

اصلاح گستگی‌های ناخواسته در متون چاپی فارسی با گسترش مورفولوژی

احسان‌الله کبیر

هادی گرایلو

دانشگاه تربیت مدرس، بخش مهندسی برق

kabir@modares.ac.ir grailu@modares.ac.ir

چکیده: در بازنگاری متون عموماً فرض براین است که آنها خراب یا آسیب دیده نیستند. با پایین آمدن کیفیت تصاویر دقت الگوریتم‌های تشخیص متون تا حد قابل توجهی پایین می‌آید. لذا اصلاح گستگی‌های متون زمینه مهمی از تحقیقات شده است. در این مقاله تصاویر دودویی متون را بر حسب نوع گستگی در آنها به پنج دسته تقسیم می‌کنیم و از عملگر گسترش مورفولوژی برای اصلاح گستگی‌های استفاده می‌کنیم. شکل و بعد ماسک در میزان موفقیت یک عملگر بسیار مهم است. ما دراین کار مناسبترین شکل و بعد ماسک را طوری که تا حد امکان بیشترین تعداد گستگی‌های متون را اصلاح کند و در عین حال اتصال ناخواسته بوجود نیاورد، به کمک تصاویر تمرین و آزمایش پیدا و ارزیابی می‌کنیم.

کلمات کلیدی: اصلاح گستگی‌های متون، عملگر گسترش، مورفولوژی.

ارزیابی می‌کنیم. بالاخره در بخش ۴ نتیجه‌گیری ارایه می‌شود.

۲. مجموعه تصاویر و دسته‌بندی آنها
برای انجام موفقیت‌آمیز اصلاح گستگی‌های ناخواسته متون یکی از شرایط لازم، داشتن تصاویری است که شامل انواع گستگی، متناسب با فراوانی آنها در نمونه‌های واقعی، باشد.

بخشهای پیوسته یک کلمه را زیر-کلمه می‌نامیم. در این تحقیق از تصاویری شامل حدود ۱۰۰۰ کلمه استفاده شده است. این کلمات حدود ۵۰۰۰ زیر-کلمه دارند. این تصاویر از نظر نوع گستگی به ۵ دسته تقسیم می‌شوند:

- دسته اول: تصاویر سالمی هستند که نه گستگی دارند و نه پیکسلهای پراکنده.
- دسته دوم: تصاویری که فقط پیکسلهای پراکنده دارند.
- دسته سوم: تصاویری که فقط گستگی دارند.
- دسته چهارم: تصاویری که هم گستگی دارند و هم شامل پیکسلهای پراکنده هستند.
- دسته پنجم: تصاویری که گستگی ندارند اما شکل ظاهری آنها به دلیل وجود پیکسلهای پراکنده نامناسب شده است.

هر الگوریتم اصلاح گستگی را که به تصاویر اعمال می‌کنیم باید ارزیابی کنیم. ارزیابی الگوریتم برای هر دسته

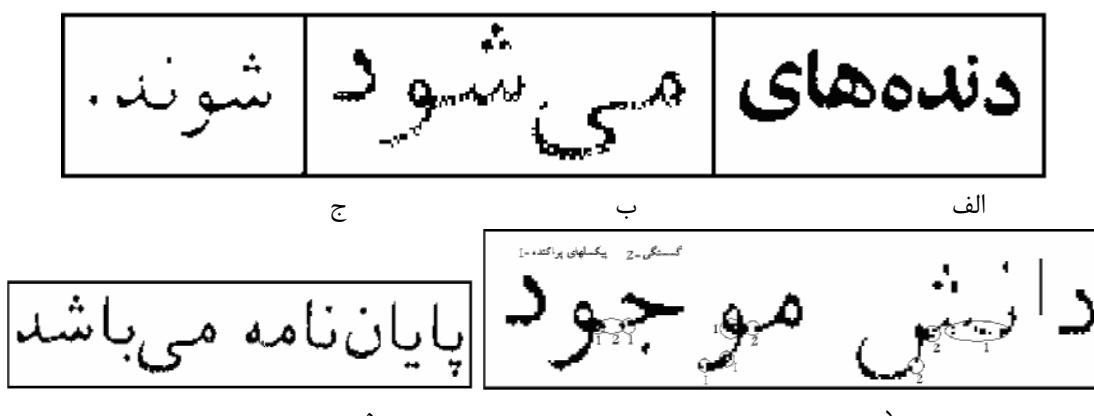
۱. مقدمه: در سیستمهای OCR که امروزه در زندگی ما اهمیت و کاربرد زیادی پیدا کرده‌اند [۱]، فرض بر سالم بودن متون است [۲]. از آنجاکه بازدهی سیستمهای OCR به کیفیت روش و دودویی شدن بستگی دارد [۳ و ۲] و با پایین آمدن کیفیت تصاویر متون حتی به اندازه کم، دقت الگوریتم‌های تشخیص متون تا حد قابل توجهی پایین می‌آید [۴ و ۵]، توجه به اصلاح آسیب‌ها مانند گستگی یا چسبیدگی متون ضروری است. لذا امروزه اصلاح گستگی یا چسبیدگی متون، زمینه مهمی از تحقیقات شده است [۶ تا ۹].

