

## شناسایی خودکار مکان مورد توجه به کمک توزیع غیر پارامتری مبتنی بر ریسک بیزی

مهناز رضوی، امیرحسین طاهری‌نیا و هادی صدوقی یزدی

### چکیده

در این مقاله روش جدیدی برای شناسایی خودکار مکان‌های مورد توجه افراد با استفاده از دوربین‌های نظارتی در فروشگاه‌های بزرگ ارائه شده است. منظور از مکان مورد توجه، ناحیه‌ای در تصویر است که افراد بیشتری رفت و آمد داشته‌اند. بدین منظور ابتدا با استفاده از روش تشخیص افراد مبتنی بر توصیف گرهای HOG، انسان از دیگر اشیاء موجود در تصویر متمایز می‌شود و هر فرد تشخیص داده شده به عنوان یک رخداد در تصویر تلقی می‌شود. سپس مورد توجه‌ترین مکان با استفاده از توزیع غیر پارامتری مبتنی بر ریسک بیزی بر روی موقعیت مکانی افراد تشخیص داده شده، بدست می‌آید. در توزیع پیشنهادی هسته جدیدی تعریف شده است که از کارایی بالایی برخوردار است. در ریسک پیشنهادی نیز تابع ضرر جدیدی تعریف شده است که صحت بالاتری نسبت به تابع ضرر مربعی در محاسبه مراکز توجه دارد و در توزیع‌های مختلف در محاسبه قله‌های توزیع بهتر عمل می‌کند. برای ارزیابی روش از فریم‌های ویدئویی که از دوربین‌های زنده نظارتی در فروشگاه‌های مختلف گرفته شده، استفاده شده است. در ارزیابی هسته پیشنهادی کارایی آن ۸۵٪ بدست آمده است و در ارزیابی صحت تابع ضرر پیشنهادی به طور میانگین بر روی داده‌های مصنوعی ۹۳٫۵٪ و بر روی داده‌های واقعی ۹۰٪ بدست آمده است که در مقایسه با روش‌های موجود بهتر است.

### کلیدواژه‌ها

مکان مورد توجه، رخداد، تابع توزیع غیر پارامتری، ریسک بیزی، دوربین‌های زنده نظارتی.

عمومی، بررسی آمار مشتریان فروشگاه‌ها و بررسی رفتار مسافران مترو کاربرد دارند [۱].

### ۱ مقدمه

امروزه تشخیص ناهنجاری در ویدئوهای نظارتی یکی از مسائل مهم و پرکاربرد در زمینه‌های مختلف است. تشخیص ناهنجاری در مکان‌هایی با حضور افراد مختلف کار چالش برانگیزی است زیرا نوع ناهنجاری در تشخیص آن اهمیت دارد. اگر سیستم نظارتی وجود داشته باشد که یک یا چندین ناظر تصاویر را به صورت برخط بررسی کنند. ناظر ویدئویی ناهنجاری در مکان‌های خلوت را ساده‌تر تشخیص می‌دهد. اما تشخیص ناهنجاری در مکان‌هایی با ازدحام جمعیت پیچیده‌تر خواهد بود. در این مقاله، مکان‌هایی که دارای ازدحام بیشتری هستند و برای تشخیص ناهنجاری نیاز به بررسی بیشتری دارند بدست می‌آیند و همچنین در فروشگاه‌های کوچک و بزرگ برای بدست آوردن مورد توجه‌ترین مکان استفاده می‌شوند.

هدف اصلی این مقاله، شناسایی مکان‌های مورد توجه مشتریان فروشگاه‌ها است و این مکان‌ها به کمک توزیع غیر

تحلیل جمعیت برای حل بسیاری از مشکلات جامعه مهم و ضروری است. ایجاد سیستمی هوشمند که در طراحی فضاها، عمومی و زیرساخت‌ها، توسعه و اصلاح آنها استفاده می‌شود. این دسته از سیستم‌های هوشمند در تشخیص ناهنجاری در مکان‌های

این مقاله در مهرماه ۱۳۹۶ دریافت، در اسفندماه ۱۳۹۷ سومین بازنگری و در فروردین‌ماه ۱۳۹۸ پذیرفته شد

<sup>۱</sup> کارشناس ارشد هوش مصنوعی، آزمایشگاه تحقیقاتی شناسایی الگو، گروه کامپیوتر، دانشکده مهندسی، دانشگاه فردوسی مشهد

رایانامه: [mah.razavi@mail.um.ac.ir](mailto:mah.razavi@mail.um.ac.ir)

<sup>۲</sup> گروه کامپیوتر، دانشکده مهندسی، دانشگاه فردوسی مشهد

رایانامه: [taherinia@um.ac.ir](mailto:taherinia@um.ac.ir)

<sup>۳</sup> گروه کامپیوتر، دانشکده مهندسی، دانشگاه فردوسی مشهد

رایانامه: [h-sadoghi@um.ac.ir](mailto:h-sadoghi@um.ac.ir)

نویسنده مسئول: امیرحسین طاهری‌نیا

۲. در مرحله بعد با ایده گرفتن از تابع تخمین چگالی هسته (KDE)<sup>۵</sup> و ریسک بیزی، توزیع افراد تشخیص داده شده بدست می آید که همان مکان‌های مورد توجه است و سپس قله‌های تابع توزیع با استفاده از روش پیشنهادی محاسبه می‌شوند که مراکز نواحی مورد توجه هستند.

## ۲ مروری بر تحقیقات گذشته

تعداد زیادی پژوهش در این زمینه انجام شده است که با توجه به موضوعات و روش‌هایی که در این مقاله از آن‌ها استفاده شده است، مقالات به سه دسته کلی تقسیم می‌شوند: دسته اول مقالات مرتبط به تشخیص افراد، دسته دوم مقالات مرتبط به توزیع غیر پارامتری و دسته سوم مقالات مرتبط به شناسایی مکان مورد توجه<sup>۶</sup> هستند.

### ۲-۱ مقالات در زمینه تشخیص افراد

تعداد زیادی پژوهش در زمینه تشخیص و ردیابی افراد در سال‌های اخیر انجام شده است. این پروژه‌ها در سطوح و با اهداف مختلف انجام شده‌اند. همانطور که در ادامه بیان خواهد شد، برای تشخیص افراد و ردیابی افراد، از یک یا چند دوربین نظارتی استفاده می‌شود. بسیاری از پروژه‌ها برای محیط‌های داخل ساختمان و بعضی برای محیط‌های خارج ساختمان و بسیاری نیز برای هردو انجام گرفته‌اند. همچنین برخی از پروژه‌ها همپوشانی بین افراد را برطرف کرده‌اند ولی برخی فرض کرده‌اند که همپوشانی بین افراد ایجاد نشود.

به طور کلی می‌توان مقاله‌های شمارش افراد را به سه دسته تقسیم کرد<sup>[۴]</sup>:

#### ۲-۱-۱ شمارش بوسیله خوشه‌بندی مسیر:

این روش‌ها شمارش افراد را بر طبق روش‌های ردیابی انجام می‌دهند. Topkayam و همکارانش<sup>[۵]</sup> خوشه‌بندی افراد را با استفاده از یک تشخیص دهنده افراد عمومی بهبود بخشیده‌اند. این روش چندین نوع ویژگی را شامل می‌شود مانند رنگ، ویژگی‌های مکانی و زمانی در خوشه‌بندی. این روش‌ها برای بعضی از کاربردها می‌توانند عملکرد را به طور نسبی بهتر کنند. گرچه که آنها معمولاً زمانبر هستند.

#### ۲-۱-۲ شمارش بوسیله رگرسیون:

این نوع از روش‌ها یک نگاهت بین ویژگی‌های سطح پایین و شمارش افراد در صحنه یاد می‌گیرند. Chan و Vasconcelos<sup>[۶]</sup> شمارش افراد را براساس روش‌های رگرسیون بیزی انجام داده‌اند که یک تابع نگاهت بین بردار ویژگی ورودی با ابعاد بالا و مقادیر عددی خروجی یاد گرفته می‌شود. روش‌های رگرسیون بیزی شامل

پارامتری با بهره‌گیری از ریسک بیزی بدست می‌آید. تحلیل توزیع غیر پارامتری به کمک ریسک بیزی این امکان را فراهم می‌آورد تا از انواع توابع ضرر موجود در حوزه‌ی حداقل ریسک استفاده گردد. همچنین با استفاده از حداقل ریسک پیشنهادی، مراکز مورد توجه افراد نیز بدست می‌آیند.

تعیین مورد توجه‌ترین مکان از نگاه مشتریان و تحلیل نتایج بدست‌آمده از ریسک پیشنهادی روی ویدئو برگرفته از یک فروشگاه فواید بسیاری دارد. از جمله: کمک به مدیران فروشگاه در تحلیل رفتار مصرف‌کننده، اندازه‌گیری ترافیک فروشگاه، کمک به بهبود طراحی فروشگاه، سنجش اثر بخشی تبلیغات، تبلیغات هدفمند، جانمایی مناسب فروشندگان برای پاسخ‌دهی بهتر به مشتریان و پیش‌بینی مکان‌هایی با چگالی بالا در آینده.

در شکل ۱، یک نمونه تصویر ذخیره شده از سایت اینسکام<sup>[۲]</sup> آمده است که جهت محاسبه پربازدیدترین مکان در فروشگاه نشان داده شده است.

به طور کلی می‌توان مراحل انجام کار در این مقاله را به صورت زیر بیان کرد:

۱. تشخیص افراد براساس توصیف‌گر HOG<sup>[۳]</sup> انجام شده است. بعلاوه با ترکیب این روش با روش‌های ساده تشخیص حرکت مانند تفریق پس‌زمینه<sup>۲</sup>، نتایج مطلوبی برای تشخیص افراد بدست آمده است. در شکل ۲ نتیجه حاصل از روش HOG نشان داده شده است. مشکل این روش این است که برای هر فرد موجود در تصویر چندین مستطیل دربرگیرنده<sup>۳</sup> به صورت همپوشان پیدا شده است که به خاطر ماهیت الگوریتم HOG است و برای دستیابی به نتیجه قابل قبول و درست می‌توان از روش‌هایی که در بخش ۲ گفته شده، استفاده کرد. یکی از این روش‌ها که در پیاده‌سازی نیز از آن استفاده شده است، روش NMS<sup>۴</sup> است.



