

بهبود بازیابی تصاویر بر اساس محتوا و بهینه‌سازی با استفاده از روش فازی

مجید محمدپور

دانشگاه آزاد اسلامی واحد علوم تحقیقات تهران (کهگیلویه و بویراحمد).

نام نویسنده مسئول:

مجید محمدپور

تاریخ دریافت: ۱۳۹۹/۱۱/۱۲

تاریخ پذیرش: ۱۳۹۹/۱/۲۳

چکیده

یکی از مسائل کاربردی در حوزه‌ی پردازش تصویر استفاده از روش متن‌گذاری در تصاویر است. این روش دارای کاربردهای بسیار زیادی در حوزه‌ی امنیتی و تجاری است. در حوزه‌ی متن-گذاری اولین قدم تبدیل دیجیتال متن و تصویر است. ما برای افزایش مقاومت و پایداری عمل متن‌گذاری از روش تبدیل فوری کسینوسی استفاده کرده‌ایم. دو پارامتر اساسی در روش نشان‌گذاری نرخ پیک سیگنال به نویز و نرخ خطای بیت است. ما برای بهبود این دو پارامتر از روش فازی برای بهینه‌سازی نتایج به‌دست آمده استفاده کرده‌ایم و از روش عصبی برای آموزش الگوریتم بهره گرفته‌ایم. نتایج شبیه‌سازی به‌دست آمده از این روش نشان می‌دهد که عملکرد روش پیشنهادی برای دو پارامتر مطرح شده بهتر از روش‌های ارائه شده‌ی قبلی است و به این ترتیب به تکرار ۱۰۰ بار از روش فازی به نتایج مطلوب دست یافته‌ایم.

واژگان کلیدی: متن‌گذاری، تبدیل فوری کسینوسی، نرخ خطای بیت، نسبت سیگنال به نویز، گراف دو لایه.

مقدمه

پردازش تصویر دیجیتال^۱ دانش جدیدی است که سابقه آن به پس از اختراع رایانه های دیجیتال باز می گردد. با این حال، این علم نوپا در چند دهه‌ی اخیر از هر دو جنبه‌ی نظری و عملی پیشرفت‌های چشمگیری داشته است. سرعت این پیشرفت به اندازه‌ای بوده است که هم اکنون و پس از این مدت نسبتاً کوتاه، به راحتی می توان ردپای پردازش تصویر دیجیتال را در بسیاری از علوم و صنایع مشاهده نمود. علاقه به روش‌های پردازش تصویر دیجیتال از دو محدوده کاربردی اصلی نشأت می گیرد که آن محدوده‌ها عبارتند از: بهبود اطلاعات تصویری به منظور تعبیر انسانی و پردازش داده‌های صحنه برای ادراک ماشینی مستقل.

چند دسته‌ی مهم از کاربردهای پردازش تصویر به شرح زیر می باشد:

الف) کاربردهای عکاسی مانند ارتقاء، بازسازی تصاویر قدیمی، بازسازی تصاویر خراب شده با نویز و بهبود ظاهر تصاویر معمولی.

ب) کاربردهای پزشکی مانند ارتقاء ویژگی‌های تصاویر اشعه ایکس، تولید تصاویر CT-scan و MRI.

ج) کاربردهای امنیتی مانند تشخیص حرکت (در دزدگیرها)، تشخیص اثر انگشت، تشخیص چهره و تشخیص امضاء.

د) کاربردهای نظامی مانند تشخیص و رهگیری خودکار اهداف متحرک یا ثابت از هوا یا از زمین.

ه) کاربردهای سنجش از راه دور مانند ارتقاء و تحلیل تصاویر هوایی و ماهواره‌ای (برداشته شده از مناطق مختلف جغرافیایی) که در کاربردهای نقشه‌برداری، کشاورزی، هواشناسی و موارد دیگر مفید هستند.

و) کاربردهای صنعتی مرتبط با خودکارسازی صنایع مانند تفکیک محصولات مختلف بر اساس شکل یا اندازه، آشکارسازی نواقص و شکستگی‌های موجود در محصولات، تعیین محل اشیاء و اجرای فرآیند تولید با استفاده از روبات‌ها و بینایی ماشینی.

ز) کاربردهای فشرده‌سازی تصویر مانند ذخیره‌سازی، ارسال تصاویر تلویزیون با کیفیت بالا و ارسال تصاویر متحرک و زنده از روی شبکه اینترنت و یا خط تلفن.

ح) موارد متفرقه دیگری نیز مانند تصویر برداری از اسناد و ارسال آن‌ها توسط دورنگار و تشخیص خودکار نویسه در ردیف کاربردهای پردازش تصویر قرار دارند.

با افزایش سریع استفاده از اطلاعات چندرسانه‌ای مانند متن، تصویر، صدا و ویدئو تقاضا برای ذخیره، جستجو و بازیابی این اطلاعات به شدت افزایش یافته است. در این بین بازیابی تصاویر مبتنی بر محتوی با توجه به کاربرد وسیع آن از دیرباز مورد توجه پژوهشگران بوده است. یک سیستم بازیابی تصویر مبتنی بر محتوی از ویژگی‌های سطح پایین مانند رنگ، شکل و بافت جهت شاخص‌گذاری تصاویر استفاده می‌کند. اگر چه بازیابی تصویر مبتنی بر محتوی در برخی از کاربردها نتایج خوبی داشته است اما با مشکلات و هزینه‌هایی مثل قطعه‌بندی تصویر، استخراج ویژگی‌های تصویری مناسب و منطبق با ادراک کاربر روبرو است. همچنین وجود فاصله معنایی بین مفاهیم سطح بالا و ویژگی‌های سطح پایین و نیز مسأله ذهنی بودن ادراک انسان‌ها، کارایی روش‌های مبتنی بر محتوی را کاهش داده است [۱].

از این‌رو با توجه به نتیجه خیلی از کاربردها، تحقیقات در زمینه بازیابی تصویر به سمت بازیابی معنایی تصویر و استفاده از بازخورد مرتبط تغییر کرده است. روش‌های زیادی تاکنون با استفاده از بازخورد مرتبط پیشنهاد شده‌اند. یک متد احتمالی یادگیری و ویژگی‌های مرتبط معرفی شد که به‌طور خودکار با استفاده از بازخوردهای کاربر ویژگی‌های مرتبط را تشخیص می‌داد. در یک شیوه بازخورد مرتبط جدید بر پایه SVM یک کلاسه معرفی شده است که مفاهیم سطح بالای کاربر را به دست آورده و از آن در پرس‌وجوهای آینده استفاده می‌کند [۲].

اگر یک جلسه بازیابی به صورت ارائه یک پرس‌وجو توسط کاربر و مجموعه تمام بازخوردهای کاربر بعد از همان پرس وجو، تعریف شود، در این صورت اکثر روش‌ها فقط قادرند نتیجه بازیابی را در خلال یک جلسه بازیابی بهبود بخشند بدون اینکه این بهبود و نتایج حاصل از آن در جلسات بعدی بازیابی مورد استفاده قرار گیرد. آن دسته از مکانیزم‌های یادگیری هم، که وجود دارد، اغلب نیازمند مدل‌های محاسباتی پیچیده و یا استفاده از یک سری کلمات کلیدی می‌باشند. مدل گراف دو لایه با بهره-

^۱ Digital Image Processing

گیری از ویژگی‌های محتوایی تصویر در لایه اول و مفاهیم سطح بالای آن‌ها در لایه دوم سعی دارد تا حدودی مشکلات ذکر شده روش‌های بازخوردی را برطرف کند.

اما روشی که ما در این مقاله ارائه می‌نمائیم استفاده از مدل گراف دو لایه است. این مدل شامل دو لایه می‌باشد. یک لایه بصری و یک لایه معنایی. در هر لایه یک گراف بدون جهت وجود دارد و هر گره در این دو گراف معرف یک تصویر در پایگاه داده تصویری می‌باشد و اتصالات بین گره‌ها وجود نوعی شباهت و همبستگی بین تصاویر را نشان می‌دهد و وزن هر یال شدت این شباهت را نشان می‌دهد. گره‌های هر گراف در هر لایه دقیقاً متناظر با گره‌های گراف دیگر می‌باشند. یعنی به ازاء هر گره در لایه بصری یک گره متناظر با آن در لایه معنایی وجود دارد. اما اتصالات بین گره‌ها در دو لایه متفاوت است. یک اتصال در لایه معنایی نمایانگر میزان شباهت دو تصویر از نظر معنایی می‌باشد، درحالی‌که یک اتصال در لایه بصری شباهت بین دو تصویر را از نقطه نظر ویژگی‌های سطح پائین تصویری نشان می‌دهد. بازیابی تصویر بر اساس این گراف دو لایه صورت می‌گیرد و نحوه توزیع اتصالات و وزن آنها کارایی بازیابی را تعیین می‌کند [۳].

اما برای بهینه‌سازی روش مذکور ما از قاعده‌ی فازی استفاده می‌کنیم. در این روش ما برای هر یک از اعضای مجموعه از معیاری تحت عنوان درجه‌ی عضویت استفاده می‌کنیم که میزان عضویت هر داده را در یک مجموعه‌ی خاص مشخص می‌کند. بر این اساس، با استفاده از رویکرد کمینه-بیشینه میزان درجه‌ی عضویت هر یک از داده‌ها را تا زمان دستیابی به نتیجه‌ی مطلوب تغییر می‌دهیم. در نهایت انتظار داریم که با استفاده از این الگوریتم بهینه‌سازی نتایج مطلوب‌تری را به‌دست آوریم.

ما در این تحقیق با توجه به استفاده از الگوریتم فازی به دنبال دستیابی به اهداف زیر هستیم:

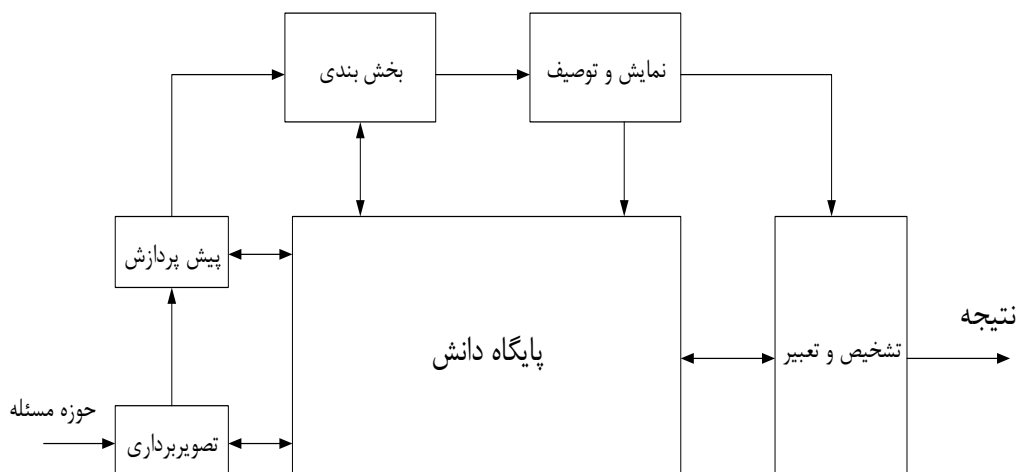
- افزایش کارایی در روش‌های بازیابی تصاویر بر اساس محتوا
- قابلیت بکارگیری روش برای کاربردها و تصاویر مختلف
- کاهش مشکلات و موانع موجود در روش‌های قبلی
- ارائه‌ی یک رویکرد جامع و کارآ برای بازیابی تصاویر بر اساس محتوا

مقاله ارائه شده شامل بخش‌های زیر است:

در بخش اول مقدمه و کلیات موضوع تشریح شده است. در بخش دوم ادبیات موضوع و مرور کارهای گذشته آورده شده است. در بخش سوم روش پیشنهادی تشریح گردیده است. بخش چهارم شامل نتایج تجربی و آزمایشات انجام شده است. بخش پنجم نتیجه‌گیری و کارهای آتی را بیان می‌کند.

باید خاطرنشان کرد که در ادامه همین بخش مراحل اصلی پردازش تصویر تشریح گردیده است.

پردازش تصویر دیجیتال محدوده وسیعی از سخت‌افزار، نرم‌افزار و مبانی نظری را در بر می‌گیرد. در این قسمت مراحل اساسی مورد نیاز برای اجرای یک پردازش روی تصویر را نام می‌بریم که در شکل ۱ نمایش داده شده است.