دراین مقاله هدف ما ارایه روشی برای اصلاح گستگی متون چاپی فارسی است [۱۰]. ما می‌خواهیم از عملگر گسترش^۱ مورفولوژی برای انجام این کار استفاده کنیم. شکل و اندازه ماسک استفاده شده برای هر عملگر مورفولوژی در کارایی آن بسیار مهم است [۹]. ما در این مقاله مناسبترین شکل و اندازه ماسک را بدست می‌آوریم. ادامه مقاله به این صورت است که در بخش ۲ مجموعه تصاویر را بر حسب نوع گستگی در آنها به پنج دسته تقسیم می‌کنیم و در مرور خصوصیات هر دسته توضیح می‌دهیم. در بخش ۳ الگوریتم پیشنهادی خود را برای اصلاح گستگی‌ها توضیح می‌دهیم و نتایج حاصل از آنرا

^۱ Dilation

می‌کنند، اما پیکسلهای پراکنده یک قسمت از حرف را به تعداد نسبتاً زیادی قطعات کوچک تقسیم می‌کنند. از آنجا که در این دسته هر دو نوع خرابی پیکسلهای پراکنده و گستینگی وجود دارند، نتایج حاصل از الگوریتم روی این دسته به صورت کمی و کیفی بیان می‌شود. در دسته پنجم همانطور که اشاره شد گستینگی نداریم اما شکل ظاهری حروف به دلایل مختلف مانند پیکسلهای پراکنده بهم ریخته است و تناسب خود را از دست داده اند (شکل ۱-۵). علت اینکه این نوع تصاویر را از دسته تصاویر سالم جدا کرده‌ایم این است که علی‌رغم اینکه گستینگی وجود ندارد اما ظاهر و تناسب حروف نامناسب شده است و این مساله یک آسیب تلقی می‌شود که باید اصلاح شود.

به طور مخصوص به آن دسته انجام می‌شود. دسته اول تصاویر سالم هستند (شکل ۱-الف)، لذا باید ارزیابی آنها به این شکل باشد که هم تعداد اتصال‌های ناخواسته شمرده شود و هم شکل ظاهری حروف بررسی کیفی شود. دسته دوم فقط شامل پیکسلهای پراکنده است (شکل ۱-ب)، لذا برای این دسته نتایج حاصل از الگوریتم را به صورت کیفی توضیح می‌دهیم. دسته سوم فقط شامل گستینگی‌ها است (شکل ۱-ج) و نتایج حاصل از الگوریتم به صورت کمی قابل بیان است. دسته چهارم هم گستینگی دارد و هم شامل پیکسلهای پراکنده هستند (شکل ۱-د). نوع این گونه از تصاویر چنان است که گستینگی‌ها عموماً پیکسلهای پراکنده قابل تفکیک است. گستینگی‌ها عموماً بین دو جزء بزرگ قرار دارند و نیز فقط یک فاصله ایجاد



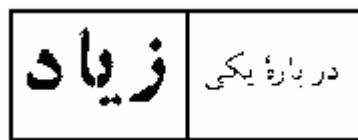
شکل ۱. نمونه‌ای از ۵ دسته تصاویر

چه حد عمومیت دارند. بنابراین تعداد تصاویر آزمایش باید بیشتر از تصاویر تمرین باشد.