شکل ۱: یک نمونه تصویر ذخیره شده از سایت اینسکام<sup>[۲]</sup>. هدف ما پاسخ به این سوال است که کدام قسمت از فروشگاه یا کدام قفسه بیشترین بازدید را داشته است.

<sup>1</sup>Histograms of Oriented Gradients

<sup>2</sup>Background subtraction

<sup>3</sup>bounding box

<sup>4</sup>Non-Maximum Suppression

<sup>5</sup>Kernel Density Estimation

<sup>6</sup>Region Of Interest (ROI)

افراد، آنها با یک روش تطبیقی با استفاده از ویژگی‌های ردیابی می‌شوند.

Maddalena و همکارانش [۱۲] با استفاده از چندین دوربین شمارش افراد را انجام داده‌اند تا شمارش افراد را در مقابل انسداد مقاوم کنند. روش ارائه شده بر اساس همپوشانی FOV<sup>۷</sup> دوربین‌ها ارائه شده است. آنها از SOM<sup>۸</sup> برای محاسبه نقشه‌های پیش‌زمینه‌ای که از زاویه‌های مختلف گرفته شده استفاده کرده‌اند و آنها را توسط افکنش<sup>۹</sup> هموگرافی روی سطح زمین ادغام کرده‌اند. شمارش افراد توسط طبقه‌بند باناظر و ردیابی افراد بدست آمده است.

Idrees و همکارانش [۱] با تشخیص انسان در جمعیت‌های انبوه، یک رویکرد جدیدی برای تشخیص انسان در جمعیت‌های انبوه پیشنهاد کرده‌اند، به طوری که بخش‌های جمعیت را محلی‌سازی کرده‌اند. روش تشخیص آنها از ویژگی‌های جمعیت اشتراکی بر اساس مشاهدات استفاده می‌کند به صورتی که هر فرد به همسایگی محلی خود شبیه است. به همین منظور از یک مدل فرم‌پذیر اجزا (DPM)<sup>۱۱</sup> استفاده کرده‌اند که مقیاس و اطمینان از تشخیص‌ها از یک تشخیص دهنده اصلی بدست آمده است، آنها مقیاس و اطمینان قبلی را با استفاده از میدان‌های مارکوف تصادفی استنباط کرده‌اند. به دلیل این که در محیط‌های شلوغ تشخیص کل بدن انسان امکان ندارد، پس تجزیه و تحلیل بر اساس اجزا باید انجام شود. آنها یک راه‌حلی برای تشخیص ترکیب اجزای بدن یک انسان ارائه داده‌اند که قادر خواهد بود افرادی که قسمت‌هایی از بدن آنها مسدود شده است را نیز تشخیص دهد. این راه‌حل اجازه می‌دهد که از چندین تشخیص دهنده استفاده شود. روش آنها فاصله بین روش‌های ابتکاری برای محیط‌های شلوغ و تجزیه و تحلیل افراد در محیط‌های خلوت را از بین می‌برد.

Dalal و Triggs [۳] از توصیف‌گرهای HOG برای تشخیص افراد استفاده کرده‌اند. برای هر تصویر گرادیان و جهت آن محاسبه می‌شود و با استفاده از طبقه‌بند SVM<sup>۱۲</sup> انسان تشخیص داده می‌شود. روش آنها فقط قادر است افرادی که ایستاده هستند را تشخیص دهد که البته اگر در مجموعه آموزشی افراد در وضعیت‌های مختلف نیز قرار داده شود، قادر خواهد بود که افراد نشسته را نیز تشخیص دهد. همانطور که در شکل ۲ (الف) مشاهده می‌شود این روش اطراف هر فرد چندین مستطیل دربرگیرنده قرار می‌دهد که این به ماهیت روش به کار گرفته شده توسط Dalal و Triggs [۳] برای تشخیص هر فرد برمی‌گردد. برای حل این مسئله Triggs پیشنهاد کرده است که از الگوریتم میانگین جابه‌جایی بعلاوه

دو روش رگرسیون پواسون (BPR)<sup>۱</sup> و گوسی (GPR)<sup>۲</sup> هستند. روش‌های BPR برای محیط‌های شلوغ بهتر عمل می‌کنند و روش‌های GPR برای محیط‌های خلوت بهتر عمل می‌کنند. همچنین دو چالش اصلی برای یادگیری یک تابع رگرسیون برای حل بعضی از مسائل وجود دارد: (۱) ویژگی‌های متناقض و ناقص (۲) داده‌های آموزشی تنک و نامتعادل.

Xu و همکارانش [۷] شمارش تعداد افراد را در یک محیط پیچیده با استفاده از روش مبتنی بر رگرسیون انجام داده‌اند. بدین منظور آنها قسمت‌های پیش‌زمینه در تصاویر را با استفاده از تقطیع مبتنی بر رنگ و خوشه‌بندی استخراج کرده‌اند و رگرسیون بر اساس چگالی و تعداد افراد در یک ناحیه ایجاد شده است.

## ۲-۱-۳ شمارش بوسیله تشخیص:

ایده اصلی این روش‌ها طراحی یک تشخیص دهنده برای تشخیص هر شخص است. برای تشخیص افراد، محاسبه اختلاف فریم‌ها یکی از ساده‌ترین روش‌ها است که پیاده‌سازی آن به آسانی انجام می‌شود. در این روش اختلاف بین فریم جاری و فریم مرجع<sup>۳</sup>، پیکسل به پیکسل یا بلوک به بلوک، انجام می‌شود. فریم مرجع به عنوان "تصویر پس‌زمینه" یا "مدل پس‌زمینه" یا "مدل محیط"<sup>۴</sup> شناخته می‌شود [۸]. ابتدا طبقه‌بند تعلق یک بلوک تصویر را به پس‌زمینه تعیین می‌کند. در مرحله بعدی وابسته به نتایج طبقه‌بندی بروزرسانی بلوک‌های متناسب با تصویر پس‌زمینه انجام می‌شود. این روش قادر است هر شیء متحرک دیگری را نیز در تصویر تشخیص دهد ولی قادر نیست موقعیت هر فرد را تشخیص دهد. یکی دیگر از روش‌های معمول برای تشخیص افراد تبدیل هاف چرخشی (CHT)<sup>۵</sup> است. این روش بر اساس تشخیص شکل‌های حلقوی از سر افراد با استفاده از دوربینی که روی سقف نصب شده است انجام می‌شود [۹]. این روش بر اساس اختلاف فریم و HOG، با توجه به اختلاف فاکتورهای مانند جابه‌جایی دوربین، روشنایی و انسداد، ارزیابی شده است. که البته این روش در محیط‌های خیلی شلوغ روش خوبی نیست.

García و همکارانش [۱۰] سیستمی جهت شمارش افراد ارائه کردند که پس از پیش پردازش توسط بانکی از فیلترهای حلقوی سر افراد تشخیص داده می‌شود و سپس ردیابی افراد توسط فیلتر کالمن انجام می‌شود.

Cai و همکارانش [۱۱] تشخیص افراد را بر اساس سر آنها و طبقه‌بند SEMB-LBP<sup>۶</sup> انجام داده‌اند. ایده سر افراد به این دلیل بوده است که در محیط‌های شلوغ اگر دوربین بر روی سقف نصب شده باشد تنها سر افراد قابل دیدن است. بعد از تشخیص سر

<sup>7</sup> Haar

<sup>8</sup>fields-of-view

<sup>9</sup>self-organizing maps

<sup>10</sup>Projection

<sup>11</sup>Deformable Parts Model

<sup>12</sup>Support Vector Machine

<sup>1</sup> Bayesian Poisson Regression

<sup>2</sup>Gaussian Process Regression

<sup>3</sup>Reference frame

<sup>4</sup>Environment model

<sup>5</sup>Circular Hough Transform

<sup>6</sup>Statistically Effective Multi-scale Block Local Binary Pattern



(ب)



(الف)

شکل ۲: نتیجه حاصل از توصیف گر HOG برای تشخیص افراد در تصاویر اینسکام. (الف): برای تشخیص افراد مشاهده شده در هر تصویر از توصیف گر HOG به علاوه یک روش ساده تشخیص حرکت (تفریق پس زمینه) استفاده شده است. (ب): برای بدست آوردن نتایج صحیح تر برای تشخیص هر فرد پس از روش تشخیص افراد با HOG از روش NMS استفاده شده است.

جابه جایی<sup>۲</sup> را پیشنهاد کردند. روش میانگین جابه جایی با مشتق گرفتن از تابع توزیع غیر پارامتری (KDE) بدست می آید.

بعد از آن cheng [۱۸] از میانگین جابه جایی به عنوان "به دنبال-مد" نام برد. این روش یکی از روش های قطعی برای خوشه بندی است که شامل روش شاخه و حد می شود و برای تقسیم بندی بهینه انجام شده است.

سپس Meer و Comaniciu [۱۹] از میانگین جابه جایی برای هموارساز، به عنوان "فیلترینگ میانگین جابه جایی" و در قطعه بندی تصویر<sup>۳</sup> به عنوان "قطعه بندی میانگین جابه جایی" استفاده کرده اند.

کاربردهای مختلفی از KDE وجود دارد. Al Boni و Gerber [۲۰] از KDE محلی برای شناسایی جرایم استفاده کرده اند. آنها برای مقابله با سه چالش KDE (پیچیدگی محاسبات، انتخاب h، انتخاب تابع هسته) از یادگیری محلی استفاده کرده اند. به طوری که کانولوشن یک پنجره ۳\*۳ از موقعیت محل مورد نظر با یک پنجره ۳\*۳ از هسته را بدست آورده اند و خروجی بدست آمده تخمین چگالی را نشان می دهد.

