

سیستم کپچای فارسی برای جلوگیری از ثبت نام خودکار رباتهای نرم‌افزاری در صفحات وب

مهدی بهلول^۱، محمد ملک‌زاده^۲،

^۱ مربی، گروه کامپیوتر، دانشکده مهندسی، دانشگاه شهید چمران اهواز

اهواز، ایران

bohlool@scu.ac.ir

^۲ دانشجوی کارشناسی، گروه کامپیوتر، دانشکده مهندسی، دانشگاه شهید چمران اهواز

اهواز، ایران

malekzadeh@scu.ac.ir

چکیده

امروزه آهنگ دولت الکترونیک در جای‌جای کشور ما به گوش می‌رسد و اکثر کارهای روزمره مردم در چند سال آینده توسط اینترنت صورت می‌گیرد. اما همواره در کنار پیشرفت تکنولوژی، کسانی هستند که سعی دارند تا به نفع خود از این پیشرفت‌ها سوءاستفاده کنند. بعنوان مثال این افراد با استفاده از برنامه‌هایی معروف به رباتهای وبگرد، که به مانند یک انسان فرمهای اینترنتی را تکمیل می‌کنند، اقدام به ثبت اطلاعات غلط و یا اطلاعات زیاد می‌کنند که این کار باعث هدر رفتن منابع و یا اختلال در کار سایتها اطلاعاتی می‌شود. برای بالا بردن صحت داده‌ها و تشخیص تفاوت میان انسان و ربات نرم‌افزاری از روشی موسوم به کپچا^۱ استفاده می‌شود. در نمونه‌ای از این روش با نمایش عکسی حاوی کلمه‌ای قابل خواندن برای انسان و غیر قابل تشخیص برای یک سیستم اتوماتیک، از تکمیل فرم توسط سیستمهای خودکار جلوگیری می‌کند. در این مقاله سعی شده یک سیستم کپچای فارسی با قابلیت اطمینان بالا ارائه گردد. تکنیکهای پردازش تصویر برای ساخت یک تصویر مغشوش بکار برده شده و با استفاده از نرم‌افزارهای تشخیص متن، صحت عملکرد این روش بررسی شده است. نتایج بدست آمده نشان می‌دهند که تصاویر تولید شده توسط انسان قابل خواندن هستند حال آنکه یک نرم‌افزار تشخیص متن در اکثر موارد قادر به تشخیص درست آنها نمی‌باشد.

کلمات کلیدی

کپچا، پردازش تصویر، تست تورینگ، تشخیص متن، امنیت شبکه، وب.

^۱ CAPTCHA (Completely Automatic Public Turing test to tell Computer and Human Apart)

1- مقدمه

نیز ارائه شده‌اند که بیشتر بر ویژگی‌های ساختاری حروف فارسی تاکید دارند تا تکنیک‌های پردازش تصویر^۴.

در این مقاله سعی شده با بهره‌گیری از ویژگی‌های زبان فارسی و نیز دستکاری تصاویر بوسیله تکنیک‌های پردازش تصویر، یک کپچای فارسی با قابلیت اطمینان بالا ارائه شود. یک تصویر شامل یک کلمه‌ی فارسی ایجاد می‌شود و پس از اعمال چند تابع پردازشی بر روی آن به کاربر فارسی زبان نشان داده می‌شود و از وی خواسته می‌شود تا آنچه را می‌بیند در کادر خاصی تایپ کند. نتیجه بدست آمده از این روش توسط نرم افزارهای تشخیص حروف مورد آزمایش قرار گرفته است و نتایج مطلوبی بدست آمده است.

ادامه مقاله بدین صورت سازماندهی شده است. در بخش دوم ویژگی‌های کپچا و انواع آن را بررسی می‌کنیم و مروری بر کارهای انجام شده خواهیم داشت. در بخش سوم خواص زبان فارسی که برای ایجاد یک کپچای قدرتمند به کار می‌آید مورد بررسی قرار گرفته است. بخش چهارم به ارائه روش پیشنهادی، نحوه ایجاد تصویر و تولید خرابی در این تصاویر اختصاص دارد. و در آخر نیز نتیجه گیری نهایی آمده است.

2- ویژگی‌های کپچا و انواع آن

کپچا تستی است کاملاً خودکار که به ما می‌گوید کاربر آن انسان است یا یک برنامه کامپیوتری؛ موفقیت این تست به دو عامل بستگی دارد، اول اینکه تا چه حد برای انسان آسان است و دوم اینکه تا چه حد برای برنامه‌های کامپیوتری مشکل است [1]. در تعریف کپچا نوعی تناقض جالب وجود دارد: ما تست کامپیوتری می‌نویسیم که هیچ برنامه کامپیوتری دیگری نتواند آن را بگذراند. از طرفی کپچا یک مسئله "برنده برنده" است چراکه اگر کسی نتواند آن را شکست دهد پس ما یک روش خوب جهت تشخیص انسان از کامپیوتر در اختیار داریم و اگر هم کسی بتواند آن را شکست دهد پس او یک مسئله سخت هوش مصنوعی را حل کرده و می‌تواند با ایده گرفتن از آن مسائل حل نشده دیگری در هوش مصنوعی را حل کرد [1].

