ارائه روشی جدید با استفاده از روش‌های ترکیب دسته‌بندی جهت شناسایی حروف دست‌نویس فارسی

مازیار کاظمی، موسسه آموزش عالی روزبهان - دانشکده علوم کامپیوتر و هوش مصنوعی - Maziar.kazemi@rouzbahan.ac.ir
محمد یوسف نژاد، دانشگاه هوا و فضای نانجینگ - دانشکده علوم و تکنولوژی کامپیوتر - myousefnezhad@ustmb.ac.ir
صابر نوریان، دانشگاه علم و صنعت ایران - دانشکده برق - snourian@elec.iust.ac.ir

چکیده: ترکیب دسته بندی، مبتنی بر انتخاب روشی، جهت انتخاب و ادغام نتایج دسته بندی پایه روی یک داده خاص، بر اساس یک معیار توافقی است. تحقیقات اخیر نشان داده است که روش‌های ترکیب دسته‌بندی نتایج بهتری نسبت به دسته‌بندی کننده تنها دارد. هدف از این مقاله، ارائه روش‌های ترکیبی دسته‌بندی برای شناسایی حروف دست‌نویس فارسی است. در این مقاله، ابتدا با استفاده از روش تکنیک آنالیز اجزای اصلی (PCA) ویژگی‌های تصاویر استخراج شده است. سپس از ماشین‌های بردار پشتیبان (SVM) و درخت تصمیم به عنوان دسته بندی کننده پایه استفاده و روش‌های ترکیبی بر روی این دسته بندها اعمال‌شده است. روش پیشنهادی با روش های شبکه های عصبی، k نزدیکترین همسایه و دسته بندهای پایه مقایسه شده است و نتایج تجربی این روش ها روی چندین مجموعه داده نشان می‌دهد که روش های ترکیبی نتایج بهتری نسبت به سایر روش ها دارد.

کلمات کلیدی: شناسایی حروف دست‌نویس فارسی، ترکیب دسته‌بندی کننده‌ها، درصد صحت، ماشین بردار پشتیبان

۱- مقدمه

امروزه بازشناسی الگو[1] در بسیاری از زمینه‌ها نقش کاربردی دارد. یکی از شاخه های بازشناسی الگو، بازخوانی نوری نویسه ها (OCR) است که به طور گسترده در ادبیات یادگیری ماشین مورد استفاده قرار می گیرد[1]. شناسایی حروف دستنویس نیز، زیر مجموعه ای از OCR است[2]. در دو دهه گذشته، فعالیت های وسیعی برای استفاده از کامپیوتر در خواندن متن‌های دست‌نویس انجام‌شده است[3] که بعضی از این روش ها مانند روش تطبیق الگوها[2] نسبت به اندازه و تغییر مکان حساس بوده و نیاز به نرمالسازی دارند بنابراین ما به دنبال الگوریتم هایی هستیم که نیاز به نرمالسازی نداشته باشند[4]. تحقیقات زیادی در زمینه OCR انجام شده است که اساسا برای شناسایی اعداد و حروف انگلیسی بوده و نتایج خوبی در این زمینه به‌دست آمده است اما درمورد زبان های فارسی و عربی به دلیل ویژگی های خاص[5] این زبان ها، الگوریتم ها و نرم افزارهای طراحی شده از دقت کمتر و قدرت تشخیص پایین تری برخوردارند. در زمینه تشخیص حروف فارسی (مخصوصا حروف دستنویس) نیز تحقیقات کمتری انجام شده و کمبودهایی در این زمینه وجود دارد. هدف از این مقاله، ارائه روشی جدید برای شناسایی حروف دست‌نویس فارسی است که از ترکیب دسته‌بندی کننده‌ها استفاده می‌کند. این روش ها در بیشتر حالت‌ها نتایج بهتری نسبت به دسته‌بندی کننده تنها دارند[6].

در ادامه مقاله ابتدا در بخش دوم به بررسی پیش زمینه های مورد نیاز پرداخته شده و در بخش سوم کارهایی که در این زمینه انجام شده بیان می شود. در بخش چهارم روش‌های دسته‌بندی پایه مورد بررسی قرار می گیرد. سپس در بخش پنجم مدل پیشنهادی ارائه می شود و در بخش ششم به ارزیابی و بررسی فواید و مشکلات مدل پیشنهاد شده پرداخته می شود و در نهایت در بخش هفتم نتایج حاصل از این مقاله و خط مشی کارهای آینده بیان می شود.

