

به نام خدا

مقدمه ای بر HCR

(زیر مجموعه ای از OCR)

گرد آوری

محمد حسن فردوسی زاده

کلمات کلیدی

پردازش تصویر، OCR، HCR، شبکه های عصبی

چکیده

در این مقاله مقدمه ای بر HCR (تشخیص نوری دست نویس) آمده است. در این مقاله در ابتدا روش هایی مختلفی که تا کنون برای این منظور استفاده شده است بررسی شده و سپس به مقایسه آنها پرداخته است.



خلاصه

بیش از سه دهه از کار بر روی سیستم های تشخیص دستنویس می گذرد. در این زمینه از روش های بسیاری بهره گرفته شده است که در میان آنها استفاده از شبکه های عصبی مصنوعی در سال های اخیر رواج بسیاری یافته است. از دلایل کاربرد وسیع شبکه های عصبی مصنوعی می توان به قابلیت تعمیم بالا، توانایی یادگیری قابل ملاحظه و استحکام نسبی در برابر اغتشاش اشاره کرد. این تحقیق نگاهی اجمالی به کاربردهای شبکه عصبی مصنوعی در بخش استخراج ویژگی و طبقه بندی حروف دارد. در این مسیر نقاط قوت و ضعف روش های مستقل از شبکه های عصبی بیان شده و چند نمونه ی تشخیص دستنویس با استفاده از شبکه های عصبی مصنوعی به همراه نتایج آن آورده شده است.

مقدمه

شناسایی الگو، شاخه‌ای از هوش مصنوعی است که با طبقه‌بندی و توصیف مشاهدات سروکار دارد. شناسایی الگو به ما کمک می‌کند داده‌ها (الگوها) را با تکیه بر دانش قبلی یا اطلاعات آماری استخراج‌شده از الگوها، طبقه‌بندی نماییم. الگوهایی که می‌بایست طبقه‌بندی شوند، معمولاً گروهی از سنجش‌ها یا مشاهدات هستند که مجموعه نقاطی را در یک فضای چند بعدی مناسب تعریف می‌نمایند. [۱]

در چند دهه گذشته بازشناسی الگوهای نوشتاری شامل حروف، ارقام و دیگر نمادهای متداول در اسناد نوشته‌شده به زبان‌های مختلف، توسط گروه‌های مختلفی از محققین مورد مطالعه و بررسی قرار گرفته است. نتیجه این تحقیقات منجر به پیدایش مجموعه‌ای از روش‌های سریع و تا حد زیادی مطمئن موسوم به «اُسی‌آر» یا «بازشناسی نوری حروف» به منظور وارد نمودن اطلاعات موجود در اسناد، مدارک، کتاب‌ها و سایر مکتوبات تایپی و حتی **دست‌نویست** به داخل رایانه شده است.

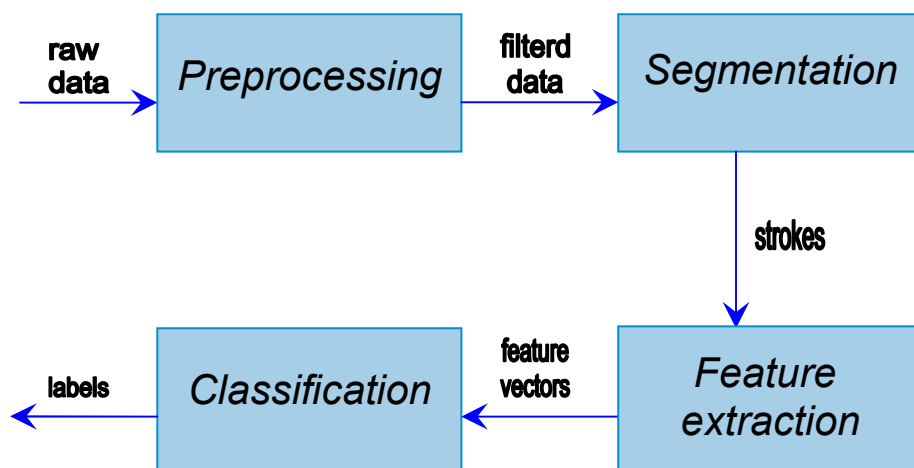
HCR یا همان تشخیص حروف دست‌نویس یکی از کاربرد های مهم بازشناسی نوری حروف است. تشخیص کاراکترهای دست‌نویس کاربردهای زیادی نظیر خواندن مبلغ چک بانکی، کد پستی، آدرس پستی و فرمهای مختلف دارد.

فصل اول شامل شرح مختصری بر HCR است، فصل دوم به بیان روش های پیشین در HCR می‌پردازد، فصل سوم به بررسی کاربرد های شبکه عصبی در HCR می‌پردازد و فصل چهارم به نتیجه گیری اختصاص دارد.

فصل اول

مقدمه ای بر HCR

هر سیستم HCR به چهار بخش زیر تقسیم می‌شود:



شکل ۱-۱ بلوک دیاگرام HCR

در زیر به توضیح هر یک از قسمت های زیر می‌پردازیم.

