



سومین کنفرانس ملی مباحث نوین در کامپیوتر و فناوری اطلاعات
3rd National Conference on Advanced Topics in Computer and Information
Technology
بیست و هشتم آذر ماه ۱۳۹۸



ارائه روشی کارآمد جهت تشخیص حروف در تصاویر دیجیتال با استفاده از فرکتال دو بعدی و الگوریتم ژنتیک

عبدالجلیل بچاری

گروه مهندسی کامپیوتر - دانشگاه آزاد اسلامی واحد ماهشهر

a.j.bachari@gmail.com

مرجان عبدیزدان

گروه مهندسی کامپیوتر - دانشگاه آزاد اسلامی واحد ماهشهر

Abdeyazdan87@yahoo.com

بهمنوش براهیم زاده

گروه مهندسی کامپیوتر - دانشگاه آزاد اسلامی واحد دزفول

Bh.berahimzadeh@gmail.com

چکیده

تشخیص حروف، یکی از چالش‌های مرتبط حوزه پردازش تصویر است که به دلیل کاربرد آن در صنایع مختلف نظیر تبدیل نوشتار فیزیکی به اسناد الکترونیکی، تشخیص پلاک خودرو و کاربردهای نظیر آن در سیستم‌های امنیتی و ... همواره مورد توجه بوده است. عملکرد سیستم‌های OCR به شدت به بخش شناخت فونت بستگی دارد. اگر چه شناخت فونت در حال تبدیل شدن به بخش جدایی ناپذیر از هر سیستم OCR در زبان‌ها است، اما هنوز هم در زبان‌های مختلف توسعه نیافته است. روش‌های مختلف از رویکردهایی نظیر تحلیل ویژگی، انتخاب ویژگی و ایجاد مدل پیش‌بینی، به منظور افزایش دقت تشخیص حروف توسعه یافته‌اند. البته با همه پژوهش‌های انجام شده، به دلیل پویایی مسئله و تنوع داده‌های تولید شده، همچنان از نظر دقت و سرعت تشخیص در کاربردهای مختلف و البته زبان‌های مختلف، زمینه پژوهش‌های فراوانی وجود دارد. یکی از رویکردهای حل مسئله، توسعه مدل تحلیل تصویر براساس مدل‌های ریاضی استخراج ویژگی به منظور ایجاد توصیف دقیق در فازهای قطعه‌بندی و تشخیص است. در این پژوهش روشی مبتنی بر الگوی فراکتال مبتنی بر بافت برای استخراج ویژگی و الگوریتم ژنتیک برای بهینه‌سازی قطعه‌بندی تصاویر در دو فاز استفاده شده است. روش پیشنهادی در فاز اول کلمات را جدا و سپس حروف را قطعه‌بندی می‌کند و در نهایت به کمک نزدیک‌ترین الگوی مشابه تشخیص را انجام می‌دهد. نتایج روش پیشنهادی براساس معیارهای صحت، دقت، بازیابی و معیار F روی دو مجموعه داده MNIST و NEOCR نشان داد که روش پیشنهادی در مقایسه با روش‌های مبتنی بر ویژگی بسیار بهتر (تقریباً ۱۱ درصد) و قابل رقابت با روش‌های مبتنی بر یادگیری عمیق است.

واژه‌های کلیدی:

تشخیص حروف، فراکتال بافت، k-نزدیک‌ترین همسایه، الگوریتم ژنتیک، MNIST, NEOCR

۱- مقدمه

پس از پیدایش رایانه، حرکت بشر به سمت مکانیزه کردن کارهای انسانی روزه روز رشد بیشتری داشته است. ایجاد توانایی خواندن برای سامانه‌های رایانه ای ازجمله مواردی است که از سال‌های پیش در رویای بشر بوده و موضوع تحقیق بسیاری از پژوهشگران را به خود اختصاص داده است. شناسایی متون دست نویس و چاپی به ماشین این امکان را می‌دهد که بتواند نوشته‌های موجود درون یک تصویر را شناسایی و به متن قابل ویرایش و قابل جستجو تبدیل کند. شناسایی کاراکترها زیرمجموعه شناسایی الگو و پردازش تصویر است. از کاربردهای سامانه‌های شناسایی کاراکترها، می‌توان به کنترل خودکار گواهینامه‌ها، سامانه‌های دریافت اسکناس، کمک به نابینایان در امور خواندن، تبدیل کتابخانه‌های قدیمی به کتابخانه‌های دیجیتال، ورود اطلاعات گذرنامه، شناسایی داده‌های موجود در ورقه‌های آزمون، دسته‌بندی نامه‌ها در اداره پست و دسته‌بندی چک‌های دست‌نویس نیز اشاره کرد [۱]. در این میان زبان انگلیسی بیشترین سهم از پژوهش‌ها را به خود اختصاص داده و پیشرفت‌های قابل توجهی نیز در این زمینه حاصل شده است. به طوری که برخی از نرم‌افزارهای ارائه شده قادرند در شرایط تعیین شده تا نزدیک به ۱۰۰ درصد نتیجه صحیح را ارائه

دهند. با این وجود تحقیق‌ها در این زمینه کماکان ادامه دارد. هر ساله روش‌های جدید زیادی ارائه می‌شوند. بیشترین کارها و پژوهش‌هایی که در زمینه‌های شناسایی تصویری کاراکترها انجام شده بر روی زبان‌های انگلیسی، چینی و ژاپنی بوده است [۲].

۱-۱- بیان مسئله

با گسترش حوزه فعالیت سامانه‌های مکانیزه، بازشناسی خودکار متون و کاراکترها مورد توجه بسیاری از پژوهشگران قرار گرفته است. تحقیقات در زمینه زبان انگلیسی پیشرفت‌های بیشتری داشته است. در این زمینه سامانه‌هایی ایجاد شده است که با ضریب خطای نزدیک به صفر، عملیات بازشناسی را در شرایط تعیین شده انجام می‌دهند. در سیستم‌های بازشناسی حروف انگلیسی، از روش‌های متفاوتی نظیر مدل مخفی مارکوف، مدل‌سازی هندسی، شبکه‌های عصبی، منطق فازی و فواصل اقلیدسی استفاده شده است [۳]. از چالش‌های عمده در این زمینه تنوع نوع قلم یا دست‌خط، وجود نویز و یا اختلالات تصویری است که ممکن است عملیات شناسایی را با مشکل مواجه کند. در سال‌های اخیر بازشناسی متون فارسی، عربی، هندی و سایر زبان‌های مشابه بیش از پیش مورد توجه قرار گرفته است و پیشرفت‌های زیادی در این بخش حاصل شده است. به دلیل پیچیدگی‌های موجود در زبان

انگلیسی، ذاتا رسیدن به نقطه مطلوب در زمینه شناسایی، راه زیادی در پیش است [۴]. در این پژوهش، روشی نوین برای بازشناسی حروف انگلیسی در تصاویر ارائه شده است. که هدف آن تشخیص حروف و اعداد انگلیسی از روی تصاویر به کمک فرکتال دو بعدی و الگوریتم ژنتیک است.

