

# یک روش مستقل از تباین برای دودویی کردن تصاویر متنی

مرتضی ولیزاده و احسان‌اله کبیر

روش‌های ارائه شده برای دودویی کردن تصاویر متنی را می‌توان به دو گروه عمده روش‌های کلی و روش‌های محلی تقسیم کرد. در روش‌های کلی با استفاده از ویژگی‌های همه نواحی تصویر، یک آستانه مشترک برای همه پیکسل‌های آن تصویر استخراج می‌شود و با مقایسه روشنایی پیکسل‌های تصویر با آستانه به دست آمده، تصویر دودویی تولید می‌شود. روش آستانه‌یابی بر مبنای خوشه‌یابی [۱]، کمینه‌کردن آنتروپی [۲] و یافتن دره هیستوگرام [۳] معروف‌ترین روش‌های آستانه‌یابی عمومی هستند. با توجه به این که در بعضی از قسمت‌های تصویر آسیب دیده، پس‌زمینه تیره‌تر از نواحی متنی موجود در سایر قسمت‌های همان تصویر است، آستانه‌یابی عمومی برای دودویی کردن تصاویر متنی آسیب دیده مناسب نیست و برای این تصاویر عموماً از روش‌های آستانه‌یابی محلی استفاده می‌شود.

در روش‌های آستانه‌یابی محلی، هر پیکسل با توجه به ویژگی‌های همسایه‌هایش برچسب متن یا پس‌زمینه می‌گیرد. در بعضی از این روش‌ها، فقط از سطوح خاکستری پیکسل‌های همسایه برای آستانه‌یابی استفاده می‌شود [۴] تا [۷] و دسته دیگر از آنها از ویژگی‌های عرض قلم [۸] تا [۱۱] و ساختار لبه‌ها [۱۲] تا [۱۴] برای بهبود کیفیت آستانه‌یابی بهره می‌برند. اخیراً از روش‌های یادگیری ماشین نیز برای آستانه‌یابی محلی تصاویر متنی استفاده می‌شود. در آنها از هر پیکسل یک بردار ویژگی استخراج شده و به یک طبقه‌بند آموزش دیده داده می‌شود و طبقه‌بند آن را به عنوان متن یا پس‌زمینه طبقه‌بندی می‌کند [۱۵]. ارزیابی روش‌های آستانه‌یابی محلی نشان می‌دهد [۱۶] و [۱۷] که هیچ یک از آنها نیز برای دودویی کردن همه انواع تصاویر آسیب دیده مناسب نیستند و هر یک از روش‌های ارائه شده برای آسیب‌های خاصی کارایی دارند. همچنین اکثر روش‌های ارائه شده دارای چندین پارامتر از پیش تعیین شده هستند که قابلیت سازگاری با شرایط مختلف تباین را ندارند. در این قسمت چند نمونه از روش‌های آستانه‌یابی محلی معروف و مشکلات آنها را بررسی می‌کنیم.

در [۷] الگوریتمی مطابق (۱) برای دودویی کردن تصاویر متنی ارائه شده است که سطح آستانه به‌طور محلی و با جابه‌جایی یک پنجره روی تصویر تعیین می‌شود. آستانه  $T$  برای پیکسل واقع در مرکز پنجره، با محاسبه میانگین و واریانس سطوح روشنایی پیکسل‌های واقع در داخل پنجره تعیین می‌شود

$$T = m + ks \quad (1)$$

که در آن  $m$  و  $s$  به ترتیب متوسط و واریانس روشنایی پیکسل‌های داخل پنجره را نشان می‌دهند.  $k$  یک مقدار ثابت است که برای تصاویر متنی ۰/۲- در نظر گرفته می‌شود. این الگوریتم در نزدیکی نواحی متنی به‌طور مؤثر متن را استخراج می‌کند ولی در نواحی‌ای که متن وجود ندارد، نویز پس‌زمینه را تقویت می‌کند به طوری که بسیاری از پیکسل‌های پس‌زمینه به‌عنوان متن در نظر گرفته می‌شوند.

در [۵]، (۱) با در نظر گرفتن سطوح روشنایی نواحی متنی و پس‌زمینه به‌صورت (۲) اصلاح شده است

چکیده: در این تحقیق یک روش مستقل از تباین برای دودویی کردن تصاویر متنی آسیب دیده ارائه می‌شود. این روش به تنظیم پارامتر توسط کاربر نیاز ندارد و برای دودویی کردن تصاویر متنی با آسیب‌های تباین کم و غیر یکنواختی روشنایی پس‌زمینه و متن مناسب است. آستانه‌یابی با روش ارائه شده شامل سه مرحله است. در مرحله اول، مستقل از تباین متن و پس‌زمینه، قسمت‌های بارز هر یک از حروف با روش بارش باران اصلاح شده استخراج می‌شود. الگوریتم بارش باران اصلاح شده برای استخراج قسمت‌های بارز حروف طراحی شده است و مشکلات روش بارش باران را شامل نمی‌شود. در مرحله دوم با استفاده از نواحی متنی استخراج شده، روشنایی نواحی متنی به‌طور محلی تخمین زده می‌شود. همچنین با توجه به این که در تصاویر متنی تعداد پیکسل‌های متن در مقایسه با پیکسل‌های پس‌زمینه ناچیز است، روشنایی پس‌زمینه با میانگین‌گیری از روشنایی تصویر اصلی به‌طور محلی تخمین زده می‌شود. در مرحله سوم برای هر پیکسل، حد آستانه با استفاده از تخمین روشنایی پس‌زمینه و متن محاسبه می‌شود. این روش آستانه‌یابی به‌طور بصری و کمی با چهار روش متداول آستانه‌یابی مقایسه شده است و نتایج ارزیابی نشان می‌دهد که روش پیشنهادی برای دودویی کردن تصاویر متنی که با دوربین گرفته شده‌اند نسبت به روش‌های متداول برتری دارد و نواحی متنی با تباین کم را به خوبی استخراج می‌کند.

کلید واژه: آستانه‌یابی محلی، تصاویر متنی آسیب دیده، مدل جریان آب، تخمین روشنایی، مستقل از تباین.

## ۱- مقدمه

در سال‌های اخیر تحقیقات زیادی در زمینه تحلیل اسناد متنی انجام شده است. به‌طور کلی یک سیستم تحلیل اسناد متنی از چهار بخش اصلی اخذ تصویر، دودویی کردن، قطعه‌بندی تصویر و بازشناسی متن تشکیل می‌شود. کیفیت دودویی کردن تصویر کارایی مراحل بعدی را تحت تأثیر قرار می‌دهد، بنابراین این بخش یکی از مراحل مهم در تحلیل اسناد متنی محسوب می‌شود. با وجود این که تحقیقات زیادی در زمینه دودویی کردن تصاویر متنی انجام شده است، هنوز دودویی کردن تصویر متنی در شرایطی که روشنایی پس‌زمینه و متن در نواحی مختلف تصویر متفاوت است، جذابیت پژوهشی دارد. از سوی دیگر با پیشرفت دوربین‌های دیجیتال، تعداد تصاویر متنی که توسط این دوربین‌ها تهیه می‌شوند به سرعت افزایش یافته است. این تصاویر دارای آسیب‌هایی از قبیل غیر یکنواخت بودن روشنایی پس‌زمینه و نواحی متنی، تباین کم و متغیر بین متن و پس‌زمینه و نویز دوربین هستند و بسیاری از روش‌های متداول برای دودویی کردن آنها مناسب نیستند.

این مقاله در تاریخ ۲۳ فروردین ماه ۱۳۸۸ دریافت و در تاریخ ۱ دی ماه ۱۳۸۸ بازنگری شد. این پژوهش با استفاده از پشتیبانی مرکز تحقیقات مخابرات ایران بر اساس قرارداد شماره ۵۰۰/۴۶۸۳/ت انجام شده است.

مرتضی ولیزاده، بخش مهندسی برق، دانشگاه تربیت مدرس، تهران، کدپستی ۱۴۱۱۵-۱۴۳ (email: valizadeh@modares.ac.ir).

احسان‌اله کبیر، بخش مهندسی برق، دانشگاه تربیت مدرس، تهران، کدپستی ۱۴۱۱۵-۱۴۳ (email: kabir@modares.ac.ir).

