

تشخیص نواحی مزاحم بصری در تصاویر به وسیله روش نمایش تنک دومرحله‌ای و وزن دار نمونه آزمون

فرزین صبوری و فرزین یغمایی

انگیزه‌ها، هدایت توجه بصری^۱ مخاطب از ناحیه‌ای به ناحیه دیگر می‌باشد چرا که عکاس مایل است که مخاطب، همواره بیشینه تمرکز و توجهش را بر سوژه‌ای معطوف کند که مورد نظر اوست. این موضوع برای سال‌های متوالی مورد علاقه محققین بوده است، برای مثال [۲] و [۳] تصویر را جهت هدایت توجه بصری بیننده، ویرایش می‌کنند و تغییر می‌دهند. مقاله [۴] برجستگی نواحی خاصی از تصویر را مطابق نظر کاربر تغییر می‌دهد تا با انجام آن، هدایت توجه بصری مخاطب و بیننده مطابق نظر کاربر صورت گیرد. چالش اصلی محققین این حوزه در تعریف ماهیت مفهوم هدایت توجه بصری مخاطب و نحوه تأثیرگذاری بر آن است. لذا محققین حوزه بینایی ماشین با تعریف و تشخیص پدیده‌ای به نام "برجستگی بصری"^۲ به عنوان مدلی محاسباتی از توجه بصری، با کمک و همفکری متخصصین سایر حوزه‌ها مانند علوم اعصاب، علوم شناختی، روان‌شناسی و ... سعی در حل این مشکل داشته‌اند.

طبق تعریف مورد توافق این متخصصین، به طور کلی برجستگی بصری عموماً به ناحیه‌ای از تصویر اطلاق می‌شود که نسبت به محیط اطرافش متمایز و مستعد در جلب توجه مخاطب باشد [۵]. متأسفانه این استعداد در جلب توجه، همواره مطابق با نظر عکاس نیست و گاه ناحیه‌ای برخلاف خواست عکاس، نظر مخاطب را از سوژه مورد نظر وی دور می‌کند. به عبارتی دیگر، حواس مخاطب به ناحیه‌ای جلب می‌شود که هدف عکاس نبوده است. چالشی که در این مقاله به آن پرداخته خواهد شد، تشخیص نواحی برجسته‌ای است که حاوی اطلاعات مفیدی نیستند و صرفاً موجب اختلال در عطف توجه کاربر به پیام یا سوژه اصلی تصویر و در نتیجه کاهش کیفیت بصری تصویر می‌شوند. مشخصه اصلی این نواحی، میزان بالای برجستگی آنها می‌باشد چرا که ما معتقدیم اگر ناحیه‌ای فاقد اطلاعات مهم باشد، اما به دلیل عدم برجستگی، نظر و توجه مخاطب را از نواحی مهم پرت نکند، نیازمند حذف و ویرایش نیست.

تحقیقات انجام‌شده بر پایه کدگذاری و بازنمایی تنک، شبکه‌های عصبی کانولوشن، روش‌های آماری و ... تمام نواحی برجسته را شناسایی کرده و تمایزی بین نواحی برجسته اصلی و شاخص نسبت به نواحی برجسته مزاحم قایل نیستند. بنابراین چالش اصلی در کاربردهای مذکور، پردازش نواحی برجسته‌ای است که در عین دارابودن استعداد در جلب توجه، از میزان پایین اطلاعات مفید برخوردارند. از این رو با وجود تلاش‌ها و تحقیقات در این حوزه، مشکلات و چالش‌هایی مانند مجموعه دادگان و عدم وجود معیاری محاسباتی و کیفی بودن تشخیص مزاحم بصری هنوز وجود دارند که یا حل نشده و یا به طور کامل به دقت بالایی دست نیافته است. لذا مهم‌ترین ویژگی که در روش پیشنهادی این مقاله برای نواحی مزاحم بصری در نظر گرفته شده است، میزان برجستگی هر

چکیده: مخاطب یک تصویر مایل است که در کوتاه‌ترین زمان، پیام اصلی تصویر را دریافت کند. از این رو سیستم بینایی انسان توجه بصری را ناخودآگاه به سمت نواحی برجسته، با فرض وجود اطلاعات مفید در آنها هدایت می‌کند. عملاً این فرض همواره صادق نبوده و در مواردی، نواحی برجسته صرفاً موجب مزاحمت بصری می‌گردند. از این رو در کاربردهای مختلف نیاز به ساز و کاری جهت تشخیص این نواحی می‌باشد تا با حذف این نواحی، حواس مخاطب از سوژه اصلی تصویر پرت نشود. همچنین نادیده گرفتن این نواحی، کمک شایانی است به روش‌هایی که بر پایه تشخیص نواحی برجسته و مهم عمل می‌کنند. بدین منظور در این مقاله، بر اساس روش‌های منطبق بر چالش عدم توازن دسته‌ها، هر قطعه از تصاویر آموزشی با توجه به ماسک آنها به ۹ دسته افزایش می‌شود که شماره هر دسته متناسب با شدت مزاحمت است. سپس ویژگی‌های مبتنی بر قطعه استخراج و دسته هر قطعه بر اساس روش نمایش تنک دومرحله‌ای و وزن دار نمونه آزمون که بر مبنای سیستم کدگذاری و بازنمایی تنک است، تعیین می‌شود. به منظور ارزیابی دقیق روش پیشنهادی و مقایسه آن با سایر روش‌ها، ۴ معیار ارزیابی با رویکردهای مختلف معرفی و پیشنهاد می‌شود. با ارزیابی و سنجش نتایج نشان داده می‌شود که روش پیشنهادی علی‌رغم زمان‌بر بودن، نسبت به کارهای پیشین دارای دقت بیشتری است.

کلیدواژه: سیستم بینایی انسان، توجه بصری، نواحی برجسته، مزاحمت بصری، سیستم کدگذاری و بازنمایی تنک.

۱- مقدمه

ویرایش تصویر همواره یکی از نیازهای عکاسان - به عنوان کاربر - و از چالش‌های قدیمی - برای متخصصین - در حوزه پردازش تصویر و بینایی ماشین بوده است. ویرایش و در حالت کلی‌تر، دستکاری کردن تصویر یک عمل بسیار جامع است که در نهایت هدف آن بهبود تصویر است. بهبود دادن تصویر با توجه به نیاز کاربر یا متخصص تعریف می‌شود و در این زمینه انگیزه‌ها متفاوت است. به عنوان مثال، انگیزه یک عکاس مناظر طبیعی از ویرایش و بهبود یک تصویر، متفاوت از انگیزه یک پزشک متخصص تصاویر MRI و پزشکی [۶] یا متخصص تصاویر سنجش از دور است. با این حال می‌توان گفت که انگیزه اصلی از ایجاد تغییر در تصویر، دستیابی به تصویری شامل اطلاعات قابل فهم و تحلیل‌پذیرتر می‌باشد. با توجه به این موضوع، مشکلات و چالش‌های پیش رو نیز متناظر با انگیزه و انتظار کاربر، متفاوت است. یکی از این

این مقاله در تاریخ ۱۵ اسفند ماه ۱۳۹۶ دریافت و در تاریخ ۱ مرداد ماه ۱۳۹۷ بازنگری شد.

فرزین صبوری، دانشکده مهندسی برق و کامپیوتر، دانشگاه سمنان، سمنان، ایران، (email: fardin.sabouri@semnan.ac.ir).

فرزین یغمایی (نویسنده مسئول)، دانشکده مهندسی برق و کامپیوتر، دانشگاه سمنان، سمنان، ایران، (email: f_yaghmaee@semnan.ac.ir).

1. Guidance of Visual Attention

2. Visual Saliency

انگیزه تفکیک نواحی برجسته به دو دسته مهم و غیر مهم، یا دو دسته مزاحم و مفید در این حوزه ایجاد می‌شود. از طرف دیگر، اخیراً بسیاری از چالش‌های قدیمی در حوزه بینایی ماشین مانند فشرده‌سازی تصویر و ویدئو، ارزیابی کیفیت تصویر و ویدئو، قالب‌دهی مجدد و تغییر اندازه تصاویر با رویکردهای مبتنی بر تشخیص نواحی برجسته، مجدداً مطرح شده‌اند. انگیزه اصلی استفاده از تشخیص برجستگی در این کاربردها، چشم‌پوشی از نواحی بی‌اهمیت و کم‌اهمیت با تمرکز بر نواحی برجسته به عنوان نواحی با ارزش، جهت رسیدن به الگوریتم‌هایی با سرعت پردازش بالا و رفتاری مشابه با مکانیسم تصمیم‌گیری در مغز انسان است. لذا [۱۴] و [۱۵] الگوریتم‌هایی بر پایه تشخیص برجستگی جهت فشرده‌سازی تصاویر و ویدئو ارائه دادند، بدین ترتیب که نواحی برجسته و مهمی که پتانسیل بالایی در جلب توجه و تمرکز مخاطب دارند، با نرخ پایین‌تری دست‌خوش کاهش حجم و کیفیت شوند.