تصاویر تمرین و آزمایش باید به نحوی انتخاب شوند که هر یک تا حد امکان شامل همه انواع تصاویر باشند. البته این که از هر دسته چند عدد در مجموعه تصاویر تمرین قرار داده شود بستگی به فراوانی هر کدام از دسته‌ها در تصاویر متی چاپی فارسی دارد. بطور مثال اگر از نظر آماری تصاویری که پیکسلهای پراکنده دارند بیشتر از تصاویری است که گستینگی دارند باید در مجموعه تصاویر تمرین تعداد تصاویر حالت اول بیشتر از حالت دوم باشد.

۱-۲. مجموعه‌های تمرین و آزمایش

مجموعه تصاویر به دو مجموعه تمرین و آزمایش تقسیم شده‌اند. تصاویر تمرین تصاویری هستند که الگوریتم اصلاح گستینگی ابتدا روی آنها پیاده‌سازی می‌شود تا بهترین مقدار یا مقادیر پارامترهای الگوریتم انتخاب شوند. در الگوریتم گسترش انواع ماسک‌ها را به تصاویر تمرین اعمال می‌کنیم. سپس از بین نتایج حاصل، مناسب‌ترین شکل و اندازه ماسک را انتخاب می‌کنیم. بعد از انتخاب بهترین پارامترها کارایی الگوریتم را روی تصاویر آزمایش می‌سنجیم تا مشخص شود که پارامترهای انتخاب شده تا



شکل ۲. دونمونه تصویر استفاده شده با بیشترین و کمترین عرض قلم

به همراه نتیجه اعمال گسترش با دو ماسک ۱۳ و ۱۷ را نشان می‌دهد. همانطور که در شکل ۴-ب و ۴-ج نشان داده شده است، ماسک ۱۳ به این دلیل که در جهت افقی نیز کشیدگی ایجاد می‌کند توانسته گستگی مشخص شده را اصلاح و پر کند، اما ماسک ۱۷ نتوانسته است. همچنین توجه می‌کنید که ماسک ۱۳ نسبت به ماسک ۱۷ توانسته فاصله گستگی‌ها را کمتر کند و شکل ظاهری حروف را بهبود دهد. شکل ۵ یک تصویر را به همراه نتیجه گسترش آن با سه ماسک ۱۰، ۱۱ و ۱۲ نشان می‌دهد. همانطور که ملاحظه می‌شود ماسک ۱۰ که بیشترین کشیدگی را در راستای قطری دارد، تمام گستگی‌ها را پر کرده است، اما دونقطه‌ها را بطور کامل به هم وصل و تبدیل به یک مستطیل کرده است و نیز حرفه موجود در حروف "و" و "ض" را پر کرده است و بالاخره اینکه تورفتگی فوقانی حرف "ح" را تقریباً پر کرده است. از طرف دیگر در تصویر مربوط به ماسک ۱۲ که کمترین میزان کشیدگی در راستای قطری را در بین سه ماسک دارد، ملاحظه می‌شود که تمام گستگی‌ها را ازین نبرده است و دو نقطه‌ها را به هم وصل کرده، اما آنها را مانند ماسک قبلی آنقدر به هم نچسبانده که تبدیل به یک مستطیل شود، بلکه حالت دو نقطه بودن تا حد خوبی مشخص است. این ماسک حرفه داخل حروف را تا حد خوبی حفظ کرده و انحنای طبیعی حروف مانند انحنای فوقانی حرف "ح" را نیز تا حدی حفظ کرده است. ماسک ۱۱ که حد متوسط کشیدگی در راستای قطری را در بین سه ماسک مذکور دارد از نظر خصوصیاتی که برای دو ماسک قبل بررسی شد حد متوسط را اعمال کرده است. یعنی از یک طرف توانسته تمام گستگی‌ها را پر کند واز طرف دیگر حرفه داخل حرف "ض" و انحنای طبیعی حروف را تا حدی حفظ کرده و دونقطه‌ها را اتصال داده است.

