



آشکارسازی لوگو در تصاویر اسناد با استفاده از روش‌های مورفولوژی و الگوریتم گسترش مرز چهارگوش ویژگی

سینا حسن‌زاده^۱، حسین پورقاسم^۲ و پیمان معلم^۳

^۱دانشجوی کارشناسی ارشد دانشگاه آزاد اسلامی واحد نجف‌آباد

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

^۳استادیار دانشکده برق دانشگاه اصفهان، p_moallem@eng.ui.ac.ir

چکیده- لوگو معمولا در تصاویر اسناد معرف منبع سند و مالکیت آن می‌باشد. روش جدیدی که در این مقاله جهت آشکارسازی لوگو پیشنهاد شده است به دو بخش تقسیم می‌شود. در بخش اول توسط عملیات مورفولوژی مکان‌های کاندید لوگو را استخراج کرده و در بخش دوم توسط الگوریتم گسترش مرز چهارگوش ویژگی مکان دقیق لوگو شناسایی می‌گردد. در هر دو بخش از طبقه بند درخت تصمیم استفاده کرده به گونه‌ای که پذیرش اشتباه لوگو تا حد ممکن پایین نگه داشته شود. این روش بر روی مجموعه وسیعی از تصاویر اسناد مورد آزمایش قرار گرفته و کارایی آن در آشکارسازی لوگو نشان داده شده است.

کلید واژه- آشکارسازی لوگو، الگوریتم گسترش مرز چهارگوش ویژگی، عملیات مورفولوژی، درخت تصمیم

در [۳] از این ویژگی که چگالی مکانی پیکسل‌های پیش

زمینه در ناحیه لوگو نسبت به نواحی دیگر در تصویر سند بیشتر می‌باشد، استفاده شده است. ابتدا یک پنجره ایجاد شده و عمل جاروب تصویر صورت می‌گیرد و در هر جاروب، چگالی مکانی محاسبه می‌گردد. بعد از جاروب کل تصویر، بیشترین مقدار چگالی مکانی نشان دهنده مکان لوگو می‌باشد. یکی از ضعف‌های این روش عدم آشکارسازی کامل لوگو است.

در [۴] روشی پیشنهاد شده است که ترکیبی از آشکارسازی و استخراج لوگو در یک قالب کاری یکپارچه می‌باشد. ابتدا جهت ادغام تمام قسمت‌های جدا از هم لوگو اندازه تصویر سند کاهش می‌یابد. سپس عمل طبقه بندی اولیه لوگو از غیر لوگو توسط طبقه بند فیشر آموزش دیده انجام شده و آشکارسازی نهایی به کمک طبقه بند‌های متوالی صورت می‌گیرد. زمانی که فاصله میان لوگو و بخش‌های دیگر سند کم باشد، ناحیه لوگو و غیر لوگو ادغام خواهند شد که یکی از ضعف‌های این روش به شمار می‌آید.

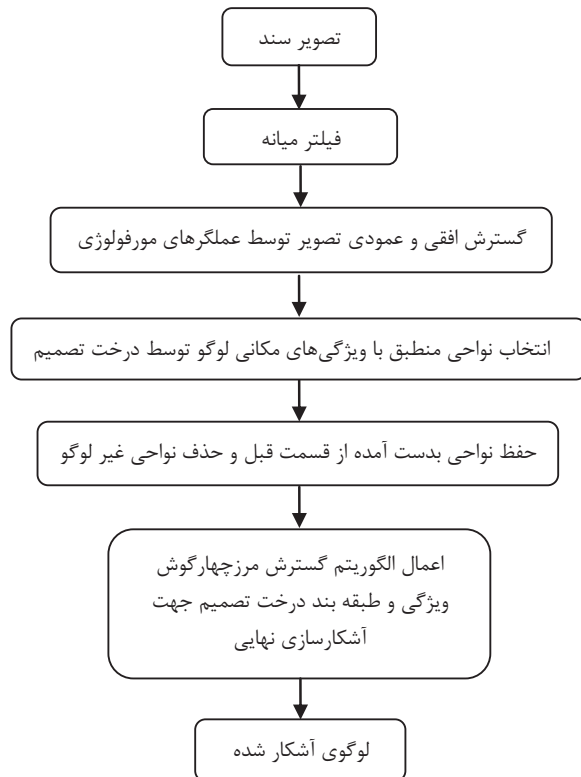
در [۵] از دو ویژگی فشردگی مکانی و یکنواختی رنگ سنجی جهت آشکارسازی لوگو استفاده می‌شود. ابتدا محتوای تصویر کاهش داده شده و فاصله میان اجزای متنی و گرافیکی لوگو توسط عملیات مورفولوژی کم می‌شود. سپس ناحیه لوگو