همچنین Akbarzadeh و همکارانش [۲۱] از KDE برای پیش بینی مسیر هدف در کاربردهای نظارتی استفاده کرده اند. آنها از موقعیت های گذشته هدف برای پیش بینی حرکت آینده هدف استفاده کرده اند و به طور کلی برای حرکت افراد پیاده ارزیابی شده است.

### ۳-۲ مقالات در زمینه شناسایی مکان مورد توجه

در زمینه شناسایی مکان مورد توجه در پردازش تصاویر کارهای زیادی انجام شده است. مقالات و پژوهش هایی که در مورد

لگاریتم مقیاس فعلی تصویر استفاده شود. البته این روش برای حل مسئله گفته شده دارای مشکلاتی است، از جمله: برای هر تصویر از پایگاه داده باید بطور جداگانه عرض باند تعیین شود و همچنین اگر افراد به هم خیلی نزدیک باشند یکی را به عنوان فرد تشخیص داده شده انتخاب می کند. به همین منظور برای حل این مساله می توان از اعمال NMS به مستطیل های دربرگیرنده استفاده کرد. البته در این روش هم باید پارامتر همپوشانی به صورت دستی مقدارهی شود و همچنین اگر افراد به هم خیلی نزدیک باشند به طوری که مستطیل ها با هم همپوشانی داشته باشند یکی را انتخاب می کند [۱۳].

Wang و همکارانش [۱۴] برای تشخیص افراد با استفاده از تصویر متحرک پیشینه<sup>۱</sup> که یک فیلتر غیر خطی استصحت را در چندین تشخیص دهنده افراد افزایش داده اند. آنها پیکسل های متحرک را با تجمع گرادیان زمانی آنها در یک MHI شناسایی می کنند.

Coniglio و همکارانش [۱۵] روش جالبی بر اساس پیشینه های شکل، HOG و چندین SVM ارائه کردند که پس زمینه را از مستطیل دربرگیرنده حذف می کند.

### ۲-۲ مقالات در زمینه توزیع غیر پارامتری

تخمین چگالی، تخمین یک تابع چگالی از داده های مشاهده شده است که دو روش پارامتری و غیر پارامتری برای تخمین چگالی وجود دارد. یک مثال برای توزیع پارامتری، توزیع نرمال با میانگین  $\mu$  و واریانس  $\sigma^2$  است. روش KDE یک نوع توزیع غیر پارامتری است، روش KDE برای اولین بار توسط Parzen [۱۶] ارائه شد. همچنین Fukunaga و Hostetler [۱۷] میانگین

<sup>۲</sup>Mean-Shift<sup>۳</sup>mode-seeking<sup>۴</sup>Image Segmentation<sup>۱</sup>Motion history image (MHI)

### ۳-۱ تشخیص افراد

در شکل ۴ مرحله تشخیص افراد به صورت نموداری نشان داده شده است. این بخش از دو فاز آموزش و آزمون تشکیل شده است. در فاز آموزش تعدادی تصاویر به عنوان مجموعه آموزشی به طبقه‌بند داده می‌شود. این مجموعه آموزشی شامل تصاویر مثبت و منفی است. تصاویر مثبت شامل افراد در حالت‌های مختلف و تصاویر منفی شامل پس‌زمینه و ... است. بر روی این مجموعه تصاویر پیش پردازش انجام می‌شود. سپس برای تمام این تصاویر توصیف‌گرهای HOG استخراج می‌شوند و این توصیف‌گرها برای تصاویر مثبت و منفی با استفاده از طبقه‌بند SVM آموزش داده می‌شوند [۳].

در مرحله آزمون برای هر تصویر ورودی از پایگاه داده موردنظر، توصیف‌گر HOG استخراج می‌شود و با طبقه‌بندی، افراد موجود در تصویر تشخیص داده می‌شوند. در شکل ۲ (الف) نتیجه حاصل از تشخیص افراد بر روی یکی از تصاویر پایگاه داده موردنظر نشان داده شده است.

### ۳-۲ خوشه‌بندی مکانی افراد

در بخش قبلی گفته شد که با استفاده از توصیف‌گر HOG افراد تشخیص داده می‌شوند و هر فرد تشخیص داده شده در تصویر به عنوان یک رخداد در نظر گرفته می‌شود. می‌توان برای حل مشکل مستطیل‌های دربرگیرنده همپوشان برای هر فرد از روش NMS استفاده کرد.

در اینجا منظور از خوشه‌بندی مکانی افراد، تشخیص افراد در تمام تصاویر مجموعه داده است به طوری که موقعیت هر فرد در تصویر مشخص می‌شود و در شکل ۳ مرحله دوم نشان داده شده است. در ادامه به تعریف رخداد و حالت‌های مختلف آن می‌پردازیم.

ROI کار کرده‌اند، به دو دسته کلی تقسیم می‌شوند، (۱) با ناظر<sup>۱</sup> (۲) بدون ناظر<sup>۲</sup>.

در حالت با ناظر که بیشتر بر روی یک شیء خاص متمرکز می‌شود، ناحیه مورد توجه را کاربر تعیین می‌کند، بدین صورت که با در نظرگرفتن یک پوشش<sup>۳</sup> از شیء مورد نظر یا بخشی از شیء مورد نظر (به کاربرد روش بستگی دارد) ناحیه مورد توجه بدست می‌آید. در حالت بدون ناظر، در بیشتر کارهای انجام شده ناحیه مورد توجه برای یک قسمتی از تصویر بدست می‌آید که ممکن است شامل همه اشیاء موجود در تصویر باشد و یا شامل هیچ شیء ای نباشد. نکته مهم در این حالت این است که باید ناحیه مورد توجه کاملاً خودکار بدست آید. کاربردهای مختلفی برای تشخیص ROI وجود دارد و یکی از بیشترین کاربرد آن در تصاویر پزشکی است، به طوری که ناحیه موردنظر پزشک برای درمان یا عمل جراحی یا موارد دیگر از بقیه قسمت‌های بدن متمایز می‌شود. Brinkmann [۲۲] از ROI به عنوان یک روش برای کاهش محاسبات استفاده کرده است و بدین صورت تعریف می‌کند: یک ناحیه مورد توجه یا ROI، یک مربع یا مستطیلی است که توسط کاربر مشخص شده است که با استفاده از آن محاسبات محدود می‌شود.

Caron و همکارانش [۲۳] روش جدیدی برای تشخیص ROI در تصاویر سیاه و سفید دیجیتال، با استفاده از قوانین توان ارائه کرده‌اند. از دو قانون Zipf یا قانون معکوس Zipf برای تشخیص ROI استفاده شده است.

Cheng و Buckles [۲۴] یک روشی برای تشخیص ROI در تصاویر ویدئویی ارائه کرده‌اند به طوری که به دنبال ناحیه‌ای هستند که مقدار تابع چگالی احتمال<sup>۴</sup> آن پایین باشد. آنها فعالیت را در هر مکانی بر اساس ساختار سه بعدی تنسور و با استفاده از توزیع احتمال فاصله‌ای بین ساختار تنسورها مدل کرده‌اند. توزیع با استفاده از توزیع غیر پارامتری KDE تخمین زده شده است.

### ۳ مفاهیم اولیه

در این بخش ابتدا اجزاء سیستم پیشنهادی ارائه شده است، سپس مباحث تئوری و تحلیلی بیان می‌گردند. اجزاء سیستم پیشنهادی از چهار فاز اصلی تشکیل شده است، در شکل ۳ مراحل اصلی کار نشان داده شده است، شامل: الف) تشخیص افراد توسط توصیف‌گر HOG (ب) خوشه‌بندی مکانی افراد (ج) خوشه‌بندی مکانی-زمانی افراد با استفاده از توزیع غیر پارامتری KDE (د) پربازدیدترین مکان با استفاده از تخمین چگالی هسته مبتنی بر ریسک بیزی، که در ادامه به آن می‌پردازیم. همچنین مباحث تئوری لازم برای محاسبه عرض باند بهینه در انتها بیان می‌شود.

<sup>1</sup>Supervised

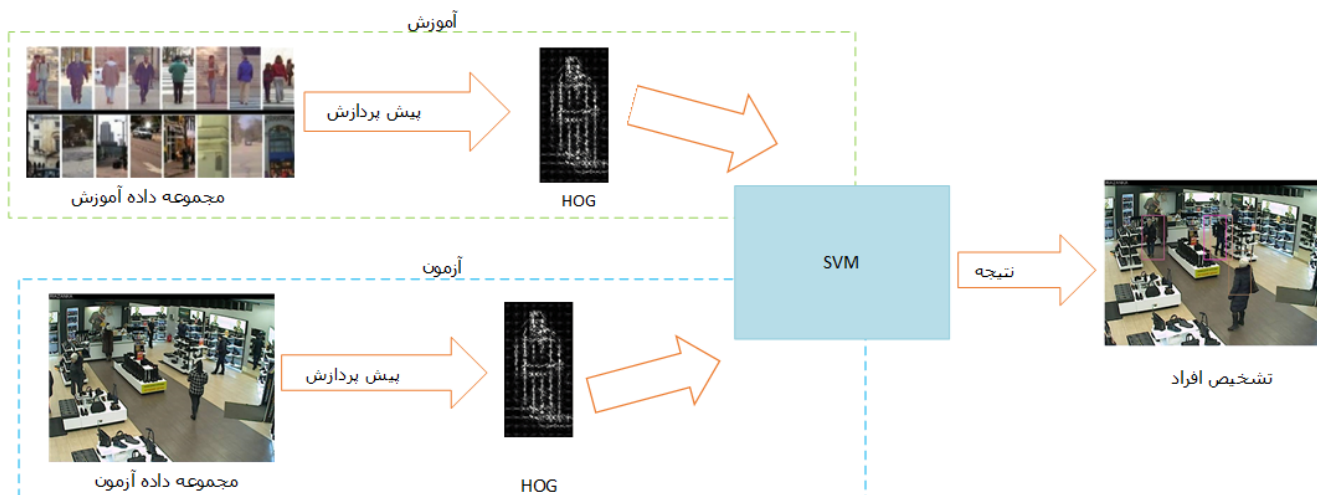
<sup>2</sup>Unsupervised

<sup>3</sup>Mask

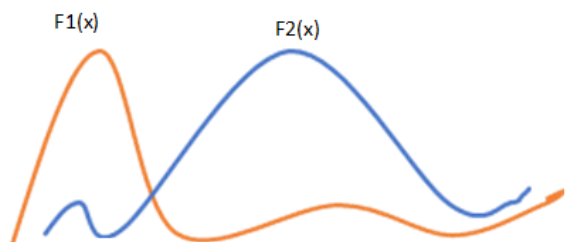
<sup>4</sup>Probability Density Function (PDF)



