

بازشناسی کاراکترهای پلاک خودرو با استفاده از ساختار اختلاط خبره‌ها

منصور نجاتی^۱

حسین پورقاسم^۲

چکیده

در این مقاله چارچوب طبقه‌بندی جدیدی برای بازشناسی کاراکترهای فارسی پلاک خودرو ارائه می‌شود. در این چارچوب، مجموعه‌ای از ویژگی‌های مختلف بر مبنای تصویرهای جهتی، آشکارساز لبه "کیرش"^۳ و میانگین‌های محلی برای کاراکترهای پلاک محاسبه می‌گردد. این ویژگی‌ها، سپس با استفاده از یک ساختار دینامیکی ترکیب طبقه‌بندها، با نام اختلاط خبره‌ها طبقه‌بندی می‌شوند. این ساختار شامل سه شبکه عصبی پرسپترون به عنوان طبقه‌بندهای پایه و یک شبکه پرسپترون دیگر به عنوان شبکه میانجی است که در یک فرایند رقابتی، تقسیم فضای ورودی بین طبقه‌بندهای پایه را یاد می‌گیرد و همزمان ضریب صلاحیتی را متناسب با داده ورودی، به خروجی هر طبقه‌بند پایه اعمال می‌نماید. الگوریتم بازشناسی پیشنهادی بر روی مجموعه داده‌ای شامل ۱۴۲۵۶ تصویر باینری کاراکترهای فارسی پلاک‌های رسمی ایران ارزیابی شده و دقت بازشناسی ۹۹/۴۲ درصد به دست آمد. نتایج تجربی نشان می‌دهند که در مقایسه با شبکه‌های عصبی پرسپترون، ساختار اختلاط خبره‌ها منجر به افزایش دقت در بازشناسی کاراکترهای فارسی پلاک خودرو می‌گردد.

کلیدواژه

بازشناسی پلاک خودرو، اختلاط خبره‌ها، پرسپترون چندلایه

^۱ دانشجوی دکتری برق، دانشگاه صنعتی اصفهان

^۲ استادیار دانشکده برق، دانشگاه آزاد اسلامی واحد نجف آباد h_pourghasem@iaun.ac.ir

^۳ Kirsch

با توجه به رشد بی‌رویه تعداد خودروها در سال‌های اخیر، استفاده از سیستم‌های هوشمند به دلیل کارایی بالا و صرفه جویی قابل توجه منابع انسانی و مالی، کاربرد وسیعی پیدا کرده‌اند. تشخیص و خواندن خودکار پلاک خودرو یک بخش اساسی در بسیاری از سیستم‌های هوشمند تصویری مانند مدیریت هوشمند پارکینگ‌ها [۲۰]، اخذ خودکار عوارض در بزرگراه‌ها [۳]، اعمال قوانین راهنمایی و رانندگی [۵ و ۴]، کنترل امنیتی مناطق انحصاری [۶] و کنترل ترافیک خودروها در جاده‌ها و اتوبان‌ها می‌باشد. سیستم‌های تشخیص پلاک خودرو (LPR^۴) عموماً دارای سه مرحله پردازشی هستند: شناسایی مکان پلاک در تصویر، بخش بندی تصویر پلاک به کاراکترهایش و بازشناسی کاراکترها. در این مقاله بر روی یک روش جدید برای مرحله بازشناسی کاراکترهای پلاک تمرکز می‌شود. الگوریتم‌های بازشناسی کاراکترها در LPR را می‌توان به دو گروه روش‌های مبتنی بر یادگیری و تطبیق الگو دسته‌بندی نمود. در گروه اول نشان داده شده است که با داشتن مجموعه داده آموزشی بزرگ، شبکه‌های عصبی مصنوعی می‌توانند به نتایج بازشناسی خوبی دست پیدا کنند و لذا رایج‌ترین روش مورد استفاده در سیستم‌های LPR هستند [۷-۱۳]. شبکه‌های عصبی پرسپترون چند لایه (MLPs^۵) از جمله شبکه‌های عصبی مصنوعی هستند که در [۸] برای شناسایی کاراکترهای فارسی پلاک‌های ایران و در [۹] برای شناسایی کاراکترهای لاتین در پلاک خودروها به کار رفته‌اند. برای آموزش این نوع شبکه‌ها از روش پس انتشار خطا استفاده می‌گردد و تعداد نرون‌های لایه مخفی نیز با آزمایش و خطا تعیین می‌شود. در [۱۰]، برای شناسایی کاراکترهای مختلف روی پلاک که شامل کاراکترهای چینی، حروف و اعداد انگلیسی هستند، از چهار شبکه عصبی پرسپترون چند لایه به ترتیب برای طبقه‌بندی حروف چینی، حروف انگلیسی، ترکیب حروف و اعداد انگلیسی و همچنین اعداد استفاده گردیده است. برای بهبود سرعت آموزش پس انتشار خطا نیز الگوریتم ارتجاع^۶ به کار رفته است. به عنوان نمونه، نگار [۱۱] دو نوع شبکه پس انتشار خطای سه لایه با تابع انتقال تانژانت سیگموئید^۷ و شبکه کوانتیزاسیون برداری یادگیر (LVQ^۸) را برای بازشناسی کاراکترها در یک سیستم LPR به کار برده و مقایسه‌ای را بین این دو نوع شبکه بر مبنای دقت بازشناسی انجام داده است.

^۴ License plate recognition

^۵ Multilayer perceptron

^۶ resilient

^۷ tan-sigmoid

^۸ Learning vector quantization

استفاده از شبکه‌های عصبی احتمالی (PNNs)^۹ برای بازشناسی پلاک خودرو اولین بار توسط آناگ نوستوپولوس^{۱۰} و همکارانش [۷] معرفی شد. در روش معرفی شده آن‌ها، از دو شبکه احتمالی، یکی برای بازشناسی حروف الفبا و دیگری برای بازشناسی کاراکترهای عددی استفاده شده است. مشابه همین روش نیز در [۱]، برای مرحله بازشناسی در یک سیستم LPR به کار رفته است. با این تفاوت که تنها از یک شبکه PNN برای بازشناسی کل کاراکترهای عددی و الفبایی موجود در پلاک استفاده شده است. تعداد نرون‌های مورد نیاز در شبکه‌های عصبی احتمالی معمولاً بیشتر از شبکه‌های استاندارد پس انتشار پیشخور^{۱۱} است. اما شبکه‌های PNN در کسری از زمان مورد نیاز برای آموزش شبکه‌های پیشخور، آموزش می‌بینند. البته این خصیصه مفید همراه با ایرادهایی مانند نیاز به حافظه بیشتر و زمان اجرای کندتر در مقایسه با شبکه‌های عصبی مرسوم می‌باشد.

