

بازشناسی قلم‌های فارسی بر اساس ترکیب ویژگی‌ها و خبره‌ها

وحید زارعی^۱، محمد اسماعیل دوست^۲، احسان مرتضوی^۳

۱مریی دانشکده مهندسی دریا، دانشگاه علوم و فنون دریایی خرمشهر V.zarei@kmsu.ac.ir

۲استادیار دانشکده مهندسی دریا، دانشگاه علوم و فنون دریایی خرمشهر

۳کارشناسی ارشد برق الکترونیک، دانشگاه شاهرود

تاریخ دریافت: ۹۴/۷/۴ تاریخ پذیرش: ۹۵/۱/۲۱

چکیده

در این مقاله، یک روش جدید بر اساس ترکیب ویژگی‌ها و طبقه‌بندها برای بازشناسی قلم‌های فارسی معرفی شده است. ویژگی‌ها از تصاویر متون مختلف که به صورت بافت‌هایی با ابعاد یکسان 128×128 پیکسل در آمده‌اند و با استفاده از روش‌های مختلفی همچون فیلتر گابور، نقاب‌های سوبل/روبرت و تبدیل‌های موجک دابیشز، هار و فیلترهای پایین‌گذر و بالا‌گذر سیلت استخراج شده‌اند. ویژگی‌های استخراج شده از این روش‌ها به صورت ۲ به ۲ با یکدیگر ترکیب گردیده‌اند. برای ارزیابی نرخ بازشناسی به کمک ویژگی‌های استخراج شده از سه شبکه عصبی MLP جهت طبقه‌بندی استفاده شده و نتایج آنها با هم ترکیب شده است. هر کدام از شبکه‌های عصبی با استفاده از مجموعه ویژگی‌های حاصل از ترکیب دو روش آموزش داده شده است. در نهایت با استفاده از الگوریتم جمعیت پرندگان، وزن‌های بهینه را به منظور کمینه شدن خطا تعیین شده‌اند. روش پیشنهادی با استفاده از پایگاه داده Hoda شامل ۲۱۰۰۰ نمونه از ۱۰ قلم فارسی مختلف، ارزیابی گردیده است. با استفاده از وزن‌های تعیین شده، ترکیب شبکه‌های عصبی صورت گرفته است. با استفاده از روش ارائه شده نرخ بازشناسی ۹۷،۴۵ درصد بدست آمده است که نسبت به دیگر روش‌های ارائه شده بهبود چشمگیری در نرخ بازشناسی نتیجه شده است.

کلیدواژه

بازشناسی قلم فارسی، تبدیل موجک، گرادیان، اختلاط خبره‌ها

مقدمه

صورت که ابتدا نوع قلم را بازشناسی کرده و در انتها سندها را تشخیص می‌دهد.

تا کنون عمل بازشناسی روی ویژگی‌های متفاوتی از قلم‌ها انجام شده است: ویژگی‌های چاپی از جمله انحاء حروف و فضای بین حروف [۲، ۳] و ویژگی‌های ساختاری [۴، ۵].

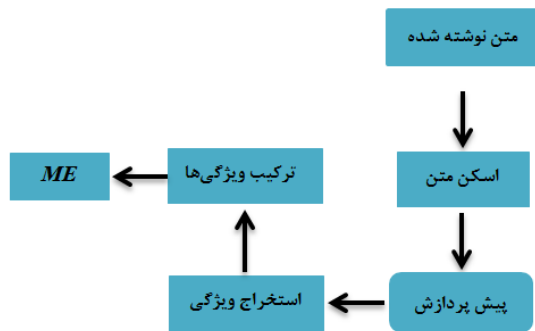
بر اساس گزارشات منتشر شده بازشناسی قلم به دو روش انجام می‌شود. شناسایی در سطح کل صفحه و یا یک پاراگراف [۶] و یا شناسایی در سطح یک خط یا لغت [۵، ۷].

در زمینه بازشناسی قلم در سال‌های اخیر در زبان‌های انگلیسی و چینی تحقیقات مفیدی صورت گرفته است. در [۳] یک روش آماری برای بازشناسی قلم بر اساس ویژگی‌های چاپی محلی ارائه شده است. در [۵] نویسنده از تصاویری که به صورت بافت در نظر گرفته شده‌اند برای ۶ قلم چینی و ۸ قلم انگلیسی استفاده کرده و با استفاده از فیلتر گابور^۳ ۱۶ کاناله ویژگی‌ها استخراج شده است.

سامانه OCR^۱ یکی از شاخه‌های شناسایی الگو می‌باشد. سامانه OCR متون موجود در تصاویر سندها را به صورت خودکار بازشناسی کرده و آنها را به متون قابل جستجو و قابل ویرایش توسط رایانه تبدیل می‌کند. امروزه از سامانه OCR برای بازشناسی مستندات چاپی استفاده می‌شود. پیچیدگی‌های این سامانه برای زبان فارسی و عربی که حروف یک کلمه به یکدیگر چسبیده است نسبت به زبان لاتین که حروف آن به صورت مجزا نوشته می‌شوند، بیشتر می‌باشد. بازشناسی قلم یکی از مسائل مهم در زمینه تحلیل و بازشناسی پرونده‌ها در سامانه OCR می‌باشد و می‌تواند به بهبود نرخ بازشناسی در این سامانه کمک کند. عملی که بازشناسی قلم را به صورت خودکار انجام می‌دهد ADP^۲ نام دارد [۱]، به این

1 Optical Character Recognition
2 Automatic Document Processing

3 Gabor filter



شکل ۱. مراحل بازشناسی قلم

استخراج ویژگی

معرفی پایگاه داده

در این مقاله از پایگاه داده HODA استفاده شده است پایگاه داده مورد استفاده ۲۱۰۰۰ نمونه از ۱۰ قلم فارسی می‌باشد. ۲۵٪ از ویژگی‌ها برای مجموعه آزمایش و ۷۵٪ برای آموزش شبکه در نظر گرفته شده است. این مجموعه نمونه شامل بافت‌های آماده‌ای است که پس از انجام پیش‌پردازش‌های لازم روی تصاویر خطوط متن آماده شده‌اند. برای اینکه تاثیر اندازه قلم و درجه تفکیک تصویر روی دقت بازشناسی کم شود، بافتها از نظر پهنای قلم نرمال شده‌اند. در شکل (۲) نمونه‌ای از بلوک‌ها با قلم‌های مورد استفاده نشان داده شده است.



