ارزیابی خودکار ویژگی‌های ایمنی عابر پیاده از تصاویر نمای خیابان با استفاده از بخش‌بندی معنایی و یادگیری بازنمایی خودنظارتی: (مطالعه‌ای در مقیاس شهری در اصفهان)

مقاله علمی – پژوهشی

رسول ذبیحیان، گروه مهندسی عمران، واحد بین المللی کیش، دانشگاه آزاد اسلامی، کیش، ایران

*__محمدرضا احدی (نویسنده مسئول)__، پژوهشکده حمل و نقل، مرکز تحقیقات راه، مسکن و شهرسازی، تهران، ایران

علی نادران، دانشکده عمران، واحد علوم و تحقیقات، دانشگاه آزاد اسلامی، تهران، ایران

*__پست الکترونیکی نویسنده مسئول: m.ahadi@bhrc.ac.ir__

دریافت:۱۴۰۵/۰۳/۲- پذیرش:۱۴۰۵/۰۸/۰۵

صفحه ۱۶-۱

چکیده

ویژگی‌های مقیاس خرد محیط ساخته‌شده به‌طور قابل‌توجهی بر ایمنی عابر پیاده تأثیر می‌گذارد؛ با این حال، ممیزی‌های میدانی متعارف پرهزینه، ذهنی و در مقیاس شهری دشوار اجرا هستند. این پژوهش یک چارچوب بینایی رایانه‌ای ترکیبی را برای ارزیابی خودکار ایمنی عابر پیاده با استفاده از تصاویر نمای خیابان توسعه و اعتبارسنجی می‌کند. این چارچوب بخش‌بندی معنایی را برای استخراج ویژگی‌های صریح و قابل‌تفسیر محیطی با یادگیری بازنمایی خودنظارتی برای ثبت ویژگی‌های بصری نهفته- از جمله وضعیت روسازی، محصوریت فضایی، کیفیت مصالح و بی‌نظمی بصری — یکپارچه می‌سازد. تحلیل تجربی ۳۶۸٬۵۰۰ تصویر نمای خیابان گردآوری‌شده از ۹۲٬۱۲۵ موقعیت زمین‌مرجع در اصفهان، ایران را در برمی‌گیرد. یک مدل SegFormer-B4 تنظیم‌شده دقیق برای کمّی‌سازی هشت ویژگی مرتبط با ایمنی به کار رفت، در حالی که یک ترانسفورمر بینایی آموزش‌دیده با DINO بازنمایی‌های بصری زمینه‌ای تولید کرد. این جریان‌های ویژگی در یک مدل رگرسیون XGBoost برای ساخت شاخص خودکار ایمنی عابر پیاده و محیط (APSEI) ادغام شدند و از نمرات ممیزی میدانی متخصصان از ۵۰۰ موقعیت لایه‌بندی‌شده به‌عنوان مرجع حقیقت استفاده شد. اعتبار بیرونی در برابر ۳٬۲۱۲ تصادف عابر پیاده-وسیله نقلیه زمین‌کدشده از طریق رگرسیون پواسون وزن‌دار جغرافیایی ارزیابی شد. مدل بخش‌بندی به میانگین نسبت اشتراک بر اجتماع ۰/۷۶۲ و دقت کلی پیکسلی ۰/۹۱۴ دست یافت. در نمونه ممیزی نگه‌داشته‌شده، APSEI توافق عالی با ارزیابی‌های متخصصان نشان داد (ICC = ۰/۸۶؛ R^2 = ۰/۸۳) و اریبی سیستماتیک آن ناچیز بود. مدل فضایی تصادفات ۷۳/۲٪ از انحراف فراوانی تصادفات عابر پیاده را تبیین کرد، در مقایسه با ۶۱/۸٪ برای مدل مبتنی بر ویژگی‌های صریح. APSEI در ۸۸٪ از واحدهای فضایی معنادار و منفی بود و افزایش یک واحدی در شاخص با کاهش ۵۵٪ در فراوانی مورد انتظار تصادفات در کریدورهای حومه‌ای و شریانی پرسرعت همراه بود. یافته‌ها نشان می‌دهد ترکیب ویژگی‌های معنایی قابل‌تفسیر با بازنمایی‌های زمینه‌ای خودنظارتی، رویکردی مقیاس‌پذیر و مبتنی بر شواهد را برای تشخیص وضعیت ایمنی عابر پیاده، شناسایی نابرابری‌های فضایی و اولویت‌بندی مداخلات زیرساختی فراهم می‌کند.

واژه‌های کلیدی: ایمنی عابر پیاده، تصاویر نمای خیابان، بخش‌بندی معنایی، یادگیری خودنظارتی، DINO، SegFormer، رگرسیون پواسون وزن‌دار جغرافیایی، ارزیابی ایمنی شهری

۱- مقدمه

مرگ‌ومیر عابران پیاده یک بحران سلامت جهانی پایدار و ناعادلانه است که ۲۳٪ از تخمین ۱/۳۵ میلیون مرگ سالانه ناشی از ترافیک جاده‌ای را تشکیل می‌دهد (سازمان جهانی بهداشت، ۲۰۱۸). این بار به‌طور نامتناسب بر جمعیت‌های آسیب‌پذیر در کشورهای کم‌درآمد و با درآمد متوسط تحمیل می‌شود، جایی که شهرنشینی سریع و غالباً برنامه‌ریزی‌نشده از ارائه زیرساخت ایمن

و اختصاصی عابر پیاده پیشی گرفته است (Naci و همکاران، ۲۰۰۹؛ Sleet و همکاران، ۲۰۲۰. فراتر از تراژدی انسانی غیرقابل‌اندازه‌گیری، هزینه اقتصادی تصادفات عابر پیاده — شامل هزینه‌های پزشکی، از دست رفتن بهره‌وری و خسارت اموال — می‌تواند تا ۳٪ از تولید ناخالص داخلی یک کشور را مصرف کند، که نیازمند تغییری پارادایمی از تحلیل واکنشی تصادفات به برنامه‌ریزی ایمنی پیشگیرانه و داده‌محور است (Chen و همکاران، ۲۰۱۹).

نقش حیاتی محیط ساخته‌شده در مقیاس خرد

مجموعه روبه‌رشدی از شواهد اپیدمیولوژیک، ارتباط علّی بین ویژگی‌های مقیاس خرد محیط ساخته‌شده و پیامدهای ایمنی عابر پیاده را به‌طور محکم تثبیت کرده است. فراتحلیل بنیادین Ewing و Cervero (2010) بر «پنج D» محیط ساخته‌شده — تراکم، تنوع، طراحی، دسترسی به مقصد و فاصله تا حمل‌ونقل عمومی — به‌عنوان تعیین‌کننده‌های بنیادین رفتار سفر و به‌تبع آن، خطر تصادف و آسیب تأکید کرد. در سطح عابر پیاده، وجود، عرض و پیوستگی پیاده‌روها؛ طراحی و خط‌کشی گذرگاه‌های عابر پیاده؛ کیفیت روشنایی خیابان؛ و چگالی موانع ثابت و پارک حاشیه‌ای، نه صرفاً امکانات زیبایی‌شناختی، بلکه اقدامات متقابل ایمنی حیاتی هستند که مستقیماً بر رفتار راننده، مواجهه عابر پیاده و شدت تعارضات وسیله نقلیه–عابر پیاده تأثیر می‌گذارند (Stoker و همکاران، ۲۰۱۵؛ Zuniga-Teran و همکاران، ۲۰۱۷). بنابراین، درک و اندازه‌گیری عینی این ویژگی‌های محیطی ریزدانه، برای هر رویکرد مؤثر «چشم‌انداز صفر» یا «سیستم ایمن» بنیادین است.

از ممیزی‌های میدانی کاربرد–محور به ارزیابی‌های خودکار تصاویر نمای خیابان

به‌طور سنتی، استاندارد طلایی برای گردآوری داده‌های محیطی در مقیاس خرد، ممیزی میدانی حضوری بوده است که از ابزارهای اعتبارسنجی‌شده نظیر سیستم بازبینی محیط عابر پیاده (PERS) یا سیاهه Irvine-Minnesota استفاده می‌کند (Day و همکاران، ۲۰۰۶). این روش‌ها در حالی که داده‌های غنی و مبتنی بر حقیقت زمینی فراهم می‌کنند، دارای محدودیت‌هایی هستند: بسیار زمان‌بر و پرهزینه برای استقرار در مقیاس بزرگ، دارای مسائل قابلیت اعتماد بین‌ارزیاب به دلیل قضاوت انسانی ذهنی، و دارای چالش‌های لجستیکی و ایمنی قابل‌توجه برای ارزیابان (Brownson و همکاران، ۲۰۰۹). ظهور تصاویر نمای خیابان (SVI) با وضوح بالا، نظیر Google Street View و سرویس‌های مشابه، یک انقلاب روش‌شناختی را برانگیخته است. SVI به‌عنوان یک نماینده مجازی قدرتمند، مقیاس‌پذیر و مقرون‌به‌صرفه از محیط فیزیکی عمل می‌کند و ممیزی از راه دور چشم‌اندازهای وسیع شهری را امکان‌پذیر می‌سازد (Rundle و همکاران، ۲۰۱۱). هم‌زمان، بلوغ بینایی رایانه‌ای، به‌ویژه بخش‌بندی معنایی مبتنی بر یادگیری عمیق، استخراج خودکار در سطح پیکسل اشیاء فیزیکی از این تصاویر را با سرعت و مقیاس فضایی بی‌سابقه‌ای ممکن ساخته و به‌طور مؤثر گلوگاه مشاهده دستی را دور می‌زند (Biljecki & Ito, 2021).