۱-۱ پیش پردازش

کلیه اعمالی که روی تصویر صورت می‌گیرند تا موجب تسهیل در روند اجرای فازهای بعدی گردد؛ مانند دوگانی کردن تصویر، حذف نویز، هموارسازی، نازکسازی، تشخیص زبان و فونت کلمات، و نظایر این‌ها. از مجموعه این پردازش‌ها، هدف‌های زیر دنبال می‌شود [۲]

- کاهش نویز
- نرمالیزه کردن داده‌ها
- فشردسازی میزان اطلاعاتی که می‌بایست محفوظ بماند.
- بازشناسی خط، زبان و فونت

۱-۱-۱ کاهش نویز

نویز ایجاد شده بواسطه دستگاه‌های اسکنر نوری منجر به ایجاد نقطه نقطه‌های لک‌مانند، قطعه خط‌های گسسته، اتصال بین خطوط، فضا‌های خالی در خطوط متن، پرشدن حفره‌های موجود در تصویر برخی حروف، و ... می‌گردد. همچنین اعوجاج‌های مختلف شامل تغییرات محلی، منحنی‌شدن گوشه‌های حروف، تغییر شکل یا خوردگی حروف را نیز باید در نظر داشت. قبل از مرحله بازشناسی حروف، لازم است که این نقایص برطرف شوند. مهم‌ترین دلیل برای کاهش نویز، کم کردن خطا در مراحل قطعه‌بندی و بخصوص بازشناسی می‌باشد. کاهش نویز همچنین سبب کم شدن اندازه فایل تصویر می‌شود که به نوبه خود، کاهش زمان مورد نیاز برای پردازش‌ها و ذخیره‌سازی‌های آینده را در پی خواهد داشت. [۳]

۱-۱-۲ نرمالیزه کردن

روش‌های نرمالیزه کردن داده‌ها به حذف تغییرات نگارشی کمک می‌کند و داده‌های استاندارد شده‌ای را نتیجه می‌دهد. روش‌های پایه نرمالیزه کردن عبارت‌اند از

الف. نرمالیزه کردن کجی متن و استخراج خطوط زمینه :

به دلیل بی‌دقتی در مرحله اسکن یا بی‌دقتی نویسنده در هنگام نگارش متن دست‌نوشته، ممکن است خطوط متن نسبت به تصویر، اندکی انحراف یا چرخش داشته باشند. این وضع ممکن است کارایی الگوریتم‌های به کار رفته در طبقات بعدی سیستم «اُسی‌آر» را تحت تأثیر قرار دهد؛ چرا که یکی از مفروضات در بیشتر روش‌های قطعه‌بندی، کج نبودن تصویر متن ورودی است و در نتیجه لازم است که این نقیصه، آشکار و تصحیح گردد. آشکارسازی خط زمینه در بسیاری از تکنیک‌های قطعه‌بندی و بازشناسی متون فارسی، عربی و لاتین، نقش اساسی دارد. [۴]

ب. نرمالیزه کردن اریب‌شدگی :

در متون چاپی فارسی و لاتین، کاراکترهای دارای قالب ایتالیک از راستای عمود، انحراف دارند. در متون دست‌نوشته نیز برخی از نویسنده‌ها حروف را به صورت زاویه‌دار می‌نویسند. این پدیده با عنوان «اریب‌شدگی» شناخته می‌شود و ممکن است دقت برخی از الگوریتم‌های قطعه‌بندی یا بازشناسی را تحت تأثیر قرار دهد و از این رو در این سیستم‌ها لازم است که در مرحله پیش‌پردازش، میزان اریب‌بودن کاراکترها شناسایی و تصحیح گردد.

ج. نرمالیزه کردن (تغییر مقیاس دادن) اندازه:

در سیستم‌های «آسی آر»، اغلب تصاویر کلمات یا حروف خیلی کوچک یا خیلی بزرگ، به یک اندازه استاندارد نرمالیزه می‌شوند تا بدین ترتیب عملیات بازشناسی، مستقل از اندازه فونت متن گردد.

د. هموارسازی کانتور: خط تشکیل‌دهنده مرز یک کاراکتر را کانتور آن کاراکتر گویند. در متون دست‌نویست، به واسطه لرزش یا حرکات ناخواسته دست نویسنده در هنگام نگارش، ممکن است که کانتور حروف ناصاف شود. این وضع در سیستم‌های بازشناسی متون چاپی و دست‌نویست نیز، به دلیل تغییر مقیاس حروف یا وجود نویز در مرحله اسکن تصاویر ممکن است ظاهر گردد. روش‌های هموارسازی کانتور، به منظور جبران این نقیصه مورد استفاده قرار می‌گیرند. به طور کلی هموارسازی کانتور، تعداد نقاط نمونه مورد نیاز برای نمایش کاراکتر را کاهش می‌دهد و در نتیجه کارایی مراحل پردازشی باقیمانده را بهبود می‌بخشد.

۱-۱-۳ فشرده‌سازی

این نکته پذیرفته شده است که تکنیک‌های کلاسیک فشرده‌سازی تصاویر که تصویر را از حوزه مکانی به حوزه‌های دیگر منتقل می‌کنند، برای بازشناسی حروف مناسب نیستند. در بازشناسی حروف، عمل فشرده‌سازی نیازمند آن دسته از تکنیک‌های حوزه مکانی است که اطلاعات شکلی را حفظ می‌نمایند. دو تکنیک متعارف فشرده‌سازی، یکی تکنیک اعمال سطح آستانه (به منظور دوگانی یا دوسطحی کردن تصاویر سطح خاکستری) و دیگری نازک‌سازی می‌باشند.