این تست عمومی است یعنی روش‌هایی که در ساخت کپچا به کار می‌روند باید در اختیار همگان قرار گیرد، پس برنامه‌ای که می‌نویسیم باید آنقدر قدرتمند باشد که دیگران حتی با آگاهی کامل از روش‌ها، نتوانند آن را شکست دهند.

کپچاها به دو دسته کلی تقسیم می‌شوند: اول، کپچاهایی که بر پایه OCR^۵ هستند، و دوم، کپچاهایی که ربطی به OCR ندارند. OCR به روش‌های گفته می‌شود که سعی دارند با پردازش بر روی تصاویر، حروف یا کلمات موجود در آن را استخراج نمایند. در کپچاهای نوع اول یک کلمه، معمولاً تصادفی، ایجاد می‌شود و آن را در پیش‌زمینه یک تصویر قرار می‌دهند، "توابع خرابی" به آن تصویر اعمال می‌کنند بطوریکه برنامه‌های OCR کنونی نتوانند آن کلمه را تشخیص دهند، در نهایت تصویر را به کاربر نشان می‌دهند و از او خواسته می‌شود تا

با پیشرفت تکنولوژی‌های مرتبط با کامپیوتر و افزایش قدرت برنامه نویسان هرروزه برنامه‌های سودمندی بر روی اینترنت به دنیا عرضه می‌شود. مردم توسط امکانات موجود در شبکه جهانی اینترنت کارهای روزمره خود از قبیل ثبت‌نام و عضویت در سازمان‌ها، آموزش، تفریح، ارسال پست الکترونیک، انجام کارهای بانکی، جستجوی مقالات، و... را انجام می‌دهند و صاحبان سایت‌های اینترنتی نیز در مقابل ارائه این خدمات کسب درآمد می‌کنند.

در مقابل، قدرت هکرها نیز افزایش یافته و برنامه‌های مخرب قدرتمندتری تولید شده است. یکی از رایج‌ترین حملات هکرها، حمله به پایگاه داده‌های سایت‌هایی است که اطلاعات دریافت شده از کاربران خود را نگهداری می‌کنند. این برنامه‌های مخرب خود را به جای یک کاربر جا می‌زنند و با ثبت اطلاعات غلط و غیرمفید منابع سایت موردنظر را به هدر می‌دهند. به اینگونه برنامه‌های مخرب ممانعت از ارائه خدمات^۱ گویند. پس نیاز است تا در دنیای مجازی بتوان تنها با یک برنامه خودکار کامپیوتری تفاوت میان یک انسان و یک برنامه کامپیوتری را تشخیص داد (بدون اینکه بخواهیم محدودیت‌های مکانی و زمانی را در نظر بگیریم).

روش‌هایی که بوسیله آنها بتوان یک انسان را از یک کامپیوتر متمایز کرد معروف به "تست تورینگ" هستند [4]، در این آزمون برای ارزیابی هوشمندی یک کامپیوتر، کامپیوتر مربوطه را در یک اتاق و یک انسانی عادی را در اتاقی دیگر قرار می‌دهند سپس یک داور که در اتاق سومی قرار دارد با پرسش سوال سعی در تشخیص انسان از کامپیوتر خواهد کرد و هرچقدر که این تشخیص سخت‌تر باشد پس کامپیوتر مورد آزمایش هوشمندتر است.

هنگامی که یک برنامه کامپیوتری قادر باشد تستی تولید کند که ارزیابی فوق را به طور خودکار انجام دهد به آن CAPTCHA می‌گویند [3]. برای اولین بار «آندر برودر»^۲ در سال ۱۹۹۷ توانست چنین برنامه‌ای برای سایت Altavista.com ارائه دهد [8]، پروفیسور لویس وون آهن^۳ و همکاران او در دانشگاه Carnegie Mellon ایده آقای "برودر" را گسترش دادند و آن را CAPTCHA نامیدند [2]. در این روش سوالی از کاربر پرسیده می‌شود و از کاربر خواسته می‌شود تا به این سوال پاسخ دهد، اگر این تست با موفقیت گذرانده شود پس کاربر مورد نظر یک انسان است.