ماشین‌های بردار پشتیبان (SVMs) نیز از دیگر روش‌های مبتنی بر یادگیری به کار رفته در بازشناسی کاراکترهای پلاک خودرو است [۱۵ و ۱۴]. در [۱۴]، از چهار طبقه‌بند مبتنی بر SVM برای شناسایی کاراکترها و اعداد ردیف بالا و پایین در پلاک خودروهای کشور کره جنوبی استفاده شده است. در هر یک از این طبقه‌بندها به تعداد کلاس‌های آن طبقه‌بند، SVM در نظر گرفته شده و با استراتژی یکی در برابر همه آموزش دیده اند. شبکه بازگشتی هاپفیلد [۱۲] و شبکه عصبی خود سازمانده مبتنی بر نگاشت‌های خود سازمانده کوهونس^{۱۲} [۱۳]، از دیگر شبکه‌های عصبی می‌باشد که به طور محدود در بازشناسی کاراکترهای پلاک به کار رفته‌اند.

سیستم‌های یادگیر در روش‌های ذکر شده فوق همگی بر روی کل فضای ورودی آموزش می‌بینند. یعنی به دنبال پیدا کردن مدلی هستند که بتوانند در کل فضای مسأله، پیش بینی‌های خوبی انجام دهند. با این حال در مسائل پیچیده، پیدا کردن یک مدل خوب بسیار مشکل است. روش مناسب‌تر در این موارد، افراز مسأله به چند زیرمسأله کوچک‌تر، حل کردن آن‌ها و سپس ترکیب جواب‌ها است. ترکیب طبقه‌بندها^{۱۳} در زمینه مسائل طبقه‌بندی با سیستم‌های یادگیر نیز از روش‌های معروفی است که برای افزایش دقت و بهبود کارایی در مسائل پیچیده طبقه‌بندی استفاده می‌شود. در این روش، از چندین طبقه‌بندی برای حل یک مسأله یکسان استفاده شده و پاسخ آن‌ها با روش مشخصی در یک پاسخ ترکیب می‌شود. این شیوه در مدلی که نمایش می‌دهد، می‌تواند انعطاف‌پذیری بیشتری داشته باشد. ترکیب طبقه‌بندها می‌تواند در دو ساختار استاتیک یا دینامیک انجام گیرد که در حالت دینامیک، سیگنال ورودی طبقه‌بندها به طور مستقیم در مکانیزم ترکیب

Probabilistic neural network^۹

Anagnostopoulos^{۱۰}

Feedforward backpropagation^{۱۱}

Kohonen^{۱۲}

Ensembles of classifiers^{۱۳}

خروجی‌ها دخالت دارد [۱۶]. اختلاط خبره‌ها که اولین بار توسط جاکوبس^{۱۴} و دیگران [۱۷ و ۱۸] معرفی گردید، مشهورترین روش در ساختار دینامیکی ترکیب طبقه بندهاست [۱۹]. اصل ضمنی در این ساختار، این است که طبقه‌بندهای سازنده قادر خواهند بود که به بخش‌های خاصی از فضای ورودی تخصص پیدا کنند و فضای ورودی به نوعی بین این طبقه‌بندهای خبره تقسیم می‌شود. در این مقاله برای افزایش نرخ بازشناسی کاراکترهای پلاک خودرو، از مدل ترکیبی بر اساس ساختار اختلاط خبره‌ها استفاده می‌شود که بر طبق دانش نویسندگان مقاله در زمان نگارش مقاله، این ساختار ترکیب طبقه‌بندها برای کاربرد مورد نظر به کار نرفته است. در این مدل، پرسپترون‌های چند لایه به عنوان بلوک‌های سازنده ساختار اختلاط استفاده می‌شوند. یک بردار ویژگی مستخرج از تصاویر باینری کاراکترهای پلاک به عنوان ورودی این طبقه بند ترکیبی به آن اعمال می‌شود. در روش پیشنهادی، از تصویرهای جهتی، خروجی آشکارساز لبه " کیرش " و میانگین‌های محلی به عنوان بردار ویژگی توصیفگر کاراکترهای باینری استفاده می‌گردد. این ویژگی‌ها با هزینه محاسباتی کمی به دست آمده و منجر به نتایج قابل قبول برای کاربرد بازشناسی کاراکترهای پلاک خودرو بصورت بلادرنگ می‌شوند. در ادامه، ابتدا در رابطه با جمع آوری مجموعه داده بحث شده و بعد از آن در مورد جزئیات استخراج ویژگی و نحوه آموزش طبقه‌بند ترکیبی مبتنی بر اختلاط خبره‌ها صحبت می‌شود. نتایج تجربی و نتیجه‌گیری نیز در بخش‌های نهایی مقاله ارائه می‌گردند.

روش پیشنهادی بازشناسی کاراکترهای خودرو مبتنی بر ساختار اختلاط خبره‌ها

روش پیشنهادی بازشناسی کاراکترهای خودرو مبتنی بر ساختار اختلاط خبره‌ها، شامل دو مرحله اصلی می باشد: در مرحله اول، استخراج ویژگی از تصویر باینری ورودی صورت می‌گیرد و مرحله دوم شامل مشخص نمودن کلاس یا برچسب تصویر ورودی با استفاده از یک ترکیب خاص از طبقه‌بندهای MLP است که در ساختار اختلاط خبره‌ها آموزش دیده‌اند. این دو مرحله در ادامه به تفصیل شرح داده خواهند شد.

Jacobs^{۱۴}

استخراج ویژگی

اولین مرحله در شناسایی کاراکترها در یک سیستم LPR، استخراج ویژگی از تصاویر ورودی است. انتخاب مجموعه ویژگی‌های مناسب، تأثیر زیادی بر نرخ طبقه‌بندی نهایی دارد. از طرفی برای کاربردهای بلادرنگ نیاز است تا الگوریتم استخراج ویژگی دارای حجم محاسباتی زیادی نباشد. در الگوریتم پیشنهادی، از تصویرهای جهتی، آشکارساز لبه "کیرش" و میانگین‌های محلی به عنوان ویژگی از کاراکترهای باینری استخراج می‌گردد که پیچیدگی کمی دارد و همچنین منجر به نتایج خوبی می‌گردند.

تصویرهای جهتی

تصویرهای جهتی یا هیستوگرام کاراکتر تصویر شده در جهت خاص، از ویژگی‌های مرسوم در بازشناسی کاراکتر نوری (OCR^{۱۵}) هستند. علاوه بر تصویرهای عمودی و افقی، ما از تصویر قطری نیز برای ایجاد تمایز بیشتر و استخراج بهتر اطلاعات هندسی-مکانی کاراکترها بهره می‌گیریم. هر المان در این تصویرهای جهتی برابر با حاصل جمع مقادیر پیکسل‌های کاراکتر باینری در آن جهت هستند. (پاسخ ۲-۲ و ۳-۲) با توجه به اینکه اندازه این ویژگی به ابعاد تصویر کاراکترها بستگی دارد، برای تولید بردارهای ویژگی هم اندازه در ورودی سیستم طبقه‌بند نیاز است که ابتدا تصاویر به یک اندازه مشخصی نرمال شده و سپس تصویرهای جهتی از آن‌ها استخراج گردد. این عمل نرمال‌سازی اندازه تصاویر باینری کاراکترها، به عنوان یک پیش پردازش و قبل از محاسبه ویژگی‌ها انجام می‌گیرد. مثالی از این تصویرهای جهتی برای یک کاراکتر در شکل ۱ ارائه شده است.