شکل ۲. نمونه‌ای از بلوک‌ها با قلم‌های مورد استفاده

همانطور که مشاهده می‌شود، قلم‌های فارسی شباهت بسیاری به هم دارند و به همین دلیل، تشخیص قلم‌های فارسی از یکدیگر بسیار دشوار است، به طوری که حتی با چشم هم به سختی می‌توان آنها را از یکدیگر تشخیص داد. استخراج ویژگی از تصاویر متنی پس از پیش‌پردازش صورت می‌گیرد. در این مقاله تصویر به صورت بافت در نظر گرفته شده است، و بعد از نرمالسازی و پیش‌پردازش صورت گرفته بر روی آنها که در این پایگاه داده انجام شده است توسط چندین روش که در ادامه توضیح داده می‌شود ویژگی‌ها استخراج شده‌اند.

همچنین محققان از روش‌های آماری مرتبه بالا نیز برای استخراج ویژگی استفاده کرده‌اند [۸]. عیب این گونه تحقیقات بدلیل استفاده از تصاویر ورودی با درجه تفکیک ۳۰۰ dpi یا ۴۰۰ dpi می‌باشد. این میزان درجه تفکیک برای بازشناسی قلم زیاد است. در سال‌های اخیر در زمینه بازشناسی قلم‌های فارسی تحقیقات خوبی صورت گرفته است. چند نمونه از تحقیق‌های صورت گرفته در این زمینه عبارتند از: برجی و حمیدی که برای تشخیص قلم‌های فارسی از فیلتر گابور استفاده کرده‌اند [۹]. راشدی و نظام-آبادی پور از ضرایب همبستگی جهت بازشناسی قلم فارسی استفاده کرده‌اند [۱۰]. همچنین خسروی و کبیر از روش مبتنی بر گرادیان برای استخراج ویژگی استفاده کرده‌اند که سرعت استخراج ویژگی و نرخ بازشناسی با استفاده از این روش نسبت به روش‌های قبلی بسیار بهتر شده است [۱۱].

در تحقیقات صورت گرفته برای بازشناسی قلم‌ها از طبقه‌بندهای متفاوتی مانند MLP^۵ [۱۱]، بیزین [۳]، نزدیکترین همسایگی [۵]، RBF^۶ [۱۲] برای بازشناسی قلم‌ها در زبان‌های مختلف استفاده شده است. در روش پیشنهادی در این تحقیق تصویر اصلی در فرایند پیش‌پردازش با حذف فضاهای سفید اضافی به یک ساختار متنی یک شکل تبدیل شده و با استفاده از تبدیل موجک‌های مادر متفاوت: Symlet, Db2, Haar و روش مبتنی بر گرادیان و فیلتر گابور ویژگی‌ها از بافت‌هایی با اندازه ۱۲۸×۱۲۸ استخراج شده‌اند. قلم‌های مورد استفاده در این مجموعه عبارتند از: زر، هما، تیر، یاقوت، ترافیک، تایم نیورومن، نازنین، میترا، لوتوس، تاها [۱۱]. بعد از استخراج برای طبقه‌بندی ابتدا ویژگی‌های استخراج شده از روش‌های مختلف با هم ترکیب شده است و سه مجموعه ویژگی به دست می‌آید. سپس با به کارگیری الگوریتم تجزیه و تحلیل مؤلفه‌های اصلی (PCA) تعداد ویژگی‌ها کاهش داده شده و در نهایت با استفاده از شبکه مرکب (ME)^۷ پیشنهاد شده که در بخش چهارم معرفی خواهد شد طبقه‌بندی و بازشناسی قلم‌ها انجام گرفته است. شکل (۱) مراحل روش مورد استفاده برای بازشناسی قلم در این مقاله را به صورت بلوک دیگرام نشان داده است.

در ادامه، ساختار مقاله به این ترتیب است: در بخش دوم پایگاه داده مورد استفاده معرفی می‌گردد. سپس استخراج ویژگی‌ها بر اساس روش‌های مورد نظر از بلوک متنی توضیح داده شده است. سپس نحوه ترکیب و کاهش ویژگی‌ها بیان می‌گردد. طبقه‌بندی و اختلاط خبره‌ها در بخش بعد توضیح داده شده است و در بخش پایانی نتیجه‌گیری و مقایسه روش پیشنهاد شده با سایر روش‌های به کار گرفته شده در این زمینه ارائه گردیده است.

- 4 Dots per inch
- 5 Multi layer perceptron
- 6 Radial basis function
- 7 Principal Component Analysis
- 8 Mixture of Experts

استخراج ویژگی با فیلتر گابور

فیلتر گابور یک فیلتر خطی است که با انتخاب صحیح پارامترهای آن می‌توان از آن برای تشخیص ویژگی‌های بافت‌ها استفاده کرد.

فیلتر گابور ترکیبی از یک تابع گوسی و نمایی می‌باشد که به صورت رابطه (۱) تعریف می‌شود:

$$h(x, y, f, \theta) = g(x, y) \cos(2\pi f(x_r + y_r)) \quad (1)$$

$$\begin{aligned} x_r &= x \cos \theta \\ y_r &= y \sin \theta \end{aligned} \quad \text{که در آن}$$

که f و θ به ترتیب فرکانس شعاعی و جهتی هستند که موقعیت کانالها در صفحه فرکانسی به وسیله آنها تعریف می‌شود و $g(x, y)$ تابع گوسین دو بعدی است که به وسیله رابطه (۲) بدست می‌آید [۱۳].

$$g(x, y) = \frac{1}{2\pi\sigma^2} \exp\left[-\frac{1}{2}\left(\frac{x^2 + y^2}{\sigma^2}\right)\right] \quad (2)$$

در این رابطه x و y مختصات پیکسل‌ها و σ انحراف معیار تابع گوسی است. هر فیلتر گابور برای یک باند مشخص از فرکانس شعاعی و جهتی تنظیم می‌شود. در این پژوهش از چهار جهت 90° ، 45° و 135° درجه استفاده گردیده است [۵].