۱ - مقدمه

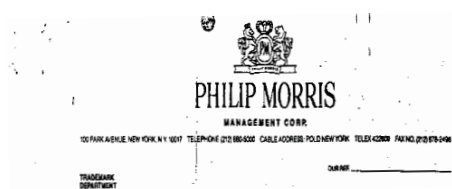
یکی از نیازهای دنیای امروز مکانیزه کردن طبقه بندی اسناد جهت صرفه جویی در زمان می‌باشد. با افزایش روزمره حجم اسناد، شناسایی لوگو به عنوان یک روش عملی و قابل اعتماد جهت شناسایی منبع اسناد و تحلیل محتوای آن به شمار می‌آید. شناسایی صحیح لوگو وابسته به آشکارسازی صحیح آن است. لوگو معمولا ترکیبی از متن و بخش‌های گرافیکی بوده که این امر عمل آشکارسازی را مورد تاثیر قرار می‌دهد. همچنین از دیگر مسائل در آشکارسازی لوگو، تنوع طرح بندی اسناد و کیفیت مختلف تصاویر اسکن شده می‌باشد.

کارهای انجام شده در زمینه لوگو مربوط به تصاویر اسناد به دو دسته آشکارسازی لوگو [۵،۶،۲،۳،۴] و شناسایی لوگو [۷،۸،۹،۱۰،۱۱،۱۲] تقسیم می‌شوند. روش [۲] از ناحیه بندی درخت تصمیم $x-y$ از نوع سلسله مراتبی بالا به پایین جهت تحلیل اسناد باینری استفاده می‌کند. پس از ناحیه بندی، از هر ناحیه ۱۶ ویژگی استخراج شده و توسط یک طبقه بند مبتنی بر قانون عمل طبقه بندی لوگو و غیر لوگو صورت می‌گیرد. چون الگوریتم درخت $x-y$ همواره خوب عمل نمی‌کند، این روش قابل اعتماد نمی‌باشد.

است. شکل ۳ نتیجه را نشان می‌دهد.



شکل ۱: مراحل آشکارسازی لوگو



شکل ۲: نمونه ای از تصویر یک سند



شکل ۳: تصویر حاصل از گسترش افقی و عمودی روی تصویر.

۲-۳- طبقه‌بند درخت تصمیم

در این مرحله ناحیه‌هایی که دارای ویژگی‌های ابعادی لوگو می‌باشند توسط طبقه‌بند درخت تصمیم انتخاب می‌شوند. این ویژگی‌ها در جدول ۱ آورده شده است. شکل ۴ درخت تصمیم به کار رفته با دو گره را نشان می‌دهد.

توسط چگالی‌های مکانی و رنگی آشکارسازی می‌گردد. این روش مختص تصاویر رنگی می‌باشد.

در [۶] به کمک الگوریتم گسترش مرز چهارگوش ویژگی، موقعیت چهارگوش در برگیرنده هر جزء به هم متصل را استخراج کرده و سپس توسط طبقه بند درخت تصمیم عمل آشکارسازی صورت می‌گیرد. در صورت وجود نویز زیاد در تصویر، سرعت و دقت الگوریتم به میزان زیادی کاهش می‌یابد.