کپچاهای موجود عموماً سعی می‌کنند با ایجاد خرابی در یک تصویر حاوی یک یا چند حرف، امکان تشخیص آن را برای یک برنامه‌ی خودکار کامپیوتری مشکل کنند و در عین حال امکان خواندن آن توسط انسان فراهم باشد. این نرم افزارها بر پایه حروف انگلیسی عمل می‌کنند و معمولاً مجموعه‌ای از کلمات بی معنی را برای تشخیص انتخاب می‌کنند. همچنین اخیراً نمونه‌های فارسی از این نوع کپچاها

علاوه بر این روش‌های دیگری نیز تا کنون ارائه شده‌اند همچون Pessimal Print Method که کاراکترها را با کیفیت پایین نمایش می‌داد تا OCR نتواند آن را تشخیص دهد [10] اما این متد نیز هک شد چراکه متخصصان توانستند در تصاویر به‌کاررفته ساختار حروف را به حالت اولیه و با کیفیت آن برگردانند [2]، یا BaffleText Method که کلماتی را تولید می‌کرد که در زبان انگلیسی موجود نبودند و توابع خرابی به تصویر اعمال می‌کرد که هرچند کار OCR را مشکل می‌کرد اما خوانایی آن را نیز برای کاربر انسانی سخت می‌کرد [11].

سعی ما بر آن است تا کپچایی بر پایه OCR ارائه دهیم که در آن از کلمات فارسی استفاده می‌شود. اخیراً نمونه‌هایی در این زمینه ارائه شده است [2] که در آنها بیشتر بر روی ویژگی‌های خاص حروف و الفبای فارسی تاکید شده است. برای مثال در نمونه‌ای که توسط آقای شیرعلی شهرضا ارائه شده [2] یک کلمه فارسی تصادفی ایجاد و بر روی یک تصویر پشت‌زمینه رنگی با ویژگی‌های خاص قرار داده می‌شود، (شکل ۵ - الف)، این روش چون از ویژگی‌های خاص الفبای فارسی و پشت‌زمینه رنگی استفاده کرده است مقاومت خوبی در مقابل OCR از خود نشان داده است.

نمونه‌ی دیگری نیز توسط همین محقق ارائه شده است که در آن از خط نستعلیق بهره گرفته شده است (شکل ۵ - ب). در پیاده سازی این روش، بانکی از تصاویر کلمات نوشته شده به شیوه نستعلیق تهیه شده، نرم افزار مذکور به طور تصادفی یکی از این تصاویر را انتخاب کرده و به کاربر نشان می‌دهد. در این روش هیچ پردازش خاصی بر روی تصویر (به جز اعمال نویز) انجام نمی‌شود و همین مسئله می‌تواند نقطه ضعف اساسی این روش باشد چرا که اخیراً روش‌هایی جهت تشخیص کلمات نستعلیق ارائه شده است [18] که می‌توانند روش مذکور را با شکست مواجه سازند.



(الف)



(ب)

شکل (5): نمونه‌ای از کپچای فارسی [2]

آنچه را که می‌بیند تایپ کند (شکل ۱). اگر متن تایپ شده درست باشد پس تست با موفقیت گذرانده شده است و کاربر یک انسان تشخیص داده می‌شود، در غیر اینصورت تست دوباره تکرار می‌شود.

* Enter the code shown: [More info](#)
This helps Yahoo! prevent automated registrations.



شکل (1): یک نمونه کپچای بر پایه OCR [7]

1-2- کارهای انجام شده در زمینه کپچای نوع اول

در این زمینه کارهای فراوانی صورت گرفته‌است، برای مثال CMU's EZ-Gimpy که توسط "وون آهن" و همکاران او در دانشگاه Carnegie Mellon ارائه شد و توسط سایت Yahoo! نیز به کار گرفته شد [9]، (شکل ۲)، این متد کلمات خود را از یک دیکشنری با حدود ۸۵۰ کلمه انتخاب می‌کرد و از آنجایی که از یک فونت خاصی استفاده می‌کرد به آسانی شکست خورد [2,9].



شکل (2): CMU's EZ-Gimpy CAPTCHA [9]

سایت Paypal.com نیز متد دیگری در این زمینه ارائه داد، (شکل ۳)، اما به خاطر وجود فاصله زیاد میان حروف متدهای قطعه‌قطعه سازی تصاویر به راحتی هک شد [9].



شکل (3): نمونه کپچای سایت paypal [9]

میکروسافت هم متد دیگری برای سرویس پست الکترونیک خود، Hotmail، ارائه داد (شکل ۴). در این متد ایجاد خمیدگی در حروف و وجود خطوط زاید کار قطعه‌قطعه‌سازی تصویر را بسیار مشکل می‌کرد [2] اما برخی مواقع خمیدگی زیاد ناخوانایی حروف را برای کاربر موجب می‌شد [9].



شکل (4): نمونه کپچای استفاده شده در Hotmail [2]

2-2- چند نمونه از کپچای نوع دوم

در کپچاهای نوع دوم هیچ وابستگی به OCR وجود ندارد و تستی ارائه می‌شود که هیچ برنامه‌بینایی ماشینی^۶ نتواند آن را بگذراند، برای مثال در نمونه‌ای از آن که به Bongo شهرت یافته دو مجموعه اشکال مختلف به همراه یک شکل سوم به کاربر نشان داده می‌شود و از وی خواسته می‌شود تا مشخص کند که این شکل سوم متعلق به کدام مجموعه است [8] در ضمن سعی می‌شود این دو مجموعه از اشکال تا حد ممکن دارای ساختار مشابه باشند تا تشخیص آن برای نرم‌افزارهای تشخیص ویژگی‌ها^۷ مشکل باشد.