الف. دوگانی (دوسطحی) کردن تصویر متن:

تصاویر دیجیتالی به یکی از سه صورت- تصاویر رنگی، تصاویر سطح خاکستری (مشابه تصویر یک تلویزیون سیاه و سفید که رنگ تصویر به صورت سیاه، سفید و طیفی از رنگ‌های خاکستری ظاهر می‌شود)، و تصاویر دوگانی یا دوسطحی (مشابه تصویر یک سند فکس‌شده که رنگ پیکسل‌های تصویر، تنها سیاه یا سفید است)- می‌باشند. به منظور کاهش حجم ذخیره‌سازی مورد نیاز و افزایش سرعت و سهولت پردازش، اغلب مطلوب است که با انتخاب یک سطح آستانه، تصاویر سطح خاکستری یا رنگی را به تصاویر دوگانی تبدیل نمود.

ب. نازک‌سازی:

با این عمل، تصویر کاراکترها به تصویری با عرض یک پیکسل تبدیل می‌شود؛ درست مثل این که کاراکترها با یک قلم نوک باریک نوشته شده باشند. نازک‌سازی در حالی که کاهش قابل‌ملاحظه‌ای در حجم داده‌ها ایجاد می‌کند، اطلاعات شکلی کاراکتر را نیز حفظ می‌نماید.

۴-۱-۱ بازشناسی خط ، زبان و فونت

بازشناسی خط، تعداد کلاس‌های مختلف نمادهایی را که باید مورد ملاحظه قرار گیرند کاهش می‌دهد. شناسایی زبان متن، به منظور به کارگیری مدل‌های متنی خاص ضرورت دارد. طبقه‌بندی فونت‌ها، کاهش تعدد شکل‌های مختلف حروف در هر کلاس که لازم است در فرایند بازشناسی لحاظ گردند را به دنبال دارد و سبب می‌شود که امر شناسایی، تنها به یک کلاس فونت محدود گردد. بازشناسی خط و زبان و فونت در کاربردهایی مانند نمایه‌سازی و دستکاری اسناد نیز مطلوب می‌باشد.

۱-۲ قطعه بندی (جدا سازی)

قطعه‌بندی مرحله‌ای بسیار مهم برای سیستم‌های «اُسی‌آر» مخصوصاً «اُسی‌آر» فارسی و عربی (که حروف کلمات به صورت سرهم نوشته می‌شوند) می‌باشد. قطعه‌بندی به دو گونه تقسیم می‌شود:

۱. قطعه‌بندی بیرونی که عبارت است از تفکیک قسمت‌های مختلف تصویر نظیر متن، گرافیک و خطوط و نیز جدا کردن بخش‌های مختلف متن مانند پاراگراف‌ها، جملات یا کلمات؛
۲. قطعه‌بندی درونی که منظور از آن، جداسازی حروف کلمات مخصوصاً در مورد کلمات سرهم نوشته‌شده در متون لاتین، یا در رسم‌الخط‌های پیوسته نظیر فارسی و عربی است. همچنین حروفی که در متن اصلی جدا بوده‌اند، اما به خاطر کیفیت پایین دستگاه اسکنر به هم چسبیده‌اند، توسط این دسته از تکنیک‌ها از یکدیگر جدا می‌شوند. [۲]

مرحله قطعه‌بندی بیرونی، بحرانی‌ترین و حساس‌ترین قسمت در حوزه تحلیل تصویر اسناد می‌باشد و یک مرحله ضروری برای سیستم‌های «اُسی‌آر» برون خط محسوب می‌شود. گرچه مبحث تحلیل اسناد با روش‌ها و تکنیک‌های خاص خود یک حوزه تحقیقاتی تا حدی متفاوت نسبت به «اُسی‌آر» است اما تقسیم‌بندی تصویر سند به نواحی متنی و غیرمتنی، یک بخش لاینفک در نرم‌افزارهای «اُسی‌آر» به حساب می‌آید. بنابراین برای افرادی که در زمینه «اُسی‌آر» تحقیق می‌نمایند، داشتن دانش عمومی از تکنیک‌های آنالیز اسناد ضرورت دارد. تحلیل پیکربندی صفحات در سه مرحله انجام می‌گیرد: مرحله اول تفکیک نواحی متنی در تصویر از نواحی غیرمتنی (شامل گرافیک و خطوط) است. مرحله دوم، تحلیل ساختاری است که با قطعه‌بندی نواحی متنی به بلوک‌هایی از تصویر سند (نظیر پاراگراف، ردیف، کلمه، و...) مرتبط می‌باشد. مرحله سوم که تحلیل عملکردی نام دارد، با استفاده از اطلاعات مکانی، اندازه و قوانین مختلف صفحه‌بندی، عملکرد هر یک از اجزای سند (نظیر عنوان، چکیده، و ...) را تعیین می‌نماید. [۳]

نقطه تمایز اصلی میان سیستم‌های «اُسی‌آر» لاتین و فارسی برای متون چاپی، در مرحله قطعه‌بندی درونی نهفته است؛ چرا که حروف کلمات در نگارش فارسی برخلاف نگارش رسمی لاتین به صورت سرهم نوشته می‌شوند و در نتیجه ضرورت انجام صحیح این مرحله برای متون فارسی و عربی نسبت به متون لاتین، اهمیت فوق‌العاده بیشتری دارد. با وجود فعالیت‌های نسبتاً چشمگیر دهه گذشته و تنوع تکنیک‌های عرضه‌شده، قطعه‌بندی متون پیوسته (بخصوص متون دست‌نوشته پیوسته) به حروف، هنوز هم یک مسئله قابل بررسی مانده است. روش‌های قطعه‌بندی حروف به سه دسته تقسیم می‌شوند:

- قطعه‌بندی صریح
- قطعه‌بندی ضمنی
- تکنیک‌های ادغام‌شده

در مواردی همچون متون فارسی که حروف به صورت سرهم نوشته می‌شوند، سه رویکرد مختلف در بازشناسی برون خط متون کلمات یا زیرکلمات وجود دارد: [۵]

- رویکرد مبتنی بر قطعه‌بندی کلمات
- رویکرد مبتنی بر بازشناسی کلمه به عنوان یک الگوی واحد ،
- رویکرد ترکیبی.