[1] Pattern Recognition
[2] template matching

۲- پیش زمینه

یک سیستم OCR از سه مرحله تشکیل شده است[7]:

I-پیش پردازش[3] : به مجموعه اعمالی که روی تصویر صورت می گیرند تا موجب تسهیل در روند اجرای فازهای بعدی گردد پیش پردازش می گویند. از اهداف این مرحله می توان به کاهش نویز، نرمالیزه کردن داده ها و فشرده‌سازی نام برد[8]. مهمترین دلیل برای کاهش نویز کم کردن خطا در مرحله بازشناسی است. همچنین سبب کاهش اندازه تصویر می شود که به نوبه خود، کاهش زمان مورد نیاز برای پردازش ها و ذخیره سازی های آینده را در پی خواهد داشت. دو روش متعارف فشرده سازی، یکی دوگانی (دوسطحی) کردن تصویر و دیگری نازک سازی می باشد[8].

II-استخراج ویژگی[4] : در این مرحله به هر الگو یک کد یا بردار ویژگی نسبت داده می شود که معرف آن الگو در فضای ویژگی می باشد و آن را از دیگر الگوها متمایز می کند. انتخاب روش استخراج ویژگی، مهمترین عامل در بازشناسی الگو است[9].

III-بازشناسی[5] : در این مرحله با استفاده از الگوریتم های بازشناسی، کاراکترها و علائم نوشتاری شناسایی می شوند[7].

۳- کارهای انجام شده

در این بخش به مروری بر کارهای انجام شده در زمینه تشخیص حروف دستنویس فارسی و عربی می پردازیم.

مسروری [10] روشی برای بازشناسی حروف مجزای دستنویس فارسی ارائه نموده است. وی ۳۳ کلاس برای حروف فارسی در نظر گرفته است. بازشناسی حروف در دو مرحله انجام می گیرد. در مرحله اول به کمک یک طبقه بندی کننده فازی « قاعده ـ پایه » که قوانین آن به وسیله مجموعه نمونه های آموزشی یادگرفته می شوند، حروف به ۸ گروه تقسیم می گردند. ویژگی های مورد استفاده در این مرحله با روش مکان های مشخصه در راستاهای ۹۰ و ۴۵ درجه از تصویر باینری حرف بدست می‌آید. در مرحله دوم، به وسیله طبقه بندی کننده های درخت تصمیم که برای هر گروه به طور خاص طراحی شده اند، بازشناسی نهایی حروف در هر گروه انجام می گیرد. در این مرحله، ویژگی های به کار رفته عموماً ساختاری هستند .

صالحیان و همکاران در [11] یک سیستم کامل برای شناسایی کلمات فارسی نستعلیق با استفاده از شبکه های عصبی ارائه کرده اند. آن ها در مرحله پیش پردازش، پس از یافتن بخش های متصل، سرکش ها و زیرکش های حروف را کشف و از تصویر حذف کردند و با استفاده از یک الگوریتم تقطیع که بر اساس کانتور بالایی و پایینی کلمه عمل می کند، تصویر کلمه را به دنباله ای از زیر کلمه ها تقسیم کردند. پس از انجام عمل تقطیع، هشت ویژگی شامل سه توصیفگر فوریه و تعدادی ویژگی ساختاری برای نمایش زیرکلمه ها در فضای ویژگی بکار برده شد. شناسایی با استفاده از یک شبکه عصبی پرسپترون چندلایه انجام شده است.

مولایی و همکاران [12] روشی به منظور بازشناسی کد های پستی و اسامی شهر های ایران که بر روی بسته های پستی به صورت دست نویس نوشته می شوند، پیشنهاد نموده اند. بردارهای ویژگی این روش، از طریق بکارگیری تبدیل موجک گسسته (DWT) با موجک اساسی هار محاسبه می گردند. در مرحله استخراج ویژگی ها، تبدیل موجک سه سطحی برروی تصویر نازک شده اعمال می شود. برای آموزش سیستم، یک شبکه عصبی MLP با قانون پس انتشار خطا مورد استفاده قرار می گیرد.

کلاکسین و همکاران [13] الگوریتمی برای بازشناسی کلمات عربی بدون تقطیع طراحی نموده اند که در آن ابتدا هر کلمه به یک تصویر قطبی نرمالیزه تبدیل می شوند و آنها سپس یک تبدیل فوریه دو بعدی به تصویر قطبی اعمال می گردد. مرحله بازشناسی با بکارگیری تابع فاصله اقلیدسی نرمالیزه شده انجام می گیرد. جهت آزمایش این الگوریتم، بیش از ۱۷۰۰ کلمه در ۴ فونت مختلف مورد استفاده قرار می گیرد.