و انحراف معیار تولید می‌شود. سپس با استفاده از چندین قاعده ساده، لبه‌های متن تشخیص داده می‌شود. این روش برای حذف پس‌زمینه‌هایی با تغییرات سریع مناسب نیست و همچنین به سه پارامتر از پیش تعیین شده نیاز دارد که به ویژگی‌های تصویر وابسته هستند. در [۲۰] از این دانش که نواحی متنی به‌طور محلی تیره‌تر از نواحی پس‌زمینه هستند، استفاده شده و یافتن پیکسل‌های متنی به یک مسئله یافتن کمینه‌های محلی تبدیل شده است. در این روش مقادیر روشنایی تصویر مانند پستی و بلندی زمین در نظر گرفته شده و از مدل بارش باران<sup>۱</sup> برای دودویی کردن تصویر متنی استفاده شده است. این روش به‌طور محلی نواحی متنی را از همسایه‌های آنها متمایز می‌کند ولی چندین مشکل اساسی دارد. مشکل اصلی آن نبودن معیاری برای توقف بارش باران است. مشکل دوم یکسان نبودن تعداد بهینه مراحل بارش برای نواحی مختلف تصویر است. همچنین حجم محاسبات این روش خیلی زیاد است.

هدف اصلی مقاله ارائه روشی وقتی برای دودویی کردن تصاویر متنی آسیب‌دیده و اصلاح آسیب‌های آنهاست. با توجه به این که اطلاعی درباره روشنایی پس‌زمینه و نواحی متنی و تباین بین آنها نداریم، در الگوریتم پیشنهادی ابتدا با استفاده از یک روش مستقل از تباین، قسمت‌های بارز هر یک از حروف استخراج می‌شود و روشنایی نواحی متنی تخمین زده می‌شود و نیز با در نظر گرفتن این که تعداد پیکسل‌های پس‌زمینه خیلی بیشتر از پیکسل‌های متنی است، با میانگین‌گیری روشنایی پیکسل‌های تصویر اصلی، روشنایی پس‌زمینه به‌طور محلی تخمین زده می‌شود. سپس آستانه‌ای به‌صورت میانگین روشنایی نواحی متنی و پس‌زمینه محاسبه شده و برای دودویی کردن تصویر استفاده می‌شود. با توجه به این که الگوریتم پیشنهادی به‌طور محلی و مستقل از تباین، آستانه‌ای با فاصله بیشینه از سطح روشنایی متن و پس‌زمینه را محاسبه می‌کند، برای دودویی کردن تصاویر متنی با تباین کم و متغیر و تصاویری که روشنایی متن و پس‌زمینه در آنها متغیر است، مناسب است. این الگوریتم به تنظیم پارامتر توسط کاربر نیاز ندارد و برای انواع گوناگونی از تصاویر متنی آسیب‌دیده مناسب است.

این مقاله از ۶ بخش تشکیل شده است. در بخش ۲ روشی مستقل از تباین برای استخراج قسمت‌هایی از نواحی متنی ارائه می‌شود. بخش ۳ به تخمین روشنایی نواحی متنی و پس‌زمینه می‌پردازد. در بخش ۴ چگونگی استخراج آستانه محلی و دودویی کردن تصویر متنی تشریح می‌شود. در بخش ۵ الگوریتم پیشنهادی ارزیابی شده و بخش ۶ به نتیجه‌گیری اختصاص داده می‌شود.

## ۲- استخراج قسمت‌های بارز نواحی متنی

در این مرحله قصد داریم مستقل از تباین بین متن و پس‌زمینه و روشنایی آنها، به‌طور محلی قسمت‌های بارز نواحی متنی را استخراج کنیم. اکثر روش‌های محلی برای دودویی کردن تصویر متنی، دارای پارامترهایی هستند که به تباین متن و پس‌زمینه وابسته هستند و برای این کاربرد مناسب نیستند، ولی با اصلاح روش بارش باران می‌توان از آن به‌عنوان روشی مستقل از تباین برای استخراج قسمت‌های بارز نواحی متنی استفاده کرد. برای بیان روش بارش باران اصلاح‌شده، ابتدا به تشریح الگوریتم بارش باران پایه می‌پردازیم و سپس تغییرات ایجادشده در آن را شرح می‌دهیم.

$$T = m \left( 1 - k \left( 1 - \frac{s}{R} \right) \right) \quad (2)$$

که در آن مقدار  $k$  برابر با ۰.۵ و مقدار  $R$  برای تصاویر با ۲۵۶ سطح خاکستری، ۱۲۸ در نظر گرفته شده است. با این تدبیر برای تصاویر متنی با تباین خوب نتایج خوبی به‌دست می‌آید ولی این روش به تباین بین متن و پس‌زمینه حساسیت زیادی دارد و نواحی متنی با تباین کم از بین می‌روند. برای این که این روش برای تصاویر مختلف کارایی خوبی داشته باشد، باید پارامترهای  $k$  و  $R$  با توجه به ویژگی‌های تصویر به‌طور محلی تنظیم شوند.

در [۱۸] برای تنظیم وقتی پارامترهای روش‌های آستانه‌یابی به‌صورت زیر عمل شده است. ابتدا بر اساس یک ویژگی که معرف تباین ناحیه‌ای است، تصویر اصلی به تصاویر کوچک‌تر تقسیم می‌شود. سپس با استخراج سه ویژگی از هر ناحیه و طبقه‌بندی آن، یکی از برچسب‌های پررنگ، کم‌رنگ، پررنگ-کم‌رنگ و پس‌زمینه به آن ناحیه زده می‌شود و بر اساس برچسب هر ناحیه پارامترهای مناسبی به آن اختصاص داده می‌شود. در این مرجع از ویژگی‌های متوسط روشنایی و اندازه گرادینان برای تعیین آستانه استفاده شده است و طبق (۳) پارامترها تنظیم می‌شوند

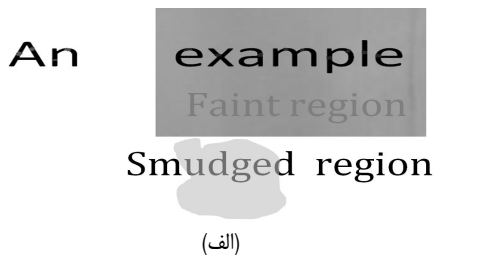
$$T(x, y) = wM(x, y) + kG_N(x, y)$$

$$\text{where } \begin{cases} \text{if background,} & w = 0, k = 0 \\ \text{if faint,} & w = 0.5, k = -0.1 \\ \text{if heavy,} & w = 0.7, k = -0.8 \\ \text{if heavy with faint,} & w = 0.7, k = -0.1 \end{cases} \quad (3)$$

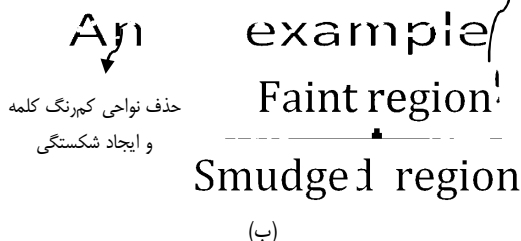
در [۶] از ترکیب آستانه‌یابی کلی و محلی برای دودویی کردن تصویر متنی استفاده شده است. در این روش اگر فاصله بین روشن‌ترین و تیره‌ترین پیکسل داخل یک پنجره محلی از یک آستانه از پیش تعیین شده بیشتر باشد، متوسط آنها به‌عنوان حد آستانه برای پیکسل واقع در مرکز آن پنجره در نظر گرفته می‌شود. در غیر این صورت از آستانه کلی برای دودویی کردن آن پیکسل استفاده می‌شود. این روش برای حالتی که تغییرات روشنایی پس‌زمینه در داخل یک پنجره زیاد باشد، کارایی ندارد و قسمت‌هایی از پس‌زمینه به‌عنوان متن در نظر گرفته می‌شود. همچنین اگر تباین بین متن و پس‌زمینه از حد آستانه از پیش تعیین شده کمتر باشد، متن به احتمال زیاد به‌عنوان پس‌زمینه در نظر گرفته شده و حذف می‌شود.