در رویکرد بازشناسی مبتنی بر قطعه‌بندی، ابتدا کلمه در مرحله جداسازی به حروف یا زیرحروف، شکسته می‌شود؛ آنگاه قطعات جداشده بازشناسی می‌شوند و از کنار هم قرار گرفتن آن‌ها، کلمه شناسایی خواهد شد. روش‌های به کار گرفته‌شده در این رویکرد به دو گروه مختلف تقسیم می‌شوند:

- تقطیع کلمه به حروف،

- تقطیع کلمه به زیرحروف.

در گروه اول، کلمه به حروف جداسازی می‌شود و با شناسایی حروف جداشده، کلمه بازشناسی می‌گردد. در گروه دوم، کلمه به زیرحروف مثل پاره‌منحنی‌ها و ساختارهای پایه دیگر جداسازی می‌شود و با شناسایی زیرحروف‌ها و ترکیب آن‌ها، کلمه بازشناسی می‌گردد. در این رویکرد نمی‌توان در ابتدا مرز حروف را به طور کامل مشخص کرد، بلکه حروف به ترتیب از ابتدا به انتهای کلمه، بازشناسی و جداسازی می‌شوند. در هیچک از دو رویکرد نخست که مبتنی بر جداسازی هستند، به شکل کلی کلمه توجهی نمی‌شود و سعی بر آن است که با بازشناسی حروف یک کلمه، آن کلمه شناخته شود. در رویکرد بازشناسی کلمه به عنوان یک الگوی واحد، تلاشی برای تقطیع کلمه به حروف و بازشناسی حروف موجود در کلمه صورت نمی‌گیرد و کلمه در قالب یک الگو بررسی می‌گردد. قطعه‌بندی غلط کاراکترها، عامل بسیاری از خطاهای «آسی‌آر» است (مانند: $m \rightarrow rn$ یا $rn \rightarrow m$). میزان دقت یک الگوریتم قطعه‌بندی به سبک نگارش حروف، کیفیت دستگاه چاپ (کاراکترهای ایتالیک لکه‌دار برای قطعه‌بندی دارای اشکال می‌باشند)، و نیز نسبت اندازه فونت به قدرت تفکیک دستگاه اسکنر (تابع گسترش نقاط و نرخ نمونه‌برداری مکانی) بستگی دارد.

نتیجه مطلوب مرحله قطعه‌بندی، تصویری است که تنها حاوی یک کاراکتر باشد و بجز پیکسل‌های پس‌زمینه، هیچ شیء دیگری را شامل نشود. اما هنگامی که اشیای چاپی، در تصویر ورودی خیلی نزدیک به هم ظاهر شوند (مانند نقشه‌های هیدروگرافی)، این منظور همواره قابل حصول نخواهد بود. غالباً در چنین حالتی دیگر کاراکترها یا اشیای چاپی، به طور تصادفی در داخل تصویر کاراکتر قرار می‌گیرند و احتمالاً ویژگی‌های استخراج‌شده را تحریف می‌نمایند. این مورد یکی از دلایلی است که بیان می‌دارد چرا هر سیستم بازشناسی حروف، یک گزینه‌ی وازدگی دارد. [۵]

۱-۳ استخراج ویژگی

این مرحله یکی از مراحل بسیار با اهمیت در سیستم‌های «اُسی آر» است؛ چرا که نتایج حاصل از این مرحله، مستقیماً بر روی کیفیت مرحله بازشناسی اثر می‌گذارد. در مرحله بازنمایی، به هر الگوی ورودی (کاراکتر یا کلمه- بر حسب آن که رویکرد سیستم، مبتنی بر کدامیک از دو گروه «قطعه‌بندی کلمات» یا «شناسایی کلمه به عنوان یک الگوی واحد» باشد)، یک کد یا بردار ویژگی نسبت داده می‌شود که معرف آن الگو در فضای ویژگی‌ها است و آن را از دیگر الگوها متمایز می‌سازد. در انتخاب بردارهای ویژگی لازم است موارد زیر مورد توجه قرار گیرند:

- بردار ویژگی هر الگو باید تا حد امکان از بردارهای ویژگی دیگر الگوها متمایز باشد (فاصله بین بردارهای ویژگی در فضای ویژگی‌ها، حداکثر باشد).
- بردار ویژگی الگوها باید تا بیشترین حد ممکن، خصوصیات شکل و ساختار الگوها را از تصویر آن‌ها استخراج نماید.
- تا حد امکان نسبت به نویز، تغییر اندازه و نوع فونت، چرخش، و دیگر تغییرات احتمالی الگوها دارای ثبات باشد.
- شرایط، نوع و خصوصیات الگوهای ورودی در انتخاب بردارهای ویژگی اثر می‌گذارند. به عنوان مثال، باید تعیین نمود که آیا حروف یا کلماتی که می‌بایست تشخیص داده شوند جهت و اندازه مشخصی دارند یا خیر، دست‌نوشته هستند یا چاپی، یا این که تا چه حد بوسیله نویز، مغشوش شده‌اند. همچنین گاهی کفایت می‌کند که سیستم، تنها جوابگوی گروه محدودی از الگوها (مثلاً الگوهای با اندازه یا نوع فونت از پیش مشخص شده) باشد.
- در مورد حروفی که به چندین شکل نوشته می‌شوند (مانند 'و' و 'a', '4' و «۴») لازم است که بیش از یک کلاس الگو به یک کاراکتر خاص تعلق یابد. همانطور که عنوان شد، بازنمایی یک مرحله بسیار مهم در حصول راندمان مناسب برای سیستم‌های بازشناسی حروف است؛ ولی برای دستیابی به عملکرد بهینه، لازم است که دیگر مراحل نیز بهینه گردند و باید توجه نمود که این مراحل، مستقل از هم نیستند. یک روش خاص استخراج ویژگی‌ها، طبیعت خروجی مرحله پیش‌پردازش را به ما دیکته می‌کند یا حداقل ما را در انتخابمان محدود می‌سازد. مراحل قطعه‌بندی و بازشناسی، دو وجه تمایز عمده میان سیستم‌های «اُسی آر» فارسی و لاتین می‌باشند. بواسطه وجود تفاوت‌های اساسی بین نحوه نگارش فارسی و لاتین، امکان اعمال مستقیم تکنیک‌های قطعه‌بندی و بازشناسی مربوط به سیستم‌های «اُسی آر» لاتین، برای متون فارسی وجود ندارد. پیچیدگی‌های مختص نگارش فارسی، بر پیچیدگی الگوریتم‌های این دو مرحله می‌افزاید. درست به همین دلیل است که بیشتر نرم‌افزارهای «اُسی آر» تجاری لاتین، قادر به پشتیبانی زبان فارسی و عربی نمی‌باشند.

۱-۴ طبقه‌بندی و بازشناسی (با یک یا چند طبقه‌بندی کننده)

این مرحله شامل روش‌هایی برای متناظر ساختن هر یک از الگوهای به دست آمده از مرحله استخراج ویژگی‌ها، با یکی از کلاس‌های فضای الگوهای مورد بحث است که از طریق کمینه ساختن فاصله بردار ویژگی‌های هر الگوی ورودی نسبت به یکی از بردارهای مرجع، انجام می‌گیرد. بردارهای مرجع، بردارهایی هستند که قبلاً از نمونه‌های آموزشی اخذ شده‌اند. تکنیک‌های عرضه شده برای این مرحله را می‌توان در روش‌های مربوط به چهار گروه عمومی مبحث شناسایی الگو، جستجو نمود [۲]:

- تطابق قالبی،
- تکنیک‌های آماری،
- تکنیک‌های ساختاری،
- شبکه‌های عصبی.
- چهار گروه فوق لزوماً مستقل یا مجزا از یکدیگر نمی‌باشند و گاهی می‌توان تکنیک‌های یک گروه را در میان تکنیک‌های مربوط به دیگر گروه‌ها یافت.

فصل دوم

شرح چند روش مستقل از شبکه عصبی

در چند نمونه از روش های پیشین در زمینه HCR با بیان عملکرد و نقاط قوت و ضعف هر کدام بیان شده است. [۶]

۲-۱ روش بر پایه قانون مرتبه ای

۲-۱-۱ شرح روش

در این روش حروف به چهار گروه ابتدا، میانه، آخر و منفصل تقسیم شده اند. سیستم شناسایی حروف به تقسیم مرتبه ای بر حسب تعداد قسمت های هر حرف تقسیم بندی می شود. رابطه بین جایگاه نقطه در مقایسه با قسمت اصلی حرف، کمک به ساخت قانونی برای تشخیص هر حرف می کند. قدرت تشخیص این روش برای حروف ۱۰۰٪ گزارش شده است که البته مشخص نیست که این درصد برای داده های آموزش است یا داده های آزمایش.

۲-۱-۲ نقاط ضعف و قوت

این روش دارای نرخ تشخیص عالی و یک راهبرد خوب تقسیم و غلبه برای کاهش کلاس ها با استفاده از قانون مرتبه ای است.

این روش به دلیل وابستگی به تعداد قسمت های اصلی حروف حساسیت زیادی نسبت به نویز دارد. بعلاوه استفاده از نسبت نقاط کمینه و بیشینه ممکن است برای یک مجموعه داده خاص بهینه باشد اما قابلیت تعمیم بالایی ندارد.

۲-۲ روش بر پایه تحلیل ساختاری حروف جدا شده

۲-۲-۱ شرح روش

در این روش برای انتخاب ویژگی حروف از یک روش تحلیل ساختاری بهره برده شده است. طبقه بند از یک درخت تصمیم استفاده می‌کند. در پیش پردازش جداسازی حروف بر اساس روش Belaid&Haton برای یافتن بیشینه انحنا استفاده شده است. بعضی از ویژگی های استخراج شده در روند جداسازی کد های جهات، انحنا و همچنین حضور نقاط است. برای یافتن بهترین پارامتر ها می‌توان درخت تصمیم را به صورت دستی تنظیم نمود که البته ممکن است به پدیده بیش یادگیری (Overfitting) منجر شود. این سیستم با نوشته های ده نفر آزموده شده است که نرخ تشخیص آن ۸۶٪ بوده است.

۲-۲-۲ نقاط ضعف و قوت

این روش دارای دقت بالا و جداسازی خودکار حروف است. اما به چرخش حساسیت داشته و بر روی مجموعه کوچکی از داده ها و کلاس های خروجی آزمایش شده است. از نگرانی های موجود در این روش امکان رخداد بیش یادگیری است.