برخی محققان از تشخیص نواحی برجسته و با ارزش در فشرده‌سازی ویدئو و خلاصه‌سازی مشاهدات بصری بهره برده‌اند [۱۶] و [۱۷]. همچنین در [۱۸] مشابه همین ایده در جهت ارزیابی و تخمین کیفیت تصویر و ویدئو به کار رفته است. در این مقالات برای نواحی برجسته و مستعد در جلب تمرکز و نگاه مخاطب، وزن و ضریب تأثیر بالاتری در نظر گرفته شده و در نتیجه میزان کیفیت نواحی با میزان برجستگی پایین، تأثیر چندانی بر خروجی الگوریتم نخواهد داشت. از سایر کاربردهای مشابه با تشخیص نواحی برجسته با فرض وجود اطلاعات مهم در این نواحی، می‌توان به قالب‌دهی مجدد تصاویر^۴ و تغییر اندازه آنها [۱۹] و [۲۰] اشاره کرد. در این مقالات به هنگام برش یک عکس، نواحی مهم و برجسته حفظ و سایر نواحی با نرخی متناسب با عدم اهمیتشان حذف خواهند شد. با توجه به آنچه گفته شد بدیهی است که هر یک از کاربردهای گفته‌شده، نیازمند تشخیص دقیق نواحی و اشیای برجسته و مهم می‌باشد. این مقالات از تشخیص برجستگی به عنوان یک پیش‌پردازش استفاده می‌کنند که طی آن میزان برجستگی و اهمیت هر پیکسل مشخص می‌شود. در مرحله دوم، نواحی برجسته به دید توده‌هایی با اطلاعات متراکم و مهم، مورد پردازش و تحلیل قرار گرفته و جهت پردازش‌های بعدی از سایر نواحی تفکیک می‌شوند. حال اگر ماهیت تصویر به گونه‌ای باشد که خروجی مرحله اول، شامل نواحی با برجستگی بصری بالا، اما فاقد اطلاعات مهم باشد، آن گاه الگوریتم در مرحله دوم دچار خطا و اشتباه می‌شود. متأسفانه مقالات فوق، تمام نواحی برجسته را شناسایی کرده و تمایزی بین نواحی برجسته‌ای که مخاطب را به هدف عکاس جلب می‌کند و یا نواحی برجسته‌ای که توجه مخاطب را از سوژه مورد نظر او دور می‌کند قابل نیستند. بنابراین چالش اصلی در کاربردهای مذکور، پردازش نواحی برجسته‌ای است که در عین دارابودن استعداد در جلب توجه، از میزان پایینی اطلاعات مفید برخوردارند. پس این مورد هم می‌تواند انگیزه مضاعفی در جهت تشخیص نواحی حواس پرت‌کن و مزاحم بصری باشد.

با این حال متأسفانه به موضوع نواحی مزاحم بصری در بینایی ماشین کمتر پرداخته شده و این مفهوم در پردازش تصویر، تنها در [۲۱] معرفی شده است. آنها به وسیله تحلیل مجموعه تصاویر موجود و ماسک (نقشه) مربوط به نواحی مزاحم بصری، ویژگی‌های مختلفی را برای هر پیکسل از جمله کانال‌های رنگ، تشخیص وجود اشیاء، متن یا چهره انسان و میزان برجستگی بر اساس روش‌های مختلف، انتخاب کرده و سپس با تجمیع

ناحیه بر اساس روش‌های مختلف است.

در این مقاله که در واقع ادامه کار روش [۶] و [۷] را دنبال می‌کند، روشی جدید برای تعیین نواحی مزاحم بصری بر اساس روش‌های منطق بر چالش عدم توازن دسته‌ها، جهت دسته‌بندی مناسب قطعه‌های تصاویر آموزشی و در ادامه ویژگی‌های مبتنی بر قطعه، استخراج و دسته هر قطعه بر اساس روش نمایش تنک دومرحله‌ای و وزن‌دار نمونه آزمون^۱ بر مبنای سیستم کدگذاری و بازنمایی تنک مشخص می‌شود.

۲- بررسی و مرور کارهای پیشین

تشخیص برجستگی از سال‌های ۱۹۹۰ به بعد یکی از موضوعات محبوب و مشهور در بینایی ماشین بوده [۸] و از آن زمان تاکنون، روش‌ها و مفاهیم بسیاری جهت تشخیص نواحی و اشیای برجسته در یک تصویر با نگرش و انگیزه‌های متفاوت جهت غلبه بر چالش‌های موجود ارائه شده است. از این رو [۹] به بررسی تشخیص نواحی برجسته در زمانی کوتاه با معرفی مفهوم پوشش فضایی^۲ پرداخته است که بر اساس آن یک بازنمایی با ابعاد پایین از تصویر ارائه می‌گردد. مقاله [۱۰] با صرف زمان بیشتر، میزان برجستگی و اهمیت را بر اساس محتوای تصویر و ناحیه تعریف می‌کند. نگرش و استدلال آنان این است که یک ناحیه برجسته است اگر نواحی مشابه تنها در نزدیکی خودش باشند و هرچه نواحی مشابه در فواصل دورتر موجود باشد، احتمال برجستگی آن ناحیه کم است. این رویکرد گرچه به دقت بالایی رسیده است اما مستلزم محاسبه میزان تشابه بین تمام تکه‌ها و نیازمند انجام محاسبات زیاد است. مقاله [۱۱] از این ایده که نواحی برجسته غالباً نزدیک به مرکز تصویر هستند، سعی در اضافه‌کردن این ویژگی مکانی به نواحی و استفاده از آن در دسته‌بندی و تشکیل نقشه برجستگی دارد. آنها همچنین برای اولین بار از هر دو مدل رنگ RGB و Lab همزمان استفاده کردند، چرا که عقیده دارند در برخی تصاویر مدل RGB و در برخی دیگر مدل Lab در تشخیص نواحی برجسته موفق است. استخراج ویژگی‌های مورد نظر آنها نیاز به محاسبات پیچیده ندارد و به این دلیل بسیار مورد توجه بوده‌اند.

مقاله [۱۲] به بررسی ویژگی‌های مؤثر بر تمایز یک تکه تصویر^۳ به عنوان یک مشخصه برای میزان برجستگی آن تکه می‌پردازد. برخلاف روش‌های پیشین که هر کدام فقط بر یکی از موارد تشخیص الگوها، کانال‌های رنگ و یا مشخصات سطح بالاتر تمرکز کرده بودند، این روش سعی در تجمیع و بهره‌برداری از تمامی این ویژگی‌های سه‌گانه دارد. الگوریتم آنها با صرف نظر از ویژگی‌های سطح بالایی مانند تشخیص صورت، تشخیص بدن انسان یا نوشته (متن)، علی‌رغم افزایش سرعت اجرای الگوریتم، تا حدودی از دقت تشخیص کاسته است. از طرفی [۱۳] با استفاده از کدگذاری و بازنمایی تنک به تشخیص برجستگی با استخراج و ترکیب دو نقشه برجستگی سراسری و محلی به صورت بازنمایی رتبه پایین پرداخته است. در این روش از الگوریتم ICA برای یادگیری بر اساس توابع پایه استفاده شده است.

اگرچه تمامی این روش‌ها با توجه به انگیزه اولیه‌شان به موفقیت‌های بسیاری دست یافته‌اند اما هیچ کدام به موضوع مزاحمت بصری یا تقسیم نواحی برجسته به دو دسته نواحی برجسته مهم و نواحی برجسته غیر مهم از نقطه نظر سیستم بینایی و مکانیسم تشخیص مغز انسان نپرداخته‌اند. لذا

1. WTPTSSR
2. Spatial Envelope
3. Patch

ابتدا مطابق با داده‌های آموزشی، واژه‌نامه D ساخته می‌شود. سپس ماتریس قطری W را که هر عنصر از قطر آن فاصله نمونه y با داده‌های آموزشی d است مطابق (۱) محاسبه می‌شود

$$W = \text{diag}([dist(y, d_1), dist(y, d_2), \dots, dist(y, d_n)]) \quad (1)$$

که در آن

$$dist(y, d_i) = \|y - d_i\|^q \quad (2)$$

n تعداد کل داده‌های آموزشی و q پارامتر تعدیل‌کننده محلی می‌باشد. در (۲) اگر $q = 0$ در نظر گرفته شود، روش نمایش تنک دومرحله‌ای و وزن‌دار نمونه آزمون به نمایش تنک دومرحله‌ای تبدیل می‌شود.

لذا بر اساس مطالب فوق در روش پیشنهادی جهت محاسبه بردار ضرایب X با توجه به ماتریس W ، (۳) مطابق با الگوریتم رگرسیون ریج^۳ به صورت زیر بازنویسی می‌شود

$$X = (DD^T + \lambda \times W^T W)^{-1} D^T y \quad (3)$$

همچنین می‌توان (۳) را با استفاده از الگوریتم LASSO [۲۱] یا سایر الگوریتم‌های مطرح و قدرتمند در سیستم کدگذاری و بازنمایی تنک استفاده می‌کرد اما علی‌رغم کاربرد و قابلیت‌های بسیار، در برخی مسایل نسبت به الگوریتم رگرسیون ریج دچار اشکال می‌شوند. عمده برتری رگرسیون ریج نسبت به LASSO در مواردی است که همبستگی بین ویژگی‌ها زیاد و همچنین تعداد نمونه‌ها بسیار بیشتر از تعداد ویژگی‌ها باشد [۲۲]. از آنجا که این دو شرط در چالش مطرح‌شده صادق است، لذا این برتری در روش پیشنهادی مد نظر قرار گرفته و از الگوریتم رگرسیون ریج استفاده شده است. به منظور آشنایی با سایر روش‌های بهینه‌سازی در روش کدگذاری و بازنمایی تنک می‌توان به [۲۳] و [۲۴] رجوع کرد.

در روش پیشنهادی پس از محاسبه بردار X بر اساس (۴) مقدار con_i برای تمام داده‌های آموزشی محاسبه می‌گردد

$$con_i = \|y - d_i x_i\|_p, \quad i = 1 \text{ to } n \quad (4)$$

در این رابطه، x_i ضرایب متناظر با داده d_i در بردار X است. به مانند روش KNN، پس از محاسبه con_i برای تمام داده‌های آموزشی، M داده که کمترین مقدار con_i را دارند به عنوان همسایه‌های نزدیک‌تر انتخاب و با قراردادن آنها در یک ماتریس، واژه‌نامه جدید (\bar{D}) با M عضو تشکیل و دسته‌بندی به روش کدگذاری تنک انجام می‌شود.