شکل ۳: فرآیند انجام کار. در این کار از تصاویر دوربین های نظارتی که در فروشگاه های مختلف تعبیه شده است، و به صورت برخط در سایت اینسکام [۲] بارگذاری شده اند، استفاده می شود. مرحله اول: پس از عملیات پیش پردازش روی تصاویر ورودی با استفاده از روش های تشخیص حرکت (تفریق پس زمینه) و توصیف گر HOG افراد موجود در هر تصویر تشخیص داده شده اند. مرحله دوم: برای تمام تصاویر مجموعه داده مورد نظر، افراد تشخیص داده شده اند. مرحله سوم: با توجه به موقعیت مکانی تمام افرادی که در مدت زمان محدود به فروشگاه رفت و آمد داشته اند، تابع توزیع غیر پارامتری KDE بدست آمده است. مرحله چهارم: با توجه به تابع توزیعی که در مرحله سوم نشان داده شده است، قله های تابع توزیع محاسبه شده است که با داشتن این قله ها پر بازدیدترین مکان ها در فروشگاه مورد نظر بدست آمده است.



شکل ۴: در این شکل مرحله آموزش و آزمون طبقه‌بند SVM نشان داده شده است [۳]. مرحله آموزش (قسمت بالا): برای تصاویر مثبت و منفی پس از اعمال پیش پردازش، توصیف‌گر HOG با استفاده از طبقه‌بند SVM یاد گرفته شده‌اند. مرحله آزمون: تصویر ورودی پس از پیش پردازش و بدست آوردن توصیف‌گر HOG به طبقه‌بند داده می‌شود و افراد از سایر اشیاء موجود در تصویر تشخیص داده می‌شوند.



شکل ۵: تاثیر رخداد  $f_2$  در  $f_1$  با توجه به میزان توانایی  $f_2$  در  $x$  محاسبه می‌شود. که میزان شباهت  $f_1$  و  $f_2$  را هم نشان می‌دهد.

### ۲-۲-۳ نحوه‌ی اجماع رخدادها

اجماع ترکیبی: تاثیر چند رخداد در  $x$  را می‌توان با اجماع چند رخداد در نظر گرفت:

$$f(x) = \sum_{i=1}^n \alpha_i f_i(x)$$

چنانچه  $f_i(x)$  از مجموعه متعامد اختیار شده باشد می‌توان ضرایب ترکیب را چنین نوشت:

$$\alpha_i = \frac{\langle f(x), f_i(x) \rangle}{\langle f_i(x), f_i(x) \rangle}$$

اجماع مجاورتی:

$$f(x) = \langle f_1(x), f_{\Omega_1}(x) \rangle f_1(x) + \dots + \langle f_n(x), f_{\Omega_n}(x) \rangle f_n(x)$$

در رابطه بالا  $f_{\Omega_i}(x)$  نزدیکترین رخداد به  $f_i(x)$  است. که می‌توان نوشت:

$$f(x) = \sum_{i=1}^n E_{f_{\Omega_i}} \{f_i(x)\} f_i(x)$$

یعنی اثر نزدیکترین همسایه رخداد  $f_i$ ، یعنی  $f_{\Omega_i}$  برای میزان اثر بخشی رخداد  $f_i(x)$  در ساخت  $f(x)$  است.

### ۳-۲-۱ رخداد

تعریف ۱: رخداد مستقل یعنی بروز متغیر تصادفی در یک موقعیت خاص.  $f$  مجموعه یا تابعی از رخداد  $x$  ( $f(x)$ ) است. در این مقاله رخداد مستقل مدنظر است [۲۵].

تعریف ۲: رخداد وابسته یعنی بروز یک رخداد تصادفی در یک حالت خاصی متأثر از متغیر تصادفی رخدادی دیگر.  $f_1$  مجموعه رخدادی است که در حالت  $x$  دیده شده‌است و از مجموعه  $f_2$  ( $f_1(x), f_2$ ) متأثر است.

برخی از حالات رخداد وابسته را که می‌توان در نظر گرفت: حالت (۱) ترکیب خطی دو رخداد. معادل علیت رخداد  $f_2$  بر  $f_1$  است.

$$f_1(x, f_2) = \alpha f_1(x) + (1 - \alpha) f_2(x)$$

که  $f_1$  و  $f_2$  رخدادهای در حالت  $x$  هستند.

حالت (۲)  $f_2(x)$  تاثیر ضریبی دارد یعنی عدم رخداد  $f_2$  در  $x$  مثل عدم رخداد  $f_1$  در  $x$  و متأثر از  $f_2$  است. بعبارتی در صورتی  $f_1$  در  $x$  رخ می‌دهد که  $f_2$  هم در  $x$  رخ داده باشد.

$$f_1(x, f_2) = f_1(x) f_2(x)$$

حالت (۳) علت  $f_1$  است و نقش تابع توزیع  $x$  را دارد.

$$f_1(x, f_2) = E_{f_2} \{f_1(x)\} = \int f_1(x) f_2(x) dx = \langle f_1(x), f_2(x) \rangle$$

اگر رخداد  $f_1$  کاملاً از  $f_2$  متأثر باشد، رابطه بالا مقدار زیادی تولید می‌کند. همانطور که در مثال شکل ۵ آمده، تاثیر مولد  $f_2$  بر  $f_1$  نشان داده شده‌است. در ادامه نحوه‌ی تعامل رخدادهای بیان شده‌است.

<sup>1</sup>Event

$$q_i = -p_i$$

در مرحله ۲، مینیمم‌سازی روی  $x$  انجام می‌شود:

$$\min_x f(x) = \min_x \sum_{i=1}^n \alpha_i q_i \frac{\|x - x_i\|^2}{h} \quad (3)$$

$$x = \frac{\sum_{i=1}^n \alpha_i p_i x_i}{\sum_{i=1}^n \alpha_i p_i} \quad (4)$$

رابطه (۴) همان رابطه‌ای است که در الگوریتم میانگین

جابه‌جایی استفاده می‌شود [۱۷] و می‌توان نوشت:

$$f_i(x) = K_1(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{h}\right) \quad (5)$$

حال در رابطه (۳) تغییراتی می‌دهیم و به سمت رابطه (۲)

حرکت می‌کنیم:

$$\min_x f(x) = \min_x \sum_{i=1}^n q_i \frac{\|x - x_i\|^2}{h} \quad (6)$$

که  $x_i$  ها نمونه‌های مسئله هستند. این رابطه شبیه به ریسک است:

$$R(x) = \int_x q(x_i) l(x, x_i) dx_i = E\{l(x, x_i)\} \\ = \sum_{i=1}^n q(x_i) l(x, x_i) \quad (7)$$

که اگر از  $x_i$  نمونه‌برداری شود و نمونه‌های یادگیر جایگزین شوند و تابع ضرر  $l_1(x, x_i) = \frac{\|x - x_i\|^2}{h}$  باشد، رابطه (۶) نتیجه می‌شود.

در شکل ۶ (الف) مفهوم تابع ضرر پیشنهادی رابطه (۸) بیان شده است.

$$l_2(x, x_i) = \frac{\|x - x_i\|^2}{h} \exp\left(-\frac{\|x - x_i\|^2}{h}\right) \quad (8)$$

به عبارتی در KDE داده‌هایی که پرت هستند از محاسبه تابع چگالی کنار گذاشته می‌شوند. نتیجه می‌شود:

$$\min_x f(x) = \min_x \sum_{i=1}^n q_i l_2(x, x_i) \quad (9)$$

$$\min_x f(x) \\ = \min_x \sum_{i=1}^n q_i \frac{\|x - x_i\|^2}{h} \exp\left(-\frac{\|x - x_i\|^2}{h}\right) \quad (10)$$

حال از رابطه (۱۰) نسبت به  $x$  مشتق گرفته می‌شود و بدست می‌آید:

$$x = \frac{\sum x_i q_i \exp\left(-\frac{\|x - x_i\|^2}{h}\right) \left(1 - \frac{\|x - x_i\|^2}{h}\right)}{\sum q_i \exp\left(-\frac{\|x - x_i\|^2}{h}\right) \left(1 - \frac{\|x - x_i\|^2}{h}\right)} \quad (11)$$

این رابطه ترکیب همه نمونه‌های یادگیر را برای محاسبه  $x$  نشان می‌دهد.

$$l_3(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{h}\right) \left(1 - \frac{\|x - x_i\|^2}{h}\right) \quad (12)$$

### ۳-۳ خوشه‌بندی مکانی- زمانی افراد

با استفاده از توزیع غیرپارامتری KDE، افرادی که در تمام فریم‌ها در طول زمان تشخیص داده شده‌اند، خوشه‌بندی می‌شوند و به صورت نقشه‌های چگالی جمعیت بر روی یکی از تصاویر مجموعه داده موردنظر نمایش داده می‌شوند. در شکل ۳ مرحله سوم نمایش سه بعدی از خوشه‌بندی مکانی- زمانی افراد نشان داده شده است.

### ۳-۴ روش پیشنهادی- بدست آوردن مورد توجه‌ترین مکان با استفاده از تخمین چگالی هسته مبتنی بر

#### ریسک بیزی

الگوریتم پیشنهادی در جدول ۱ آورده شده است و مراحل کار به صورت گام به گام بر روی داده‌های واقعی (تصاویر ویدئویی) بیان شده است.

برای سهولت در نوشتن روابط ریاضی تابع توزیع در تمام روابط به صورت یک بعدی در نظر گرفته شده است.