اگر برای هر چهار جهت، ۲ فرکانس شعاعی وجود داشته باشد در مجموع یک فیلتر گابور ۸ کاناله ایجاد می‌شود. مقادیری که برای λ و σ با آزمون و خطا بدست آمده به صورت زیر می‌باشد.

$$\lambda = \frac{1}{f} = [2.7, 4.1]$$

$$\sigma = [1.512, 2.965]$$

فیلتر گابور را با تصویر ورودی کانولوشن می‌کنیم و خروجی که برای هر کدام از کانال‌ها بدست می‌آید را به ۱۶ زیر بلوک مساوی با اندازه 32×32 تقسیم کرده و برای هر کدام از زیر بلوک‌ها میانگین و انحراف معیار را محاسبه می‌کنیم. برای هر کانال، ۳۲ ویژگی استخراج می‌شود که در مجموع طول بردار ویژگی‌ها برای یک تصویر برابر با ۲۵۶ می‌باشد.

روش مبتنی بر گرادیان

در این مقاله روش گرادیان برای یک بافت، با بکارگیری نقاب‌های سوبل و روبرت برای بازشناسی قلم بررسی گردیده و دامنه و فاز هر تصویر ورودی محاسبه شده‌اند. در روش گرادیانی، از اطلاعات دامنه و فاز، برای استخراج ویژگی استفاده می‌شود [۱۱]. روش مبتنی بر گرادیان را می‌توان به وسیله دو نقاب سوبل و روبرت پیاده سازی کرد که هر کدام به صورت جداگانه در زیر توضیح داده شده است.

نقاب سوبل

نقاب سوبل (SF) به صورت افقی و عمودی از اطلاعات ۸ پیکسل که در همسایگی هم قرار دارند، گرادیان را بدست می‌آورد و با استفاده از گرادیان‌های به دست آمده ویژگی‌ها استخراج می‌گردند.

مراحل استخراج ویژگی با این روش به صورت زیر می‌باشد:

۱- بلوک ورودی را با استفاده از یک پنجره 4×4 به ۱۶ زیر بلوک با اندازه 32×32 تقسیم می‌کنیم [۱۱].

۲- گرادیان تصویر در دو جهت با توجه به عملگر مشتق‌گیر سوبل در راستای x و y برای هر کدام از پیکسل‌ها به صورت رابطه (۳) و (۴) محاسبه می‌شود.

$$g(x, y) = [g_x, g_y]^T \quad (3)$$

که

$$g_x(x, y) = f(x+1, y-1) + 2f(x+1, y) + f(x+1, y+1) - f(x-1, y-1) - 2f(x-1, y) - f(x-1, y+1) \quad (4)$$

$$g_y(x, y) = f(x-1, y+1) + 2f(x, y+1) + f(x+1, y+1) - f(x-1, y-1) - 2f(x, y-1) - f(x+1, y-1) \quad (5)$$

که در این رابطه $f(x, y)$ مقدار پیکسلی است که در موقعیت (x, y) قرار گرفته است.

۳- برای هر پیکسل دامنه و فاز گرادیان را با توجه به رابطه (۶) و (۷) محاسبه می‌کنیم.

$$P(x, y) = \tan^{-1}(g_x, g_y) \quad (6)$$

$$A(x, y) = \sqrt{g_x^2 + g_y^2} \quad (7)$$

۴- کل زوایای محاسبه شده را به ۱۶ مقدار، سطح‌بندی می‌کنیم.

۵- برای هر کدام از ۱۶ زاویه سطح‌بندی شده، یک بردار ویژگی ۱۶ تایی از رابطه (۸) بدست می‌آید.

$$f_\theta = \sum_{x_\theta, y_\theta} A(x, y) \quad (8)$$

که در این رابطه f_θ بردار ویژگی متناظر با زاویه θ می‌باشد. همچنین x_θ و y_θ موقعیت کل پیکسل‌هایی است که زاویه آنها با هم برابر و مساوی θ می‌باشد.

نقاب روبرت

مراحل استخراج ویژگی برای نقاب روبرت مشابه نقاب سوبل می‌باشد، با این تفاوت که در مرحله ۲ گرادیان تصویر با توجه به

(۵)، اعمال تبدیل موجک سیملت به تصویری از قلم لوتوس نشان داده شده است که تصاویر خروجی از فیلترهای بالا گذر و پایین گذر نشان دهنده ضرایب جزئیات و تقریب می‌باشند.

ضرایب بکار برده شده برای فیلترهای پایین‌گذر و بالاگذر Symlet عبارتند از: [۱۴]

$$\begin{aligned} Lo_D &= [-0.1294, 0.2241, 0.8365, 0.4830] \\ Hi_D &= [-0.4830, 0.8365, -0.2241, -0.1294] \end{aligned}$$

تعداد مراحل تجزیه در تبدیل موجک خیلی مهم می‌باشد تعداد مراحل تجزیه در این مقاله ۳ بار در نظر گرفته شده‌است. تصاویر قلم‌ها دارای اندازه ۱۲۸×۱۲۸ می‌باشند و با گرفتن هر بار موجک، اندازه ماتریس ضرایب نصف اندازه تصویر اصلی می‌شود. روش استخراج ویژگی به این صورت می‌باشد که در ابتدا تصاویر ورودی توسط تبدیل موجک تجزیه شده و ضرایب تقریب و جزئیات آن بدست می‌آید. ماتریس‌های ضرایب تقریب و جزئیات بدست آمده را به ۱۶ زیر بلوک مساوی تقسیم کرده و برای هر کدام از این زیر بلوک‌ها میانگین و انحراف معیار محاسبه می‌گردد. بعد از استخراج ویژگی‌ها برای ارزیابی میزان دقت هر روش از یک طبقه بند MLP برای کلاس‌بندی استفاده شده‌است که در جدول ۱ نتایج بازشناسی برای هر کدام از این روش‌ها بررسی و با یکدیگر از نظر زمان استخراج ویژگی‌ها، میزان دقت یا نرخ بازشناسی و طول بردار ویژگی مقایسه شده‌اند. همانطور که مشاهده می‌شود نرخ بازشناسی با استفاده از ویژگی‌های سوبل ۸۹،۸۷ درصد بدست آمده‌است که نسبت به روش‌های دیگر از نظر زمان و نرخ بازشناسی بهتر بوده است. زمان استخراج ویژگی با استفاده از فیلتر گابور ۱۱۰ms است که نسبت به سایر روش‌ها بسیار بالا می‌باشد، دلیل این امر حجم محاسبات زیاد فیلتر گابور بوده به صورتی که برای n کانال نیاز به n کانولوشن می‌باشد.