مشکل اصلی این روش، مقایسه نمونه آزمایشی با تمام داده‌های موجود جهت تعیین ماتریس W و همچنین محاسبه میزان تشابه آنها مطابق (۴) است. لذا این روش در مواردی که تعداد داده‌های آموزشی زیاد باشد، بسیار زمان‌بر خواهد بود. یک راه مؤثر غلبه بر این مشکل و رفع محدودیت‌های سخت‌افزاری به هنگام پیاده‌سازی، حذف تعدادی از داده‌های کلاس‌های بزرگ‌تر مانند کلاس ۱ و ۲ می‌باشد. در آزمایشات انجام‌شده مشاهده گردید حذف تعدادی از داده‌هایی که فاصله بیشتری از مرکز کلاس‌های خود دارند موجب افزایش چشم‌گیر سرعت می‌شود. راه دیگر غلبه بر این مشکل، کاهش ابعاد هر داده می‌باشد که این مورد در تحقیقات آتی می‌تواند مورد بررسی قرار گیرد.

در شکل ۱ بلوک دیاگرام مراحل و نحوه اجرای الگوریتم به طور کامل قابل مشاهده است. همان طور که گفته شد روش پیشنهادی مقاله شامل بلوک‌های متفاوتی است و در این بین تغییر قطعه‌بندی، استخراج ویژگی و

این ویژگی‌ها در هر قطعه از تصویر، به ویژگی‌های مبتنی بر قطعه دست یافته‌اند. آنها در ادامه به وسیله الگوریتم LASSO که بر پایه کدگذاری تنک^۱ می‌باشد [۲۲]، نسبت به کاهش ویژگی‌ها و دسته‌بندی قطعات مختلف از نظر میزان مزاحمت بصری اقدام کردند. گرچه در [۲۳] به تفصیل در مورد نواحی مزاحم بصری بحث نشده است اما از آن جهت که نویسندگان آن مقاله برای اولین بار این موضوع را مطرح و در این راستا اقدام به تهیه مجموعه دادگان کردند، لذا مطالعات و مقاله آنها از اهمیت بالایی برخوردار است.

در این مقاله با توجه به چالش‌هایی مانند عدم توازن دسته‌ها که در ادامه تشریح خواهند شد، از روش نمایش تنک دومرحله‌ای و وزن‌دار نمونه آزمون به منظور تشخیص نواحی مزاحم بصری استفاده شده است. نتایج پیاده‌سازی نشان می‌دهد که این روش از نظر معیارهای مختلف دارای عملکرد بهتری نسبت به کارهای پیشین بوده و تا حد زیادی بر چالش‌های موجود غلبه خواهد کرد. بخش‌های مختلف به این صورت سازمان‌دهی می‌شود: در بخش ۳ روش پیشنهادی مبتنی بر کدگذاری و بازنمایی تنک ارائه می‌شود. در بخش ۴ ضمن معرفی معیارهای مورد استفاده، نتایج و اعتبارسنجی روش پیشنهادی در مقایسه با سایر روش‌ها بررسی خواهد شد.

۳- روش پیشنهادی مقاله

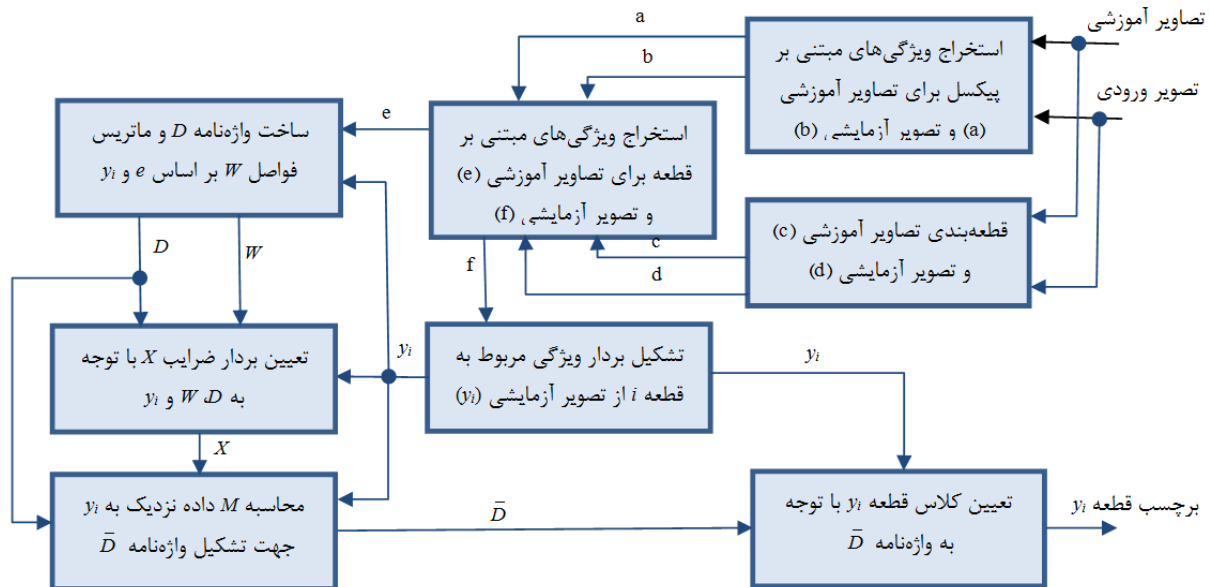
در ابتدای این بخش مروری اجمالی بر روش نمایش تنک دومرحله‌ای و وزن‌دار نمونه آزمون ارائه می‌گردد. در ادامه، به این دلیل که در روش پیشنهادی واژه‌نامه مربوط به کدگذاری و بازنمایی تنک، از کنار هم قرار دادن بردار ویژگی قطعات هر تصویر ساخته می‌شود، لازم است نحوه استخراج و انتخاب ویژگی‌های مورد نظر و چالش‌های مربوط، تشریح گردد. جهت آشنایی با سایر روش‌های ساخت واژه‌نامه در سیستم کدگذاری و بازنمایی تنک می‌توان به [۲۳] و [۲۴] رجوع کرد. در نهایت نحوه دسته‌بندی نمونه‌های آزمایشی توسط روش پیشنهادی بیان می‌شود.

۱-۳ روش نمایش تنک دومرحله‌ای و وزن‌دار نمونه آزمون

در حوزه شناسایی الگو، ثابت شده که ساختار محلی داده‌ها در دسته‌بندی آنها بسیار مهم است. از این رو در [۲۵] اشاره شده که در برخی موارد، اطلاعات محلی داده می‌تواند به دسته‌بندی توسط کدگذاری تنک کمک شایانی کند، لذا [۲۶] با ارائه روش نمایش تنک دومرحله‌ای و وزن‌دار نمونه آزمون سعی در بهره‌گرفتن از اطلاعات محلی داده‌های آموزشی در سیستم کدگذاری تنک دارد. در این روش، ابتدا به ترتیبی که در ادامه بیان خواهد شد، M داده نزدیک‌تر به نمونه آزمایشی را تعیین و سپس فقط بر اساس این داده‌ها، کدگذاری تنک اجرا می‌شود. این روش از الگوریتم نمایش تنک دومرحله‌ای^۲ معرفی شده در [۷] الهام گرفته شده است اما با محاسبه فاصله هر داده آموزشی از نمونه آزمایشی، وزنی را به عنوان میزان تأثیر آن داده آموزشی در نظر گرفته و بدین ترتیب با بهره‌گیری از اطلاعات محلی، دسته‌بندی دقیق‌تری نسبت به نمایش تنک دومرحله‌ای انجام می‌دهد.

به طور خلاصه، نحوه دسته‌بندی و روابط به کار گرفته شده در روش نمایش تنک دومرحله‌ای و وزن‌دار نمونه آزمون بدین صورت است که در

1. Sparse Coding and Representation
2. TPTSSR



شکل ۱: بلوک دیاگرام مراحل مختلف روش پیشنهادی.

تشخیص وجود اشیای مختلف در تصویر [۳۰] است.

• ویژگی‌های منحصر به نواحی مزاحم بصری.

در دسته سوم با توجه به اهمیت میزان برجستگی هر پیکسل و رویکردهای گسترده در این حوزه، این ویژگی بر اساس روش‌های [۸]، [۹]، [۱۱]، [۱۲] و [۳۱] استخراج و با توجه به زمان اجرای پایین آنها، از توانمندی تمامی آنها بهره‌برداری شده است. با توجه به آنچه پیش‌تر هم گفته شد، میزان اهمیت هر پیکسل رابطه معکوسی با میزان مزاحمت آن دارد. لذا این ویژگی نیز مبتنی بر رویکرد [۱۰] استخراج شده است. در این مقاله اشاره شده که صرفاً میزان برجستگی در جلب توجه افراد دخیل نیست و اهمیت اطلاعات موجود در آن ناحیه نیز بسیار مؤثر و تعیین‌کننده است. از طرفی [۳۲] با معرفی مشخصه منحصر به فردی یا غیر معمول بودن^۱، در پی دسته‌بندی نواحی از این نقطه نظر بوده است. این رویکرد مبتنی بر محتوا در مورد نواحی مزاحم هم می‌تواند مؤثر و قابل اتکا باشد.

در ادامه روش پیشنهادی جهت تجمیع ویژگی‌ها، می‌توان از روش‌های مختلفی مانند میانگین‌گیری، محاسبه کمینه یا بیشینه، مد و میانه مقادیر هر مجموعه استفاده کرد. لذا با تحلیل نتایج آزمایشات مختلف، در روش پیشنهادی از روش محاسبه میانه مقادیر، جهت تولید ویژگی‌های مبتنی بر قطعه، پس از قطعه‌بندی تصویر بر اساس روش MCG [۳۳]، جهت استخراج ویژگی‌های پیکسل‌های عضو هر قطعه و تجمیع آنها برای تولید ویژگی‌های مبتنی بر قطعه، استفاده شده است.