شکل ۱: مراحل اصلی پردازش تصویر دیجیتال

مرحله اول این فرآیند، تصویربرداری^۲-یعنی به دست آوردن تصویر دیجیتال-است. انجام دادن چنین کاری نیازمند یک حسگر تصویربردار^۳ و قابلیت دیجیتال سازی سیگنال خروجی حسگر می باشد. پس از این که تصویر دیجیتال به دست آمد، مرحله بعدی پیش پردازش آن است. وظیفه اصلی پیش پردازش، بهبود تصویر به روش هایی است که امکان توفیق سایر پردازش ها را نیز افزایش دهد. پیش پردازش، به طور معمول به روش هایی برای ارتقاء تمایز، حذف نویز و جداسازی آن نواحی که زمینه شان نشان دهنده احتمال وجود اطلاعات حرفی-عددی است، می پردازد. مرحله بعدی به بخش بندی^۴ می پردازد. در تعریف وسیع، بخش بندی فرایندی است که تصویر ورودی را به قسمت ها یا اجزای تشکیل دهنده اش تقسیم می کند. به طور کلی بخش بندی یکی از مشکل ترین کارها در پردازش تصویر دیجیتال است. از طرفی یک شیوه قوی بخش بندی، تا حد زیادی فرایند را به حل موفق مساله نزدیک می کند [۴]. از طرف دیگر الگوریتم های ضعیف یا خطا دار بخش بندی، تقریباً همیشه باعث خرابی اتفاقی می شوند. خروجی مرحله بخش بندی معمولاً، داده های پیکسلی خام است که یا مرز یک ناحیه یا تمام نقاط درون آن ناحیه را تشکیل می دهند. در هر دو حالت باید داده ها را به شکل مناسب برای پردازش رایانه ای تبدیل نمود. اولین تصمیمی که باید گرفته شود این است که آیا داده ها باید به صورت مرز یا به صورت یک ناحیه کامل نمایش داده شود. نمایش مرزی وقتی مفید است که مشخصات خارجی شکل نظیر گوشه ها یا خمیدگی ها مورد نظر باشد. نمایش ناحیه ای وقتی مفید است که خواص درونی بخش های تصویر نظیر بافت یا استخوان بندی شکل مورد توجه باشد. در هر حال در بعضی کاربردها هر دو نمایش به کار می رود. انتخاب یک روش نمایش، تنها قسمتی از راه تبدیل داده های خام به شکل مناسب برای پردازش بعدی رایانه ای است. توصیف^۵، که انتخاب ویژگی^۶ نیز خوانده می شود، به استخراج ویژگی هایی که مقداری از اطلاعات کمی مورد نظر را به ما می دهند یا برای تشخیص گروهی از اشیاء از گروه دیگر، اساسی هستند، می پردازد. مرحله آخر شکل ۱ شامل تشخیص و تعبیر است. تشخیص^۸ فرایندی است که بر اساس اطلاعات حاصل از توصیف گر ها یک برجسب را به یک شی منتسب می کند. تعبیر^۹ شامل انتساب معنا به یک مجموعه از اشیاء تشخیص داده شده است. دانش به شکل پایگاه داده دانش^{۱۰} در درون سامانه پردازش تصویر، ذخیره می شود. این دانش ممکن است تنها دانستن محل نواحی دارای جزئیات مورد علاقه باشد. بنابراین جستجوی مورد نیاز برای آن اطلاعات محدود می شود [۴]. پایگاه دانش ممکن است کاملاً پیچیده باشد، نظیر فهرست به هم مرتبط تمام نقایص اصلی ممکن در یک مساله بازرسی مواد یا یک پایگاه داده تصویری که حاوی تصاویر ماهواره ای تفکیک بالا از یک منطقه در ارتباط با کاربردهای آشکارسازی تغییر^{۱۱} باشد. پایگاه دانش علاوه بر هدایت عمل هر واحد پردازش، بر تعامل بین واحدها نیز نظارت می کند. این نمودار نشان می دهد که ارتباط بین واحدهای پردازش اغلب براساس دانش قبلی در مورد نتیجه پردازش است. این پایگاه دانش نه تنها عمل هر واحد را هدایت می کند، بلکه به عملیات بازخورد^{۱۲} بین واحدها نیز کمک می کند [۵].

ادبیات موضوع و مرور کارهای گذشته

افزایش آسان دسترسی به تصاویر دیجیتال در اینترنت و رشد ابزارهای قدرتمند برای ویرایش تصاویر دیجیتال سبب اهمیت موضوع تصاویر دیجیتال در دنیای امروز شده است. بنابراین، ویرایش تصاویر دیجیتال از اهمیت بالایی برخوردار است.

² Image acquisition

³ Imaging sensor

⁴ Segmentation

⁵ Eventual failure

⁶ Description

⁷ Feature selection

⁸ Recognition

⁹ Interpretation

¹⁰ Knowledge base

¹¹ Change detection

¹² Feed back

یک روش برای حل این مسئله استفاده از بازشناسی تصاویر بر اساس محتوا است. بازشناسی به عنوان تکنیکی قدرتمند اطلاعات اضافی در تصویر با حفظ کیفیت تصویر تعریف می‌شود.

یک بازشناسی دیجیتال می‌تواند قابل رویت و غیرقابل رویت باشد. یک بازشناسی قابل رویت اساساً شامل یک پیغام قابل رویت یا آرم شرکتی در تصویر است. از طرف دیگر، بازشناسی غیرقابل رویت مشابه تصویر اصلی به تصویر اضافه می‌شود. وجود یک بازشناسی غیرقابل رویت می‌تواند با استفاده از الگوریتم‌های آشکارسازی و استخراج مشخص شود [۴].

در اکثر کاربردهای بازشناسی در تصاویر دیجیتال، تصویر بازشناسی شده قبل از دریافت پردازش می‌شود. پردازش می‌تواند شامل مقایسه‌ی بی‌اتلاف، نویز افزایشی، بهبود و فیلترینگ تصویر باشد. یک راهکار بازشناسی ممکن است به صورت عمدی و غیرعمدی فرآیند پردازش را با مشکل مواجه کند. تصویر بازشناسی شده سپس تصویر بازشناسی شده‌ی تحت حمله نامیده می‌شود. روش اولیه در برابر حملات شامل روش سعی و خطا برای تخمین ترکیب مقادیر پیکسل‌ها می‌باشد که بیشترین اثر را بر روی آشکارکننده برای حداقل اعوجاج تصویر و سپس تخمین برای حذف متن دارد.

الگوریتم‌های بازشناسی تصاویر دیجیتال نیازمندی‌هایی همچون غیرمقاوم بودن و غیربصری بودن را شامل می‌شوند. غیرمقاوم بودن به معنای قابلیت آشکارکننده برای استخراج بازشناسی از تصویر است [۴].

با توجه به حوزه‌ای که تصویر بازشناسی شده است، تکنیک‌های بازشناسی دیجیتال می‌تواند در حوزه‌های طیفی و فضایی دسته‌بندی شود. در روش حوزه‌ی فضایی، یک متن با استفاده از اصلاح مستقیم پیکسل‌ها در تصویر وارد می‌شود. اولین روش فضایی تصویر را به دو گروه تقسیم می‌کند و سپس یک مقدار ثابت را به یک گروه از تصاویر وارد می‌کند. روش اصلاحی دیگر قرار دادن پیکسل‌های تصویر در موقعیت فضایی است. روش‌های برپایه‌ی روش فضایی نواقصی همچون ضعیف بودن در برابر فشرده‌سازی JPEG، نویز افزایشی و تبدیل‌های تصویر دارد.

روش‌های حوزه‌ی طیفی، تصویر را به حوزه‌ی فرکانس تبدیل می‌کنند و ضریب فرکانسی مدوله‌شده را به تصویر وارد می‌کنند. در کل، روش‌های حوزه‌ی طیفی نسبت به روش‌های حوزه‌ی فضایی مقاوم‌تر هستند. مزیت اصلی تکنیک‌های برپایه‌ی تبدیل آن است که تبدیل‌های تصویر انرژی تجزیه‌ی خوبی دارند و بیشتر انرژی تصویر می‌تواند در یک ناحیه‌ی محدود در حوزه‌ی تبدیل قرار گیرد. توابع پایه‌ی تبدیل منجر می‌شوند که این ضرایب اطلاعات دریافتی بیشتری از تصویر داشته باشند. محققین از تبدیلات مختلفی برای کاربردهای بازشناسی استفاده کرده‌اند. بعضی از تبدیلات مهم در ادامه ذکر می‌شود. تبدیل فوریه‌ی گسسته^{۱۳} اولین تبدیلی است که در بازشناسی مورد استفاده قرار گرفته است و می‌تواند چرخش، مقیاس و ... را بهبود بخشد. تبدیل کسینوسی گسسته^{۱۴} برای بازیابی متن در تصویر مورد استفاده قرار می‌گیرد. در طرحی که از DCT استفاده کرده است، یک تصویر به دو باند فرکانسی تقسیم شده است و متن در یک باند فرکانسی پائین به تصویر افزوده شده است [۶].

در این بخش به ارائه‌ی انواع روش‌های ارائه شده در حوزه‌ی بازشناسی متن در تصاویر دیجیتال می‌پردازیم. در ابتدا به بیان مقدمه‌ای از بحث حاضر پرداخته و در نهایت هر یک از روش‌های مطرح شده را مورد بررسی قرار می‌دهیم. در دنیای امروزی، اینترنت روشی برای دسترسی به اطلاعات و ایده‌های نوین است. اینترنت بستر مناسبی را برای موضوعات چندرسانه‌ای فراهم می‌کند. اما مسئله‌ای که در این بین وجود دارد، امنیت اطلاعات است. بازشناسی شاخه‌ای از امنیت اطلاعات برای قرار دادن مشخصات اطلاعاتی در یک بستر دیجیتال است. در این بخش ما یک روش هیبرید برای طرح-های بازشناسی بر مبنای تبدیل DCT را ارائه می‌نمائیم.

روش هیبرید برای الگوریتم بازشناسی

Sum و همکارانش [۷] یک روش هیبرید را برای طرح بازشناسی ارائه داده‌اند. از مشخصه‌ی ماتریس متعامد برای بازشناسی متن استفاده می‌شود. برای بازشناسی بزرگترین ضریب از ماتریس متعامد انتخاب می‌شود. البته در این حین از روش-

¹³ Discrete Fourier Transform(DFT)

¹⁴ Discrete Cosine Transform(DCT)

های اصلاحی نیز استفاده می‌گردد. پس از آن معکوس تبدیل DCT برای ساخت مجدد تصویر بازیابی شده انتخاب می‌گردد. به دلیل آنکه بزرگترین ضریب متعامد در کل فرآیند پردازش تصویر حضور دارد، تصویر بازیابی شده نمی‌تواند بطور کامل ساخته شود. کیفیت تصویر بازیابی شده می‌تواند به صورت تدریجی تعیین شود. برای استخراج متن، تبدیل DCT بکار می‌رود و بر مبنای بزرگترین ضریب، آزمایش صورت می‌گیرد. پس از آن متن استخراج می‌شود.

روش بازیابی متن

در گام اول تصویر اصلی در بلوک‌های جداگانه تجزیه می‌شود. در گام دوم تبدیل DCT به انجام می‌رسد. در گام سوم بزرگ‌ترین ضریب $S_i(1,1)$ از هر المان S استخراج می‌شود و از یک ضریب کیفیت Q برای انتخاب آن استفاده می‌گردد.

$$Y_i = S_i(1,1) \bmod Q$$

در گام بعدی وقتی ضریب وزن‌دهی صفر است ($w_i = 0$):

$$\text{if } Y_i < \frac{3Q}{4} \rightarrow S'_i(1,1) = S_i(1,1) + \frac{Q}{4} - Y_i$$

$$\text{else } S'_i(1,1) = S_i(1,1) + \frac{5Q}{4} - Y_i$$

وقتی که ضریب وزن‌دهی برابر یک است ($w_i = 1$):

$$\text{if } Y_i < \frac{3Q}{4} \rightarrow S'_i(1,1) = S_i(1,1) - \frac{Q}{4} + Y_i$$

$$\text{else } S'_i(1,1) = S_i(1,1) + \frac{3Q}{4} - Y_i$$

در گام آخر معکوس تبدیل DCT به کار برده می‌شود.

روش استخراج متن

در گام اول موقعیت تصویر متن‌گذاری شده در بلوک‌ها قرار داده می‌شود. در گام دوم تبدیل DCT صورت می‌گیرد. در گام سوم بزرگ‌ترین ضریب $S'_i(1,1)$ از هر المان S استخراج می‌شود.

$$Z = S'_i(1,1) \bmod Q$$

در گام بعدی اگر $Z < \frac{Q}{2}$ بود، متن استخراج شده دارای مقدار بیت صفر است. در غیراین صورت مقدار این بیت برابر ۱ در نظر گرفته می‌شود.

روش اصلاحی برای بازیابی

در [۸] از تکنیک LSR برای اصلاح روش بازشناسی متن استفاده شده است. به این دلیل از LSR استفاده شده است. این تکنیک به راحتی مورد استفاده قرار می‌گیرد و نتایج آن از شرایط مطلوبی برخوردار است.

اما روشی که برای بازیابی بر مبنای تکنیک LSR مورد استفاده قرار می‌گیرد، به صورت زیر است [۹]:

- ابتدا تصویر اصلی به چند فریم تبدیل می‌شود
- بر مبنای تعداد متن‌هایی که نیاز به پنهان‌سازی داریم، تصویر به شکل باینری تبدیل می‌شود.
- این تعداد سپس به صورت یک پیام در LSR ذخیره می‌شود.
- برای هر فریم پیغام در LSR ذخیره می‌شود.
- در انتها همه‌ی فریم‌ها مجدداً ترکیب شده و تصویر اصلی را ایجاد می‌کنند.

استفاده از چگالی انرژی در بازیابی

انرژی یک تصویر در حوزه‌ی فضایی در میان همه‌ی نمونه‌ها توزیع می‌شود. با این حال انرژی در حوزه‌ی تبدیل تنها به مقدار کمی از نمونه‌ها خلاصه می‌شود. چگالی انرژی^{۱۵} برای مقایسه‌ی تبدیلات استفاده می‌شود. چگالی انرژی، نرخ انرژی یک تصویر از حوزه‌ی تبدیل به حوزه‌ی فضایی است [۱۰ و ۱۱].