با توجه به متونی که در اختیار داشتیم بیشترین فراوانی انواع آسیب‌ها اول مربوط به گستگی و بعد از آن گستگی همراه با پراکندگی پیکسلها و درنهایت پیکسلهای پراکنده بود. از بین ۵۰۰۰ زیر-کلمه حدود ۸۰۰ زیر-کلمه در مجموعه تصاویر تمرين و بقیه در مجموعه تصاویر آزمایش قرار دارند. در تهیه تصاویر تمرين و آزمایش سعی شده است که از عرض قلم‌های مختلف استفاده شود. در شکل ۲ دو تصویر نشان داده شده است که کمترین و بیشترین عرض قلم را در بین تصاویر مورد استفاده در این تحقیق دارند.

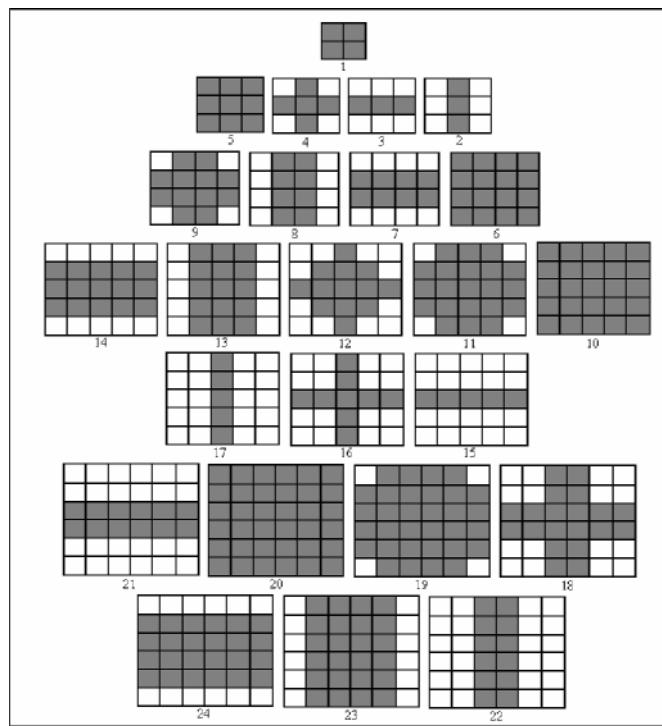
۳. اصلاح گستگی به کمک عملگر گسترش

لغت مورفولوژی عموماً معرف شاخه‌ای از زیست‌شناسی است که با شکل و ساختار حیوانات و گیاهان سروکار دارد. درینجا مورفولوژی ریاضی به عنوان یک ابزار مناسب در پردازش تصویر به‌ویژه تصویر دوسطحی استفاده می‌شود.^[۹]

۱-۳ ماسکهای استفاده شده برای تعیین مناسبترین شکل و اندازه

در عملگر گسترش از ۲۴ ماسک با اشکال و ابعاد مختلف استفاده شد. این آزمایشها روی تصاویر تمرين انجام شد تا مناسبترین ماسک انتخاب شود. این ماسکهادر شکل ۳ به ترتیب اندازه‌شان نشان داده شده‌اند.

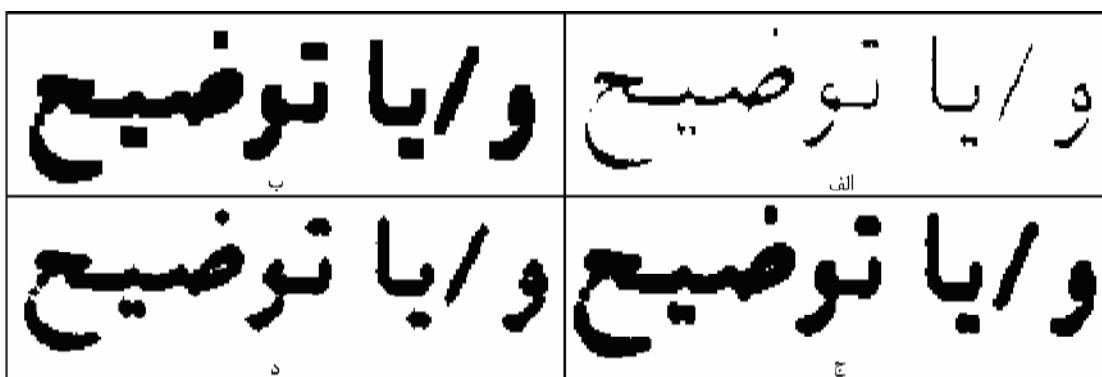
در طراحی هر ماسک منظور خاصی در نظر گرفته شده است. ایده کلی در طراحی ماسک این است که در جهتی که یک ماسک طول بیشتری داشته باشد در همان جهت کشیدگی بیشتر ایجاد می‌کند. به طور مثال ماسکهای ۳، ۷، ۱۴، ۱۵، ۲۴ و ۲۱ بیشتر در راستای افقی کشیدگی ایجاد می‌کنند. تفاوت این ماسکها باهم یکی ابعاد آنها و دیگری نسبت کشیدگی افقی به کشیدگی عمودی آنها است. این تفاوت‌های ظریف می‌توانند مهم باشند. شکل ۴ یک تصویر