۲-۱- اهمیت پژوهش

بازشناسی نوری حروف در حوزه‌های گوناگون مانند تشخیص پلاک خودرو، امنیت (مانند تصدیق گذرنامه)، بارکد و سیستم‌های بانکی (چک)، نقش مهمی ایفا می‌کند و تا کنون روش‌ها و سیستم‌های متعددی برای تشخیص کاراکترها ارائه شده است. از مهم‌ترین چالش‌های سیستم‌های بازشناسی موجود، این است که دارای تعداد زیادی نمونه آموزشی و در نتیجه حجم بالایی از محاسبات هستند، زیرا هنگام آموزش هر یک از کاراکترها به چندین نمونه آموزشی از آن کاراکتر، نیاز دارد. همچنین اگر این سیستم‌ها برای تشخیص کاراکترها با تنوع فونت‌های بیشتری توسعه یابند، خطای تشخیص، افزایش می‌یابد. بنابراین به سیستم بازشناسی ای نیاز داریم که قادر باشد علاوه بر کم بودن تعداد نمونه‌های آموزشی، در مقابل تغییر فونت، مقاوم بوده و خطای تشخیص آن در ازای افزودن فونت‌های جدید، تغییر آن چنانی نداشته باشد [۵].

۳-۱- اهداف پژوهش

هدف اصلی این پژوهش یافتن یک راه‌حل جدید برای تشخیص کلمات از تصاویر دیجیتال با استفاده از الگوریتم ژنتیک و فراکتال دوبعدی می‌باشد. هدف علمی این پژوهش، استفاده از این راه‌کار در جهت پیشبرد شناسایی متن از تصاویر با استانداردهای مختلف است.

هدف کاربردی این پژوهش در استخراج متن از تصویر، افزایش امنیت می‌باشد. مهم‌ترین رسانه مورد استفاده در عصر حاضر تصاویر هستند. بیشتر تصاویر حاوی انواع متن می‌باشند که بیانگر اطلاعات خاصی هستند. مانند تصاویری که از ماهواره از خیابان‌ها یا تصاویر خودروها و غیره تهیه می‌شود. برای دسترسی اطلاعات در این تصاویر، متن‌ها باید استخراج شوند.

۴-۱- ساختار پژوهش

این پژوهش در شش فصل تنظیم شده است. در فصل اول ضمن بررسی بیان مسئله، اهداف و فرضیات مسئله نیز بیان شده است.

در فصل دوم مبانی تحقیق و در فصل سوم کارهای مشابه قبلی که در زمینه تشخیص حروف انجام شده‌اند، مرور می‌شود. در فصل چهارم روش پیشنهادی بیان شده و در فصل پنجم نتایج مشاهده شده بررسی می‌شود. در نهایت در فصل ششم ضمن نتیجه‌گیری کلی، پیشنهاداتی برای کارهای آتی بیان خواهد شد.

۲- کلیات تحقیق

در این فصل، به مفاهیم مبنایی موجود و نیز پیشینه تحقیق پرداخته شده است.

۲-۱- تشخیص کاراکتر از متن

در این بخش مفاهیم مربوط به تشخیص کاراکتر از متن بررسی می‌شود. مفاهیم بررسی شده در این بخش عمدتاً ساختار تحقیق را مشخص می‌کند و محدود مسئله مورد بررسی بهتر بیان خواهد شد.

۲-۱-۱- اهمیت تشخیص حروف دست‌نویس

تشخیص حروف دست‌نویس کاربرد بسیاری دارد؛ برای مثال در مرتب کردن بسته‌های پستی، دفتر کار خودکار، خواندن و مرتب کردن حواله‌های بانکی و غیره. از تشخیص اعداد هم در کارهای مالی بسیار استفاده می‌شود. بنابراین سیستم باید از یک نرخ دقت بالا، سرعت زیاد و کاربرد آسان بهره‌مند باشد. تا به حال چندین روش ارائه شده است اما به دلیل اینکه دقت مناسبی نداشته‌اند مورد قبول واقع نشده‌اند. در یکی از این روش‌ها ترکیبی از شبکه عصبی و پرسپترون برای تشخیص ارقام دست‌نویس در یک رشته عددی استفاده شده است. مروری بر تکنیک‌های نرمال‌سازی و استخراج الگو صورت گرفته است. یک نگاهت جدید برای تشخیص حروف عربی بیان شده است. این سیستم صرف‌نظر از اندازه، جهت و موقعیت الگوی ورودی کار می‌کند [۶].

۲-۱-۲- برون خط / درون خط

به‌طور کلی، تشخیص دست خط به دودسته طبقه‌بندی شده است:

- برون خطی (آنلاین)
- درون خطی (آفلاین)

برای تشخیص دست خط، ورودی سیستم ما یک تصویر اسکن شده شامل متن پاراگراف دست‌نوشته است. تصویر باید در قالب تعیین شده باشد. پیش‌پردازش گام اول در تشخیص دست خط است. پیش‌پردازش چندین گام را به کار می‌گیرد: قطعه‌بندی خط، اصلاح پایه و غیره. برای مثال ورودی تصویر متن خاص در را گرفته

و خروجی کلمات تقسیم شده را می دهد [۸].

۲-۱-۳- تشخیص کاراکتر نوری

تشخیص کاراکتر نوری معمولاً به عنوان فرآیند تشخیص کاراکتر خارج از خط مطرح می شود، به این معنی که سیستم تصاویری را تشخیص می دهد که از کاراکترهای غیرپویا تشکیل می شوند. از سوی دیگر، تشخیص کاراکتر دست خط درون خطی نیاز به شناسایی حرکت های یک قلم دارد. در تشخیص کاراکتر برون خطی، ماشین کاراکترهای چاپ شده و اعداد دست نویس می تواند با دقت بالاتر از ۹۹٪ شناخته شود. تشخیص کامل تنها با بررسی انسان می تواند تصدیق شود. در حال حاضر چندین نرم افزار تجاری و ابزارهای موجود برای این کاربرد خاص وجود دارد. با این حال، تمرکز این تحقیق ها بر فرآیند تشخیص کاراکتر دست خط برون خطی (آفلاین) است [۹].

۲-۲- الگوریتم ژنتیک

گلدبرگ (از شاگردان هلند) توانست در سال ۱۹۸۹ با انتشار کتاب «الگوریتم ژنتیک در جستجو، بهینه سازی و آموزش ماشین»، چارچوب و قواعد حاکم بر الگوریتم ژنتیک را ارائه کند و در سال ۱۹۹۰ همگرایی الگوریتم ژنتیک ثابت شد. الگوریتم های ژنتیک، هیوریستیک های جستجو و الگوریتم های بهینه سازی هستند که به طور موازی اجرا می شوند و از اصل داروین در رابطه با انتخاب طبیعی و تکثیر ژنتیکی الهام گرفته شده اند. به عبارت دیگر این الگوریتم ها، تکنیک های بهینه سازی بر اساس انتخاب و ترکیب مجدد (باز ترکیبی) راه حل های امیدبخشی می باشند. هدف از اعمال الگوریتم ژنتیک بر روی این مسائل، بررسی توانایی این الگوریتم در به دست آوردن جواب های بهینه و نیز بررسی نحوه همگرایی و عملکرد آن به ازای تنظیمات مختلف است. در الگوریتم های ژنتیک سنتی، راه حل ها به عنوان رشته های باینری ۰ و ۱ نمایش داده می شوند. کدگذاری ها با توجه به نوع مسئله، می تواند باینری یا حقیقی باشند [۱۳].