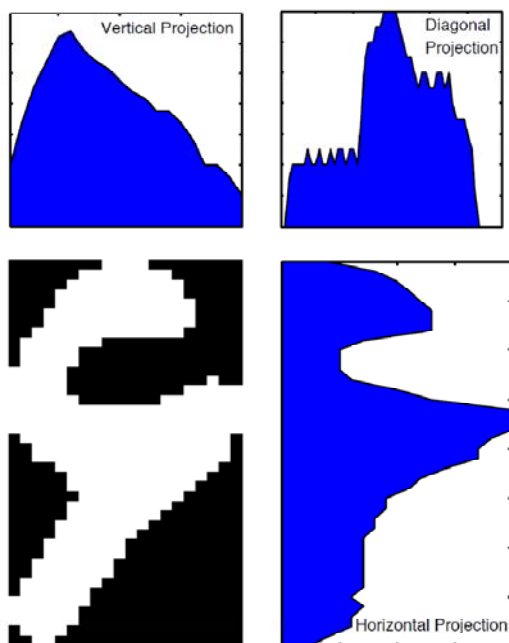
آشکارساز لبه "کیرش"

آشکارسازهای لبه تفاضلی مرتبه اول به لحاظ شناسایی محلی یک قطعه خط و محاسبه سریع، عملگرهای مناسبی هستند. آشکارسازهای لبه "کیرش"، "پریویت"^{۱۶} و "سوبل"^{۱۷} از جمله آشکارسازهای تفاضلی مرتبه اول لبه می‌باشند. البته در میان این آشکارسازهای لبه، آشکارساز لبه "کیرش" به دلیل در نظر گرفتن همه هشت همسایگی هر پیکسل، لبه‌های در چهار جهت اصلی را دقیق تر از دیگر آشکارسازها شناسایی می‌کند [۲۰]. الگوریتم غیرخطی ارتقاء لبه توسط "کیرش" به صورت زیر تعریف می‌شود [۲۱]:

^{۱۵} Optical character recognition

^{۱۶} Prewitt

^{۱۷} Sobel



شکل ۱. مثالی از تصویرهای جهتی عمودی، افقی و قطری برای یک کاراکتر عددی فارسی.

$$G(i, j) = \max \left\{ 1, \max_{k=0}^7 (|5S_k - 3T_k|) \right\}, \quad (۱)$$

که

$$S_k = A_k + A_{k+1} + A_{k+2}, \quad (۲)$$

$$T_k = A_{k+3} + A_{k+4} + A_{k+5} + A_{k+6} + A_{k+7}. \quad (۳)$$

در رابطه ۱، $G(i, j)$ برابر با گرادبان پیکسل (i, j) بوده، زیرنویس‌های A به پیمانه 8^{18} ارزیابی شده و A_k همسایگی هشت تایی پیکسل (i, j) می‌باشد که مانند شکل ۲

تعریف می گردند. ما در این مقاله، بردارهای ویژگی جهتی را برای چهار جهت افقی (H)، عمودی (V)، قطری راست (R) و قطری چپ (L) بصورت زیر محاسبه می نماییم:

$$G(i, j)_H = \max(|5S_0 - 3T_0|, |5S_4 - 3T_4|),$$

$$G(i, j)_V = \max(|5S_2 - 3T_2|, |5S_6 - 3T_6|),$$

$$G(i, j)_R = \max(|5S_1 - 3T_1|, |5S_5 - 3T_5|),$$

$$G(i, j)_L = \max(|5S_3 - 3T_3|, |5S_7 - 3T_7|). \quad (۴)$$

محاسبات فوق می تواند به سادگی با کانوالو کردن ماسک های لبه "کیرش" با تصویر باینری کاراکتر و سپس نگه داشتن بیشترین مقادیر در هر جهت انجام گیرد. به بیان دیگر، برای تولید تصویر لبه جهت دار G_i در جهت $i \in \{H, V, R, L\}$ ، کانوالوشن تصویر کاراکتر با دو ماسک "کیرش" متناظر با جهت i محاسبه شده و سپس برای هر مکان پیکسل در دو تصویر حاصل شده از کانوالوشن، مقدار قدرمطلق بزرگتر در G_i ذخیره می گردد. شکل ۳، ماسک های "کیرش" به کار رفته برای محاسبه تصاویر لبه جهت دار را نشان می دهد.

A_0	A_1	A_2
A_7	(i, j)	A_3
A_6	A_5	A_4

شکل ۲. تعریف همسایگی های هشت تایی پیکسل (i, j) .

5	5	5
-3	0	-3
-3	-3	-3

-3	-3	-3
-3	0	-3
5	5	5

(ب)

-3	-3	5
-3	0	5
-3	-3	5

5	-3	-3
5	0	-3
5	-3	-3

(الف)

-3	5	5
-3	0	5
-3	-3	-3

-3	-3	-3
5	0	-3
5	5	-3

(د)

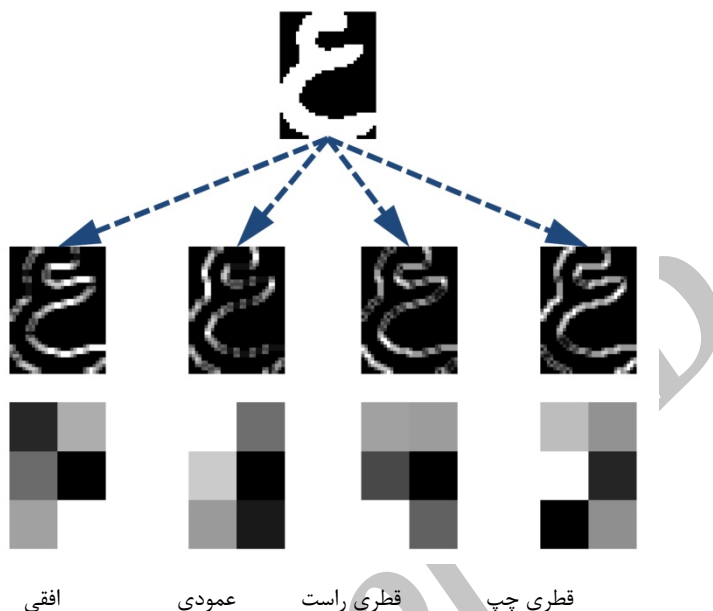
-3	-3	-3
-3	0	5
-3	5	5

5	5	-3
5	0	-3
-3	-3	-3

(ج)

شکل ۳. ماسک‌های آشکارساز لبه "کیرش"، (الف) عمودی، (ب) افقی، (ج) قطری چپ، (د) قطری راست.