اگر تابع توزیع  $f(x)$  را به صورت اجماع رخدادهای در نظر بگیریم، در این صورت  $f(x) = \sum_{i=1}^n \alpha_i f_i(x)$  بیشینه احتمال رخداد در  $x$  ناشی از  $n$  رخداد  $f_i$  چنین محاسبه می‌شود [۱۶]:

$$\max_x f(x) = \max_x \sum_{i=1}^n \alpha_i f_i(x) \quad (1)$$

چنانچه رخداد در  $x$  را با تابع مثبت گوسی نشان دهیم، به صورت  $f_i(x) = \exp\left(-\frac{\|x - x_i\|^2}{h}\right)$  می‌شود که در  $x_i$  محتمل‌ترین حالت رخداد است و در  $x$  های دور از  $x_i$  احتمال رخداد کاهش می‌یابد. حال محتمل‌ترین محل وقوع رخداد به صورت زیر محاسبه می‌شود:

$$\max_x f(x) = \max_x \sum_{i=1}^n \alpha_i \exp\left(-\frac{\|x - x_i\|^2}{h}\right) \quad (2)$$

جهت محاسبه بیشینه، مزدوج تابع  $f_i(x)$  قرار داده می‌شود و از بهینه‌سازی HQ [۲۶] استفاده شده است:

$$\max_x f(x) \\ = \max_x \sum_{i=1}^n \alpha_i \sup_{p_i < 0} \left\{ \frac{\|x - x_i\|^2}{h} p_i - \varphi(p_i) \right\}$$

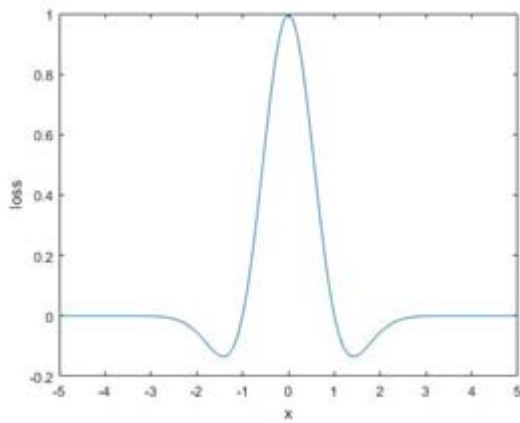
که در این رابطه  $\varphi(p_i)$  تابع مزدوج  $\exp\left(-\frac{\|x - x_i\|^2}{h}\right)$  است.

$$\max_x f(x) \\ = \max_{x, p_i < 0} \sum_{i=1}^n \alpha_i \left( -\frac{\|x - x_i\|^2}{h} p_i - \varphi(p_i) \right)$$

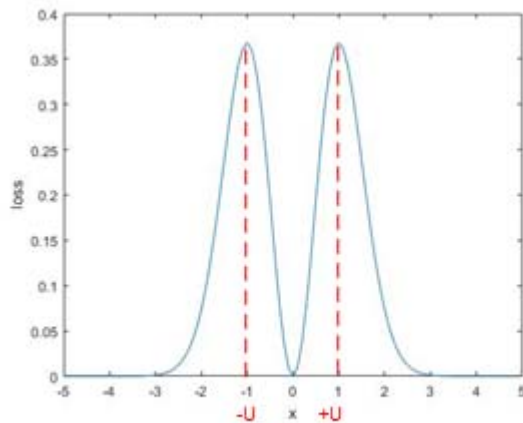
در مرحله ۱، بیشینه‌سازی روی  $p_i$  انجام می‌شود:

$$p_i = -\exp\left(-\frac{\|x - x_i\|^2}{h}\right)$$





(ب)



(الف)

شکل ۶: تابع ضرر. (الف): این شکل بیان می‌کند که اگر  $x = 0$  باشد ضرری بوجود نمی‌آید و در  $x = |U|$  بیشترین ضرر ایجاد می‌شود و هر چه  $x_i$  بزرگتر از  $|U|$  می‌شود، ضرر کمتر خواهد شد، این موضوع برای حذف داده‌های پرت است. (ب): این شکل بیان می‌کند که اگر  $x$  حول  $x_i$  باشد به طور کامل لحاظ می‌شود ولی همین که دور شود وزن آن منفی می‌شود.

می‌توان نوشت:

$$\begin{aligned} MISE(\hat{f}) &= \int E[\hat{f}(x) - f(x)]^2 dx \\ &= \int MSE_x(\hat{f}) dx \\ &= \int [E[\hat{f}(x)] - f(x)]^2 dx \\ &\quad + \int var \hat{f}(x) dx \end{aligned} \quad (17)$$

در رابطه (۱۱)،  $q_i$  نمونه‌هایی را که حول  $x$  هست را نگه می‌دارد و  $l_3(x, x_i)$  وزن هر نمونه است. در شکل ۶ (ب) مفهوم تابع ضرر در رابطه (۱۲) بیان شده است.

با توجه به رابطه (۵) و (۱۰) می‌توان نوشت:

$$K_2(x, x_i) = \exp^2\left(-\frac{\|x - x_i\|^2}{h}\right) \quad (13)$$

مورد توجه‌ترین مکان‌ها با استفاده از رابطه (۱۱) به صورت تکراری بدست می‌آید.

### ۳-۵ محاسبه عرض باند بهینه

در روابط نوشته شده در بخش قبل پارامتر  $h$  نشان دهنده عرض باند است و در مقدار تابع توزیع موثر است، به همین دلیل باید مقداری بهینه برای آن در نظر گرفت. بدین منظور ابتدا باید میانگین مربعات خطا<sup>۱</sup> و انتگرال میانگین مربعات خطا<sup>۲</sup> تعریف شوند [۲۷].

### ۳-۵-۲ محاسبه بایاس و واریانس

در رابطه (۱۵) داریم [۲۷]:

$$MSE_x(\hat{f}) = bias_h(x)^2 + var \hat{f}(x) \quad (18)$$

$$bias_h(x) = E[\hat{f}(x)] - f(x) \quad (19)$$

$$\begin{aligned} bias_h(x) &= \frac{1}{2} h^2 f''(x) k_2 \\ &\quad + \text{higher order terms in } h \end{aligned} \quad (20)$$

$$\int bias_h(x)^2 dx \approx \frac{1}{4} h^4 k_2 \int f''(x)^2 dx \quad (21)$$

$$var \hat{f}(x) \approx n^{-1} h^{-1} f(x) \int K(t)^2 dt \quad (22)$$

### ۳-۵-۱ میزان اختلاف: میانگین مربعات خطا و انتگرال

#### میانگین مربعات خطا

در اینجا اختلاف بین  $f$  و  $\hat{f}$  را بدست می‌آید. یک روش معمول برای انجام این کار میانگین مربعات خطا است که به صورت زیر تعریف می‌شود [۲۷]:

$$MSE_x(\hat{f}) = E[\hat{f}(x) - f(x)]^2 \quad (14)$$

با استفاده از خواص ابتدایی میانگین و واریانس داریم:

$$MSE_x(\hat{f}) = [E[\hat{f}(x)] - f(x)]^2 + var \hat{f}(x) \quad (15)$$

و همچنین برای انتگرال میانگین مربعات خطا داریم:

$$MISE(\hat{f}) = E \int [\hat{f}(x) - f(x)]^2 dx \quad (16)$$

<sup>1</sup>Mean Square Error(MSE)

<sup>2</sup>Mean Integrated Square Error(MISE)

## جدول ۱: الگوریتم پیشنهادی

ورودی: تصاویر ویدئویی

- گام ۰: ایجاد یک ماتریس  $mat$ :  $1 \times 2$  و  $N$ : تعداد فریم‌ها
- گام ۱: گرفتن یک تصویر به عنوان ورودی:  $frame = 1$
- گام ۲: اعمال الگوریتم HOG برای تشخیص افراد
- گام ۳: اگر فرد تشخیص داده شد موقعیت فرد به ماتریس اضافه می‌شود:  $mat(x,y)$
- گام ۴:  $frame = frame + 1$
- گام ۵: اگر  $frame \leq N$  بازگشت به گام ۱
- گام ۶: محاسبه تابع توزیع دو بعدی بر روی افراد تشخیص داده شده
- گام ۷: رسم تابع توزیع گام ۶
- گام ۸: محاسبه الگوریتم میانگین جابه‌جایی با مشتق گرفتن از تابع توزیع گام ۶
- گام ۹: محاسبه قله‌های تابع توزیع گام قبل

خروجی: استخراج تابع توزیع (نقشه چگالی جمعیت) و محاسبه قله‌های تابع توزیع

نوع مجموعه داده مختلف برای ارزیابی نتایج استفاده شده است. دوربین‌های نظارتی<sup>۱</sup> تعبیه شده در فروشگاه‌ها به صورت برخط تصاویر را در سایت اینسکام [۲] بارگذاری می‌کند. در این سایت می‌توان فروشگاه‌های مشابه دیگری در هر کشوری که به این سایت وصل هستند را نیز مشاهده کرد. پس از مشاهده این تصاویر، ۱۵۰۰ تصویر متوالی از هر فروشگاه مورد بررسی قرار گرفته است.