Xie و همکارانش از PSHT برای متن‌گذاری در تصاویر دیجیتال استفاده نموده‌اند. آن‌ها تصاویر را بلوک‌های جزءبندی کرده‌اند و تطبیق پارامتریک Slant برای هر بلوک را مورد استفاده قرار داده‌اند. آن‌ها از تکنیک حوزه‌ی طیفی در روش بازیابی و استفاده از متوسط همبستگی بین متن استخراج‌شده و متن اصلی استفاده کرده‌اند. انتخاب پارامترهای تبدیل یکی از بخش‌های اصلی متن‌گذاری در حوزه‌ی تبدیل می‌باشد. Xie از یک روش بهینه‌ی برمبنای تقریب جزئی استفاده کرده است. نتایج گزارش شده توسط وی نشان می‌دهد که در طرح متن‌گذاری PSHT عملکرد بهتری را در همه‌ی درجه‌های فشرده‌سازی دارا می‌باشد [۱۲].

روش پیشنهادی

در این قسمت می‌خواهیم به بررسی روش پیشنهادی برای بازیابی تصاویر بر اساس محتوا بپردازیم. برای این منظور در ابتدا تئوری تبدیل را به منظور تبدیل تصویر به تصویر دیجیتال ارائه می‌نمائیم. سپس روش بهینه‌سازی پیشنهادی را مطرح می‌کنیم و در انتها با ارائه روش پیشنهادی، نتایج شبیه‌سازی را ارائه می‌نمائیم. در شکل ۲ بلوک دیاگرام روش پیشنهادی آورده شده است.

همان‌طور که از شکل ۲ مشخص است، الگوریتم و روش پیشنهادی برای بازیابی متن از تصویر از چندین گام اصلی تشکیل شده است که به قرار زیر است:

- گام اول: مطابق با هر الگوریتم در حوزه‌ی پردازش تصویر، لازم است که تصویر به یک تصویر دیجیتال تبدیل شود. برای این کار الگوریتم تبدیل مویک مورد استفاده قرار گرفته است. تبدیل مویک یکی از انواع تبدیل‌های تصویر است که برای استخراج ویژگی از تصویر مناسب است.

- گام دوم: در این مرحله یک متن از پیش تعیین‌شده در تصویر جاسازی می‌شود. برای این کار لازم است که متن نیز به صورت یک متن تبدیل شده به دیجیتال قرار داشته باشد. برای جایگذاری متن در تصویر از روش جاگذاری پیکسل استفاده می‌شود.

- گام سوم: در این مرحله پارامترهای کیفیت تصویر محاسبه می‌شود. سه پارامتر حساسیت فرکانسی، حساسیت لبه‌ای و حساسیت شفافیت را به عنوان پارامترهای ارزیابی کیفیت تصویر در نظر گرفته‌ایم. در صورت برآورده نشدن هر یک از این پارامترها لازم است که از الگوی بهینه‌سازی استفاده شود.

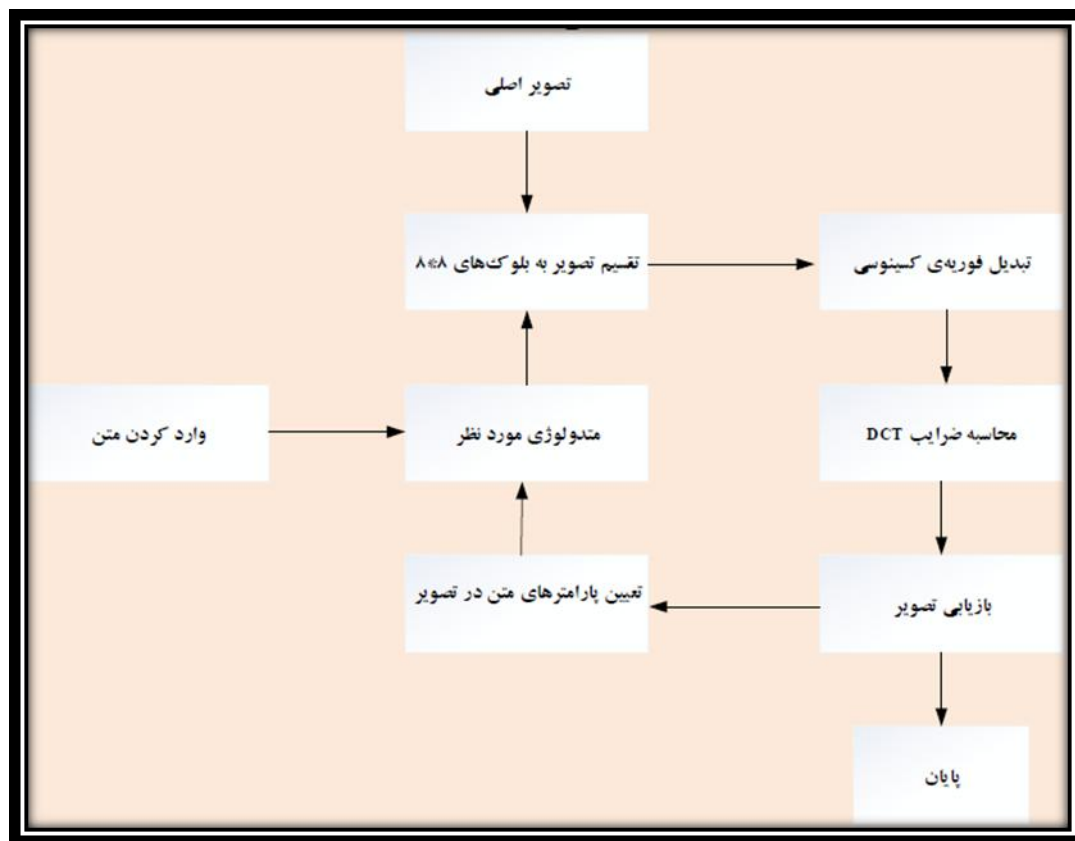
- گام چهارم: در این مرحله الگوریتم استخراج متن از تصویر مورد استفاده قرار می‌گیرد. البته پیش از آن باید توجه داشته باشیم که تصویر متن‌گذاری شده باید با استفاده از معکوس تبدیل مویک به یک تصویر حوزه‌ی زمان تبدیل شود و به عنوان ورودی الگوریتم استخراج متن در نظر گرفته شود. الگوی استخراج متن بر اساس ضرایب تبدیل مویک قرار دارد.

- گام پنجم: در این مرحله متن استخراج شده با متن مرجع مقایسه می‌شود. در صورت وجود خطا الگوریتم بهینه‌سازی عصبی-فازی مورد استفاده قرار می‌گیرد. الگوریتم بهینه‌سازی مستقیماً بر روی تبدیل مویک پیاده‌سازی می‌شود تا ضرایب به دست آمده از تبدیل مویک بهینه باشد.

نکته‌ی اساسی در این قسمت آن است که عنوان داشته باشیم الگوریتم بهینه‌سازی یک الگوریتم آموزش‌پذیر است و در نهایت در صورت دستیابی به اهداف مفروض می‌توان این الگو را برای نمونه‌های دیگر از استخراج متن نیز مورد استفاده قرار داد.

¹⁵ Energy Concentration(EC)

در ادامه‌ی این فصل هر یک از گام‌های بیان شده به صورت مفصل و با جزئیات مورد بررسی قرار خواهد گرفت.



شکل ۲: بلوک دیاگرام روش پیشنهادی

تبدیل کسینوسی گسسته

اما در این قسمت به ارائه‌ی تبدیل مورد نیاز برای بخش تصویر و متن می‌پردازیم. مطابق با آنچه که در بلوک دیاگرام شکل ۲ نیز آورده شده است، قدم اول در هر فرآیند پردازشی تبدیل تصویر از حوزه‌ی زمان به حوزه‌ی فرکانس است. مطابق با آنچه که بیان شد، رویکرد اصلی در این تحقیق آن است که در ابتدا یک متن در تصویر جاسازی شود و سپس با یک رویکرد بهینه متن از تصویر استخراج شود و با متن اصلی مقایسه شده و صحت آن مورد ارزیابی قرار گیرد.

دلیل آن که در فرآیند پردازش تصویر تصویر از حوزه‌ی زمان به حوزه‌ی فرکانس تبدیل می‌شود آن است که اساساً در حوزه‌ی تصویر مقادیر پیکسل‌ها اهمیت دارد و هر فعل و انفعال ریاضی بر روی پیکسل‌های تصویر صورت می‌گیرد. بنابراین تبدیل این کمک را می‌کند که بتوانیم مقادیر پیکسل‌های تصویر را به دست آوریم.

اما همانطور که بیان شد، برای تبدیل تصویر از حوزه‌ی زمان به حوزه‌ی فرکانس استفاده می‌شود که اصلی‌ترین و متداول‌ترین آن‌ها تبدیل فوری کسینوسی است. این تبدیل در ادامه مورد بررسی قرار خواهد گرفت.

تبدیل کسینوسی گسسته برای تبدیل سیگنال از حوزه‌ی زمان به حوزه‌ی فرکانس مورد استفاده قرار می‌گیرد. باید توجه داشته باشیم در صورتی که ضرایب فرکانسی مورد استفاده در جاسازی تصاویر دارای مقادیر پائین باشند، در این صورت جاسازی تصاویر با مشکلاتی همراه است. اما در صورتی که از ضرایب فرکانسی با مقادیر بالا استفاده کنیم، می‌توانیم در برابر حملات صورت گرفته امنیت بیشتری را ایجاد کنیم. تا آنجا که الگوریتم به دست آمده مقاوم بوده و پایداری بیشتری را در برابر حملات از خود نشان دهد.

تبدیل کسینوسی استفاده شده در پردازش تصاویر دیجیتال برای یک تصویر A با ابعاد $N \times N$ از روابط زیر به دست می‌آید:

$$B_{pq} = \alpha p \alpha q \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} A_{mn} \left(\frac{\cos(2\pi m + 1)p}{2M} \right) \left(\frac{\cos(2\pi n + 1)q}{2N} \right); \quad (1)$$

$$0 \leq p \leq M-1, 0 \leq q \leq N-1$$

$$\alpha p = \begin{cases} \frac{1}{\sqrt{M}} & ; p = 0 \\ \frac{2}{\sqrt{M}} & ; 1 \leq p \leq M-1 \end{cases} \quad (2)$$

$$\alpha q = \begin{cases} \frac{1}{\sqrt{N}} & ; q = 0 \\ \frac{2}{\sqrt{N}} & ; 1 \leq q \leq N-1 \end{cases} \quad (4)$$

p و q متغیرهایی هستند که به ترتیب دارای مقادیر $1 \leq p \leq M-1$ و $1 \leq q \leq N-1$ می‌باشند. بنابراین معکوس تبدیل اشاره شده از رابطه‌ی ۵ به دست می‌آید:

$$A_{mn} = \alpha p \alpha q \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} B_{pq} \left(\frac{\cos(2\pi m + 1)p}{2M} \right) \left(\frac{\cos(2\pi n + 1)q}{2N} \right); \quad (5)$$

$$0 \leq p \leq M-1, 0 \leq q \leq N-1$$

همان‌طور که از معادلات ۱ تا ۴ نیز مشخص است، برای تبدیل فوری‌ی کسینوسی قدم اول تعیین ضرایب تبدیل است. در رابطه‌ی ۱ نحوه‌ی محاسبه‌ی ضرایب آورده شده است. این فرم یک ساختار استاندارد است. اما با به دست آمدن ضرایب، این ضرایب در معادله‌ی ۴ وارد شود، رابطه‌ی تبدیل فوری‌ی برای یک تصویر نمونه‌ی A به دست می‌آید. نکته‌ی با اهمیت در این قسمت آن که در ادامه و در بحث بهینه‌سازی عملیات متغیرهای ما همین ضرایب فوری‌ی هستند. با تغییر این ضرایب فوری‌ی، مقادیر پیکسل‌ها تغییر کرده و در نهایت می‌توانیم پارامترهای تصویر و متن را تغییر داده و بر اساس اهداف متصور و از قبل تعیین شده، تصویر و متن را اصلاح نمائیم. این رویکرد در ادامه مورد بحث و بررسی قرار خواهد گرفت. اما با این معرفی به بخش جاسازی متن و استخراج متن از تصویر می‌رسیم. در بخش بعدی این رویکرد مورد بحث و بررسی قرار خواهد گرفت.

الگوریتم جاسازی متن در تصویر

الگوریتم‌های متن‌گذاری و بازیابی متن به صورت متداول در تحقیقات قبلی ارائه شده است. باید توجه داشته باشیم که هدف ما از ارائه‌ی این مقاله ارائه‌ی الگوریتمی برای بهبود متن استخراج شده از تصویر است. بنابراین الگوریتم جاسازی و بازیابی متن مطابق با تحقیقات قبلی صورت گرفته است و در اینجا به صورت مختصر به آن اشاره می‌شود.

الگوریتم جاسازی

الگوریتم جاسازی پیشنهادی در گام‌های زیر خلاصه می‌شود:

- وارد کردن تصویر اصلی به الگوریتم
- تقسیم تصویر به دو بخش 8×8
- استفاده از تبدیل فوری‌ی کسینوسی برای تبدیل تصویر به تصویر دیجیتال
- محاسبه‌ی ضریب DCT برای تصویر
- وارد کردن متن به متدولوژی مورد نظر
- وارد کردن متن به تصویر. در صورتی که پارامترهای حساسیت در مقدار مطلوبی قرار داشته باشد، الگوریتم خاتمه می‌یابد، در غیر این صورت با روش ارائه شده پیکسل‌های متن در تصویر اصلی تغییر می‌کند.

