ادغام شبکه‌های عصبی بر اساس یادگیری با همبستگی منفی در بازشناسی برون‌خط کلمات دست‌نویس

سیدعلی‌اصغر عباس‌زاده آرانی و احسان‌اله کبیر

چکیده: در این تحقیق، یک روش طبقه‌بندی جمعی بر اساس یادگیری با همبستگی منفی برای بازشناسی کل‌نگر کلمات دست‌نویس با حجم محدود پیشنهاد می‌شود. در این روش مجموعه داده آموزشی پس از پیش‌پردازش و استخراج ویژگی به طبقه‌بندهای پایه پرسپترون چندلایه اعمال می‌شود. سپس شبکه‌های عصبی پایه به روش یادگیری با همبستگی منفی، آموزش داده شده و از این طریق گوناگون می‌شوند. هنگامی که داده‌های آزمایشی پس از استخراج ویژگی به طبقه‌بندهای پایه اعمال می‌شوند، هر طبقه‌بند پایه خروجی نسبتاً متفاوتی را تولید می‌کند. با ادغام خروجی طبقه‌بندهای پایه، خروجی نهایی سیستم به دست می‌آید. برای آزمایش روش پیشنهادی از سه نوع ویژگی شامل ویژگی‌های مبتنی بر منطقه‌بندی، گرادیان تصویر و کد زنجیره‌ای کانتور استفاده شده است. در آزمایش‌هایی که روی ۷۷۵ تصویر از نام ۳۱ مرکز استان کشور، از مجموعه داده "ایرانشهر" انجام شده است، استفاده از ویژگی‌های مبتنی بر گرادیان و آموزش ٦ شبکه پرسپترون با همبستگی منفی و ادغام آنها از طریق رأی‌گیری، میانگین نرخ بازشناسی برابر با ۹٦/۱۰ درصد را به دست داده است. سپس خطاهای روش پیشنهادی تحلیل و ردیابی شده است.

کلیدواژه: **ادغام طبقه‌بندها، بازشناسی کلمات دست‌نویس، پرسپترون چندلایه، یادگیری با همبستگی منفی.**

۱- مقدمه

بازشناسی دست‌نوشته به صورت برون‌خط در کاربردهای متعددی مانند تعیین آدرس پستی [۱]، پردازش چک‌های بانکی [۲]، بازشناسی یادداشت‌های دست‌نویس [۳] و تولید کتابخانه‌های دیجیتال اسناد تاریخی [۴] مطرح است. مقاله‌ها و پژوهش‌ها در زمینه بازشناسی دست‌نویس در زبان‌های لاتین، چینی و ژاپنی بسیار زیاد است و این در حالی است که در مورد زبان‌های عربی و فارسی کارهای کمتری انجام شده است.

به طور کلی دو رویکرد در زمینه بازشناسی کلمات وجود دارد: رویکرد جزءنگر و کل‌نگر. در روش‌های جزءنگر، یک کلمه به صورت مجموعه‌ای از واحدهای کوچک‌تر مانند حروف یا زیرحروف در نظر گرفته می‌شود. پس از قطعه‌بندی کلمه به واحدهای تشکیل‌دهنده آن، به کمک بازشناسی این واحدها کلمه تشخیص داده می‌شود. در روش کل‌نگر، کلمات از روی شکل عمومی آنها و بدون شکستن آنها به حروف تشخیص داده می‌شوند. در این روش هر کلمه در فرهنگ کلمات به عنوان یک کلاس جداگانه در نظر گرفته می‌شود. روش‌های کل‌نگر از متداول‌ترین روش‌های بازشناسی کلمات در کاربردهایی با مجموعه واژگان محدود و ثابت هستند. از آنجایی که روش‌های کل‌نگر، کلمه را به اجزای تشکیل‌دهنده آن نمی‌شکنند، می‌توانند ویژگی‌های منحصر به یک کلمه را به خوبی مدل کنند. به عنوان مثال، اثر متقابل بین حروف کنار هم در یک کلمه که می‌تواند شکل ظاهری آنها را تغییر دهد، در روش کل‌نگر در نظر گرفته می‌شود.

در این تحقیق، یک سیستم بازشناسی چندگانه مبتنی بر آموزش با همبستگی منفی (NCL)[1] برای بازشناسی کل‌نگر کلمات دست‌نویس بدون قید پیشنهاد شده است. در این روش هر یک از طبقه‌بندهای پایه یک شبکه عصبی پرسپترون چندلایه با قاعده یادگیری انتشار به عقب است که از طریق اضافه‌نمودن یک ترم جریمه (شباهت خروجی شبکه‌های پایه) به تابع خطای سیستم، خطای آنها متفاوت و گوناگون می‌شود. از این طریق هر یک از طبقه‌بندهای پایه روی بخشی از فضای ورودی تمرکز نموده و این فضا بین آنها تقسیم می‌شود. این روش در گروه رویکردهای صریح ایجاد گوناگونی بین طبقه‌بندهای پایه قرار می‌گیرد.

در ادامه این تحقیق مطالب زیر ارائه خواهد شد. در بخش ۲ کارهای پیشین در زمینه بازشناسی کلمات دست‌نویس فارسی و عربی مرور می‌شوند. در بخش ۳ روش پیشنهادی به طور مفصل تشریح شده و در نهایت در بخش ۴ مراحل انجام آزمایش‌ها و نتایج آنها گزارش می‌شود. سپس در بخش ۵ به بحث در مورد نتایج به دست آمده در مرحله قبل پرداخته و نهایتاً در بخش ۶ نتیجه‌گیری می‌شود.

۲- مروری بر کارهای پیشین در زمینه بازشناسی کلمات دست‌نویس فارسی

در این بخش مروری بر کارهای انجام‌شده در زمینه بازشناسی کلمات دست‌نویس فارسی و عربی ارائه می‌شود.

در [۵] بازشناسی کلمات دست‌نویس برای یک مجموعه از نام‌های صد شهر ایران انجام شده است. در این پژوهش از تعمیم الگوریتم وربش زمانی پویا (DTW)[2] به دو بعد جهت مقایسه کلمات استفاده شده و میزان شناسایی در بهترین حالت‌ها و با استفاده از ویژگی مکان مشخصه برابر با ۷۸/۸ درصد بوده است.

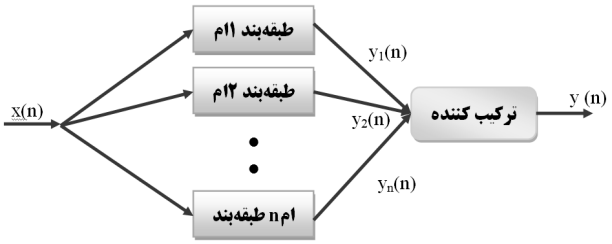
مدل مارکوف پنهان (HMM)[3] در بازشناسی کلمات دارای موفقیت‌های چشمگیری بوده است. در این روش غالباً ویژگی‌ها پس از پیش‌پردازش از طریق یک پنجره لغزان استخراج می‌شوند. دهقان و همکارانش در [۶]، یک سیستم بازشناسی بر اساس چندی‌سازی برداری[4] از طریق نگاشت ویژگی خودسامان (SOFM)[5] و مدل مارکوف پنهان

این مقاله در تاریخ ۱۹ اسفند ماه ۱۳۹۳ دریافت و در تاریخ ۲۴ اسفند ماه ۱۳۹۳ بازنگری شد.

سیدعلی‌اصغر عباس‌زاده آرانی، دانشکده مهندسی برق و کامپیوتر، دانشگاه تربیت مدرس، تهران، (email: saa.abbaszadeh@modares.ac.ir).

احسان‌اله کبیر، دانشکده مهندسی برق و کامپیوتر، دانشگاه تربیت مدرس، تهران، (email: kabir@modares.ac.ir).

1. Negative Correlation Learning
2. Dynamic Time Warping
3. Hidden Markov Model
4. Vector Quantization
5. Self Organization Feature Map



شکل ۱: ادغام طبقه‌بندها (هر طبقه‌بند با توجه به مقدار پارامترهایش به پاسخ متفاوتی برای مسأله می‌رسد و انتظار می‌رود با ترکیب پاسخ‌ها، دقت طبقه‌بندی افزایش پیدا کند).

گسسته برای خواندن نام شهرها در آدرس‌های پستی ارائه کرده‌اند. در این سیستم پس از تکنیک‌های پیش‌پردازش شامل باینری‌کردن، حذف نویز، تصحیح شیب و تخمین خط زمینه، کد زنجیره‌ای کانتور کلمه با یک پنجره لغزان از راست به چپ روبش می‌شود و بردارهای ویژگی استخراج می‌شوند. برای ایجاد کتاب کد و همچنین هموارسازی توزیع‌های احتمال مشاهدات، از نگاشت ویژگی خودسامان استفاده شده است. سپس برای نام هر شهر، یک مدل مارکوف پنهان جداگانه از طریق الگوریتم بام- ولش کلاسیک آموزش داده می‌شود. مؤلفین [۶] در ادامه کار خود یک سیستم مشابه پیشنهاد داده‌اند با این تفاوت که در آن از چندی‌سازی برداری فازی استفاده شده است [۷]. در این روش برای تولید یک کتاب کد فازی، خوشه‌بندی c میانگین فازی، FCM، به کار می‌رود. مجموعه داده‌ای که این گروه برای بازشناسی کلمات دست‌نویس فارسی به کار برده‌اند شامل ۱۷۰۰۰ تصویر از نام ۱۹۸ شهر است. دو روش پیشنهادی آنها روی این مجموعه داده در بهترین حالت به ترتیب نرخ بازشناسی ۶۵ و ۶۷/۲ درصد را به دست داده‌اند.