به عنوان آخرین مرحله استخراج ویژگی "کیرش"، هر یک از تصاویر لبه جهت‌دار G_i ، $i \in \{H, V, R, L\}$ با اندازه 21×14 پیکسل به بلوک‌های زیرتصویر غیرهمپوشان 7×7 تجزیه شده و با انباشتن مقادیر پیکسل‌های هر بلوک، به یک تصویر ویژگی 3×2 فشرده می‌شود. بدین ترتیب انرژی محلی لبه‌های جهتی به عنوان ویژگی استخراج می‌گردند. این کار شبیه به تجزیه تصویر با استفاده از موجک‌ها و نگهداری ضرایب تقریب است. هر یک از این تصاویر فشرده شده به صورت یک بردار بازچیده می‌شوند و در کنار هم تشکیل یک بردار ویژگی با ابعاد $4 \times 6 = 24$ را می‌دهند. شکل ۴ مثالی از اعمال ماسک‌های لبه "کیرش" برای تولید چهار تصویر لبه جهت‌دار و تصاویر فشرده شده متناظر با آن‌ها را نمایش می‌دهد. بردار ویژگی "کیرش" برای دو کاراکتر متفاوت نیز در شکل ۵ ارائه گردیده است. در این شکل برای مثال مشاهده می‌شود که المان‌های آخر بردار ویژگی کاراکتر ۶، دارای مقادیر بزرگتری نسبت به المان‌های متناظر در بردار ویژگی کاراکتر ۲ هستند و دلیل آن این است که کاراکتر ۶ دارای لبه‌های قطری بیشتری نسبت به کاراکتر ۲ می‌باشد.



شکل ۴. مثالی از آشکارسازی لبه "کیرش". سطر دوم نشان دهنده تصاویر لبه جهت دار و سطر سوم تصاویر فشرده شده متناظر آن را نشان می دهد.

میانگین های محلی

میانگین های محلی از دیگر ویژگی هایی است که برای در نظر گرفتن مشخصه های سراسری تصویر ورودی استفاده می شود. برای استخراج این ویژگی، ابتدا تصویر به بلوک های غیرهمپوشان تجزیه شده و میانگین مقادیر پیکسل ها برای هر بلوک محاسبه می شود. از آن جا که تصویر ورودی یک تصویر باینری است. مقادیر محاسبه شده در واقع نشان دهنده درصدی از پیکسل های هر بلوک است که در برگیرنده کاراکتر می باشد. از طرفی به خاطر این که مقادیر مربوط به بلوک ها در کنار هم تشکیل بردار ویژگی را می دهند، اطلاعات مکانی نیز در آن لحاظ شده است. نحوه بلوک بندی در روش پیشنهادی به دو صورت انجام می گیرد. در حالت اول تصویر 21×14 ورودی مانند بخش (ب) به ۶ بلوک غیر همپوشان 7×7 پیکسل تجزیه شده و میانگین هر بلوک محاسبه می شود. در حالت دوم، یک بلوک 7×7 ، b_c ، در مرکز تصویر در نظر گرفته شده و سه ویژگی $f_{c,1}$ ، $f_{c,2}$ و $f_{c,3}$ برای آن محاسبه می شود:

$$f_{c,1} = \frac{E_b}{w_b \times h_b} \quad (5)$$

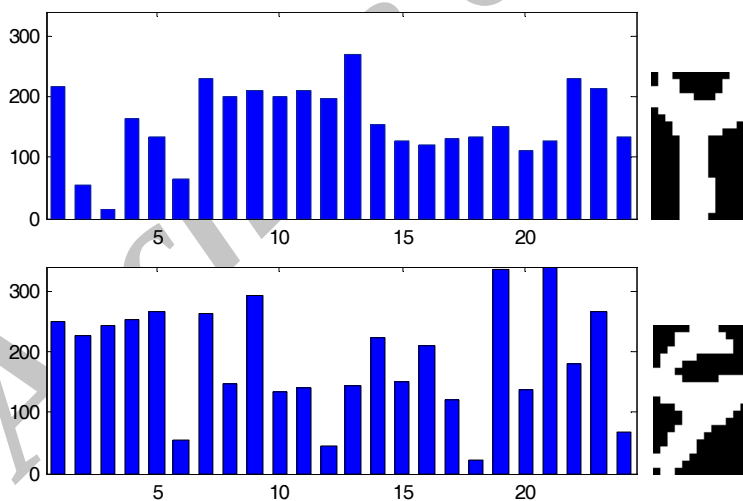
$$f_{c,2} = \frac{E_b}{E_I} \quad (6)$$

$$f_{c,3} = \frac{E_b}{w_I \times h_I} \quad (7)$$

که w_b و h_b به ترتیب برابر با پهنا و ارتفاع بلوک مرکزی، w_I و h_I برابر با پهنا و ارتفاع تصویر ورودی بوده و E_I و E_b نیز انرژی بلوک مرکزی و تصویر ورودی می‌باشند:

$$E_b = \sum_{i=1}^{h_b} \sum_{j=1}^{w_b} \delta(b_c(i, j) - 1), \quad E_I = \sum_{i=1}^{h_I} \sum_{j=1}^{w_I} \delta(I(i, j) - 1) \quad (8)$$

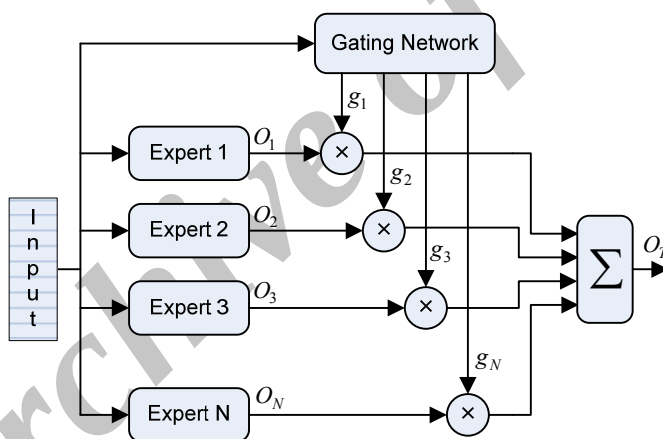
که $\delta(.)$ نشان دهنده تابع دلتا می‌باشد.



شکل ۵. بردار ویژگی مبتنی بر آشکارساز لبه "کیرش" برای دو کاراکتر عددی متفاوت.