پس از تجمیع ویژگی‌های مبتنی بر پیکسل، نیاز است تا ویژگی‌های منحصر به هر قطعه نیز استخراج شوند. بدین منظور ویژگی‌های مکانی و توپولوژیکی هر قطعه، شامل مختصات مکانی آن قطعه در تصویر، محیط، مساحت، طول و عرض، قطر، فاصله از مرکز، جهت یا زاویه و ... نیز در نظر گرفته شده است. از این رو جهت تشکیل بردار ویژگی نهایی، در کنار سایر ویژگی‌های تجمیع‌شده برای هر قطعه، از این دست ویژگی‌ها نیز استفاده شده است. نتایج نشان می‌دهد که اضافه کردن این ویژگی‌های مکانی، کمک شایانی به دقت دسته‌بندی انجام‌شده می‌کند.

مراحل مبتنی بر قطعه پیاده‌سازی شده و در ادامه بقیه جزئیات استفاده‌شده از پیاده‌سازی‌های دیگر مقالات در بلوک‌های روش پیشنهادی به طور کامل تشریح شده است.

۳-۲ نحوه استخراج ویژگی جهت ساخت واژه‌نامه

از مهم‌ترین چالش‌های موجود در مسأله تشخیص نواحی برجسته و مزاحم، به مانند سایر مسایل دسته‌بندی، نحوه بازنمایی تصاویر بر اساس ویژگی‌ها می‌باشد. همین امر سبب شده که اکثر محققین در این زمینه به دنبال بهترین روش‌های بازنمایی ویژگی تصاویر باشند. نکته مهم در استخراج ویژگی جهت تشخیص نواحی مزاحم بصری، برجسته یا موارد مشابه این است که این مسایل متفاوت از مسایل مطابقت تصاویر یا دسته‌بندی اشیای است. به عنوان مثال در مسأله تشخیص چهره، هدف یافتن بیشترین تطابق بین تصویر ورودی و مجموعه تصاویر برجسته‌شده می‌باشد اما در زمینه تشخیص برجستگی، یک شیء در یک تصویر می‌تواند برجسته باشد ولی همان شیء یا ناحیه، در تصویری دیگر ممکن است برجسته نباشد. این موضوع نشان از اهمیت ماهیت تصویر و ترکیب صحنه آن دارد. در مورد تشخیص نواحی مزاحم، موضوع از این هم چالش‌برانگیزتر است زیرا حتی با فرض تشخیص نواحی برجسته، نیاز است این نواحی مجدداً از نظر مکانیسمی مشابه سیستم بینایی انسان مورد تحلیل قرار گرفته و میزان اهمیت یا مزاحمت آنها تعیین شود. لذا در این مقاله سعی شده در کنار لحاظ کردن میزان برجستگی هر ناحیه بر اساس رویکردهای مختلف، از ویژگی‌های مؤثر دیگری نیز بهره‌گیری شود. مطمئناً سایر محققین با آزمایشات دقیق‌تر می‌توانند ویژگی‌های مؤثر و قدرتمندتری استخراج و به حل این چالش کمک کنند.

با توجه به توضیحات داده‌شده، در روش پیشنهادی به طور کلی می‌توان ویژگی‌های مبتنی بر پیکسل استفاده‌شده را به سه دسته تقسیم کرد:

- ویژگی‌های دستی مانند شدت کانال‌های رنگ و احتمال رخداد آنها در اندازه پنجره‌های متفاوت بر اساس [۲۸] در دو مدل رنگ RGB و Lab، تشخیص کنج و لبه‌ها [۲۹] و فاصله هر پیکسل از مرکز و مرزهای تصویر می‌باشد.
- ویژگی‌های سراسری تصویر مانند سطح افق تصویر [۱۰] و [۲۹] و

جدول ۱: تعداد قطعات عضو هر دسته با توجه به پارامتر t .

| پارامتر t | دسته | ۱ | ۲ | ۳ | ۴ | ۵ | ۶ | ۷ | ۸ | +۹ | مجموع | میانگین قطعه در تصویر |
|-------------|------|-------|------|------|------|------|-----|-----|-----|-----|-------|-----------------------|
| ۰.۴ | | ۵۶۷۴۹ | ۶۷۲۴ | ۳۳۲۲ | ۲۳۵۵ | ۱۲۶۶ | ۷۶۱ | ۵۲۹ | ۳۱۴ | ۲۲۹ | ۷۲۲۴۹ | ۱۷۹ |
| ۰.۶ | | ۲۲۰۶۴ | ۴۶۲۲ | ۲۰۹۳ | ۷۰۱ | ۴۷۷ | ۲۹۴ | ۲۳۹ | ۱۸۳ | ۹۱ | ۳۰۷۶۴ | ۸۱ |

۳-۳ دسته‌بندی قطعات تصویر از نظر میزان مزاحمت

در این مقاله، پس از استخراج ویژگی‌های مربوط به داده‌های آموزشی و نمونه آزمایشی، با قراردادن بردار ویژگی‌های مربوط به قطعات داده‌های آموزشی به عنوان مؤلفه‌های هر قطعه از داده‌های آموزشی، واژه‌نامه تشکیل می‌شود. در نهایت، برچسب هر قطعه از ماسک مربوط استخراج شده و جهت تخمین برچسب نمونه‌های آزمایشی در روش کدگذاری تنک، مورد استفاده قرار می‌گیرند.

چالش قابل ذکر در این مرحله این است که با فرض دسته‌بندی قطعات در دو دسته نواحی مزاحم و نواحی غیر مزاحم، بیش از ۹۵ درصد قطعات در دسته نواحی غیر مزاحم قرار خواهند گرفت. این مسأله موجب عدم توازن دسته‌ها و ایجاد چالش‌های مخصوص به خود در الگوریتم‌های دسته‌بندی از جمله اغلب روش‌های مبتنی بر دسته‌بندی با کدگذاری تنک خواهد شد، زیرا روش کدگذاری تنک سعی دارد که نمونه آزمایشی را با ترکیب خطی تعداد محدودی از داده‌های آموزشی بازنمایی کند. حال اگر تعداد داده‌های یک دسته - تعداد مؤلفه‌های یک دسته در واژه‌نامه - بیشتر باشد، آن گاه احتمال این که هر داده‌ای، فارغ از نوع برچسبش به وسیله مؤلفه‌های دسته بزرگ‌تر با خطای کمتری بازنمایی شود، بالا خواهد بود. این موضوع در بخش ۳-۴ به هنگام استفاده از روش OMP [۲۴] مشهود است. جهت غلبه بر مشکل عدم توازن دسته‌ها در حوزه داده‌کاو، راهکارهای مختلفی از جمله تکرار داده‌های دسته کوچک‌تر و یا حذف تعدادی از داده‌های موجود در دسته بزرگ‌تر با رویکردهای مختلف ارائه گردیده که از جمله این رویکردها، حذف داده‌های دور از مرکز دسته است [۳۴]. متأسفانه نتایج حاصل از پیاده‌سازی‌ها نشان می‌دهد که اجرای این ایده کمک قابل توجهی به افزایش دقت دسته‌بندی OMP نخواهد کرد. اما روش نمایش تنک دومرحله‌ای و وزن دار نمونه آزمون از آنجا که به اطلاعات محلی داده‌ها توجه ویژه‌ای دارد، لذا هنگام تشکیل واژه‌نامه \bar{D} ، فقط از M داده نزدیک به نمونه آزمایشی مطابق (۴) و با در نظر گرفتن فاصله آنها استفاده کرده و با حذف بسیاری از داده‌های دور و غیر مشابه با نمونه آزمایشی، تا حد زیادی مشکل دسته‌بندی در این وضعیت را حل می‌کند.

۴- نتایج مقایسه و ارزیابی روش پیشنهادی

به منظور تحلیل دقیق‌تر نتایج، لازم است در ابتدا آشنایی مختصری با مجموعه دادگان استفاده‌شده و نحوه دسته‌بندی قطعات آنها ایجاد گردد و پس از آن معیارهای استفاده‌شده جهت ارزیابی و مقایسه نتایج معرفی می‌شود. سپس نتایج روش پیشنهادی این مقاله با روش‌های مبتنی بر نمایش تنک دومرحله‌ای [۶]، OMP و نتایج [۲۱] به عنوان اصلی‌ترین مقاله این حوزه مقایسه شده است. در نهایت خروجی هر روش بر روی چند تصویر از مجموعه دادگان مختلف مشاهده خواهد شد.

۴-۱ معرفی مجموعه دادگان و نحوه برچسب‌گذاری

قطعات تصاویر موجود در آن

نویسندگان [۲۱] در سال ۲۰۱۵ دو مجموعه دادگان شامل تصاویر و ماسک‌های مربوط به نواحی مزاحم تهیه کرده‌اند. مجموعه دادگان اول حاوی تصاویر شخصی افراد بوده و تعداد آن ۳۷۸ تصویر است که در دسترس عموم قرار نگرفته است. مجموعه دادگان دوم حاوی ۴۰۳ تصویر با اندازه متوسط حدود ۱۰۲۴×۷۶۸ بوده که برخلاف اکثر مجموعه دادگان معرفی‌شده برای تشخیص برجستگی موجود که فاقد پیچیدگی صحنه و تعداد بالایی اشیای موجود در هر تصویر هستند، دارای ابعاد بالا و تعداد قطعات زیاد نیز هستند که موجب چالش‌های به خصوصی خواهد شد. متأسفانه این تنها مجموعه دادگان موجود در این حوزه می‌باشد و از این رو نتایج این مقاله تنها روی آن اعتبارسنجی و با کارهای پیشین مقایسه شده است. جهت آشنایی بیشتر با این مجموعه دادگان و دسته‌بندی اشیای موجود در تصاویر از لحاظ محتوایی، می‌توان به [۲۱] رجوع کرد.