شکاف پژوهشی: فراتر از شمارش پیکسل به‌سوی درک صحنه

با وجود این پیشرفت‌ها، شکاف معناداری بین برچسب‌گذاری خودکار صحنه‌های خیابانی و درک کل‌نگر از ایمنی عابر پیاده و کیفیت محیطی باقی مانده است. اکثریت مطالعات پیشگام در این حوزه بر مدل‌های بخش‌بندی معنایی برای کمّی‌سازی مساحت نسبی ویژگی‌های گسسته، مانند «شاخص دید سبز» یا «عامل دید آسمان» متکی بوده‌اند و این معیارها را به ایمنی ادراک‌شده یا قابلیت پیاده‌روی مرتبط ساخته‌اند (Naik و همکاران، ۲۰۱۴؛ Li و همکاران، ۲۰۱۸؛ Zhang و همکاران، ۲۰۱۸). با این حال، این رویکرد اساساً تقلیل‌گرایانه است. نخست، صرف شمارش پیکسل نسبت به چیدمان فضایی، بافت و وضعیت عناصر بی‌تفاوت است؛ یک پیاده‌رو ترک‌خورده و ناهموار و یک پیاده‌رو به‌خوبی نگهداری‌شده ممکن است برچسب معنایی یکسانی داشته باشند. دوم، این روش‌های صریح مبتنی بر پیکسل قادر به ثبت ویژگی‌های بصری نهفته پیچیده‌ای نیستند که عمیقاً بر ایمنی دنیای واقعی تأثیر می‌گذارند - نظیر «حس محصوریت» ناشی از نمای ساختمان‌ها، پیچیدگی ادراک‌شده یک صحنه، کیفیت مصالح روسازی یا سطح بی‌نظمی بصری-ویژگی‌هایی که در هیچ طبقه‌بندی معنایی به‌صراحت تعریف نشده‌اند (Yin & Wang, 2016؛ Aultman-Hall, 2016 & Harvey). در نهایت، پیوند تجربی مستقیم بین این شاخص‌های صرفاً بصری و فراوانی تصادفات عابر پیاده به‌طور عینی ثبت‌شده،

به‌ویژه در بافت کلان‌شهرهای جنوب جهانی، به‌طور انتقادی کم‌بررسی مانده و تأثیر ترجمه‌ای آن‌ها را برای برنامه‌ریزی ایمنی حمل‌ونقل محدود کرده است (Arellana و همکاران، ۲۰۲۰). این پژوهش مستقیماً به این شکاف می‌پردازد و یک چارچوب روش‌شناختی نوین پیشنهاد می‌کند که دقت ساختاری بخش‌بندی معنایی را با درک زمینه‌ای و ضمنی ارائه‌شده توسط یادگیری بازنمایی خودنظارتی مدرن (SSL) ترکیب می‌کند.

پرسش‌ها و اهداف پژوهش

برای پر کردن شکاف‌های دانشی شناسایی‌شده، این مقاله یک بررسی تجربی در مقیاس شهری در اصفهان، ایران انجام می‌دهد. این پژوهش توسط سه پرسش اصلی هدایت می‌شود:

ادغام روش‌شناختی: چگونه می‌توان یک چارچوب بینایی رایانه‌ای ترکیبی، با یکپارچه‌سازی بخش‌بندی معنایی پیشرفته با یادگیری بازنمایی خودنظارتی، را برای استخراج مجموعه‌ای کل‌نگر از ویژگی‌های ایمنی عابر پیاده از تصاویر نمای خیابان عملیاتی کرد؟ این مجموعه شامل عناصر فیزیکی صریح (مانند پیاده‌روها، گذرگاه‌ها، موانع، عرض مؤثر مسیر) و زمینه‌های بصری ضمنی (مانند محصوریت، کیفیت مصالح) است.

اعتبارسنجی تجربی با تصادفات: آیا شاخص خودکار جامع ایمنی عابر پیاده و محیط (APSEI) نشان‌دهنده ارتباط آماری معنادار و فضایی استوار با فراوانی تصادفات عابر پیاده–وسیله نقلیه ثبت‌شده رسمی در سراسر شهر است؟

حقیقت زمینی و مقیاس‌پذیری: چگونه ارزیابی APSEI از کیفیت محیطی و ایمنی با یک ممیزی میدانی سنتی انجام‌شده توسط متخصص مقایسه می‌شود و از این طریق دقت آن اعتبارسنجی شده و پتانسیل آن به‌عنوان یک ابزار پشتیبانی برنامه‌ریزی واقعاً مقیاس‌پذیر و عینی تثبیت می‌شود؟

۲- پیشینه تحقیق

۲-۱-شاخص‌های ایمنی عابر پیاده و ممیزی‌های محیطی: تحول و محدودیت‌ها

ارزیابی نظام‌مند محیط‌های عابر پیاده ریشه در طراحی شهری و بهداشت عمومی دارد و از مشاهده کیفی به ابزارهای ممیزی ساختاریافته تحول یافته است. کار بنیادین اولیه به‌دنبال عملیاتی‌سازی مفاهیم قابلیت پیاده‌روی و ایمنی از طریق ابزارهایی مانند سیستم بازبینی محیط عابر پیاده (PERS) و سیاهه Irvine-Minnesota بود که نیازمند ارزیابان آموزش‌دیده برای امتیازدهی به ده‌ها ویژگی در مقیاس خرد بودند (Day و همکاران، ۲۰۰۶ ؛ Millington و همکاران، ۲۰۰۹). هم‌زمان، شاخص‌های ترکیبی مانند Walk Score به دلیل دسترسی کلان‌مقیاس به امکانات محبوبیت یافتند، اما به دلیل غفلت کامل از طراحی مقیاس خرد، زیبایی‌شناسی و زیرساخت حیاتی ایمنی به‌شدت مورد انتقاد قرار گرفتند (Carr و همکاران، ۲۰۱۱؛ Ram, 2018 & Hall]). نقدهای بنیادی این روش‌های سنتی انسان–محور در سه نکته همگرا می‌شوند: آن‌ها با ذهنیت بین‌ارزیاب قابل‌توجهی همراه‌اند، هزینه‌بر و زمان‌برند، و در مقیاس شهری غیرقابل‌اجرا هستند.

۲-۲- برتری تصاویر نمای خیابان (SVI) در تحلیل‌های شهری و حمل‌ونقل

دسترسی به پایگاه‌های داده عظیم و زمین‌مرجع تصاویر نمای خیابان، به پیشگامی Google Street View، برای علوم شهری تحول‌آفرین بوده و به‌طور مؤثر یک دوقلوی دیجیتال از شبکه‌های جاده‌ای سیاره ایجاد کرده است. مرور جامع Biljecki و Ito (2021) نقش SVI را به‌عنوان یک منبع داده مکانی حیاتی رسمیت بخشید. در حمل‌ونقل، SVI از کاربرد اولیه خود به‌عنوان یک جعبه‌ابزار ممیزی مجازی (Rundle و همکاران، ۲۰۱۱) فراتر رفته و به ستون فقرات خطوط لوله یادگیری عمیق تبدیل شده است. مطالعات قدرت آن را در پیش‌بینی حجم عابر پیاده (Yin و همکاران، ۲۰۱۵)، نقشه‌برداری بی‌نظمی‌های فیزیکی شهری نظیر چاله‌ها و تابلوهای آسیب‌دیده (Ibrahim و همکاران، ۲۰۲۰) و استخراج ویژگی‌های هندسی دقیق حریم راه نشان داده‌اند.

۲-۳- بخش‌بندی معنایی برای استخراج ویژگی‌های صریح در تحلیل ایمنی

در خط مقدم این تلاش رمزگشایی، بخش‌بندی معنایی قرار دارد؛ یک وظیفه یادگیری عمیق که هر پیکسل در یک تصویر را به یک طبقه از پیش تعریف‌شده مانند «راه»، «پیاده‌رو»، «وسیله نقلیه» یا «پوشش گیاهی» طبقه‌بندی می‌کند. معماری‌هایی نظیر PSPNet (Zhao و همکاران، ۲۰۱۷) و DeepLabv3+ (Chen و همکاران، ۲۰۱۸) معیارهای عملکرد را در مجموعه‌داده‌های صحنه شهری پیچیده نظیر Cityscapes

تعیین کرده‌اند. کار بنیادین Naik و همکاران (۲۰۱٤) از رویکردی مبتنی بر پیکسل برای ایجاد «Streetscore» برای ایمنی ادراک‌شده استفاده کرد، در حالی که Li و همکاران (۲۰۱۸) یک «شاخص دید سبز» اصلاح‌شده را عملیاتی کردند. مطالعات بعدی مجموعه غنی‌تری از ویژگی‌های فیزیکی مرتبط با ایمنی را استخراج کرده‌اند (Aghaabbasi و همکاران، ۲۰۲۰؛ Koo و همکاران، ۲۰۲۲). با این حال، محدودیت حیاتی پابرجاست: این هستی‌شناسی سطح-پیکسل ذاتاً نسبت به زمینه کور است.