در [۴] ابتدا با استفاده از آستانه‌یابی و درون‌یابی، روشنایی پس‌زمینه تخمین زده می‌شود و برای هر پیکسل اختلاف روشنایی آن در پس‌زمینه تخمینی و تصویر اصلی محاسبه می‌شود. اگر این اختلاف از یک حد آستانه وقتی بیشتر باشد آن پیکسل به‌عنوان متن در نظر گرفته می‌شود و در غیر این صورت به‌عنوان پس‌زمینه برچسب زده می‌شود. در این روش، آستانه‌یابی و تخمین پس‌زمینه برای حالتی که تباین متن و پس‌زمینه کم است به خوبی انجام نمی‌شود و نواحی متنی از بین می‌روند. همچنین تعیین آستانه وقتی در این روش به سه پارامتر از پیش تعیین شده نیاز دارد که تنظیم مناسب این پارامترها وابسته به تصویر است.

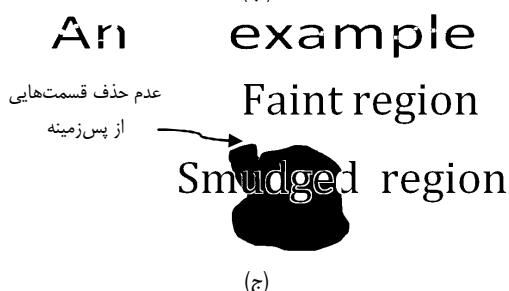
در [۱۹] از این قاعده استفاده شده است که در لبه‌های اشیا فرکانس محلی بالاست. ابتدا لبه‌های اشیا تشخیص داده می‌شوند و سپس نواحی بین این لبه‌ها پر می‌شوند. برای استخراج لبه‌های نواحی متنی، یک تصویر کمکی با در نظر گرفتن همسایه‌های هشت‌تایی هر پیکسل ساخته می‌شود و برای نواحی کوچک تصویر، متوسط و انحراف معیار روشنایی تصویر کمکی محاسبه می‌شود. با درون‌یابی این مقادیر دو تصویر میانگین



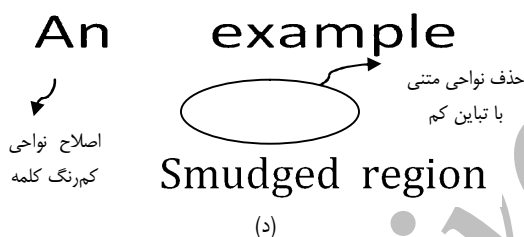
(الف)



(ب)



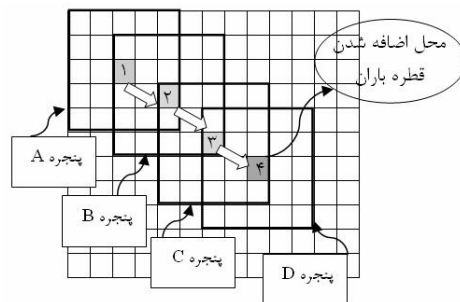
(ج)



(د)

شکل ۲: (الف) تصویر آسیب‌دیده مصنوعی، (ب) تصویر دودویی بعد از ۵ مرحله بارش، (ج) تصویر دودویی بعد از ۱۵ مرحله بارش و (د) تصویر دودویی بعد از ۲۵ مرحله بارش.

و همچنین برای بررسی اثر تغییرات سریع روشنایی پس‌زمینه، یک لکه بر روی نوشته‌ها ایجاد شده است. برای نشان دادن حذف قسمت‌های کم‌رنگ حروف و ایجاد شکستگی در آنها، قسمت‌هایی از حروف کم‌رنگ شده است. در مثال شکل ۲ مشاهده می‌شود که در روش بارش باران نه تنها معیاری برای توقف بارش باران وجود ندارد، برای تصاویری که ویژگی‌های آن در نواحی مختلف تصویر متفاوت است، تعداد مناسب تکرار الگوریتم بارش برای نواحی مختلف تصویر متفاوت است و اگر به درستی تعیین نشود ممکن است باعث حذف نواحی متنی شود و یا قسمت‌هایی از پس‌زمینه به‌عنوان متن معرفی شوند. برای مثال در شکل ۲-ب تعداد تکرار الگوریتم بارش برای استخراج عبارت "Faint Region" مناسب است ولی برای حذف کامل پس‌زمینه کافی نیست و در شکل ۲-د تعداد تکرار الگوریتم بارش برای حذف کامل پس‌زمینه و اصلاح شکستگی‌های عبارت "An example" مناسب است ولی برای استخراج عبارت "Faint Region" مناسب نیست و باعث حذف آن می‌شود. با توجه به این که عملکرد روش پیشنهادی برای دودویی کردن تصاویر متنی وابستگی زیادی به کیفیت حذف پس‌زمینه و حفظ قسمت‌هایی از هر حرف دارد، باید روش بارش باران به‌نحوی اصلاح شود که قسمت‌هایی از هر حرف را استخراج و پس‌زمینه را به‌طور کامل حذف کند و تا حد ممکن مستقل از تباين بین متن و پس‌زمینه باشد. البته فرض می‌کنیم که تباين



شکل ۱: فرایند جستجو برای یافتن کمینه محلی.

## ۱-۲ روش بارش باران

در روش بارش باران [۲۰]، روشنایی هر پیکسل از تصویر متنی به‌عنوان ارتفاع یک نقطه از سطح زمین در نظر گرفته می‌شود و بارش باران بر روی همه نقاط سطح زمین انجام می‌شود. قطره‌های باران به سمت نواحی با ارتفاع کمتر حرکت می‌کنند و در آن نواحی جمع می‌شوند. شکل ۱ چگونگی حرکت قطره باران برای یافتن نقطه کمینه محلی و قرار گرفتن در آن را نشان می‌دهد.

برای بیان چگونگی حرکت قطره باران از مثال شکل ۱ استفاده می‌کنیم. فرض کنید که قطره باران در نقطه ۱ فرود آید، ارتفاع نقاط داخل پنجره‌ای به ابعاد  $W \times W$  و مرکزیت نقطه ۱ بررسی می‌شود. اگر نقطه ۱ کمینه محلی باشد، قطره به همان نقطه اضافه می‌شود و در غیر این صورت، مرکز پنجره جستجو به نقطه کمینه محلی انتقال داده می‌شود (نقطه ۲) و کمینه محلی جدید در داخل پنجره جدید جستجو می‌شود. این الگوریتم تا وقتی که مرکز پنجره به‌عنوان کمینه محلی معرفی شود (نقطه ۴) ادامه پیدا می‌کند. در مثال ارائه‌شده بعد از ۴ بار جستجوی کمینه محلی، مرکز پنجره به‌عنوان کمینه محلی معرفی شده و قطره باران به آن نقطه از زمین اضافه می‌شود. با اضافه‌شدن قطره باران به یک نقطه از زمین، ارتفاع آن به اندازه یک واحد افزایش پیدا می‌کند. با بارش باران بر روی همه نواحی، آب در نواحی پست جمع شده و باعث افزایش ارتفاع آن نواحی می‌شود. فرایند بارش به تعداد مشخصی که از پیش تعیین شده است، تکرار و سپس متوقف می‌شود.

پس از توقف بارش باران، چاله‌های آب استخراج و متوسط ارتفاع سطح آب هر چاله به کل نقاط آن اختصاص داده می‌شود و ارتفاع سطح هر نقطه پس از بارش باران به‌دست می‌آید. با کم کردن ارتفاع سطح اولیه هر نقطه از ارتفاع سطح آن پس از بارش باران، مقدار آب جمع‌شده بر روی هر نقطه از سطح زمین به‌دست می‌آید. با دودویی کردن مقدار آب جمع‌شده بر روی سطح زمین، تصویر دودویی حاصل می‌شود. تعداد مراحل بارش باران مهم‌ترین پارامتر این الگوریتم است که متأسفانه روشی برای تعیین آن ارائه نشده است.