بعد از این مراحل الگوریتم معکوس DCT برای هر بخش اعمال می‌شود تا تصویر به شکل اصلی تبدیل شود. اما مطابق با آنچه که در شکل ۳ آورده شده است، در صورتی که بخواهیم الگوریتم جاگذاری متن در تصویر را به صورت یک شبه‌کد در نظر بگیریم، می‌توانیم به صورت زیر این شبه‌کد را نمایش دهیم:

```

1. INPUT: input image A and input text T
2. Calculation size(A) = Plen(A)
3. Divide Plen(A) to 8 & set Aj = {A1, A2, ..., A8}
4. DCT for A & set Bj = {B1, B2, ..., B8}
5. DCT for T & set Tj = {T}
6. A{Tj}j = TjUAj
7. AT = DCT - 1(A{Tj}j)
8. Calculate Sensitive Parameter AT(P)
AT(P) = {AT(F), AT(E), AT(T)}
∀ AT(P) ∈ {F, E, T}
Stop Algorithm
9. Else
    ANN, weighting factor Bj
    Fuzzy logic, membership level Bj
10. Go to step 4
11. OUTPUT: Image texture A(T)

```

شکل ۳: شبه‌کد الگوریتم جاسازی

در رویکرد پردازش تصویر اولین اقدام و عملکردی که باید مد نظر قرار گیرد، بحث تبدیل تصویر به حوزه‌ی فرکانس است یا به اصطلاح رایج تبدیل به یک تصویر دیجیتال تبدیل شود. همان‌طور که اشاره شد روش‌های مختلفی در این زمینه وجود دارد که رایج‌ترین آنها استفاده از تبدیل فوری‌ی کسینوسی است. در این تبدیل در ابتدا ضرایب فوری‌ی به دست می‌آید و سپس تصویر بر اساس این ضرایب به حوزه‌ی فرکانس تبدیل می‌شود. بعد از این مرحله و پس اینکه پیکسل‌های تصویر و متن به دست آمد، طول تصویر محاسبه می‌شود و تصویر به ۸ قسمت تبدیل می‌شود. در مرحله‌ی بعدی پیکسل‌های متن در پیکسل‌های تصویر اصلی جاسازی می‌شود. در مرحله‌ی بعدی از معکوس تبدیل فوری‌ی استفاده می‌شود و تصویر متن‌گذاری شده به حوزه‌ی زمان تبدیل می‌شود. در این مرحله پارامترهای حساسیت فرکانسی (F)، حساسیت لبه‌ای (E) و حساسیت شفافیت (T) محاسبه می‌شود و با مقادیر مطلوب که ابتدائاً و به صورت اختیاری در نظر گرفته‌ایم، مقایسه می‌شود. اگر مقادیر مورد نظر برآورده شود، الگوریتم خاتمه می‌یابد در غیر این صورت رویکرد ترکیبی عصبی-فازی مورد استفاده قرار می‌گیرد. در این روش ضرایبی تحت عنوان ضرایب وزن-عضویت به ضرایب فوری‌ی اعمال می‌شود و این مرحله تا زمان دستیابی به مقادیر مطلوب حساسیت ادامه می‌یابد.

الگوریتم بازیابی

الگوریتم استخراج متن‌گذاری در گام‌های زیر خلاصه شده است:

- وارد کردن تصویر متن‌گذاری شده به الگوریتم
- تقسیم تصویر به دو بلوک 8×8
- استفاده از تبدیل فوری‌ی کسینوسی برای تبدیل تصویر به یک تصویر دیجیتال
- محاسبه‌ی ضرایب DCT برای تصویر
- محاسبه‌ی پارامترهای حساسیت برای متن
- در صورتی که مقادیر صحیح باشد، پایان الگوریتم در غیر این صورت استفاده از روش فازی مورد نظر برای اصلاح متن-

گذاری

بلوک دیاگرام الگوریتم بازیابی متن در شکل ۴ نشان داده شده است. اما مطابق با آنچه که در شکل ۴ نشان داده شده است، در زیر شبه کد مورد نظر بر اساس این بلوک دیاگرام آورده شده است که مراحل اساسی روش بازیابی متن از تصویر را نشان می‌دهد.

```

1. INPUT: Image texture A(T)
2. Calculation size (A) = Plen(A)
3. Divide Plen(A) to 8 & set Aj = {A1, A2, ..., A8}
4. DCT for A & set Bj = {B1, B2, ..., B8}
5. DCT for T & set Tj = {T}
6. A{T}j = TjUAj
7. AT = DCT - 1(A{T}j)
8. Calculate Sensitive Parameter AT(P)
AT(P) = {AT(F), AT(E), AT(T)}
∀ AT(P) ∈ {F, E, T}
Stop Algorithm
9. Else
    ANN, weighting factor Bj
    Fuzzy logic, membership level Bj
10. Go to step 4
11. OUTPUT: Image (A)

```

شکل ۴: شبه کد الگوریتم بازیابی

همان‌طور که از شبه‌کد فوق مشخص است، ورودی این الگوریتم یک تصویر متن‌گذاری شده است. در گام نخست تصویر به ۸ بخش تقسیم می‌شود. سپس در مرحله‌ی بعدی تبدیل فوریه‌ی کسینوسی مورد استفاده قرار می‌گیرد تا مقادیر پیکسل‌های تصویر متن‌گذاری شده به‌دست آید. حال بر اساس قاعده‌ی جداسازی لبه‌ای و متوسط آماری مقادیر پیکسل‌ها، متن از تصویر استخراج می‌شود. اما برای سنجش کیفیت و دقت متن استخراج شده کفایت پارامترهای حساسیت برای آنها محاسبه شود. در صورت مناسب نبودن مقادیر این پارامترها، لازم است الگوریتم بهینه‌سازی پیشنهادی مورد استفاده قرار گیرد. در ابتدا الگوریتم عصبی بر اساس قاعده‌ی جست‌وجوی گلوبال محدوده‌ی پیکسل‌های متن را به‌دست می‌آورد و سپس الگوریتم فازی بر اساس قاعده‌ی جست‌وجوی محلی بهترین مقدار را برای این پیکسل‌ها به‌دست می‌آورد. ماحصل روش‌های عصبی و فازی با بکارگیری ضرایب عضویت و وزن‌دهی بر روی تصویر و متن، اجرا می‌شود. در نهایت خروجی الگوریتم یک متن استخراج شده است تصویر با کیفیت مناسب و مطلوب است. اما بعد از این که به معرفی الگوریتم‌های متداول برای جاسازی و بازیابی متن از تصویر اشاره کردیم و رویکرد اجرایی آن‌ها مشخص شد، در بخش بعدی رویکرد بهینه‌سازی برای بهبود مقوله‌ی جاسازی و بازیابی متن از تصویر به‌همراه رویکرد بهینه‌سازی هر روش برای مسئله‌ی مورد بحث ما ارائه می‌شود.

رویکرد بهینه‌سازی

در ارتباط با رویکرد مورد بحث که همان مقوله‌ی جاسازی و بازیابی بهینه‌ی متن از تصویر است، مشخص شد که الگوریتم دو الگوریتم متداول برای جاسازی و بازیابی متن از تصویر وجود دارد که این سناریو نمی‌تواند به عنوان یک سناریوی بهینه مطرح باشد. برای حل این مسئله در این تحقیق و سایر تحقیقات دیگر رویکرد بهینه‌سازی و استفاده از الگوهای بهینه‌سازی مطرح شده است.

اما قبل از اینکه بخواهیم مهم‌ترین سناریوهای بهینه‌سازی و همین‌طور سناریوی پیشنهادی در این تحقیق را مطرح کنیم، لازم است رویکرد بهینه‌سازی اجرایی بر روی مدل و همچنین متغیرها و محدودیت‌های بهینه‌سازی را مطرح کنیم.

در هر الگوریتم بهینه‌سازی لازم است که توابع هدف، متغیرها و محدودیت‌های بهینه‌سازی مشخص شود. در این راستا، در تحقیق مورد بررسی و برای جاسازی و بازیابی متن از تصویر، متغیرها همان مقادیر پیکسل‌های تصویر و متن هستند. محدودیت‌ها بازه‌ی تغییرات قابل قبول برای پیکسل‌ها هستند. به عبارت دیگر با توجه به ماهیت تصویر و متن، نمی‌توان مقدار یک پیکسل را هم برای متن و هم برای تصویر از بازه‌ی مشخصی بیشتر تغییر داد. اما سومین مسئله، توابع هدف هستند. توابع هدف پارامترهایی هستند که می‌خواهیم با بهینه‌سازی و تغییر متغیرها به آنها برسیم. توابع هدف را پیک سیگنال به نویز برای تصویر، نرخ خطای بیت، امنیت، سرعت و زمان پردازش در نظر گرفته‌ایم که بخش بعدی به همراه پارامترهای کیفیت ارائه خواهد شد.

اما برای بهینه‌سازی باید توجه داشته باشیم که تغییر مقادیر پیکسل‌ها با استفاده از یک رویکرد مشخص برای هر الگوریتم بهینه‌سازی صورت می‌گیرد. یعنی در هر الگوریتم بهینه‌سازی، بر اساس یک ساختار ریاضی مشخص ضرایبی تولید و بر روی مقادیر پیکسل‌های تصویر وارد می‌شود که منجر به تغییر پیکسل‌ها و دست‌یابی به توابع هدف مربوطه شود. این فرآیند تولید ضرایب تکرارپذیر است و تا زمان دست‌یابی به مقادیر مطلوب از نقطه‌نظر توابع هدف ادامه خواهد یافت. اما بعد از ارائه‌ی این مقدمه، به بعضی از مهم‌ترین الگوریتم‌های بهینه‌سازی قابل استفاده برای مسئله‌ی مطرح می‌پردازیم و در بخش نتیجه‌گیری و ارائه‌ی نتایج، به‌عنوان معیار مقایسه با الگوریتم پیشنهادی مان مقایسه می‌کنیم.

الگوریتم بهینه‌سازی ژنتیک

یکی از انواع الگوریتم‌های بهینه‌سازی، الگوریتم ژنتیک می‌باشد [۱۳]. برای تغییر متغیرها در الگوریتم‌های بهینه‌سازی، از یک سری ضرایب استفاده می‌شود که در الگوریتم ژنتیک این ضرایب با نام کروموزوم‌ها شناخته می‌شوند. کروموزوم‌ها دنباله‌ای از اعداد هستند که بر اساس قواعد جهش و ترکیب، در هر مرحله به‌روزرسانی می‌شوند. طول کروموزوم‌ها اختیاری است اما در رویکرد این تحقیق و برای بکارگیری الگوریتم ژنتیک، طول هشت برای کروموزوم در نظر گرفته شده است. اما برای کروموزوم‌ها لازم است که یک مقدار اولیه در نظر گرفته شود. در اینجا این مقدار اولیه را برابر صفر در نظر گرفته‌ایم. به این ترتیب، در هر مرحله با توجه به رویکرد جهش و ترکیب که ساختار و اجرای آن در فصل دوم توضیح داده شده است، این مقادیر به‌روزرسانی می‌شود.

اما لازم است این توضیح کلی را بیان کنیم که بهینه‌سازی ژنتیک یک الگوریتم محلی است که در مرحله‌ی بازیابی متن از تصویر و بر روی پیکسل‌های متن و تصویر وارد می‌شود.

الگوریتم بهینه‌سازی PSO

یکی دیگر از الگوریتم‌های بهینه‌سازی محلی الگوریتم PSO است [۱۴]. این الگوریتم برای بهینه‌سازی از بکارگیری ضرایب سرعت ذره برای رویکرد بهینه‌سازی استفاده می‌کند. الگوریتم جمعیت ذرات را به‌وجود می‌آورد. هر ذره نمایان‌گر راه حل پتانسیلی به مشکل بهینه خواهد بود. بگذارید s به اندازه‌ی گروه باشد. این ویژگی‌ها با نشانه‌های زیر بیان می‌شود:

x_i : موقعیت رایج ذره

v_i : سرعت معمول ذره

y_i : بهترین موقعیت خصوصی ذره

بهترین موقعیت ذره که با ذره‌ی i در ارتباط است، زمانی بهترین خواهد شد که میزان بالاترین تناسب در نظر گرفته شود.

سمبل F را برای نشان دادن تابع هدف، که کم‌کم کوچک‌تر می‌شود استفاده خواهیم کرد. معادله به‌روزرسانی برای بهترین موقعیت اختصاصی تابع‌ها با وابستگی به گام زمانی به‌صورت زیر نمایش داده می‌شود:

(۶)

$$y_i(t+1) = \begin{cases} y_i(t) & \text{if } f(x_i(t+1)) \geq f(y_i(t)) \\ x_i(t+1) & \text{if } f(x_i(t+1)) < f(y_i(t)) \end{cases}$$

در PSO دو مقوله وجود دارد که مدل‌های G_{best} و L_{best} نامیده می‌شود. تفاوت بین دو الگوریتم بر اساس مجموعه‌ای از ذرات است که با این ذرات به وجود آمده مستقیماً تعامل برقرار می‌کنند که سمبل Y برای نشان دادن این تعامل استفاده می‌شود.

درباره‌ی جزئیات دو مدل بطور کامل در زیر بحث خواهد شد. تعریف Y همانطور که در مدل G_{best} استفاده می‌شود، در معادله ۷ آورده شده است:

(۷)

$$\hat{y}(t) \in \{y_1(t), y_2(t), \dots, y_s(t)\} | f(\hat{y}(t)) = \min \in \{f(y_1(t)), f(y_2(t)), \dots, f(y_s(t))\}$$

توجه داشته باشید که این تعریف بیان‌کننده‌ی این مطلب است که $\hat{y}$ بهترین موقعیتی است که تاکنون توسط هر ذره کشف شده است.