۲-٤-پارادایم یادگیری بازنمایی خودنظارتی (SSL) در بینایی رایانه

یک تغییر پارادایم در یادگیری بصری مستقیماً به این محدودیت‌ها می‌پردازد. یادگیری خودنظارتی (SSL) طبقه جدیدی از الگوریتم‌ها را تعریف می‌کند که بازنمایی‌های بصری غنی و همه‌منظوره را از پیکره‌های عظیم تصاویر بدون برچسب، با حل وظایف پیش‌متنی مشتق‌شده از ساختار خود داده، یاد می‌گیرند. چارچوب‌های محوری شامل SimCLR (Chen و همکاران، ۲۰۲۰)، Momentum Contrast (He و همکاران، ۲۰۲۰) و DINO (Caron و همکاران، ۲۰۲۱) هستند که نشان دادند ترانسفورمرهای بینایی آموزش‌دیده بدون هیچ برچسبی می‌توانند بازنمایی‌های فضایی و معنایی فوق‌العاده‌ای بیاموزند.

۲-۵-ترکیب و نوآوری پژوهش حاضر

ترکیب این جریان‌های ادبیات، شکافی روشن را آشکار می‌کند. از یک سو، پژوهش ایمنی عابر پیاده با موفقیت از بخش‌بندی معنایی برای کمّی‌سازی ساده و صریح ویژگی‌ها استفاده کرده است. از سوی دیگر، جامعه بینایی رایانه ابزارهای SSL توانمند برای درک عمیق و ضمنی صحنه را توسعه داده، اما ترجمه آن‌ها به یک مسئله واقعی ایمنی حمل‌ونقل تقریباً وجود ندارد. این مطالعه یکپارچه‌سازی نخستین از نوع خود است و یک چارچوب ترکیبی پیشنهاد می‌کند که «اسکلت» ساختاری قابل‌تفسیر از بخش‌بندی معنایی را با «گوشت» زمینه‌ای غنی ارائه‌شده توسط ویژگی‌های خودنظارتی آموزش‌دیده با DINO ترکیب می‌کند.

۳- منطقه مطالعه و داده‌ها

۳-۱-اصفهان به‌عنوان مطالعه موردی نماینده

۳-۱-۱- ویژگی‌های جغرافیایی، جمعیتی و شبکه

اصفهان، مرکز استان اصفهان، سومین شهر پرجمعیت ایران است با جمعیت کلان‌شهری بیش از ۲/۲ میلیون نفر که حدود ۵۵۱ کیلومتر مربع وسعت دارد. بافت شهری این شهر پالیمپسستی از دوران‌های برنامه‌ریزی متمایز است که محیطی منحصربه‌فرد و ناهمگن برای مطالعه خطر عابر پیاده فراهم می‌کند. این تنوع شامل: (۱) یک هسته قرون‌وسطایی متراکم و حفاظت‌شده که با خیابان‌های باریک و پرپیچ‌وخم و حجم بالای فعالیت عابر پیاده حول میدان نقش‌جهان مشخص می‌شود؛ (۲) مناطق مدرنیستی شبکه‌ای گسترده با خیابان‌های شریانی عریض که اغلب جریان وسایل نقلیه را در اولویت قرار می‌دهند (مانند خیابان چهارباغ)؛ و (۳) سکونتگاه‌های غیررسمی پیرامونی و شهرهای جدید که به‌سرعت در حال رشد و اغلب کم‌خدمت هستند.

۳-۱-۲- الگوهای فضایی تصادفات عابر پیاده-وسیله نقلیه

تحلیل اکتشافی مقدماتی داده‌های فضایی (ESDA) سوابق تصادفات ثبت‌شده رسمی، الگویی عمیقاً نگران‌کننده و از نظر فضایی ناهمگن از آسیب‌ها و مرگ‌ومیر عابران پیاده را آشکار می‌کند. با استفاده از دوره تجمیع پنج‌ساله (۲۰۱۷–۲۰۲۲)، برآورد چگالی کرنل (KDE) برای شناسایی نقاط داغ آماری معنادار انجام شد. تحلیل نشان می‌دهد تصادفات عابر پیاده به‌طور تصادفی توزیع نشده‌اند، بلکه خودهمبستگی فضایی قوی نشان می‌دهند (Global Moran's I = ۰/۳٤، p < ۰/۰۰۱) و در سه گونه غالب خوشه‌بندی می‌شوند:

نخست، نقاط داغ کریدورهای شریانی؛ دوم، هسته‌های مناطق تجاری؛ سوم، خوشه‌های پیرامونی حومه‌ای. نکته حیاتی آنکه، یک الحاق فضایی بین مکان‌های تصادف و نقاط نمونه‌برداری SVI آشکار می‌کند که بیش از ۷۲٪ تصادفات عابر پیاده در فاصله ۵۰ متری از یک قطعه خیابانی که تصویر آن گردآوری شده بود، رخ داده‌اند.

۳-۲- گردآوری و پیش‌پردازش تصاویر نمای خیابان (SVI)

۳-۲-۱- پروتکل نظام‌مند گردآوری تصویر مبتنی بر API

یک مجموعه‌داده SVI جامع در مقیاس شهری با پرس‌وجوی برنامه‌نویسی‌شده از یک API سرویس نقشه ایجاد شد. شبکه معابر کل منطقه کلان‌شهری اصفهان از OpenStreetMap (OSM) استخراج و به همه انواع راه‌های قابل‌دسترسی برای وسایل نقلیه فیلتر شد و یک شبکه پایه ۴٬۸۶۲ کیلومتری خط مرکزی تولید کرد. برای دستیابی به پوشش نظام‌مند مقیاس خرد، نقاط نمونه‌برداری به‌صورت دینامیک در امتداد هر قطعه خیابان با فواصل ثابت ۵۰ متر تولید شدند که به مجموعه هدف اولیه ۹۷٬۲۴۰ مختصات جغرافیایی منحصربه‌فرد انجامید. برای هر نقطه مختصات، چهار تصویر ثابت از طریق API درخواست شد که با جهت‌های دید اصلی (۰°، ۹۰°، ۱۸۰°، ۲۷۰°) مطابقت داشت. ابعاد تصویر روی ۱۹۲۰×۱۰۸۰ پیکسل تنظیم شد. پس از حذف خطاهای فراخوانی API، مجموعه‌داده نهایی شامل ۳۶۸٬۵۰۰ تصویر با وضوح بالا در ۹۲٬۱۲۵ نقطه زمین‌مرجع اعتبارسنجی‌شده بود که نرخ پوشش ۹۴/۷٪ از چارچوب نمونه‌برداری برنامه‌ریزی‌شده را نمایندگی می‌کند.

۳-۲-۲- خط لوله پیش‌پردازش چندمرحله‌ای تصویر

برای تضمین سازگاری هندسی و رادیومتریک برای خط لوله بینایی رایانه، یک پروتکل پیش‌پردازش سه‌مرحله‌ای روی تصاویر SVI خام اعمال شد. نخست، یک مرحله تصحیح هندسی به‌صورت دینامیک لوگو، فلش‌های ناوبری و سایر عناصر ثابت رابط کاربری را با استفاده از یک مدل تشخیص شیء از پیش آموزش‌دیده ماسک و حذف کرد. دوم، نرمال‌سازی رادیومتریک با استفاده از یک تکنیک تطبیق هیستوگرام در برابر یک تصویر مرجع انجام شد. سوم، برای اهداف تحلیلی، چهار تصویر جهت‌دار برای هر نقطه در یک نوار پانوراما به‌هم پیوسته و برای برخی تحلیل‌ها به میدان‌های دید همپوشان مجزا تقسیم شدند: مسیر عابر پیاده، سواره‌رو و چشم‌انداز مقابل.

۳-۳- گردآوری و پردازش زمینی داده‌های تصادفات عابر پیاده

متغیر وابسته برای تحلیل پیامد ایمنی از یک پایگاه داده دقیق و زمین‌کدشده تصادفات عابر پیاده–وسیله نقلیه مشتق شده است. داده‌ها از سیستم اطلاعات تصادفات متمرکز پلیس راهور اصفهان در بازه پنج‌ساله از ۲۱ مارس ۲۰۱۷ تا ۲۰ مارس ۲۰۲۲ به دست آمد. هر رکورد شامل مختصات فضایی، برچسب زمانی و فیلدهای شدت طبقه‌بندی‌شده (فوتی، جرح شدید، جرح خفیف) است. سوابق تصادفات خام تحت یک فرایند دقیق زمین‌کدگذاری و پاک‌سازی قرار گرفتند. مجموعه‌داده تحلیلی نهایی شامل ۳٬۲۱۲ تصادف عابر پیاده زمین‌کدشده است که از این میان ۱۷۸ مورد فوتی و ۸۷۶ مورد منجر به جراحات شدید ناتوان‌کننده بود.