رابطه (۹) و (۱۰) محاسبه می‌شود [۱۱]. در شکل (۳) نقاب‌های روبرتز و سوبل نشان داده شده‌اند.

$$g_x(x,y) = f(x,y) - f(x+1,y+1) \quad (9)$$

$$g_y(x,y) = f(x+1,y) - f(x,y+1) \quad (10)$$

تفاوت نقاب‌های سوبل و روبرتز، در تعداد همسایگی‌های به کار رفته به منظور آشکارسازی جهت‌ها می‌باشد. در نقاب سوبل از تعداد همسایگی‌های بیشتری استفاده شده است. در شکل (۴) یک نمونه از بلوک‌های ورودی و خروجی متناظر با آن تصویر بعد از اعمال نقاب روبرتز نمایش داده شده است.

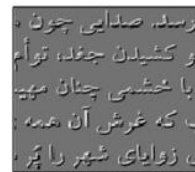
1	2	1
0	0	0
1-	2-	1-

1	0	1-
2	0	2-
1	0	1-

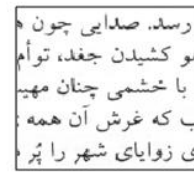
1	0
0	1-

0	1
1-	0

شکل ۳. نقاب‌های سوبل و روبرتز



(ب)

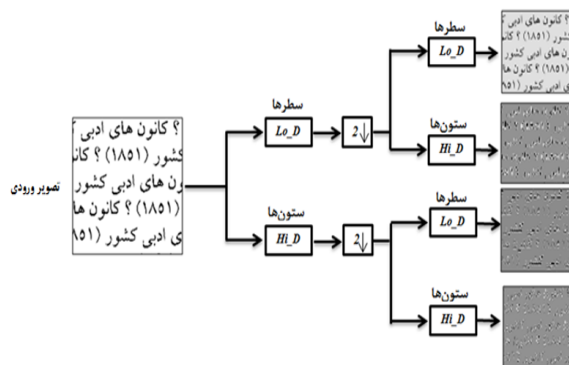


(الف)

شکل ۴: اعمال نقاب روبرتز به یک بلوک از تصویر ورودی، (الف) یک نمونه از بلوک ورودی، (ب) خروجی نقاب روبرتز

تبدیل موجک

تبدیل موجک دارای قدرت بزرگنمایی بوده و به صورت فشرده اطلاعاتی در مورد سیگنال و جزئیات آن در اختیار می‌گذارد. تبدیل موجک، ترکیبی از دو فیلتر پایین‌گذر و بالاگذر می‌باشد که در هر مرحله بر روی تصویر اعمال می‌شود و چهار زیر مجموعه مختلف به عنوان ضرایب تبدیل موجک تصویر بدست می‌آیند. ضرایب خروجی از فیلتر پایین‌گذر، شکل اولیه سیگنال را دنبال می‌کنند و به آنها ضرایب تقریب گفته می‌شود. همچنین ضرایب خروجی فیلتر بالاگذر، جزئیات فرکانس‌های بالای سیگنال را در بر دارند که به آنها ضرایب جزئیات گفته می‌شود و شامل جزئیات عمودی، افقی و قطری از تصویر می‌باشند. در این مقاله از تبدیلات موجک متفاوت Symlet، Db1 و Haar استفاده شده‌است. در شکل



شکل ۵. اعمال تبدیل موجک Symlet به تصویر ورودی از قلم لوتوس

جدول ۱. نرخ بازشناسی برای هر کدام از این روش‌های استخراج ویژگی

نرخ بازشناسی	زمان استخراج ویژگی برای هر نمونه	طول بردار ویژگی کاهش یافته	طول بردار ویژگی	ویژگی
۸۹,۸۷ %	۱۵ ms	۱۲۰	۲۵۶	SF
۸۹,۱۵ %	۱۲ ms	۱۲۰	۲۵۶	RF
۸۰,۷۴ %	۱۱۰ ms	۱۲۰	۲۵۶	Gabor
۷۵,۷۳ %	۳۶ ms	۱۲۰	۳۸۴	Haar
۷۲,۵۸ %	۳۷ ms	۱۲۰	۳۸۴	Db1
۷۴,۶۲ %	۳۲ ms	۱۲۰	۳۸۴	symlet

اصلی (PCA [۱۵]) تعداد ویژگی‌ها کاهش یافته‌اند. نحوه کاهش ویژگی با این الگوریتم به این صورت است که ابتدا ویژگی‌ها از ویژگی با بالاترین اهمیت برای جداسازی کلاسها تا ویژگی با کمترین اهمیت توسط PCA مرتب شده‌اند و بعد با استفاده از طبقه‌بند KNN ارزیابی ویژگی‌ها صورت پذیرفته است که در این مرحله با اضافه شدن هر ویژگی به ترتیب از با اهمیت ترین آنها تغییرات نرخ بازشناسی توسط الگوریتم بررسی گردیده که در نهایت فقط ویژگی‌هایی که در بازشناسی تاثیرگذار بوده‌اند انتخاب گردیده‌اند. با این کار حجم اطلاعات کم شده و اطلاعات زائد حذف گردیده‌اند و علاوه بر اِشغال کمتر حافظه، زمان پردازش آموزش شبکه‌های عصبی (خبره‌ها) کاهش می‌یابد. تعداد ویژگی‌ها برای هر کدام از مجموعه ویژگی‌های حاصل از ترکیب ویژگی‌ها ۶۴۰ عدد می‌باشد که پس از کاهش ویژگی این تعداد به ترتیب به ۲۰۰، ۲۰۰ و ۲۵۰ ویژگی برای SSF، DRF و HGF^{۱۱} رسیده‌اند. این مجموعه ویژگی‌های به دست آمده در بخش بعدی برای طبقه‌بندی و بازشناسی قلم‌ها مورد استفاده قرار گرفته‌اند.