الگوریتم از دو دنباله‌ی تصادفی مستقل استفاده می‌کند: $r_1 \approx U(0,1)$ ، $r_2 \approx U(0,1)$. که این دنباله‌ها برای تاثیر گذاشتن روی طبیعت تصادفی الگوریتم استفاده می‌شود. مقادیر r_1 و r_2 با ثابت‌های $C_1 \geq 0$ و $C_2 \leq 2$ مقیاس‌بندی شده‌اند. این ثابت‌ها را ضریب تسریع‌کننده می‌نامیم که روی اندازه‌ی ماکزیمم گامی تاثیر می‌گذارد که ذره بتواند یک تکرار داشته باشد. گام سرعت به‌روزرسانی بویژه برای هر بعد مجزا شده است. $j \in 1, n$ است، بنابراین $v_{i,j}$ بیان‌کننده‌ی این است که j^{th} بعد محور سرعت با ذره i^{th} در ارتباط می‌باشد. معادله‌ی به‌روزرسانی از قرار زیر است:

(۸)

$$v_{i,j}(t+1) = v_{i,j}(t) + c_1 r_{1,j}(t) [y_{i,j}(t) - x_{i,j}(t)] + c_2 r_{2,j}(t) [\hat{y}_j(t) - x_{i,j}(t)]$$

از تعریف به‌روزرسانی سرعت روشن است که C_1 اندازه‌ی گام ماکزیمم را در مسیر بهترین ذره‌ی کلی تنظیم می‌کند و C_2 اندازه‌ی گام در جهت بهترین موقعیت شخصی آن ذره را تنظیم می‌کند. میزان $v_{i,j}$ به برد محدود شده است تا ذره مکان جست‌وجو را ترک نکند. اگر مکان جست‌وجو با حوزه‌های $[-x_{max}, x_{max}]$ تعریف شود، سپس اندازه‌ی v_{max} نوعاً به گونه‌ای قرار داده می‌شود که $v_{max} = k \times x_{max}$ که در آن $[0.1 \leq k \leq 1.0]$. موقعیت هر ذره با استفاده از بردار سرعت جدید برای ذره به‌هنگام می‌شود. بنابراین:

(۹)

$$x_i(t+1) = x_i(t) + v_i(t+1)$$

الگوریتم شامل استفاده تکراری از معادلات به روز شده که در بالا ارائه شد می‌باشد. مطابق با توضیحات بیان شده، در این الگوریتم ضرایب با توجه به رویکرد به‌روزرسانی در هر مرحله بر روی مقادیر پیکسل‌ها وارد می‌شوند و این به‌روزرسانی و تکرارهای متوالی در نهایت با برآورده شدن مقادیر هدف مورد انتظار، بهینه‌سازی خاتمه می‌یابد. توجه شود که مقدار اولیه‌ی سرعت برای ذره‌ها در این الگورتیم اختیاری است که در تحقیق ما، این مقدار را برابر صفر در نظر گرفته‌ایم. در PSO و با توجه به مسئله‌ی ما، ذرات همان پیکسل‌ها هستند.

الگوریتم بهینه‌سازی خوشه‌بندی

دو الگوریتم قبلی الگوریتم‌های محلی در حوزه‌ی بهینه‌سازی بودند. اما الگوریتم خوشه‌بندی الگوریتم بهینه‌سازی سراسری یا گلوبال است و همانطور که در فصل دوم نیز بیان شد، برای دسته‌بندی اطلاعات به‌صورت قاعده‌مند مورد استفاده قرار می‌گیرد. در مطالعه‌ی فصل دوم انواع الگوریتم‌های خوشه‌بندی و قواعد آن مشخص شده است. اما در ارتباط با تحقیق مورد بررسی، از خوشه‌بندی می‌توان برای تبدیل تصویر متن‌گذاری شده به‌صورت بخش‌های کوچکتر استفاده کرد. در این راستا برای تصویر متن‌گذاری شده و برای جست‌وجو برای یافتن پیکسل‌های متن، خوشه‌بندی این مدل را به تعداد دلخواه از دسته‌های پیکسل تبدیل می‌کند. معیار خوشه‌بندی می‌تواند پارامترهای مختلفی باشد که در این تحقیق موقعیت پیکسل‌ها به عنوان معیار خوشه‌بندی در نظر گرفته شده است. تصویر را به ۶۴ خوشه‌ی دلخواه تبدیل کرده و رویکرد استخراج متن از تصویر بر روی آن اجرا شده است.

الگوریتم بهینه‌سازی عصبی

اما یکی دیگر از الگوریتم‌های بهینه‌سازی سراسری الگوریتم عصبی است که این الگوریتم به عنوان الگوریتم ابتدایی برای روش پیشنهادی مطرح است. در الگوریتم بهینه‌سازی برای اجرای بهینه‌سازی از ضرایب وزن‌دهی استفاده می‌شود. روش عصبی مورد استفاده در این تحقیق الگوریتم عصبی RBF است که رویکرد به‌روزرسانی ضرایب وزن‌دهی در آن بر اساس معادله‌ی دایره است. این الگوریتم به همراه مثال آموزش و آزمایش در فصل دوم به تفصیل مورد بررسی قرار گرفت.

اما در اجرای این تحقیق، روش عصبی برای تغییر درجه‌ی عضویت پارامتری را با نام ضریب وزن‌دهی مورد استفاده قرار می‌دهد. ضریب وزن‌دهی در هر مرحله بر مقدار درجه‌ی عضویت ورودی وارد شده و با تغییر آن، درجه‌ی عضویت جدید را تولید می‌کند. بعد از هر مرحله ضریب وزن‌دهی به‌روزرسانی شده و وارد مرحله‌ی بعدی می‌شود. در روابط زیر نحوه‌ی به‌روزرسانی ضریب وزن‌دهی و همچنین به‌روزرسانی درجه‌ی عضویت نشان داده شده است.

$$\omega_{new} = \omega_{old} + \alpha(x_{old}) + b$$

$$x_{new} = x_{old} + \alpha(1 - \omega_{new}) + b$$

در روابط فوق ω ضریب وزن‌دهی، x درجه‌ی عضویت، α یک مقدار ثابت بین صفر و یک و b بایاس شبکه‌ی عصبی است که معمولاً مقداری برابر با یک دارد.

مقدار اولیه برای وزن را صفر در نظر گرفته شده است. به این ترتیب این الگوریتم با توجه به مقدار پارامترهای هدف، فرآیند بهینه‌سازی را صورت می‌دهد. توجه شود که در فاز نخست، الگوریتم عصبی مورد استفاده قرار گرفته و با توجه به رویکرد بهینه‌سازی سراسری، به مقدار هدف نزدیک می‌شود و در ادامه با به‌کارگیری الگوریتم فازی که یک الگوریتم بهینه‌سازی محلی است، پاسخ بهینه حاصل خواهد شد.

الگوریتم فازی

اما الگوریتم دیگر برای بهینه‌سازی الگوریتم فازی است که فاز نهایی بهینه‌سازی در رویکرد پیشنهادی را تشکیل می‌دهد. این الگوریتم یک الگوریتم بهینه‌سازی محلی است که معمولاً در کنار یک الگوریتم بهینه‌سازی سراسری عملکرد مطلوبی را از خود به نمایش می‌گذارد.

اما در این قسمت به روش سیستم فازی مورد نظرمان می‌پردازیم. همان‌طور که در بخش قبلی نیز ذکر کردیم، هدف ما استفاده از ترکیبی از روش‌های فازی و عصبی برای بهینه‌سازی پارامترهای جاسازی و بازیابی متن از تصویر است. در ارتباط با روش اصلی مورد نظرمان برای جاسازی و بازیابی در بخش بعدی توضیح می‌دهیم. اما در این بخش به استفاده از روش فازی می‌پردازیم.

روش فازی یکی از روش‌های بهینه‌سازی بر مبنای هوش مصنوعی است. در این روش فضای نمونه جست‌وجو شده و به صورت تصادفی مقدار مورد نظر به‌دست می‌آید. در ارتباط با روش فازی ذکر این نکته از اهمیت بالایی برخوردار است که نقطه‌ی

شروع الگوریتم تعریف یک درجه‌ی عضویت برای هر یک از پیکسل‌های تصویر و متن است. به عبارت دیگر در این روش، هر پیکسل بر مبنای مقداری که از نظر تصویر دارد، یک درجه تخصیص داده می‌شود. بر مبنای این درجه، فضای تصویر مورد بررسی قرار می‌گیرد. در هر مرحله مقدار این درجه‌ی عضویت تعریف شده برای هر پیکسل، به صورت تصادفی و در یک محدوده‌ی مشخص که شعاع عضویت نامیده می‌شود، تغییر می‌کند. سپس در هر زمان که تابع هدف ما به بهترین جواب دست یابد، درجه‌ی عضویت در نظر گرفته شده برای هر پیکسل به آن پیکسل اضافه می‌شود و با این شرایط متن از تصویر استخراج می‌شود. توجه داشته باشیم که تابع هدف ما در این الگوریتم میزان پیک سیگنال به نویز است و بر این اساس هر زمان که این پارامتر در بهترین مقدار خود قرار داشته باشد، شرط خاتمه‌ی الگوریتم برآورده می‌شود.

اما الگوریتم مورد استفاده‌ی ما برای بهینه‌سازی فازی به صورت زیر است:

ما در الگوریتم فازی پیشنهادی، الگوریتم فازی درختی پیشنهاد داده‌ایم که دارای قابلیت‌هایی همچون سرعت و دقت بالای پردازش است. یک بخش فازی یک بخش تعمیم‌یافته از crisp می‌باشد. crisp تنها عضویت کامل و بدون عضویت را شامل می‌شود، در حالیکه فازی عضویت جزئی را نیز مورد پذیرش قرار می‌دهد. در یک بخش crisp، عضویت یک پیکسل x از یک تصویر A به وسیله‌ی تابع عضویت $u_A(x)$ بیان می‌گردد:

$$u_A(x) = \begin{cases} 1 & ; \text{if } x \in A \\ 0 & ; \text{if } x \notin A \end{cases}$$

تئوری فازی این فرضیه را با استفاده از تعریف یک پیکسل تعمیم می‌دهد که می‌تواند مقداری در بازه‌ی 0 تا 1 داشته باشد. یک تابع عضویت به صورت مرسوم به شکل زیر تعریف می‌شود:

$$u_A: X \rightarrow [0,1]$$

جائیکه x اشاره به پیکسل‌های یک تصویر مشخص دارد. فرض کنیم که A و B دو بخش فازی با توابع عضویت $u_A(x)$ و $u_B(x)$ می‌باشند. عملگرهای فازی می‌تواند به صورت زیر تعریف شود:

$$u_{A \cap B}(x) = u_A(x) \tau u_B(x)$$

جائیکه τ یک عملگر t-norm است. τ یک تابعی از $[0,1] * [0,1] \rightarrow [0,1]$ است و باید شرایط زیر را برآورده کند:

$$a \tau 1 = a$$

$$a \tau b = b \tau a$$

$$a \tau b \geq c \tau d \text{ if } a \geq c, \quad b \geq d$$

$$a \tau b \tau c = a \tau (b \tau c) = (a \tau b) \tau c$$

اجتماع:

$$u_{A \cap B}(x) = u_A(x) \rho u_B(x)$$

جائیکه ρ یک عملگر s-norm است. ρ یک تابعی از $[0,1] * [0,1] \rightarrow [0,1]$ است و باید شرایط زیر را برآورده کند:

$$a \rho 1 = a$$

$$a \rho b = b \rho a$$

$$a \rho b \geq c \rho d \text{ if } a \geq c, \quad b \geq d$$

$$a \rho b \rho c = a \rho (b \rho c) = (a \rho b) \rho c$$

برش- α :

$$A_\alpha(x) = \{x \in X | u_A(x) \geq \alpha\}$$

جائیکه A_α یک برش- α از بخش فازی A است. بنابراین A_α تعامل تمام المان‌های مجموعه‌ی فراگیر X است که درجه‌ی عضویت در A بزرگتر یا مساوی α است. بنابراین عملگر فازی مورد استفاده در الگوریتم، می‌تواند از قواعد فازی استخراج شود. الگوریتم فازی درختی پیشنهادی از چندین فرضیه‌ی فازی برای یافتن بخش‌های فازی از داده‌های کاملی تشکیل شده است. الگوریتم برای ایجاد ساختار درختی برای نواحی فازی طراحی شده است.

ساختار فازی درختی می‌تواند برای استخراج متن بر مبنای قاعده کمک نماید. با این حال، هنگامی که درجه‌ی عضویت را بسط می‌دهیم، این فرآیند نسبت به فرآیند اصلی پیچیده‌تر می‌شود. الگوریتم به صورت زیر توصیف می‌شود:

ورودی: یک پایگاه داده‌ی کمی شامل N پیکسل، بخش توابع عضویت و ترشولد کمینه‌ی از پیش تعیین شده می‌باشد.

خروجی: متن استخراج شده

گام ۱: تبدیل مقادیر کمی v_{ij} از هر بخش I_j در زامین پیکسا به یک بخش فازی s_{ij} به صورت $s_{ij} = \left(\frac{f_{ij1}}{R_{j1}} + \dots + \frac{f_{ijn}}{R_{jn}} \right)$ با استفاده از توابع عضویت داده شده، جاییکه h تعداد نواحی فازی می‌باشد.

گام ۲: محاسبه‌ی عددی هر ناحیه‌ی فازی R_{j1} :

$$count_{j1} = \sum_{i=1}^n f_{ij1}$$

گام ۳: یافتن $max - count_{j1} = \max_{f=1}^h (count_{ij})$ برای $j = 1 - m$ ، جاییکه h تعداد نواحی فازی برای I_j و m تعداد پیکسل‌ها می‌باشد.