یافتن ویژگی‌های بهینه از دیگر حوزه‌های پژوهش در زمینه بازشناسی کلمات فارسی و عربی بوده است. در [۸] از تبدیل موجک M باندی بر روی تصویر کلمات استفاده شده و کارایی سیستم را نسبت به ویژگی‌های مبتنی بر فوریه- موجک و گشتاورهای زرنیک افزایش داده است. در این مرجع از فاصله ماهالانوبیس برای طبقه‌بندی استفاده شده است. مجموعه لغات به کار گرفته شده در این تحقیق شامل صد کلمه است که توسط ۱۲ نویسنده مختلف نوشته شده‌اند. نتیجه بازشناسی در بهترین حالت ۹۶ درصد به دست آمده است.

۲-۱ مروری بر ادغام طبقه‌بندها

بسیاری از الگوریتم‌های یادگیری در حقیقت یک نوع جستجوی محلی انجام می‌دهند که ممکن است در کمینه‌های محلی گرفتار شوند. اجرای الگوریتم‌های یادگیری تحت شرایط متفاوت، یادگیری جمعی، روشی است برای آن که بتوان تقریب بهتری از یک طبقه‌بند بهینه را فراهم کرد (شکل ۱).

یک مجموعه از ماشین‌های یادگیری از دسته‌ای از یادگیرنده‌ها استفاده می‌کند تا راه حل‌های جزئی را برای حل مسئله فرا گیرند و سپس این راه حل‌ها را جهت شکل‌دادن یک راه حل نهایی یک‌پارچه می‌سازد. از این روش در پژوهش‌ها تحت عناوین شورای ماشین‌های یادگیرنده، طبقه‌بندهای مرکب یا سیستم‌های طبقه‌بندی چندگانه نام برده می‌شود. به یک ماشین یادگیرنده تنها اصطلاحاً یادگیرنده پایه اطلاق می‌شود [۹].

برای آن که ترکیب خروجی طبقه‌بندها مفید باشد، باید نیازهای خاصی برآورده شوند. اولین نیاز برای طراحی سیستم طبقه‌بندی مرکب، به کارگیری یک چارچوب ریاضی مناسب برای قاعده ترکیب است تا به طریقی از نقاط قوت طبقه‌بندها استفاده و از نقاط ضعف آنها پرهیز شود. دومین نیاز، وجود تعدادی طبقه‌بند پایه با کارایی قابل قبول و نسبتاً متفاوت از یکدیگر است [۱۰]. از این ویژگی به عنوان گوناگونی طبقه‌بندهای پایه نام برده می‌شود.

مهم‌ترین بخش طراحی یک سیستم بازشناسی چندگانه، ایجاد طبقه‌بندهای پایه دقیق و گوناگون است. روش‌های ایجاد گوناگونی بر مبنای نحوه ایجاد تفاوت در فرایند یادگیری به دو دسته ضمنی[1] و صریح[2] تقسیم می‌شوند. روش‌های ضمنی با تغییرات ضمنی در فرایند یادگیری طبقه‌بندهای پایه سعی در گوناگون‌کردن آنها دارند. در این روش‌ها، طی یادگیری معیاری از گوناگونی بررسی نمی‌شود و لذا تضمینی برای متفاوت‌شدن طبقه‌بندها وجود ندارد بلکه فقط امیدوار هستیم که خطاهای طبقه‌بندهای ایجادشده با یکدیگر متفاوت باشند. مهم‌ترین روش‌های ضمنی برای ایجاد گوناگونی عبارتند از استفاده از طبقه‌بندهای پایه متفاوت، بازنمایی متفاوت الگوها و استفاده از مجموعه آموزشی متفاوت.

روش‌های صریح با تحت تأثیر قراردادن مسیر یادگیری، طبقه‌بندهای پایه متفاوتی را ایجاد می‌کنند. این روش‌ها در فرایند یادگیری طبقه‌بندها، معیاری از گوناگونی را اعمال کرده و بر اساس آن مسیر یادگیری طبقه‌بندها را در فضای یادگیری، تغییر می‌دهند تا طبقه‌بندهایی گوناگون در خطا ایجاد شوند.

۲-۲ ادغام طبقه‌بندها در بازشناسی دست‌نوشته

استفاده از ادغام طبقه‌بندهای چندگانه برای بازشناسی دست‌نوشته به دهه ۹۰ باز می‌گردد. در [۱۱] پژوهشگران بر اساس سطوح اطلاعاتی که از طریق طبقه‌بندهای مختلف به دست می‌آیند، روش‌هایی را برای ادغام طبقه‌بندهای پایه پیشنهاد نموده‌اند. در آزمایش‌هایی که برای بازشناسی ارقام دست‌نویس انجام شده، روش‌های پیشنهادی نسبت به طبقه‌بندهای تنها بهبود معناداری را در نرخ بازشناسی نشان می‌دهد. در [۱۲] پژوهشگران ابتدا سه روش بازشناسی کلمات دست‌نویس با استراتژی مختلف را با یکدیگر مقایسه و در نهایت برای رسیدن به نتایج بهتر ادغام آنها را پیشنهاد کرده‌اند. در روش اول، یک تکنیک کل‌نگر برای بازشناسی تصاویر کلمات به کار می‌رود. دو روش دیگر هر چند مبتنی بر روش‌های جزءنگر هستند اما برای افزایش گوناگونی بین خروجی طبقه‌بندهای پایه از الگوریتم‌های قطعه‌بندی و طبقه‌بندی مختلف استفاده می‌شود.

روش‌های متنوعی برای ایجاد طبقه‌بندهای پایه گوناگون وجود دارند که می‌توان به روش‌های بگینگ[3]، بوستینگ[4]، بگینگ نصف- نصف[5] و زیرفضای تصادفی[6] اشاره کرد که همه این روش‌ها در [۱۳] برای بازشناسی برون خط کلمات دست‌نویس به کار رفته‌اند.

برای ادغام طبقه‌بندهای پایه در یک روش جمعی نیز روش‌های مختلفی پیشنهاد شده است. روش‌های ادغام مثل روش فضای دانش- رفتار [۱۴]، روش شمارش بوردا [۱۵]، شبکه عصبی پرسپترون [۱۶] و [۱۷]، روش‌های مبتنی بر رأی‌گیری [۱۸]، تئوری دمپسترشفر [۱۹] و [۲۰]، کلیشه تصمیم [۲۱] و بهینه‌سازی گروه ذرات [۲۲] از جمله روش‌هایی هستند که در بازشناسی دست‌نوشته از طریق ادغام طبقه‌بندها پیشنهاد شده‌اند.

1. Implicit
2. Explicit
3. Bagging
4. Boosting
5. Half & Half Bagging
6. Random Subspace

استفاده از روش‌های ضمنی برای گوناگون‌کردن خطای طبقه‌بندهای پایه در بازشناسی کلمات دست‌نویس بسیار متداول است. در بسیاری از پژوهش‌ها با استفاده از طبقه‌بندهای پایه با معماری و قاعده یادگیری متفاوت، گوناگونی ایجاد می‌شود. به عنوان مثال در [۲۳] ادغام شبکه‌های عصبی و مدل‌های مارکوف پنهان با بیشترین حاشیه مورد بررسی قرار گرفته است. این پژوهش نشان می‌دهد که با وجود این که هر یک از طبقه‌بندهای پایه پیشرفت‌های چشمگیری در بازشناسی دست‌نوشته داشته‌اند اما ادغام آنها در بازشناسی بهبود مؤثرتری ایجاد می‌کند.

استفاده از ویژگی‌های مختلف روشی دیگر برای ایجاد گوناگونی به صورت ضمنی است [۲۴]. در [۲۵]، دو تکنیک استخراج ویژگی نگاشت‌های لبه بهبودیافته و منطقه‌بندی چندگانه برای بازشناسی نویسه‌های شکسته پیشنهاد شده است. نتایج آزمایش نشان داده‌اند که این دو ویژگی، مستقل و مکمل یکدیگر هستند و به همین دلیل برای به دست آوردن نرخ بازشناسی بیشتر، ادغام خروجی طبقه‌بندهای پایه‌ای که با استفاده از مجموعه ویژگی‌های متفاوت و به طور جداگانه آموزش داده شده‌اند پیشنهاد گردیده است. این گروه پژوهشگر در [۲۶] هم‌پوشانی خطای طبقه‌بندهای آموزش‌دیده از طریق ویژگی‌های مبتنی بر ساختار، منطقه‌بندی و تقعر را بررسی نموده‌اند که نشان می‌دهد این ویژگی‌ها خطاهای متفاوتی در بازشناسی اعداد دست‌نویس ایجاد می‌کنند. ادغام از طریق شبکه پرسپترون چندلایه، بهترین نتیجه را در این پژوهش‌ها به دست داده است. در [۲۲] به هر یک از طبقه‌بندهای پایه ویژگی‌های استخراج‌شده از بخشی از تصویر ورودی اعمال می‌شود. این روش باعث ایجاد گوناگونی بین طبقه‌بندهای پایه که از نوع SVD[1] هستند می‌شود. سپس طبقه‌بندهای پایه از طریق یک روش مبتنی بر الگوریتم بهبودیافته گروه ذرات ترکیب می‌شوند. یکی از مزایای این روش که برای طبقه‌بندی ارقام دست‌نویس فارسی و عربی طراحی شده این است که به اندازه مجموعه آموزشی حساسیت کمی دارد.