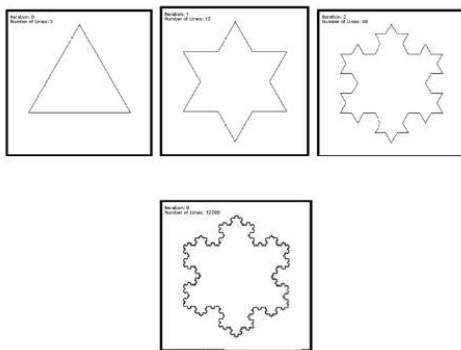
۲-۳- فراکتال

فراکتال یک شکل هندسی ناهموار یا تکه تکه است که می تواند به قسمتهایی تقسیم شود به طوری که هر کدام (حداقل به طور تقریبی) نسخه کوچک تری از کل شکل باشد. فراکتال یک شاخه از ریاضیات است که توسط بنویت مندلبرات پی ریزی شد. مندلبرات

کلمه فراکتال را برای اشیایی که در همه مقیاس ها خودشان بودند انتخاب کرد. او در کتاب The Fractal Geometry Of Nature که در سال ۱۹۸۲ منتشر شده است شاخه جدیدی از ریاضیات را معرفی نمود و اشیای ریاضی با خصوصیات خودسانی که معمولاً چیزهای عجیب یا هیولاهای ریاضی تلقی می شدند را به رسمیت شناخت [۱۸].

۲-۳-۱- منحنی برف دانه کخ

شکل (۲-۱) سه مرحله از ایجاد منحنی برف دانه کخ را نشان می دهد (این روند بر روی سه ضلع یک مثلث به صورت هم زمان پیاده سازی شده است). روند کار با یک قطعه آغاز می شود که دو نقطه را به هم وصل می کند ($n=0$). در مرحله دوم، هر قطعه از مولد با یک نسخه کوچک تر از خودش جایگزین می گردد و با بی نهایت تکرار این روند فراکتال کامل به دست می آید [۱۹].



شکل ۲-۱: تصاویری از مراحل به دست آمدن برف دانه کخ.

تعریف مولد یک فراکتال: یک منحنی است که تبدیل اولین مرحله را مشخص می کند. در هر مرحله هر قطعه از منحنی فراکتال در حال ایجاد، با یک مقیاس مناسب از مولد جایگزین می گردد. محیط و مساحت برف دانه کخ: در هر مرحله، هر قطعه از منحنی برف دانه کخ با یک قطعه منحنی مولد به طول $\frac{4}{3}$ طول اولیه قطعه جایگزین می شود.

۳- پیشینه تحقیق

این فصل پیشینه پژوهش رو مطرح می کند. در این فصل ابتدا به مرور کارهای انجام شده پرداخته می شود و سپس جدیدترین کارهایی که برای مقایسه نهایی استفاده می شود، به صورت مشروح بیان خواهند شد.

۳-۱- بررسی کلی تحقیقات پیشین

چن و همکاران در سال ۱۹۹۷ ترکیبی از شبکه‌ی عصبی هم‌گذار و پرسپترون برای تشخیص ارقام دست‌نویس در یک‌رشته عددی استفاده شده است. در این مقاله در مرحله استخراج ویژگی‌ها روش قاب‌بندی مورد استفاده قرار گرفته است [۲۵].

در پژوهش لیو و همکاران در سال ۲۰۰۲، آموزش فعال شبکه عصبی برای دسته‌بندی کاراکتر بکار رفته است. روش موردنظر بر اساس این ایده بیان شده است که برای انجام یک کار که احتیاج به دانش متخصص دارد، با داشتن k متخصص و با رأی مستقل و شیوه مناسب برای ترکیب آرا آن‌ها، به نتایج بهتری خواهیم رسید. در دسته‌بندی ایده این است که k دسته‌بند مختلف Φ_1, \dots, Φ_k برای تصمیم‌گیری در مورد $d \in c$ بکار گرفته می‌شود و سپس خروجی‌ها را با روش مناسبی باهم ترکیب می‌کند [۲۶].

۳-۲- بررسی مرتبط‌ترین تحقیقات جهت مقایسه با روش

پیشنهادهای

وایدا و همکاران در سال ۲۰۱۷، از روش عصبی فازی مبتنی بر شبکه عصبی سه لایه پرسپترون برای شناسایی حروف دست‌نویس استفاده شده است. این سیستم به دلیل توانایی غلبه بر تغییرات بااهمیت است. این پژوهش با استفاده از داده‌های MNIST کار خود را ارزیابی کرده و به دقت ۹۸٫۷ درصد دست یافته است [۴۵].

گودفلو و همکاران در سال ۲۰۱۳، روش مبتنی بر شبکه کانولوشنی ارائه کردند که با دقت ۹۸٫۳ درصد روی مجموعه داده

۴- روش پیشنهادی

هدف اصلی این پژوهش پیدا کردن یک راه‌حل جدید برای تشخیص کلمات از تصاویر دیجیتال با استفاده از الگوریتم ژنتیک و فراکتال دوبعدی می‌باشد. از کاربردهای تشخیص دست خط می‌توان به متن نوشته شده یا چاپ شده جهت تبدیل متن قابل ویرایش ماشین تایپ اشاره کرد. با استفاده از روش تشخیص می‌توان وسایل خودکار کتاب‌خوان برای نابینایان طراحی کرد.

۴-۱- کلیات روش پیشنهادی

همان‌طور که در شکل (۴-۱) نشان داده شده است برای پیاده‌سازی:

- ابتدا تصویر از ورودی دریافت می‌شود و روی آن پیش پردازش اولیه صورت می‌گیرد.

MNIST عمل می‌کند. این روش مبتنی بر لایه‌های درهم پیچش است و می‌تواند هم به صورت برخط و هم برون خط عملیات تشخیص را انجام دهد. پ این روش با مشکل زمان آموزش مواجه است [۵۲].

روش	هدف	الگوریتم	مزایا	معایب
گودفلو و همکاران [۵۲] ۲۰۱۳	حفظ دقت و افزایش سرعت	روش مبتنی بر شبکه کانولوشنی عمیق CNN	دقت بالا در کنار باتک فیلتر کامل	سرعت آموزش بسیار کم است و برای به روز رسانی نیاز به آموزش دوباره طولانی دارد که با ماهیت پویای مسئله همخوانی ندارد
صبور و همکاران [۵۳] ۲۰۱۷	افزایش دقت	شبکه عمیق CapsNet	دقت بالا و رفع ایرادات روش CNN	سرعت آموزش بسیار کم، عدم پوشش مناسب برای همه مجموعه داده
وایدا و همکاران [۲۵] ۲۰۱۷	افزایش دقت تشخیص با تحلیل فازی فضای مسئله	روش عصبی فازی مبتنی بر شبکه عصبی سه لایه پرسپترون	دقت بالا	نسبت دقت بدست آمده با معماری پیچیده استفاده شده همخوانی ندارد.

جدول ۳-۱: مقایسه روش‌های پیشین که از مجموعه داده MNIST استفاده کرده‌اند.