طبقه‌بند اختلاط خبره‌ها

برای مرحله طبقه‌بندی در الگوریتم پیشنهادی بازشناسی کاراکترهای پلاک، ما از ترکیب خبره‌ها استفاده می‌نماییم. اختلاط خبره‌ها از مشهورترین روش‌های ترکیب طبقه‌بندهاست که در آن سیگنال ورودی بطور مستقیم در بکار انداختن مکانیزمی که خروجی طبقه‌بندها را در یک خروجی کلی ترکیب می‌کند، تأثیر دارد [۲۲]. ایده اصلی اختلاط خبره‌ها بر مبنای اصل تقسیم و غلبه^{۱۹} است که بر طبق آن، یک مسئله پیچیده محاسباتی به تعدادی مسائل ساده تر تقسیم شده و در نهایت حل مسئله اصلی با ترکیب حل مسائل کوچک‌تر امکان‌پذیر خواهد شد. اختلاط خبره‌ها اولین بار توسط "ژاکوبس" و همکارانش [۱۷] معرفی گردید. مدل پیشنهادی آنها از طبقه‌بندهای خطی ساده‌ای به عنوان طبقه‌بندهای پایه تشکیل شده است که خروجی آنها بوسیله یک شبکه میانجی ترکیب می‌شود. این طبقه‌بندها به‌طور همزمان آموزش داده می‌شوند، به طوری که فضای ویژگی ورودی به چند زیرفضا تقسیم شده و هر طبقه‌بند پایه در مدل سازی زیرفضای خودش خوب عمل می‌کند. ساختار این مدل در شکل ۶ نشان داده شده است.



شکل ۶. مدل اختلاط خبره‌ها در ترکیب طبقه‌بندها. این مدل از طبقه‌بندهای پایه (خبره‌ها) و یک شبکه میانجی تشکیل شده است. طبقه‌بندهای پایه برای یادگیری الگوهای ورودی رقابت کرده و شبکه میانجی نقش یک میانجی‌گر در این رقابت را ایفا می‌کند.

در مدل اختلاط خبره‌های "ژاکوبس"، طبقه‌بندهای پایه و شبکه میانجی دارای ساختار خطی هستند درحالی‌که برای حل مسائل پیچیده‌تر، باید از ساختارهای توانمندتری برای آنها استفاده

^{۱۹} Divide and conquer

کرد [۱۹]. در این مقاله به منظور بهبود کارایی، از یک مدل اصلاح شده اختلاط خبره‌ها که طبقه بندهای پایه و شبکه میانجی در آن دارای ساختار پرسپترون چند لایه هستند، استفاده می‌گردد.

آموزش اختلاط خبره‌ها

هر طبقه‌بند پایه در مدل اختلاط خبره‌های بهبود یافته، یک شبکه پرسپترون با یک لایه مخفی است که خروجی O_i را بصورت تابعی از بردار ورودی x ، وزن‌های لایه مخفی و لایه خروجی و تابع فعال‌ساز سیگموئیدی تولید می‌کند. به وسیله شبکه میانجی، وزن g_i به هر یک از این خروجی‌ها داده می‌شود (شکل ۶). شبکه میانجی از دو بخش تشکیل شده است: بخش اول آن یک شبکه پرسپترون چند لایه با تابع فعال‌سازی سیگموئید بوده و بخش دوم هم عملگر غیرخطی Softmax است. بخش اول شبکه میانجی، خروجی O_g را به عنوان خروجی شبکه پرسپترون چند لایه تولید کرده و تابع Softmax را برای به دست آوردن وزن‌های g_i به آنها اعمال می‌نماید:

$$g_i = \frac{\exp(O_{g_i})}{\sum_{j=1}^N \exp(O_{g_j})}, \quad i = 1, \dots, N \quad (9)$$

که N برابر با تعداد طبقه‌بندهای پایه موجود در مدل اختلاط خبره‌ها است. وزن‌های g_i مقادیر نامنفی بوده و حاصل جمع آنها ۱ خواهد بود و می‌توانند به عنوان احتمال تولید خروجی مطلوب y توسط طبقه بند i -ام تعبیر گردند. خروجی نهایی کل شبکه از ترکیب وزن‌دار خروجی طبقه‌بندهای پایه براساس خروجی شبکه میانجی محاسبه می‌شود:

$$O_T = \sum_{i=1}^N O_i g_i \quad (10)$$

مرحله آموزش شبکه، از الگوریتم پس انتشار خطا^۲ استفاده شده و وزن شبکه‌های پرسپترون چند لایه توسط این الگوریتم آموزش می‌بینند. برای هر طبقه بند پایه i و شبکه میانجی، وزن‌ها برطبق روابط زیر به روزرسانی می‌گردند.

$$\Delta w_y = \eta_e h_i (y - O_i)(O_i(1 - O_i))O_{hi}^T \quad (11)$$

$$\Delta w_h = \eta_e h_i w_y^T (y - O_i)(O_i(1 - O_i))(O_{hi}(1 - O_{hi}))x^T \quad (12)$$

$$\Delta w_{yg} = \eta_g (h - g)(O_g(1 - O_g))O_{hg}^T \quad (13)$$

^۲Error back propagation

$$\Delta w_{hg} = \eta_g w_{yg}^T (h - g)(O_g (1 - O_g))(O_{hg} (1 - O_{hg}))x^T \quad (14)$$

η_e و η_g به ترتیب نرخ یادگیری برای طبقه‌بندهای پایه و شبکه میانجی است. وزن های w_h و w_y به ترتیب وزن های لایه ورودی به لایه مخفی و لایه مخفی به لایه خروجی در طبقه‌بندهای پایه هستند. w_{yg} و w_{hg} نیز به ترتیب وزن های لایه ورودی به لایه مخفی و لایه مخفی به لایه خروجی در شبکه میانجی می‌باشند. بردارهای O_{hg} و O_{hi} ، خروجی لایه مخفی در طبقه‌بندهای پایه و شبکه میانجی بوده و O_{hg}^T و O_{hi}^T نشان دهنده ترانهاده آن‌ها است. h_i در این روابط برابر با تخمینی از احتمال پسین^{۲۱} تولید خروجی مطلوب y توسط طبقه بند i -ام است و طبق رابطه زیر به دست می‌آید.

$$h_i = \frac{g_i \exp\left(-\frac{1}{2}(y - O_i)^T (y - O_i)\right)}{\sum_{j=1}^N g_j \exp\left(-\frac{1}{2}(y - O_j)^T (y - O_j)\right)} \quad (15)$$

در طول فرایند آموزش، طبقه‌بندهای پایه برای یادگیری هر الگوی ورودی با هم رقابت می‌کنند و شبکه میانجی نرخ اصلاح وزن بیشتری به طبقه‌بند برنده اختصاص می‌دهد. از این رو در طول زمان، شبکه میانجی فضای ورودی را بر طبق عملکرد طبقه‌بندهای پایه تقسیم‌بندی می‌کند [۲۳].