در بیشتر مراجع، گوناگونی هم برای مجموعه‌های ویژگی و هم برای نوع طبقه‌بندها در نظر گرفته می‌شود. به عنوان مثال در [۲۷] و [۲۸] یک سیستم برخط برای بازشناسی متون دست‌نویس ارائه شده که دو مجموعه ویژگی متفاوت، مستخرج از داده‌های برخط و برون‌خط را برای آموزش دو طبقه‌بند پایه متفاوت (مدل مارکوف پنهان و BLSTM[2]) به کار می‌برد. در [۲۹] یک سیستم طبقه‌بندی چندگانه برون‌خط برای بازشناسی دست‌نوشته‌های عربی ارائه شده است. این سیستم علاوه بر به کارگیری دو سیستم بازشناسی جداگانه مبتنی بر شبکه ART[3] فازی و توابع پایه شعاعی (RBF)[4]، از دو مجموعه ویژگی متفاوت مبتنی بر گشتاورهای هو و زرنیک برای آموزش استفاده می‌کند. جهت ایجاد تصمیم نهایی در این سیستم از طرح‌های ادغام متفاوت استفاده می‌شود که بهترین آنها، حدود ۶ درصد از بهترین طبقه‌بند پایه روی داده IFN/ENIT، نرخ بازشناسی بیشتری به دست می‌دهد. در [۳۰] مجموعه ویژگی‌هایی که وجوه مختلف تصویر نویسه‌های شکسته را برای اهداف بازشناسی توصیف می‌کنند، به طبقه‌بندهای گوناگون در یک روش جمعی اعمال می‌شوند. نتایج آزمایشی در این پژوهش نشان داده که با ادغام بهترین طبقه‌بندهای پایه و مجموعه‌های ویژگی می‌توان نرخ بازشناسی را به طور مؤثری نسبت به روش‌های بازشناسی موجود بهبود داد.

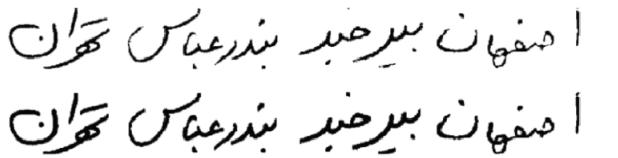
ویژگی مشترک بین تمام روش‌های ضمنی این است که در همه آنها با فراهم‌نمودن شرایط ایجاد گوناگونی بین طبقه‌بندهای پایه، سعی می‌شود که خطای آنها در فضای داده‌های ورودی متفاوت شوند. در این روش‌ها، وجوه مختلفی از داده ورودی به طبقه‌بندهای پایه اعمال می‌شود و از این طریق می‌توان امیدوار بود که خطاهای طبقه‌بندهای پایه با یکدیگر متفاوت باشند. آن گاه از طریق یک روش ادغام مناسب می‌توان تمام فضای داده ورودی را پوشش داد اما تضمینی برای به وجود آمدن چنین شرایط بهینه‌ای وجود ندارد. این مهم‌ترین عیب روش‌های ضمنی ارزیابی می‌شود. در مقابل، روش‌های صریح در فرایند یادگیری طبقه‌بندهای پایه دخالت نموده و با اعمال معیار مناسبی از گوناگونی، روند یادگیری آنها را تحت تأثیر قرار می‌دهند. لذا از این منظر روش‌های صریح به روش‌های ضمنی برتری دارند. به عنوان مثال در [۳۱] یک سیستم طبقه‌بندی چندگانه بر اساس معیارهای گوناگونی برای بازشناسی دست‌نوشته‌های عربی پیشنهاد شده که در آن ابتدا مجموعه بزرگی از طبقه‌بندهای پایه تشکیل می‌شود. سپس زیرمجموعه‌ای بهینه از میان این مجموعه انتخاب می‌شود و بدین منظور از ۶ معیار گوناگونی مختلف استفاده می‌شود. در [۳۲] روشی مبتنی بر الگوریتم ژنتیک با اهداف چندگانه سلسله مراتبی برای بازشناسی دست‌نوشته پیشنهاد شده که در آن مجموعه‌ای از طبقه‌بندهای پایه بهینه بر اساس کمینه‌کردن خطا و تعداد ویژگی‌ها تولید و سپس در راستای بیشینه‌نمودن دقت سیستم و یک معیار گوناگونی، ادغام می‌شوند.

روش‌های جریمه‌ای از مهم‌ترین روش‌های صریح برای ایجاد گوناگونی در طبقه‌بندهای پایه هستند. روش‌های جریمه‌ای با اضافه‌کردن یک مؤلفه جریمه به تابع خطای طبقه‌بندهای پایه مسیر یادگیری آنها را متفاوت می‌سازند. مؤلفه جریمه معمولاً یک مؤلفه مبتنی بر شباهت خطای طبقه‌بندها است. روزن در [۳۳] روشی را معرفی می‌نماید که در آن شبکه‌ها به صورت متوالی آموزش می‌بینند و حاصل‌ضرب بایاس شبکه‌های پایه به عنوان مؤلفه جریمه در نظر گرفته می‌شود. ایراد عمده روش روزن این است که چون شبکه‌ها به صورت متوالی آموزش می‌بینند، فرایند یادگیری سیستم مرکب طولانی است. با استفاده از روش NCL، در [۳۴] و [۳۵] همه شبکه‌ها به صورت هم‌زمان آموزش می‌بینند. در این روش هر شبکه در طی یادگیری نه تنها خطایش را نسبت به خروجی مطلوب کاهش می‌دهد، بلکه تفاوتش را با شبکه‌های دیگر نیز افزایش می‌دهد و پارامترهای یادگیریش را بر اساس کمینه‌شدن خطا و بیشینه‌شدن گوناگونی تنظیم می‌کند.

۳- بازشناسی کلمات دست‌نویس از طریق آموزش طبقه‌بندها با ایجاد همبستگی منفی بین آنها

برای بازشناسی کلمات دست‌نویس ابتدا مراحل پیش‌پردازش شامل آستانه‌گیری، کاهش نویز، عملیات هموارسازی و یکنواخت‌سازی پهنای قلم روی تصاویر کلمات ورودی اعمال می‌شود و سپس ویژگی‌های مناسب از آنها استخراج می‌شود. برای ارزیابی بهتر روش پیشنهادی از طیف متنوعی از ویژگی‌های مطرح‌شده در مراجع، از نظر عملکرد و پیچیدگی محاسباتی استفاده شده است. در این پژوهش ویژگی منطقه‌ای مبتنی بر چگالی نقاط سیاه، ویژگی مبتنی بر گرادیان تصویر مطرح‌شده در [۳۶] و ویژگی منطقه‌ای مبتنی بر کد زنجیره‌ای کانتور، از تصاویر کلمات استخراج می‌شوند. ویژگی‌های به دست آمده در آموزش سیستم طبقه‌بندی چندگانه به کار گرفته می‌شوند.

1. Singular Value Decomposition
2. Bidirectional Long Short-Term Memory
3. Adaptive Resonance Theory
4. Radial Basis Functions



شکل ۲: فرایند پیش‌پردازش شامل هموارسازی تصویر و یکنواخت‌سازی پهنای قلم.

۳-۱ پیش‌پردازش

در این مرحله تمام تصاویر کلمات مورد آزمایش، ابتدا با یک مقدار آستانه مناسب آستانه‌گیری شده و پس از حذف نویز حاشیه سفید آنها حذف می‌شوند. سپس از یک روش مناسب برای یکنواخت‌سازی پهنای قلم استفاده می‌شود. بدین منظور، ابتدا پهنای قلم از طریق روشی که در [۳۷] مطرح شده است، تخمین زده می‌شود. در این روش ابتدا تصویر کلمه ستون به ستون جاروب شده و فراوان‌ترین اندازه پاره‌خط‌های سیاه در ستون‌ها به عنوان پهنای قلم در نظر گرفته می‌شود. آن گاه در هر ستون از تصویر پاره‌خط‌هایی که اندازه آنها کمتر از پهنای قلم است از دو طرف گسترش داده می‌شوند تا به پهنای قلم برسند. سپس یک عملگر گسترش با عنصر ساختاری ۳×۳ روی تصویر اعمال می‌شود. شکل ۲ تصاویر چند کلمه را قبل و بعد از هموارسازی و یکنواخت‌سازی پهنای قلم نشان می‌دهد.