۳-۴-ممیزی میدانی حقیقت زمینی توسط متخصص برای اعتبارسنجی مدل

برای آموزش شاخص خودکار جامع ایمنی عابر پیاده و محیط (APSEI) و اعتبارسنجی کمّی معیارهای خودکار در برابر کیفیت محیطی ادراک‌شده انسانی، یک ممیزی میدانی حقیقت زمینی اختصاصی انجام شد. یک نمونه تصادفی لایه‌بندی‌شده از ۵۰۰ نقطه خیابانی از ۹۲٬۱۲۵ موقعیت نمونه‌برداری SVI گرفته شد و لایه‌ها بر اساس سلسله‌مراتب راه و فراوانی تاریخی تصادفات تعریف شدند. یک تیم سه‌نفره از ارزیابان آموزش‌دیده برنامه‌ریزی شهری و ایمنی حمل‌ونقل به‌طور مستقل هر یک از ۵۰۰ موقعیت را بازدید و ارزیابی کردند. ابزار ممیزی یک سیاهه استاندارد و عینی اقتباس‌شده از ابزارهای ممیزی مقیاس خرد بود (Day و همکاران، ۲۰۰۶). برای هر نقطه، ارزیابان ۱۵ عنصر گسسته شامل پیوستگی پیاده‌رو، عرض مؤثر، کیفیت سطح، کفایت روشنایی، دید خط‌کشی گذرگاه، حائل از ترافیک و ایمنی ادراک‌شده کلی را روی یک مقیاس لیکرت ۵–نقطه‌ای امتیازدهی کردند. قابلیت اعتماد بین‌ارزیاب بین سه ارزیاب، اندازه‌گیری‌شده با ضریب همبستگی درون‌رده‌ای (ICC)، ۰/۸۷ (۹۵٪CI : ۰/۸۴–۰/۹۰) بود که نشان‌دهنده توافق عالی است.

۴- روش‌شناسی

این بخش چارچوب محاسباتی چهارمرحله‌ای نوینی را که برای ارزیابی خودکار ویژگی‌های ایمنی عابر پیاده در مقیاس شهری توسعه یافته، شرح می‌دهد. همه مدل‌های یادگیری عمیق در PyTorch پیاده‌سازی و روی یک خوشه محاسباتی مجهز به چهار GPU NVIDIA A100 آموزش داده شدند.

٤-١- مرحله یک: استخراج ویژگی‌های صریح ایمنی از طریق بخش‌بندی معنایی

٤-١-١- معماری مدل و منطق انتخاب

ما معماری SegFormer (Xie و همکاران، ۲۰۲۱) را به‌عنوان پشتیبان این وظیفه انتخاب می‌کنیم. برخلاف رمزگذارهای سنتی شبکه عصبی پیچشی (CNN)، رمزگذار ترانسفورمر سلسله‌مراتبی و چندمقیاسه SegFormer هم ویژگی‌های درشت با وضوح بالا و هم ویژگی‌های ریز با وضوح پایین تولید می‌کند. رمزگشای سبک MLP آن از نظر محاسباتی کارآمد است و آن را برای استنتاج در مقیاس بزرگ روی بیش از ۳٦۰٬۰۰۰ تصویر مناسب می‌سازد.

٤-١-٢- راهبرد یادگیری انتقالی و تنظیم دقیق

یک پارادایم یادگیری انتقالی برای تطبیق یک مدل از پیش آموزش‌دیده روی یک مجموعه‌داده شهری عمومی بزرگ با ویژگی‌های بصری خاص اصفهان استفاده شد. واریانت SegFormer-B4 که روی مجموعه‌داده Cityscapes (Cordts و همکاران، ۲۰۱٦) پیش‌آموزش‌دیده بود، به‌عنوان پیکربندی وزن اولیه به کار رفت. یک نمونه تصادفی از ۲٬۰۰۰ تصویر پانوراما از پیکره SVI اصفهان به‌صورت دستی در سطح پیکسل توسط سه پژوهشگر آموزش‌دیده با استفاده از ابزار LabelMe حاشیه‌نویسی شد و یک مجموعه‌داده اختصاصی «Isfahan-PedSafety» ایجاد شد. مدل سپس برای ۸۰ دوره روی این مجموعه‌داده با یک تابع هزینه آنتروپی متقاطع وزن‌دار آموزش دید. بهینه‌ساز AdamW با نرخ یادگیری اولیه ٦-e۵ و برنامه بازپخت کسینوسی استفاده شد.

٤-١-٣- تعاریف طبقات هدف و محاسبه شاخص در سطح پیکسل

مدل تنظیم‌شده دقیق برای بخش‌بندی قابل‌اعتماد یک طبقه‌بندی پالایش‌شده از هشت طبقه هدف آموزش دید که به‌طور خاص برای ارتباط برقرارشده یا فرض‌شده آن‌ها با خطر تصادف عابر پیاده و کیفیت محیطی ادراک‌شده انتخاب شدند. جدول ۱ به‌طور رسمی هر طبقه و شاخص صریح و قابل‌کمّی‌سازی مشتق‌شده از نسبت‌های پیکسلی آن در تصویر را تعریف می‌کند.

جدول ۱. طبقات هدف بخش‌بندی معنایی و شاخص‌های صریح مشتق‌شده

نام طبقه	تعریف عملیاتی برای حاشیه‌نویسان	شاخص مشتق‌شده	روش محاسبه در هر تصویر
پیاده‌رو	هر سطح سنگ‌فرش‌شده برجسته یا مشخص‌شده برای استفاده انحصاری عابر پیاده	نسبت پیاده‌رو (SR)	N_sidewalk / N_horizontal_plane
گذرگاه عابر پیاده	ناحیه عبور عابر پیاده خط‌کشی‌شده روی سواره‌رو	امتیاز حضور گذرگاه (CPS)	Binary (else 0 ,۵۰۰ < 1 if N_crosswalk)
فضای سبز	شاخ و برگ شامل درختان، درختچه‌ها و چمن	شاخص دید سبز (GVI)	N_green / N_total (Li et al., 2018)
مانع ثابت	عناصر عمودی دائمی (تیرهای برق، بولاردها، کیوسک‌ها، جعبه‌های گل)	چگالی مانع (OD)	N_obstacle / (N_sidewalk + ε)
پارک حاشیه‌ای	وسایل نقلیه ثابت در امتداد جدول یا سواره‌رو	نسبت تجاوز پارک (PER)	N_parking / (N_sidewalk + N_road)
روشنایی خیابان	تیرهای چراغ برق، بدون توجه به وضعیت روشنایی	شاخص تعداد روشنایی (LCI)	تعداد مراکز نمونه گسسته / تصویر
علائم/چراغ ترافیکی	علائم و چراغ‌های رو به بیننده، مرتبط با تیرها	امتیاز کنترل ترافیک (TCS)	تعداد مراکز نمونه گسسته / تصویر
عرض مؤثر مسیر	عرض پیوسته بدون مانع در صفحه پیاده‌رو	شاخص عرض مؤثر مسیر (EPW)	حداقل فاصله پیکسلی افقی بین موانع در ماسک پیاده‌رو

ε یک ثابت کوچک برای جلوگیری از تقسیم بر صفر است. عرض مؤثر مسیر (EPW) با اعمال یک الگوریتم تبدیل فاصله روی ماسک دودویی پیاده‌رو، پس از تفریق ماسک موانع ثابت و وسایل نقلیه پارک‌شده‌ای که به آن تجاوز می‌کنند، محاسبه می‌شود. حداقل عرض مقطع شناسایی و با عرض قاب تصویر نرمال‌سازی می‌شود. برای هر یک از ۹۲٬۱۲۵ نقطه نمونه‌برداری، چهار شاخص جهت‌دار برای هر طبقه میانگین‌گیری شد تا یک نمایه محیطی همه‌جهته ایجاد شود.

٤-٢-مرحله دو: تقطیر زمینه بصری ضمنی از طریق یادگیری خودنظارتی

٤-٢-١- پیش‌آموزش با DINO روی پیکره SVI اصفهان

ما الگوریتم DINO (خود-تقطیر بدون برچسب) (Caron و همکاران، ۲۰۲۱ [٤]) را برای آموزش یک ترانسفورمر بینایی (ViT-S/16) از ابتدا روی کل پیکره بدون برچسب SVI اصفهان (یک زیرمجموعه تصادفی از ۱۲۰٬۰۰۰ پانوراما) به کار بردیم. چارچوب DINO با ایجاد دو «نمای تقویت‌شده» تصادفی متفاوت از یک تصویر ورودی کار می‌کند. شبکه «دانش‌آموز» برای تطبیق توزیع احتمال خروجی شبکه «معلم» آموزش می‌بیند. این فرایند خود-تقطیر، ViT را مجبور می‌کند یک فضای نهفته بسیار ساختاریافته و معنادار بدون هیچ حاشیه‌نویسی انسانی بیاموزد. پیش‌آموزش برای ۳۰۰ دوره با اندازه دسته ۲۵٦ انجام شد.