## ۲-۲ مشکلات روش بارش باران

در این روش اگر تعداد تکرار الگوریتم بارش باران کم باشد، قسمت‌های کم‌رنگ حروف حذف می‌شوند و در حروف شکستگی ایجاد می‌شود و همچنین قسمت‌هایی از پس‌زمینه به‌عنوان متن در نظر گرفته می‌شوند. اگر تعداد مراحل بارش باران زیاد باشد، ممکن است بعضی از نواحی متنی که تباين آنها با پس‌زمینه کم است، حذف شوند. برای نشان دادن مشکلات بیان‌شده از شکل ۲-الف استفاده می‌کنیم و به‌ازای مقادیر مختلف تکرار الگوریتم بارش باران، تصویر دودویی را تولید می‌کنیم. در این تصویر مصنوعی نوشته‌های پررنگ و کم‌رنگ وجود دارد



(الف)



(ب)

شکل ۴: ناحیه‌ای از پس‌زمینه که آب آن به ناحیه متنی کشیده شده است. (الف) ناحیه جستجوی مربعی و (ب) ناحیه جستجوی پیشنهادی.

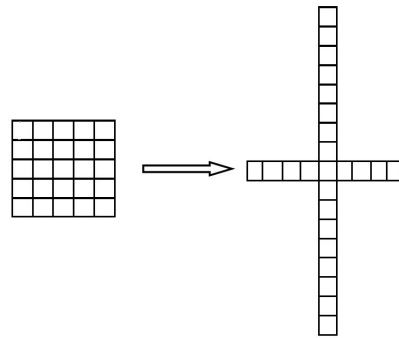
دارد و با توجه به نزدیکی حروف در راستای افقی، ناحیه جستجو در این راستا کشیدگی کمتری دارد. در شکل ۴ مساحت قسمتی از پس‌زمینه که آب آن به ناحیه متنی سرازیر می‌شود، برای ناحیه جستجوی مربعی و ناحیه جستجوی پیشنهادی نشان داده شده است.

در شکل ۴ رنگ خاکستری نشان‌دهنده نواحی پس‌زمینه است که آب آنها به نواحی متنی سرازیر شده است. مشاهده می‌شود که با پنجره جستجوی پیشنهادی، مساحت ناحیه‌ای از پس‌زمینه که در رقابت با ناحیه متنی قرار می‌گیرد افزایش پیدا کرده و باعث افزایش میزان آب جمع‌شده در نواحی متنی و در نتیجه افزایش تمایز بین متن و پس‌زمینه می‌شود.

### ۲-۳-۳ اختصاص میزان آب یکسان به همه نقاط یک چاله

در الگوریتم بارش باران به همه نواحی یک چاله سطح آب یکسان اختصاص داده می‌شود، بنابراین پس از کم کردن ارتفاع سطح اولیه از ارتفاع سطح پس از بارش باران، مقدار آب در نواحی مختلف یک چاله متفاوت خواهد بود. این امر باعث می‌شود تا قسمت‌هایی از پس‌زمینه که تیره‌تر از همسایه‌های خود هستند، به‌عنوان متن در نظر گرفته شوند. ولی ما به همه نواحی یک چاله، متوسط آب آن چاله را اختصاص می‌دهیم. این کار باعث می‌شود تا میزان آب نواحی تیره پس‌زمینه با میزان آب سایر نواحی پس‌زمینه که با نواحی تیره چاله مشترکی تشکیل می‌دهند، متوسط‌گیری و تعدیل شود. در نتیجه نواحی تیره پس‌زمینه در مرحله آستانه‌یابی به‌عنوان متن در نظر گرفته نمی‌شوند. شکل ۵ تفاوت روش پیشنهادی را با روش پایه بارش باران به‌طور آشکار نشان می‌دهد. پس از آستانه‌یابی یک تفاوت عمده بین روش پیشنهادی و روش پایه بارش باران ایجاد می‌شود. در روش پیشنهادی به همه نقاط یک چاله برچسب یکسان متن یا پس‌زمینه زده می‌شود ولی در روش پایه ممکن است بعضی از نقاط یک چاله برچسب متن و سایر نقاط آن برچسب پس‌زمینه بگیرند. ممکن است این نگرانی به‌وجود آید که اگر نواحی متنی با قسمتی از پس‌زمینه چاله‌های مشترک تشکیل دهند، با الگوریتم پیشنهادی حذف خواهند شد. خوشبختانه چنین حالتی تقریباً هیچ وقت به‌وجود نمی‌آید. دلیل عدم اتصال چاله‌های متنی به پس‌زمینه را توأم با چگونگی یافتن چاله‌ها بررسی می‌کنیم. برای یافتن چاله‌ها به‌صورت زیر عمل می‌شود.

با کم کردن ارتفاع سطح اولیه از ارتفاع سطح پس از بارش باران، مقدار آب جمع‌شده در هر نقطه به‌دست می‌آید که آن را با  $WA^1$  نشان می‌دهیم



شکل ۳: تغییر شکل ناحیه جستجو.

بین متن و پس‌زمینه از ۱۰ سطح خاکستری بیشتر است که برای تصاویر با ۲۵۶ سطح خاکستری معمولاً این فرض برقرار است.

## ۲-۳ اصلاح روش بارش باران

همان‌طور که در شکل ۲-ب نشان داده شد، اگر تعداد مراحل بارش باران کم باشد، نواحی متنی با تباین کم حفظ می‌شوند ولی قسمت‌هایی از پس‌زمینه به‌عنوان متن معرفی می‌شوند. بنابراین در این مقاله برای جلوگیری از حذف نواحی متنی با تباین کم و افزایش سرعت، فقط از یک مرحله بارش باران برای استخراج نواحی متنی استفاده می‌شود و برای حذف پس‌زمینه تمهیدات زیر در نظر گرفته می‌شود.

### ۲-۳-۱ هموارسازی تصویر متنی

در روش‌های دودویی کردن تصاویر متنی نمی‌توان از فیلترهای پایین‌گذر برای حذف نویز استفاده کرد، زیرا این کار باعث مات‌شدن لبه‌های حروف و چسبیدن آنها به یکدیگر می‌شود. ولی با توجه به این که هدف ما در این مرحله دودویی کردن تصویر متنی نیست و فقط می‌خواهیم قسمت‌هایی از متن را استخراج کنیم، با استفاده از یک پنجره  $5 \times 5$  تصویر متنی را هموار می‌کنیم و از آن به‌عنوان یک تصویر کمکی برای استخراج قسمت‌های بارز متن استفاده می‌کنیم. برای تولید تصویر دودویی نهایی از تصویر فیلترنشده استفاده خواهیم کرد. ابعاد پنجره فیلتر هموارساز با در نظر گرفتن عرض قلم در مجموعه تصاویر انتخاب شده است. ابعاد این فیلتر باید از عرض قلم کوچک‌تر باشد تا حداقل قسمت‌هایی از هر حرف وجود داشته باشد که سطوح خاکستری آنها تحت تأثیر سطوح خاکستری پس‌زمینه قرار نمی‌گیرند. در تصاویر استفاده‌شده عرض قلم بین ۶ تا ۸ متغیر است.

### ۲-۳-۲ تغییر شکل ناحیه جستجو

برای افزایش میزان آبی که در یک مرحله بارش باران از نواحی پس‌زمینه به داخل نواحی متنی سرازیر می‌شود، ناحیه جستجوی کمینه محلی را مطابق شکل ۳ تغییر می‌دهیم.

با این تغییر شکل ناحیه جستجو، تعداد مقایسه‌ها برای یافتن کمینه محلی مانند حالت قبل، ۲۵ عدد باقی می‌ماند ولی میزان آبی که با یک مرحله بارش باران به نواحی متنی سرازیر می‌شود افزایش پیدا می‌کند و باعث افزایش تمایز بین نواحی متنی و پس‌زمینه می‌شود.