مجموعه دادگان مورد استفاده، ماسک مربوط به هر تصویر، میزان مزاحمت بصری ایجادشده توسط هر پیکسل را در سطوح خاکستری نشان می‌دهد. لذا با اعمال یک حد آستانه، می‌توان پیکسل‌ها را در دسته‌های مختلف دسته‌بندی کرد. در این مقاله، جهت گسسته‌سازی میزان مزاحمت هر پیکسل، قطعات تصاویر که با استفاده از روش MCG به دست آمده، با اعمال حدود آستانه مناسب به ۱۰ دسته مختلف مطابق با ماسک مربوط افزای شده است. اما به دلیل این که کلاس ۱۰ (سطح ۲۲۵ تا ۲۵۵) در مجموعه دادگان مورد استفاده هیچ عضوی ندارد، لذا دسته ۱۰ و ۹ ادغام شده و کلاس +۹ نامیده خواهد شد.

از سوی دیگر، روش قطعه‌بندی MCG قادر است با تغییر یک حد آستانه (t) تعبیه‌شده در الگوریتم، با حساسیت‌های مختلف قطعه‌بندی تصاویر را انجام دهد. لازم به ذکر است که مقدار t در این روش با تعداد قطعات ایجادشده توسط الگوریتم نسبت معکوس دارد. از آنجا که مقدار حد آستانه ۰.۴ نتایج دقیق‌تری حاصل کرد، تمام آزمایشات روش پیشنهادی و سایر روش‌ها بر اساس این حد آستانه صورت گرفته است.

در جدول ۱ که برچسب هر دسته متناسب با شدت مزاحم‌بودن اعضای آن دسته است، تعداد قطعات هر دسته با توجه به دو مقدار مختلف t مشاهده می‌شود. مطابق آنچه که در بخش ۳-۳ گفته شد، ملاحظه می‌شود که دسته یک بیشترین اعضا را دارد که با توجه به ماهیت تصاویر موجود در مجموعه دادگان امری منطقی می‌باشد.

۴-۲ معرفی معیارهای استفاده‌شده جهت اعتبارسنجی

در این مقاله، جهت اعتبارسنجی آزمایشات و نتایج روش‌های مختلف از ۴ معیار مختلف استفاده می‌شود. بدین ترتیب که ابتدا هر روش بر اساس معیارهای شناخته‌شده صحت^۲ و MSE سنجیده می‌شود. سپس با

k دسته مطابق (۷) تعریف کرد

$$AMSE = \frac{1}{k} \sum_{i=1}^k MSE_i \quad (7)$$

۳-۴ نتایج حاصل از آزمایشات مطابق روش‌های مختلف

لازم به ذکر است که چون در روش نمایش تنک دومرحله‌ای و وزن دار نمونه آزمون پارامتر M جهت ساخت واژه‌نامه \bar{D} و پارامتر q در (۴) دو پارامتر مهم و مؤثر در نتایج است لذا در آزمایشات انجام شده، مقادیر مختلفی برای پارامتر M و q در نظر گرفته شده تا اثرگذاری این پارامترها در نتایج نهایی ملاحظه و مقدار بهینه بر اساس معیارهای مختلف انتخاب گردد. به منظور نزدیک شدن به نقاط بهینگی M و q ابتدا با در نظر گرفتن $q=0$ دسته‌بندی به مانند روش نمایش تنک دومرحله‌ای و بر حسب مقادیر مختلف M مانند [۶] انجام می‌شود تا مقدار بهینه M تعیین گردد. سپس با تثبیت این پارامتر، به ازای سایر مقادیر q نیز مجدداً دسته‌بندی انجام می‌شود تا به ازای پارامتر M به دست آمده در مرحله قبل، این پارامتر نیز بهینه گردد. نتایج حاصل از آزمایشات بر اساس روش نمایش تنک دومرحله‌ای به روش یکی به کنار^۲ با در نظر گرفتن مقادیر مختلف پارامتر M بر طبق معیارهای Acc ، AUC ، MSE و $AMSE$ به ترتیب در شکل‌های ۲ و ۳ ملاحظه می‌شود. با توجه به جدول ۱ و موضوع نامتوازن بودن کلاس‌ها، مشخص است که با در نظر گرفتن مقادیر کوچک M ، احتمال اثرپذیری روش نمایش تنک دومرحله‌ای از داده‌های نویز و پرت بسیار بالا خواهد بود، لذا با توجه به عدم دقت بالای ماسک‌های موجود در مجموعه دادگان به کار رفته، دقت روش پیشنهادی با در نظر گرفتن مقادیر کوچک همسایگی پایین بوده است. اما با افزایش مقدار M تا حدود ۲۰۰ دقت رو به افزایش بوده است اما با توجه به این که کوچک‌ترین دسته تعداد ۲۲۹ عضو دارد، لذا در نظر گرفتن مقادیر بزرگ‌تر از ۲۵۰ برای M باعث کاهش دقت در دسته‌بندی داده‌های متعلق به این دسته می‌شود. از این رو با افزایش M و عبور آن از مرز تعداد داده‌های موجود در هر دسته، معیار $AMSE$ با افزایش آبی مواجه می‌شود. با تحلیل آزمایشات صورت‌گرفته، مشاهده می‌شود که با در نظر گرفتن مقادیر بزرگ برای M (۲۵۰) و بزرگ‌تر از آن، عملکرد روش پیشنهادی به روش‌های کلاسیک دسته‌بندی اسپارس مانند روش رگرسیون ریبج نزدیک شده و مطابق آنچه گفته شد تمامی نمونه‌های آموزشی به دسته‌های بزرگ‌تر (دسته ۱ و ۲) متعلق می‌گردند. به طور کلی می‌توان گفت با در نظر گرفتن مقادیر بسیار بزرگ M (بزرگ‌تر از ۲۰۰)، عملکرد روش پیشنهادی به روش OMP بر اساس هر دو معیار نزدیک خواهد شد و تنها دلیل اختلافات جزئی در عملکرد آنها به الگوریتم‌های بهینه‌سازی به کار رفته برمی‌گردد. با توجه به نتایج می‌توان این گونه استدلال کرد که برای پارامتر M باید مقداری نزدیک به تعداد داده‌های موجود در کوچک‌ترین دسته در نظر گرفت تا هم احتمال تأثیر نویز و اشکالات مجموعه دادگان کاهش یافته و هم مجدداً مشکل عدم توازن دسته‌ها پیش نیاید.

با توجه به جدول ۱ و شکل‌های ۲ و ۳ می‌توان تعداد همسایگی ۱۵۰ را به عنوان یک مقدار نسبتاً بهینه (بهینه محلی) بر اساس هر ۴ معیار دانست. با این فرض میزان دقت روش نمایش تنک دومرحله‌ای و وزن دار نمونه آزمون با در نظر گرفتن مقادیر مختلف برای q در (۲) در شکل‌های ۴ و ۵ ملاحظه می‌شود. با توجه به شکل‌های ۴ و ۵ مشخص

توجه به عدم توازن اعضای هر کلاس، جهت قابل اتکاتر کردن نتایج معیار صحت، از معیار AUC و جهت قابل اتکاتر کردن نتایج معیار MSE معیاری جدید به نام $AMSE$ معرفی و نتایج هر آزمایش بر اساس این معیار نیز مورد تحلیل قرار می‌گیرد.

۴-۲-۱ معیار صحت

معیار صحت از معیارهای متداول و شناخته شده در یادگیری ماشین و شناسایی الگو است. با فرض وجود تعداد N نمونه آزمایشی، صحت دسته‌بندی در حالتی که برچسب T نمونه به درستی تشخیص داده شود، به سادگی از (۵) حاصل می‌شود

$$Acc = \frac{T}{N} \quad (5)$$

۴-۲-۲ معیار AUC

در مواردی که توزیع داده‌ها غیر متوازن است نیاز به معیاری دقیق‌تر با رویکردی متفاوت از معیار صحت احساس می‌شود زیرا با فرض وجود اکثر داده‌ها در یک دسته، اگر به صورت پیش‌فرض هر نمونه آزمایشی به این دسته تعلق داده شود، آن گاه صحت دسته‌بندی بسیار بالا خواهد بود. این در حالی است که هیچ نمونه‌ای به دسته کوچک‌تر متعلق نشده است. در چالش مطرح شده در این مقاله نیز به دلیل غیر مزاحم بودن اکثر قطعات وجود در یک تصویر، اگر هر قطعه غیر مزاحم تشخیص داده شود معیار صحت بسیار بالا خواهد بود، لذا مشخص است تکیه بر این معیار به تنهایی، موجب بروز اشتباه در ارزیابی روش‌ها می‌گردد.

در این مقاله از آنجا که برچسب تخمین زده شده برای هر نمونه آزمایشی یک عدد طبیعی در بازه یک تا ۹ خواهد بود، لذا معیار AUC با فرض وجود n_i داده در k دسته مختلف در حالتی که تعداد t_i داده از کلاس i درست دسته‌بندی شود برابر با (۶) خواهد بود

$$AUC = \frac{\sum_{i=1}^k Acc_i}{k} \quad (6)$$

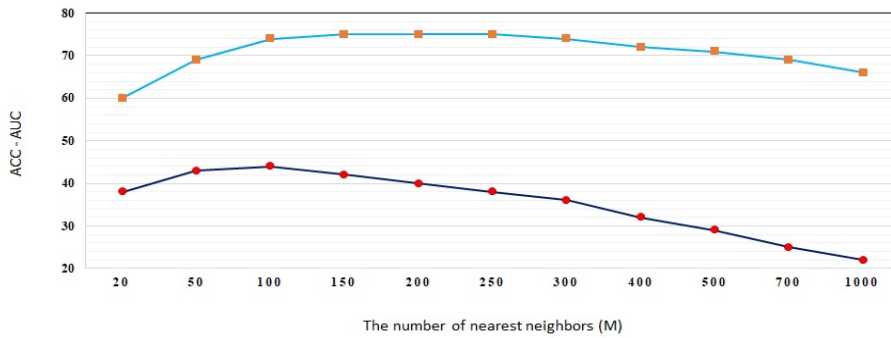
که Acc_i طبق (۵) صحت دسته‌بندی دسته i و برابر با t_i/n_i است.