## ۴-۲ معیارهای ارزیابی

کارایی (عملکرد) هسته با رابطه ۲۶ اندازه‌گیری می‌شود [۲۷]. کارایی هسته پیشنهادی (رابطه ۲۴) ۸۵٪ است که مقدار قابل قبولی است. نتایج محاسبه بازده روی هسته‌های مختلف و مشهور نشان می‌دهد که تفاوت بین هسته‌های مختلف بر اساس MISE خیلی کم است. و کاملاً مطلوب و درست است که انتخاب هسته را بر مبنای ملاحظات دیگری قرار دهیم. (مانند درجه تمایز و محاسبات) [۲۷]

$$eff(K) = \frac{3}{5\sqrt{5}} \left\{ \int t^2 K(t) dt \right\}^{-\frac{1}{2}} \left\{ \int K(t)^2 dt \right\}^{-1}$$

برای مقایسه بهتر دو روش میانگین جابه‌جایی و روش پیشنهادی از معیار عملکرد از قبیل صحت که به صورت زیر (روابط ۲۷ و ۲۸) تعریف می‌شود، استفاده شده است:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$TP = \frac{N_C}{N_T}, FN = \frac{N_I}{N_T}$$

$$\int var \hat{f}(x) \approx n^{-1} h^{-1} \int K(t)^2 dt \quad (23)$$

در رابطه (۱۷) و برای بدست آوردن مقدار بهینه  $h$  از دید مینیمم سازی MISE داریم:

$$MISE(\hat{f}) = \frac{1}{4} h^4 k_2 \int f''(x)^2 dx + n^{-1} h^{-1} \int K(t)^2 dt \quad (24)$$

و با استفاده از قانون Silverman [۱۶] (لم 4a) داریم:

$$h_{opt} = k_2^{-\frac{2}{5}} \left\{ \int K(t)^2 dt \right\}^{\frac{1}{5}} \left\{ \int f''(x)^2 dx \right\}^{-\frac{1}{5}} n^{-\frac{1}{5}} \quad (25)$$

## ۴ نتایج

در این بخش نتایج ارزیابی و مجموعه داده‌های مختلف بررسی شده‌اند. در بخش ۴-۱ مجموعه داده استفاده شده معرفی شده است. در بخش ۴-۲ معیارهای ارزیابی آورده شده است. در بخش ۴-۳ شرایط پیاده‌سازی و در بخش ۴-۴ نتیجه ارزیابی روی مجموعه داده‌های گفته شده بیان شده است.

## ۴-۱ معرفی مجموعه داده

مجموعه داده‌ها به دو دسته مجموعه داده‌های مصنوعی یک بعدی و مجموعه داده‌های واقعی تقسیم می‌شود. در مجموعه داده مصنوعی یک بعدی از ۶ دسته مجموعه داده مختلف استفاده شده است. هر دسته شامل ۷۰۰ نمونه است که به صورت تصادفی ایجاد شده‌اند. سپس توزیع غیرپارامتری KDE برای این مجموعه داده‌ها بدست آمده‌اند و هر توزیع از سه نوع عرض باند مختلف (عرض باند بهینه، بیشتر از عرض باند بهینه، کمتر از عرض باند بهینه) ایجاد شده است. برای مجموعه داده واقعی از ۵

<sup>1</sup>Surveillance cameras

### ۳-۴ جزئیات پیاده‌سازی

در این مقاله، برای پیاده‌سازی روش پیشنهادی لازم است تعدادی از پارامترها مقداردهی شوند. پارامتر اول عرض باند بهینه است که در تئوری بر طبق رابطه (۲۵) است و در پیاده‌سازی با تقریب زدن مناسب بعضی از مقادیر از رابطه زیر محاسبه شده‌است: (۲۹)

$$h_i = \sigma_i \left\{ \frac{4}{(d+2)n} \right\}^{\frac{1}{d+4}} \quad i = 1, 2, \dots, d$$

جدول ۲ و جدول ۳، به صورت شهودی و تقریبی با توجه به تعداد نمونه‌ها بدست آمده است. همچنین در این جدول‌ها صحت برای هر حالت از میانگین صحت‌ها در هر حالت عرض باند مشخص شده، محاسبه شده است.

### ۴-۴ ارزیابی

شکل ۸ نشان داده شده‌است. برای مقایسه بهتر از سه حالت عرض باند استفاده شده است: در حالت عرض باند بهینه، کمتر از عرض باند بهینه و بیشتر از عرض باند بهینه. همانطور که مشاهده می‌شود

در حالت بیشتر از عرض باند بهینه روش پیشنهادی از روش میانگین جابه‌جایی در محاسبه قله‌های تابع توزیع بهتر عمل کرده‌است. ولی در حالت کمتر از عرض باند بهینه روش پیشنهادی نتوانست بعضی از قله‌ها را پیدا کند. همچنین عرض باند بهینه با روشی که در بخش ۳-۴ گفته شد، محاسبه شده است. در شکل ۹ تابع توزیع غیرپارامتری تخمین چگالی هسته (KDE) بر روی مجموعه تصاویر ورودی نشان داده شده‌است. هدف از این شکل مشخص شدن کار انجام شده روی داده‌های واقعی بوده‌است.

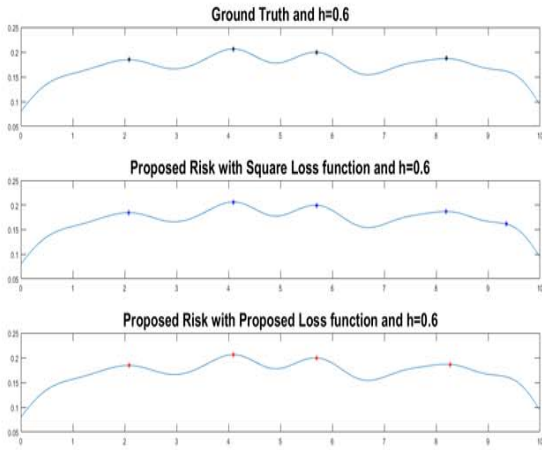
در شکل ۱۰ تابع توزیع بر روی تصاویر گرفته شده از مکان‌های مختلف بدست آمده‌است. در این شکل به طور میانگین بر روی ۱۵۰۰ تصویر از مجموعه داده بررسی شده است و قله‌های تابع توزیع محاسبه شده است. مجموعه داده‌های (الف) و (د) شکل ۱۰ در روز دیگری نیز بررسی شده‌اند و نتایج در شکل ۱۱ نشان داده شده است. این کار بدین دلیل بوده است که نشان داده شود آیا مکان‌های مورد توجه در هر دو زمان یکی بوده است یا نه. به عبارتی دنبال یک ارتباط معنایی بین توزیع‌های یک پایگاه داده در زمان‌های مختلف بودیم. همانطور که مشاهده می‌شود تقریباً در هر دو زمان مکان‌های مورد توجه مشابه بوده است.

جدول ۲: تابع ضررهای مختلف اعمال شده بر روی توزیع‌های یک بعدی (داده‌های مصنوعی) به همراه صحت آنها. در این جدول از ۳ نوع تابع توزیع مختلف برای مقایسه بین تابع ضرر مربعی و تابع ضرر پیشنهادی استفاده شده‌است.

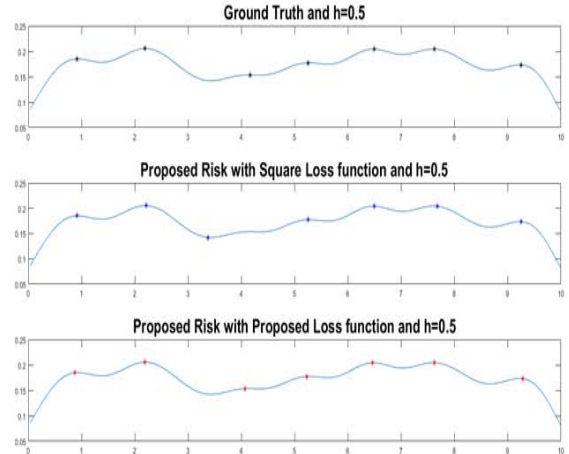
تابع ضرر	صحت (عرض باند بهینه)	صحت (بیشتر از عرض باند بهینه)	صحت (کمتر از عرض باند بهینه)
تابع ضرر مربعی	٪۱۰۰	٪۶۷	٪۸۸
تابع ضرر پیشنهادی	٪۱۰۰	٪۸۷.۵	٪۹۳

جدول ۳: تابع ضررهای مختلف اعمال شده بر روی توزیع‌های دو بعدی (داده‌های واقعی) به همراه صحت آنها. در این جدول از ۳ نوع تابع توزیع مختلف برای مقایسه بین تابع ضرر مربعی و تابع ضرر پیشنهادی استفاده شده‌است.

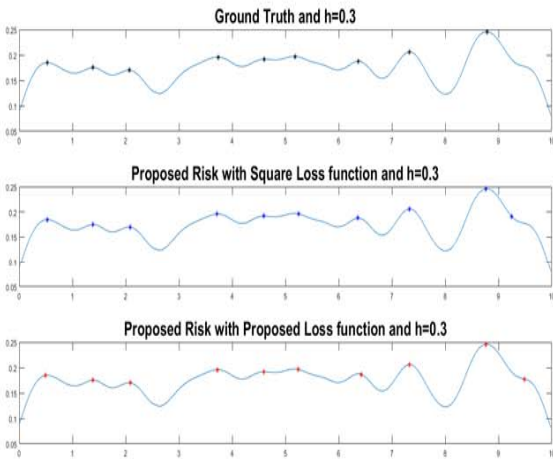
تابع ضرر	صحت (عرض باند بهینه)	صحت (بیشتر از عرض باند بهینه)	صحت (کمتر از عرض باند بهینه)
تابع ضرر مربعی	٪۱۰۰	٪۸۳	٪۵۷
تابع ضرر پیشنهادی	٪۱۰۰	٪۱۰۰	٪۷۱