روشی که پیشنهاد شده است در راستای بهبود [۶] می‌باشد. به طوریکه قبل از اعمال الگوریتم گسترش مرز چهارگوش ویژگی، به کمک عملیات مورفولوژی تعداد مکان‌هایی را که این الگوریتم می‌بایست بروی آن‌ها عمل کند را کاهش داده که منتهی به بهبود دقت و سرعت خواهد شد.

ساختار این مقاله بدین صورت خواهد بود که بخش دوم به بیان روش پیشنهادی جهت آشکارسازی لوگو می‌پردازد. در فصل سوم نتایج عملی مقاله و در فصل چهارم نتیجه گیری نهایی ارائه شده است.

۲- روش پیشنهادی جهت آشکارسازی لوگو

مراحل مختلف روش پیشنهادی جهت آشکارسازی لوگو در شکل ۱ نشان داده شده است. با استناد به روش [۶]، مکان لوگو معمولاً در قسمت یک پنجم بالای تصویر قرار دارد. شکل ۲ نمونه‌ای از تصویر یک سند می‌باشد.

۲-۱- فیلتر میانه

پس از خواندن تصویر ورودی به آن فیلتر میانه اعمال می‌شود تا نویزهای نقطه‌ای احتمالی موجود را از بین ببرد. قابل ذکر است به دلیل کیفیت مختلف تصاویر اسناد، مجاز به استفاده محدود از آن می‌باشیم در غیر این صورت احتمال از بین رفتن بخشی از لوگو وجود خواهد داشت.

۲-۲- گسترش افقی و عمودی تصویر

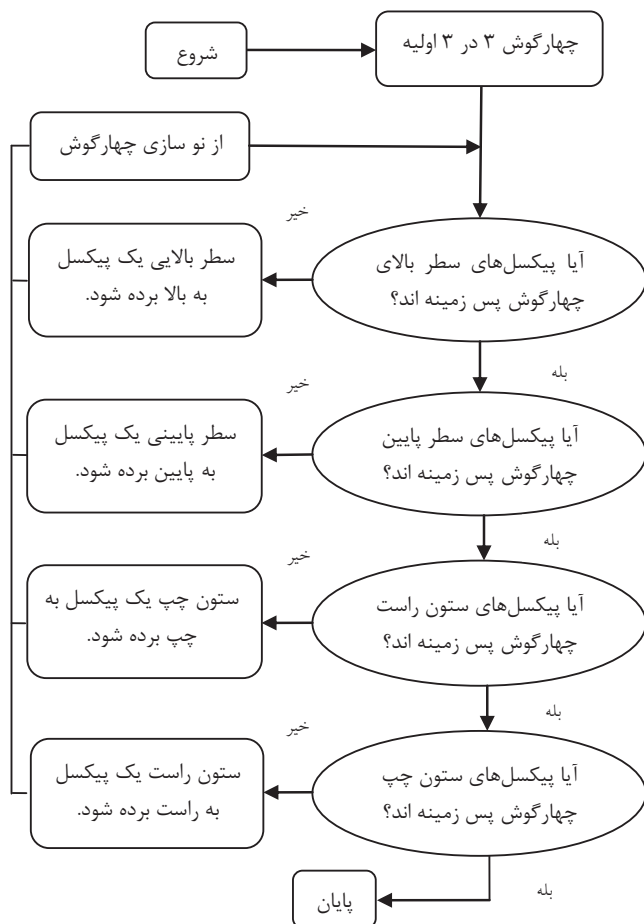
این عمل جهت اتصال بخش‌های جدا از هم لوگو صورت می‌گیرد. عمل گسترش به صورت زیر تعریف می‌شود:

$$A \oplus B = \{x | [(\hat{B})_x \cap A] \subseteq A\} \quad (1)$$

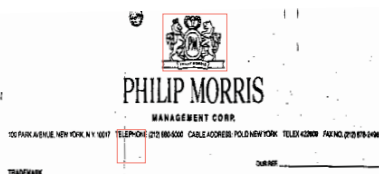
گسترش A با B مجموعه تمام x ‌هایی است که $(\hat{B})_x$ و A لا اقل در یک عنصر غیر صفر هم‌پوشانی داشته باشند. $(\hat{B})_x$ قرینه B حول مرکز خودش و سپس انتقال این قرینه با x



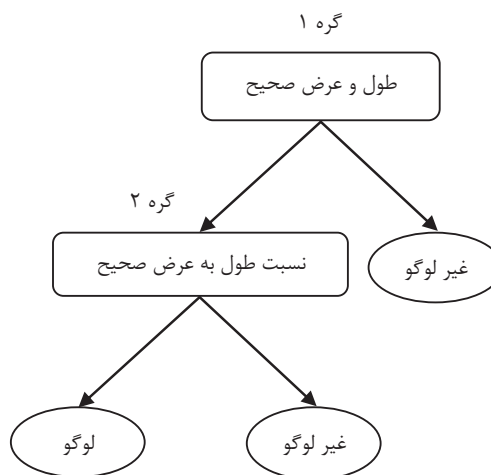
شکل ۵: دو ناحیه نمونه کاندید لوگو



شکل ۶: روند کار روش گسترش مرز چهار گوش ویژگی [۶]



شکل ۷: نواحی به دست آمده به عنوان لوگو



شکل ۴: طبقه بند درخت تصمیم جهت انتخاب نواحی کاندید لوگو

۴-۲ حذف نواحی غیر لوگو

در این مرحله به ازاء هر ناحیه کاندید لوگو، تمام نواحی به غیر از آن ناحیه را حذف می‌کنیم. با این کار، الگوریتم گسترش مرز چهار گوش ویژگی فقط روی چند ناحیه کاندید لوگو اعمال خواهد شد. شکل ۵ نتیجه را برای دو ناحیه نمونه از نواحی کاندید لوگو نشان می‌دهد. بر اساس آزمایش‌های انجام شده تعداد این نواحی از چهار ناحیه فراتر نخواهد رفت.

۵-۲ الگوریتم گسترش مرز چهار گوش ویژگی

بعد از یافتن نواحی کاندید لوگو به طور جداگانه، از الگوریتم فوق برای یافتن مکان دقیق لوگو استفاده می‌کنیم. روند کلی الگوریتم فوق بدین صورت است که ابتدا از طریق جاروب بالا به پایین تصویر سند، اولین پیکسل پیش زمینه‌ای که یافت شد دور آن یک چهار گوش ۳ در ۳ تشکیل می‌شود. این چهار گوش را چهار گوش ویژگی می‌نامند [۶] و طبق تعریف کوچکترین پنجره‌ای است که ناحیه کاندید لوگو را در بر می‌گیرد و چهار گوشه آن شامل پیکسل‌های پس زمینه بوده و دارای کمترین مساحت می‌باشد. این چهار گوش با گسترش مرز ناحیه کاندید لوگو بزرگ شده تا اینکه تمام ناحیه را در برگیرد. در روش پیشنهادی به جای جاروب کلی تصویر، الگوریتم مذکور فقط بروی نواحی بدست آمده از مرحله ۲-۴ اعمال می‌گردد. سپس توسط طبقه بند درخت تصمیم ذکر شده در بخش ۲-۳، پذیرش اشتباه لوگو تا حد ممکن کاهش داده می‌شود. شکل ۶ روند کار روش [۶] و شکل ۷ نتیجه نهایی را نشان می‌دهد.