در روش دیگر موسوم به Pix که از نوع دوم است یک مجموعه تصاویر (معمولاً^۴ تا ۶ تصویر) به کاربر نشان داده می‌شود و از وی پرسیده می‌شود که "این تصاویر چه چیزی را نشان می‌دهند؟" [12].

نمونه دیگر "کپچای صوتی" است که شامل یک پایگاه داده‌ها از کلمه‌هایی است که به‌فرد صوت نویز دار ذخیره شده‌اند. در این کپچا به صورت تصادفی یکی از این کلمه‌ها برای کاربر پخش می‌شود و از وی خواسته می‌شود تا آنچه را می‌شنود تایپ کند [8].

هرچند که شکست کپچاهای نوع دوم سخت‌تر از نوع اول است اما گذراندن آن توسط کاربر انسانی نیز مشکل‌تر است و از طرفی بار محاسباتی که به پایانه‌های ارائه دهنده این نوع کپچا اعمال می‌شود از نوع اول بیشتر است.

3- الفبای فارسی، کپچای فارسی

در روش ارائه شده، سعی بر آن بوده که از ویژگیهای ساختاری الفبا و زبان فارسی برای بالا بردن قابلیت تشخیص انسان و پایین آوردن احتمال تشخیص OCR استفاده شود. پس ابتدا نگاهی خواهیم داشت بر ساختار و ویژگی‌های الفبای فارسی که در ایجاد تصاویر مطلوب به ما کمک شایانی می‌کنند.

الفبای فارسی شامل ۳۲ حرف است که هر کدام از آن‌ها می‌توانند با ۲ تا ۴ حالت مختلف در یک کلمه ظاهر شوند، حالت‌های مختلف قرار گرفتن حروف در یک کلمه امتحان شده است و ۱۱۸ حالت مختلف از کل این ۳۲ حرف بدست آمده است. برای مثال حرف "ح" را در نظر بگیرید: حمید، محمود، صبح؛ مشاهده می‌کنیم که حرف "ح" به سه شکل متفاوت در این سه کلمه ظاهر شده است. وابستگی حالت حرف به مکان قرار گرفتن در کلمه کار تشخیص ویژگی‌ها را برای OCR فارسی مشکل می‌کند [13].

حروف فارسی دارای اندازه‌های متفاوتی هستند که کار قطعه‌سازی آنها را مشکل می‌سازد، برای مثال حروف "گ، ه، و، د" هر کدام اندازه خاصی دارند.

بسیاری از حروف فارسی مشابه هستند و تنها تفاوتشان در یک یا چند نقطه و یا علائم خاصی است (جدول ۱). از آنجاییکه تشخیص

نقطه از نویز برای OCR مشکل است می‌توان از این ویژگی برای ایجاد تصویر کارآمدتر استفاده کرد.

جدول (۱): حروف مشابه در الفبای فارسی [2]

پ	ب	ت	ث	ج	چ	ح	خ	د	ذ	ر	ز	س	س	ص	ض	ط	ظ	ع	غ	ف	ق	ک	گ
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

در زبان فارسی می‌توان بسیاری از کلمات را به‌صورت کشیده نوشت که زیبایی نوشته را بالا می‌برد و از طرفی کار OCR را مشکل می‌کند [13]. برای نمونه: "می‌آید"، یا "محسبت".

4- کپچای فارسی

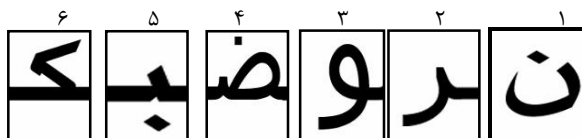
در این بخش به بررسی روش ارائه شده در این مقاله پرداخته می‌شود.

4-1- نحوه ایجاد تصاویر در کپچای فارسی

مرحله اول در تولید تصویر کپچا، ایجاد تصویر پایه است. برای اینکار نیاز است که حروف مطلوب شناسایی شود و در یک پیش زمینه مناسب قرار گیرد. شرح روش در زیر آمده است.

4-1-1- حروف

پایگاه داده‌ای شامل تمامی حروف الفبای فارسی و تمامی حالات مختلفی که این حروف می‌توانند در یک کلمه ظاهر شوند به صورت تصاویر هم اندازه ایجاد شده است، این تصاویر به چهار دسته تقسیم می‌شوند: "حروف اول چسبان"، "حروف آخر چسبان"، "حروف وسط"، "حروف غیر چسبان". در شکل (۶) چند نمونه از این تصاویر را مشاهده می‌کنیم.



شکل (۶): چند نمونه از تصاویر موجود در پایگاه داده حروف، به ترتیب از راست به چپ: ۱. غیر چسبان، ۲-۴. آخر چسبان، ۵، ۶. وسط.