٤-٢-٢-استخراج بردار ویژگی و کاهش ابعاد

پس از پیش‌آموزش، مدل آموزش‌دیده ViT-S/16 صرفاً به‌عنوان یک استخراج‌کننده ویژگی به کار رفت. برای هر یک از ۳٦۸٬۵۰۰ تصویر کامل، بردار توکن [CLS]-بُعدی از ۷٦۸ آخرین بلوک ترانسفورمر استخراج شد که نمایانگر یک توصیف‌گر معنایی کل‌نگر و به‌شدت فشرده از زمینه بصری ضمنی صحنه است. برای تسهیل کاوش و تفسیر این فضای ویژگی ۷٦۸-بُعدی، ما از تقریب و تصویربرداری منیفلد یکنواخت (UMAP) (McInnes و همکاران، ۲۰۱۸) برای کاهش ابعاد غیرخطی استفاده کردیم. UMAP به دلیل حفظ برتر ساختار داده جهانی و مقیاس‌پذیری محاسباتی، بر t-SNE ترجیح داده شد. بردارهای SSL۷٦۸-بُعدی به یک جاسازی دوبُعدی برای تجسم تصویر شدند و خوشه‌های نوظهوری را آشکار کردند که فرض می‌شود با گونه‌شناسی‌های متمایز محیطی شهری مطابقت دارند.

٤-٣-مرحله سه: ساخت شاخص خودکار ایمنی عابر پیاده و محیط (APSEI)

٤-٣-٤- مهندسی ویژگی ترکیبی

برای هر نقطه نمونه‌برداری، یک بردار ویژگی نهایی جامع با الحاق دو جریان داده متمایز ایجاد شد: (۱) بردار ویژگی صریح (۸ بُعد): هشت شاخص اسکالر محاسبه‌شده از بخش‌بندی معنایی؛ و (۲) بردار ویژگی ضمنی (۷٦۸ بُعد): توکن [CLS] ۷٦۸-بُعدی از مدل ViT-DINO. این فرایند به یک بردار ویژگی ترکیبی ۷۷٦-بُعدی در هر موقعیت انجامید که یک امضای دیجیتال کل‌نگر ترکیب‌کننده اندازه‌گیری‌های فیزیکی قابل‌تفسیر با معناشناسی بصری غنی و ظریف را فراهم می‌کند.

٤-٣-٤- مدل رگرسیون نظارت‌شده برای پیش‌بینی نمره ایمنی

برای زمین‌گیر کردن این ویژگی‌های تصویری در معیاری از ایمنی و کیفیت دنیای واقعی، یک مدل پیش‌بین نظارت‌شده ساختیم. متغیر هدف، نمره ممیزی متخصص پیوسته و میانگین‌گیری‌شده (روی مقیاس ۱-۵) از ۵۰۰ نقطه ممیزی میدانی حقیقت زمینی بود. بردارهای ویژگی ترکیبی ۷۷٦-بُعدی از این ۵۰۰ موقعیت، همراه با نمرات ممیزی متناظرشان، داده آموزش و آزمون را تشکیل دادند. مجموعه‌داده به‌طور تصادفی به مجموعه آموزش (۷۰٪، n= ۳۵۰) و مجموعه آزمون نگه‌داشته‌شده (۳۰٪، n=۱۵۰) تقسیم شد. ما یک رگرسیور XGBoost (Chen & Guestrin, 2016 [6]) آموزش دادیم. ابرپارامترها با استفاده از جستجوی شبکه‌ای اعتبارسنجی متقاطع ۵-تایی روی مجموعه آموزش تنظیم شدند. در نهایت، مدل XGBoost آموزش‌دیده برای پیش‌بینی نمره APSEI برای همه ۹۲٬۱۲۵ نقطه نمونه‌برداری در اصفهان به کار رفت.

٤-٤- مرحله چهار: راهبرد اعتبارسنجی چندهدفه

٤-١-١- اعتبارسنجی تجربی با داده‌های عینی تصادفات (تطبیق تصادفات)

آزمون اعتبارسنجی بیرونی اولیه شامل برقراری یک پیوند آماری بین APSEI صرفاً مشتق‌شده از تصویر و پیامد دنیای واقعی تصادفات عابر پیاده-وسیله نقلیه است. واحد تحلیل به‌عنوان

حائل ۵۰۰ متری قطعه راه تعریف شد. برای هر قطعه، متغیر وابسته فراوانی تجمیع‌شده پنج‌ساله تصادفات عابر پیاده و متغیر مستقل اصلی میانگین نمره APSEI بود. برای کنترل اثرات مخدوش‌کننده، متغیرهای مواجهه شامل حجم میانگین ترافیک روزانه سالانه (AADT)، تراکم جمعیت مسکونی و شاخص آنتروپی اختلاط کاربری زمین گنجانده شد. ما یک مدل رگرسیون پواسون وزن‌دار جغرافیایی (GWPR) اعمال کردیم (Hadayeghi و همکاران، ۲۰۰۳).

٤-۲-۲- اعتبار معیاری با ممیزی میدانی متخصص (همبستگی ممیزی)

جریان اعتبارسنجی دوم توافق مستقیم بین شاخص خودکار و استاندارد طلایی تثبیت‌شده، هرچند ذهنی، ارزیابی متخصص انسانی را ارزیابی می‌کند. با استفاده از ۱۵۰ نقطه آزمون نگه‌داشته‌شده از ممیزی متخصص، ما ضریب همبستگی درون‌رده‌ای (ICC) با استفاده از یک مدل اثرات مختلط دوطرفه برای توافق مطلق و ضریب تعیین (R^2) بین نمرات APSEI پیش‌بینی‌شده و نمرات میانگین ممیزی میدانی را محاسبه کردیم (Li, 2016 [8 & Koo]). یک نمودار Bland-Altman نیز برای بازرسی بصری حدود توافق و اریبی سیستماتیک بین دو روش ساخته شد.

۵- نتایج

۵-۱- عملکرد مدل بخش‌بندی معنایی

مدل تنظیم‌شده دقیق SegFormer-B4 عملکرد پیشرفته‌ای در تجزیه چشم‌اندازهای خیابانی پیچیده اصفهان نشان داد. پس از ۸۰ دوره تنظیم دقیق روی مجموعه‌داده ۲٬۰۰۰ تصویری Safety Isfahan-Ped، مدل به میانگین نسبت اشتراک بر اجتماع (mIoU)۰/۷۶۲ و دقت کلی پیکسلی ۰/۹۱٤ روی مجموعه اعتبارسنجی ۲۰٪ نگه‌داشته‌شده دست یافت.

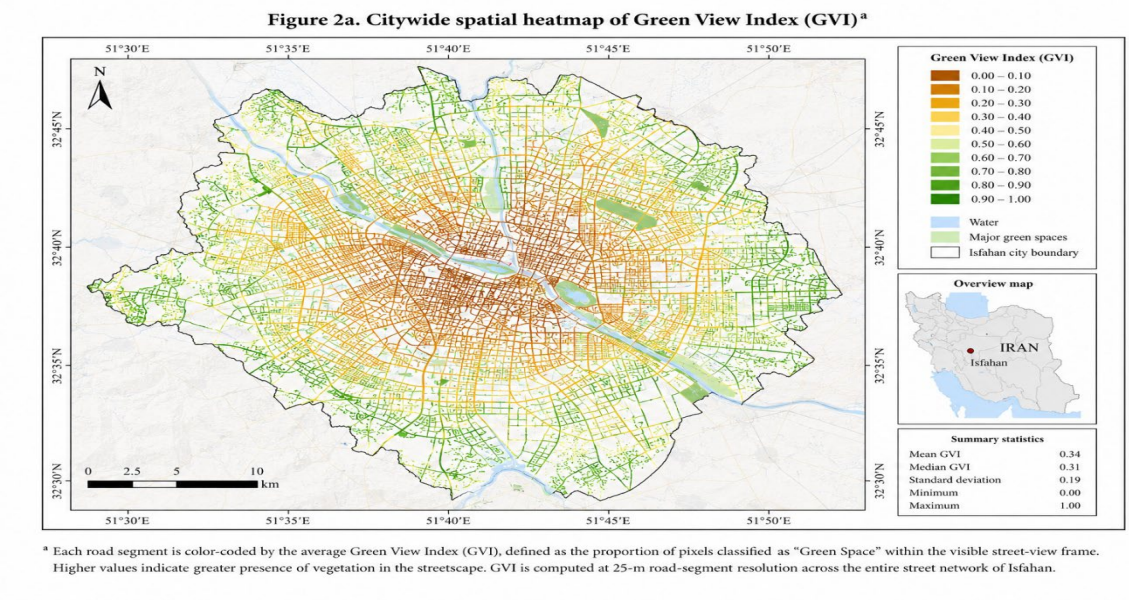
IoU و F1–امتیاز هر طبقه، که در جدول ۲ تفصیل یافته، عملکرد بسیار استواری را در همه طبقات حیاتی ایمنی نشان می‌دهد.