گام ۴: بررسی اینکه آیا تعداد $max - count_{j1}$ بزرگتر یا برابر مقدار کمینه‌ی $n \times s$ است یا خیر. اگر این شرایط برآورده شود، ناحیه‌ی فازی در بخش نواحی متوالی (L_1) قرار می‌گیرد:

$$L_1 = \{max - R_j | max - count_{j1} \geq n \times s, 1 \leq j \leq m\}$$

گام ۵: ساخت جدول عنوان با حفظ L_1

گام ۶: حذف نواحی فازی که در L_1 قرار دارند.

گام ۷: کامل کردن نواحی باقیمانده در هر تصویر به وسیله‌ی مقادیر عضویت‌شان

گام ۸: تعیین مقدار اولیه‌ی عضویت از درخت فازی

گام ۹: وارد کردن پیکسل‌های تصویر تبدیل شده به درخت فازی

در نواحی فازی در جدول عنوان به صورت یک به یک و از پائین به بالا پردازش می‌شوند. شماره‌ی پیکسل‌ها شامل ناحیه‌ی فازی می‌باشد. به دلیل اینکه هر شاخه در درخت فازی مطابق با مقادیر عضویت نواحی فازی در یک تصویر ساخته می‌شود، شماره‌ی پیکسل k ام فازی به وسیله‌ی عملگر اشتراک فازی می‌تواند به دست آید. اگر مقدار کلی عضویت یک پیکسل K فازی برای همه‌ی تصاویر بزرگتر یا برابر مقدار کمینه‌ی از پیش تعیین شده باشد، این پیکسل از فرآیند الگوریتم خارج می‌شود.

فرض می‌کنیم که یک تصویر برای ساخت درخت فازی و برای استخراج بخش‌های فازی $\{f_1:v_1, f_2:v_2, \dots, f_n:v_n\}$ استفاده شده است. جاییکه f_n داده و v_n مقدار عضویت است. فرض می‌کنیم که دو تصویر $\{f_1:v_1, f_2:v_2\}$ و $\{f_1:v_3, f_2:v_4\}$ وجود دارد. برای هر ساختار درختی، مقادیر عضویت در تصویر سپس مطابق با ساختار درختی تطبیق می‌یابد. بنابراین اگر $v_1 > v_3$ و $v_4 > v_3$ باشد، چهار گره‌ی درختی برای ساخت درخت تولید می‌شود. برای ساخت درخت، از رویدادهای متناوب f_1 و f_2 ، تنها دو پیکسل تولید می‌شود، چون که مقدار عضویت f_1 برابر $(v_1 + v_3)$ است و مقدار عضویت f_2 برابر $(v_2 + v_4)$ است. فرض می‌کنیم که تصویر شامل m پیکسل است.

بنابراین مطابق روش فازی فوق، فرآیند الگوریتم تا زمانی ادامه می‌یابد که نرخ پیک سیگنال به نویز در بالاتری مقدار ممکن قرار داشته باشد. در این حالت الگوریتم خاتمه می‌یابد و متن از تصویر استخراج می‌شود.

اما بعد از معرفی انواع الگوریتم‌های بهینه‌سازی و به‌ویژه الگوریتم بهینه‌سازی مورد استفاده که ترکیبی از الگوریتم عصبی و فازی است، در قسمت بعدی توابع هدف به عنوان معیارهای بهینه‌سازی مورد بررسی قرار خواهد گرفت.

پارامترهای هدف

بعد از معرفی روش پیشنهادی برای جاسازی متن در تصویر و بازیابی آن و همچنین ارائه‌ی الگوریتم‌های بهینه‌سازی و رویکرد اجرای آن‌ها، در این قسمت پارامترهای سنجش و هدف برای تحقیق مشخص می‌شود. این پارامترها معیارهای سنجش و ارزیابی روش پیشنهادی را نسبت به روش‌های دیگر مشخص می‌کنند که در ادامه به بررسی آن‌ها می‌پردازیم.

پارامتر هدف نسبت پیک سیگنال به نویز

اصلی‌ترین و شاید مهم‌ترین معیار عملکرد در حوزه‌ی پردازش سیگنال و پردازش تصویر، نسبت پیک سیگنال به نویز یا PSNR می‌باشد. این پارامتر مشخص می‌کند که در یک تصویر تا چه سطحی نویز وجود دارد و بنابراین هرچه میزان این پارامتر بالاتر باشد، الگوریتم پردازشی مناسب‌تر است. معمولاً اگر رویکرد پردازشی از فیلترینگ مناسب و از قالب‌بندی مناسبی برخوردار باشد، مقدار PSNR آن بالاتر خواهد بود.

برای سنجش و مقایسه‌ی این پارامتر برای روش‌های مختلف، یک نویز تصادفی به تصاویر وارد می‌شود و به‌صورت نسبی PSNR در پایان پردازش محاسبه می‌شود. این پارامتر معمولاً به‌صورت لگاریتمی تعریف می‌شود:

$$PSNR = \log\left(\frac{P_{signal}}{P_{Noise}}\right) \quad (10)$$

پارامتر هدف نرخ خطای بیت

در حوزه‌ی پردازش دیجیتال یا پردازش در حوزه‌ی فرکانس، با توجه به آن که تحقیق ما یک الگوریتم آموزشی است که با یک تصویر اولیه مقایسه می‌شود، بنابراین لازم است که تمامی پیکسل‌های متن استخراجی با پیکسل‌های متن اصلی یکسان باشد. میزان خطای بیت‌های متن استخراجی نسبت به متن اصلی را نرخ خطای بیت می‌نامند و با BER نمایش می‌دهند. طبیعی است که هرچه میزان این پارامتر کمتر باشد، الگوریتم از عملکرد مناسب‌ترین برخوردار است. توجه شود که این پارامتر به‌صورت صفر و یک محاسبه می‌شود. یعنی یک بیت یا نسبت به متن اصلی یکسان است یا متفاوت است. نسبت تعداد بیت‌های خطا نسبت به بیت‌های اصلی را برای محاسبه‌ی BER به‌کار می‌بریم. توجه شود که بیت در این‌جا همان پیکسل‌ها می‌باشد.

پارامتر هدف زمان پردازش

یکی از مهم‌ترین پارامترهایی که در حوزه‌ی پردازش تصویر از اهمیت بالایی برخوردار است، پایداری در برابر حملات است. برای ارزیابی این پارامتر از معیار همبستگی نرمالیزه استفاده می‌شود. همبستگی نرمالیزه پارامتری است که تعیین می‌کند آیا پیکسل‌های تصویر همچنان دارای همبستگی هستند یا خیر.

حساسیت شفافیت

پارامترهای حساسیت می‌تواند معیار مناسبی برای سنجش اثر نویز بر روی تصاویر باشد. هنگامی که می‌خواهیم یک متن را در یک تصویر جاسازی نمائیم لازم است که مشخص کنیم که اثر نویز بر این فرآیند به چه میزان است. حساسیت شفافیت به-وسیله‌ی رابطه‌ی زیر سنجیده می‌شود:

$$L_k = X_{DC,k} / \text{mean}(X_{DC}) \quad (11)$$

ضریب DTC در k امین بلوک است و به‌معنای تعداد همه‌ی ضرایب در متن است.

حساسیت فرکانسی

اگر ما تصویر را به بلوک‌های 8×8 تقسیم کنیم و DCT را برای هر بلوک به‌صورت جداگانه اجرا کنیم، بنابراین یک ماتریس 8×8 از ضرایب DCT تشکیل می‌شود. این ماتریس به سه قسمت تقسیم می‌شود: فرکانس بالا (H)، فرکانس پائین (L) و

فرکانس میانی (M). توجه داشته باشیم که ضریب انرژی تصویر در ضرایب فرکانس پائین قرار داده می‌شود. در صورتی که ضرایب فرکانس پائین تغییر داده شوند، تصویر اصلی دچار تغییر می‌شود. البته در صورتی که ضرایب فرکانس بالا تغییر داده شوند، آن‌گاه هنگامی که فرآیند فشرده‌سازی تصویر را صورت دهیم، ممکن است متن از تصویر اصلی محو شود. دلیل این امر هم شاید آن باشد که فشرده‌سازی سبب حذف ضرایب DCT در حالت فرکانس بالا می‌شود.

با توجه به دلایل مطرح شده در حالت فرکانس بالا، ما از ضرایب DCT در فرکانس میانی برای جاسازی تصویر استفاده می‌کنیم.

حساسیت لبه‌ای

لبه در تصاویر با توجه به عملکرد ترشولد آشکار می‌شود. در Toolbox متلب از تابع $graythresh()$ برای آشکارسازی لبه‌ای استفاده شده است. از رابطه‌ی زیر حساسیت لبه‌ای محاسبه می‌شود:

$$E_k = graythresh(f) \quad (12)$$

که f بلوک مورد نظر است.

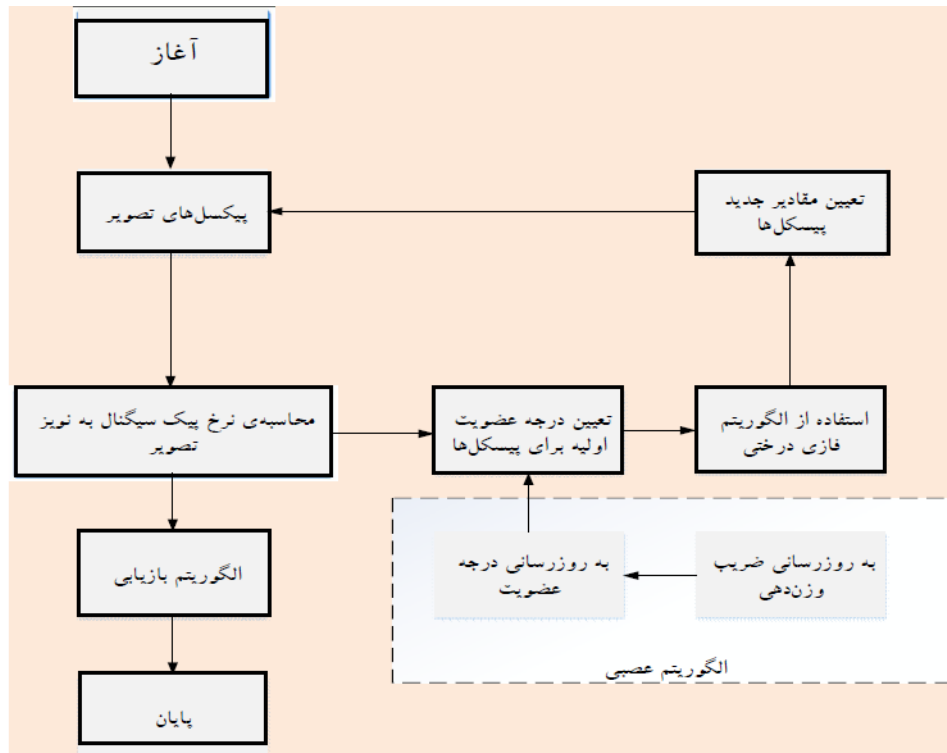
متدولوژی مورد نظر

روش پیشنهادی ما بر اساس ترکیب روش‌های عصبی و فازی قرار دارد. توجه کنید که الگوریتم عصبی یکی از الگوریتم‌های بهینه‌سازی سراسری است که محدوده‌ی جواب را با سرعت مشخص می‌کند. از طرف دیگر الگوریتم فازی یک الگوریتم بهینه‌سازی محلی است که جواب دقیق را در یک مجموعه‌ی داده‌ی محدود به دست می‌آید. نقص الگوریتم عصبی دقیق نبودن جواب نهایی است و برای الگوریتم فازی گستره‌ی جست‌وجوی محدود به عنوان نقص وجود دارد. بنابراین ترکیب این دو الگوریتم با اولویت اجرای الگوریتم عصبی در ابتدا می‌تواند در مدت زمان کوتاهی جواب بهینه را به صورت دقیق حاصل نماید.

همان‌طوری که ذکر نمودیم، در روش فازی هدف ما تغییر مقادیر درجه‌ی عضویت برای بهینه‌سازی پارامتر پیک سیگنال به نویز است. اما باید ببینیم که این پارامتر به چه صورت تغییر می‌کند. ما برای تغییر مقدار درجه‌ی عضویت هر پیکسل روش عصبی RBF را مورد استفاده قرار می‌دهیم. روش عصبی در هر مرحله مقدار درجه‌ی عضویت را تغییر می‌دهد و مقدار جدید را در روش فازی قرار می‌دهد. روش فازی نیز مقدار جدید درجه‌ی عضویت را در پیکسل‌ها وارد کرده و مقدار PSNR را محاسبه می‌کند.

بر این اساس بلوک دیاگرام شکل ۵ را ارائه می‌کنیم که الگوریتم پیشنهادی ما را برای فرآیند استخراج متن، نمایش می‌دهد.

اما باید در این قسمت به این نکته اشاره داشته باشیم که در روش پیشنهادی ما، روش عصبی را برای بهینه‌سازی تعیین مقدار درجه‌ی عضویت مورد استفاده قرار می‌دهیم. همچنین روش فازی مورد استفاده‌ی ما در این مقاله استفاده از روش درخت فازی است. این الگوریتم نسبت به الگوریتم‌های مشابه قبلی دارای تفاوت بنیادی از نظر تعیین پارامترها است. به این صورت که معیار ارزیابی روش فازی درجه‌ی عضویت پیکسل‌ها است و برای تغییر درجه‌ی عضویت از روش عصبی استفاده شده است.