روش	هدف	الگوریتم	مزایا	معایب
فان و همکاران [۵۴] ۲۰۱۳	حفظ دقت و افزایش سرعت	استفاده از SIFT برای استخراج ویژگی	سرعت زیاد و حجم محاسبات پایین	تحلیل ناکافی ویژگی‌ها، دقت کم به نسبت روش‌های مشابه
چین و همکاران [۲۷] ۲۰۱۸	افزایش دقت تشخیص	شبکه عمیق CNN	دقت بالا	پیچیدگی مدل و زمان آموزش بسیار زیاد
لیو و همکاران [۵۵] ۲۰۱۹	افزایش دقت و اصلاح خودکار در زمان	شبکه عمیق بازگشتی LSTM جدید	دقت بالا و رویه اصلاح خودکار درون LSTM	زمان ساخت مدل بسیار زیاد و سخت افزار مورد نیاز برای اجرا در اختیار هر فردی نیست

جدول ۳-۲: مقایسه روش‌های پیشین که از مجموعه داده NEOCR استفاده کرده

- تصویر به بلوک‌های کوچک و مساوی تقسیم می‌شود (به کمک بعد فراکتال). این عمل باعث بالا رفتن سرعت تشخیص و سادگی کار می‌شود.
- الگوریتم ژنتیک مکان کاراکترها را شناسایی می‌کند
- بعد فراکتال کاراکترها را تشخیص می‌دهد و آن‌ها را استخراج می‌کند.

نویز گوسی موجود در تصاویر را کاهش داده و تصویر آماده استفاده برای مراحل بعد می‌شود.

۲-۲-۴- تغییر اندازه تصویر و بلوک‌بندی

پس از پیش‌پردازش، تصویرها استاندارد سازی و بلوک‌بندی می‌شوند. نکته مهم این است که در آزمایش‌های مبتنی بر مجموعه داده تصاویر یک اندازه هستند اما در آزمایش‌های موردی یک مرحله جداگانه برای یکسان‌سازی اندازه تصاویر نیز انجام می‌گیرد. در این مرحله، تصویر به ابعاد مستطیلی که معادل میانگین اندازه کل مجموعه داده است تغییر اندازه داده می‌شوند. سپس یک بلوک‌بندی اولیه به منظور موازی سازی اجرای استخراج ویژگی انجام می‌شود (هدف این است که سرعت اعمال استخراج ویژگی افزایش یابد).

۲-۲-۳- استخراج ویژگی اولیه

پس از آماده‌سازی در دو مرحله قبل، تصویر جدید را با استفاده از فراکتال مبتنی بر یافت قطعه‌بندی کرده و نقاط موردنظر (متن) در تصویر، علامت‌گذاری می‌شود. در این پژوهش روش فراکتال دوبعدی به این صورت است که ابتدا با روش DCD بعد فراکتال تعیین می‌شود. در این روش بعد توسعه داده شده توسط روش MD تعریف می‌شود. برای اندازه‌گیری سطح به دست آمده در هر سطح از مرکز توسعه داده شده وابسته که به شعاع r می‌باشد برای هر مجموعه E این سطح توسعه داده شده (سطحی که توسط فراکتال توسعه داده می‌شود) از رابطه (۱-۴) محاسبه می‌شود.

$$DCD(d) = \lim_{d \rightarrow 0} (n - \frac{\log V(d)}{\log(d)}) \quad (1-4)$$

که در این رابطه:

DCD: بعد توسعه داده شده مجموعه E

D: ماکزیمم شعاع توسعه داده شده

N: بعد فضای تصویر

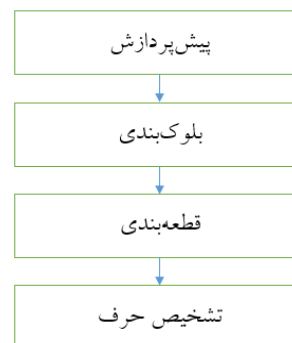
$$V = \{x \in \mathbb{R}^n: \exists a \in E / dis(a, x) \leq r\} \quad (2-4)$$

Dis(a,x): فاصله بین نقطه a, x

V(d): حجم n بعدی مجموعه d

این مرحله آن قدر تکرار می‌شود تا کمینه شعاع به دست آید. سپس شکل با شعاع بدست آمده توسعه داده شده، داده می‌شود.

از تصویر به دست آمده از مرحله پیش‌پردازش با استفاده از روش فراکتال (با توجه به این که فراکتال دوبعدی موردنظر است)



شکل ۱-۴: روند کلی الگوریتم پیشنهادی

۲-۲-۴- شرح روش پیشنهادی

در این بخش اجزای روش پیشنهادی به صورت مشروح بیان می‌شود. بلوک دیاگرام روش پیشنهادی به صورت جزئی در شکل (۲-۴) آمده است.

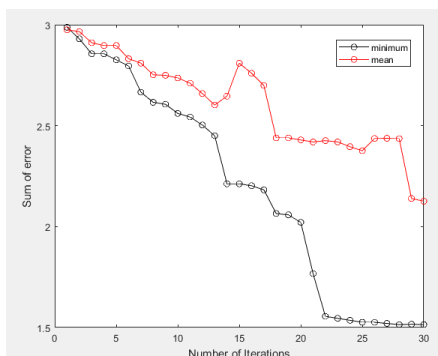


شکل ۲-۴: فلوچارت روش پیشنهادی

۲-۲-۱- پیش‌پردازش

پس از ورودی تصویر پیش‌پردازش اولیه روی آن صورت می‌گیرد. هدف از این فاز، استانداردسازی تصویر ورودی است. به همین منظور، ابتدا تصویر به حوزه‌ی خاکستری منتقل می‌شود سپس با استفاده از یک فیلتر میانه و یک فیلتر گوسی، نویز فلفل و نمک و

در روابط بالا منظور از True Positive یا TP، مواردی هستند که در مجموعه داده به عنوان کلاس مثبت (مثلا حرف K) مشخص شده‌اند و سیستم به درستی آن‌ها را تشخیص داده است (به درستی حرف K تشخیص داده است)، به همین شکل False Positive با FP، مواردی هستند که در مجموعه داده به عنوان کلاس مثبت (مثلا کاملا حرف K) مشخص شده‌اند و سیستم به نادرست آن‌ها را تشخیص داده است (مثلا حرف H تشخیص داده است). از سوی دیگر True Negative یا TN، مواردی هستند که در مجموعه داده به عنوان منفی (مثلا هر کلاسی غیر از حرف K) مشخص شده‌اند و سیستم به درستی آن‌ها را منفی تشخیص داده است. در نهایت False Negative یا FN، مواردی هستند که در مجموعه داده به عنوان منفی (مثلا هر کلاسی غیر از حرف K) مشخص شده‌اند و سیستم به نادرست آن‌ها را مثبت (حرف K) تشخیص داده است.