شکل ۳. ماسکهای استفاده شده در این مقاله برای تعیین بهترین شکل و اندازه ماسک

| | | |
|-------------------------|-------------------------|-------------------------|
| نباید نوشته شود. | نباید نوشته شود. | نباید نوشته شود. |
|-------------------------|-------------------------|-------------------------|

شکل ۴. الف- تصویر اصلی ب- عملکرد ماسک ۱۳ ج- عملکرد ماسک ۱۷



شکل ۵. الف- تصویر اصلی ب- نتیجه ماسک ۱۰.

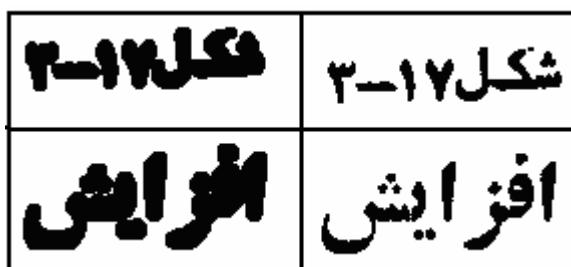
ج- نتیجه ماسک ۱۲ د- نتیجه ماسک ۱۱

تمرین اعمال کردیم. نتیجه بدست آمده درمورد شکل مناسب ماسک که تاحد ممکن گستنگی‌ها را پرکند و درعین حال اتصال ناخواسته بوجود نیاورد و شکل ظاهری

۲-۳ یافتن ماسک مناسب عملگر گسترش را با ماسکهای نشان داده شده در شکل ۳ برای تعیین مناسبترین شکل و ابعاد ماسک، به تصاویر

البته برای متون با برخی قلم‌های خاص که حروف ضخیم و نزدیک به هم نوشته می‌شوند، ماسک مذکور باعث می‌شود که شکل ظاهری حروف خراب شود و علاوه بر این تعداد اتصال‌های ناخواسته نسبت به معمول کمی بالاتر رود (شکل ۶). چندنمونه از نتایج بدست آمده از اعمال عملگر گسترش با ماسک مناسب برای تصاویر آزمایش در شکل ۷ نشان داده شده است.

حروف را حفظ کند، این است که ماسک باید مربعی باشد. در راستای افقی و عمودی به یک اندازه کشیدگی ایجاد کند، اما کشیدگی آن در راستاهای قطری بین نصف تا سه‌چهارم کشیدگی در راستای افقی یا عمودی باشد. در مورد ابعاد ماسک نیز این نتیجه بدست آمده که ابعاد ماسک باید حدود $1/5$ برابر عرض قلم باشد. ما عرض قلم را از مرجع [۱۱] بدست آورده‌یم.



شکل ۶. نمونه‌هایی از نتایج ماسک با ابعاد $1/5$ برابر عرض قلم.
ستون راست: تصاویر اصلی و ستون چپ: نتیجه گسترش

| | |
|-------------------------|-------------------------|
| نتایج ارائه می‌گردند. | نتایج ارائه می‌گردند. |
| دققت شود و از حاشیه‌روی | دققت شود و از حاشیه‌روی |
| اور درایو و سیله‌ای | اور درایو و سیله‌ای |
| چند کنترل مکانیکی | چند کنترل مکانیکی |
| دو تیغه مجاور هم از نوع | دو تیغه مجاور هم از نوع |
| گشتاور | گشتاور |