شکل ۵: الگوریتم پیشنهادی نهایی برای بازیابی متن روش افزایش نرخ پیک سیگنال تصویر نسبت به نویز

برای توضیح روش مورد نظر باید ذکر کنیم که یک بردار مشخصه‌ی X از نشان داده شده است که $Y \in \gamma$ بردار متن با طول m است و y_i برابر نامین المان از Y است. مقدار این المان بین -1 و $+1$ قرار دارد که $+1$ نشان می‌دهد که نشام برای تصویر X به درستی استخراج شده است یا خیر. همچنین -1 بیانگر نامناسب بودن متن برای X است. برای هر متن دسته‌بندی r ساخته شده است. این دسته‌بندی‌ها می‌تواند مدل‌های مختلفی داشته باشد.

احتمال آن که متن Y به درستی از تصویر اصلی استخراج شده است، از طریق رابطه‌ی زیر به دست می‌آید:

$$P(Y|X) = \frac{1}{Z(X)} \exp(-E(Y|X)) \quad (13)$$

جایی که $E(Y|X)$ تابع انرژی می‌باشد و از طریق رابطه‌ی زیر به دست می‌آید:

(14)

$$E(Y|X) = \sum_i^r \sum_j^{m_i} w_{ij} f_{ij}(X, y_j) + \sum_i^{m_i} \sum_j^3 w_{ijk} \delta(y_i, y_j, k) f'_{ij}(y_i, y_j)$$

جائیکه $f_{ij}(X, y_j)$ اندازه‌ی احتمالی است که نتیجه‌ی نامین دسته‌بندی برای نامین متن مثبت باشد.

اندازه‌ی همبستگی بین نامین و نامین کلمه است و $\delta(y_i, y_j, k)$ به صورت زیر به دست می‌آید:

(15)

$$\delta(y_i, y_j, k) = \begin{cases} 1 ; y_i = -1, y_j = -1, k = 0 \\ 1 ; y_i = -1, y_j = +1, k = 1 \\ 1 ; y_i = +1, y_j = -1, k = 2 \\ 1 ; y_i = +1, y_j = +1, k = 3 \\ 0 ; otherwise \end{cases}$$

$Z(X)$ تابع بخش‌بندی است و به صورت زیر محاسبه می‌شود:

$$Z(X) = \sum_{y \in Y} \exp(-E(X|Y)) \quad (16)$$

اما در روش پیشنهادی از یک فرآیند آموزشی بهره گرفته شده است که به صورت زیر می باشد:

فرآیند آموزشی دارای دو گام اصلی است: تعیین $f_{ij}(X, y_j)$ و $f'_{ij}(y_i, y_j)$ ، تعیین پارامترهای W . در ابتدا دسته بندی r با کرنل های مختلف برای هر دسته آموزش داده می شود. سپس $f_{ij}(X, y_j)$ بر مبنای فاصله از x تعیین علامت می شود. همبستگی بین λ امین و λ امین کلمه است و از طریق کوواریانس بین λ امین و λ امین پیکسل از نشان به دست می آید. بنابراین از تطبیق نرمالیزه برای تعیین جفت های پیکسل های متن استفاده می شود. برای پیکسل های p و q ، داده های نرمالیزه به صورت زیر تعیین می شود:

$$NormMI(p, q) = \frac{MI(p, q)}{\min\{H(p), H(q)\}} \quad (17)$$

جائیکه $MI(p, q)$ داده های متقابل از پیکسل های p و q است و $H(p)$ آنترپی حاشیه ای از متن p است که به صورت زیر تعریف می شود:

$$MI(p, q) = \sum P(y_p, y_q) \log \frac{P(y_p, y_q)}{P(y_p)P(y_q)}$$

$$H(p) = - \sum P(y_p) \log P(y_p)$$

آزمایشات و نتایج تجربی

در این قسمت به بررسی نتایج به دست آمده از پیاده سازی روش مورد نظر می پردازیم. برای این منظور یک تصویر و یک متن را برای این تصویر مطابق با شکل زیر در نظر می گیریم.



شکل ۶: یک تصویر و یک متن

اولین مرحله همان طور که بیان شد تبدیل فوریه ی کسینوسی است که باید بر روی تصویر و متن مورد استفاده قرار گیرد. در تصویر فوق هم تصویر اصلی و هم متن مورد نظر قرار داده شده است. بر این اساس در جدول ۱ مقادیر پیکسل های تصویر بعد از تبدیل فوریه آورده شده است.

جدول ۱: مقادیر پیکسل های تصویر بعد از تبدیل فوریه

11	28	85	111	114	155	156	157	144	12
109	156	112	110	97	96	92	85	19	15
140	156	160	175	188	159	128	99	28	18
85	100	120	132	142	110	100	98	39	22
111	149	175	188	156	139	125	101	44	21
106	121	148	164	144	131	119	98	45	20
92	102	121	141	126	116	101	91	42	18

90	93	100	102	142	181	129	88	61	25
110	115	132	192	191	199	152	108	91	36
60	65	72	85	92	78	68	59	49	45
21	30	39	55	62	70	79	66	56	60
27	65	103	128	130	136	125	115	92	85

جدول فوق یک نمونه از مقادیر پیکسل‌ها و متن است. اما در این مرحله مطابق با شکل ۷۸ جدول پیکسل‌ها به ۸ قسمت مساوی تقسیم می‌شود.

11	28	85	111	114	155	156	157	144	12
109	156	112	110	97	96	92	85	19	15
140	156	160	175	188	159	128	99	28	18
85	100	120	132	142	110	100	98	39	22
111	149	175	188	156	139	125	101	44	21
106	121	148	164	144	131	119	98	45	20
92	102	121	141	126	116	101	91	42	18
90	93	100	102	142	181	129	88	61	25
110	115	132	192	191	199	152	108	91	36
60	65	72	85	92	78	68	59	49	45
21	30	39	55	62	70	79	66	56	60
27	65	103	128	130	136	125	115	92	85

شکل ۷: تقسیم بندی پیکسل‌ها به ۸ قسمت مساوی

اما در مرحله‌ی بعدی مقادیر پیکسل‌های متن در پیکسل‌های تصویر وارد می‌شود که در جدول ۲ نمونه‌ای از این مورد نشان داده شده است.

جدول ۲: مقادیر پیکسل‌های متن در پیکسل‌های تصویر

11	28	85	111	114	155	156	157	144	12
109	156	112	110	97	96	92	85	19	15
140	156	160	175	188	159	128	99	28	18
85	100	120	132	142	110	100	98	39	22
111	149	175	188	156	139	125	101	44	21
106	121	148	164	144	131	119	98	45	20
92	102	121	36	42	39	28	91	42	18
90	93	100	102	52	55	129	88	61	25
110	115	132	192	191	199	152	108	91	36
60	65	72	85	92	78	68	59	49	45
21	30	39	55	62	70	79	66	56	60
27	65	103	128	130	136	125	115	92	85

اما در مرحله‌ی بعدی مقادیر پارامترهای حساسیت و پارامترهای تصویر محاسبه می‌شود که در بخش سوم نتایج این قسمت آورده شده است. اما فرض بر آن است که مقادیر مورد نظر مطلوب نیست. بنابراین از روش عصبی-فازی برای تغییر ضرایب فوریه استفاده می‌شود.

مطابق با جدول ۳ لازم است مقادیر اولیه برای اجرای الگوریتم نیز در نظر گرفته شود. در جدول زیر پارامترهای مرتبط با شبیه‌سازی مشخص گردیده است.

جدول ۳: پارامترهای شبیه‌سازی الگوریتم پیشنهادی

ردیف	پارامتر	مقدار
۱	تعداد خوشه	تبدیل تصویر به دو خوشه‌ی ۸*۸
۲	خطای الگوریتم عصبی	خطای مطلوب ۰/۰۰۱ در نظر گرفته شده است
۳	وزن اولیه برای شبکه‌ی عصبی	وزن اولیه صفر در نظر گرفته شده است
۴	بیت اولیه برای متن	صفر

اما مرحله‌ی بعدی اختصاص به تبدیل فوریه کسیونسی دارد. در جدول ۴ مقادیر پیکسل‌ها بعد از اجرای این تبدیل آورده شده است.

جدول ۴: مقادیر پیکسل‌ها بعد از اجرای تبدیل فوریه کسیونسی

Pixel	Hex	Dec	Pixel	Hex	Dec	Pixel	Hex	Dec	Pixel	Hex	Dec
1,1	129	96	1,2	194	64	1,3	145	32	1,4	111	26
1,6	61	97	1,7	41	65	1,8	121	93	2,1	96	48
2,2	162	98	2,3	142	66	2,4	165	92	2,5	151	88
2,6	155	88	2,7	128	78	2,8	123	96	3,1	142	58

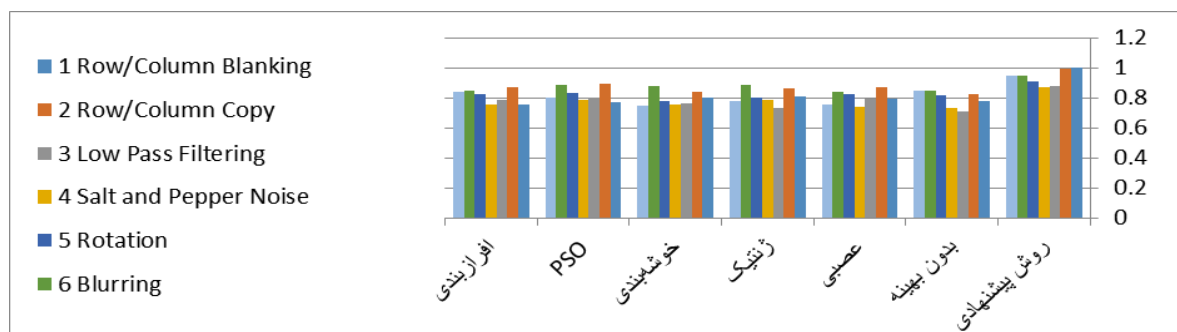
اما بعد از اینکه تبدیل فوریه‌ی کسیونسی مورد استفاده قرار گرفت و مقادیر پیکسل‌های تصویر متن‌گذاری شده به‌دست آمد، روش استخراج متن از تصویر مورد استفاده قرار می‌گیرد. در جدول ۵ مقایسه‌ی زمانی مربوط به پردازش قبل از بکارگیری روش فازی و بعد از آن، با توجه به تکرارهای متوالی متن داده شده است. همان‌طور که ملاحظه می‌گردد، با بکارگیری روش فازی زمان پردازش به‌میزان قابل‌توجهی کاهش یافته است.

جدول ۵: مقایسه‌ی زمان پردازش با و بدون بکارگیری روش فازی

نمونه	روش پیشنهادی	بدون بهینه	عصبی	ژنتیک	خوشه‌بندی	PSO	افرازبندی
۱	6.6nsec	11nsec	9.8nsec	9.3nsec	10.2nsec	10.0nsec	9.7nsec
۲	7.1nsec	11.7nsec	10.2nsec	9.8nsec	9.9nsec	10.3nsec	10.0nsec
۳	7.0nsec	10.6nsec	9.5nsec	9.8nsec	9.8nsec	9.6nsec	10.1nsec
۴	6.3nsec	10.2nsec	9.6nsec	9.5nsec	10.1nsec	9.6nsec	9.8nsec

نتایج شبیه‌سازی می‌دهد که الگوریتم ارائه شده یک الگوریتم مقاوم در برابر حملات است. بازیابی متن از تصویر در شرایطی در نظر گرفته شده است که حملات مختلف به تصویر اصلی صورت پذیرفته است. برای اینکه اثر نویز پس‌زمینه حذف شود، تصویر متن‌گذاری شده از یک فیلتر پائین‌گذر عبور داده می‌شود. همچنین از یک فیلتر میان‌گذر برای از بین بردن اثر حملات استفاده شده است.

در شکل ۸ مقادیر همبستگی نرمالیزه شده آورده شده است.

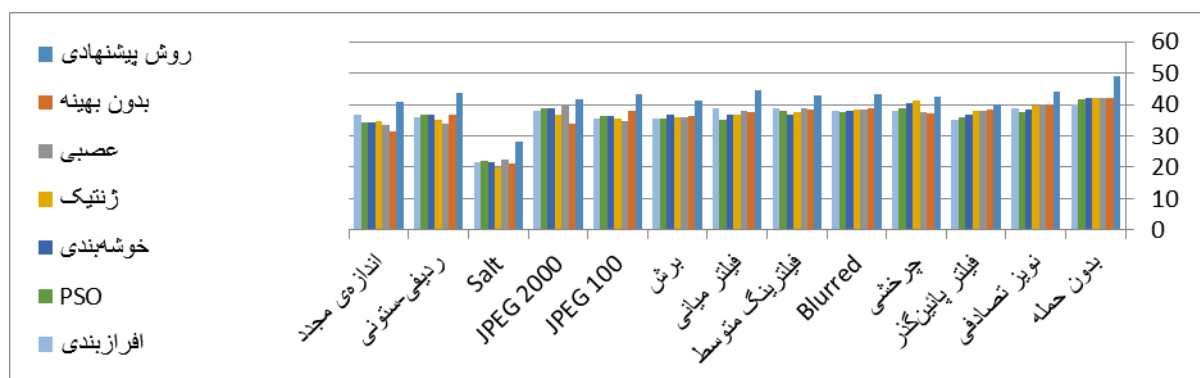


شکل ۸: مقایسه‌ی همبستگی نرمالیزه در برابر حملات

جدول ۶: سنجش پایداری در برابر حملات

ردیف	نوع حملات	مقدار همبستگی
۱	Bit Plane Removal	1
۲	Sharpening	0.9984
۳	RC Blank	0.9979
۴	Intensity Transformation	0.9974
۵	RC Copying	0.942
۶	Image Contrast Attack	0.9921
۷	Blurring	0.9504
۸	Gaussian Noise	0.9473

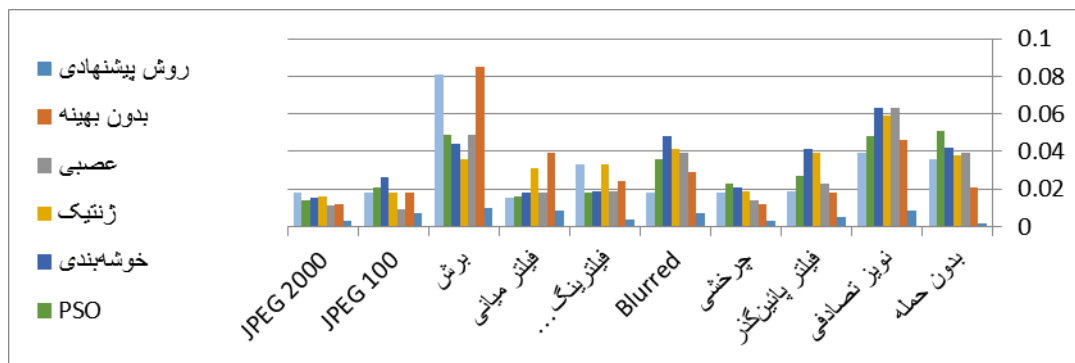
در جدول ۶ همبستگی نرمالیزه شده به عنوان مقیاسی برای سنجش پایداری آورده شده است. دو پارامتر اصلی برای سنجش الگوریتم جاسازی و جایابی استفاده PSNR و BER است. ما این ادعا را داریم که مقدار این دو پارامتر در روش پیشنهادی ما نسبت به سایر روش‌های قبلی از شرایط مطلوب‌تری برخوردار است. در شکل ۹ مقدار PSNR به دست آمده از روش ما نشان داده شده است. البته باید به این نکته اشاره داشته باشیم که نتایج به دست آمده با اجرای ۱۰۰ تکرار روش فازی حاصل شده است.