نتایج تجربی

جمع‌آوری داده‌ها

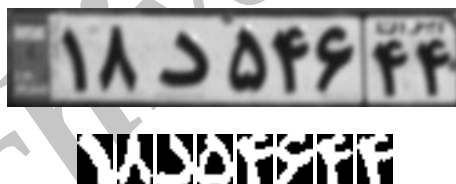
برای آموزش و ارزیابی طبقه‌بند معرفی شده در این مقاله، نیاز به مجموعه داده‌ای از تصاویر کاراکترهای پلاک می‌باشد. ما برای این منظور از یک مجموعه تصاویر سطح خاکستری^{۲۲} خودروهای سبک و سنگین که توسط واحد تصویربرداری یک سیستم LPR جاده‌ای اخذ شده بود استفاده می‌کنیم و به این وسیله یک مجموعه داده‌ای شامل ۱۴۲۵۶ تصویر باینری از کاراکترهای فارسی پلاک‌های رسمی ایران تولید نمودیم. تصاویر باینری این مجموعه، همگی توسط الگوریتم بخش بندی پلاک استفاده شده در یک سیستم LPR عملی (واقعی) که هم اکنون به طور آزمایشی در یکی از بزرگراه‌های شهر اصفهان در حال کار است، به دست آمد. سپس با گروه‌بندی آن‌ها به

Posterior probability^{۲۱}
Gray-scale^{۲۲}

دو گروه داده‌های آموزشی و آزمایشی، برای ارزیابی الگوریتم پیشنهادی مان مورد استفاده واقع می‌گیرد. شکل ۷، دو نمونه از تصاویر سطح خاکستری به کار رفته برای تولید مجموعه داده کاراکترها را نشان می‌دهد که ناحیه بزرگ شده پلاک نیز در گوشه پایین و سمت راست این تصاویر نمایش داده شده است. پلاک‌های رسمی ایران هر یک دارای ۸ کاراکتر بوده که یکی از آن‌ها حرف و مابقی عدد هستند و به طور کلی ۲۵ نوع کاراکتر مختلف در پلاک‌های رسمی ایران وجود دارد. تصویر یک پلاک رسمی ایران همراه با تصاویر باینری متناظر با کاراکترهایش که همگی به ابعاد یکسانی نرمال سازی شده‌اند، در شکل ۸ آورده شده است.



شکل ۷. دو نمونه از تصاویر اخذ شده توسط یک سیستم LPR نصب شده در بزرگراه. این تصاویر برای تولید مجموعه داده کاراکترها استفاده شده است.



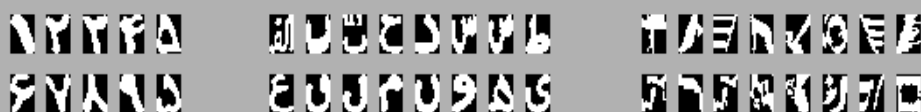
شکل ۸. نمونه پلاک رسمی ایران و تصاویر باینری متناظر با کاراکترهایش در اندازه نرمال شده.

همانطور که در بخش ۲-۱ اشاره گردید، برای داشتن بردارهای ویژگی هم اندازه نیاز است تا قبل از استخراج ویژگی از تصاویر باینری کاراکترها، ابعاد آن‌ها را در طی یک مرحله پیش پردازش به ابعاد مشخصی نرمال کنیم. برای این منظور، همه تصاویر مجموعه داده را به اندازه 21×14 پیکسل تغییر اندازه دادیم. این اندازه بر این اساس انتخاب شده است که هم حجم محاسباتی استخراج ویژگی زیاد نشود و هم دقت طبقه‌بندی کاراکترها تنزل پیدا نکند. تصاویر مجموعه داده را می‌توان به سه دسته کلی کاراکترهای عددی، کاراکترهای الفبایی و غیر کاراکترها تقسیم کرد. تعدادی از تصاویر این سه گروه در شکل ۹ نشان داده شده است. از این مجموعه داده، ۷۰۰۰ نمونه از تصاویر

به طور تصادفی انتخاب شده و به عنوان مجموعه آموزش در نظر گرفته شد. مابقی تصاویر نیز در مرحله آزمایش مورد استفاده قرار گرفتند. در مرحله استخراج ویژگی، سه نوع ویژگی معرفی شده در این مقاله، برای تمام تصاویر مجموعه داده، محاسبه شده و یک بردار ویژگی به اندازه 102×1 برای هر تصویر تولید و ذخیره گردید. این بردارها به عنوان ورودی لایه اول طبقه‌بندهای پایه و شبکه میانجی به کار می‌روند.

مقایسه با روش های ترکیب طبقه‌بندها

برای مرحله طبقه‌بندی کاراکترها، یک ساختار اختلاط خبره‌ها شامل سه شبکه پرسپترون سه لایه به عنوان طبقه‌بندهای پایه و یک شبکه پرسپترون سه لایه به عنوان شبکه میانجی استفاده گردید. برای هر یک از طبقه‌بندهای پایه، تعداد ۱۵ نرون در لایه مخفی و ۲۶ نرون در لایه خروجی و نرخ یادگیری $\eta_e = 0.5$ انتخاب شد. برای شبکه میانجی نیز از ۴ نرون در لایه مخفی، ۳ نرون در لایه خروجی و نرخ یادگیری $\eta_g = 0.1$ استفاده شد. تعداد نرون‌های لایه مخفی و نرخ‌های یادگیری فوق بعد از انجام آزمایشات متعدد با توپولوژی‌ها و نرخ‌های یادگیری مختلف برای طبقه‌بندهای پایه و شبکه میانجی انتخاب شدند. تعداد نرون‌های لایه خروجی طبقه‌بندهای پایه، برابر با تعداد کلاس‌های تصاویر است که شامل ۲۵ کاراکتر مختلف موجود در پلاک‌های رسمی ایران و یک کلاس برای تصاویر غیر کاراکتر می‌باشد. با استفاده از مدل اختلاط خبره‌ها با پارامترها و توپولوژی‌های اشاره شده در بالا، نرخ شناسایی برابر با ۹۹/۴۲٪ بر روی مجموعه تصاویر آزمایش به دست آمد.



(الف)

(ب)

(ج)

شکل ۹. تعدادی از تصاویر مجموعه داده بکار رفته در آزمایشات. (الف) کاراکترهای عددی. (ب) کاراکترهای حرفی. (ج) غیر کاراکترها.

به منظور مقایسه کارایی روش اختلاط خبره‌ها در بازشناسی کاراکترهای پلاک خودرو، یک شبکه پرسپترون چند لایه نیز به عنوان طبقه‌بند پیاده‌سازی شده و بر روی مجموعه داده یکسانی با مدل اختلاط خبره‌های استفاده شده در این مقاله، مقایسه گردید. این شبکه پرسپترون دارای یک لایه

مخفی بوده که با انتخاب‌های مختلفی برای تعداد نرون‌های لایه میانی و نرخ یادگیری، با استفاده از الگوریتم پس انتشار خطا آموزش داده شد و بهترین توپولوژی برای مقایسه با روش پیشنهادی مقاله مورد استفاده قرار گرفت. نتایج تجربی نشان می‌دهند که سه شبکه پرسپترون در ساختار ترکیبی اختلاط خبره‌ها، عمل بازشناسی را بهتر از یک شبکه پرسپترون انجام می‌دهند. جزئیات مربوط به پارامترها، بهترین توپولوژی‌ها و نرخ‌های شناسایی بر روی مجموعه آزمایش در جدول ۱ خلاصه شده است. با توجه به توپولوژی طبقه‌بند پرسپترون چند لایه و توپولوژی طبقه‌بندهای پایه در ساختار اختلاط خبره مشاهده می‌گردد که در ساختار ترکیبی، با استفاده از طبقه‌بندهای کوچک‌تر توانسته‌ایم به دقت بیشتری برسیم.