---

[3] Preproccesing
[4] Feature Extraction
[5] Recognition

۴- دسته بندهای پایه

انتخاب نوع مدل دسته بندی، یکی از مراحل بازشناسی الگو است. دسته بند به عنوان هسته اصلی یک سامانه بازشناسی الگو مطرح است[14]. دسته بند، هر الگوی ناشناخته را بر اساس ویژگی های آن، به یکی از کلاس های شناخته شده نسبت می دهد. دسته بندهای متعددی برای بازشناسی الگو پیشنهاد شده اند که می توان به دسته بند بیز، شبکه های عصبی، ماشین بردار پشتیبان[6] و دسته بند کمترین فاصله اشاره کرد. پس از انتخاب مدل، باید پارامترهای آن را مشخص کرد. پارامترها در طول فرآیند یادگیری تعیین می شوند پس از آن که مدل کامل شد، با استفاده از نمونه های آزمایشی، سامانه بازشناسی الگو اعتبارسنجی می شود. ما در این بخش به توضیحاتی در مورد دسته بندهای SVM و درخت تصمیم می پردازیم.

۴-۱ ماشین بردار پشتیبان(SVM)

SVM، بیشترین حاشیه بین دو کلاس را در یک فضای ویژگی خاص پیدا می کند[15]. بنابراین ابرصفحه ای را انتخاب می کند که فاصله ی آن از نزدیک ترین داده ها در هر دو طرف جداکننده ی خطی، ماکزیمم باشد. تابع تصمیم گیری برای جداکردن داده ها، با استفاده از بردارهای پشتیبان که نزدیک ترین داده های آموزشی به ابرصفحه جداکننده هستند، تعیین می شود. در واقع ابرصفحه بهینه در SVM جداکننده ای بین بردارهای پشتیبان است. مرز تصمیم گیری معمولا با استفاده از یک تابع خطی که دو کلاس را از هم جدا می کند به دست می آید اما در بعضی از حالات این مرز به صورت خطی قابل تفکیک نیست، بنابراین داده ها را به فضای ویژگی با ابعاد بالاتر انتقال می دهند که در این فضا داده ها به صورت خطی قابل جداسازی هستند[16]. برای انتقال، ابتدا تابع $\varphi(x)$ را برای نگاشت به فضای دیگر پیدا می کنیم. با استفاده از تابع کرنل $K(x,x_i)=\varphi(x_i)^T\varphi(x)$ ، تابع تصمیم گیری SVM به این صورت نمایش داده می شود:

$$f(x)=\sum_{i=1}^{n}\alpha_i y_i K(x,x_i)+b \qquad (1)$$

در رابطه (1) $f(x)$ خروجی تصمیم، y_i برچسب داده های x_i و x داده ای است که قرار است دسته بندی شود. پارامترهای α_i و b در طول آموزش مقداردهی می شوند. در SVM می خواهیم مسئله زیر را بهینه سازی کنیم:

$$\min \frac{1}{2}||w||^2+C\sum_{i=1}^{n}\epsilon_i \qquad subject\ to\ y_i(w^T\varphi(x_i)+b)\geq 1-\epsilon_i \qquad (2)$$

کرنل های زیادی وجود دارد اما برای مسئله OCR، بهترین کرنل تابع شعاع مبنا[7] (RBF) می باشد:

$$K(x,x_i)=\exp\left(-\gamma||x-x_i||^2\right),\gamma>1 \qquad (3)$$

در رابطه (2) C پارامتر جریمه و در رابطه (3) γ پارامتر کرنل می باشد.

اگرچه SVM یک دسته بند دو کلاسه است، SVM چند کلاسه از ترکیب دسته بندهای دو کلاسه حاصل می شود[17].

SVM کاربردهای زیادی دارد اما شهرت آن به خاطر موفقیتش در تشخیص حروف دست نویس است که با شبکه های عصبی پیچیده قابل قیاس می باشد[18].

۴-۲ درخت تصمیم

استفاده از درخت های تصمیم به عنوان دسته بند در حال حاضر کاملا کاربردی شده است. درخت تصمیم بر اساس آنالیز داده های ورودی و به منظور پیدا کردن یک ویژگی به عنوان مبنای تصمیم گیری برای هر نود به کار می رود. در هر نود ویژگی های مختلف داده ها بررسی شده و یک ویژگی که در صورت انتخاب باعث کاهش بی نظمی[8] (آنتروپی) می شود، گزینش می شود[19].

---
[6] Support Vector Machine
[7] Radial basis function
[8] Entropy

درخت تصمیم گیری به عنوان یک دسته بند چندمنظوره در یادگیری ماشین در نظر گرفته می شود و همچنین به عنوان یک دسته بند ناپایدار محسوب می شود. این بدان معنی است که می توانیم با آموزش های متوالی مجموعه داده ها با همان مقدار اولیه، راه حل های مختلف را پوشش دهیم[20].