از آنجایی که کلمه‌های موجود در تصاویر می‌بایست با این حروف ساخته شوند پس این حروف هر کدام متناسب با تقسیم بندی فوق در یک کلاس خاص قرار می‌گیرند تا هنگام انتخاب حروف برای ایجاد یک کلمه تصادفی بتوان ساختار مناسب‌تری ایجاد کرد. تصاویر این بانک را

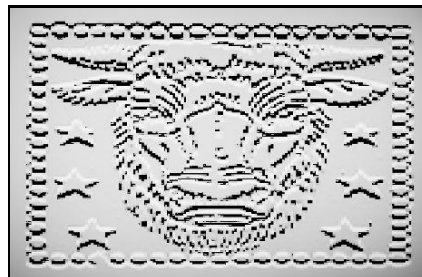
در اندازه‌های یکسانی از لحاظ تعداد پیکسل ایجاد شده‌اند تا در هنگام تشکیل کلمه بهتر در کنار یکدیگر قرار گیرند.

مشخص بودن قلم به کار رفته در یک تصویر کمک خوبی به OCR است تا ساختار حروف را بدرستی استخراج کند بنابراین تصاویر این حروف را با قلم‌های متفاوتی ایجاد می‌شوند تا وابسته به یک قلم خاص نباشند.

4-1-2- پشت‌زمینه

برای پشت‌زمینه، تصاویر مناسبی با اندازه‌های یکسان که دارای ویژگی‌های خاصی هستند انتخاب می‌شود. اکثر OCR های کنونی ابتدا سعی می‌کنند تا کلمه را از پشت‌زمینه‌ی آن جدا کنند و سپس به کار تشخیص حروف بپردازند، هرچه بتوانیم این کار را برای OCR مشکل‌تر سازیم کپی‌های بهتری ارائه داده‌ایم [13].

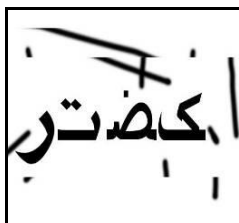
پشت‌زمینه‌هایی که دارای خطوط سیاه و سفید و اشکال ناموزون باشند، نویزی شکل و از لحاظ رنگ‌بندی طیف گسترده‌ای داشته باشند انتخاب مناسب‌تری هستند تا اینکه از پشت‌زمینه‌ی کاملاً سفید یا کاملاً سیاه استفاده کنیم، (شکل ۷). پس تعدادی تصویر پشت‌زمینه‌ی مناسب انتخاب، و در یک بانک اطلاعاتی ذخیره می‌شوند.



شکل (7): یک پشت‌زمینه‌ی مناسب

در این مرحله، الگوریتم سعی می‌کند: اول، حروفی که به هم مشابه ترند انتخاب کند. دوم، حروفی که شامل نقطه و علائم خاص هستند انتخاب کند. سوم، حروفی با اندازه‌های متفاوت انتخاب کند (بنا به مطالب ارائه شده در بخش ۳).

حروف انتخابی را بر روی تصویر پشت‌زمینه به ترتیب قرار داده می‌شود تا تصویر نهایی شامل یک کلمه‌ی فارسی باشد که از نظر ساختار الفبا و زبان فارسی درست باشد هرچند که معنای خاصی نداشته باشد. در شکل‌های (۸، ۹) چند نمونه از این تصاویر آمده است. تصاویری همانند شکل (۹) چندان با ساختار زبان فارسی سازگار نیستند و برنامه سعی می‌کند تا آنجایی که می‌تواند از ساخت چنین تصاویری پرهیز کند.



شکل (9): یک کلمه نامناسب



شکل (8): یک کلمه مناسب

تصاویر ساخته شده تا حدودی به هدف موردنظر نزدیک هستند اما باید دقت کرد که OCRها هرروز قوی‌تر می‌شوند. برنامه باید پردازش‌های خاصی روی تصاویر اعمال کند تا ساختار حروف به کار رفته و خود کلمه را به نحوی کنترل شده بر هم بزند و یا بر کل تصویر توابعی اعمال کند که تشخیص OCR با مشکل روبرو شود. این توابع را "تابع خرابی" گوئیم. بخش بعد به این موضوع اختصاص دارد.

4-2- اعمال توابع خرابی به تصویر

همانطور که در بخش قبل گفته شد، تصویر بدست آمده می‌بایست تخریب شود تا توسط یک برنامه خودکار قابل تشخیص نباشد. این کار با توابعی معروف به "توابع خرابی" انجام می‌شود. این توابع را در رده‌های مختلفی، بر اساس تاثیر آنها بر تصویر، تقسیم بندی می‌شوند و برنامه پس از ایجاد تصویر به صورت تصادفی یکی از این رده‌ها را انتخاب و بر تصویر اعمال می‌کند. دلیل رده‌بندی و انتخاب تصادفی توابع این است که: همواره لازم نیست همه "توابع خرابی" به تصویر اعمال شوند چراکه باعث می‌شوند خوانایی کلمه برای کاربر از بین برود و از طرفی وقت‌گیر هم هست. و همچنین مشخص و ثابت بودن توابع خرابی به OCRها کمک شایانی می‌کند و یا به اشخاص نفوذگر اجازه می‌دهد با اعمال عکس این تابع ثابت، اثر آن را از بین ببرد.