پیاده سازی و مقایسه کارایی

در این بخش جزئیات پیاده سازی و مقایسه کارایی طرح‌های پیشنهادی با سایر کارهای ارائه شده در این زمینه ارائه می‌گردد. ابتدا ساختار شبکه مرکب که جهت طبقه بندی به کار گرفته شده است معرفی می‌گردد و در ادامه نتایج به دست آمده با استفاده از شبکه پیشنهادی آورده شده است و با سایر کارها مقایسه گردیده است.

طبقه‌بندی و اختلاط خبره‌ها

طبقه‌بندی به عنوان آخرین مرحله شناسایی قلم‌ها می‌باشد. در این مقاله از اختلاط خبره‌ها (ME) [۱۶] جهت طبقه بندی قلم‌ها و بازشناسی آنها استفاده گردیده است. برای طبقه‌بندی داده‌ها از سه شبکه عصبی به عنوان شبکه‌های پایه استفاده شده است که هر کدام بر روی مجموعه ویژگی‌هایی که به روش‌های مختلف در مرحله قبل از مجموعه داده‌ها استخراج گردیده‌اند آموزش داده شده‌اند و بعد از آموزش شبکه‌ها در لایه خروجی به منظور کمینه کردن خطا از الگوریتم جمعیت پرندگان (PSO) استفاده گردیده است. برای ایجاد این شبکه‌ها روش استخراج ویژگی و در نتیجه بردارهای ویژگی متفاوت بوده‌اند که به همین دلیل خطاهای شبکه‌ها نیز متفاوت می‌باشند و ترکیب این شبکه‌ها بهبود نرخ بازشناسی را نتیجه می‌دهد. در این مقاله از دو ساختار برای ترکیب شبکه‌های عصبی که شبکه‌های پایه را تشکیل می‌دهند استفاده گردیده است که این ساختارها در شکل ۶ نمایش داده

ترکیب و کاهش ویژگی‌ها

بعد از استخراج ویژگی‌ها با روش‌های ذکر شده، ویژگی‌هایی که از روشهای متفاوت به دست آمده‌اند را با یکدیگر به صورت ۲ به ۲ ترکیب کرده و سه مجموعه ویژگی جدید به دست می‌آوریم که هر کدام حاصل ترکیب دو مجموعه ویژگی از ویژگی‌های قبلی می‌باشد. در این مقاله با دو روش گرادیانی و سه روش تبدیل ویولت و گابور شش مجموعه ویژگی ایجاد گردیده است که ماهیت روشهای استخراج ویژگی به وسیله تبدیل‌های ویولت با سه روش دیگر متفاوت می‌باشند. به همین دلیل برای ترکیب، هر کدام از مجموعه ویژگی‌های استخراج شده به وسیله روش‌های تبدیل ویولت را با یک مجموعه ویژگی دیگر که از روش‌های فیلتر گابور و گرادیانی استخراج گردیده‌اند ترکیب گردیده‌اند.

اولین مجموعه، حاصل ترکیب ویژگی‌های استخراج شده توسط نقاب روبرتز و ویژگی‌های حاصل از موجک Db1 است که DRF نامیده شده است. این دو ویژگی تفاوت ذاتی در نحوه استخراج دارند. ویژگی روبرتز به صورت قطری از اطلاعات ۴ پیکسل که در همسایگی هم قرار دارند، گرادیان گرفته و ویژگی موجک Db1 از اطلاعات افقی، عمودی و قطری تصویر استفاده می‌کند.

مجموعه دوم از ترکیب ویژگی‌های حاصل از فیلتر گابور با ویژگی‌های حاصل از موجک هار به دست می‌آید (HGF). و در نهایت از ترکیب دو ویژگی گرادیان مبتنی بر نقاب سوبل و موجک Symlet حاصل شده است. با ترکیب ویژگی‌ها با این روش نرخ بازشناسی در مقایسه با هر کدام از مجموعه ویژگی‌های تکی بهبود یافته است که در جدول ۲ این نتایج ارائه گردیده‌اند.

بعد از ترکیب ویژگی‌ها که خود باعث بهبود در نرخ بازشناسی قلم‌ها می‌گردد با استفاده از الگوریتم تجزیه و تحلیل مؤلفه‌های

- 9 Sobel_Symlet Feature
- 10 Daubechies-Roberts Feature
- 11 Haar-Gabor Feature

نتایج به دست آمده از پیاده سازی و مقایسه

نتایج به دست آمده مربوط به دقت برای هر کدام از طبقه‌بندی‌های پایه به صورت تکی که بر روی یک مجموعه ویژگی ترکیب شده در مرحله قبل آموزش دیده‌اند و همچنین نتایج حاصل از ترکیب طبقه‌بندی‌های پایه با استفاده از دو ساختار ذکر شده در جدول ۲ آورده شده اند.

جدول ۲. نرخ بازشناسی بعد از ترکیب ویژگی‌ها و با استفاده از ترکیب طبقه‌بندی‌ها

نرخ بازشناسی	زمان طبقه‌بندی	طول بردار ویژگی کاهش یافته	طول بردار بردار ویژگی	ویژگی‌ها
٪۹۱،۵۴	۸۷۳ ثانیه	۲۰۰	۶۴۰	SSF
٪۹۲،۰۵	۶۴۴،۲۵ ثانیه	۲۰۰	۶۴۰	DRF
٪۸۸،۰۵	۱۴۸۹ ثانیه	۲۵۰	۶۴۰	HGF
٪۹۵،۸۸	۴۱۲۰ ثانیه	×	×	ME1
٪۹۷،۴۵	۵۲۳۷ ثانیه	×	×	ME2