شکل ۵-۱: همگرایی الگوریتم ژنتیک در مرحله قطعه‌بندی

ماتریس‌هایی با ابعاد ۱۶ در ۱۶ از تصویر تولید می‌شود و در این تصاویر هر جا پیکسلی ۱ با ۱ جمع می‌شود و به این ترتیب تعداد ۱ها در آن قسمت تشخیص تشخیص داده می‌شود. با داشتن تعداد کل پیکسل‌های انتخاب شده عدد D را محاسبه کرده با توجه به این شعاع و روش DCD تصویر توسعه داده شده داده می‌شود. در این بخش از تابع استاندارد استخراج ویژگی‌های فراکتال استفاده شده است. کد تابع آن در ادامه آمده است.

۳-۴- معیار ارزیابی سیستم

نتایج نهایی سیستم بر مبنای پارامترهای صحت (رابطه ۴-۳)، دقت (رابطه ۴-۴)، بازیابی (رابطه ۴-۵) و معیار F (رابطه ۴-۶) گزارش می‌شود.

$$Precision = \frac{\text{number of True Positive}}{\text{number of True Positive} + \text{number of False Positive}} \quad (4-3)$$

$$Accuracy = \frac{\text{number of True Positive} + \text{number of True Negative}}{\text{Total population}} \quad (4-4)$$

$$Recall = \frac{\text{number of True Positive}}{\text{number of True Positive} + \text{number of False Negative}} \quad (4-5)$$

$$F \text{ measure} = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (4-6)$$

۵- تحلیل نتایج

۵-۱- همگرایی الگوریتم ژنتیک

در ابتدا شکل (۵-۱) نشان دهنده یک بار اجرای الگوریتم ژنتیک و همگرایی آن در فاز قطعه‌بندی است. ورودی این مرحله، تصویری از مجموعه داده NEOCR است (شکل (۴-۵)). این شکل نشان می‌دهد که روش پیشنهادی، در مرحله بهینه‌سازی خود در تکرار ۱۲۸ام به همگرایی رسیده است.

همچنین شکل (۵-۲) نشان دهنده یک بار اجرای الگوریتم ژنتیک و همگرایی آن در فاز تشخیص حروف است. این شکل نشان می‌دهد که روش پیشنهادی، در مرحله بهینه‌سازی خود در تکرار ۱۲۳ام به همگرایی رسیده است. در این مرحله نیز ورودی برای تصویر نشان داده شده در شکل (۴-۵) است. برای هر دو شکل (۵-۱) و (۵-۲) جدول (۵-۱) ملاک پارامترهای الگوریتم ژنتیک است.

4 6 8 7 5 3 2 2

پ) خروجی واقعی

شکل (۳-۵) نمونه تشخیص یک تصویر ورودی به نگاشت خروجی برای مجموعه داده MNIST

شکل (۴-۵) نمونه تشخیص یک تصویر ورودی به نگاشت خروجی را برای روش پیشنهادی نشان می‌دهد. شکل (۴-۵ الف) نمونه ورودی برای مجموعه داده NEOCR و شکل (۴-۵ ب) خروجی روش پیشنهادی و شکل (۴-۵ پ) خروجی واقعی برای داده ورودی است. نتایج نشان می‌دهد که تمامی حروف به درستی تشخیص داده شده است.



الف) نمونه ورودی برای مجموعه داده NEOCR

FINSTERE GASSL

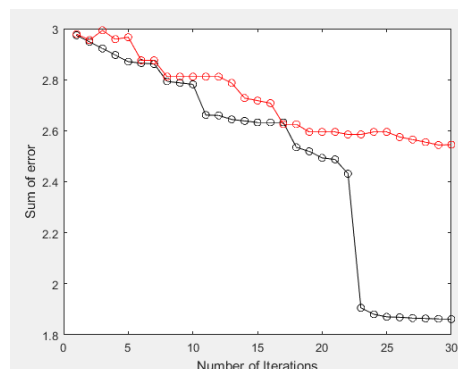
ب) خروجی روش پیشنهادی

FINSTERE GASSL

پ) خروجی واقعی

شکل (۴-۵) نمونه تشخیص یک تصویر ورودی به نگاشت خروجی برای مجموعه داده NEOCR

در جدول (۱-۵) مقایسه خروجی روش پیشنهادی برای مجموعه داده MNIST و روش‌های مشابه را نشان می‌دهد.

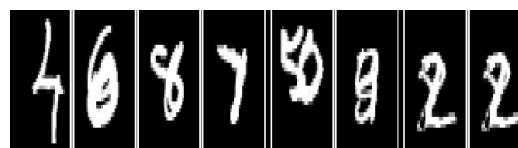


شکل ۲-۵: همگرایی الگوریتم ژنتیک در کمینه‌سازی اختلاف حرف ورودی با حروف دیگر

۲-۵- ارزیابی نتایج

در این بخش نتایج ارزیابی روش پیشنهادی بیان شده است. ابتدا نتایج منفرد و سپس نتایج مقایسه روی مجموعه داده‌های شرح داده شده در فصل چهارم بیان می‌شود.

شکل (۳-۵) نمونه تشخیص یک تصویر ورودی به نگاشت خروجی را برای روش پیشنهادی نشان می‌دهد. شکل (۳-۵ الف) نمونه ورودی برای مجموعه داده MNIST و شکل (۳-۵ ب) خروجی روش پیشنهادی و شکل (۳-۵ پ) خروجی واقعی برای مجموعه داده ورودی است. نتایج نشان می‌دهد که برای عدد سه ورودی، خروجی اشتباه و باقی حالات صحیح است.



الف) نمونه ورودی برای مجموعه داده MNIST

4 6 8 7 5 8 2 2

ب) خروجی روش پیشنهادی

روش	دقت	صحت	بازایی	معیار F
روش پیشنهادی	۹۹/۳ درصد	۹۸/۷ درصد	۹۹/۶ درصد	۹۹/۱۴ درصد
گودفلو و همکاران [۵۲] ۲۰۱۳ در	۹۸/۱ درصد	-	-	۹۸/۵۲ درصد
صبور و همکاران [۵۳] ۲۰۱۷ در	۹۹/۲۳ درصد	-	-	۹۸/۸۹ درصد
وایدا و همکاران [۴۵] ۲۰۱۷ در	۹۸/۷۰ درصد	-	-	۹۸/۲۳ درصد

جدول (۱-۵) خروجی روش پیشنهادی برای مجموعه داده MNIST

با مشاهده جدول (۱-۵) می توان مشاهده کرد که روش پیشنهادی بهتر از سه روش دیگر مبتنی بر یادگیری عمیق عمل می کند. اما مشکل روش پیشنهادی این است که چند فاز متوالی و طولانی دارد در مقابل دو روش مبتنی بر یادگیری عمیق بدون نیاز به پیش پردازش خاصی مسئله را حل می کنند. از سوی دیگر مزیت روش پیشنهادی توانایی بازنمایی ویژگی های استخراج شده است درحالی که همه عملیات روش های پیشین به صورت جعبه سیاه است و هیچ توصیفی از نحوه عملکرد نیست.

در مقایسه با روش نیز، روش پیشنهادی علاوه بر نشان دادن دقت بهتر، ثابت می کند تحلیل فراکتال نتایج بهتری به نسبت تحلیل فازی به دنبال دارد.

در جدول (۵-۲) مقایسه خروجی روش پیشنهادی برای مجموعه داده NEOCR و روش های مشابه را نشان می دهد.

