recognition rate with using definite and indefinite clustering methods", *2th International Conference on Challenges and New Solutions in Industrial Engineering and Management and Accounting*, (In Persian) (1400/2021).
19. Khosravi, H. and Kabir, E. "Introducing a very large dataset of handwritten Farsi digits and a study on their varieties", *Pattern Recognition Letters*, **28**(10), pp. 1133-1141 (2007).
20. He, Q. "A review of clustering algorithms as applied in IR", Technical. Report, IUIUCLIS-1999/6+IRG (1999).
21. Web, A.R., *Statistical Pattern Recognition*, John Wiley & Sons (2004).
22. Legendre, P. and Legendre, L., *Numerical Ecology*, 2nd Edition, Elsevier, Amsterdam (1998).
23. Fung, G. "A comprehensive overview of basic clustering algorithms", *IEEE Transactions on Information Theory*, **27**(1), pp. 49-60 (2001).
24. Mahabadi, A., Kazemian, A. and Torkamani, A. "Fuzzy recognition of persian handwritten digit", *The CIS Journal on Computer Science and Engineering*, **4**(1), pp. 19-25 (In Persian) (1385/2006).

25. Nabavi kerizi, S.H. Ebrahimpor, R. and Kabir, E. "Application of classifier combinations in handwritten digit recognition", *The 3th Iranian International Conference on Machine Vision and Image Processing*, pp. 115-119 (2004) (In Persian) (1383/2004).

26. Rosenblatt, F. "A probabilistic model for information storage and organization in the brain", *Psychological Review*, **65**(6), pp. 386-408 (1958).

27. Glover, F. "Future paths for integer programming and links to artificial intelligence", *Computers and Operations Research*, **13**(5), pp. 533-549 (1968).

28. Ghanbari, N. "Attribute selection for handwritten digit recognition", Master Thesis, Birjand University (In Persian) (1390/2011).

29. Abdi, M.J. and Salimi, H. "Farsi handwriting recognition with mixture of RBF experts based on particle swarm optimization", *International Journal of Information Science and Computer Mathematics*, **2**(2), pp. 129-136, (In Persian) (1389/2010).

30. Rezaie, A. and Kimiyaei, F. "Detection of handwritten digits using fuzzy logic", *International Journal Information Technology in Engineering Design*, **10**(2), pp. 23-37 (In Persian) (1397/2018).

31. Sarvaramini, F., Nasrollahzadeh, A. and Soryani, M. "Persian handwritten character recognition using convolutional neural network", *In Electrical Engineering (ICEE), Iranian Conference on*, pp. 1676-1680 IEEE (2018).