۳-۲ مروری بر ادغام طبقه‌بندها

برای آن که بهبود روش پیشنهادی بهتر خود را نشان دهد در این تحقیق از چند نوع ویژگی پرکاربرد استفاده شده است. یکی از این روش‌ها، روش منطقه‌ای مبتنی بر چگالی نقاط سیاه است [۳۸]. در این روش یک شبکه ۸×۶ روی تصویر کلمه قرار می‌گیرد و تصویر را منطقه‌بندی می‌نماید. در هر منطقه تعداد پیکسل‌های سیاه‌رنگ به کل پیکسل‌ها محاسبه می‌شود. بنابراین به ازای هر تصویر ۴۸ ویژگی به دست می‌آید. در استخراج ویژگی مبتنی بر گرادیان، از عملگر سوبل برای محاسبه مؤلفه‌های x و y گرادیان تصویر استفاده می‌شود. در این روش، زاویه گرادیان در ۸ جهت فریمن[1] کوانتیزه شده و برای هر جهت یک صفحه جهت[2] ساخته می‌شود. سپس روی هر صفحه جهت ۶۴ نقاب گوسی با مراکزی که در کل صفحه به طور یکنواخت توزیع شده‌اند، اعمال می‌شود. از این طریق از هر صفحه جهت ۶۴ و از هر تصویر ۵۱۲ ویژگی مبتنی بر گرادیان استخراج می‌شود [۳۶]. در نهایت از تکنیک آنالیز مؤلفه‌های اساسی (PCA)[3] برای کاهش بعد از ۵۱۲ به ۵۰ استفاده می‌شود. سومین ویژگی بر اساس کد زنجیره‌ای کانتور استخراج شده است [۳۸]. در این روش ابتدا کل تصویر با یک شبکه ۶×۴ منطقه‌بندی می‌شود و سپس در هر منطقه، کدهای زنجیره‌ای پیکسل‌های کانتور در ۱۶ جهت استخراج می‌شوند. بر اساس فراوانی کدها در هر یک از جهات، برای هر منطقه ۱۶ ویژگی محاسبه می‌شود. بنابراین از هر تصویر در مجموع ۳۸۴ ویژگی به دست می‌آید که از طریق تحلیل مؤلفه‌های اساسی به ۵۰ ویژگی کاهش می‌یابد. جدول ۱ به طور خلاصه ابعاد ویژگی‌های به کار رفته در این پژوهش را نشان می‌دهد.

۳-۳ آموزش طبقه‌بندها از طریق یادگیری با همبستگی منفی

هر یک از طبقه‌بندهای پایه یک شبکه عصبی پرسپترون چندلایه با یک لایه پنهان است که از طریق پس‌انتشار خطا آموزش می‌بینند. در این روش، خروجی نورون‌های لایه خروجی و پنهان، تابعی از داده ورودی، وزن اتصالات بین نورونی و نوع تابع فعالیت است. تابع فعالیت در این پژوهش، تابع سیگموئید در نظر گرفته شده است. در روش آموزش از طریق پس‌انتشار خطا در هر مرحله از آموزش، داده‌های آموزشی که همان ویژگی‌های تصاویر کلمات دست‌نویس هستند به شبکه‌های پایه اعمال شده و سپس از طریق محاسبه تابع خطا میزان تغییر پارامترهای شبکه تعیین می‌شود. خطای هر لایه به جز لایه آخر با توجه به بردار خطای لایه بعد محاسبه می‌گردد. لذا تابع خطا نقش به سزایی در آموزش طبقه‌بندهای پایه دارد. در روش یادگیری با همبستگی منفی، یک ترم جریمه که میزان شباهت خروجی طبقه‌بندهای پایه را محاسبه می‌نماید به تابع خطا افزوده می‌شود

$$E_i = \frac{1}{2}(O_i - y)^2 + \lambda P_i \qquad (1)$$

که در آن E_i تابع خطا، y بردار خروجی مطلوب و λ ضریب وزنی مؤلفه جریمه P_i است. مؤلفه جریمه به گونه‌ای انتخاب می‌شود که اشتراک خطا برای طبقه‌بندهای پایه نوعی خطا محسوب شود. مؤلفه جریمه در روش یادگیری با همبستگی منفی به شکل زیر است

$$P_i = -\left\| O_i - \bar{O} \right\|^2 \qquad (2)$$

که در آن $\bar{O}$ میانگین خروجی طبقه‌بندهای پایه است. لذا روابط اصلاح وزن‌های شبکه‌های عصبی از طریق پس‌انتشار خطا عبارتند از

$$\Delta w_{y,i} = \eta_e[(y - O_i) - \lambda \nabla_{O_i} P_i].(O_i(1 - O_i))O_{hi}^T \qquad (3)$$

$$\Delta w_{h,i} = \eta_e w_y^T[(y - O_i) - \lambda \nabla_{O_i} P_i]. (O_i(1 - O_i))O_{h_i}(1 - O_{h_i})x^T \qquad (4)$$

$$\nabla_{O_i} P_i = -2(1 - \frac{1}{M})(O_i - \bar{O}) \qquad (5)$$

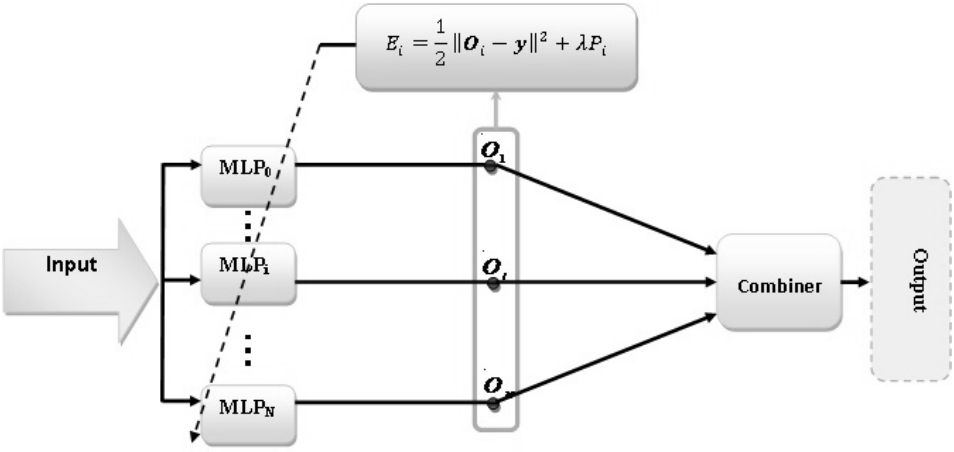
که در آن η_e نرخ یادگیری، λ پارامتر کنترل روش آموزش با همبستگی منفی، w_h و w_y به ترتیب وزن‌های ورودی به لایه پنهان و لایه پنهان به لایه خروجی هستند. O_h^T ترانهاده O_h است که نماینده خروجی لایه پنهان شبکه است.

پس از آموزش طبقه‌بندهای پایه از طریق آموزش با همبستگی منفی، خروجی آنها از طریق یک قاعده مناسب ادغام می‌شوند. در این تحقیق قواعد ادغام میانگین، بیشینه، حاصل‌ضرب، رأی‌گیری و کلیشه تصمیم آزمایش می‌شوند. بلوک دیاگرام روش پیشنهادی در شکل ۳ نشان داده شده است.

۴- نتایج آزمایش‌ها

در این بخش ابتدا مجموعه داده انتخابی معرفی شده و سپس نتایج آزمایش‌هایی که برای دسته‌بندی این مجموعه داده انجام شده‌اند گزارش می‌شود. جهت مقایسه بهتر، ویژگی‌های استخراج‌شده به طبقه‌بندهای مرکب، روش‌های جمعی توانمند و روش پیشنهادی اعمال می‌شوند.

1. Freeman
2. Direction Plane
3. Principal Component Analysis



شکل ۳: روش پیشنهادی برای بخش طبقه‌بندی (در مرحله اول طبقه‌بندهای پایه از طریق روش NCL آموزش داده می‌شوند و سپس از طریق یک روش مناسب ادغام می‌شوند).

جدول ۱: ابعاد ویژگی‌های استخراج‌شده از کلمات دست‌نویس در این پژوهش و ابعاد آنها.

نوع ویژگی	اندازه اولیه	اندازه پس از کاهش بعد
ویژگی منطقه‌ای مبتنی بر چگالی نقاط سیاه	۴۸	۴۸
ویژگی مطرح‌شده در [۳۶]	۵۱۲	۵۰
ویژگی منطقه‌ای مبتنی بر کد زنجیره‌ای کانتور	۳۸۴	۵۰

جدول ۲: نرخ بازشناسی از طریق برخی از روش‌های ادغام موجود (میانگین و انحراف معیار).