۵- مدل پیشنهادی

هدف از این مقاله، استفاده از روش های ترکیب دسته بندی ها به منظور شناسایی حروف دست نویس فارسی و مقایسه آن با روش های معمول دسته بندی است. ما از چهار روش ترکیبی Bagging ، Boosting، Random Forest و ECOC[9] در این مسئله استفاده کردیم که در ادامه مختصرا این روش ها را شرح می دهیم.

معمولا مدل های مختلف، پیشبینی های متفاوت با نتیجه های متفاوت ایجاد می کنند. بنابراین می توانیم برای دستیابی به نتایج بهتر، تعداد مدل ها را با استفاده از زیرمجموعه داده های متفاوت توسعه دهیم و یا از شرایط مختلف در روش مدل سازی انتخابی استفاده کنیم[21].

اگر تعدادی دسته‌بند پایه داشته باشیم می‌توان با ترکیب نتایج آن‌ها به‌دقت بالاتری رسید. ایده روش های ترکیبی به این صورت است که مجموعه‌ای از دسته‌بندی کننده‌ها را با داده‌های آموزشی ایجاد کرده و میزان صحت را با انجام عملیات رای گیری بر روی نتایج آن ها بدست می آوریم. ایده کلی در شکل زیر نشان داده شده است:

نحوه محاسبه خطای این روش با فرض داشتن N دسته‌بند که خطای هر یک P است از رابطه زیر محاسبه می‌شود:

$$P(error) = \sum_{k=\frac{N}{2}+1}^{N} \binom{N}{k} P^k (1-p)^{N-k} \quad (4)$$

در واقع مفهوم ترکیب این است که خروجی چندین مدل با هم مخلوط شده تا به تصمیم گیری بهتر برسند[21].

Bagging ۵-۱

این روش با دستکاری مجموعه داده های آموزشی ایجاد می گردد. در این روش مجموعه داده های آموزشی به صورت مضاعف ساخته می شوند و بدین صورت عمل می کند که دسته بند پایه با داده های آموزشی متفاوتی آموزش می بیند تا در نتایج اندکی تفاوت ایجاد گردد. داده‌ها در این مرحله با روش نمونه‌برداری تصادفی با جایگزینی انتخاب می‌شوند در نتیجه با این که دسته بندها یکسان هستند ولی به دلیل انتخاب تصادفی نمونه ها، نتایج متفاوت خواهد بود. در نهایت با انجام عملیات رای گیری روی خروجی های بدست آمده از دسته بندها، نتیجه نهایی حاصل می گردد. این روش برای الگوریتم های ناپایدار، که تغییری کوچک در ورودی، دچار تغییر در نتیجه آن ها می شود مناسب است. روش های متفاوتی برای عملیات رای گیری وجود دارد که می توان به میانگین گیری، ماکزیمم و مینیمم گیری، استفاده از میانه و ... اشاره کرد. ما در اینجا از روش ماکزیمم گیری استفاده می کنیم که به صورت زیر می باشد:

$$\sum_{j=1}^{L} d_{j,k} = max_{m=1}^{c} \sum_{j=1}^{L} d_{j.m} \quad (5)$$

در رابطه (5) L تعداد دسته بندها، c تعداد کلاس ها و d خروجی دسته بند می باشد[22].

Boosting ۵-۲

این روش نیز با دستکاری مجموعه داده های آموزشی ایجاد می گردد. هدف این است که در یک دسته بند ضعیف تغییراتی ایجاد کنیم تا به یک دسته بند قوی تبدیل شود. سعی می کنیم داده های آموزشی را در هر مرحله طوری تغییر دهیم که دسته بندهای ایجاد شده مکمل یکدیگر باشند و نقاط ضعف را پوشش دهند. بدین منظور به هر نمونه وزنی اختصاص می دهیم که در ابتدا این وزن به صورت یکنواخت خواهد بود. پس از آموزش مدل، نمونه های آموزشی را آزمایش می کنیم. داده هایی که درست دسته

[9] Error Correcting Output Coding

بندی شده اند در مرحله بعد با احتمال کمتری ظاهر می شوند و داده هایی که اشتباه دسته بندی شده اند با احتمال بیشتری در مرحله بعد ظاهر می شوند که این احتمالات با استفاده از وزن داده ها محاسبه می شود. Boosting تمایل دارد تا به درجه صحت بالا دست پیدا کند اما این کار ریسک بیش برازش[10] را در پی دارد.