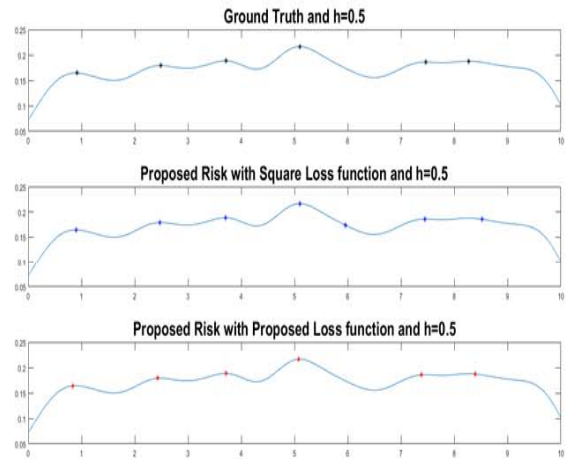
(ب) مجموعه داده دوم و عرض باند = 0.6



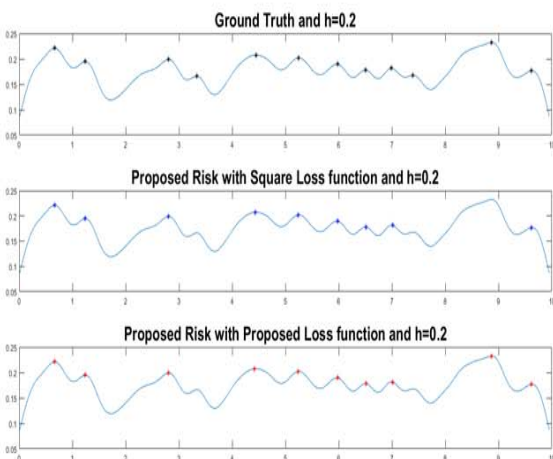
(الف) مجموعه داده اول و عرض باند = 0.5



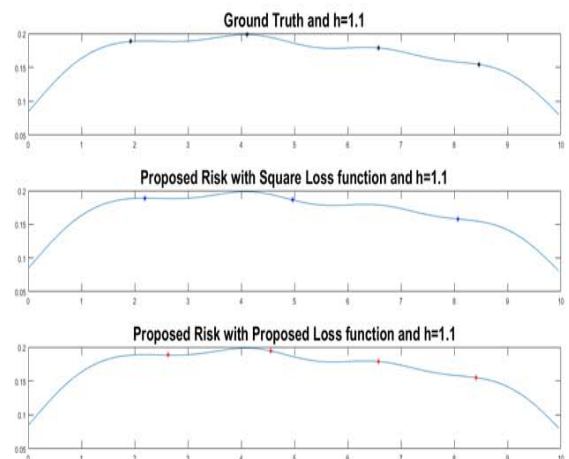
(د) مجموعه داده چهارم و عرض باند = 0.3



(ج) مجموعه داده سوم و عرض باند = 0.5

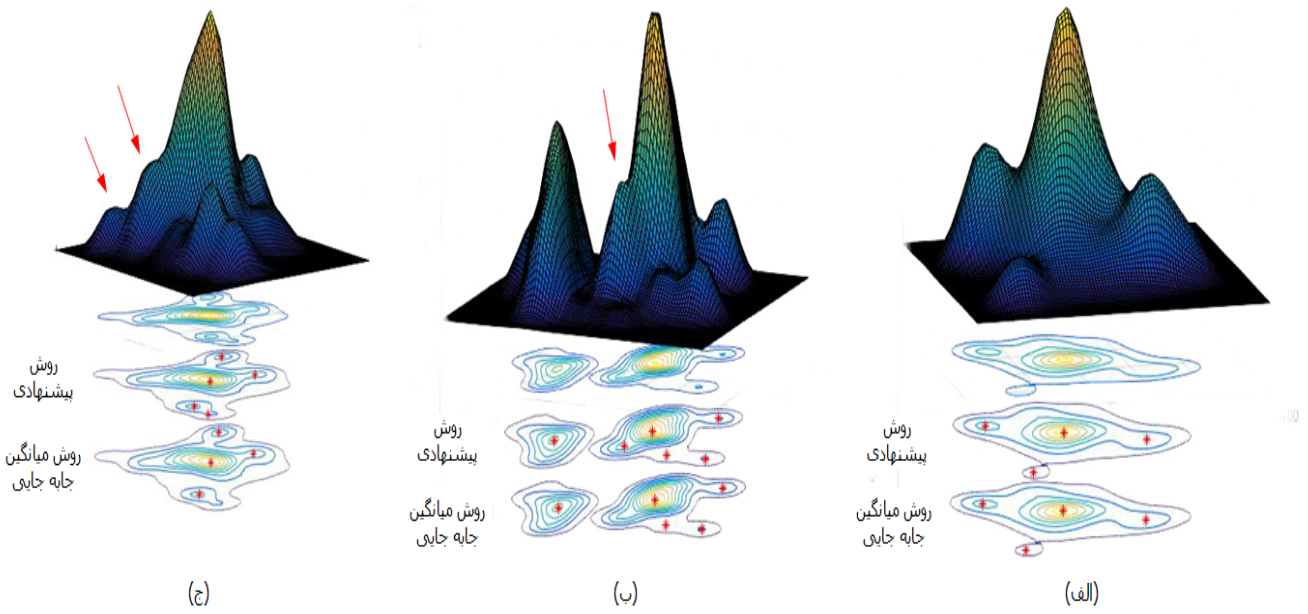


(و) مجموعه داده ششم و عرض باند = 0.2

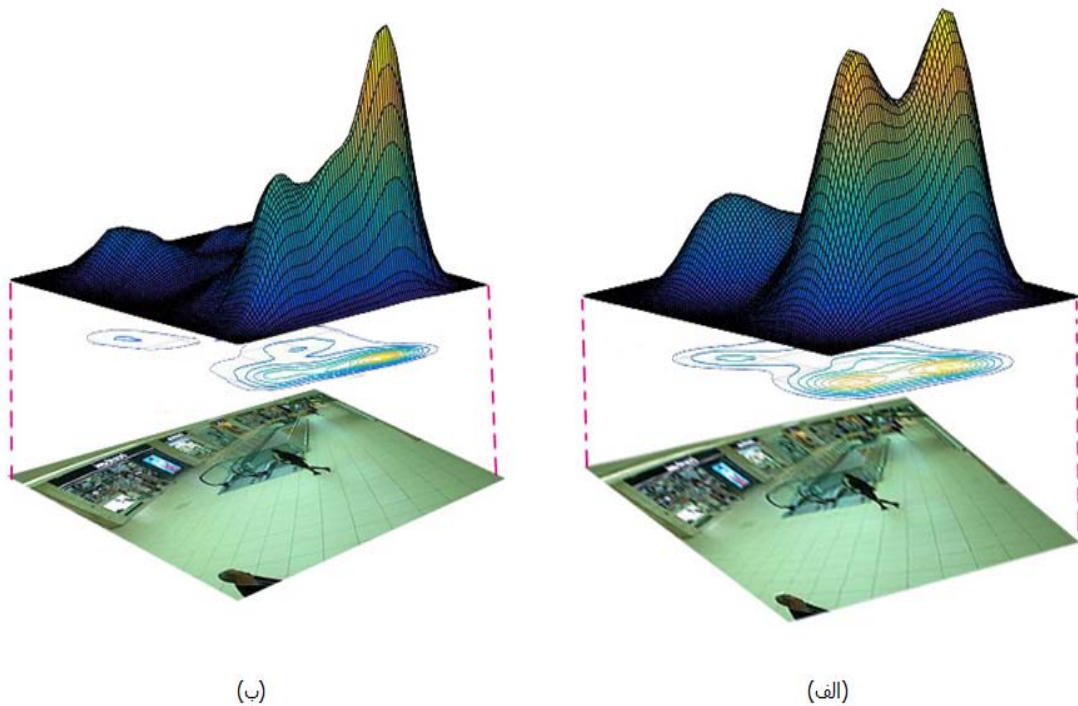


(ه) مجموعه داده پنجم و عرض باند = 1.1

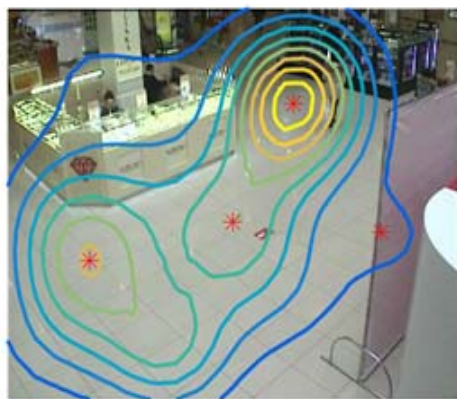
شکل ۷: تابع توزیع غیر پارامتری روی داده‌های مصنوعی یک بعدی و قله‌های آن. در ردیف بالای هر شکل قله‌های درست و در ردیف وسط، قله‌های بدست آمده با روش میانگین جابه‌جایی و تابع ضرر مربعی و در ردیف آخر قله‌های بدست آمده با روش پیشنهادی و تابع ضرر پیشنهادی نشان داده شده است.



شکل ۸: تابع توزیع دوبعدی و مشخص کردن قله‌های آن با دو روش میانگین جابه‌جایی و روش پیشنهادی. (الف) در حالت عرض باند بهینه: هر دو روش تمام قله‌ها را پیدا کرده‌اند. (ب) توزیعی یکنواخت با کمی نویز است، همانطور که ملاحظه می‌شود روش میانگین جابه‌جایی قله نشان داده شده را نمی‌تواند پیدا کند ولی در روش پیشنهادی این قله تشخیص داده می‌شود. (ج) در توزیع یکنواخت‌تر بیشتر شده است و باز هم عملکرد روش ارائه شده بهتر است ولی همانطور که ملاحظه می‌شود دو تا از قله‌ها را هیچ کدام از دو روش نتوانسته‌اند پیدا کنند.



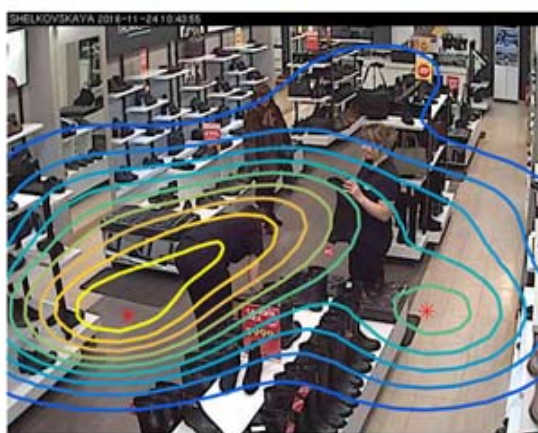
شکل ۹: تابع توزیع غیر پارامتری KDE بر روی ۱۵۰۰ تصویر متوالی از مجموعه داده مورد نظر. در این تصاویر پس از این که افراد بر اساس مکان آنها تشخیص داده شده‌اند، توزیع غیر پارامتری رسم شده است.