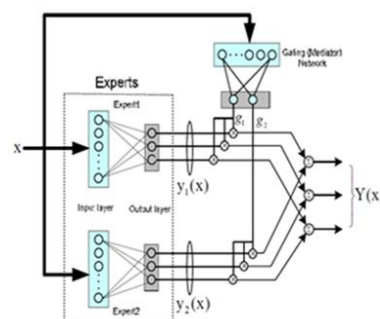
نتایج ذکر شده بعد از ۱۰ مرتبه اجرای الگوریتم‌ها و میانگین گرفتن از آنها حاصل گردیده است. مشاهده می‌شود که ترکیب طبقه‌بندی‌ها با ساختار دوم که به هر نرون خروجی یک وزن داده شده است با ۹۷،۴۵ درصد دقت بالاترین میزان دقت را داشته که نسبت به بهترین طبقه‌بندی تکی یعنی DRF به میزان ۵،۴ درصد بهبود یافته است. و بعد از آن نیز شبکه مرکب با ساختار اول با ۹۵،۸۸ درصد بالاترین دقت را داشته است که نسبت به بهترین طبقه‌بندی تکی به میزان ۳،۸۳ درصد بالاتری داشته است. با توجه به این جدول همان‌طور که گفته شد شبکه مرکب با ساختار دوم بهترین دقت و بعد از آن شبکه مرکب با ساختار اول قرار دارد ولی از نظر زمان آموزش شبکه مرکب اول به میزان قابل توجهی بهتر از شبکه مرکب با ساختار دوم می‌باشد، که دلیل آن نیاز به تکرار کمتر الگوریتم PSO برای یافتن وزن‌های بهینه می‌باشد [۱۷]. قابل ذکر است که برای یافتن وزن‌های بهینه با الگوریتم PSO برای ساختار اول از ۲۰۰ تکرار و برای ساختار دوم ۴۰۰ تکرار استفاده گردیده است.

نتایج حاصل از شبکه مرکب برای ساختار اول و دوم به ترتیب به صورت ماتریس همجوشی در جدول‌های ۳ و ۴ آورده شده است. در این جدول‌ها نرخ بازشناسی و همچنین درصدی از نمونه‌ها را که اشتباه تشخیص داده برای هر کدام از قلم‌ها به طور مجزا مشخص گردیده است و در نهایت نرخ بازشناسی کلی ارائه گردیده است. که با توجه به جدول‌های ارائه شده نرخ بازشناسی ۹۵،۸۸ درصد برای ساختار مرکب اول و ۹۷،۴۵ برای شبکه مرکب ساختار دوم به دست آمده است.

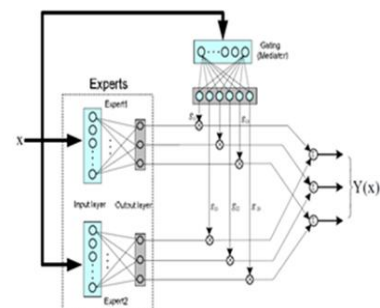
در ساختار اول که در شکل ۶-الف نمایش داده شده است خروجی شبکه‌های پایه توسط ۳ وزن که به آنها داده شده است با هم ترکیب گردیده‌اند. این وزن‌ها که در شکل با g نشان داده شده‌اند با استفاده از الگوریتم PSO جهت کمینه کردن مجموع مربعات خطا در مرحله آموزش شبکه به دست آمده‌اند. هر ذره در الگوریتم به صورت برداری چند بعدی می‌باشد که در ابتدا به صورت تصادفی انتخاب شده اند و طی هر تکرار الگوریتم به روز می‌گردند. در ساختار اول این ذرات در یک فضای سه بعدی برای پیدا کردن ۳ وزن مربوط به خروجی هر کدام از خبره‌ها در کمینه شدن تابع خطا به کار گرفته شده است و در ساختار دوم که در شکل ۶-ب نشان داده شده است، ذرات در فضای ۳۰ بعدی برای پیدا کردن ۳۰ وزن بهینه مورد استفاده قرار گرفته شده اند. که این ۳۰ وزن به نرون‌های خروجی هر ۳ شبکه پایه اعمال گردیده و در نهایت نرون‌های متناظر با هم ترکیب گردیده‌اند. شبکه‌های پایه یا خبره‌ها^{۱۱} در این مقاله شبکه عصبی ۳ لایه از نوع پیش‌رونده^{۱۳} و دارای ۴۰ نرون در لایه مخفی هستند. از تابع سیگموئید که در رابطه (۱۱) نشان داده شده به عنوان تابع فعال‌ساز در نرون‌های میانی و خروجی استفاده شده است.

$$\varphi(x) = \frac{1}{1 + e^{-kx}} \quad \text{with } k > 0 \quad (11)$$

پارامتر K در تابع سیگموئید، ۰،۳ در نظر گرفته شده و نرخ آموزش و گشتاور به ترتیب ۰،۵ و ۰،۸ در نظر گرفته شده اند.



الف



ب

شکل ۶. شبکه مرکب (الف) ساختار اول (ب) ساختار دوم [۱۸]