شکل ۷. چند نمونه از تصاویر حاصل از گسترش با ماسک مناسب به تصاویر آزمایش
ستون راست: تصاویر اصلی. ستون چپ: نتیجه گسترش

ناخواسته بوجود آمد که همگی مربوط به عرض قلم‌های خاص هستند که حروف در آنها ضخیم نوشته می‌شوند. دسته دوم از تصاویر متون شامل ۱۹۰۰ زیرکلمه است که فقط شامل پیکسلهای پراکنده هستند. نتیجه قابل پیش‌بینی که بدست آمد این است که تا زمانی که فاصله پیکسلهای پراکنده از هم از حدود ابعاد ماسک بیشتر نشود، همانطور که در شکل ۸ نشان داده شده است، الگوریتم گسترش آنها را به هم وصل می‌کند اما در غیر اینصورت جزء گستته به وجود خواهد آمد.

۳-۳ نتایج آزمایش

همانطور که در بخش ۲ اشاره شد تصاویر را به پنج دسته تقسیم کرده‌ایم که تعداد تصاویر هر دسته متناسب با فراوانی آنها در نمونه‌های واقعی تصاویر متنی است. در دسته اول فقط الگوهای سالم وجود دارد. در این دسته بطور متوسط ۷۰۰ زیر کلمه وجود دارد. برای ارزیابی باید تعداد اتصال‌های ناخواسته بیان شود و شکل ظاهری حروف بررسی شود. پس از گسترش با ماسک مناسب، ۸ اتصال

| | |
|------------------------------|--|
| نتایج ارائه می‌گردند. | |
|------------------------------|--|

شکل ۸. عملکرد الگوریتم گسترش روی دسته دوم، الف- تصویر اصلی ب- بزرگ شده یکی از کلمات تصویر الف که در آن دو ناحیه با پیکسل پراکنده مشخص شده است. ج- موقفیت الگوریتم در پرکردن پیکسلهای پراکنده و نیز باقی ماندن برخی پیکسلهای پراکنده که فاصله‌شان از حدود ابعاد ماسک بیشتر است.

وجود دارد. در این زیرکلمات حدود ۳۲۵۰ گستگی بدون احتساب پیکسلهای پراکنده وجود دارد. پس از اعمال الگوریتم گسترش، حدود ۸۰/۸ درصد از آنها اصلاح و پر شدند. از نظر کیفی تصاویر اولیه که به خاطر پیکسل پراکنده شکل و تناسب اولیه خود را از دست داده بودند، پس از اعمال الگوریتم، شکل و تناسب بهتری پیدا کردند(شکل ۹).

دسته سوم از تصاویر شامل ۷۰۰ زیر کلمه و فقط شامل گستگی است. حدود ۲۱۰۰ گستگی در این دسته وجود دارد که پس از اعمال الگوریتم حدود ۸۱ درصد از آنها اصلاح و پر شدند. برای حدود ۴ درصد از تعداد کل زیر کلمه ها نیز اتصال ناخواسته بوجود آمد. دو نمونه از تصاویر این دسته به همراه نتیجه گسترش آنها در ردیف دوم شکل ۷ نشان داده شده است.

دسته چهارم شامل تصاویری است که هم گستگی و هم پیکسل پراکنده دارند. در این دسته حدود ۱۲۰۰ زیر کلمه

| | |
|-------------------------------|-------------------------------|
| از نتیجه به دانش موجود | از نتیجه به دانش موجود |
| ببخش نتایج | ببخش نتایج |
| روش کار که ممکن است | روش کار که ممکن است |

شکل ۹. چند نمونه از نتایج دسته چهارم، پیکسلهای پراکنده به خوبی پر شده‌اند و شکل ظاهری حروف بهتر شده است.

مورفولوژی این به هم ریختگی شکل ظاهری را نیز به میزان نسبتاً خوبی اصلاح کرده است. در این دسته پس از گسترش حدود ۳ درصد از زیر کلمات، اتصال ناخواسته بوجود آمد.