۲-۳ روش ساختاری و فازی

۲-۳-۱ شرح روش

این روش ترکیبی از دو تکنیک ساختاری و فازی است. تحلیل ساختاری کلاس های متفاوت را تمیز می‌دهد و منطق فازی قادر به تشخیص دستخط های مختلف در یک کلاس واحد است. در زمینه نتیجه آزمایش هیچ گزارشی ارائه نشده است.

۲-۳-۲ نقاط ضعف و قوت

این سیستم نتایج یادگیری عالی دارد و نقطه ضعف اصلی آن در این است که قوانین منطق فازی به صورت ابداعی است.

۲-۴- روش تطبیق قالب و برنامه نویسی پویا

۲-۴-۱ شرح روش

پژوهندگان در این روش با استفاده از تطبیق قالب و برنامه نویسی پویا نشان دادند که چگونه می توان خطا را در یک سیستم تشخیص برخط، کمینه کرد. در این روش بانک نمونه های اولیه پس از هموار سازی و هنجار سازی، بوجود می آید. بهترین نتیجه با ۹ نمونه اولیه و با دقت ۹۶٪ برای یک نویسنده حاصل شده است.

۲-۴-۲ نقاط ضعف و قوت

این روش نتایج آزمایشی بسیار خوبی دارد. اما ضعف آن از آنجا ناشی می شود که نمونه های مختلف آن دستخط های متفاوت از یک نفر هستند که باعث می شود تنوع در میان دستخط ها وجود نداشته باشد و نتایج قابلیت تعمیم پائینی دارد.

۲-۵- روش تطبیق قالب مرتبه ای و طبقه بندی بر اساس نزدیک ترین همسایه

۲-۵-۱ شرح روش

در این روش با استفاده از ویژگی های ثابت سعی شده است تا تعداد کلاس ها بر اساس تطبیق قالب کاهش یابند. ویژگی های ثابت عبارتند از:

- تعداد نقاط
- جایگاه نقاط در مقایسه با قسمت اصلی حرف
- تعداد قسمت های غیر اصلی حرف
- شیب قسمت غیر اصلی حرف

طبقه بند نزدیک ترین همسایه با استفاده از قسمت اصلی (کرسی) حرف نزدیک ترین کلاس را مشخص می کند. دقت تشخیص به نسبت طول رشته ارائه شده تغییر می کند. در مورد یک رشته ۶۰ کاراکتری از هفت نویسنده متفاوت دقت ۸۴٪ گزارش شده است. وزن دهی دستی ویژگی ها، دقت را تا ۹۳٪ بالا برده است.

۲-۵-۲ نقاط ضعف و قوت

مانند بسیاری از سیستم های دیگر دارای نتایج خوبی است همچنین شبیه به بسیاری از آنها به نوبت حساس بوده و قابلیت تعمیم خوبی ندارد. در ضمن مجموعه داده های آزمایش آن تنها دارای ۶۰ کاراکتر بوده اند.

فصل سوم

کاربردهای شبکه های عصبی مصنوعی در HCR

عمده کاربرد شبکه های عصبی در ۲ بخش زیر است:

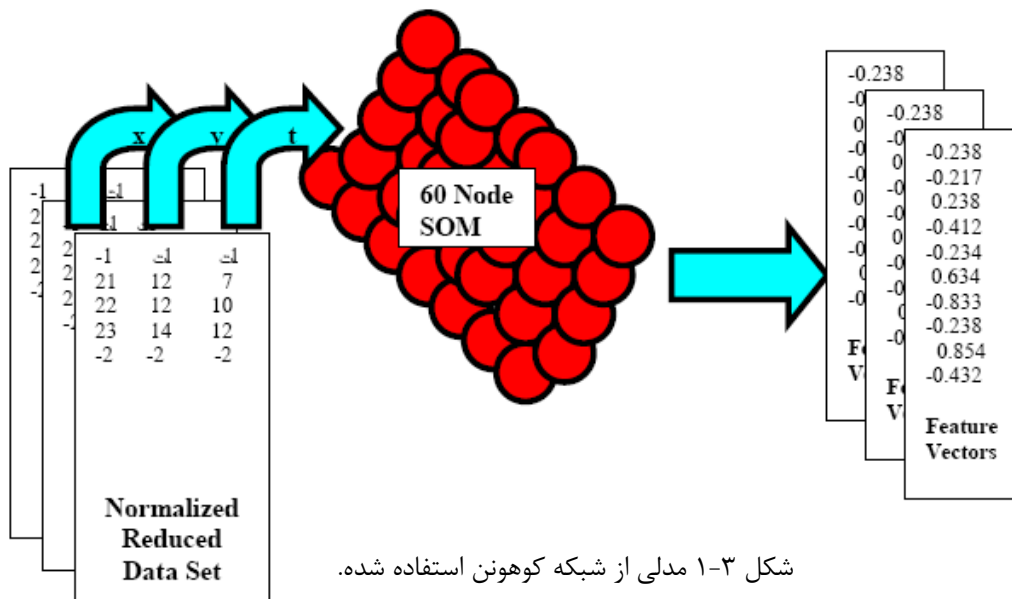
- استخراج ویژگی
- طبقه بندی

۳-۱ استخراج ویژگی

یک کاربرد شبکه های عصبی در زمینه استخراج ویژگی است. شاید مناسبترین شبکه برای استفاده در این قسمت شبکه کوهون باشد. این شبکه که یک شبکه بدون ناظر است دارای مزایای مختلفی است که از جمله آنها می توان به این مورد اشاره کرد که ویژگی هایی را که ممکن است به صورت دستی نتوان استخراج نمود و یا دارای حجم فوق العاده زیاد باشد را بدست می دهد.