شکل ۹: چند نمونه لوگوی به درستی آشکار نشده به دلیل وجود نویز زیاد

همانطور که قبلا ذکر شد به دلیل کیفیت مختلف تصاویر اسناد مجاز به استفاده محدود از فیلتر می‌باشیم. به همین دلیل بعضی از لوگوهایی که در اطراف آن‌ها نویز زیادی وجود دارد به درستی آشکار نمی‌شوند. این امر بدان علت است که چون در اطراف لوگو نویز زیادی وجود دارد چهارگوش ویژگی این نویزها را قسمتی از لوگو در نظر گرفته و با گسترش نویز چهارگوش ویژگی نیز بزرگ می‌شود، به طوریکه محدوده‌ای بزرگتر از ناحیه لوگو را در بر می‌گیرد که در بیشتر موارد توسط طبقه‌بند درخت تصمیم به عنوان ناحیه لوگو طبقه بندی می‌شود ولی طبق تعریف یک عمل آشکارسازی موفق به حساب نمی‌آیند. شکل ۹ چند نمونه از لوگوهایی را که به دلیل وجود نویز زیاد در اطراف لوگو با موفقیت آشکار نشده‌اند نشان می‌دهد.

نتایج بدست آمده در آشکارسازی لوگو با استفاده از روش پیشنهادی به کمک نرم افزار مطلب، شامل ۸۲ درصد صحت و ۹۴ درصد دقت می‌باشد. در جدول ۲ مقایسه عملکرد میان روش پیشنهادی و روش [۶] نشان داده شده است. یکی دیگر از ضعف های روش [۶] مربوط به اسنادی است که در اطراف لوگو کادر وجود دارد. در این حالت الگوریتم گسترش مرز چهارگوش ویژگی دچار خطا می‌شود. چون چهارگوش ویژگی به جای گسترش در اطراف لوگو، به دور کادر گسترش یافته و در نتیجه لوگو را از دست می‌دهد. شکل ۱۰ یک نمونه از این اسناد را نشان می‌دهد که در آن آشکارسازی لوگو توسط روش [۶] دچار خطا شده است. شکل ۱۱ آشکارسازی موفق لوگو توسط روش پیشنهادی در این گونه اسناد نشان می‌دهد. در صورتی که در روش پیشنهادی این مشکل وجود ندارد. این بدان علت است که ابتدا نواحی کاندید لوگو به کمک عملیات مورفولوژی و طبقه بند درخت تصمیم انتخاب شده، سپس الگوریتم گسترش مرز چهارگوش ویژگی در داخل کادر و به طور مستقیم بروی لوگو اعمال می‌شود.

روش پیشنهادی نسبت به روش [۶] زمان پردازش کمتری نیاز دارد. زمان بر بودن الگوریتم گسترش مرز چهارگوش ویژگی با نگاهی به روند کار این الگوریتم در شکل ۶ مشاهده می‌شود. با کارگیری عملیات مورفولوژی و طبقه بند درخت تصمیم، تعداد مکان‌هایی را که الگوریتم گسترش مرز چهارگوش ویژگی

همانطور که در شکل ۷ دیده می‌شود دو ناحیه به عنوان ناحیه لوگو انتخاب شده است که پذیرش اشتباه در اینجا یک مورد می‌باشد. چون ناحیه‌ای که به اشتباه پذیرش شده است در پایگاه داده لوگو وجود ندارد، در بخش شناسایی لوگو به طور خودکار حذف خواهد شد.