۴-۲-۳ معیار خطای میانگین مربعات

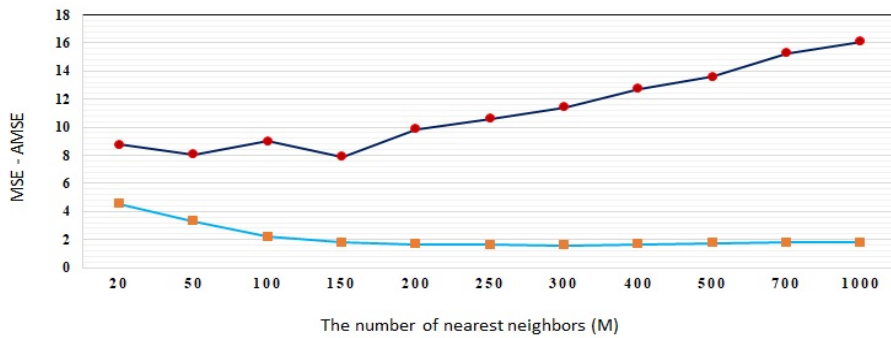
با توجه به این که قطعات تصاویر موجود در ۹ دسته به گونه‌ای افزاز شده‌اند که شماره هر کلاس متناسب با میزان مزاحمت اعضای آن دسته است، لذا اعتبارسنجی نتایج بر اساس معیار خطای میانگین مربعات (MSE) معتبرتر از معیارهای صحت و AUC می‌باشد چرا که در این دو معیار، اگر نمونه با برچسب i به اشتباه به هر یک از ۸ کلاس دیگر متعلق شود، ارزش آن دسته‌بندی مطلقاً صفر خواهد بود. اما در معیار MSE اگر نمونه با برچسب i به دسته نزدیک‌تری متعلق شود، با ارزش‌تر از حالتی است که به دسته دورتر متعلق شود.

۴-۲-۴ معیار متوسط خطای میانگین مربعات (AMSE)

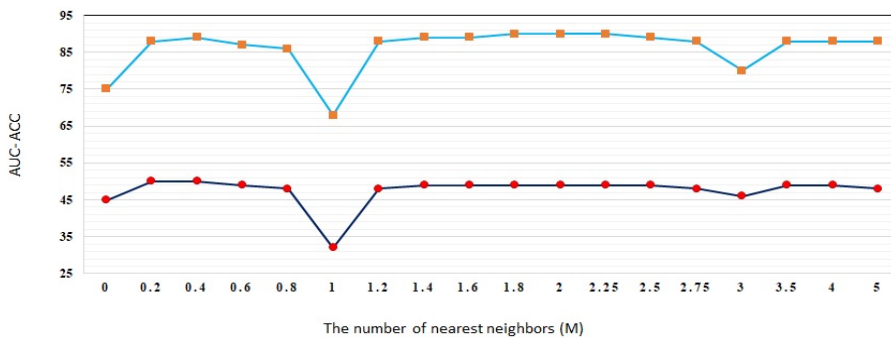
به دلیل عدم توازن شدید دسته‌ها و دلایلی که در بخش ۴-۲-۲ ذکر شد، در ابتدا برای اعضای هر کلاس ابتدا معیار MSE محاسبه و سپس از ۹ مقدار MSE به دست آمده برای هر کلاس میانگین گرفته می‌شود تا با معیار دقیق‌تری نیز نتایج به دست آمده ارزیابی گردند. بدین منظور اگر برای داده‌های متعلق به کلاس شماره i معیار MSE محاسبه و MSE_i نامیده شود، آن گاه می‌توان معیار دیگری به نام $AMSE$ با فرض وجود



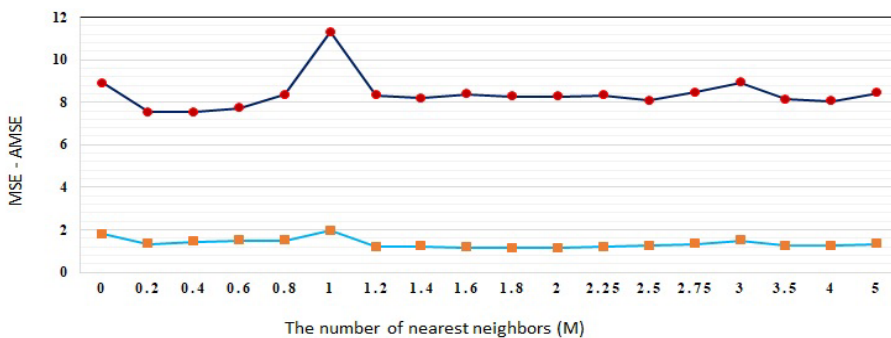
شکل ۲: عملکرد روش مبتنی بر نمایش تنک دومرحله‌ای [۶] بر طبق معیارهای صحت (بالا) و AUC (پایین) بر اساس تغییرات همسایگی M .



شکل ۳: عملکرد روش مبتنی بر نمایش تنک دومرحله‌ای [۶] بر طبق معیارهای MSE (پایین) و AMSE (بالا) بر اساس تغییرات همسایگی M .



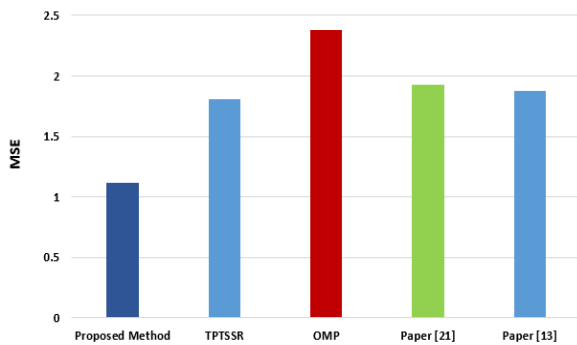
شکل ۴: عملکرد روش پیشنهادی بر طبق معیارهای صحت (بالا) و AUC (پایین) بر اساس تغییرات پارامتر q .



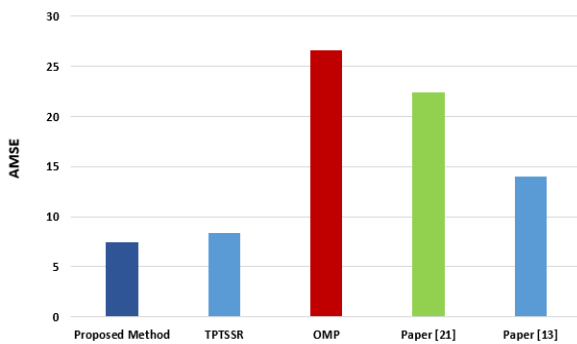
شکل ۵: عملکرد روش پیشنهادی بر طبق معیارهای MSE (پایین) و AMSE (بالا) بر اساس تغییرات پارامتر q .

که تکیه بر روش‌های تشخیص برجستگی به تنهایی نمی‌تواند خروجی مناسبی در زمینه تشخیص نواحی مزاحم بصری تولید کند. در شکل‌های ۶ و ۷ نتایج حاصل از روش پیشنهادی با در نظر گرفتن مقادیر بهینه M و q ، در مقایسه با روش‌های نمایش تنک دومرحله‌ای [۶]، OMP، روش [۳۵] و [۲۱] به عنوان تنها مقاله مطرح در این حوزه و استفاده از روش [۱۳] بر اساس کدگذاری و بازنمایی تنک جهت تعیین مزاحمت بصری، به ترتیب بر اساس معیارهای معرفی شده مقایسه شده‌اند. طبق نتایج شکل‌های ۶ و ۷ مشخص است که روش OMP اکثر قطعات را به دسته

است که بالاترین مقدار صحت به دست آمده برابر ۹۰ درصد و این میزان برای معیار AUC برابر با ۴۹ درصد است. همچنین بهترین دقت بر اساس معیارهای MSE و AMSE به ترتیب برابر با ۱/۱۳ و ۷/۵۲ می‌باشد. گرچه [۳۵] در زمینه تشخیص برجستگی است اما در بخش مقایسه نتایج، نتایج حاصل از اجرای این روش نیز لحاظ شده است. مشخص است که به دلیل نزدیک بودن دو مفهوم برجستگی و مزاحمت بصری همچنان که در بخش استخراج ویژگی نیز ذکر شد، در برخی از تصاویر این روش می‌تواند نتیجه مطلوب داشته باشد اما با توجه به نتایج کلی مشخص است

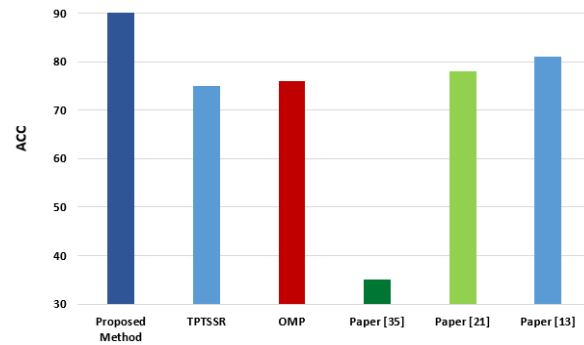


(الف)

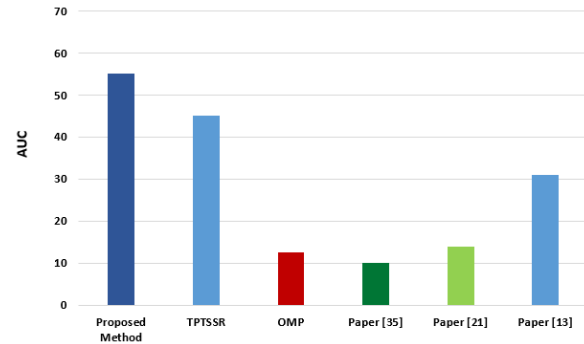


(ب)

شکل ۷: عملکرد روش پیشنهادی در قیاس با روش‌های نمایش تنک دومرحله‌ای [۶]، OMP و [۲۱] و [۱۳] بر اساس معیارهای (الف) MSE و (ب) AMSE.