۵-۳ Random Forest

این روش با دستکاری ویژگی های ورودی ایجاد می گردد. در این روش برای هر مجموعه داده، یک زیرمجموعه از ویژگی ها انتخاب می شوند که این زیرمجموعه ها می توانند به صورت تصادفی و یا برحسب یک حوزه تعریف انتخاب شوند. این روش برای ترکیب درخت های تصمیم طراحی شده است که پیشبینی های چندین درخت تصمیم را با هم ترکیب می کند به طوری که هر درخت با یک بردار انتخابی تولید شده است. هر بردار انتخابی با توزیع احتمال یکنواخت ایجاد می گردد. این امر برخلاف رویکرد مورد استفاده در Adaboost است که دارای احتمالات متفاوت بود. اگر در روش Bagging از درخت تصمیم به عنوان دسته بند مبنا استفاده کنیم یک Random Forest خواهیم داشت.

۵-۴ ECOC

این روش با دستکاری برچسب کلاس ها ایجاد می گردد و زمانی که تعداد کلاس ها خیلی زیاد باشد از آن استفاده می شود. این روش ابتدا با اختصاص کد به هر یک از کلاس ها آغاز می شود. این کد می تواند به صورت تصادفی و یا بر اساس یک حوزه معین تعریف شده انتخاب شود. سپس با استفاده از آن کد به عنوان برچسب هر کلاس، آموزش را آغاز می کنیم. چون کدهای ما به صورت دودویی هستند پس ما برای هر بیت این کد، یک دسته بند دودویی در نظر می گیریم که آن را آموزش می دهیم. مثلا اگر کد ما برای هر کلاس ۱۲ بیتی باشد، باید ۱۲ دسته بند دودویی را آموزش دهیم. پس از آموزش، برای هر نمونه آزمایشی، کدی تولید می گردد که با استفاده از این کد و کدی که به هر کلاس اختصاص داده بودیم، فاصله همینگ را محاسبه می کنیم. کمترین فاصله محاسبه شده متعلق به هر کلاسی بود آن را به عنوان خروجی در نظر می گیریم.

۶-ارزیابی

در این بخش نتایج به کارگیری روش پیشنهادی روی مجموعه داده‌های مختلف و پارامترهای مورد استفاده گزارش می شود. همانطور که قبلا ذکر گردید هر سیستم بازشناسی الگو از سه مرحله تشکیل شده است. در مرحله پیش پردازش عملیاتی بر روی تصاویر انجام شد تا کیفیت آن ها افزایش یابد. در مرحله استخراج ویژگی، از پنج روش مختلف که در مقالات برای بازشناسی حروف و اعداد دست نویس استفاده شده است بهره بردیم. این روش ها عبارتند از:آنالیز اجزای اصلی(PCA) [23]، گرادیان تصویر[24]، فیلترهای گاوسی[25] ، ضرایب DCT [26] و گشتاورها و ماتریس GLCM [7]. پس از آموزش داده ها با استفاده از روش های استخراج ویژگی مذکور و مقایسه نتایج حاصل، به این نتیجه رسیدیم که PCA بهترین روش برای این منظور می باشد و ما از این روش در ادامه استفاده کردیم.

۶-۱ مجموعه داده ها

روش پیشنهادی بر روی ۱۰ مجموعه داده مورد آزمایش قرار گرفته است. هر مجموعه شامل ۳۲۰۰ نمونه از حروف دستنویس فارسی می باشد که هر مجموعه دربردارنده ۱۰۰ نمونه برای هر حرف فارسی است. از هر مجموعه ۲۲۴۰ نمونه مربوط به فاز آموزش (۷۰ نمونه از هر حرف) و ۹۶۰ نمونه مربوط به فاز آزمایش (۳۰ نمونه برای هر حرف) است. هر تصویر شامل یک پس زمینه سفید می باشد که حروف در وسط آن قرار دارد. ابتدا برای افزایش کیفیت تصاویر را به دودویی تبدیل کردیم و به منظور کاهش حجم محاسبات، حروف را از این پس زمینه جدا کرده و سپس آن را به یک اندازه استاندارد تبدیل می کنیم.