ناحیه جستجو برای یافتن کمینه محلی به‌نحوی انتخاب شده است که نواحی متنی برای جذب آب با نواحی پس‌زمینه بیشتری رقابت کنند ولی با نواحی متنی مجاور رقابتی نداشته باشند تا این که یک حرف پررنگ در رقابت با حروف کم‌رنگی که در مجاورت آن قرار دارند، باعث حذف آنها نشود. با توجه به این که فاصله بین خطوط متن از هم زیاد است (حدود ۵۰ تا ۱۰۰ پیکسل)، ناحیه جستجو در راستای عمودی کشیدگی زیادی

## ۲-۳-۴ آستانه‌یابی و استخراج نواحی بارز متن

با آستانه‌یابی مقدار آب به‌دست آمده با الگوریتم پیشنهادی،  $PWA^1$  تصویر دودویی به‌دست می‌آید که در آن قسمت‌های بارز نواحی متنی استخراج شده و پس‌زمینه با دقت خوبی حذف می‌شود. باید به این نکته توجه داشت که مقدار آبی که با الگوریتم پیشنهادی به‌دست می‌آید از لحاظ فیزیکی تعبیر خاصی ندارد و با میزان آب واقعی که بر روی سطح زمین جمع می‌شود متفاوت است.

در روش پایه بارش باران از الگوریتم آستانه‌یابی کلی استفاده شده است که بر اساس کمینه‌کردن تغییرات درون کلاسی کار می‌کند [۱]. این روش آستانه‌یابی برای الگوریتم پایه بارش باران به‌خوبی کارآمد است، زیرا تعداد مراحل بارش باران زیاد است و بین نواحی متنی و پس‌زمینه به اندازه کافی تمایز ایجاد شده است. ولی برای روش پیشنهادی که فقط از یک مرحله بارش باران استفاده می‌کند، در مواقعی که مساحت نواحی متنی نسبت به پس‌زمینه خیلی کم باشد ممکن است حد آستانه را کمتر از مقدار بهینه تخمین بزند و به تبع آن قسمتی از پس‌زمینه به‌عنوان متن در نظر گرفته شود. برای جلوگیری از چنین حالت‌هایی از دانش اولیه‌ای که درباره میزان آب نواحی متنی و پس‌زمینه داریم، استفاده می‌کنیم.

نواحی متنی از دو طرف با پس‌زمینه در ارتباط هستند و تجربه نشان می‌دهد که مساحت نواحی پس‌زمینه که آب آنها به نواحی متنی سرازیر می‌شود، از دو برابر نواحی متنی بیشتر است. بنابراین متوسط مقدار آب جمع‌شده در نواحی متنی پس از یک بار بارش باران، از سه واحد بیشتر می‌شود در حالی که مقدار متوسط آب در نواحی پس‌زمینه کمتر از این مقدار است. این دانش می‌تواند طبق (۵) با مقدار آستانه کلی ترکیب شود و نواحی متنی را استخراج کند. شرط  $\max(3, Th)$  در (۵) باعث می‌شود تا در مواقعی که آستانه کمتر از حد بهینه تخمین زده می‌شود از دانش اولیه به‌عنوان آستانه استفاده شود

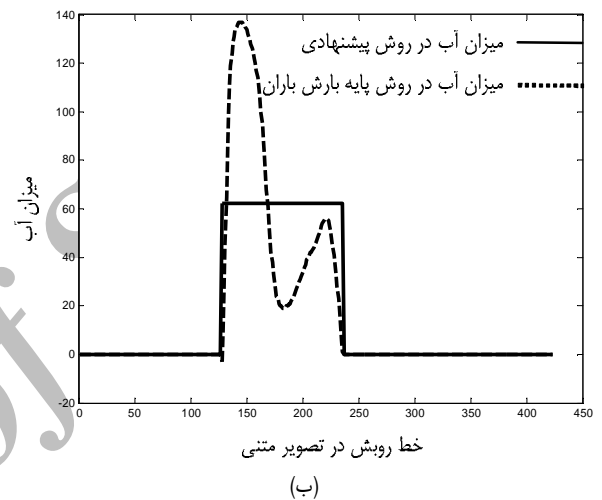
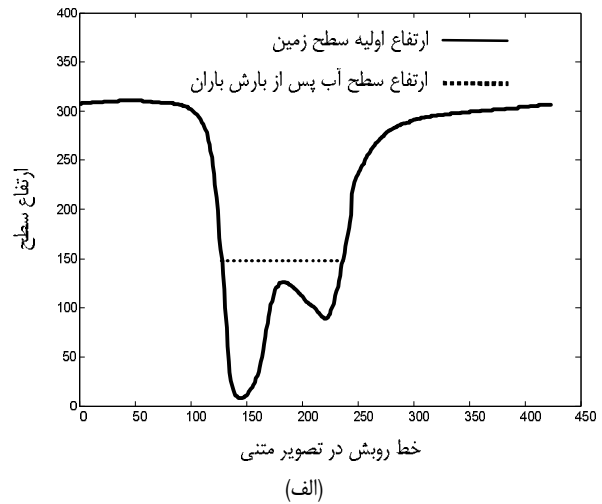
$$PTR(x, y) = \begin{cases} 1 & PWA(x, y) \geq \max(3, Th) \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

در این رابطه  $PTR$  بیانگر تصویری است که در آن قسمت‌های بارز نواحی متنی استخراج شده است و  $Th$  معرف آستانه کلی به‌دست آمده با روش [۱] است.

شکل ۶ نتایج استخراج قسمت‌های بارز نواحی متنی تصویر شکل ۲-الف را برای روش‌های بارش باران پایه و پیشنهادی پس از یک مرحله بارش نشان می‌دهد. مشاهده می‌شود که روش بارش باران پیشنهادی پس از یک مرحله بارش باران به‌خوبی نواحی متنی را استخراج و پس‌زمینه را حذف می‌کند. در حالی که روش بارش باران پایه با یک مرحله بارش قادر به استخراج متن و حذف پس‌زمینه نیست. پس از استخراج قسمت‌های بارز هر حرف، از آنها برای تخمین روشنایی نواحی متنی استفاده می‌شود.

## ۳- تخمین محلی روشنایی متن و پس‌زمینه

هدف اصلی الگوریتم ارائه‌شده برای استخراج قسمت‌های بارز نواحی متنی، تخمین روشنایی متن است. برای این منظور به‌صورت زیر عمل می‌شود.



شکل ۵: (الف) ارتفاع اولیه سطح زمین و سطح آب پس از بارش باران و (ب) مقایسه میزان آب نقاط داخل یک چاله با روش پیشنهادی و روش پایه بارش باران.

و تصویر دودویی کمکی  $B$  را برای یافتن چاله‌ها مطابق (۴) تولید می‌کنیم

$$B(x, y) = \begin{cases} 1 & WA(x, y) > 0 \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

که در آن  $x$  و  $y$  نشان‌دهنده مختصات پیکسل است. با برچسب زدن به عناصر متصل به هم در تصویر  $B$ ، چاله‌های موجود در تصویر استخراج می‌شوند. حال به بررسی این موضوع می‌پردازیم که تحت چه شرایطی ممکن است چاله متناظر با نواحی متنی به نواحی پس‌زمینه متصل شود.

اگر در نقاط متناظر با پیکسل‌های پس‌زمینه که در همسایگی ناحیه متنی قرار دارند حداقل یک قطره آب جمع شود، چاله متناظر با نواحی متنی ممکن است به پس‌زمینه متصل شود. با توجه به این که این نقاط برای جذب آب در رقابت با نواحی متنی هستند، فقط زمانی آب بر روی آنها جمع می‌شود که چاله‌های متنی لبریز شده باشند. اما به دلیل این که فقط یک بار بارش باران انجام می‌شود. احتمال لبریز شدن نواحی متنی تقریباً صفر است. زیرا طبق فرض مسئله، اختلاف سطح ارتفاع نواحی متنی و پس‌زمینه به‌طور محلی بیشتر از ۱۰ واحد است و تجربه نشان می‌دهد که با یک بار بارش چاله‌هایی به عمق ۱۰ واحد هرگز لبریز نمی‌شوند.

دلیل این که در روش پایه بارش باران نمی‌توان به همه نواحی چاله‌ها مقدار آب یکسانی اختصاص داد، این است که در آن روش تعداد مراحل بارش باران زیاد است و چاله‌های متنی لبریز شده به پس‌زمینه متصل می‌شوند و از بین می‌روند.