جدول ۲. عملکرد بخش‌بندی هر طبقه روی مجموعه اعتبارسنجی Isfahan-PedSafety

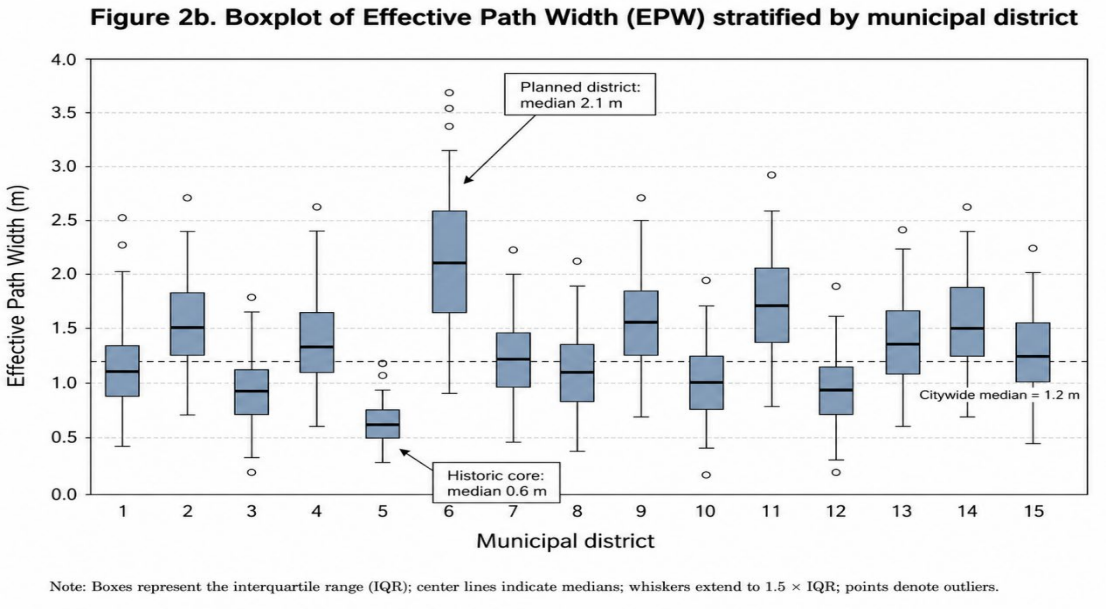
طبقه هدف	IoU	دقت	فراخوانی	F1–امتیاز
پیاده‌رو	۰,۸٤۳	۰,۹۱٦	۰,۹۱۸	۰,۹۱۷
گذرگاه عابر پیاده	۰,۷۷۲	۰,۸۹۱	۰,۸۵۳	۰,۸۷۱
فضای سبز	۰,۸۱۱	۰,۹۰٤	۰,۸۸۹	۰,۸۹٦
مانع ثابت	۰,٦۸٤	۰,۸۲۳	۰,۸۰۸	۰,۸۱۵
پارک حاشیه‌ای	۰,۷۵۱	۰,۸٦۸	۰,۸۵۳	۰,۸٦۰
روشنایی خیابان	۰,٦۹۲	۰,۸٤۵	۰,۷۹۳	۰,۸۱۸
علائم/چراغ ترافیکی	۰,٦۵۳	۰,۸۰۵	۰,۷۷۷	۰,۷۹۱
راه/پس‌زمینه	۰,۸۷٤	۰,۹٤۱	۰,۹۲٦	۰,۹۳۳
میانگین ماکرو	**۰,۷٦۰**	**۰,۸۷٤**	**۰,۸۵۲**	**۰,۸٦۳**

۵-۲-تحلیل بصری و آماری ویژگی‌های محیطی صریح

استنتاج در مقیاس شهری مدل بخش‌بندی روی همه ۹۲٬۱۲۵ نقطه یک پایگاه داده فضایی غنی و چندبُعدی از هشت شاخص صریح تولید کرد که نابرابری‌های محیطی stark را در سراسر بافت شهری اصفهان آشکار می‌کند. میانگین در مقیاس شهری شاخص دید سبز (GVI) مقدار پایین ۰/۰۸۷ (SD = ۰/۰۶۵) با توزیع به‌شدت چوله به راست بود. «بیابان‌های سبز» شدید (GVI < ۰/۰۲) در هسته تاریخی متراکم و گستره‌های وسیع، خشک و صناعی و توسعه‌های مسکونی پیرامونی به سمت جنوب‌شرق متمرکز شده‌اند. تحلیل عرض مؤثر مسیر (EPW) نشان داد که میانه EPW در سراسر شهر فقط ۱/۲ متر بود و ۳۵٪ از همه نقاط نمونه‌برداری EPW کمتر از ۰/۸ متر داشتند.



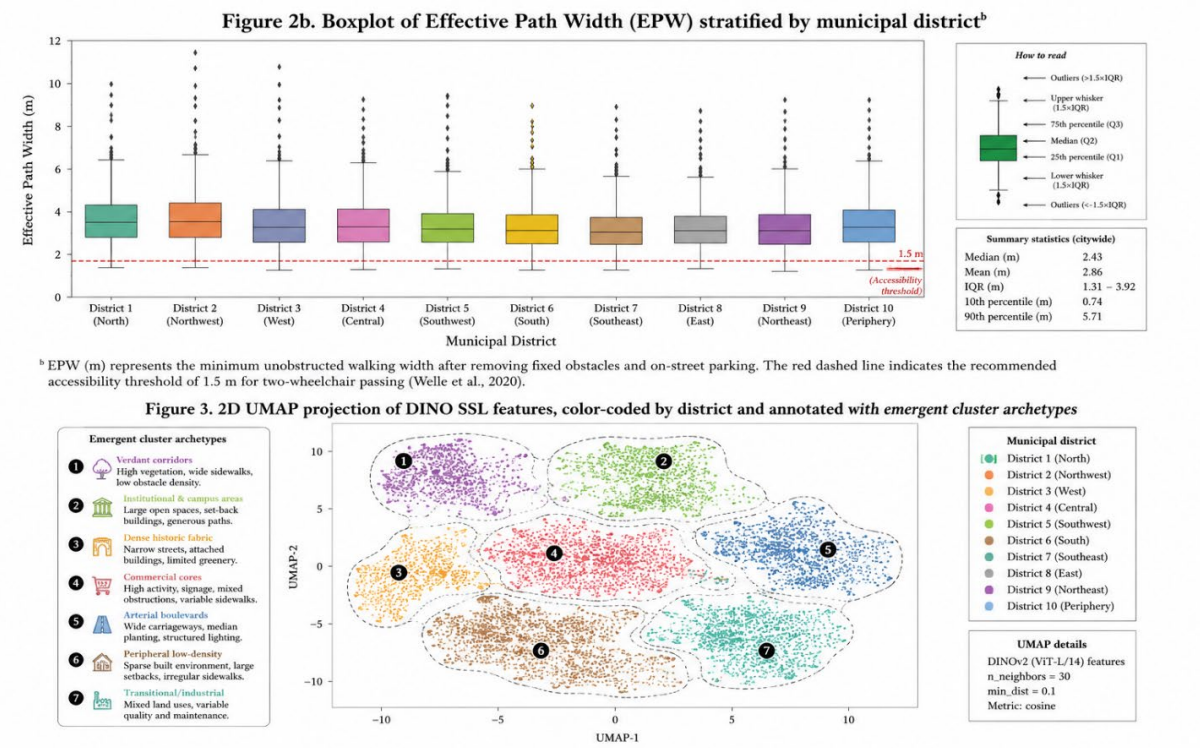
شکل ۲ الف. نقشه حرارتی فضایی در مقیاس شهری شاخص دید سبز (GVI)