روش	دقت	صحت	بازایی	معیار F
روش پیشنهادی	۷۸/۷۱ درصد	۷۶/۴ درصد	۸۵/۷۲ درصد	۸۰/۱ درصد
فان و همکاران [۵۴] ۲۰۱۳ در	۷۳/۷ درصد	-	-	۶۹/۴۴ درصد
چین و همکاران [۳۷] ۲۰۱۸ در	۷۶/۲۹ درصد	-	-	۷۸/۹۲ درصد
لیو و همکاران [۵۵] ۲۰۱۹ در	-	۸۴/۵۰ درصد	۷۷/۱ درصد	۸۰/۰۶ درصد

جدول (۲-۵): خروجی روش پیشنهادی برای مجموعه داده

NEOCR

با مشاهده جدول (۲-۵) می توان مشاهده کرد که روش پیشنهادی بهتر از روش مبتنی بر استخراج ویژگی SIFT عمل می کند. علت این موضوع توصیف بهتر روش پیشنهادی توسط الگوریتم فراکتال است. مشکل روش پیشنهادی در قیاس با این روش این است که چند فاز متوالی و طولانی دارد در مقابل روش [۵۴] بسیار سریع است. البته لازم به ذکر است که هدف تشخیص برخط نبوده است.

در مقایسه با روش [۵۵] روش پیشنهادی در دو معیار بازایی و معیار F بهتر عمل کرده است که نشان می دهد روش پیشنهادی نرخ پوشش بهتری روی همه حروف دارد در حالی که به صورت میانگین نیز متعادل تر از روش [۵۵] عمل می کند. از سوی دیگر نرخ صحت در روش [۵۵] بیشتر است که نشان از دقت تشخیص برای حروف و نرخ پایین FP دارد. البته روش [۵۵] از یک معماری بسیار پیشرفته مبتنی بر LSTM استفاده می کند که مرتبه زمانی را به شدت افزایش می دهد و به هیچ وجه معیار قیاس با روش پیشنهادی که از معماری غیرعمیق استفاده می کند نیست، همچنین روش [۵۵] توصیف پذیری مناسبی برای بردار ویژگی ندارد. همچنین روش پیشنهادی بسیار بهتر از روش [۳۷] عمل کرده است، جایی که این روش مبتنی بر سیستم CNN گرچه پیچیدگی بیشتری دارد اما به دلیل تنوع ساختار در تصاویر ورودی، توانایی یادگیری همه حالات را به وضوح نداشته است.

به صورت کلی و براساس نتایج حاصل از جداول (۱-۵) و (۲-۵) مهم ترین دلیل نتایج حاصل از روش پیشنهادی را می توان دو مورد در نظر گرفت. اولین مورد این است که مسئله (داده ها) ورودی ذاتا توزیعی دارد که توسط الگوریتم فراکتال قابل توصیف است. یعنی تصاویر و اجزای درون آن به خوبی توسط الگوریتم فراکتال توصیف می شوند که این از خواص نوشتار و خط است که در فصل مبانی به توصیف آن پرداخته شده است.

مورد دیگر این است که برخلاف روش های معمول روش پیشنهادی، به چند سطح قطعه بندی سعی در تشخیص دقیق محل حروف دارد و در نتیجه تاثیرگذاری همسایگی های غیر مرتبط با هر جزء تصویر را حذف می کند. این موضوع کمک می کند که بردار توصیفی برای هر حرف کمترین اطلاعات اضافی یا نویز را مرحله تشخیص نهایی داشته باشد.

۶- نتیجه گیری

اگر چه شناخت حروف در حال تبدیل شدن به بخش جدایی ناپذیر از هر سیستم OCR در زبان ها است، اما هنوز هم در زبان های مختلف توسعه نیافته است. می توان مشکلات سیستم های قبلی را به صورت کلی بیان نمود: شناسایی دقیق قلم در تصویر، شناسایی دقیق مکان کاراکتر در تصویر، تشخیص دقیق کارکترهای موجود در تصویر، سرعت اجرای استخراج کاراکتر از تصویر. در این سیستم

با پیاده‌سازی فرکتال دو بعدی و الگوریتم ژنتیک دقت بالا در پردازش متن از تصویر در مقایسه با روش‌های پیشین مورد ارزیابی نشان داده است. نوآوری روش پیشنهادی در استفاده از الگوریتم ژنتیک و فراکتال در تقسیم‌بندی شکل و نیز تشخیص کاراکتر مربوطه است. مشاهده شده است که با استفاده از فرکتال مناسب می‌توان حروف تصاویر دست‌نویس را به خوبی تشخیص داد که البته الگوریتم ژنتیک در بهبود آن موثر است. نتایج شبیه‌سازی بیان‌گر این مسئله بوده است. در این تحقیق تشخیص حروف و کلمات در تصاویر رنگی نیز انجام شده است که توانایی روش پیشنهادی است. ویژگی‌های روش پیشنهادی تشخیص و جداسازی متن از تصویر، تشخیص کلمات و جداسازی کلمات از تصویر است که توسط توابع و ویژگی‌های پردازش تصویر صورت گرفت و از الگوریتم ژنتیک و فراکتال دو بعدی نیز برای تشخیص کلمات و حروف استفاده شد. تابع هدف بر اساس مقدار اختلاف بین حروف و کلمات پایگاه داده با حروف و کلمات جداسازی شده مدل‌سازی شده است.

با توجه به نتایج می‌توان اینگونه بیان کرد که استفاده از تحلیل فراکتال، برای تشخیص حروف بهتر از روش‌های مبتنی بر الگوریتم‌های فازی عمل می‌کند، علت این موضوع ساختار ذاتی حروف دست‌نویس است که بیشتر مبتنی بر توزیع فراکتال‌ها هستند تا توابع فازی؛ البته روش پیشنهادی دقت ۱۰۰ درصد ندارد و شاید بهتر باشد از ترکیب تحلیل ویژگی فراکتال در کنار دیگر روش‌های تحلیل استفاده کرد.

مورد دیگر این است که روش پیشنهادی، از دو فاز اصلی تشخیص قطعه کلمات و سپس قطعه‌بندی حروف بهره می‌گیرد که این موضوع باعث شده است که پوشش تشخیص حروف مقدار بهتری نسبت به روش‌های معمول شود. از سوی دیگر روش‌های مبتنی بر یادگیری عمیق که به صورت یک مرحله‌ای انجام می‌شوند، توانایی تعمیم کمتری را به دلیل پیچیدگی ورودی مسئله دارند.

دیگر موردی که می‌توان از ارزیابی‌های نتیجه گرفت، تعادل مناسب روش پیشنهادی است. دیده شد که روش پیشنهادی، در مجموعه‌داده تشخیص حروف صحت کمتری نسبت به پژوهشی مبتنی بر یادگیری عمیق نشان داد. اما پوشش بهتر و در نهایت میانگین هارمونیک (معیار F) بالاتر، موجب شد که در ارزیابی در

جایگاه بهتری قرار بگیرد.