12 Experts
13 feed forward

جدول ۳. ماتریس همجوشی ترکیب اول

MRR	تایم	تاهما	تیتتر	هما	زر	یاقوت	ترافیک	نازنین	میترا	لوتوس	ماتریس همجوشی ترکیب اول
٪۱۴،۱۲	٪۰،۴۴	٪۰،۱۴	۰	۰	٪۱۰،۰۸	٪۰،۱۴	٪۰،۱۴	٪۲،۷۴	٪۰،۴۴	٪۸۵،۸۸	لوتوس
٪۶،۶۷	٪۰،۱۶	٪۰،۱۶	۰	٪۰،۳۲	٪۱،۹	۰	٪۰،۱۶	٪۳،۴۹	٪۹۳،۳۳	٪۰،۴۷	میترا
٪۷،۵۸	٪۰،۶۶	٪۰،۳۶	۰	۰	٪۲،۶۹	٪۰،۱۷	٪۰،۱۷	٪۹۲،۴۲	٪۲،۶۹	٪۰،۸۴	نازنین
٪۳،۴۹	٪۰،۳۲	٪۰،۹۵	۰	۰	٪۰،۹۵	٪۰،۱۶	٪۹۶،۵۱	٪۰،۶۳	٪۰،۳۲	٪۰،۱۶	ترافیک
٪۲،۳۶	٪۰،۱۶	۰	۰	٪۰،۱۶	٪۰،۳۱	٪۹۷،۶۴	٪۰،۴۷	٪۰،۹۴	٪۰،۱۶	٪۰،۱۶	یاقوت
٪۵،۹۸	۰	٪۰،۱۸	۰	۰	٪۹۴،۰۲	٪۰،۱۸	۰	٪۱،۷۹	٪۱،۲۵	٪۲،۵۸	زر
٪۰،۴۸	۰	۰	۰	٪۹۹،۵۲	۰	۰	٪۰،۱۶	۰	٪۰،۱۶	٪۰،۱۶	هما
٪۳،۵۳	۰	٪۰،۱۵	٪۹۶،۴۷	۰	٪۰،۳۱	۰	٪۰،۳۱	٪۱،۵۴	٪۱،۲۲	۰	تیتتر
٪۲،۵۸	٪۰،۴۹	٪۹۷،۴۲	۰	۰	۰	۰	٪۱،۱۰	٪۰،۱۶	٪۰،۴۹	٪۰،۳۴	تاهما
٪۲،۵۸	٪۹۶،۸۵	٪۰،۱۶	۰	۰	٪۰،۳۲	٪۰،۴۷	٪۰،۴۷	٪۱،۴۱	٪۰،۱۶	٪۰،۱۶	تایم
٪۹۵،۸۸											نرخ بازشناسی

جدول ۴. ماتریس همجوشی ترکیب دوم

MRR	تایم	تاهما	تیتتر	هما	زر	یاقوت	ترافیک	نازنین	میترا	لوتوس	ماتریس همجوشی ترکیب دوم
٪۱۳،۲	٪۰،۳	۰	۰	۰	٪۸،۴۲	۰	۰	٪۲،۲۸	٪۰،۲	٪۸۸،۸۰	لوتوس
٪۵،۷۴	۰	٪۰،۱۳	۰	۰	٪۲،۱۶	۰	٪۰،۱۳	٪۳،۰۲	٪۹۴،۲۶	٪۰،۳۰	میترا
٪۷،۴۳	٪۰،۵۹	٪۰،۳۶	۰	۰	٪۲،۴۷	٪۰،۳۶	٪۰،۳۶	٪۹۲،۵۷	٪۲،۴۷	٪۰،۸۲	نازنین
٪۲،۲۴	۰	٪۰،۶۳	۰	۰	٪۰،۸۶	٪۰،۲۵	٪۹۷،۷۶	٪۰،۲۵	٪۰،۲۵	۰	ترافیک
٪۲،۰۴	٪۰،۱۶	٪۰،۱۶	۰	٪۰،۱۶	٪۰،۳۱	٪۹۷،۹۶	٪۰،۳۱	٪۰،۷۸	۰	٪۰،۱۶	یاقوت
٪۶،۴۸	۰	٪۰،۱۸	۰	۰	٪۹۳،۵۲	٪۰،۱۸	٪۰،۱۸	٪۱،۸۰	٪۱،۴۲	٪۲،۷۲	زر
٪۱،۱	۰	۰	۰	٪۹۸،۹۰	٪۰،۳۱	۰	٪۰،۴۷	٪۰،۱۵	٪۰،۱۵	۰	هما
٪۲،۴۸	۰	۰	٪۹۷،۵۲	۰	۰	۰	٪۰،۳۱	٪۱،۲۴	٪۰،۹۲	۰	تیتتر
٪۲،۵۲	٪۰،۴۸	٪۹۷،۴۸	۰	۰	٪۰،۱۶	۰	٪۰،۷۸	٪۰،۱۶	٪۰،۶۳	٪۰،۳۱	تاهما
٪۲،۵۱	٪۹۸،۷۸	۰	۰	۰	٪۰،۱۶	٪۰،۳۱	٪۰،۳۱	٪۱،۴۱	٪۰،۱۶	٪۰،۱۶	تایم
٪۹۷،۴۵											نرخ بازشناسی

جدول ۵. مقایسه نتایج به دست آمده با روش به کار گرفته شده توسط سایر محققین

نرخ بازشناسی (%)	زمان مورد نیاز برای هر نمونه (میلی ثانیه)	طول بردار ویژگی ها	تعداد داده ها	پایگاه داده استفاده شده	روش به کار گرفته شده
۹۴،۱۶	۲۵	۵۱۲	۲۱۰۰۰	HODA	SRF + Boosting[11]
۸۱،۷	۳۲۹	۳۲۹	۲۱۰۰۰	HODA	Gabor [5]
۹۶،۹	-	۲۹۷	۱۰۰۰	Specific	Correlation Coefficients[10]
۹۸،۷	-	۴۸	۱۰۰۰	Specific	Gabor [19]
۹۵،۸۸	۱۶۱	-	۲۱۰۰۰	HODA	ساختار مرکب اول پیشنهادی
۹۷،۴۵	۲۴۹	-	۲۱۰۰۰	HODA	ساختار مرکب دوم پیشنهادی

در جدول ۵ مقایسه ای بین نتایج به دست آمده در مقاله ارائه شده با سایر پژوهشهای انجام شده در این زمینه انجام شده است. همانطور که دیده می شود نتایج بیانگر بهبود میزان دقت در کار ارائه شده نسبت به سایر کارهای انجام شده می باشد.

کارهای ارائه شده در [۵] و [۱۱] که بر روی پایگاه داده یکسانی با این کار انجام گردیده است از دقت پایین تر نسبت به ساختارهای پیش نهادی در این کار را دارد. کارهای ارائه شده در [۱۰] و [۱۹] نیز بر روی پایگاه داده دیگری پیاده سازی انجام

در جدول ۵ مقایسه ای بین نتایج به دست آمده در مقاله ارائه شده با سایر پژوهشهای انجام شده در این زمینه انجام شده است. همانطور که دیده می شود نتایج بیانگر بهبود میزان دقت در کار ارائه شده نسبت به سایر کارهای انجام شده می باشد.