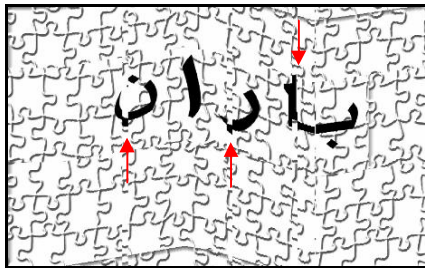
4-1-3- ایجاد تصویر

با مهیا بودن تصاویر مربوط به حروف و پشت‌زمینه و داشتن توابع لازم و مورد نیازی که در محیط نرم‌افزاری چون MATLAB پیاده‌سازی می‌شوند می‌توان تصاویر مطلوب را ایجاد کرد. الگوریتم برنامه به شکلی که در زیر توضیح داده‌ایم سازماندهی شده است.

- یک تصویر پشت‌زمینه انتخاب می‌شود.
- بین ۱ تا ۷ حرف از تصاویر موجود در "بانک حروف" به ترتیب انتخاب می‌شود، باید دقت کرد که کلمه‌ای مطلوب است که خواندن، درک و تایپ آن برای کاربر فارسی‌زبان راحت باشد؛ پس باید حروف طبق قواعد خاصی انتخاب شوند. برای مثال آوردن یک "حرف وسط" در آغاز کلمه و یا پایان کلمه با یک "حرف اول چسبان" با ساختار زبان فارسی ناسازگار و در نتیجه برای کاربر فارسی‌زبان ناخوانا است.

4-2-3- ساختارشکنی با شیفت

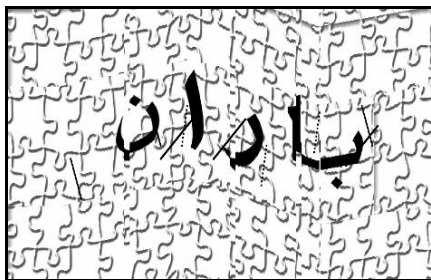
هر حرف ساختار و ویژگی خاصی دارد که با تشخیص آن‌ها بهتر می‌توان یک حرف را از تصویر استخراج کرد [15]. پس برای جلوگیری از این مسئله می‌بایست ساختار حروف موجود در کلمه را به نحوی کنترل شده بر هم زد. این تابع با اعمال شیفت به نقاط مختلف تصویر این کار را انجام می‌دهد، شکل (۱۲).



شکل (12): ساختارشکنی حروف با شیفت

4-2-4- خط خطی کردن تصویر

کشیدن خطوط با زوایای مختلف در نقاط مختلف تصویر یک ایده سنتی اما همچنان کاراست. اکثر OCR ها ابتدا تا آنجاکه می‌توانند زواید موجود در تصویر را از بین می‌برند و کلمه را از تصویر پشت‌زمینه جدا می‌کنند سپس به تشخیص حروف یا کلمه می‌پردازند [15] پس با اضافه کردن این خطوط گره دیگری در کار OCR ایجاد خواهیم کرد، شکل (۱۳).



شکل (13): کشیدن خطوط در تصویر

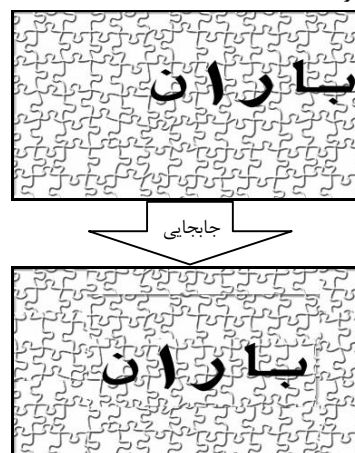
4-2-5- ایجاد حرکت در تصویر^{۱۱}

"قطعه‌قطعه‌سازی" تصویر یکی از کاربردی‌ترین مباحث در OCR است [13-15] که با هموارسازی^{۱۲} تصویر و ایجاد حرکت دچار مشکل می‌شود. با اعمال "فیلترهای حرکتی" به تصویر، شکل (۱۴)، به طور کنترل شده می‌توان به این مهم دست پیدا کرد. البته باید توجه داشت که این نوع خرابی در تصویر، بینایی را برای کاربر انسانی با مشکل مواجه خواهد کرد و در الگوریتم پیشنهادی سعی می‌شود تا حد ممکن از این تابع استفاده نشود.

در نهایت، می‌توان با دسته‌بندی مناسب، با تعداد توابع کمتری تصویر مناسب‌تری ایجاد کرد. در بخش بعد به تفصیل این توابع را بررسی خواهیم کرد. توابع خرابی که در اینجا بررسی می‌شوند همگی بوسیله نرم‌افزار MATLAB پیاده‌سازی شده‌اند. یک تصویر نمونه مطابق با آنچه در بخش پیشین توضیح داده شد ایجاد شده و اثر آنها را بر این تصویر مشاهده و تحلیل خواهد شد.