(الف)



(ب)

شکل ۸: عملکرد روش پیشنهادی در قیاس با روش‌های نمایش تنک دومرحله‌ای [۶]، OMP و [۳۵]، [۲۱] و [۱۳] بر اساس معیارهای (الف) صحت و (ب) AUC.

یک متناظر کرده است زیرا در حالتی که تمام داده‌ها به دسته شماره یک تعلق گیرد، آن گاه مقادیر معیارهای ACC و AUC به ترتیب برابر با $78/56$ و $11/11$ و میزان دقت بر اساس معیارهای MSE و $AMSE$ به ترتیب برابر با $2/25$ و $22/6$ خواهند بود. همچنین با در نظر گرفتن ماتریس W در (۳) دقت روش پیشنهادی به شکل چشم‌گیری نسبت به روش نمایش تنک دومرحله‌ای افزایش یافته است.

از طرفی استفاده از روش کدگذاری و بازنمایی تنک [۱۳] در زمینه تشخیص مزاحمت بصری تصویر نشان می‌دهد با توجه به نامتوازن بودن دسته‌های تصاویر مجموعه دادگان، تفکیک درستی بین این دسته‌ها قرار نمی‌دهد و اختصاص به دسته‌های اول و دوم باعث افزایش مقادیر معیارهای ACC و AUC شده است. در این روش به دلیل بررسی برجستگی از دید سراسری و محلی، قابلیت تفکیک‌پذیری بالاتری نسبت به اجزا و هم‌رنگی در پس‌زمینه را نشان می‌دهد و میزان نواحی تشخیص داده شده به عنوان مزاحم بصری را بیشتر از حد واقعی و اشتباه تشخیص می‌دهد. با توجه به شکل ۲ و ۳ همان طور که مشهود است با افزایش تعداد همسایگی نتایج روش پیشنهادی و روش [۱۳] به یک محدوده نزدیک می‌شوند که این امر نشان از برتری عملکرد روش پیشنهادی در مواجهه با تصاویر با دسته‌های نامتوازن را نشان می‌دهد. از طرفی به این دلیل که نتایج [۳۵] بر اساس معیارهای MSE و $AMSE$ بسیار ضعیف بود، از ذکر نتایج این روش به منظور عدم برهم خوردن مقیاس نمودارها خودداری شده است. این مقادیر به ترتیب برابر $19/38$ و $81/8$ می‌باشد.

۴-۴ خروجی هر روش بر تصاویر مختلف

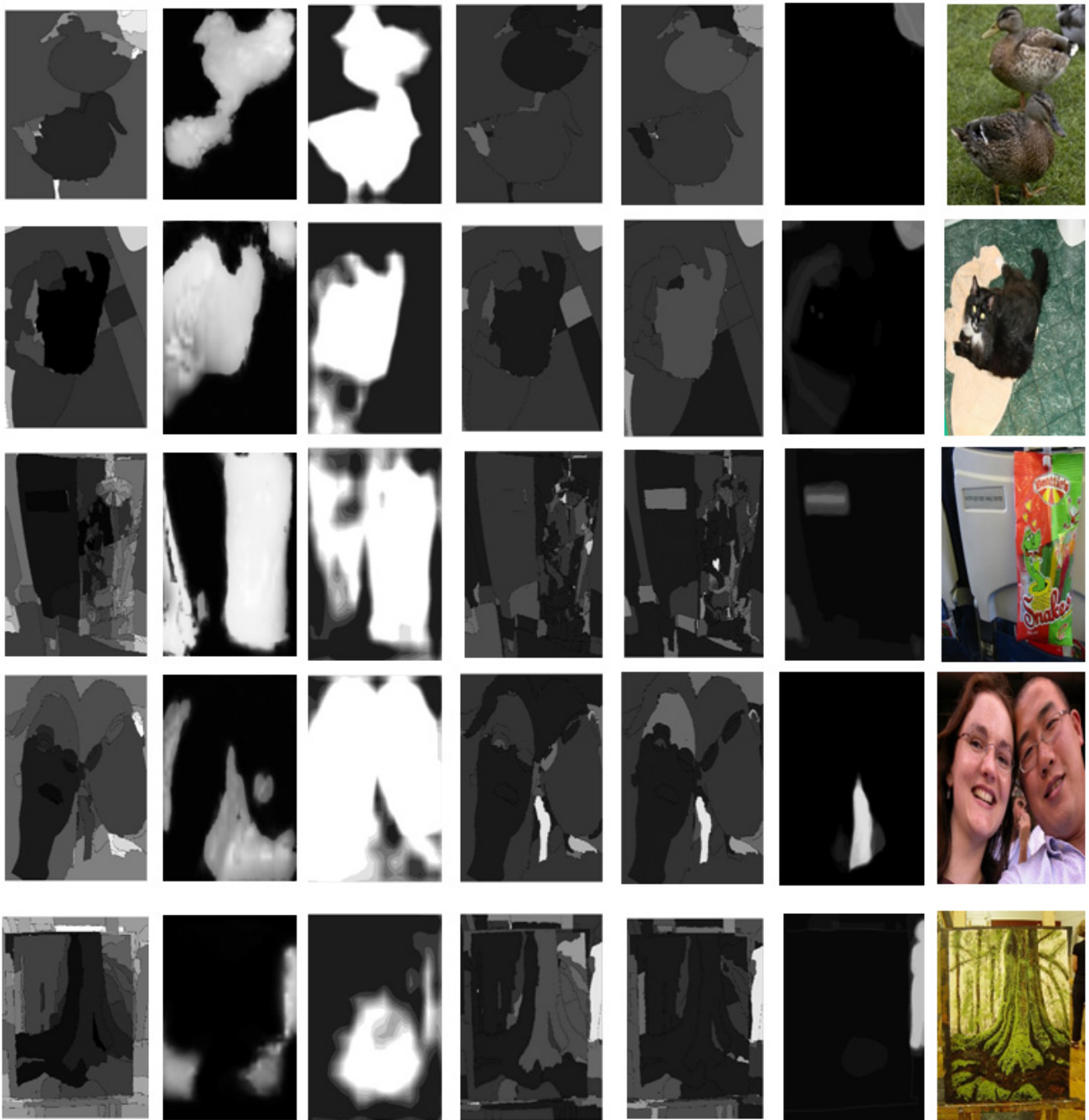
در این بخش خروجی روش‌های مختلف روی چند تصویر از مجموعه دادگان [۲۱] و مجموعه دادگان تهیه‌شده توسط نویسندگان این مقاله آماده شده است. در شکل ۸ چند تصویر از مجموعه دادگان استفاده‌شده، ماسک نواحی مزاحم، خروجی‌های روش پیشنهادی، روش نمایش تنک دومرحله‌ای [۶]، روش [۳۵]، روش [۱۳] و [۲۱] ملاحظه می‌شود.

۵- بحث و نتیجه‌گیری

در این مقاله مفهومی جدید و کمتر شناخته شده به نام نواحی مزاحم بصری معرفی و برای اولین بار تعریفی جامع از آن ارائه گردید. مشخصه اصلی این نواحی میزان برجستگی بالا و عدم بهره‌مندی از اطلاعات مفید و مهم است. گرچه در حوزه تشخیص نواحی برجسته و توجه بصری روش‌های بسیاری با رویکردهای مختلف ارائه شده است اما در زمینه نواحی مزاحم بصری چالش‌های حل‌نشده بسیاری وجود دارد. از مهم‌ترین آنها مدل‌کردن مکانیسمی مشابه عملکرد قشر بینایی و مغز انسان از نظر توجه بصری می‌باشد، لذا استخراج ویژگی‌های جدید با رویکرد تشخیص نواحی مزاحم، جز با تحلیل دقیق مکانیسم تشخیص و تصمیم‌گیری در مغز انسان میسر نخواهد بود.

در این مقاله جهت دسته‌بندی قطعات از سیستم کدگذاری و بازنمایی تنک بر اساس روشی منطبق بر چالش عدم توازن دسته‌ها استفاده شد. در روش پیشنهادی این مقاله، با استفاده از روش نمایش تنک دومرحله‌ای و وزن‌دار نمونه آزمون که خود بر پایه سیستم کدگذاری و بازنمایی تنک است، پس از بهینه‌کردن پارامترهای مختلف این روش، دقت بالایی حاصل شده است. مشکل اصلی روش استفاده‌شده زمان‌بر بودن آن است که با کاهش ابعاد هر داده یا حذف تعدادی از آنها بر اساس رویکردهای مختلف می‌توان تا حدود زیادی بر این مشکل غلبه کرد.

جهت ارزیابی و مقایسه روش پیشنهادی با سایر روش‌ها از ۴ معیار مختلف با رویکردهای مختلف جهت مشخص شدن نقاط قوت و ضعف هر روش استفاده شده است. نتایج نشان می‌دهد روش پیشنهادی این مقاله از نظر دقت بر طبق ۳ معیار از معیارهای معرفی‌شده نسبت به کارهای پیشین برتری دارد. با این حال از آنجا که مفهوم مزاحم بصری به تازگی در پردازش تصویر و بینایی ماشین مطرح شده است، تا کنون مطالعات گسترده‌ای روی آن صورت نگرفته و راهکاری دقیق و بلادرنگ برای



شکل ۸: تصاویر موجود در (الف) مجموعه دادگان [۲۱]، (ب) ماسک نواحی مزاحم، (ج) خروجی روش پیشنهادی، (د) خروجی روش نمایش تنک دومرحله‌ای [۶]، (ه) خروجی [۳۵]، (و) خروجی روش [۱۳] و (ز) خروجی [۲۱].