1. Proposed Water Amount
2. Partial Text Region

### ۳-۱ پیاده‌سازی سریع الگوریتم‌های تخمین روشنایی

به‌منظور کاهش حجم محاسبات، از الگوریتم ارائه‌شده در [۲۱] برای تخمین روشنایی پس‌زمینه و متن استفاده می‌کنیم. در این مرجع از تصویر تجمعی برای محاسبه میانگین همسایگی‌ها در یک تصویر استفاده شده است. روشنایی تصویر تجمعی در هر پیکسل به‌صورت مجموع روشنایی پیکسل‌های بالا و سمت چپ آن پیکسل تعریف می‌شود و با (۸) به‌صورت ریاضی بیان می‌شود

$$I(x, y) = \sum_{i=-1}^x \sum_{j=-1}^y g(i, j) \quad (۸)$$

به‌دست آوردن تصویر تجمعی با یک بار رویش تصویر و با انجام سه عمل جمع به‌ازای هر پیکسل قابل اجراست و محاسبه متوسط روشنایی پیکسل‌های واقع در داخل پنجره‌ای به ابعاد  $(2w+1) \times (2w+1)$  طبق (۹) انجام می‌شود که مستقل از ابعاد پنجره، به‌ازای هر پیکسل ۳ عمل جمع و یک عمل تقسیم مورد نیاز است. در این رابطه  $M$  بیانگر متوسط روشنایی پیکسل‌های داخل پنجره است. این الگوریتم بدون هیچ تغییری برای تخمین روشنایی پس‌زمینه قابل اجراست. تخمین روشنایی پس‌زمینه برای هر پیکسل با ۶ عمل جمع و ۱ عمل تقسیم انجام می‌شود.

تخمین روشنایی متن اندکی متفاوت است زیرا فقط از روشنایی پیکسل‌های متنی استفاده می‌شود. برای پیاده‌سازی آن ابتدا با ضرب درایه به درایه تصاویر  $g$  و  $PRT$ ، تصویری حاصل می‌شود که روشنایی نواحی متنی بارز از تصویر اصلی جایگزین و روشنایی سایر پیکسل‌ها صفر می‌شود. حال با اعمال الگوریتم [۲۱] بر تصویر حاصل، مجموع روشنایی پیکسل‌های متنی داخل پنجره به‌دست می‌آید. به‌همین ترتیب با استفاده از تصویر  $PRT$ ، تعداد پیکسل‌های متنی داخل پنجره نیز به‌دست می‌آید. با تقسیم مجموع روشنایی پیکسل‌های متنی داخل پنجره بر تعداد آنها، روشنایی متن تخمین زده می‌شود. با این الگوریتم برای تخمین روشنایی متن برای هر پیکسل ۱۲ عمل جمع و ۱ عمل تقسیم مورد نیاز است.

### ۴- محاسبه آستانه و دودویی کردن تصویر

در این تحقیق برای هر پیکسل از تصویر با توجه به تخمین روشنایی نواحی متنی و پس‌زمینه، طبق (۱۰) یک آستانه جداگانه محاسبه می‌شود

$$T(x, y) = \begin{cases} \frac{BM(x, y) + FM(x, y)}{2} & FM(x, y) \geq 0 \\ 0 & \text{otherwise} \end{cases} \quad (۱۰)$$

بنابراین در آن نواحی که متن وجود دارد، متوسط روشنایی تخمینی پس‌زمینه و متن به‌عنوان آستانه در نظر گرفته می‌شود و در نواحی دور از متن، آستانه برابر با صفر قرار داده می‌شود تا به‌طور کامل پس‌زمینه را حذف کند.

تصویر دودویی با استفاده از (۱۱) تولید می‌شود

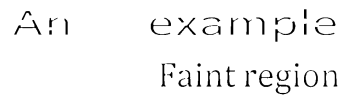
$$\text{Binary}(x, y) = \begin{cases} 1 & \text{if } g(x, y) \geq T(x, y) \\ 0 & \text{otherwise} \end{cases} \quad (۱۱)$$

که در آن Binary نشان‌دهنده تصویر دودویی نهایی است. در شکل ۷ تصویر دودویی حاصل از الگوریتم پیشنهادی برای تصویر شکل ۲-الف نشان داده شده است.

$$M(x, y) = \frac{I(x+w, y+w) - I(x+w, y-w) - I(x-w, y+w) + I(x-w, y-w)}{(2w+1) \times (2w+1)} \quad (۹)$$



Smudged region  
(الف)



Smudged region  
(ب)

شکل ۶: استخراج قسمت‌هایی از نواحی متنی تصویر شکل ۲-الف، (الف) با روش بارش باران پایه پس از یک مرحله بارش و (ب) با روش بارش باران پیشنهادی.

فرض کنید که روشنایی تصویر اصلی با  $g$  و روشنایی تخمینی برای نواحی متنی با  $FM^1$  نشان داده می‌شود و از ناحیه‌ای مربعی به ابعاد  $(2w_1+1) \times (2w_1+1)$  برای تخمین روشنایی متن در مرکز پنجره استفاده می‌شود. ابتدا وجود پیکسل متنی در داخل ناحیه مربعی بررسی می‌شود. اگر  $\sum_{i=-w_1}^{w_1} \sum_{j=-w_1}^{w_1} PRT(x+i, y+j) > 0$  برقرار باشد، نشان‌دهنده وجود پیکسل متنی در داخل ناحیه مربعی است و از (۶) برای تخمین روشنایی متن استفاده می‌شود

$$FM(x, y) = \frac{\sum_{i=-w_1}^{w_1} \sum_{j=-w_1}^{w_1} g(x+i, y+j) \times PRT(x+i, y+j)}{\sum_{i=-w_1}^{w_1} \sum_{j=-w_1}^{w_1} PRT(x+i, y+j)} \quad (۶)$$

نشان‌دهنده این است که در داخل ناحیه مربعی، پیکسل متنی وجود ندارد و روشنایی ناحیه متنی به‌صورت  $FM(x, y) = -1$  تعریف می‌شود. تفسیر (۶) این است که از متوسط روشنایی نواحی متنی داخل یک ناحیه کوچک به‌عنوان تخمین روشنایی متن استفاده می‌شود.

همچنین روشنایی پس‌زمینه که با  $BM^2$  نمایش داده می‌شود، با استفاده از (۷) تخمین زده می‌شود

$$BM(x, y) = \frac{\sum_{i=-w_1}^{w_1} \sum_{j=-w_1}^{w_1} g(x+i, y+j)}{(2w_1+1) \times (2w_1+1)} \quad (۷)$$

برای تخمین روشنایی پس‌زمینه از پیکسل‌های متنی و پس‌زمینه داخل یک پنجره با ابعاد  $(2w_1+1) \times (2w_1+1)$  استفاده می‌شود. دلیل پذیرفتن این تقریب آن است که تعداد پیکسل‌های متنی در مقایسه با پیکسل‌های پس‌زمینه خیلی کم است و تأثیر چندانی در تقریب پس‌زمینه ندارد.

در این مقاله  $w_1 = 25$  و  $w_1 = 50$  در نظر گرفته شده است. همان‌طور که از روابط مشخص می‌شود برای تخمین روشنایی پس‌زمینه هر پیکسل ۱۰۰۰۰ جمع و برای تخمین روشنایی متن در هر پیکسل حداقل ۲۵۰۰ جمع لازم است و با این حجم محاسبات، الگوریتم پیشنهادی کاربردی نخواهد بود و باید روشی برای کاهش میزان محاسبات پیدا شود.

1. Foreground Model
2. Background Model

er known as Old Faithful has never failed to erupt on time.  
 dard dialects are used as educational models in schools.  
 s never been such wide access to news from so many media.  
 ruction of a city on the ruins of a previous settlement is common.  
 rimitive mammals did not develop until the Triassic Period.

sting that fewer mental health disorders are experienced by people over the age  
 nged instruments, the tones are produced by a bow when it is played across

(الف)

1. The geyser known as Old Faithful has never failed to erupt on time.
2. Only standard dialects are used as educational models in schools.
3. There has never been such wide access to news from so many media.
4. The construction of a city on the ruins of a previous settlement is common.
5. The first primitive mammals did not develop until the Triassic Period.