برای مقایسه ساختار اختلاط خبره‌ها با ساختارهای ترکیبی دیگر، روش‌های دیگر ترکیب طبقه‌بندها شامل روش رأی‌گیری، قواعد بیشینه، میانگین و حاصل ضرب نیز بر روی مجموعه داده پیاده‌سازی شدند. جدول ۲، تغییرات نرخ بازشناسی را نسبت به تعداد نرون‌های لایه میانی طبقه‌بندهای پایه برای بهترین نتایج حاصل از این روش‌های ترکیبی را نشان می‌دهد. همان‌طور که در این جدول مشاهده می‌شود، نرخ بازشناسی روش اختلاط خبره‌ها بر روی نمونه‌های آزمایش در بهترین آزمایش ۹۹/۴۲ درصد است که نسبت به بالاترین نرخ بازشناسی دیگر روش‌های ترکیبی پیاده‌سازی شده بر روی همین مجموعه داده، در حدود ۱/۳۳ درصد بهتر است.

جدول ۱. نرخ بازشناسی و مشخصات ساختار اختلاط خبره‌ها و شبکه پرسپترون چندلایه

مدل شبکه	پرسپترون چندلایه	اختلاط خبره‌ها
توپولوژی	102:40:26	شبکه‌های پایه: 102:15:26 شبکه میانجی: 102:4:3
نرخ یادگیری	0.5	شبکه‌های پایه: 0.5 شبکه میانجی: 0.1
نرخ بازشناسی (%)	97.52	99.42

جدول ۲. تغییرات نرخ بازشناسی نسبت به تعداد نرون های لایه مخفی در طبقه بندی های پایه برای روش های ترکیب خروجی طبقه بندی ها بر مبنای رأی گیری، قواعد بیشینه، میانگین و حاصل ضرب و مقایسه آن ها با ساختار اختلاط خبره ها.

نرخ بازشناسی کاراکترها (%)							تعداد نرون های لایه میانی
۴۰	۳۵	۳۰	۲۵	۲۰	۱۵	۱۰	
۹۶/۷۶	۹۷/۹۱	۹۵/۳۷	۹۶/۰۱	۹۳/۲۱	۹۲/۳۰	۹۰/۷۲	رأی گیری
۹۷/۶۲	۹۶/۶۹	۹۴/۴۵	۹۵/۳۳	۹۱/۳۶	۸۹/۸۰	۸۸/۳۴	میانگین
۹۸/۰۹	۹۸/۰۱	۹۷/۲۰	۹۶/۹۱	۹۶/۷۵	۹۴/۰۶	۹۲/۹۸	بیشینه
۸۶/۱۰	۸۱/۳۷	۷۶/۰۱	۸۰/۱۱	۸۰/۶۵	۸۳/۳۹	۸۰/۹۲	حاصل ضرب
۹۷/۵۲	۹۷/۳۶	۹۵/۳۰	۹۵/۹۷	۹۲/۷۱	۹۰/۴۱	۸۷/۶۳	پرسپترون چندلایه
۹۹/۳۶	۹۹/۴۰	۹۹/۴۲	۹۹/۳۶	۹۹/۴۰	۹۹/۴۲	۹۹/۱۳	اختلاط خبره ها

مقایسه با روش های موجود

در این بخش، روش پیشنهادی بازشناسی کاراکتر را با سه روش ارائه شده در مراجع [۱۱، ۱۸]، بر روی مجموعه تصاویر آزمایشی مقایسه می گردد. روش بازشناسی ارائه گردیده در [۱۱]، شامل استخراج ویژگی از تصاویر با استفاده از تبدیل پرتو پنکه ای^{۲۳} و سپس طبقه بندی آنها با یک شبکه LVQ است. تبدیل پرتو پنکه ای برای تولید تصاویر جهتی از یک تصویر در جهت های مختلف به کار می رود. این عمل با دوران منبع پرتو حول مرکز تصویر با گام دروان از پیش تعیین شده و تولید یک تصویر جهتی در هر دوران انجام می گیرد. گام دوران پیشنهاد شده در [۱۱] برابر با ۱۰ درجه بوده و ۳۶ تصویر جهتی متناظر با ۳۶ دوران منبع پرتو به دست می آید. رزولوشن هر تصویر جهتی برابر با ۱۵ نمونه انتخاب شده و بنابراین یک بردار ویژگی به اندازه ۵۴۰ المان برای هر تصویر کاراکتر تولید می شود. شبکه LVQ استفاده شده در مرحله طبقه بندی، دارای ۵۲ نرون در لایه رقابتی و ۲۶ نرون در لایه خطی است. در [۸]، از دو شبکه پس انتشار خطای سه لایه برای طبقه بندی کاراکترهای عددی و حروف پلاک استفاده شده است. بردارهای ویژگی به اندازه ۱۶ المان به عنوان ورودی این دو شبکه بکار رفته اند. این بردارهای ویژگی از تصاویر باینری کاراکترها استخراج شده، به طوری که هر تصویر به ۱۶ بلوک غیرهمپوشان تقسیم شده و نسبت پیکسل های سیاه به کل مساحت بلوک به عنوان ویژگی آن بلوک محاسبه می گردد. سومین روشی که آن را بررسی می نماییم، روش ارائه شده در [۱] است. در این روش از شبکه PNN برای طبقه بندی کاراکترهای

^{۲۳}Fan-beam transform

پلاک استفاده شده است. تعداد نرون‌های لایه مخفی در این شبکه برابر با تعداد نمونه‌های آموزشی می‌باشد که این تعداد در مجموعه داده ما برابر با ۷۲۰۷ نمونه است. الگوی خود تصاویر باینری کاراکترها نیز به عنوان ورودی شبکه لحاظ شده است. بر طبق [۱]، بعد از بخش بندی پلاک به کاراکترهایش، تصویر هر کاراکتر به اندازه 12×9 پیکسل تغییر اندازه داده شده و یک بردار الگو با قراردادن ستون‌های تصویر در زیر هم تشکیل می‌گردد.

همه روش‌های اشاره شده فوق و روش پیشنهادی بر روی مجموعه داده یکسانی مورد آزمایش قرار گرفته شد. مشخصات طبقه‌بندها به همراه نرخ بازشناسی کسب شده در جدول ۳ ارائه گردیده است. نتایج به دست آمده در این بخش نیز دلالت بر برتری روش پیشنهادی دارد. بهترین دقت بعد از روش پیشنهادی، مربوط به شبکه PNN می‌باشد. اما این شبکه نیاز به حافظه زیادی دارد.

جدول ۳. مقایسه روش پیشنهادی با روش‌های دیگر بازشناسی کاراکترها.