شکل ۲ ب. نمودار جعبه‌ای عرض مؤثر مسیر (EPW) لایه‌بندی‌شده بر اساس منطقه شهرداری

۵-۳-تجسم فضای ویژگی یادگرفته‌شده خودنظارتی

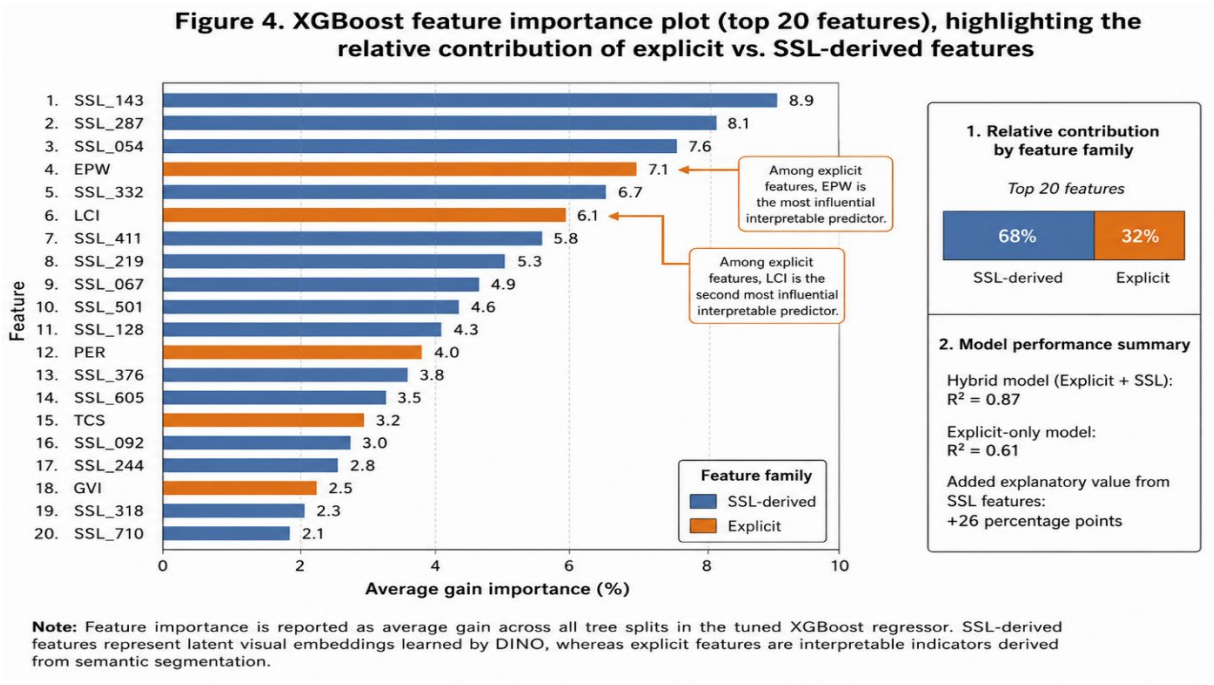
پیش‌آموزش DINO روی پیکره بدون برچسب SVI اصفهان یک فضای نهفته بسیار ساختاریافته تولید کرد که به‌طور خودبه‌خود گونه‌شناسی‌های محیطی ظریف را بدون هیچ برچسب آموزشی صریح ثبت کرد. تصویر دوبُعدی UMAP از بردارهای توکن [CLS]ـ۷۶۸-بُعدی، مشتق‌شده از یک نمونه تصادفی ۲۰٬۰۰۰ تصویری، خوشه‌های متمایز و به‌خوبی جدا شده‌ای را آشکار می‌کند که با کهن‌الگوهای بصری منسجم مطابقت دارند. مهم‌تر از همه، یک خوشه «محیط تخریب‌شده» (خوشه C) ظاهر شد که از تصاویر مناطق پیرامونی و صناعی پر شده است. کاوش در این تصاویر آشکار می‌کند که مدل DINO یاد گرفته ترکیب خاصی از بافت‌های بصری سطح-پایین - آسفالت ترک‌خورده، مسیرهای خاکی غیرسنگ‌فرش، توده‌های زباله بدون ساختار و لبه‌های ساختمانی فرسوده - را به یک بازنمایی نهفته منسجم تبدیل کند.



شکل ۳. تصویر دوبُعدی UMAP از ویژگی‌های SSL DINO، رنگ‌آمیزی‌شده بر اساس منطقه و حاشیه‌نویسی‌شده با کهن‌الگوهای خوشه‌ای نوظهور

۵-۵. تحلیل اهمیت ویژگی برای مدل APSEI

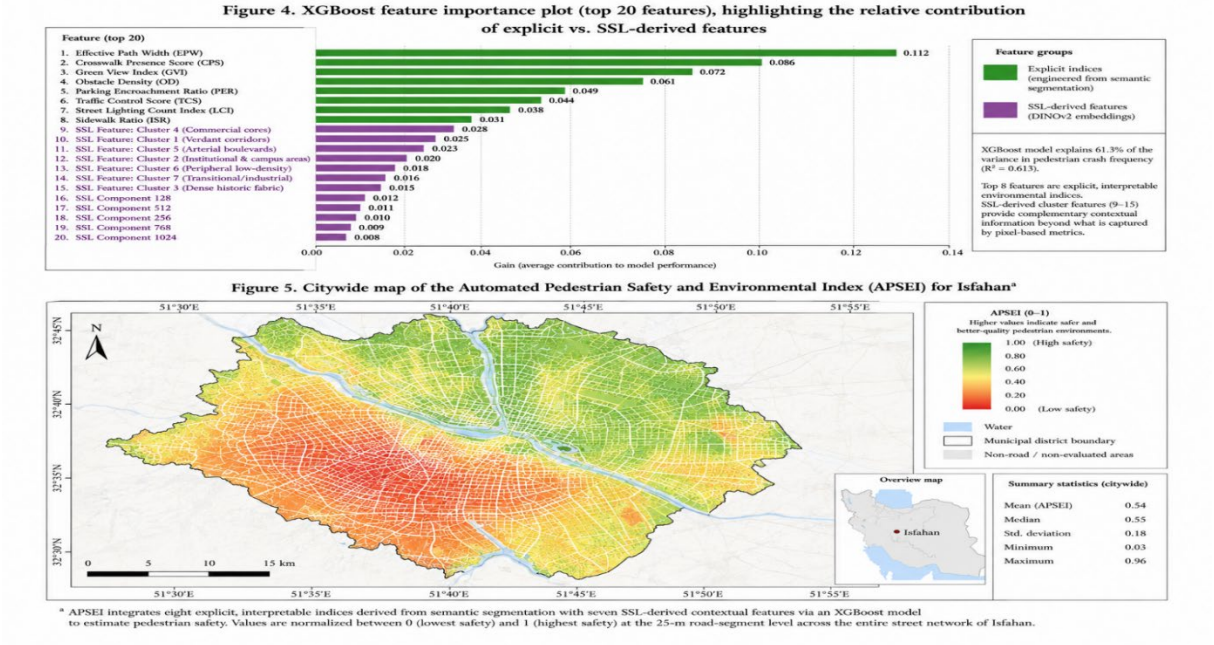
مدل XGBoost آموزش‌دیده برای پیش‌بینی نمرات ممیزی متخصص از بردار ویژگی ترکیبی ۷۷۶-بُعدی، دقت پیش‌بینی بالایی نشان داد. از میان ویژگی‌های صریح، عرض مؤثر مسیر (EPW) و شاخص تعداد روشنایی (LCI) پیش‌بین‌های غالب بودند و با هم ۱۸٪ از بهره کل را به خود اختصاص دادند. با این حال، چشمگیرترین یافته، غلبه جمعی ویژگی‌های ضمنی مشتق‌شده از SSL است. در مجموع، ۵۰ بُعد ویژگی SSL برتر بیش از ۴۷٪ از بهره پیش‌بینی کل مدل را سهم داشتند. مدل مبتنی صرفاً بر ویژگی‌های بخش‌بندی صریح به R^2 تنها ۰/۶۱ دست یافت، در حالی که مدل ترکیبی کامل شامل ویژگی‌های SSL به R^2 = ۰/۸۷ رسید.



شکل ٤. نمودار اهمیت ویژگی XGBoost (۲۰ ویژگی برتر)

۵-۵- نقشه در مقیاس شهری شاخص خودکار ایمنی عابر پیاده و محیط (APSEI)

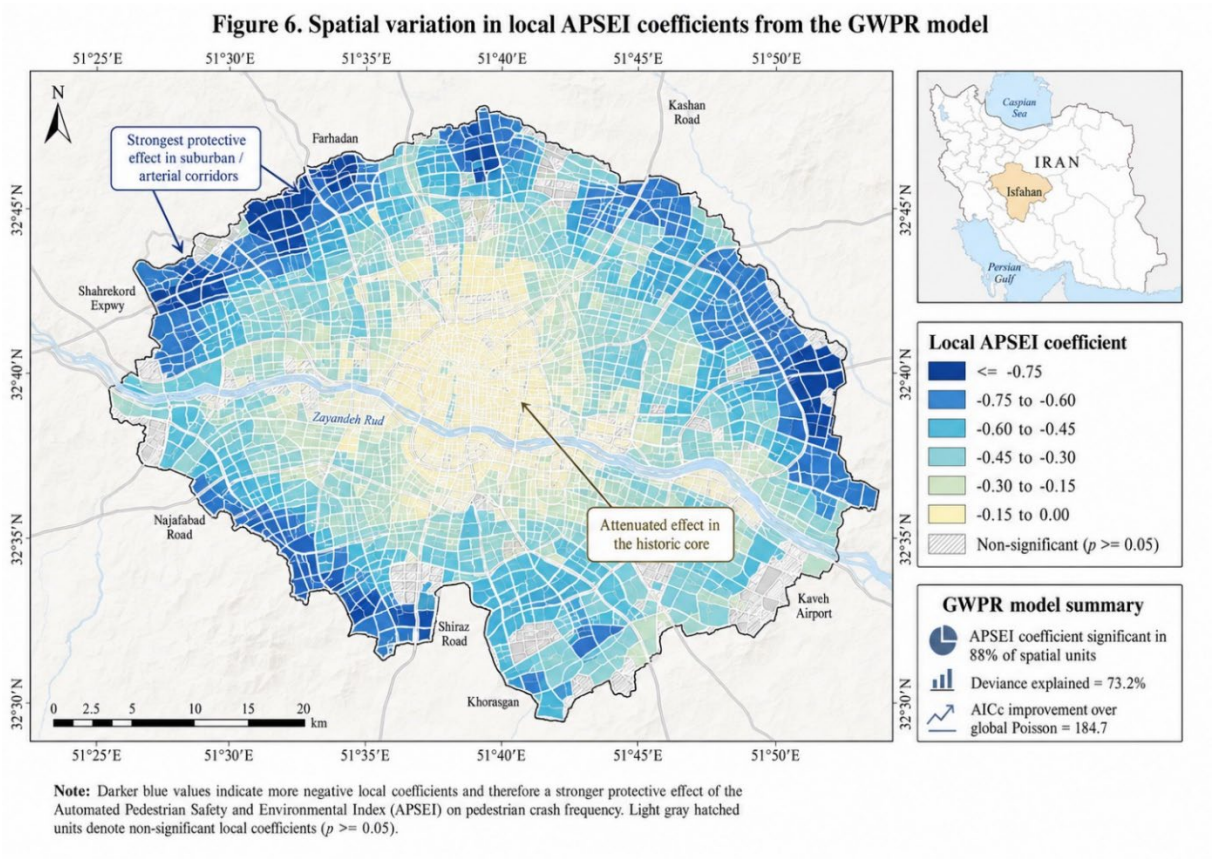
مدل XGBoost آموزش‌دیده برای پیش‌بینی نمره APSEI برای همه ۹۲٬۱۲۵ نقطه نمونه‌برداری مستقر شد و اولین سطح پیوسته با وضوح بالا از ایمنی عابر پیاده و کیفیت محیطی برای کل شهر اصفهان را تولید کرد. نمرات APSEI در مقیاس ۱-۵ از ۱/۱۸ تا ٤/۷۲ متغیر است با میانگین در مقیاس شهری ۲/۸۱ (SD =۰/۸۳). خطرناک‌ترین کریدورها (APSEI < ۱/۸) بزرگراه‌های اصلی مانند بلوار کاوه و کمربندی‌های پرسرعت اطراف شهر را ترسیم می‌کنند.