اشکال اصلی روش پیشنهادی، زمانبر بودن دو فاز مربوط به قطعه‌بندی است. اجرای دو مرحله‌ای الگوریتم ژنتیک به منظور جداسازی کلمات و حروف، سربرار زیادی در مجموعه‌داده‌های بزرگ ایجاد می‌کند. برای مثال در مجموعه‌داده MNIST که ۱۰ هزار داده آزمایش دارد، متوسط سی ثانیه برای تشخیص هر تصویر نیاز است. نتیجه اینکه ۵۰۰۰ دقیقه معادل ۸۳ ساعت برای آزمایش همه حالا در این مجموعه‌داده زمان لازم است. در حالی که در روش‌های سبکتر، این زمان تا ۵ ثانیه برای هر تصویر کمتر است. موضوع دیگر در مورد روش پیشنهادی عملکرد نامناسب در تصاویر با موارد زاید (نویز ذاتی) است که دقت را در مجموعه داده NEOCR کاهش داده است. این مجموعه‌داده که تصاویر عمدتاً برگرفته از دنیای واقعی را دارد، همراه با نویز ذاتی در ۲۵ درصد تصاویر است که کاهش دقت را در روش پیشنهادی به همراه دارد. در نهایت روش پیشنهادی به دلیل نوع تحلیل ویژگی، توانایی مدیریت چرخش‌ها را دارد که این موضوع تنها در روش مبتنی بر CapsNet رعایت شده است و از نقاط ضعف روش‌های مبتنی بر CNN است.

۶-۱- کارهای آتی

در این پیشنهادی‌هایی برای پژوهشگران مبتنی بر تحقیق انجام شده ارائه می‌شود. این پیشنهادها شامل بهبودهایی در فاز استخراج ویژگی، قطعه‌بندی و تشخیص نهایی است.

- استفاده از شبکه‌های عمیق در ترکیب با استخراج ویژگی فراکتال بافت
- استفاده از روش‌های بهینه‌سازی فرااکتشافی در ترکیب با فراکتال بافت
- استفاده از ویژگی‌های حوزه فرکانس در کنار ویژگی‌های مکانی برای تشخیص بهتر حروف
- استفاده از فازهای پیش پردازش مبتنی بر شکل‌شناسی برای اصلاح شکل حروف و اعداد دست‌نویس
- استفاده از روش‌های تعقیب دست‌خط برای تشخیص بهینه کلمات

- بررسی روش پیشنهادی برای تشخیص حروف در مجموعه داده‌های شبه عربی

منابع:

- [۱] دهقانی، علی "بازشناسی حروف مجزای دست نویس فارسی با استفاده از مدل پنهان مارکف با چگالی پیوسته و . ایده ترکیب چند سیستم خبره"، پایان نامه کارشناسی ارشد، دانشگاه شیراز، دانشکده تحصیلات تکمیلی، ۱۳۷۹.
- [2] Chaudhuri A, Mandaviya K, Badelia P, Ghosh SK. Optical Character Recognition Systems for English Language. In Optical Character Recognition Systems for Different Languages with Soft Computing 2017 (pp. 85-107). Springer, Cham.
- [3] Rahman AF, Fairhurst MC. Multiple classifier decision combination strategies for character recognition: A review. Document Analysis and Recognition. 2003 Jul 1;5(4):166-94.
- [4] Naseer A, Zafar K. Comparative analysis of raw images and meta feature based Urdu OCR using CNN and LSTM. Int J Adv Comput Sci Appl. 2018 Jan 1;9(1):419-24.
- [5] Jain M. *Unconstrained Arabic & Urdu Text Recognition using Deep CNN-RNN Hybrid Networks* (Doctoral dissertation, International Institute of Information Technology Hyderabad).
- [6] Ackley HS, inventor; Hand Held Products Inc, assignee. Optical character recognition systems and methods. United States patent application US 15/793,407. 2019 Apr 25.
- [7] Nazif A, 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. 2001 May;31(2).
- [8] Sonkusare M, Sahu N. A survey on handwritten character recognition (HCR) techniques for English alphabets. Adv Vis Comput Int J. 2016 Mar;3(1):1-2.
- [9] Ackley HS, inventor; Hand Held Products Inc, assignee. Optical character recognition systems and methods. United States patent application US 15/793,407. 2019 Apr 25.
- [10] Jain A, Nandakumar K, Ross A. Score normalization in multimodal biometric systems. Pattern recognition. 2005 Dec 1;38(12):2270-85.
- [11] Witten IH, Frank E, Hall MA, Pal CJ. Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann; 2016 Oct 1.
- [12] Doermann D. Handbook of document image processing and recognition. Tombre K, editor. New York Incorporated: Springer; 2014 May 21.
- [13] Grefenstette JJ, editor. Genetic algorithms and their applications: proceedings of the second international conference on genetic algorithms. Psychology Press; 2013 Aug 21.
- [14] Dasgupta D, Michalewicz Z, editors. Evolutionary algorithms in engineering applications. Springer Science & Business Media; 2013 Jun 29.
- [15] Simon D. Evolutionary optimization algorithms. John Wiley & Sons; 2013 Jun 13.
- [16] Bäck T, Fogel DB, Michalewicz Z, editors. Evolutionary computation 1: Basic algorithms and operators. CRC press; 2018 Oct 3.
- [17] Črepinšek M, Liu SH, Mernik M. Exploration and exploitation in evolutionary algorithms: A survey. ACM Computing Surveys (CSUR). 2013 Jun 1;45(3):35.
- [18] Varma M, Garg R. Locally invariant fractal features for statistical texture classification. In 2007 IEEE 11th international conference on computer vision 2007 Oct 14 (pp. 1-8). IEEE.
- [19] Chen SS, Keller JM, Crownover RM. On the calculation of fractal features from images. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1993 Oct;15(10):1087-90.