(ب)

شکل ۸: دو نمونه از تصاویر استفاده شده برای ارزیابی با نرم افزار نویسه خوان، (الف) تصویر متنی با پس زمینه غیر یکنواخت و (ب) تصویر متنی با تباین کم.

تعداد حروفی که با استفاده از آستانه یابی کلی استخراج نمی شود در مقایسه با روش پیشنهادی خیلی زیاد است، روش پیشنهادی در مجموع بر این روش آستانه یابی نیز برتری دارد. دو نمونه از تصاویر استفاده شده برای این ارزیابی در شکل ۸ نشان داده شده است.

### ۵-۲ ارزیابی بصری

هدف اصلی این تحقیق ارائه روشی برای دودویی کردن تصاویر متنی است به نحوی که ایجاد شکستگی در حروف و به هم چسبیدن آنها را کمینه کند، همچنین پس زمینه را به طور مناسب حذف کند و نواحی با تباین کم را مانند سایر نواحی دودویی کند. لذا این معیارها برای ارزیابی بصری روش پیشنهادی استفاده می شوند.

برای ارزیابی بصری، ۳۰ تصویر متنی آسیب دیده با روش های آستانه یابی مختلف دودویی شده و مقایسه می شوند. در شکل ۹ نتیجه دودویی کردن یک نمونه تصویر متنی آسیب دیده با روش های مختلف نشان داده شده است. نتایج ارزیابی نشان می دهد روش آستانه یابی کلی [۱] نواحی تیره پس زمینه را سیاه می کند و متن موجود در این نواحی به طور کامل از بین می روند. روش [۷] متن را به خوبی استخراج می کند ولی در نواحی که متن وجود ندارد نویز پس زمینه تقویت شده و به عنوان متن برچسب زده می شود. روش [۵] برای استخراج نواحی متنی با تباین کم مناسب نیست و باعث ایجاد شکستگی در حروف و یا حذف کامل آنها می شود. روش بارش باران [۲۰] برای استخراج متن از تصاویری که تغییرات روشنایی پس زمینه آنها خیلی زیاد است، مناسب نیست. زیرا مقدار آب زیادی در نواحی تیره پس زمینه جمع می شود و پس از آستانه یابی مقدار آب، این نواحی برچسب متن می گیرند و همچنین نواحی متنی با تباین کم حذف می شوند. روش پیشنهادی به تباین بین متن و پس زمینه وابستگی ندارد و با کیفیت خوبی نواحی متنی را استخراج و پس زمینه را حذف می کند. تصاویری که با دوربین گرفته شده اند و تباین بین متن و پس زمینه در نواحی مختلف آنها متفاوت است با روش پیشنهادی به خوبی دودویی می شود در حالی که سایر روش های بررسی شده قادر به دودویی کردن این تصاویر نیستند.

### ۶- نتیجه گیری

در این مقاله روشی مستقل از تباین برای دودویی کردن تصاویر متنی ارائه شد که از چندین مرحله تشکیل می شود. در مرحله اول با ایجاد تغییرات در مدل بارش باران و با یک مرحله بارش قسمت های بارز نواحی

## An example Faint region

## Smudged region

شکل ۷: تصویر دودویی با استفاده از الگوریتم پیشنهادی برای تصویر شکل ۲-الف.

جدول ۱: مقایسه کمی روش پیشنهادی با ۴ روش معروف آستانه یابی.

روش پیشنهادی	Kim [۲۰]	Sauvola [۵]	Niblack [۷]	Otsu [۱]	
۵۳۹۷	۴۹۰۷	۴۳۴۱	۵۲۴۴	۴۸۲۵	N۱
۹	۱۰۶	۷۰۷	۳۱	۳۷	N۲
۲	۳۹۵	۳۳۰	۱۳۳	۵۴۶	N۳
۱۳	۱۹	۶۲	۳۵۷	۲	N۴

### ۵- نتایج تجربی و مقایسه با کارهای گذشته

متأسفانه روش مشخصی برای ارزیابی روش های دودویی کردن تصاویر متنی وجود ندارد و در مراجع مختلف از روش های متفاوتی استفاده شده است. در [۲۰] و [۲۲] از معیار ارزیابی بصری و در [۴] از دقت بازشناسی متن در تصویر دودویی توسط نرم افزارهای نویسه خوان برای ارزیابی استفاده شده است.

در این مقاله برای ارزیابی الگوریتم پیشنهادی آن را با چهار روش معروف آستانه یابی مقایسه می کنیم. پارامترهای از پیش تعیین شده این روش ها مطابق مقادیر ارائه شده در مقاله های اصلی تنظیم شده اند. مقایسه بر مبنای هر دو روش ارزیابی بصری و دقت بازشناسی با استفاده از نرم افزار نویسه خوان انجام می شود.

### ۵-۱ ارزیابی با استفاده از دقت بازشناسی حروف

برای انجام مقایسه، تصاویر آسیب دیده با روش های مختلف دودویی شده و به نرم افزار نویسه خوان ۹ ABBYY FineReader داده می شوند. مقایسه با توجه به خروجی نرم افزار انجام می شود. در این مقاله از معیارهای زیر برای ارزیابی کمی روش های دودویی کردن تصاویر متنی استفاده می شود که سه معیار N۱-N۳ از [۲۳] الهام گرفته شده است و معیار N۴ را برای بررسی میزان حذف پس زمینه ما پیشنهاد کردیم:

۱ N: تعداد حروفی که به درستی استخراج و بازشناسی می شوند.

۲ N: تعداد حروفی که قسمت هایی از آنها استخراج می شود ولی به اشتباه بازشناسی می شوند.

۳ N: تعداد حروفی که استخراج نمی شوند.

۴ N: تعداد لکه های پس زمینه که به عنوان حرف بازشناسی می شوند.

برای این آزمایش از ۱۰ سند متنی استفاده شده است که دارای ۵۴۰۸ حرف هستند و نتایج ارزیابی در جدول ۱ نشان داده شده است.

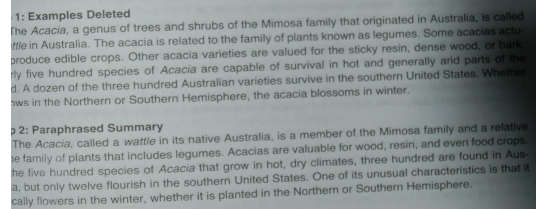
در الگوریتم پیشنهادی تعداد حروفی که به طور صحیح استخراج و بازشناسی شده اند از سایر روش های بررسی شده بیشتر است و تعداد حروفی که استخراج نشده و یا به طور اشتباه بازشناسی شده اند از سایر روش ها کمتر است. تنها در حذف قسمت هایی از پس زمینه که به عنوان حرف بازشناسی می شوند، روش [۱] نتیجه بهتری می دهد که به خاطر ماهیت آستانه یابی کلی است. زیرا در روش آستانه یابی کلی، قطعه هایی از پس زمینه که به عنوان متن معرفی می شوند خیلی بزرگ تر از حروف هستند و توسط نرم افزار نویسه خوان تشخیص داده می شوند. با توجه به این که



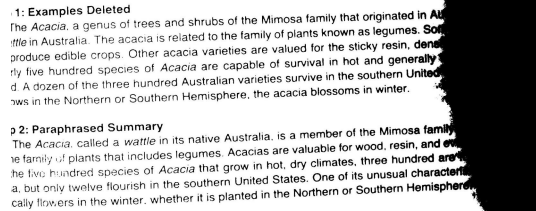
پیشنهادی در مقایسه با روش‌های آستانه‌یابی بررسی شده از کیفیت بهتری برخوردار است.