در این قسمت دو نمونه را مورد بررسی قرار می دهیم :

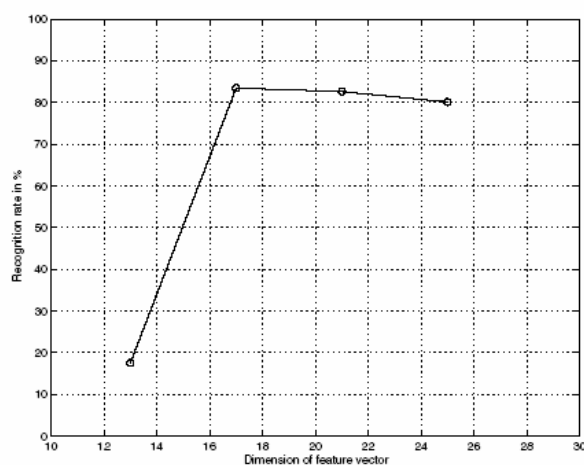
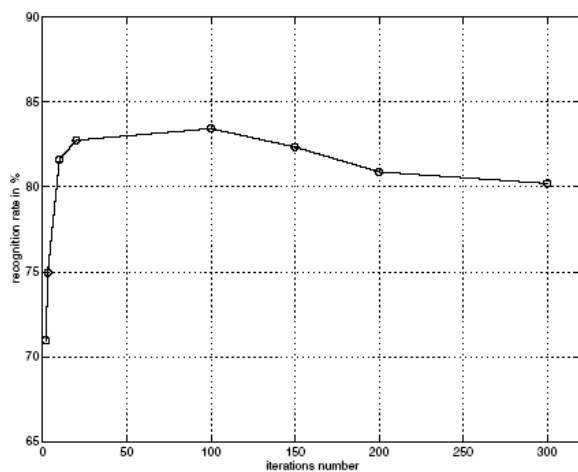
- در نمونه اول مورد بررسی پارامترهای شبکه کوهون به صورت تجربی تعیین شده است. در این نمونه با توجه به اینکه تشخیص بصورت برخط صورت می گیرد از مشخصات X ، Y و t (زمان) بعنوان ورودیهای شبکه کوهون استفاده شده است. با توجه به آزمایشات انجام شده نگارنده پی برده است که دو شبکه کوهون با ۷۰ نرون جواب قابل قبول خواهد داشت (نسبت به یک کوهون ۶۰ نرونی). همچنین تعداد تکرارها برای یادگیری ۵۰۰۰ انتخاب شده است که البته با بالا بردن تعداد تکرارها به ۷۵۰۰ تکرار، بهبود چندانی در نتایج حاصل نشده است. [۶]



شکل ۳-۱ مدلی از شبکه کوهون استفاده شده.

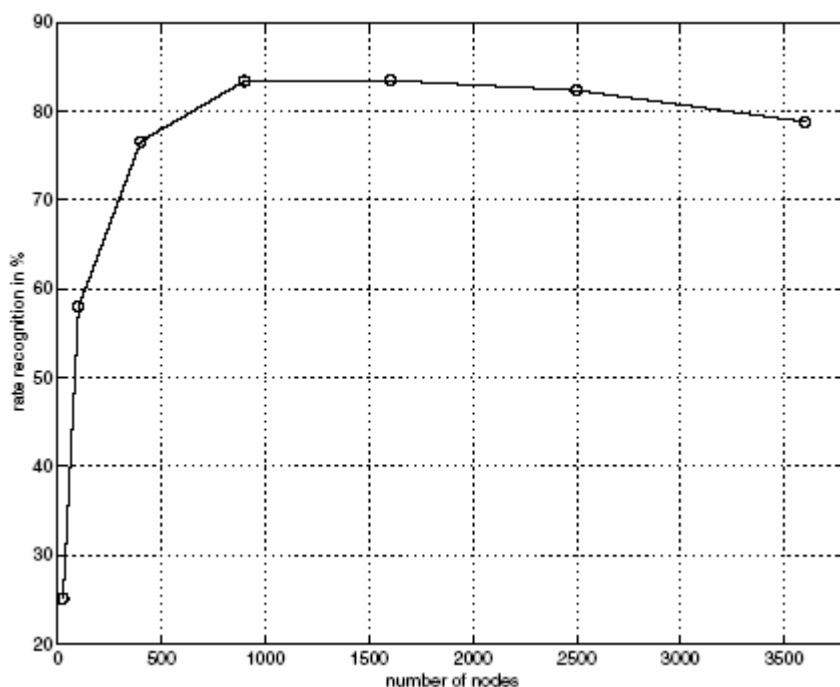
- در نمونه دوم نیز از یک شبکه کوهون برای استخراج ویژگی استفاده شده است. با توجه به اینکه این نمونه هم یک سیستم برخط است در آن از ضرایب سری فوریه مشخصات مکانی قلم نوری یعنی X و

Y بعنوان ورودیهای شبکه کوهونن استفاده شده است. در این روش ابعاد ورودی، تعداد نودها و تعداد تکرار برای شبکه کوهونن بصورت تجربی و با آزمایش بدست آمده است که پس از چند آزمایش و با ثابت نگه داشتن دو متغیر و تغییر متغیر سوم بهترین پاسخ با کمیت‌های زیر حاصل شده است :
ورودی شبکه کوهونن ۱۷ بعد، تعداد نودها ۱۶۰۰ و ۱۰۰ بار تکرار.[۷]