- [3] A. Zramdini and R. Ingold, "Optical font recognition using typograohical feature," IEEE Trans, Patern Anal, Machine intell, vol. 20, p. 877_882, 2001.
- [4] J. Villegas-Cortez and C. Avile' s-Cruz, "Font recognition by invariant moment of global textures," presented at the Internat. Workshop on Very Low Bit-Rate Video-Coding, 2005.
- [5] Y. Zhu, T. Tan, and Y. Wang, "Font recognition based on global texture analysis," IEEE Trans. Pattern Anal. Machine Intell, vol. 23, p. 1192_1200, 2001.
- [6] H. Shi, Pavlidis, T., "Font Recognition and Contextual Processing for More Accurate Text Recognition," presented at the Fourth Int. Conferenc e Document Analysis and Recognition, in:proc, 1997.
- [7] P. S. Ramanathan. R, Thaneshwaran. Arun. L, Nair. S. Valliappan,N, "Tamil Font Recognition Using Gabor Filters and Support Vector Machines," presented at the International Conference on Advances in Computing, Control, and Telecommunication Technologies, 2009.
- [8] C. Avile's-Cruse, R. Rangel-Kuoppa, M. Reyes-Ayala, A. Andrade-Gonzalez, and R. Escarela-Perez, "High-order statistical texture analysis – font recognition applied," Pattern Recognition Lett, vol. 26, p. 135_145, 2005.
- [9] A. Borji and M. Hamidi, "Support vector machine for Persian font recognition," presented at the Internat. J. Intell. Technol, 2004.
- [10] E. Rashedi, Nezamabadi-pour, H., Saryzadi, S.,, "Farsi font recognition using correlation coefficients (in Farsi)," presented at the 4th Conf. on Machine Vision and Image Processing, Ferdosi Mashhad, 2007.
- [11] H. Khosravi and E. Kabir, "Farsi font recognition based on Soberly–Roberts features," Pattern Recognition Let, vol. 31, p. 75_82, 2010.
- [12] M. Sami Ben, Abderrazak, Z, Abdellatif, B, "New features using fractal multi-dimensions for generalized Arabic font recognition," Pattern Recognition Lett, 2010.
- [13] M. Ha, H. , X. Tian, D., and Z. ZHANG, R., "Optical font recognition based on Gabor filter," presented at the Proceedings of the Fourth International Conference on Machine Learning and Cybernetics, Guangzhou, August 2005.
- [14] L. Zhang, Y. Lu, and C. Lim Tan, "Italic Font Recognition Using Stroke Pattern Analysis on Wavelet Decomposed Word Images," presented at the Pattern Recognition IEEE, 2004.
- [15] R. O. Duda, P. E.hart, and D. G. stork, pattern classification. new york: Wiley-Interscience Publication, 2001.
- [16] L. I. Kuncheva, Combining Pattern Classifiers Methods and Algorithms. Canada: A JOHN WILEY & SONS, INC., PUBLICATION, 2004.

گردیده است و با توجه به اینکه تعداد نمونه‌ها و انواع قلم‌های آن با پایگاه داده این مقاله متفاوت است به صورت خام قابل مقایسه نیست. کار ارائه شده در [۱۹] بر روی پایگاه داده ۱۰۰۰ نمونه‌ای خاص بهترین جواب را ارائه کرده است در صورتیکه روش پیشنهادی بر روی پایگاه داده HODA با ۲۱۰۰۰ نمونه بهترین نرخ بازشناسی را نتیجه داده است.

نتیجه‌گیری

در این مقاله ویژگی‌ها با استفاده از نقاب‌های سوبل/روبرتز، فیلتر گابور و تبدیل موجک استخراج شده و با یکدیگر به صورت دو به دو ترکیب شده‌اند. سپس به منظور کم شدن حجم محاسبات و حافظه مورد نیاز، با استفاده از الگوریتم PCA تعداد ویژگی‌ها کاهش داده شده‌اند، در نهایت با استفاده از ترکیب طبقه‌بندهای پایه دقت بالایی حاصل گردیده است که نسبت به هر کدام از شبکه‌های منفرد به میزان قابل توجهی بهبود داشت. برای ترکیب شبکه‌ها دو ساختار متفاوت پیشنهاد گردیده است. که در ساختار اول با استفاده از PSO به هر کدام از خبره‌ها یک وزن داده شده و در ساختار دوم به هر کدام از نرون‌های خروجی توسط PSO یک وزن داده شده است. در ساختار اول زمان کمتری برای مرحله آموزش شبکه صرف شده است ولی در مقایسه با ساختار دوم دقت کمتری داشته است با توجه به اینکه این تفاوت در دقت زیاد نمی‌باشد و زمان مرحله آموزش این دو ساختار اختلاف چشمگیری دارد در جاهایی که زمان اهمیت دارد می‌توان از این ساختار استفاده کرد و در جاهایی که دقت اهمیت خیلی بیشتری نسبت به زمان دارد می‌توان از ساختار دوم استفاده کرد. نتایج به دست آمده در مقایسه با سایر پژوهش‌های انجام شده در این زمینه نشان دهنده بهبود میزان دقت در بازشناسی نسبت به سایر کارهای انجام شده می‌باشد.

تشکر و قدردانی

این مقاله مستخرج از طرح تحقیقاتی اجرا شده با شماره قرارداد ۶۵ مورخه ۱۳۹۳/۴/۱۱ از محل اعتبارات ویژه پژوهشی دانشگاه علوم و فنون دریایی می‌باشد.

مرجع‌ها

- [1] G. Nagy, "Twenty Years of Document Image Analysis in PAMI," IEEE Trans.Pattern Anal.Machine Intell, vol. 22, p. 38_26, 2000.
- [2] B. Cooperman, "Producing Good Font Attribute Determination Using Error-Prone Information," SPIE, vol. 3, pp. 50–57, 1997

- [17] J. L. Fernandez-Martinez and E. Garcia-Gonzalo, "Stochastic Stability Analysis of the Linear Continuous and Discrete PSO Models," IEEE Trans, Pattern Anal, Machine intell, vol. 15, pp. 405-423, 2010.
- [18] B. nabavi, "Classifier Combination," phd, Educate teacher Tehran, 2008.
- [19] H.Nezamaabdi-pour, S.Saryazdi, A.Ebrahimi, Font Recognition based on the Gabor, filters Conference: 9th Annual CSI Computer Conf., CSICC'04,, At Tehran, Feb2004.