## مراجع

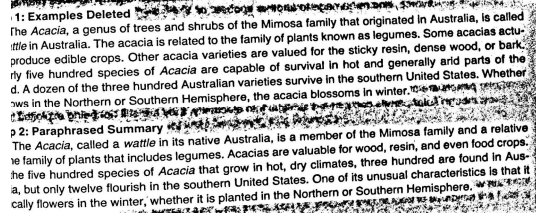
- [1] N. Otsu, "A threshold selection method from grey level histogram," *IEEE Trans. Syst. Man Cybernetics.*, vol. 9, no. 1, pp. 62-66, Jan. 1979.
- [2] J. N. Kapur, P. K. Sahoo, and A. K. C. Wong, "A new method for graylevel picture thresholding using the entropy of the histogram," *Computer Vision, Graphics and Image Processing*, vol. 29, no. 3, pp. 273-285, Mar. 1985.
- [3] J. S. Weszka and A. Rosenfeld, "Histogram modification for threshold selection," *IEEE Trans. on Systems, Man, Cybernetics*, vol. 9, no. 1, pp. 38-52, Jan. 1979.
- [4] B. Gatos, I. Pratikakis, and S. J. Perantonis, "Adaptive degraded document image binarization," *Pattern Recognition*, vol. 39, no. 3, pp. 317-327, Mar. 2006.
- [5] J. Sauvola and M. Pietikainen, "Adaptive document image binarization," *Pattern Recognition*, vol. 33, no. 2, pp. 225-236, Feb. 2000.
- [6] J. Bernsen, "Dynamic thresholding of grey-level images," in *Proc. of the 8th Int. Conf. on Pattern Recognition*, pp. 1251-1255, Paris, France, Oct. 1986.
- [7] W. Niblack, *An Introduction to Digital Image Processing*, Prentice Hall, Englewood Cliffs, NJ, pp. 115-116, 1986.
- [8] Y. Yang and H. Yan, "An adaptive logical method for binarization of degraded document images," *Pattern Recognition*, vol. 33, no. 5, pp. 787-807, May 2000.
- [9] M. Kamel and A. Zhao, "Extraction of binary character/graphics images from grayscale document images," *Graphical Models Image Processing*, vol. 55, no. 3, pp. 203-217, May 1993.
- [10] J. M. White and G. D. Rohrer, "Image segmentation for optical character recognition and other applications requiring character image extraction," *IBM J. Research Development*, vol. 27, no. 1, pp. 400-411, Feb. 1983.
- [11] O. D. Trier and T. Taxt, "Improvement of 'Integrated Function Algorithm' for binarization of document images," *Pattern Recognition Letters*, vol. 16, no. 3, pp. 277-283, Mar. 1995.
- [12] S. Rodtook and Y. Rangsaneri, "Adaptive thresholding of document images based on Laplacian sign," in *Proc. Int. Conf. on Information Technology: Coding and Computing*, pp. 501-505, 2-4 Apr. 2001.
- [13] Q. Chena, Q. Suna, P. A. Heng, and D. Xia, "A double-threshold image binarization method based on edge detector," *Pattern Recognition*, vol. 41, no. 4, pp. 1254-1267, Apr. 2008.
- [14] X. Ye, M. Cheriet, and C. Y. Suen, "Stroke-model-based character extraction from gray-level document images," *IEEE Trans. on Image Processing*, vol. 10, no. 8, pp. 1152-1161, Aug. 2001.
- [15] S. Huang, M. Ahmadi, and M. A. Sid-Ahmed, "A hidden Markov model-based character extraction method," *Pattern Recognition*, vol. 41, no. 9, pp. 2890-2900, Sep. 2008.
- [16] O. D. Trier and T. Taxt, "Evaluation of binarization methods for document images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 17, no. 3, pp. 312-315, Mar. 1995.
- [17] M. Sezgin and B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation," *J. of Electronic Imaging*, vol. 13, no. 1, pp. 146-165, Jan. 2004.
- [18] Y. Chen and G. Leedham, "Decompose algorithm for thresholding degraded historical document images," in *IEE Proc. Vis. Image Signal Processing*, vol. 152, pp. 702-714, Dec. 2005.
- [19] J. R. Parker, "Gray level thresholding in badly illuminated images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 13, no. 8, pp. 813-819, Aug. 1991.
- [20] I. K. Kim, D. W. Jung, and R. H. Park, "Document image binarization based on topographic analysis using a water flow model," *Pattern Recognition*, vol. 35, no. 1, pp. 265-277, Jan. 2002.
- [21] F. Shafait, D. Keysers, and T. M. Breuel, "Efficient implementation of local adaptive thresholding techniques using integral images," in *Document Recognition and Retrieval XV*, 2008.
- [22] E. Badekas and N. Papamarkos, "Optimal combination of document binarization techniques using a self-organizing map neural network," *Engineering Applications of Artificial Intelligence*, vol. 20, no. 1, pp. 11-24, Feb. 2007.
- [23] Y. Solihin and C. G. Leedham, "Integral ratio: a new class of global thresholding techniques for handwriting images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 761-768, Aug. 1999.



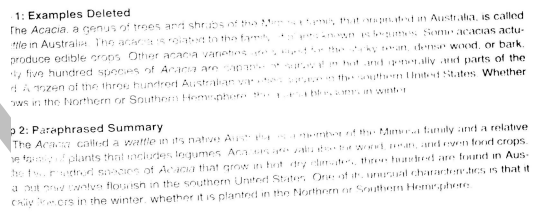
(الف)



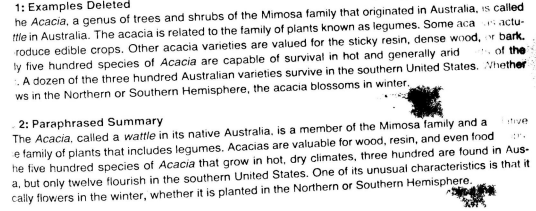
(ب)



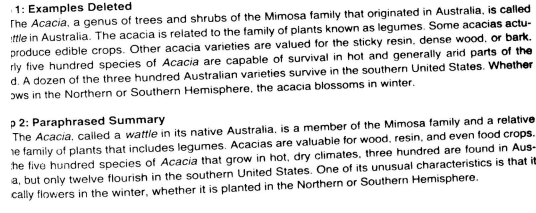
(ج)



(د)



(هـ)



(و)

شکل ۹: (الف) تصویر آسب‌دیده، (ب) تصویر دودویی با روش [۱]، (ج) تصویر دودویی با روش [۲]، (د) تصویر دودویی با روش [۳]، (هـ) تصویر دودویی با روش باران [۴] و (و) تصویر دودویی با روش پیشنهادی.

متنی استخراج می‌شود. در مرحله دوم با استفاده از قسمت‌های بارز نواحی متنی، روشی متن به‌طور محلی تخمین زده می‌شود. همچنین با میانگین‌گیری محلی روشی تصویر و با در نظر گرفتن این که تعداد پیکسل‌های متنی در مقایسه با پس‌زمینه خیلی کم است، روشی پس‌زمینه نیز تخمین زده می‌شود. در مرحله سوم برای هر پیکسل، آستانه به‌صورت میانگین روشی تخمینی پس‌زمینه و متن تعریف می‌شود. این روش آستانه‌یابی از لحاظ بصری و کمی با چند روش معروف آستانه‌یابی مقایسه شده و نتایج ارزیابی نشان می‌دهد که برای تصاویر دوربینی روش



**احسان‌اله کیپر** کارشناسی ارشد پیوسته خود را در مهندسی برق و الکترونیک از دانشکده فنی دانشگاه تهران و دکترای خود را در مهندسی سیستم‌های الکترونیک از دانشگاه اسکس در انگلستان، به ترتیب در سالهای ۱۳۶۴ و ۱۳۶۹ دریافت کرد. او اکنون استاد دانشکده مهندسی برق و کامپیوتر دانشگاه تربیت مدرس است. زمینه‌های پژوهشی مورد علاقه ایشان عبارتند از: بازشناسی الگو، به ویژه بازشناسی متون چاپی و دستنویس

**مرتضی ولی زاده** کارشناسی و کارشناسی ارشد خود را در مهندسی برق و الکترونیک به ترتیب از دانشگاه صنعتی شاهرود و دانشگاه تربیت مدرس در سالهای ۱۳۸۴ و ۱۳۸۶ دریافت کرد. او از سال ۱۳۸۶ در دانشگاه تربیت مدرس به عنوان دانشجوی دکتری مشغول به تحصیل است. زمینه‌های پژوهشی مورد علاقه ایشان عبارتند از: پردازش تصویر، پردازش سیگنال، بازشناسی الگو و آستانه‌یابی تصاویر متنی.

Archive of SID