۳- نتایج عملی

پایگاه داده‌ای که در این مقاله استفاده شده است، tobacco-۸۰۰ [۱۳] می‌باشد. این پایگاه داده شامل ۱۲۹۰ تصویر اسکن شده اسناد بوده که ۴۱۶ عدد از این تعداد حاوی لوگو می‌باشد. از این تعداد، ۱۵۰ تصویر جهت آموزش و ۲۶۶ تصویر جهت آزمایش انتخاب شده است. بر طبق تعریف روش [۴]، عمل آشکارسازی موفق خواهد بود اگر و تنها اگر ناحیه آشکار شده بیش از ۷۵ درصد پیکسل های لوگوی پایه و کمتر از ۱۲۵ درصد ناحیه لوگوی پایه را شامل شود. جهت ارزیابی روش پیشنهادی، از دو معیار صحت و دقت آشکارسازی روش [۴] تعریف شده به صورت زیر استفاده می‌کنیم:

$$(۲) \quad \text{تعداد لوگوهای به درستی آشکار شده} = \frac{\text{تعداد کل لوگوهای آزمایش شده}}{\text{صحت}}$$

$$(۳) \quad \text{تعداد لوگوهای به درستی آشکار شده} = \frac{\text{تعداد کل لوگوهای آشکار شده}}{\text{دقت}}$$

همانطور که قبلا گفته شد از درخت تصمیم جهت طبقه بندی لوگو از غیر لوگو و کاهش پذیرش اشتباه لوگو استفاده می‌شود. هدف اولیه این می‌باشد که تا حد ممکن لوگویی از دست نرود، در حالیکه تعداد پذیرش اشتباه نیز پایین نگه داشته شود. درخت تصمیم با تعداد ۱۰۰ نمونه لوگو و ۱۰۰ نمونه غیر لوگو آموزش داده می‌شود. جدول ۱ قوانین تصمیم گیری درخت طبقه بند و شکل ۸ چند نمونه از لوگوهایی که به درستی توسط روش پیشنهادی آشکار شده است را نشان می‌دهد.



شکل ۸: چند نمونه از لوگوی های آشکار شده توسط روش پیشنهادی

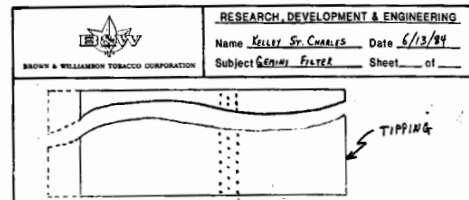
۴- نتیجه گیری

در این مقاله روشی سریع جهت آشکارسازی لوگو در تصاویر اسناد ارائه گردید. در این روش پس از حذف نویز، عمل گسترش افقی و عمودی توسط عملگرهای مورفولوژی بر روی تصویر سند انجام می‌گیرد. سپس توسط یک طبقه بند درخت تصمیم عمل طبقه‌بندی اولیه لوگو از غیر لوگو صورت گرفته و بعد از حذف نواحی غیر لوگو، الگوریتم گسترش مرز چهارگوش ویژگی به نواحی کاندید لوگو اعمال می‌شود. سرانجام توسط طبقه بند درخت تصمیم ذکر شده عمل طبقه بندی نهایی صورت می‌گیرد. بر طبق آزمایش‌های انجام شده، این روش نسبت به سایر روش‌های آشکارسازی لوگو در تصاویر اسناد سریع تر می‌باشد.