	ویژگی منطقه‌ای مبتنی بر چگالی نقاط سیاه	ویژگی مطرح‌شده در [۳۶]	ویژگی منطقه‌ای مبتنی بر کد زنجیره‌ای کانتور
میانگین	۷۸٫۹۷ (۲۳٫۶۴)	**۹۴٫۰۶** (۱۱٫۵۳)	۹۲٫۲۶ (۱۷٫۲۷)
بیشینه	۷۸٫۱۹ (۸۵٫۲۴)	۹۳٫۹۴ (۱۳٫۰۳)	۹۰٫۴۵ (۱۱٫۹۵)
حاصل‌ضرب	۷۹٫۱ (۲۳٫۶۴)	**۹۴٫۰۶** (۱۱٫۵۳)	**۹۲٫۵۲** (۱۵٫۹۴)
رأی‌گیری	**۴۸٫۷۹** (۱۵٫۷)	۹۳٫۲۹ (۶٫۱۵)	۸۹٫۸۱ (۱۳٫۳۹)
کلیشه تصمیم	۷۷٫۴۲ (۲۷٫۷۴)	۹۳٫۸۱ (۹٫۰۷)	۹۲٫۱۳ (۱۱٫۱۱)

٤-١ معرفی مجموعه داده

مجموعه داده‌ای که برای انجام آزمایش‌ها انتخاب شده‌اند شامل تصاویر نام مراکز استان‌های ایران است که از پایگاه داده "ایرانشهر" انتخاب شده‌اند. این پایگاه داده شامل ۵۰۲ شهر کشور ایران است که از میان آنها تصاویر نام ۳۱ مرکز استان انتخاب شده و بنابراین مجموعه داده انتخابی دارای ۳۱ کلاس است. برای نام هر شهر ۲۵ نمونه در دسترس است که توسط ۲۵ نویسنده مختلف نوشته شده‌اند. بنابراین مجموعه داده انتخاب‌شده به طور کلی دارای ۷۷۵ نمونه است. برای ارزیابی روش پیشنهادی از روش ارزیابی k بخشی استفاده شده که در آن $k = ۵$ است و لذا در هر مرحله از ۲۵ تصویر در هر کلاس ۱۵ تصویر برای آموزش، ۵ تصویر برای ارزیابی و ۵ تصویر برای آزمایش به کار می‌رود.

همه تصاویر در کیفیت ۹۶ نقطه در اینچ و به فرم سطح خاکستری اسکن شده‌اند. برخی از این تصاویر در شکل ۴ نمایش داده شده‌اند.

٤-٢ طبقه‌بندی از طریق ادغام طبقه‌بندها

جهت طبقه‌بندی از طریق سیستم‌های طبقه‌بندی چندگانه، ۶ طبقه‌بند پایه که شبکه‌های عصبی پرسپترون با یک لایه پنهان و دارای ۳۵ گره در لایه میانی هستند، از طریق روش پس‌انتشار به عقب با نرخ یادگیری ۰٫۱ آموزش داده شده‌اند. وزن‌های اولیه در این طبقه‌بندها به طور تصادفی و مستقل از یکدیگر تخصیص داده شده است. این عمل به طور ضمنی طبقه‌بندهای پایه را گوناگون می‌کند. خروجی طبقه‌بندها در مرحله آزمایش از طریق روش‌هایی نظیر میانگین، بیشینه، حاصل‌ضرب، رأی‌گیری و کلیشه تصمیم ادغام شده‌اند. نرخ بازشناسی هر یک از روش‌های ادغام مطرح‌شده در جدول ۲ نشان داده شده است.

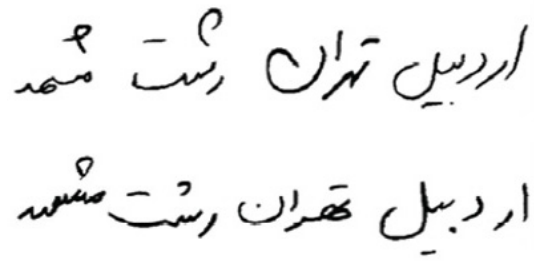
در ادبیات مرتبط با روش‌های جمعی، رویکردهایی مطرح شده‌اند که توانمندی خود را در کاربردهای مختلف نشان داده‌اند. از این دسته می‌توان به روش‌های آدابوست[1]، بگینگ، زیرفضای تصادفی و تعمیم پشته‌ای[2] اشاره نمود. این روش‌ها به صورت ضمنی، طبقه‌بندهای پایه را گوناگون نموده و سپس ادغام می‌نمایند. در روش تعمیم پشته‌ای از یک ادغام‌کننده آموزش‌پذیر استفاده می‌شود که معمولاً یک شبکه عصبی است. نرخ بازشناسی از طریق طبقه‌بندهای جمعی مطرح‌شده در جدول ۳ نمایش داده شده است.

خواننده می‌تواند برای مطالعه جزئیات روش‌های ادغام و طبقه‌بندهای جمعی توانمندی که در این بخش به کار رفته‌اند به [۳۹] یا [۴۰] و [۴۱] مراجعه نماید و از بیان شرح عملکرد آنها برای اختصار خودداری شده است.

٤-٣ طبقه‌بندی از طریق طبقه‌بندهای با همبستگی منفی

در روش‌های پیشین (رویکردهای ضمنی)، طبقه‌بندهای پایه به صورت مستقل از یکدیگر آموزش می‌دیدند و در طی این فرایند با یکدیگر تعاملی نداشتند، لذا تضمینی وجود ندارد که هر طبقه‌بند بخش متفاوتی از فضای ورودی را پوشش بدهد. اما در روش یادگیری با همبستگی منفی به عنوان

1. AdaBoost
2. Stacked Generalization



شکل ۴: نمونه‌هایی از نام‌های مراکز استان‌های ایران.

یک رویکرد صریح، همه طبقه‌بندهای پایه به طور هم‌زمان و موازی با یکدیگر آموزش می‌بینند و در تعامل با یکدیگر همبستگی خطای خود را کاهش می‌دهند. برای استفاده از این روش، شبکه‌های عصبی با یک لایه پنهان و ۳۵ نورون در لایه میانی، به عنوان طبقه‌بندهای پایه به کار گرفته شده است. پارامترهای یادگیری همانند آزمایش قبل تنظیم شده و مقدار پارامتر کنترل بهینه، λ^*، با استفاده از داده‌های ارزیابی به دست آمده است. در شکل ۵ برای ویژگی منطقه‌ای مبتنی بر چگالی نقاط سیاه، نرخ بازشناسی داده‌های ارزیابی نسبت به پارامتر کنترل، ملاحظه می‌شود.

این نمودار برای دو ویژگی دیگر نیز روند مشابهی دارد. پژوهش‌هایی هستند که نشان می‌دهند همبستگی ضعیفی بین گوناگونی و عملکرد مجموعه وجود دارد [۴۲] و [۴۳]. در مقابل برخی مطرح می‌کنند که بررسی گوناگونی، یکی از محورهای اساسی در ادغام طبقه‌بندها است. شکل ۵ نشان می‌دهد با افزایش مقدار λ در رویکرد یادگیری با همبستگی منفی، نرخ بازشناسی با کمی نوسان روند صعودی دارد و این افزایش تا مقدار آستانه‌ای ادامه پیدا می‌کند. این مقدار آستانه همان مقدار پارامتر کنترل بهینه روش یادگیری با همبستگی منفی، λ^*، است. پس از این مقدار میزان نرخ بازشناسی به طور ناگهانی افت می‌کند و دلیل این پدیده آن است که با افزایش بیش از حد پارامتر کنترل، فرایند گوناگون‌سازی طبقه‌بندهای پایه، مانع یادگیری مناسب آنها می‌شود. لذا کارایی سیستم نسبت به پارامتر λ حساس است و این حساسیت با افزایش تعداد طبقه‌بندهای پایه (و در نتیجه افزایش تعداد پارامترهای آن) زیاد می‌شود. مقدار پارامتر کنترل بهینه روش یادگیری با همبستگی منفی، λ^*، برای ویژگی‌های مبتنی بر منطقه‌بندی، گرادیان و کد زنجیره‌ای و با ساختارهایی دارای ۷، ۶ و ۴ طبقه‌بند به ترتیب برابر با ۰٫۸۵، ۱٫۱ و ۰٫۹۵ است. این مقادیر در حقیقت بهترین موازنه را بین خطای طبقه‌بندهای پایه و گوناگونی آنها برقرار می‌کند.

پس از یافتن تعداد طبقه‌بندهای پایه و پارامتر کنترل در حالت بهینه، ساختار با استفاده از این پارامترها آموزش داده می‌شود. جدول ۴ نرخ بازشناسی روش یادگیری با همبستگی منفی را در حالت بهینه نمایش می‌دهد. در این جدول، نتایج حاصل از روش‌های ادغام و ویژگی‌های مختلف به طور مجزا نمایش داده شده است. برای هر ویژگی تعداد طبقه‌بندهای پایه و پارامتر کنترل بهینه نیز نشان داده شده است.

۵- بررسی نتایج و ردیابی علل خطا

با نگاهی به جدول ۴ ملاحظه می‌شود که به ازای پارامتر کنترل بهینه یعنی $\lambda = \lambda^*$، نرخ بازشناسی به طور قابل توجهی افزایش می‌یابد. این افزایش نشان می‌دهد از طریق موازنه مناسب بین افزایش گوناگونی بین طبقه‌بندهای پایه و دقت آنها، می‌توان ترکیبی از طبقه‌بندها را به طور بهینه به دست آورد. در این جدول ملاحظه می‌شود که برای ویژگی‌های مبتنی بر منطقه‌بندی و کد زنجیره‌ای کانتور، ادغام‌کننده میانگین و برای ویژگی‌های مبتنی بر گرادیان، قاعده رأی‌گیری بهترین نتایج را در پی داشته‌اند. این روش افزایش گوناگونی بین طبقه‌بندهای پایه را تضمین می‌کند که این امر از مزایای رویکرد صریح ایجاد گوناگونی است.