- [20] Zhu F, Cheng Q, Xue S, Li C, Hartley W, Wu C, Tian T. Influence of natural regeneration on fractal features of residue microaggregates in bauxite residue disposal areas. *Land degradation & development*. 2018 Jan;29(1):138-49.
- [21] Li X, Li Z, Wang E, Liang Y, Li B, Chen P, Liu Y. Pattern recognition of mine microseismic and blasting events based on wave fractal features. *Fractals*. 2018 Jun 11;26(03):1850029.
- [22] Xu G, Li Z, Li P. Fractal features of soil particle-size distribution and total soil nitrogen distribution in a typical watershed in the source area of the middle Dan River, China. *Catena*. 2013 Feb 1;101:17-23.
- [23] Zhiznyakov AL, Privezentsev DG, Zakharov AA. Using fractal features of digital images for the detection of surface defects. *Pattern Recognition and Image Analysis*. 2015 Jan 1;25(1):122-31.
- [24] Du JX, Zhai CM, Wang QP. Recognition of plant leaf image based on fractal dimension features. *Neurocomputing*. 2013 Sep 20;116:150-6.
- [25] Chen K, Wang L, Chi H. Methods of combining multiple classifiers with different features and their applications to text-independent speaker identification. *International Journal of Pattern Recognition and Artificial Intelligence*. 1997 May;11(03):417-45.
- [26] Liu Y, Yao X, Zhao Q, Higuchi T. An experimental comparison of ensemble learning methods on decision boundaries. In *Proceedings of the 2002 International Joint Conference on Neural Networks. IJCNN'02 (Cat. No. 02CH37290)* 2002 May 12 (Vol. 1, pp. 221-226). IEEE.
- [27] Kesorn K, Phawapoothayanchai P. Optical Character Recognition (OCR) enhancement using an approximate string matching technique. *Engineering and Applied Science Research*. 2018 Dec 19;45(4):282-9.
- [28] Wemhoener D, Yalniz IZ, Manmatha R. Creating an improved version using noisy OCR from multiple editions. In *2013 12th International Conference on Document Analysis and Recognition* 2013 Aug 25 (pp. 160-164). IEEE.
- [29] Hajiannezhad AA, Mozaffari S. Fractal and multi-fractal dimensions for farsi/arabic font type and size recognition. In *2011 7th Iranian Conference on Machine Vision and Image Processing* 2011 Nov 16 (pp. 1-4). IEEE.
- [30] Moussa SB, Zahour A, Benabdelhafid A, Alimi AM. Fractal-based system for Arabic/Latin, printed/handwritten script identification. In *2008 19th International Conference on Pattern Recognition* 2008 Dec 8 (pp. 1-4). IEEE.
- [31] Lopes R, Betrouni N. Fractal and multifractal analysis: a review. *Medical image analysis*. 2009 Aug 1;13(4):634-49.
- [32] Hajiannezhad A, Mozaffari S. Font recognition using variogram fractal dimension. In *20th Iranian Conference on Electrical Engineering (ICEE2012)* 2012 May 15 (pp. 634-639). IEEE.
- [33] Moussa SB, Zahour A, Benabdelhafid A, Alimi AM. New features using fractal multi-dimensions for generalized Arabic font recognition. *Pattern Recognition Letters*. 2010 Apr 1;31(5):361-71.
- [34] Moussa SB, Zahour A, Benabdelhafid A, Alimi AM. New features using fractal multi-dimensions for generalized Arabic font recognition. *Pattern Recognition Letters*. 2010 Apr 1;31(5):361-71.
- [35] Prasad BK. Moment of Inertia-Based Approach to Recognize Arabic Handwritten Numerals. In *Innovations in Electronics and Communication Engineering* 2019 (pp. 253-260). Springer, Singapore.
- [36] Prasad BK, Sanyal G. Multiple hidden Markov model post processed with support vector machine to recognize English handwritten numerals. *International Journal on Artificial Intelligence Tools*. 2018 Aug 14;27(05):1850019.
- [37] Jain A, Sharma BK. Analysis of Activation Functions for Convolutional Neural Network based MNIST Handwritten Character Recognition. *International Journal of Advanced Studies of Scientific Research*. 2018;3(9).
- [38] Amrhein C, Clematide S. Supervised OCR Error Detection and Correction Using Statistical and Neural Machine Translation Methods. *Journal for Language Technology and Computational Linguistics (JLCL)*. 2018;33(1):49-76.

- [39] Kaur EK, Banga VK. Number plate recognition using OCR technique. *International Journal of Research in Engineering and Technology*. 2013 Sep;2(09):286-90.
- [40] Qadri MT, Asif M. Automatic number plate recognition system for vehicle identification using optical character recognition. In *2009 International Conference on Education Technology and Computer* 2009 Apr 17 (pp. 335-338). IEEE.
- [41] Du S, Ibrahim M, Shehata M, Badawy W. Automatic license plate recognition (ALPR): A state-of-the-art review. *IEEE Transactions on circuits and systems for video technology*. 2012 Jun 7;23(2):311-25.
- [42] Kunwar R, Ramakrishnan AG. Online handwriting recognition of tamil script using fractal geometry. In *2011 International Conference on Document Analysis and Recognition* 2011 Sep 18 (pp. 1389-1393). IEEE.
- [43] Sadri J, Yeganehzad MR, Saghi J. A novel comprehensive database for offline Persian handwriting recognition. *Pattern Recognition*. 2016 Dec 1;60:378-93.
- [44] Moradi V, Razzazi F, Behrad A. Recognition of Handwritten Persian Two-digit Numerals Using a Novel Hybrid SVM/HMM algorithm. *Majlesi Journal of Electrical Engineering*. 2016 Sep 26;10(3).
- [45] Vaidya A, Malathi D, Kumaran N, Mishra S, Jayaseeli JD. Fuzzy Filtered Neural Network Approach towards Handwritten Numeral Recognition. *International Journal of Applied Engineering Research*. 2017;12(18):7115-21.
- [46] Shariatmadari S, Al-Máadeed S, Akbari Y, Rida I, Emadi S. Off-line Persian Signature Verification using Wavelet-based Fractal Dimension and One-class Gaussian Process. In *2018 NASA/ESA Conference on Adaptive Hardware and Systems (AHS)* 2018 Aug 6 (pp. 168-173). IEEE.
- [47] Subrahmanyam MS, Kumar VV, Reddy BE. A Robust Zonal Fractal Dimension Method for the Recognition of Handwritten Telugu Digits. *International Journal of Image, Graphics & Signal Processing*. 2018 Sep 1;10(9).
- [48] Kia MM, Alzubi JA, Gheisari M, Zhang X, Rahimi M, Qin Y. A novel method for recognition of persian alphabet by using fuzzy neural network. *IEEE Access*. 2018 Nov 12;6:77265-71.
- [49] Hadidi G, Delavari H. Persian Handwritten Words Detection Based on Features Extraction and Fuzzy Algorithm. *Electrical and Electronics: An International Journal*. 2015;4(2):93-104.
- [50] Jlaiel MB, Kanoun S, ALIMI AM, Mullot R. Three decision levels strategy for Arabic and Latin texts differentiation in printed and handwritten natures. In *Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)* 2007 Sep 23 (Vol. 2, pp. 1103-1107). IEEE.
- [51] Sigari MH, Pourshahabi MR, Pourreza HR. Offline handwritten signature identification and verification using multi-resolution gabor wavelet. *International Journal of Biometrics and Bioinformatics (IJBB)*. 2011 Oct;5(4):234-48.
- [52] Goodfellow IJ, Bulatov Y, Ibarz J, Arnoud S, Shet V. Multi-digit number recognition from street view imagery using deep convolutional neural networks. *arXiv preprint arXiv:1312.6082*. 2013 Dec 20.
- [53] Sabour S, Frosst N, Hinton GE. Dynamic routing between capsules. In *Advances in neural information processing systems* 2017 (pp. 3856-3866).
- [54] Quy Phan T, Shivakumara P, Tian S, Lim Tan C. Recognizing text with perspective distortion in natural scenes. In *Proceedings of the IEEE International Conference on Computer Vision* 2013 (pp. 569-576).
- [55] Liu Y, Jin L, Zhang S, Luo C, Zhang S. Curved scene text detection via transverse and longitudinal sequence connection. *Pattern Recognition*. 2019 Jun 1;90:337-45.
- [56] Cohen G, Afshar S, Tapson J, van Schaik A. EMNIST: an extension of MNIST to handwritten letters. *arXiv preprint arXiv:1702.05373*. 2017 Feb 17
- [57] Nagy R, Dicker A, Meyer-Wegener K. NEOCR: A configurable dataset for natural image text recognition. In *International Workshop on Camera-Based Document Analysis and Recognition* 2011 Sep 22 (pp. 150-163). Springer, Berlin, Heidelberg.