برخی اشکالات تنها مجموعه دادگان موجود، در آینده مجموعه دادگانی جامع، متنوع و دقیق‌تر در حوزه تشخیص نواحی مزاحم بصری تهیه و در اختیار سایر محققین قرار دهند.

مراجع

[1] N. Senthilkumar and J. Thimmiraja, "Histogram equalization for image enhancement using MRI brain images," in *Proc. World Congress on Computing and Communication Technologies WCCCT'14*, pp. 80-83, Trichirappalli, India, 27 Feb.- 1 Mar. 2014.

[2] V. A. Mateescu and I. V. Bajic, "Visual attention retargeting," *IEEE MultiMedia*, vol. 23, no. 1, pp. 82-91, Jan. 2016.

[3] T. V. Nguyen, et al., "Image re-attentionizing," *IEEE Trans. on Multimedia*, vol. 15, no. 8, pp. 1910-1919, Dec. 2013.

[4] R. Mechrez, E. Shechtman, and L. Zelnik-Manor, "Saliency driven image manipulation," in *Proc. Conf. on Applications of Computer*

تشخیص آن ارائه نشده است، لذا این چالش می‌تواند در آینده بیشتر مورد مطالعه و بررسی قرار گیرد.

لازم به ذکر است که تشخیص نواحی مزاحم بصری، کاربردهای بسیاری دارد که مواردی از آن در بخش مقدمه ذکر شد. اما از کاربردهای مهم دیگر، بهبود کیفیت بصری تصویر با رویکرد حذف یا ترمیم این نواحی می‌باشد که می‌تواند موجب افزایش زیبایی تصویر گردد. همچنین استخراج مشخصات این نواحی در یک تصویر، برای مثال تعداد این نواحی، محل قرارگیری، مساحت و ... می‌تواند به عنوان ویژگی‌هایی مؤثر در کارهای آتی جهت تعیین میزان زیبایی یک تصویر، به عنوان پارامتری مؤثر مورد استفاده قرار گیرد.

از مهم‌ترین مشکلات موجود در این زمینه نبود مجموعه دادگان جامع و دقیق است. از این رو نویسندگان این مقاله سعی دارند علاوه بر رفع

- [23] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210-227, Feb. 2009.
- [24] Z. Zhang, Y. Xu, J. Yang, X. Li, and D. Zhang, "A survey of sparse representation: algorithms and applications," *IEEE Access*, vol. 3, pp. 490-530, May 2015.
- [25] K. Yu, T. Zhang, and Y. Gong, "Nonlinear learning using local coordinate coding," in *Proc. of the 22nd International Conference on Neural Information Processing Systems, NIPS'09*, pp. 2223-2231, Vancouver, BC, Canada, 7-10 Dec. 2009.
- [26] Z. Liu, J. Pu, M. Xu, and Y. Qiu, "Face recognition via weighted two phase test sample sparse representation," *Neural Processing Letters*, vol. 41, no. 1, pp. 43-53, Jan. 2015.
- [27] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," *J. of the Royal Statistical Society, Series B (Statistical Methodology)*, vol. 67, no. 2, pp. 301-320, Aug. 2005.
- [28] T. Judd, K. Ehinger, F. Durand, and A. Torralba, "Learning to predict where humans look," in *Proc. IEEE 12th Int. Conf. on Computer Vision*, pp. 2106-2113, Kyoto, Japan, 29 Sept.- 2 Oct. 2009.
- [29] E. Rosten, R. Porter, and T. Drummond, "Faster and better: a machine learning approach to corner detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 32, no. 1, pp. 105-119, Jan. 2010.
- [30] C. L. Zitnick and P. Dollar, "Edge boxes: locating object proposals from edges," in *Proc. European Conf. on Computer Vision, ECCV'14*, pp. 391-405, Zurich, Switzerland, 6-12 Sept. 2014.
- [31] M. M. Cheng, N. J. Mitra, X. Huang, P. H. Torr, and S. M. Hu, "Global contrast based salient region detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 37, no. 3, pp. 569-582, Mar. 2015.
- [32] M. Wang, J. Konrad, P. Ishwar, K. Jing, and H. Rowley, "Image saliency: from intrinsic to extrinsic context," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition, CVPR'11*, pp. 417-424, Colorado Springs, CO, USA, 20-25 Jun. 2011.
- [33] K. K. Maninis, J. Pont-Tuset, P. Arbelaez, and L. Van Gool, "Convolutional oriented boundaries: from image segmentation to high-level tasks," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 40, no. 4, pp. 819 - 833, Apr. 2018.
- [34] Y. Sun, A. K. Wong, and M. S. Kamel, "Classification of imbalanced data: a review," *International J. of Pattern Recognition and Artificial Intelligence*, vol. 23, no. 4, pp. 687-719, Nov. 2009.
- [35] X. Li, et al., "DeepSaliency: multi-task deep neural network model for salient object detection," *IEEE Trans. on Image Processing*, vol. 25, no. 8, pp. 3919-3930, Oct. 2016.
- [5] A. Borji and L. Itti, "State-of-the-art in visual attention modeling," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 35, no. 1, pp. 185-207, Jan. 2013.
- [6] F. Sabouri and F. Yaghmaee, "Visual distractors detecting in images using TPTSSR," in *Proc. 10th Iranian Conf. on Machine Vision and Image Processing, MVIP'17*, pp. 87-92, Isfahan, Iran, 22-23 Nov. 2017.
- [7] Y. Xu, D. Zhang, J. Yang, and J. Y. Yang, "A two-phase test sample sparse representation method for use with face recognition," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 21, no. 9, pp. 1255-1262, Sept. 2011.
- [8] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 20, no. 11, pp. 1254-1259, Nov. 1998.
- [9] A. Oliva and A. Torralba, "Modeling the shape of the scene: a holistic representation of the spatial envelope," *International J. of Computer Vision*, vol. 42, no. 3, pp. 145-175, May-Jun. 2001.
- [10] S. Goferman, L. Zelnik-Manor, and A. Tal, "Context-aware saliency detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 34, no. 10, pp. 1915-1926, Oct. 2012.
- [11] A. Borji and L. Itti, "Exploiting local and global patch rarities for saliency detection," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition, CVPR'12*, pp. 478-485, Providence, RI, USA, 16-21 Jun. 2012.
- [12] R. Margolin, A. Tal, and L. Zelnik-Manor, "What makes a patch distinct?," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 1139-1146, Portland, OR, USA, 23-28 Jun. 2013.
- [13] Q. Fan and C. Qi, "Saliency detection based on global and local shortterm sparse representation," *Neurocomputing*, vol. 175, no. ???, pp. 81-89, Oct. 2016.
- [14] N. Dhavale and L. Itti, "Saliency-based multifoveated MPEG compression," in *Proc. 7th Int. Symp. on Signal Processing and Its Applications*, vol. 1, pp. 229-232, Paris, France, 4-4 Jul. 2003.
- [15] C. Guo and L. Zhang, "A novel multiresolution spatiotemporal saliency detection model and its applications in image and video compression," *IEEE Trans. on Image Processing*, vol. 19, no. 1, pp. 185-198, Jan. 2010.
- [16] G. Evangelopoulos, K. Rapantzikos, A. Potamianos, P. Maragos, A. Zlatintsi, and Y. Avrithis, "Movie summarization based on audiovisual saliency detection," in *Proc. 15th IEEE Int. Conf. on Image Processing, ICIP'08*, pp. 2528-2531, San Diego, CA, USA, 12-15 Oct. 2008.
- [17] D. Simakov, Y. Caspi, E. Shechtman, and M. Irani, "Summarizing visual data using bidirectional similarity," *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition, CVPR'08*, 8 pp., Anchorage, AK, USA, 23-28 Jun. 2008.
- [18] L. Zhang, Y. Shen, and H. Li, "VSI: a visual saliency-induced index for perceptual image quality assessment," *IEEE Trans. on Image Processing*, vol. 23, no. 10, pp. 4270-4281, Oct. 2014.
- [19] R. Achanta and S. Susstrunk, "Saliency detection for content-aware image resizing," in *Proc. 16th IEEE Int. Conf. on Image Processing, ICIP'09*, pp. 1005-1008, Cairo, Egypt, 7-10 Nov. 2009.
- [20] Y. Fang, et al., "Saliency-based stereoscopic image retargeting," *Information Sciences*, vol. 372, pp. 347-358, 1 Dec. 2016.
- [21] O. Fried, E. Shechtman, D. B. Goldman, and A. Finkelstein, "Finding distractors in images," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition, CVPR'15*, pp. 1703-1712, Boston, MA, USA, 7-12 Jun. 2015.
- [22] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. of the Royal Statistical Society, Series B (Methodological)*, vol. 58, no. 1, pp. 267-288, Jan. 1996.

فردین صبوری تحصیلات خود را در مقطع کارشناسی در رشته مهندسی کامپیوتر در سال ۱۳۹۲ و مقطع کارشناسی ارشد را در سال ۱۳۹۷ در رشته مهندسی کامپیوتر گرایش هوش مصنوعی دانشگاه سمنان به پایان رسانده است. زمینه‌های تحقیقاتی مورد علاقه ایشان عبارتند از: پردازش تصویر و یادگیری ماشین.

فرزین یغمایی در سال ۱۳۷۵ مدرک کارشناسی مهندسی کامپیوتر خود را از دانشگاه صنعتی امیرکبیر و در سال‌های ۱۳۸۱ و ۱۳۸۹ به ترتیب مدارک کارشناسی ارشد و دکتری هوش مصنوعی را از دانشگاه صنعتی شریف دریافت کرده است. ایشان در حال حاضر دانشیار دانشکده مهندسی برق و کامپیوتر دانشگاه سمنان می‌باشد. از زمینه‌های علمی مورد علاقه ایشان می‌توان پردازش تصویر و ویدئو، پردازش زبان فارسی و تشخیص الگو را نام برد.