برای تحلیل بهتر عملکرد روش پیشنهادی، معیارهای ارزیابی فراخوانی[1] و دقت[2] را بررسی می‌کنیم [۴۴]. معیار فراخوانی برای یک کلاس بیان می‌کند که طبقه‌بند چه نسبتی از داده‌های متعلق به آن کلاس را درست بازشناسی می‌کند. معیار دقت برای یک کلاس بیان می‌کند که چند درصد از الگوهایی که طبقه‌بند آنها را در آن کلاس قرار داده واقعاً متعلق به آن کلاس هستند. نمودارهای شکل ۶ مقادیر فراخوانی و دقت را برای شهرهای مختلف نشان می‌دهد.

با بررسی نمودار فوق می‌توان شهرهایی را یافت که هم فراخوانی و هم دقت خوب دارند مثل "خرم‌آباد" و شهرهای دیگری هستند که هم فراخوانی و هم دقت آنها نسبتاً کم است مثل "زنجان". شهرهایی وجود دارند که هر چند فراخوانی خوبی دارند اما دقت آنها پایین است، به این معنا که الگوهایی که سیستم آنها را به این کلاس‌ها اختصاص می‌دهد کاملاً قابل اطمینان نیستند مثل "ارومیه". شهرهایی نیز هستند که هر چند فراخوانی آنها کم است اما دقت آنها صد در صد است، به این معنا که می‌توان به دسته‌بندی الگوها به این کلاس‌ها توسط سیستم کاملاً اطمینان کرد مثل "اصفهان".

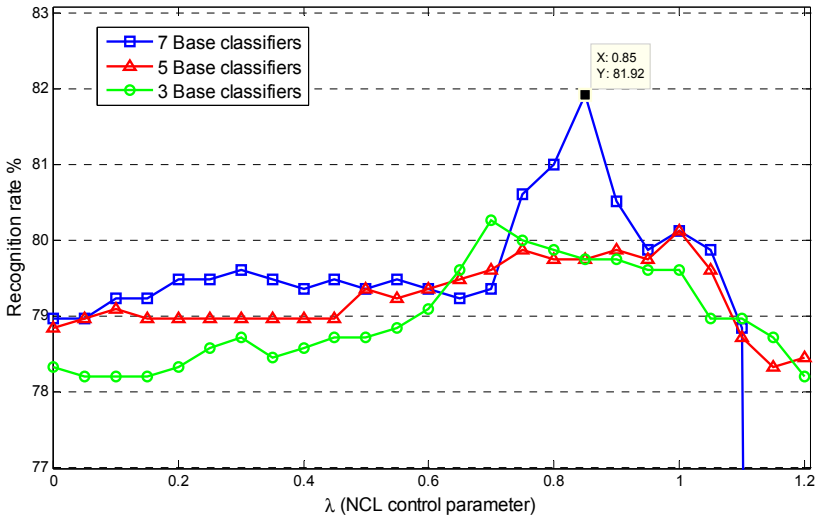
اکنون به بررسی خطاهای رویکرد پیشنهادی می‌پردازیم:

- الف) وجود "الف" و یک حرف با دایره مستقیم مثل "ن" و "س" در انتهای اکثر کلمات خطاهایی را در بازشناسی کلمات مشابه ایجاد می‌کند.
- ب) برخی از کلمات کاملاً بدخط نوشته شده‌اند و چون با دیگر الگوهای کلاس خود متفاوتند باعث ایجاد خطا می‌شوند.
- ج) گاهی نیمه ابتدایی کلمات شباهت زیادی با یکدیگر دارند.
- د) قراردادن "ن" به شکل مستقیم و معکوس در بالای حرف "ا" نیز باعث اشتباه در بازشناسی می‌شود.
- ه) "ن" یک دایره مستقیم است و نوشتن آن به شکل دایره معکوس (مانند حرف "ج") باعث بروز خطا می‌شود.
- و) خطاهایی نیز وجود دارند که وجود آنها از نظر نگارندگان توجیه‌پذیر نیست.

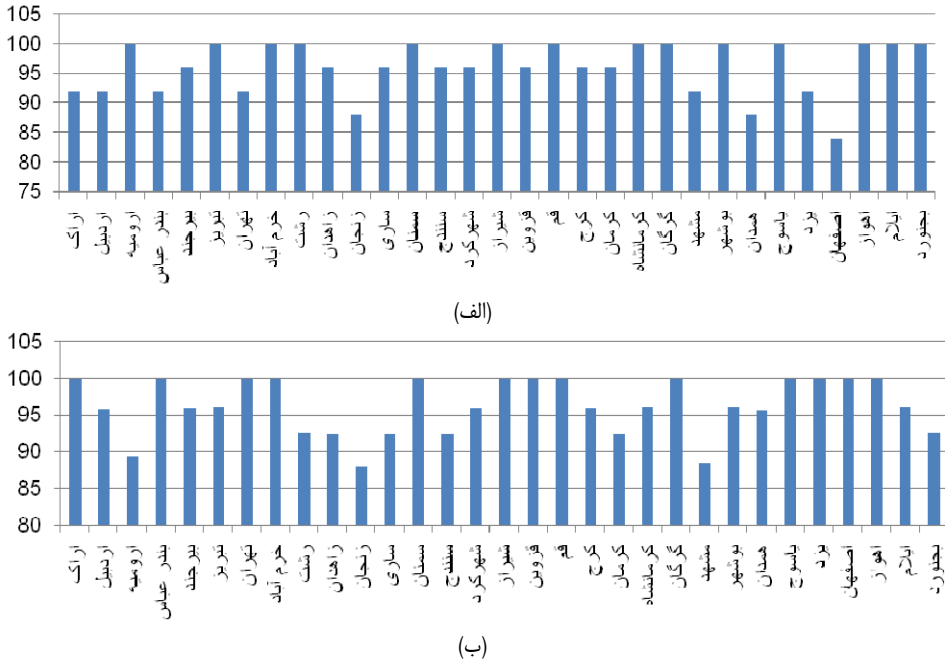
شکل ۷ نمونه‌هایی از خطاهای مطرح‌شده را نشان می‌دهد.

با توجه به نوع خطاهایی که در بالا مطرح شد به نظر می‌رسد با اصلاح مراحل پیش‌پردازش و استخراج ویژگی می‌توان خطاهای سیستم بازشناسی را کاهش داد. به عنوان مثال می‌توان با شناسایی "ا" و "ن" در انتهای کلمات، آنها را از فرایند استخراج ویژگی حذف نمود. در مورد کلماتی که حرف "ن" در بالای حرف "ا" قرار می‌گیرد یا کلماتی که "ن" با دایره معکوس نوشته شده است می‌توان کلاس جداگانه‌ای تعریف نمود. بسیاری از مشکلات موجود به دلیل کل‌نگر بودن روش بازشناسی به وجود آمده‌اند. این گونه خطاها را می‌توان با استفاده از روش‌های پس‌پردازش مناسب اصلاح نمود. در مورد کلماتی که از نظر ویژگی‌های آماری با دیگر الگوهای کلاس خود متفاوتند، می‌توان ویژگی‌های ساختاری استخراج نمود و نتایج حاصل از طبقه‌بندی آنها را با ویژگی‌های آماری ادغام نمود.

1. Recall
2. Precision



شکل ۵: یافتن مقدار λ^* (پارامتر کنترل NCL جهت ایجاد طبقه‌بندهای با همبستگی منفی از طریق داده ارزیابی).



(الف)

(ب)

شکل ۶: (الف) معیار فراخوانی و (ب) معیار دقت برای هر یک از کلاس‌های مسأله.

جدول ۳: نرخ بازشناسی از طریق برخی از روش‌های جمعی توانمند (میانگین و انحراف معیار).

	ویژگی منطقه‌ای مبتنی بر چگالی نقاط سیاه	ویژگی مطرح‌شده در [۳۶]	ویژگی منطقه‌ای مبتنی بر کد زنجیره‌ای کانتور
روش آدابوست	۷۷٫۶۵ (۱۱٫۱)	۹۳٫۲۹ (۶٫۱۵)	**۹۰٫۹٤** (۱۴٫۲۵)
بگینگ	**۸۰٫٦٤** (۱۲٫۹)	**۹۳٫٤۱** (۵٫۷۱)	۹۰٫۴۲ (۱۴٫۲۸)
تعمیم پشته‌ای	۷۹٫۳۵ (۱۶٫۰۴)	۹۳٫۲۹ (۶٫۱۵)	۸۹٫۵۲ (۱۴٫۶۳)
زیرفضای تصادفی	۷۸٫۹۷ (۱۳٫۶۵)	۹۱٫۶۱ (۱۳٫۹۳)	۸۹٫۹۱ (۲۳٫۸)

جدول ۴: نرخ بازشناسی سیستم طبقه‌بندی چندگانه یادگیری با همبستگی منفی برای ویژگی‌های مختلف و روش‌های ادغام مختلف (میانگین و انحراف معیار).