[10] Overfitting

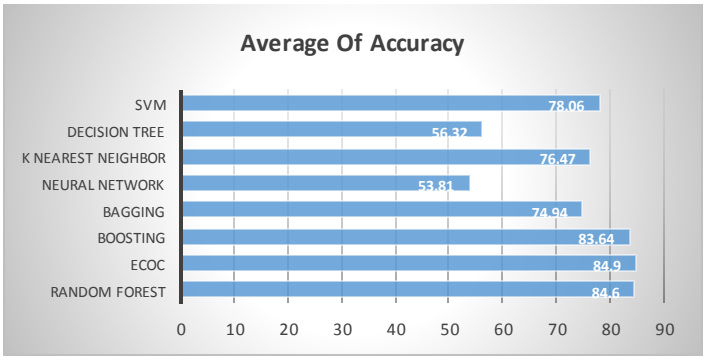
۲-۶ نتایج آزمایش

روش پیشنهادی در محیط MATLAB (ver 8.1) پیاده سازی و مورد آزمایش قرار گرفته و نتایج آزمایش روی میانگین ۱۰ بار اجرای مستقل برنامه برای هر تکنیک گزارش شده است. نتایج حاصل از این آزمایش در جدول (۱) آمده است. در این مقاله ما از چهار روش برای ترکیب استفاده کردیم. در روش Random Forest(RF) از درخت تصمیم (DT) به عنوان دسته بند پایه و در روش های Bagging، Boosting و ECOC از SVM به عنوان دسته بند پایه استفاده شده است.

جدول 1: مقایسه روش های مختلف دسته بندی

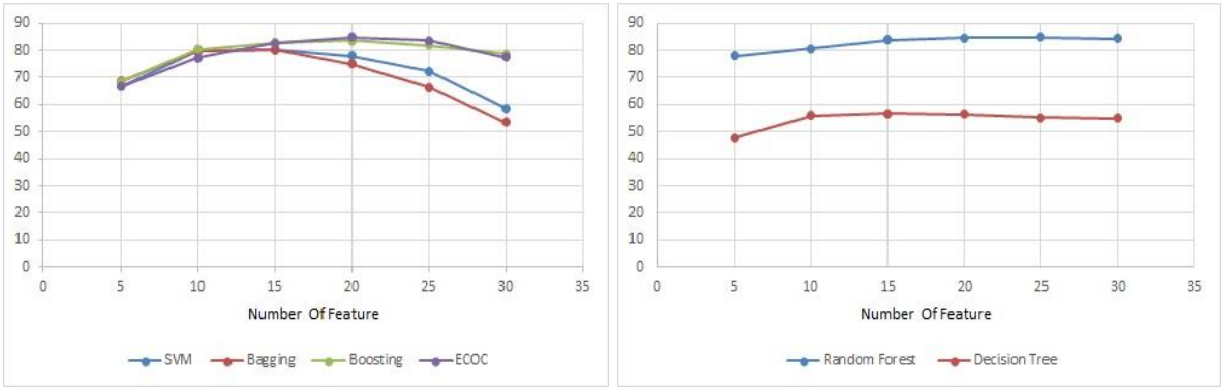
	Primary methods				***Classification Ensemble methods***			
	SVM	**DT**	**KNN**	**NN**	**Bagging**	**Boosting**	**ECOC**	**RF**
Random 1	77.01	52.27	72.92	54.06	72.71	81.57	83.3	**84.02**
Random 2	78.42	56.94	78.65	55.94	75.68	83.69	84.94	**85.39**
Random 3	79	59.11	76.15	53.13	76.55	82.42	**87.57**	86.03
Random 4	79.38	60.12	79.17	60.63	79.47	86.74	**87.56**	83.98
Random 5	76.51	56.34	71.77	54.38	69.89	81.87	84.3	**84.57**
Random 6	77.26	52.87	77.19	51.88	75.68	**83.4**	83.21	81.66
Random 7	79.31	55.67	78.54	51.25	78.92	85.73	**86.61**	84.78
Random 8	79.06	56.50	76.56	51.25	75.72	82.75	84.05	**84.83**
Random 9	79.14	58.67	77.29	59.38	73.22	84.78	84.73	**86.67**
Random 10	75.56	54.74	76.46	46.25	71.59	83.42	82.74	**84.05**

در جدول (۱) بهترین نتیجه ها به صورت پررنگ نشان داده شده است. با توجه به جدول (۱)، RF نتایج حاصل از درخت تصمیم را به صورت قابل توجهی بهبود بخشیده است. با توجه به نتایج حاصل، درخت تصمیم برای این مسئله دسته بند ضعیفی محسوب می شود و ما با استفاده از ترکیب آن ها، به نتایج بهتری رسیدیم. روش های Boosting و ECOC نتیجه SVM را بهبود بخشیده اند اما روش Bagging عملکرد ضعیفتری نشان داده است. دلیل این امر این است که روش Bagging برای دسته بندهای ناپایدار با اغتشاش کوچک عملکرد خوبی دارد. به دسته بندهایی که با تغییری کوچک در ورودی، خروجی آن ها نیز دچار تغییر می شود ناپایدار می گویند و چون SVM پایدار است در نتیجه این روش در خروجی تغییرات زیادی ایجاد نمی کند. برای ارزیابی دقیق تر کارایی این روش ها، از دسته بندهای دیگری برای مقایسه استفاده شده است. همانطور که در جدول (۱) آمده است، روش های ترکیبی نتایج بهتری نسبت به دسته بندهای شبکه های عصبی(NN)، K نزدیکترین همسایه(KNN) و دسته بندهای پایه SVM و DT دارد. در شکل ۲ میانگین درصد صحت هر تکنیک آورده شده است:



شکل 2 : میانگین درصد صحت برای هر تکنیک

همانطور که در شکل ۲ مشخص است، سه روش ترکیبی Boosting، RF و ECOC بهترین میانگین را دارا می باشند. همانطور که گفته شد این روش برای دسته بندهای ضعیف بسیار خوب عمل می کند ولی هرچقدر دسته بندها قوی تر باشند تاثیر این روش کمتر خواهد شد. در مورد دسته بندهای خیلی قوی، نمی توان به نتیجه واحدی رسید. با توجه به زمان اجرایی روش های ترکیبی، می توان نتیجه گرفت که این روش ها، برای دسته بندهای قوی مناسب نیست اما این نتیجه گیری قطعی نیست و بنا به این که از چه دسته بندی به عنوان دسته بند پایه استفاده کنیم، ممکن است این روش ها مناسب باشند. همانطور که گفته شد در مرحله استخراج ویژگی از روش PCA برای ایجاد بردار ویژگی استفاده شده است. PCA یک روش آماریست که برای هر نمونه، تعدادی از ضرایب با بزرگترین مقدار ویژه انتخاب می شود. نتایجی که تاکنون ارائه شد حاصل از ۲۰ ضریب PCA می باشد. در شکل ۳ میانگین نتایج حاصل از تکنیک های مختلف با تعداد ویژگی های متفاوت نشان داده شده است:



(a) **(b)**

شکل 3 : میانگین نتایج حاصل از هر تکنیک با تعداد ویژگی های متفاوت

با توجه به شکل (a) روش های ECOC و Boosting نتایج خوب و مشابهی ارائه می دهند. علاوه بر اینکه این روش ها نتایج حاصل از SVM را بهبود می بخشند بلکه می توانند هزینه حاصل از انتخاب ویژگی را کاهش دهند. همانطور که در شکل (a) پیداست روش های ECOC و Boosting از تعداد ۱۵ تا ۳۰ ویژگی نتایج مشابهی ایجاد می کنند در حالیکه در SVM شاهد نتایج متفاوتی هستیم. در شکل (b) روش RF به طور چشمگیری نتایج حاصل از درخت تصمیم را بهبود می بخشد این امر نشان دهنده ی قدرت روش های ترکیبی برای تبدیل دسته بندهای ضعیف به دسته بندهای قوی می باشد.

۷- نتیجه گیری

در این مقاله از روش های ترکیبی دسته بندی ها به منظور شناسایی حروف دست نویس فارسی استفاده شده است. از آنجایی که هیچ دسته بندی وجود ندارد که برای تمام مسائل جواب بهینه را تولید کند پس ما می توانیم با ترکیب نتایج آن ها به جواب مطلوب برسیم. روش های مختلفی برای ترکیب وجود دارد که ما از ۴ روش Random Forest(RF)، Bagging، Boosting و ECOC در این مقاله استفاده کردیم. نتایج تجربی روش های ترکیبی برای مسئله بازشناسی حروف دست نویس بر روی ۱۰ مجموعه داده مختلف و مقایسه آن با روش هایی مثل شبکه های عصبی، K نزدیکترین همسایه و دسته بندهای پایه نشان می دهد که این روش نسبت به روش های متداول دسته بندی برتری دارد. نتایج نشان می دهد که روش های ترکیبی برای دسته بندهای ضعیف تر بسیار مناسب می باشد و خروجی را به صورت قابل توجهی بهبود می بخشد. از بین این ۴ روش، می توان RF را بهتر از بقیه دانست زیرا علاوه بر اینکه نتایج بهتری تولید می کند، از نظر زمان اجرایی نیز بهتر از سایر روش هاست. ضمن اینکه این روش برای تعداد ویژگی های متفاوت نتایج تقریبا مشابه و ثابتی ارائه می دهد. برای کارهای آینده پیشنهاد می شود به جای تغییر در داده های یک دسته بند، داده ها را ثابت در نظر گرفته و دسته بندها را تغییر دهیم که این تغییر می تواند در پارامترهای یک دسته بند خاص باشد و یا از دسته بندهای متفاوت استفاده کنیم.