نمودار رابطه بین تعداد تکرار و نرخ تشخیص

نمودار رابطه بین ابعاد ورودی و نرخ تشخیص



نمودار رابطه بین تعداد نودها و نرخ تشخیص

۳-۲ طبقه بندی

کاربرد دیگر شبکه های عصبی در بخش طبقه بندی است. به این صورت که ویژگی های استخراج شده از نمونه ها به عنوان ورودی به یک شبکه عصبی (مثلا MLP) داده می شود. این شبکه وظیفه جدا سازی کلاس های متفاوت را به عهده خواهد داشت. در این بخش به ذکر چند نمونه از کارهای انجام شده می پردازیم :

- می دانیم که MLP برای حل مسائل غیر خطی مناسب است. در نمونه مورد بررسی ما که تشخیص دستنویس عربی است، با توجه به اینکه مسئله غیر خطی است، طبقه بندی بوسیله یک پرسپترون امکان پذیر نبوده و به همین دلیل از یک شبکه عصبی پرسپترون چند لایه استفاده شده است. در مثال بررسی شده خروجی های شبکه کوهونن که وظیفه استخراج ویژگی را بر عهده دارد، به شبکه MLP داده می شود تا از آن برای طبقه بندی استفاده شود. همچنین برای بهبود عملکرد سیستم تشخیص دستخط از روش هایی مانند هرس کردن شبکه استفاده شده است. در این مثال بهترین نتیجه حاصل شده برای داده های آزمایش با استفاده از الگوریتم هایی برای بهبود جواب به ۷۹٪ می رسد. [۶]

- در این مثال مانند مثال قبلی پژوهنده از یک شبکه MLP استفاده نموده است. این شبکه دارای ۸ نرون در لایه ورودی، ۴۰ نرون در لایه میانی و ۷۰ نرون در لایه خروجی است. قابل ذکر است که در این مثال [مرحل] جداسازی وجود نداشته و کل کلمه به شبکه عصبی مصنوعی که بعنوان طبقه بند استفاده شده، داده می شود. مجموعه داده های آزمایش ۷۰ کلمه مختلف نوشته شده توسط ۱۰۰ شخص است. نرخ تشخیص شبکه حدود ۶۶٪ گزارش شده است. [۸]

فصل چهارم

نتیجه گیری

با توجه به گسترش استفاده از شبکه های عصبی در سیستم های تشخیص دستنویس، آگاهی از نقاط قوت و ضعف این شبکه ها می تواند در استفاده بهینه از آنها راهگشا باشد. از مزایای استفاده از شبکه های عصبی مصنوعی در قسمت استخراج ویژگی می توان به توانایی این شبکه ها در یافتن ویژگی های نا محسوس و یا با حجم فوق العاده زیاد اشاره کرد. در قسمت طبقه بندی توانایی پردازش موازی در این شبکه ها باعث افزایش سرعت پاسخ دهی می شود. همچنین این شبکه ها به دلیل قابلیت تعمیم زیاد، توانایی تطبیق با دستخط های مختلف را دارا می باشد. البته استفاده از شبکه های عصبی مصنوعی معایبی نیز دارد که از آن جمله می توان به بیش یادگیری (Overfitting) اشاره کرد. رخداد این پدیده باعث می شود که شبکه در چند الگوی خاص کاملاً خبره گردد اما توانایی تشخیص الگو های دیگر را از دست بدهد. (کاهش قابلیت تعمیم) یکی دیگر از معایب شبکه های عصبی تجربی بودن تعیین شاخص های آموزش است. این امر باعث می شود فرآیند دستیابی به شاخص های بهینه زمانبر شود.

فهرست منابع

- 1-Theodoridis, Sergios; Koutroumbas, Konstantinos; Smith Ricky, Pattern Recognition, Academic Press, 1st edition, January 15, 1999.
- 2-Arica Nazif ,Yamin-Vural Fatos T., “An overview of character recognition based focused on off-line handwriting”, IEEE Transactions on Systems, Man, and Cybernetics-Part C: Applications and Reviews, Vol. 31, No. 2, May 2001.
- 3-O'Gorman, Lawrence; Kasturi, Rangachar, Document Image Analysis, IEEE Computer Society Press, Los Alamitos, 1995.
- 4-Nagy, G., “Twenty years of document image analysis in PAMI”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, pp. 38-62, Jan 2000.
- ۵-عزمی، رضا. بازشناسی متون چاپی فارسی، پایان نامه دکتری، به راهنمایی: کبیر، احسان الله، دانشکده فنی و مهندسی، دانشگاه تربیت مدرس، ۱۳۷۸.
- 6-T. Klassen, "Towards Neural Network Recognition of Handwritten Arabic Letters", MS. Thesis, Dalhousie University, Halifax, Nova Scotia, 2001.
- 7-N. Mezghani, A. Mitiche, M.Cheriet, "On-line recognition of handwritten Arabic characters using A Kohonen neural network", Proceedings of the Eighth International Workshop on Frontiers in Handwriting Recognition (IWFHR'02) 0-7695-1692-0/02 © 2002 IEEE
- 8-S.Alma'adeed, "Recognition of Off-Line Handwritten Arabic Words Using Neural Network", Proceedings of the Geometric Modeling and Imaging— New Trends (GMAI'06) 0-7695-2604-7/06 © 2006 IEEE