	ویژگی منطقه‌ای مبتنی بر چگالی نقاط سیاه	ویژگی مطرح‌شده در [۳۶]	ویژگی منطقه‌ای مبتنی بر کد زنجیره‌ای کانتور
تعداد طبقه‌بندهای پایه	۷	۶	۴
مقدار λ بهینه	۰٫۸۵	۱٫۱	۰٫۹۵
میانگین	**۸۱٫٦۵** (۲۱٫۴۴)	۹۵٫۹۷ (۱۱٫۳۲)	**۹۳٫٦۵** (۴٫۲۹)
بیشینه	۸۱٫۱۳ (۱۴٫۶۵)	۹۵٫۰۶ (۱۰٫۲۸)	۹۲٫۳۵ (۷٫۶۲)
حاصل‌ضرب	۷۹٫۳۲ (۱۷٫۳۲)	۹۲٫۸۷ (۱۶٫۷۷)	۹۱٫۹۷ (۱۱٫۴۵)
رأی‌گیری	۸۱٫۵۲ (۱۹٫۰۲)	**۹٦٫۱۰** (۱۰٫۳۲)	۹۲٫۳۵ (۷٫۶۲)
کلیشه تصمیم	۷۹٫۱۹ (۲۹٫۸۴)	۹۵٫۵۸ (۱۰٫۵۳)	۹۲٫۲۳ (۸٫۶۶)

شکل ۷: "الف" تا "و" نمونه‌هایی از انواع ۱تا ۶ خطاهای بازشناسی کلمات دست‌نویس در این پژوهش.

جدول ۵: نرخ بازشناسی هر یک از طبقه‌بندهای پایه در روش جمعی NCL (ویژگی‌های مبتنی بر گرادیان).

درصد بازشناسی	
۴۸٪	طبقه‌بند اول
۷۵/۷۴٪	طبقه‌بند دوم
۵۲/۵۲٪	طبقه‌بند سوم
۴۳/۷۴٪	طبقه‌بند چهارم
۶۸/۱۳٪	طبقه‌بند پنجم
۶۱/۱۶٪	طبقه‌بند ششم
۹۶/۱۰٪	ادغام از طریق رأی‌گیری

نکته‌ای که در اینجا لازم به تذکر است این است که در طبقه‌بند مرکبی که در این پژوهش پیشنهاد شد، هر طبقه‌بند پایه به بخشی از فضای ورودی پرداخته و در آن بخش از فضای ورودی خبره شده است. به درصد بازشناسی هر یک از طبقه‌بندهای پایه در جدول ۵ توجه کنید، ملاحظه می‌شود که درصد بازشناسی طبقه‌بندهای پایه به طرز معناداری نسبت به سیستم ادغام‌شده کم است. آزمایش‌ها نشان می‌دهد هر چه پارامتر کنترل ترم جریمه بزرگ‌تر می‌شود، نرخ بازشناسی طبقه‌بندهای پایه کوچک‌تر می‌گردد. دلیل این پدیده آن است که با افزایش λ، طبقه‌بندهای پایه محلی‌تر شده و کارایی آنها نسبت به کل فضای مجموعه آزمایشی کاهش پیدا می‌کند. این خبرگی محلی از پیش تعیین شده نیست و توسط خود ساختار هدایت می‌شود. از این طریق در فرایند یادگیری نه تنها ماهیت داده ورودی بلکه ماهیت ماشین‌های یادگیرنده پایه نیز در نظر گرفته می‌شود.

از این بررسی ملاحظه می‌شود با وجود این که نرخ بازشناسی هر طبقه‌بند پایه برای کل داده‌های آزمایشی کم است، اما ادغام آنها بسیار مؤثر واقع شده است. اگر از یک روش ادغام مناسب استفاده شود، از این ویژگی، یعنی خبرگی محلی هر طبقه‌بند پایه در یک زیرفضا از فضای کلی الگوهای ورودی، می‌توان برای تعمیم این روش برای بازشناسی کلمات دست‌نویس با فرهنگ لغاتی شامل چند صد کلمه استفاده نمود. زیرا از این طریق حجم بزرگ داده‌ها بین طبقه‌بندهای پایه تقسیم

می‌گردد و سپس در نهایت بر اساس قاعده "تقسیم و غلبه" کل مسأله حل می‌شود. این موضوع می‌تواند در پژوهش‌های آینده مورد بررسی بیشتری واقع شود.

٦- نتیجه‌گیری

در این تحقیق ابتدا روش‌های مطرح در بازشناسی برون‌خط کلمات دست‌نویس فارسی/عربی بررسی شدند. آن گاه روش جدیدی بر اساس سیستم‌های طبقه‌بندی چندگانه مطرح شد که در آن طبقه‌بندهای پایه از طریق آموزش با همبستگی منفی گوناگون می‌شوند. در این روش، طبقه‌بندهای پایه ضمن این که طی آموزش خطای خود را کاهش می‌دهند، شباهت خطای خود را نیز با طبقه‌بندهای دیگر کاهش می‌دهند. در آزمایش‌هایی که با استفاده از تصاویر نام مراکز استان‌های ایران از پایگاه داده "ایرانشهر" انجام شد، روش‌های رایج ادغام طبقه‌بندها و روش آموزش با همبستگی منفی با یکدیگر مقایسه شدند. در این آزمایش‌ها ویژگی‌هایی مبتنی بر منطقه‌بندی، گرادیان و کد زنجیره‌ای کانتور استخراج و به ورودی طبقه‌بندها اعمال شدند. نتایج آزمایش‌ها نشان می‌دهد که روش جمعی آموزش با همبستگی منفی نسبت به تکنیک‌های رایج ادغام در هر سه ویژگی نرخ بازشناسی بیشتری داشته است. این مقدار برای ویژگی‌های مبتنی بر گرادیان تصویر به طور میانگین برابر با ۹۶/۱۰ درصد است، ضمن این که طبقه‌بندی ویژگی‌های مبتنی بر گرادیان نتایج بهتری به دست داده است. با توجه به خبرگی محلی هر طبقه‌بند در یک زیرفضا از کل فضای الگوهای ورودی، استفاده از یک ادغام‌کننده مناسب که این ویژگی را در نظر گیرد نتایج بهتری به دست خواهد داد. بسیاری از خطاهای موجود در این پژوهش ناشی از رویکرد کل‌نگری است که به کار رفته است. با بهره‌گیری از روش‌های مناسب پس‌پردازش می‌توان این خطاها را کاهش داد.

مراجع

[1] A. Brakensiek and G. Rigoll, "Handwritten address recognition using hidden markov models," *Reading and Learning*, vol. 2956, pp. 103-122, 2004.

[2] R. Palacios, A. Gupta, and P. S. Wang, "Handwritten bank check recognition of courtesy amounts," *Int'l J. of Image and Graphics*, vol. 4, no. 2, pp. 203-222, Apr. 2004.

[3] M. Ye, P. A. Viola, S. Raghupathy, H. Sutanto, and C. Li, "Learning to group text lines and regions in freeform handwritten notes," in *Proc. 9th Int. Conf. on Document Analysis and Recognition, IEEE Computer Society*, vol. 1, pp. 28-32, Sep. 2007.

[4] V. Govindaraju and H. Xue, "Fast handwriting recognition for indexing historical documents," in *Proc. of 1st Int'l Workshop on Document Image Analysis for Libraries,* pp. 314-320, 2004.

[۵] ک. مسروری، ا. کبیر و ک. بدیع، "شناسایی کلمات دست‌نویس فارسی با استفاده از الگوریتم DTW،" *فصل‌نامه علمی پژوهشی دانشگاه شاهد*، شماره ۲۵، صص. ۵۹-۶۸، ۱۳۷۸.

[6] M. Dehghan, K. Faez, M. Ahmadi, and M. Shridhar, " Handwritten Farsi (Arabic) word recognition: a holistic approach using discrete HMM," *Pattern Recognition*, vol. 34, no. 5, pp. 1057-1065, May 2001.

[7] M. Dehghan, K. Faez, M. Ahmadi, and M. Shridhar, " Unconstrained Farsi handwritten word recognition using fuzzy vector quantization and hidden Markov models," *Pattern Recognition Letters,* vol. 22, no. 2, pp. 209-214, Feb. 2001.

[8] A. Broumandnia, J. Shanbehzadehb, and M. Rezakhah, "Persian/Arabic handwritten word recognition using M-band packet wavelet transform," *Image and Vision Computing*, vol. 26, no. 6, pp. 829-842, Jun. 2008.

[9] G. Brown, *Diversity in Neural Network Ensembles*, Ph.D. Thesis, University of Birmingham, 2003.

[10] L. I. Kuncheva, Combining Pattern Classifiers: Methods and Algorithms, Wiley-Interscience, 2004.

[11] L. Xu, A. Krzyzak, and C. Y. Suen, "Methods of combining multiple classifiers and their applications to handwriting recognition," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 22, no. 3, pp. 418-435, May/Jun. 1992.