مراجع

[1] Vivt Pham, Dung., "Online handwriting recognition using multi convolution neural networks", The Ninth International Conference on Simulated Evolution And Learning (SEAL), 2012.
[2] Ghods V., Kabir E., "A Study on Farsi Handwriting Styles for Online Recognition", Malaysian Journal of Computer Science. Vol. 26(1), 2013.
[3] Kumar Garg N., Kaur L., Jindal M. K., "The Hazards in Segmentation of Handwritten Hindi Text", International Journal of Computer Applications(IJCA), 2011.
[4] Al-Jawfi R., "Handwriting Arabic Character Recognition LetNet Using Neural Network", The International Arab Journal of Information Technology, Vol. 6, No. 3, July 2009.
[5] Addakiri K., Bahaj M., "On-line Handwritten Arabic Character Recognition using Artificial Neural Network", International Journal of Computer Applications(IJCA), Volume 55 - Number 13,2012.
[6] Hassanzadeh M., Ardeshir G., "A New Classifiers Ensemble Method For Handwritten Pen Digits Classification", International Research Journal of Applied and Basic Sciences ,Vol 5,pp.1092-1096, 2013.
[7] Hashemi O.,Seyedein A., "Recognizing offline handwritten Farsi words with the emphasis on detecting a city name", 7th Conference on Machine Vision and Image Processing,Tehran,2011.
[8] Arica Nazif ,Yamin-Vural Fatos T., "An overview of character recognition based focused on off-line handwriting", IEEE Transactions on Systems, Vol. 31, No. 2, May 2001.
[9] Singh P., BudhirajaI S., "Offline Handwritten Gurmukhi Numeral Recognition using Wavelet Transforms", International Journal of Modern Education and Computer Science (IJMECS), Vol.4, No.8, August 2012.
[10] Masroori K., "The Persian handwritten letters recognition with fuzzy method" ,Ph.D. Thesis, University Of Tarbiat Modarres, 2000.
[11] Salehian N., Yazdchi M., Karimian A. R., "Identification of Persian words nstaliq using artificial neural network", Scientific research of electrical engineering Majlesi, 2008.
[12] Mowlaei A., Faez K., Haghighat A.T., "Feature extraction with wavelet transformfor recognition of isolated handwritten Farsi/Arabic characters and numerals", 14th International Conference on Digital Signal Processing, Vol. 2, pp. 923-926, July 2002.
[13] Khorsheed M.S., Clocksin W.F., "Multi-font Arabic word recognition using spectralfeatures", 15th International Conference on Pattern Recognition, vol. 6, pp. 3574 - 3577, June 2000.
[14] L.I Kuncheva, "combining Pattern Classifier: Methods and algorithm", Published by John Wiley&Sons. Inc.,2004.
[15] A. Ben-Hur and J. Weston. "A users guide to support vector machines. Technical report", Department of Computer Science. Colorado State University, 2010.
[16] B. Keshari and S.M. Watt. "Hybrid mathematical symbol recognition using support vector machines", In Proceedings of the Ninth International Conference on Document Analysis and Recognition, 2007.
[17] I.-S. Oh and C. Suen, "A class-modular feed-forward neural network for handwriting recognition", Pattern Recognition, 35:229–244, 2002.
[18] Cheng L. Liu, K. Nakashima, H. Sako, H. Fujisawa, "Handwritten digit recognition: benchmarking of state-of-the-art techniques", Pattern Recognition, Vol. 36, pp. 2271-2285, 2003.
[19] Khosravi H., Yaghmaei M., "The introduction of a new structure of decision tree and use it with Clustering (FCM) to generate fuzzy rules", 15th International Conference of Computer Society of Iran, Tehran, 2009.
[20] Yang T., "Computational Verb Decision Trees", International Journal of Computational Cognition, 34-46, 2006.
[21] Alizadeh H., Yousefnezhad M., Minaei-Bidgoli B., "Wisdom of Crowds Cluster Ensemble", Intelligent Data Analysis, IOS Press, Vol. 19(3), 2015.
[22] Stepenosky N., Green D., Kounios J., and Christopher M. Clarkand Robi Polikar, "Majority vote and decision template based ensemble classifiers trained on event related potentials for early diagnosis of alzheimer's disease", In ICASSP, 2006.
[23] Nahvi M., Rafiei M., Ebrahimpoor R., Kabir E., "The combination of two-class classifiers for the recognition of handwritten Persian digits",16th Iranian Conference on Electrical Engineering, , 2008.
[24] Khosravi H., Kabir E., "Recognition of handwritten Persian digits introduce two fast and efficient features", Fourth Conference on Machine Vision and Image Processing, Ferdosi University, 2006.
[25] Soltanzade H., Rahmati M., "Persian handwritten digits recognition using gradient and support vector machine classifier", 9th Annual Conference of Computer Society of Iran,Sharif University, 2003.
[26] C-C. Liu, K-W. Chuang, "An outdoor time simulation schame based on support vector regression with radial basis function on DCT domain", Image and Vision Computing, 2009.