4-2-1- جابجایی^۸

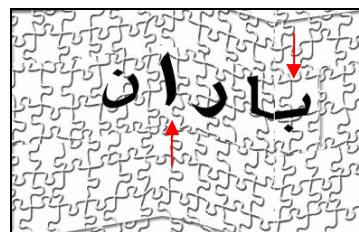
کلمه ساخته شده در تصویر به نقاط مختلف تصویر انتقال داده می‌شود چراکه ثابت بودن مکان کلمه در تصویر کمک زیادی به OCR است زیرا دیگر نیازی به جستجوی مکان کلمه ندارد [13]. "تابع جابجایی" با انتقال کلمه به هر نقطه‌ای از تصویر این کار را انجام می‌دهد که در شکل (۱۰) مشهود است.



شکل (10): جابجایی کلمه در تصویر

4-2-2- خمیدگی^۹

چون قرار گرفتن کلمه بر روی خط مستقیم به OCR کمک می‌کند تا بهتر بتواند خط پایه^{۱۰} کلمه را پیدا کند [14] پس اگر بتوانیم خمیدگی در تصویر اعمال کنیم بیشتر به هدف خود نزدیک خواهیم شد. البته این خمش باید کنترل شده باشد به نحوی که ساختار کلمه حفظ شود (شکل ۱۱).



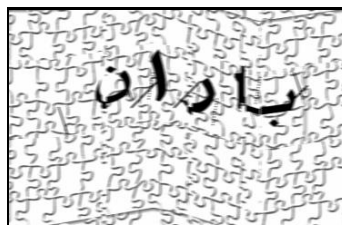
شکل (11): خمیدگی کلمه در تصویر

تشخیص" توسط برنامه‌های OCR کنونی و "خوانایی کلمه برای کاربر" است که روش پیشنهادی با بیان توابع به‌کاررفته اصل "باز" بودن روش، با اعمال توابع خرابی گوناگون و استفاده از ویژگی‌های خاص زبان فارسی اصل دوم و در نهایت با کنترل خرابی‌های اعمال شده به تصاویر اصل "خوانایی کلمه برای کاربر" را رعایت کرده است.

دقت روش ارائه شده توسط نرم‌افزارهای تشخیص متن بررسی و صحت عملکرد آن با اطمینان بالا تأیید شده است، لیکن با توجه به تازه‌گی مطلب نیاز به آزمایش‌های بیشتر در این زمینه احساس می‌شود. سعی خواهد شد با پیاده‌سازی دقیق‌تر کپچای فارسی و به-کارگیری در اینترنت به نقاط ضعف آن پی برده شود.

مراجع

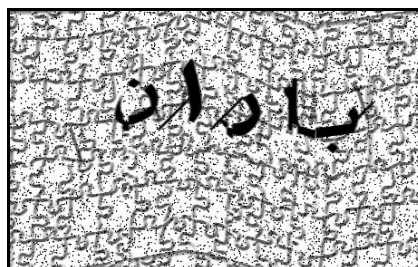
- [1] Von Ahn, L., Blum, Manuel J.Hopper, Nicholas Langford, Jhon, "CAPTCHA: Using Hard AI Problems For Security," Carnegie Mellon University, In Proc. of Eurocrypt 2003, pp.294-311.
- [2] Shirali-Shahreza, M. H., Shirali-Shahreza, M., "Persian/Arabic Baffletext CAPTCHA," Journal of Universal Computer Science, vol. 12, no. 12(2006), 1783-1796.
- [3] Datta, Ritendra, Li, Jia, Z.Wang, James, "IMAGINATION: A Robust Image-based CAPTCHA Generation System," The Pennsylvania State University and University Park, USA
- [4] Turing, A, "Computing Machinery and Intelligence," Mind,59(236):433_460, 1950
- [5] Von Ahn, Luis Blum, Manuel, Langford, Jhon, "Telling Humans and Computers Apart Automatically," Communications of the ACM, February 2004, Vol. 47, No. 2
- [6] Mori, Greg, Malik, Jitendra, "Recognizing Object in Adversarial Clutter: Breaking a Visual CAPTCHA," Proceedings of the IEEE CS Society Conference on Computer Vision and Pattern Recognition (CVPR'03), Madison, WI,2003, 134-141.
- [7] Lopresti, Daniel, "Leveraging the CAPTCHA Problem," Computer Science & Engineering, Lehigh University, Bethlehem, PA 18015, USA.
- [8] Von Ahn, Luis Blum, Manuel, Langford, Jhon, "Telling Humans and Computers Apart (Automatically) or How Lazy Cryptographers do AI," Computer Science Dept., Carnegie Mellon University, Pittsburgh, PA 15213.
- [9] Baird, Henry. S, Riopka, Terry, "ScatterType: a Reading CAPTCHA Resistant to Segmentation Attack," Computer Science & Engineering, Lehigh University, Accepted for publication in Proceedings, IS & T/SPIE Document Recognition & Retrieval XII Conference, san Jose, CA, January 16-20, 2005.
- [10] Coates A.L et al: "Pessimial Print: A Reverse Turing Test," in proceedings of 6th International Conference on Document Analysis and Recognition, Seattle, WA, USA, 2001, 1154-1158.
- [11] Chew, M, Baird, Henry. S, "Baffle Text: a Human Interactive Proof," Proceedings of the 10th SPIE/IS&T Document Recognition and Retrieval Conference (DRR2003), Sanat Clara, CA, 2003, 305-316.
- [12] Von Ahn, Luis Blum, Manuel, Langford, Jhon, The Captcha Project homepage: <http://www.captcha.net>.