[12] B. Plessis, A. Sicsu, L. Heutte, E. Menu, E. Lecolinet, O. Debon, and J. V. Debon, "A multi-classifier combination strategy for the recognition of handwritten cursive words," in *Proc. Int. Conf. on Document Analysis and Recognition*, pp. 642-645, Oct. 1993.

[13] S. Gunter and H. Bunke, "Ensembles of classifiers for handwritten word recognition," *International J. on Document Analysis and Recognition*, vol. 5, no. 1, pp. 224-232, Jul. 2003.

[14] Y. S. Huang and C. Y. Suen, "A method of combining multiple experts for the recognition of unconstrained handwritten numerals," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 17, no.1, pp. 90-94, Jan. 1995.

[15] B. Verma, P. Gader, and W. Chen, "Fusion of multiple handwritten word recognition techniques," *Patter Recognition Letters*, vol. 22, no. 9, pp. 991-998, Jul. 2001.

[16] N. Farah, M. T. Khadir, and M. Sellami, "Artiffcial neural network fusion: application to Arabic words recognition," in *Proc. European Symposium on Artificial Neural Network*, pp. 151-156, Apr. 2005.

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

[18] H. El Abed and V. Margner, "A framework for the combination of different arabic handwritten word recognition systems," in *Proc. 20th Int. Conf. on Pattern Recognition*, pp. 1904-1907, Aug. 2010.

[19] Y. Kessentini, T. Burger, and T. Paquet, "Evidential combination of multiple HMM classifiers for multi-script handwritting recognition," *Computational Intelligence for Knowledge-Based Systems Design, Lecture Notes in Computer Science*, vol. 6178, pp. 445-454, 2010.

[20] T. Burger, Y. Kessentini, and T. Paquet, "Dempster-shafer based rejection strategy for handwritten word recognition," in *Proc. Int. Conf. on Document Analysis and Recognition*, pp. 528-532, Sep. 2011.

[21] R. Ebrahimpour, R. D. Vahid, and B. Mazloom Nezhad, "Decision templates with gradient based features for Farsi handwritten word recognition," *International J. of Hybrid Information Technology*, vol. 4, no. 1, pp. 1-12, Jan. 2011.

[22] H. Salimi and D. Giveki, "Farsi/Arabic handwritten digit recognition based on ensemble of SVD classifiers and reliable multi-phase PSO combination rule," *International J. on Document Analysis and Recognition*, vol. 16, no. 4, pp. 371-386, Dec. 2012.

[23] V. Frinken, T. Peter, A. Fischer, H. Bunke, T. M. T. Do, and T. Artieres, "Improved handwriting recognition by combining two forms of hidden markov models and a recurrent neural network," *Computer Analysis of Images and Patterns, Lecture Notes in Computer Science*, vol. 5702, pp. 189-196, 2009.

[24] Y. C. Chim, A. A. Kassim, and Y. Ibrahim, "Dual classifier system for hand printed alphanumeric character recognition," *Pattern Analysis and Application*, vol. 1, no. 3, pp. 155-162, Sep. 1998.

[25] R. M. O. Cruz, G. D. C. Cavalcanti, and T. I. Ren, "An ensemble classifier for offline cursive character recognition using multiple feature extraction techniques," in *Proc. Int. Joint Conf. on Neural Networks*, 8 pp., Jul. 2010.

[26] R. M. O. Cruz, G. D. C. Cavalcanti, and T. I. Ren, "Handwritten digit recognition using multiple feature extraction techniques and classifier ensemble," in *Proc. 17th Int. Conf. on Systems, Signals, and Image Processing*, pp. 215-218, 2010.

[27] M. Liwicki and H. Bunke, "Combining diverse on-line and off-line systems for handwritten text line recognition," *Pattern Recognition*, vol. 42, no. 12, pp. 3254-3263, Dec. 2009.

[28] M. Liwicki, H. Bunke, J. A. Pittman, and S. Knerr, "Combining diverse systems for handwritten text line recognition," *Machine Vision and Applications*, vol. 22, no. 1, pp. 39-51, Jan. 2011.

[29] C. Leila, K. Maamar, and C. Salim, "Combining neural networks for Arabic handwriting recognition," in *Proc. 10th Int. Symp. on Programming and Systems*, pp. 74-79, Apr. 2011.

[30] A. Rahman and B. Verma, "Effect of ensemble classifier composition on offline cursive character recognition," *Information Processing and Management*, vol. 49, no. 4, pp. 852-864, Jul. 2013.

[31] N. Azizi, N. Farah, and M. Sellami, "Ensemble classifier construction for Arabic handwritten recognition," in *Proc. Int. Workshop on Systems, Signal Processing, and their Applications*, pp. 272-274, May 2011.

[32] L. S. Oliveira, M. Morita, and R. Sabourin, "Feature selection for ensembles applied to handwriting recognition," *International J. of Document Analysis and recognition*, vol. 8, no. 4, pp. 262-279, Sep. 2006.

[33] B. E. Rosen, "Ensemble learning using decorrelated neural networks," *Connection Science*, vol. 8, no. 3-4, pp. 373-384, Jul. 1996.
[34] Y. Liu and X. Yao, "Ensemble learning via negative correlation," *Neural Networks*, vol. 12, no. 10, pp. 1399-1404, Dec. 1999.
[35] Y. Liu and X. Yao, "Simultaneous training of negatively correlated neural networks in an ensemble," *IEEE Trans. on Systems, Man, and Cybernetics-Part B: Cybernetics*, vol. 29, no. 6, pp. 716-725, Dec. 1999.
[36] C. L. Liu, K. Nakashima, H. Sako, and H. Fujisawa, "Handwritten digit recognition: investigation of normalization and feature extraction techniques," *Pattern Recognition*, vol. 37, no. 2, pp. 265-279, Feb. 2004.
[37] R. Azmi and E. Kabir, "A new segmentation technique for omnifont Farsi text," *Pattern Recognition Letters*, vol. 22, no. 2, pp. 97-104, Feb. 2001.
[38] Ø. D. Trier, A. K. Jain, and T. Taxt, "Feature extraction methods for character recognition-a survey," *Pattern Recognition*, vol. 29, no. 4, pp. 641-662, Apr. 1995.
[39] R. Polikar, "Ensemble based systems in decision making," *Circuits and Systems Magazine, IEEE*, vol. 6, no. 3, pp. 21-45, Third Quarter 2006.
[۴۰] س. ح. نبوی کریزی و ا. کبیر، "ترکیب طبقه‌بندها: ایجاد گوناگونی و قواعد ترکیب،" *نشریه علمی پژوهشی انجمن کامپیوتر ایران*، جلد ۳، شماره ۳- الف، صص. ۹۵-۱۰۷، پاییز ۱۳۸۴.
[۴۱] م. ع. باقری، غ. ع. منتظر و ا. کبیر، "سیستم‌های دسته‌بند چندگانه روش‌های طراحی و قواعد ترکیب شورا،" *دوفصل‌نامه پردازش علایم و داده‌ها*، جلد ۸، شماره ۲، صص. ۲۹-۵۶، ۱۳۹۰.
[42] L. I. Kuncheva and C. J. Whitaker, "Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy," *Machine Learning*, vol. 51, no. 2, pp. 181-207, May 2003.
[43] D. Ruta and B. Gabrys, "Analysis of the correlation between majority voting error and the diversity measures in multiple classifier systems," in *Proc. of the 4th Int. Symp. on Soft Computing*, 7 pp., Paisley, UK, Jun. 2001.
[۴۴] م. صنیعی آباده، س. محمودی و م. طاهرپور، *داده‌کاوی کاربردی، نیاز دانش*، صص. ۱۸۷-۱۸۹، چاپ اول ۱۳۹۱.

سیدعلی‌اصغر عباس‌زاده آرانی تحصیلات خود را در مقاطع کارشناسی و کارشناسی ارشد مهندسی برق- الکترونیک به‌ترتیب در سال‌های ۱۳۸۸ و ۱۳۹۰ در دانشگاه شهید رجایی تهران به پایان رسانده است و هم‌اکنون دانشجوی مقطع دکتری رشته مهندسی برق- الکترونیک در دانشگاه تربیت مدرس تهران می‌باشد. زمینه‌های تحقیقاتی مورد علاقه ایشان عبارتند از: بازشناسی کلمات دستنویس فارسی، شبکه‌های عصبی، ترکیب طبقه‌بندها و بازشناسی الگو.

احسان‌اله کبیر در آبان ۱۳۳۷ در تهران به دنیا آمد. او کارشناسی ارشد پیوسته خود را در مهندسی برق و الکترونیک از دانشکده فنی دانشگاه تهران و دکترای خود را در مهندسی سیستم‌های الکترونیک از دانشگاه اسکس در انگلستان، به‌ترتیب در سال‌های ۱۳۶۴ و ۱۳۶۹ دریافت کرد. او اکنون استاد دانشکده مهندسی برق و کامپیوتر دانشگاه تربیت مدرس است. زمینه پژوهشی مورد علاقه او تحلیل و بازشناسی مستندات است.