(ب)



(الف)



(د)

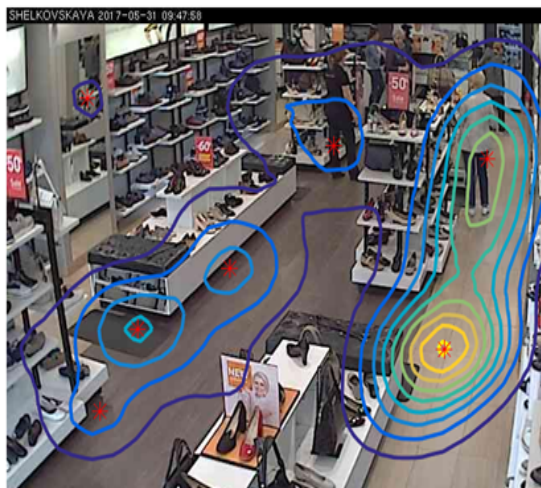


(ج)



(ه)

شکل ۱۰: نتایج روی داده‌های واقعی. (الف) تا (ه) مجموعه داده‌های دوبعدی واقعی ۱ تا ۵ را نشان می‌دهد که بعد از تشخیص افراد در ۱۵۰۰ فریم متوالی، تابع توزیع بر روی تصاویر رسم شده است و سپس با استفاده از روش پیشنهادی، قله‌های تابع توزیع بدست آمده‌است. این قله‌ها نشان می‌دهد که در کدام مکان‌ها افراد بیشتری رفت و آمد کرده‌اند و کدام مکان‌ها بیشتر مورد توجه افراد بوده است.



(ب)



(الف)

شکل ۱۱: نتایج روی داده‌های واقعی مجموعه داده ۱ و ۴ در روز دیگر

## مراجع

- [1] Idrees, H., K. Soomro, and M. Shah, *Detecting humans in dense crowds using locally-consistent scale prior and global occlusion reasoning*. IEEE transactions on pattern analysis and machine intelligence, 2015. 37(10): p. 1986–1998.
- [2] *insecam*. 2017; Available from: <https://www.insecam.org/>
- [3] Dalal, N. and B. Triggs. *Histograms of oriented gradients for human detection*. in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*. 2005. IEEE.
- [4] Gao, C., et al., *People counting based on head detection combining Adaboost and CNN in crowded surveillance environment*. Neurocomputing, 2016. 208: p. 108–116.
- [5] Topkaya, I.S., H. Erdogan, and F. Porikli. *Counting people by clustering person detector outputs*. in *Advanced Video and Signal Based Surveillance (AVSS), 2014 11th IEEE International Conference on*. 2014. IEEE.
- [6] Chan, A.B. and N. Vasconcelos, *Counting people with low-level features and Bayesian regression*. IEEE Transactions on Image Processing, 2012. 21(4): p. 2160–2177.
- [7] Xu, J., et al. *Efficient people counting with limited manual interferences*. in *Digital Image Computing: Techniques and Applications (DICTA), 2014 International Conference on*. 2014. IEEE.
- [8] Lin, H.-H., T.-L. Liu, and J.-H. Chuang, *Learning a scene background model via classification*. IEEE Transactions on Signal Processing, 2009. 57(5): p. 1641–1654.

## ۵ نتیجه‌گیری و کارهای آینده

در این مقاله، برای شناسایی مکان مورد توجه افراد از توزیع غیر پارامتری مبتنی بر حداقل ریسک بیزی استفاده شده است. پس از تشخیص افراد بر اساس توصیف گرهای HOG، برای افراد توزیع غیر پارامتری رسم شده است. سپس قله‌های تابع توزیع به عنوان مکان‌هایی که بیشترین بازدید را داشته‌اند (افراد بیشتری در آنجا رفت‌وآمد داشته‌اند) بدست آمده است.

نوآوری این مقاله در محاسبه قله‌های تابع توزیع با روش جدیدی است و همچنین در این روش تابع ضرر جدیدی تعریف شده است که با تابع ضرر مربعی که در روش میانگین جابه‌جایی استفاده شده، مقایسه شده است. نتایج بدست آمده نشان می‌دهد که روش پیشنهادی بهتر از روش میانگین جابه‌جایی به غیر از مواردی در حالت کاهش عرض باند بهینه، عمل می‌کند.

روش پیشنهادی در این مقاله در هر جایی که انسان رفت‌وآمد دارد استفاده می‌شود. در این مقاله از فروشگاه‌های خاصی که تصاویر آن در سایت اینسکام [۲] بارگذاری شده، استفاده شده است.

می‌توان برای گسترش کار در آینده از چند زاویه دیگر تغییراتی ایجاد کرد. در قسمت تشخیص افراد می‌توان از الگوریتم‌هایی با دقت بالاتر استفاده کرد. همچنین با توجه به این که ویدئوها از سایت اینسکام به صورت زنده دریافت می‌شود می‌توان نتایج را نیز به صورت زنده ارسال کرد، به طوری که در مواقعی که فروشگاه خیلی بزرگتر از فروشگاه موردنظر است، از تراکم جمعیت در هر قسمتی مطلع شد و اقدام لازم انجام شود. همچنین می‌توان در ریسک بیزی از تابع ضررهای مختلف دیگری نیز استفاده کرد.

- [24] Cheng, G. and B.P. Buckles, *A nonparametric approach to region-of-interest detection in wide-angle views*. Pattern Recognition Letters, 2014. **49**: p. 24-32.
- [25] Papoulis, A. and S.U. Pillai, *Probability, random variables, and stochastic processes*. 2002: Tata McGraw-Hill Education.
- [26] Boyd, S. and L. Vandenberghe, *Convex optimization*. 2004: Cambridge university press.
- [27] Silverman, B.W., *Density estimation for statistics and data analysis*. Vol. 26. 1986: CRC press.
- [9] Raghavachari, C., et al., *A Comparative Study of Vision Based Human Detection Techniques in People Counting Applications*. Procedia Computer Science, 2015. **58**: p. 461-469.
- [10] García, J., et al., *Directional people counter based on head tracking*. IEEE Transactions on Industrial Electronics, 2013. **60**(9): p. 3991-4000.
- [11] Cai, Z., et al. *Counting people in crowded scenes by video analyzing*. in *Industrial Electronics and Applications (ICIEA), 2014 IEEE 9th Conference on*. 2014. IEEE.
- [12] Maddalena, L., A. Petrosino, and F. Russo, *People counting by learning their appearance in a multi-view camera environment*. Pattern Recognition Letters, 2014. **36**: p. 1.134-25
- [13] Felzenszwalb, P., D. McAllester, and D. Ramanan. *A discriminatively trained, multiscale, deformable part model*. in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*. 2008. IEEE.
- [14] Wang, Y., et al., *Improving pedestrian detection using motion-guided filtering*. Pattern Recognition Letters, 2017. **96**: p. 106-112.
- [15] Coniglio, C., et al., *People silhouette extraction from people detection bounding boxes in images*. Pattern Recognition Letters, 2017. **93**: p. 182-191.
- [16] Parzen, E., *On estimation of a probability density function and mode*. The annals of mathematical statistics, 1962. **33**(3): p. 1065-1076.
- [17] Fukunaga, K. and L. Hostetler, *The estimation of the gradient of a density function, with applications in pattern recognition*. IEEE Transactions on information theory, 1975. **21**(1): p. 32-40.
- [18] Cheng, Y., *Mean shift, mode seeking, and clustering*. IEEE transactions on pattern analysis and machine intelligence, 1995. **17**(8): p. 790-799.
- [19] Comaniciu, D. and P. Meer, *Mean shift: A robust approach toward feature space analysis*. IEEE Transactions on pattern analysis and machine intelligence, 2002. **24**(5): p. 603-619.
- [20] Al Boni, M. and M.S. Gerber. *Automatic Optimization of Localized Kernel Density Estimation for Hotspot Policing*. in *Machine Learning and Applications (ICMLA), 2016 15th IEEE International Conference on*. 2016. IEEE.
- [21] Akbarzadeh, V., C. Gagné, and M. Parizeau. *Kernel density estimation for target trajectory prediction*. in *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*. 2015. IEEE.
- [22] Brinkmann, R., *The art and science of digital compositing: techniques for visual effects, animation and motion graphics*. 2008: Morgan Kaufmann.
- [23] Caron, Y., et al. *Power law dependencies to detect regions of interest*. in *Discrete Geometry for Computer Imagery*. 2003. Springer.



**مهناز رضوی** تحصیلات مقطع کارشناسی خود را در سال ۱۳۹۲ در رشته مهندسی کامپیوتر- نرم افزار به پایان رساند و مدرک کارشناسی ارشد خود را در رشته مهندسی کامپیوتر-هوش مصنوعی در سال ۱۳۹۷ از دانشگاه فردوسی مشهد دریافت نمود. زمینه‌های تحقیقاتی مورد علاقه او پردازش تصویر، تحلیل محتوای ویدئو، بینایی کامپیوتر و شناسایی الگو است.



**امیرحسین طاهری نیا** تحصیلات خود را در مقطع کارشناسی در رشته مهندسی کامپیوتر در سال ۱۳۸۳ در دانشگاه فردوسی مشهد و مقاطع کارشناسی ارشد و دکتری را در رشته هوش مصنوعی در سالهای ۱۳۸۵ و ۱۳۹۰ در دانشگاه صنعتی شریف به پایان برده‌اند. ایشان از سال ۱۳۹۱ استادیار گروه کامپیوتر دانشگاه فردوسی مشهد هستند و زمینه‌های تحقیقاتی مورد علاقه ایشان پردازش تصویر و کاربردهای آن، پنهان سازی اطلاعات و ردیابی تصاویر است.



**هادی صدوقی یزدی** مدرک کارشناسی خود را در سال ۱۳۷۳ در رشته مهندسی برق از دانشگاه فردوسی مشهد و مدرک کارشناسی ارشد و دکتری مهندسی برق را به ترتیب در سالهای ۱۳۷۵ و ۱۳۸۴ از دانشگاه تربیت مدرس دریافت نمود. زمینه‌های تحقیقاتی مورد علاقه ایشان شناسایی الگو و سیستم‌های هوشمند است و در حال حاضر استاد گروه مهندسی کامپیوتر دانشگاه فردوسی مشهد هستند.