روش	نوع طبقه بند	توپولوژی	نرخ بازشناسی (%)
[۱۱] Nagare	LVQ	۵۴۰:۲۲۰:۲۶	۸۶/۹
[۸] Broumandnia et al.	BPNN	شبکه اعداد ۱۶:۲۰:۹ شبکه حروف ۱۶:۲۰:۱۶	۹۵/۳ ۹۲/۸
[۱] Anagnostopoulos et al.	PNN	۱۰۸:۷۲۰:۷:۲۶	۹۷/۳
روش پیشنهادی	Mixture of MLP experts	شبکه های پایه ۱۰۲:۱۵:۲۶ شبکه میانجی ۱۰۲:۴:۳	۹۹/۴

نتیجه‌گیری

در این مقاله، یک مدل اختلاط خبره‌ها مبتنی بر شبکه‌های پرسپترون چند لایه را برای بازشناسی کاراکترهای پلاک خودروها معرفی نمودیم که در آن، فضای مسئله به چندین زیر فضا برای خبره‌ها تقسیم شده و خروجی آن‌ها توسط یک شبکه میانجی ترکیب می‌شود. تصویرهای جهتی، آشکارساز لبه "کیرش" و میانگین‌های محلی برای استخراج یک مجموعه ویژگی مقاوم با حجم محاسباتی کم از تصاویر باینری کاراکترهای پلاک استفاده شدند. روش بازشناسی پیشنهادی بر روی مجموعه داده بزرگی از تصاویر باینری کاراکترهای پلاک مورد آزمایش قرار گرفت و نرخ بازشناسی بالایی به دست آمد. مقایسه این ساختار ترکیب طبقه‌بندها با روش‌هایی که تنها از یک طبقه‌بند برای مدل

کردن فضای ویژگی می کنند، نشان دهنده عملکرد بهتر ساختار ترکیبی اختلاط خبره ها است. عملکرد بهتر این ساختار نسبت به ساختارهای استاتیک ترکیب طبقه بندها مانند روش های خطی و غیرخطی مرتبه آماری نیز طی آزمایشات مختلف نشان داده شد. علاوه بر این، الگوریتم بازشناسی کاراکتر پیشنهادی با چند الگوریتم مختلف بر روی مجموعه داده یکسان مقایسه گردید و مشاهده شد که الگوریتم پیشنهادی نرخ بازشناسی بالاتری را فراهم می کند.

مراجع

- [1] C. Anagnostopoulos, I. E. Anagnostopoulos, V. Loumos and E. Kayafas, "A license plate-recognition algorithm for intelligent transportation system applications," IEEE Trans. Intell. Transp. Syst., vol. 7, no. 3, Sep. 2006, pp. 377–392.
- [2] T. Sirithinaphong and K. Chamnongthai, "The recognition of car license plate for automatic parking system," in Proc. 5th Int. Symp. Signal Processing and its Applications, 1998, pp. 455–457.
- [3] R. A. Lotufo, A. D. Morgan, and A. S. Johnson, "Automatic number plate recognition," in Inst. Elect. Eng. Colloquium on Image Analysis for Transport applications, 1990, pp. 611–616.
- [4] P. Davies, N. Emmott, and N. Ayland, "License plate recognition technology for toll violation enforcement," in Inst. Elect. Eng. Colloquium Image Analysis for Transport Applications, 1990, pp. 711–715.
- [5] F. Martín, M. García, L. Alba, "New methods for automatic reading of VLP's (Vehicle License Plates)," in Proc. of IASTED International Conference on SPPRA, 2002.
- [6] S. Draghici, "A neural network based artificial vision system for license plate recognition," Int. J. Neural systems, vol. 8, 1997, pp. 113–126.
- [7] C. Anagnostopoulos, E. Kayafas, V. Loumos, "Digital image processing and neural networks for vehicle license plate identification", Journal of Electrical Engineering, vol. 1, no.2, 2000, pp. 2–7.
- [8] A. Broumandnia, M. Fathy, "Application of pattern recognition for Farsi license plate recognition," in Proc. Int. Conf. GVIP, Cairo, Egypt, Jan. 2005, pp. 25–31.
- [9] J. A. G. Nijhuis, M. H. ter Brugge, K.A. Helmholt, J. P. W. Pluim, L. Spaanenburg, R. S. Venema, M. A. Westenberg, "Car license plate recognition with neural networks and fuzzy logic", in Proc. IEEE International Conference on Neural Networks, vol. 5, 1995, pp. 2232–2236.
- [10] L. Fuliang, G. Shuangxi, "Character Recognition System Based on Back-Propagation Neural Network", International Conference on Machine Vision and Human-Machine Interface (MVHI), 24-25 April 2010, pp. 393–396.
- [11] A. P. Nagare, "License plate character recognition system using neural network," International Journal of Computer Applications, vol. 25, no. 10, July 2011.

- [12] H. K. Sulehria, Y. Zhang, D. Irfan, A. K. Sulehria, "Vehicle number plate recognition using mathematical morphology and neural networks", WSEAS Trans. on Comp., vol. 7, no. 6, June 2008, pp. 781-790.
- [13] S. L. Chang, L. S. Chen, Y. C. Chung, and S. W. Chen, "Automatic License Plate Recognition", IEEE Trans. on Intelligent Transportation Systems, vol. 5, no. 1, 2004, pp. 42-53.
- [14] K.K.Kim K.I.Kim J.B.Kim H. J.Kim, "Learning-Based Approach, for License Plate Recognition", in proc. IEEE Signal Processing Society Workshop, Neural Networks for Signal Processing, vol. 2, 2000, pp. 614 - 623.
- [15] K. Parasuraman, and P.S. Subin, "SVM based license plate recognition system," in proc. International Conference on Computational Intelligence and Computing Research, 28-29 Dec. 2010.
- [16] R. Ebrahimpour, E. Kabir, H. Esteky, M. R. Yousefi, "A mixture of multilayer perceptron experts network for modeling face/nonface recognition in cortical faceprocessing regions", Intelligent Automation and Soft Computing, vol. 14, no. 2, pp. 145-156, 2008.
- [17] R. Jacobs, M. Jordan, S. Nowlan, and GE Hinton, "Adaptive mixtures of local experts", Neural Computation, vol. 3, no. 1, 1991, pp. 79-87.
- [18] L. Kuncheva, "Combining Pattern Classifiers: Methods and Algorithms," Wiley, 2004.
- [19] R. Ebrahimpour, E. Kabir, M. R. Yousefi, "Face detection using mixture of MLP experts", Neural Processing Letters, vol. 26, 2007, pp. 69-82.
- [20] S. W. Lee, "Off-Line recognition of totally unconstrained handwritten numerals using multilayer cluster neural network", IEEE Transactions On Pattern Analysis and Machine Intelligence, vol. 18, no. 6, June 1996, pp. 648-652.
- [21] W. K. Pratt, Digital Image Processing, New York: Wiley, 1978.
- [22] S. Haykin, Neural networks: a comprehensive foundation, Prentice Hall, USA, 1999.
- [23] M.N. Dailey, G.W. Cottrell, "Organization of face and object recognition in modular neural networks," Neural Networks, vol. 12, pp. 1053-1073, 1999.