شکل ۹: مقدار PSNR برای روش پیشنهادی و سایر روش‌ها

با مقایسه‌ی نتایج به دست آمده از شکل ۹ با نتایج ارائه شده در بخش‌های قبل مشاهده می‌کنیم که با روش فازی پارامتر از عملکرد بهتری برخوردار است.

یکی دیگر از پارامترهایی که برای مقایسه می‌توانیم از آن بهره بگیریم پارامتر BER است. در شکل ۹ مقادیر به‌دست آمده از این پارامتر ارائه شده است.



شکل ۱۱: مقدار BER برای روش پیشنهادی و سایر روش‌ها

با مشاهده‌ی شکل ۱۱ به این نتیجه می‌رسیم که روش پیشنهادی در ارتباط با نرخ خطای بیت از عملکرد مطلوبی برخوردار است و می‌توانیم ادعا کنیم که استفاده از روش فازی برای جاسازی متن در تصاویر می‌تواند در بهبود کیفیت تصویر تاثیر مستقیمی داشته باشد. با ارائه‌ی نتایج ارائه شده در این تحقیق، مشخص می‌گردد که الگوریتم پیشنهادی نسبت به روش‌های مشابه دیگر عملکرد بهتری را ارائه داده است. این بهبود عملکرد در حالت‌های مختلف و در برابر حملات مختلف عملکرد مناسب‌تری را از خود نشان می‌دهد و بنابراین می‌توانیم این ادعا را داشته باشیم که این الگوریتم از شرایط مناسب‌تری برخوردار است.

نتیجه‌گیری و کارهای آتی

تکنیک‌های استخراج دیجیتال این امکان را به کاربران می‌دهد که بتوانند علائمی همچون لوگو، علائم تجاری و یا اطلاعاتی همچون حق چاپ را بدون تغییر در متن اصلی تصویر، در تصویر جاسازی نمایند. وقتی تصویر بازبینی می‌شود، می‌توان به راحتی علائم را استخراج نمود. برای حفظ مالکیت اصلی تصاویر، متن‌گذاری سخت در حوزه‌ی فرکانس قرار می‌گیرد. دلیل این انتخاب آن است که علائمی که در این حوزه قرار می‌گیرند نسبت به حوزه‌ی فضایی پایدارتر هستند.

برای اینکه بتوانیم استخراج را با دقت بالایی انجام دهیم، همچنین بتوانیم از مقیاس‌بندی با گام‌های کمتر برای استخراج استفاده کنیم، بهتر و مناسب‌تر است که از روش تبدیل فوریه‌ی کسینوسی استفاده کنیم. این روش هم در اجرا با سهولت بیشتری همراه است و هم اینکه استخراج را با پایداری بیشتری صورت می‌دهد. از طرف دیگر در تبدیل فوریه‌ی کسینوسی نسبت به تبدیل فوریه‌ی موجک روابط ریاضی آسان‌تر بوده و از حجم محاسبات پائین‌تری برخوردار است. ما از این روش برای تبدیل فوریه استفاده می‌کنیم.

پنهان‌نگاری از طریق PSNR تصویر اندازه‌گیری می‌شود. PSNR بالاتر این حجم را نشان می‌دهد که بازدهی پنهان‌نگاری در حد مطلوبی قرار دارد. همچنین ناپایداری از طریق هبستگی تصویر اصلی و تصویر بازیابی شده اندازه‌گیری می‌شود. تکنیک‌های مختلفی در این ارتباط ارائه شده‌اند. همانطور که اشاره شد، دو پارامتر PSNR و ناپایداری از اهمیت ویژه‌ای برخوردار هستند. برای بهبود این دو پارامتر از روش‌های فازی استفاده کردیم. روش فازی را برای طراحی فیلترینگ مناسب و به تبع آن افزایش PSNR بکار برده‌ایم. همچنین در هنگام جاسازی علائم در تصاویر، روش فازی دارای قابلیت‌های بالایی است و می‌تواند پنهان‌نگاری را بهبود بخشد.

اما روش اصلی ما در این مقاله بر اساس گام‌های زیر صورت گرفته است:

- در ابتدا به‌وسیله‌ی تبدیل فوریه‌ی کسینوسی تبدیل تصویر دیجیتال را به انجام رساندیم. با این تبدیل تصویر به از حوزه‌ی فرکانس به حوزه‌ی زمان تبدیل می‌شود.

• مرحله‌ی بعدی تعریف پارامترهای مرتبط به جاسازی استخراج است. ما سه پارامتر حساسیت شفافیت، حساسیت فرکانسی و حساسیت لبه‌ای را به عنوان معیارهای سنجش عمل جاسازی و بازیابی استخراج در نظر گرفتیم. این سه پارامتر بعداً در پیاده‌سازی روش فازی برای بهینه‌سازی به‌عنوان توابع هدف در نظر گرفته می‌شوند.

• بعد از اینکه تصویر به حوزه‌ی زمان تبدیل شد، ابعاد تصویر اصلی و ابعاد استخراج تعیین می‌شوند. این عمل برای آن است که پیکسل‌های تصویر اصلی و استخراج را به‌صورت ماتریسی تبدیل کنیم و بتوانیم متن را در تصویر اصلی بر اساس این پیکسل‌ها جایگزین کنیم.

• جایگزینی متن در مکان مشخص شده در تصویر اصلی. در این مرحله بر اساس ماتریسی که برای تصویر اصلی و متن تعیین کردیم، متن را جایگزین نمودیم. درایه‌های ماتریس تعیین‌کننده محل قرارگیری متن خواهند بود.

• اما برای اینکه بتوانیم روش فازی را برای بهینه‌سازی سه پارامتر بیان شده بکار ببریم، لازم است در ابتدا برای هر پیکسل متن که در تصویر اصلی جایگزین شده است، پارامتر وزن‌دهی را تعیین کنیم. بعد از این مرحله، با تغییر میزان پارامتر وزن‌دهی هر یک از پارامترها را سنجش نمودیم.

پارامترهای حساسیت می‌تواند معیار مناسبی برای سنجش اثر نویز بر روی تصاویر باشد. هنگامی که می‌خواهیم یک متن را در یک تصویر جاسازی نمائیم لازم است که مشخص کنیم که اثر نویز بر این فرآیند به چه میزان است.

روش ما بر این اساس قرار دارد که در ابتدا تصویر اصلی را به صفحات قرمز، سبز و آبی تقسیم می‌کنیم و تبدیل کسینوسی گسسته را برای هر بلوک 8×8 بکار می‌بریم. سپس پارامترهای حساسیت شفافیت، حساسیت فرکانسی و ... را برای آن به‌دست می‌آوریم. آموزش شبکه‌ی عصبی برای مقادیر حساسیت فرکانسی، ورودی که از خروجی سیستم فازی دریافت می‌دارد صورت می‌گیرد و پارامترهای حساسیت مورد استفاده در فرآیند جاسازی با تبدیل حوزه‌ی فرکانس ضرایب DCT به‌دست می‌آید و بنابراین ضرایب هر بلوک در انتها بدون تغییر باقی می‌ماند. با این توضیحات الگوریتم جاسازی را بیان می‌کنیم.

از داده‌هایی که به‌عنوان داده‌های ورودی به شبکه‌ی عصبی داده شده است، تنها در یک مورد شبکه به درستی آموزش ندیده است. بنابراین می‌توان ادعا کرد که خطای الگوریتم در حد بسیار پائینی است.

نتایج شبیه‌سازی نشان می‌دهد که الگوریتم ارائه شده یک الگوریتم مقاوم در برابر حملات است. بازیابی متن از تصویر در شرایطی در نظر گرفته شده است که حملات مختلف به تصویر اصلی صورت پذیرفته است.

اما برای مقایسه‌ی آماری روش پیشنهادی نسبت به سایر روش‌های مرتبط با مسئله‌ی مورد نظر باید اشاره داشته باشیم که در ارتباط با سرعت پردازش، افزایش امنیت، نرخ سیگنال به نویز و نرخ خطای بیت مقایسه‌ای بین روش پیشنهادی و سایر روش‌های بهینه‌سازی دیگر است که در جدول زیر به‌صورت آماری مقایسه‌ای بر اساس درصد بهبود پارامترهای مطرح شده نسبت به روش‌های بهینه‌سازی دیگر آورده شده است.

جدول ۷: نتایج آماری مقایسه‌ای بر اساس درصد بهبود پارامترهای مطرح شده نسبت به روش‌های بهینه‌سازی دیگر

پارامتر/روش	عصبی	ژنتیک	خوشه‌بندی	PSO	افرازبندی
سرعت پردازش	٪۱۲	٪۱۱	٪۱۷	٪۲۲	٪۱۶
امنیت	٪۲۷	٪۲۴	٪۳۲	٪۳۵	٪۱۹
PSNR	٪۱۹	٪۱۴	٪۱۸	٪۲۰	٪۱۷
BER	٪۹	٪۱۰	٪۸	٪۱۲	٪۱۱

منابع و مراجع

- [1] Sunitha, S., RamaSathish, A., "Extended Image Features for User Intention Refined Image Search", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 4, Issue 2, pp. 278-282, February 2014.
- [2] Huu, Q., N., Thu, H., N., T., Quoc, T., N., "An Efficient Content Based Image Retrieval Method for Retrieving Images", International Journal of Innovative Computing, Information and Control, Volume 8, No. 4, pp. 2823-2836, April 2012.
- [3] Khokhar, A., Talwar, R., "Content-based Image Retrieval: Feature Extraction Techniques and Applications", International Journal of Computer Applications, (IJCS), pp. 9-14, 2012.
- [4] Nascimento, M., A., Sridhar, V., Li, X., "Effective and Efficient Region-based Image Retrieval", Journal of Visual Languages and Computing, 14, pp. 151-179, 2003.
- [5] Neetu Sharma. S., Paresh Rawat, S., Jaikaran Singh, S., "Efficient CBIR using Color Histogram Processing", Signal and Image Processing : An International Journal (SIPIJ) Vol. 2, No. 1, pp. 94-112, March 2011.
- [6] Chang, R., Lin, S., Y., Ho, J., M., Fann, C., W., Wang, Y., C., "A Novel Content Based Image Retrieval System Using K-Means / KNN with Feature Extraction", ComSIS Vol. 9, No. 4, Special Issue, pp. 1645-1661, December 2012.
- [7] Yasmin, M., Sharif, M., Sajjad, M., "Use of Low Level Features for Content Based Image Retrieval: Survey", Research Journal of Recent Sciences, Vol. 2(11), pp. 65-75, November 2013.
- [8] Santhosh, D., Trueman, T., E., "Artificial Neural Network Technique for CBIR Based on Query Image Feature Extraction", International Journal of Innovative Research in Computer, Vol. 2, Special Issue 3, pp. 223-230, July 2014.
- [9] Bai TZ, Hou XB, an improved image matching algorithm based on sift. Trans Beijing Inst Technol 33(6):622-628 Bay H, Ess A, Tuytelaars T (2006) Surf: speeded up robust features, 2013.
- [10] Cen YG, Chen XF, Cen LH, Compressed sensing based on the single layer wavelet transform for image processing. J Commun 31(8A):51-55, 2010.
- [11] Chen TH, Horng G, Lee WB, a publicly verifiable copyrightproving scheme resistant to malicious attacks. IEEE Trans Ind Electron 52(1):327-334, 2005.
- [12] Jiang N, Wdem: weighted dynamics and evolution models for energy-constrained wireless sensor networks. Phys A: Stat Mech Appl 404:323-331, 2014.
- [13] Holland J. *Genetic algorithms and the optimal allocation of trials*. SIAM J. Comput. 2, 88-105, 1979.
- [14] Kennedy J, Eberhart R. C. *Particle Swarm Optimization*, Proceedings of the 4th IEEE International Conference on Neural Networks, pp. 1942-1948, 1995.