شکل (14): ایجاد حرکت در تصویر

4-2-6- ایجاد نویز در تصویر

همانطور که اشاره شد یکی از ویژگی‌های الفبای فارسی متفاوت بودن برخی حروف آن در تنها چند نقطه است. با اعمال نویزهای مختلف همچون نویز نمک-فلفلی به تصویر، شکل (۱۵)، می‌توان تشخیص نویز از نقطه را برای OCR مشکل کرد [16]. از آنجایی که اعمال این نوع خرابی در تصاویر هم بار محاسباتی و پیچیدگی پایینی دارد و هم کارایی تصاویر را بالا می‌برد یکی از توابع متداول در ساخت تصاویر مطلوب است.



شکل (15): اعمال نویز نمک-فلفلی به تصویر

و می‌توان چندین تابع دیگر چون مقیاس پذیری^{۱۳}، چرخش، متمم-سازی و... نیز به تصویر اعمال کرد که اثری که هرکدام بر تصویر اعمال می‌کنند می‌تواند خلی در کار OCR ایجاد کند. لیکن متفاوت بودن این خرابی‌ها در هر بار اجرا و اعمال همزمان چند خرابی به طور کنترل شده کار برنامه‌های تشخیص خودکار را بسیار مشکل یا تقریباً غیرممکن می‌کند.

5- نتیجه گیری

روشی که در این مقاله ارائه شد سیستمی اتوماتیک جهت تشخیص کاربران فارسی زبان از برنامه‌های مخربی است که قصد به هدر دادن منابع سایت‌ها را دارند. این روش در آینده‌ای نزدیک و با همگانی شدن استفاده از اینترنت در کشورمان برای کاربران فارسی زبانی که با زبان لاتین آشنایی چندانی ندارند و یا زبان فارسی را ترجیح می‌دهند، مناسب خواهد بود.

در این مقاله سعی شد با توجه به اصول اولیه کپچا، سیستمی قابل اطمینان برای جلوگیری از ورود خودکار به سایت‌های اینترنتی ارائه شود. سه اصل اساسی در طراحی کپچا "باز" بودن روش، "مشکل بودن

- [13] Johnston, Alexander, "Classifying Persian Characters with Artificial Neural Networks and Inverted Complex Zernike Moments," 15th June 2005.
- [14] Salmani Jelodar, M, Fadaeieslam, M.J, Mozayani, N, Fazeli, "A Persian OCR System Using Morphological Operators," Transactions on Engineering, Computing and Technology V4 February 2005 ISSN 1305-5313, Manuscript received January 21,2005.
- [15] Broumandnia, A, Fathy, M, "Application of pattern recognition for Farsi license plat recognition," Department of Computer, Islamic Azad University Branch of Tehran South, Department of Computer, Iran University of science and technology.
- [16] M.Lorigo, Liana, Govindaraju, Venu, "Off-line Arabic Handwriting Recognition: A Survey," Member & Senior Member in IEEE, To Appear in I120
- [17] IEEE Transactions on Pattern Analysis and Intelligence.
- [18] Zaheer Ahmad, Jehanzeb Khan Orakzai, Inam Shamsheer, Awais Adnan, "*Urdu Nastaleeq Optical Character Recognition*", Proceedings of world academy of Science, Engineering and Technology, volume 26 December 2007 ISSN 1307-6884.

[۱۹] م. شیرعلی شهرضا، "تفکیک انسان از ماشین به کمک خط نستعلیق"، نشریه مهندسی برق و مهندسی کامپیوتر ایران، سال ۵، شماره ۲، تابستان ۱۳۸۶.

زیرنویس‌ها

-
- ¹ Denial of Services
 - ² Andrei Broder
 - ³ Luis Von Ahn
 - ⁴ Image Processing
 - ⁵ Optical Character Recognition
 - ⁶ Machine Vision
 - ⁷ Feature Detection
 - ⁸ Translate
 - ⁹ Curving
 - ¹⁰ Base Line
 - ¹¹ Motioning
 - ¹² Smoothing
 - ¹³ Scaling