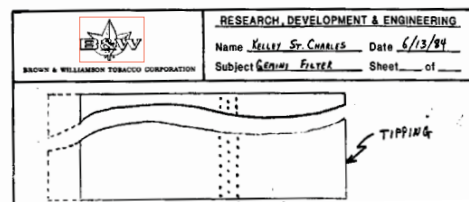
مراجع

- [۱] The university of Maryland (UMD) logo database, <http://lampsrv.umi.acs.umd.edu/projdb/project.php?id=۴۷>.
- [۲] S. Seiden, M. Dillencourt, S. Irani, R. Borrey, and T. Murphy, "Logo detection in document images". In *Proc. Int'l conf. Imaging science, Sys, and Tech*, pp. ۴۴۶-۴۴۹, ۱۹۹۷.
- [۳] T. Pham, "Unconstrained logo detection in document images", *Pattern recognition*, Vol. ۳۶, pp. ۳۰۲۳-۳۰۲۵, ۲۰۰۳.
- [۴] G. Zhu and D. Doermann, "Automatic document logo detection", In *conference on document analysis and recognition*, pp. ۸۶۴-۸۶۸, ۲۰۰۷.
- [۵] Z. Ahmed and H. Fella, "Logos extraction on picture documents using shape and color density", pp. ۲۴۹۲-۲۴۹۶, ۲۰۰۸.
- [۶] H. Wang, Y. and Chen, "Logo detection in document images based on boundary extension of feature rectangles", IEEE computer society, 1st Int'l conference on document analysis and recognition, pp. ۱۳۳۵-۱۳۳۹, ۲۰۰۹.
- [۷] D. Doermann, E. Rivlin, and I. Weiss, "Applying algebraic and differential invariants for logo recognition", *Machine vision and application*, pp. ۷۳-۷۶, ۱۹۹۶.
- [۸] P. Suda, C. Bridoux, B. Kammerer and G. Maderlechner, "logo and word matching using general approach to signal registration", In *Proc. Int'l conf. document analysis and recognition*, pp. ۶۱-۶۵, ۱۹۹۷.
- [۹] M. Gori, M. Maggini, S. Marinai, J.Q. Sheng and G. Soda, "Edge-backpropagation for noisy logo recognition", *Pattern recognition*, Vol. ۳۶, pp. ۱۰۳-۱۱۰, ۲۰۰۳.
- [۱۰] J. Chen, M.K. Leung and Y. Gao, "Noisy logo recognition using line segment hausdorff distance", *Pattern recognition*, Vol. ۳۶, pp. ۹۴۳-۹۵۵, ۲۰۰۳.
- [۱۱] K. Zyga, J. Schroeder, and R. Price, "Logo recognition using retinal coding", In *Proc. 3rd Asilomar Conf. Signals, system and Computers*, Vol. ۲, pp. ۱۵۴۹-۱۵۵۳, ۲۰۰۴.

به آن‌ها اعمال می‌شود کاهش داده که منجر به صرفه جویی در زمان شده است. جدول ۲ متوسط زمان پردازش جهت آشکارسازی لوگو را به کمک روش پیشنهادی و روش [۶] نشان می‌دهد.



شکل ۱۰: عدم آشکارسازی لوگو توسط روش [۶]



شکل ۱۱: آشکارسازی موفق لوگو توسط روش پیشنهادی

جدول ۱: قوانین تصمیم‌گیری درخت طبقه‌بند تصمیم در هر گره

شماره گره	ویژگی	محدوده مقدار
گره اول	عرض چهارگوش عرض تصویر سند	[۰,۰۴۰۰,۳۹]
	ارتفاع چهارگوش ارتفاع تصویر سند	[۰,۰۳۲۰,۲۲]
گره دوم	نسبت طول به عرض چهارگوش	[۰,۳۹۰۴,۹۵]

جدول ۲: مقایسه عملکرد روش پیشنهادی و روش [۶]

روش	صحت آشکارسازی	دقت آشکارسازی
روش پیشنهادی	٪ ۸۲	٪ ۹۴
روش [۶]	٪ ۸۰/۴	٪ ۹۳/۳

جدول ۲: مقایسه زمان پردازش دو روش برای آشکارسازی لوگو

زمان پردازش (ms)	روش پیشنهادی	روش [۶]
	۳۷۵۰	۷۲۰۰



- [۱۲] J. Neumann, H. Samet, and A. Soffèr, "Integration of local and global shape analysis for logo classification", *Pattern Recognition Letters*, Vol. ۲۳, No. ۱۲, pp. ۱۴۴۹-۱۴۵۷, ۲۰۰۲.
- [۱۳] Tobacco^{۸۰۰} complex Document image database,
<http://www.umiac.umd.edu/~zhugy/Tobacco^{۸۰۰}.html.ms>.