دسته پنجم شامل تصاویری است که پارگی ندارند اما شکل ظاهری حروف یا کلمات در آنها بهم خورده است. این دسته شامل حدود ۵۰۰ زیر کلمه است. شکل ۱۰ چند نمونه از نتیجه گسترش را نشان می‌دهد. گسترش

| | |
|----------------------------|----------------------------|
| پایان‌نامه می‌باشد. | پایان‌نامه می‌باشد. |
| است که | است که |
| ارائه اختصاری | ارائه اختصاری |

شکل ۱۰. چند نمونه از نتایج دسته پنجم. الگوریتم گسترش توانسته تاحد خوبی شکل ظاهری حروف را بهتر کند.

می‌کند. پس از پیکسلهای پراکنده، عملگر گسترش برای گستینگی‌هایی که قلم ضخیم نداشته باشند نیز نسبتاً خوب عمل می‌کند. همانطور که شکل ۸ نشان می‌دهد الگوریتم گسترش برای قلمهای ضخیم چندان خوب عمل نمی‌کند و اتصال ناخواسته بوجود می‌آورد.

به عنوان یک راهکار آینده شاید به جز عرض قلم، لازم باشد نوع قلم نیز مشخص شود تا شکل و ابعاد ماسک را وفقی انتخاب کنیم. همچنین یک راهکار دیگر این است که به کمک اسکلت حروف راستای اجزای گستینه را پیدا کنیم و سپس ابعاد و شکل ماسک را مناسب با راستای اجزای دو طرف گستینگی طراحی کنیم.

[8] J. Wang and H. Yan, "Mending broken handwriting with a macrostructure analysis method to improve recognition," *Pattern Recognition Letters*, vol.20, pp. 855-864, 1999.

[9] R. C. Gonzalez and R. E. Woods, "Digital Image Processing," Prentice Hall, 2001.

[۱۰] هادی گرایلو، "اصلاح گستینگی حروف در تصاویر دودویی"، پایان نامه کارشناسی ارشد، دانشگاه تربیت مدرس، بخش مهندسی برق، تابستان ۸۴

[۱۱] م. ر. احمدزاده و ا. کبیر، "شکستن کلمات تایپ شده فارسی به حروف"، کنفرانس کاربرد کامپیوتو در ایران، ص. ۱۷۴، اصفهان، دیماه ۱۳۷۰.

۴. نتیجه‌گیری و پیشنهاد راهکار

در این مقاله چگونگی انتخاب ماسک که شامل انتخاب شکل و ابعاد آن است، بیان شد. عملگر گسترش به همراه ماسک انتخاب شده به مجموعه تصاویر آزمایش اعمال شد. از بررسی مثالها و نتایج بدست آمده برای هر یک از دسته‌های تصاویر، قابلیت عملگر گسترش با ماسک مناسب مشخص شد.

عملکرد عملگر گسترش برای دسته دوم و چهارم که شامل پیکسلهای پراکنده هستند، نسبت به دسته‌های دیگر بهتر است. عملگر مذکور پیکسلهای پراکنده را به طور مناسب پر می‌کند و شکل ظاهری حروف را تا حد خوبی حفظ

مراجع

- [1] A. P. Whichello and H. Yan, "Linking broken character borders with variable sized masks to improve recognition," *Pattern Recognition*, vol. 29, no. 8, pp. 1429-1435, 1996.
- [2] Y. Lu, "Machine printed character segmentation-An overview," *Pattern Recognition*, vol. 28, no. 1, pp. 67-80, 1995
- [3] H. S. Baird, "The state of the art of document image degradation modeling," Proc. 4th IAPR Workshop on Document Analysis Systems , DAS 2000, Dec. 2000.
- [4] H. S. Baird, "Document Image Defect Models and Their Uses," Proc. 2nd Int'l Conf. On Document Analysis and Recognition, Tsukuba Science City, Japan, pp. 62-67, Oct. 1993
- [5] R. G. Casey and G. Nagy, "Recursive segmentation and classification of composite character patterns," Proc 6th Int'l Conf. on Pattern Recognition, Munich, Germany, pp. 1023-1026, 1982.
- [6] J. Rocha and T. Pravidis, "A shape analysis model with application to a character recognition system," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 16, no. 4, pp. 393-404, 1994.
- [7] A. P. Whichello and H. Yan, "Reconstruction of character skeletons using gabor filter features," *Electronic Letters*, vol. 31, pp. 1911-1912, 1995.