شکل ۵. نقشه در مقیاس شهری شاخص خودکار ایمنی عابر پیاده و محیط (APSEI) برای اصفهان

۵-۶- نتایج اعتبارسنجی تجربی تطبیق تصادفات

مدل رگرسیون پواسون وزن‌دار جغرافیایی (GWPR) شواهد قوی و از نظر فضایی صریح برای اعتبار بیرونی APSEI فراهم می‌کند. مدل GWPR بهبود قابل‌توجهی در نیکویی برازش نسبت به یک مدل پواسون جهانی استاندارد نشان داد (کاهش AICc به میزان ۱۸۴/۷) و ۷۳/۲٪ از انحراف فراوانی تصادفات عابر پیاده را تبیین کرد. ضریب APSEI در ۸۸٪ از واحدهای فضایی منفی و از نظر آماری معنادار ($p < ۰/۰۰۱$) بود. اثر محافظتی نمره بالاتر APSEI در کریدورهای حومه‌ای و شریانی پرسرعت قوی‌ترین است. در این محیط‌ها، افزایش یک واحدی در APSEI با کاهش ۵۵٪ در فراوانی مورد انتظار تصادفات همراه است (IRR = ۰/۴۵، $p < ۰/۰۰۱$).

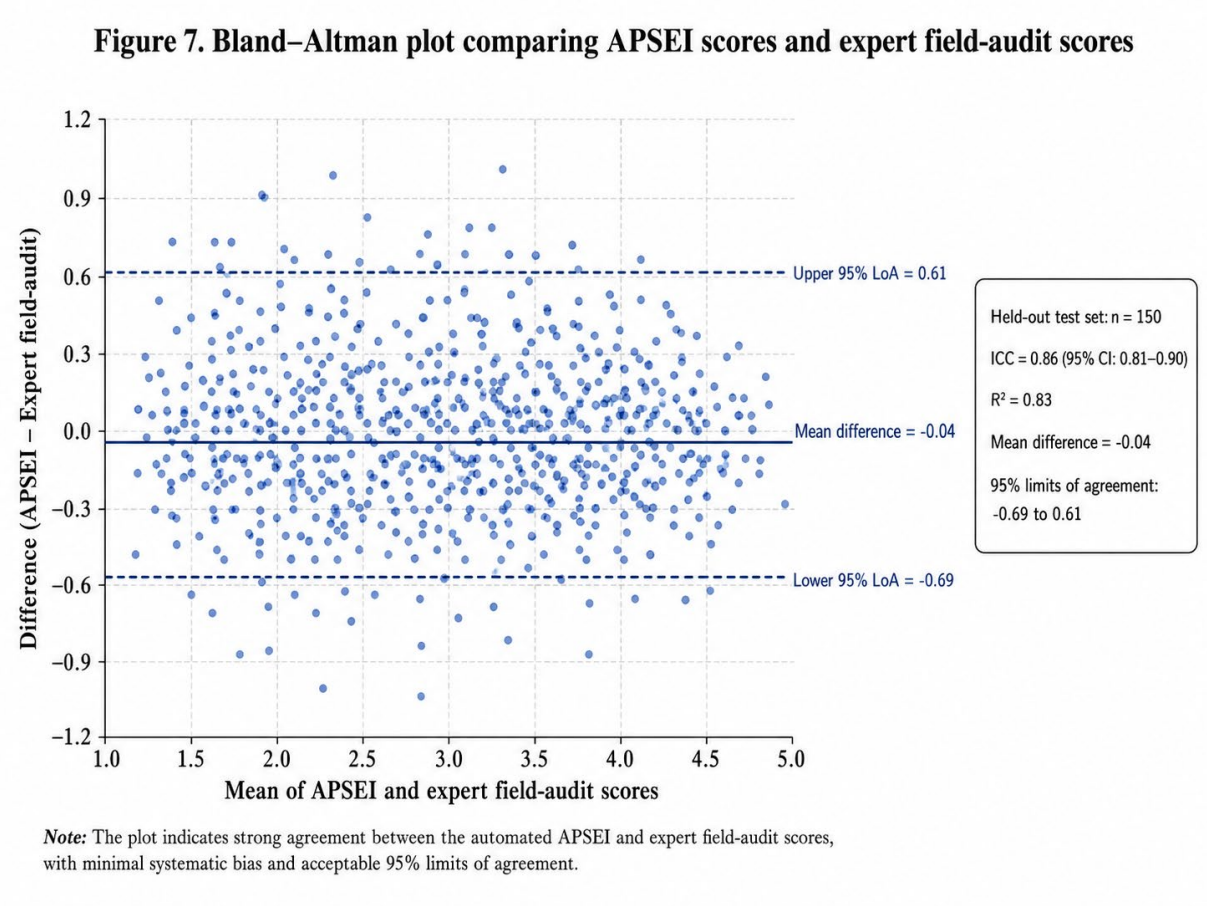


شکل ۶. تنوع فضایی ضرایب محلی APSEI از مدل GWPR

۵-۷-نتایج اعتبارسنجی ممیزی میدانی متخصص

اعتبار معیاری APSEI خودکار در برابر ارزیابی استاندارد طلایی متخصص انسانی با استفاده از مجموعه آزمون نگه‌داشته‌شده ۱۵۰ نقطه ممیزی تثبیت شد. ضریب همبستگی درون‌رده‌ای (ICC) برای توافق مطلق ۰/۸۶ (۹۵٪CI: ۰/۸۱–۰/۹۰) بود. ضریب تعیین $R^2 = ۰/۸۳$ بود.

نمودار Bland-Altman هیچ اریبی سیستماتیک معناداری را نشان نمی‌دهد، با میانگین تفاوت تنها −۰/۰۴ واحد و حدود توافق ۹۵٪ در محدوده ±۰/۶۵.



شکل ۷. نمودار Bland-Altman مقایسه نمرات APSEI و نمرات ممیزی میدانی متخصص

تفسیر یافته‌های کلیدی: قدرت پیش‌بینی زمینه بصری ضمنی

مهم‌ترین یافته نظری و عملی این پژوهش، برتری اثبات‌شده بازنمایی‌های خودنظارتی در پیش‌بینی هم ایمنی ادراک‌شده متخصص و هم پیامدهای واقعی تصادفات است. تحلیل اهمیت ویژگی از مدل XGBoost نشان داد که ۵۰ بُعد ویژگی برتر مشتق‌شده از SSL در مجموع بیش از ۴۷٪ از قدرت پیش‌بین مدل را به خود اختصاص دادند. این نتیجه کمّی شواهد تجربی قوی برای یک بینش حیاتی فراهم می‌کند: ادراک انسانی از ایمنی محیطی عابر پیاده، و خود تعیین‌کننده‌های محیطی خطر تصادف، کمتر توسط صرف حضور عناصر هدایت می‌شوند و بیشتر توسط وضعیت، پیکربندی و زمینه آن‌ها.

مشارکت‌های روش‌شناختی: عینیت، مقیاس‌پذیری و یک استاندارد طلایی جدید

این پژوهش یک مشارکت روش‌شناختی سه‌گانه به حوزه نوظهور علوم شهری محاسباتی برای ایمنی حمل‌ونقل دارد. نخست، به‌طور مستقیم به «گلوگاه ذهنیت» مستندشده ممیزی‌های میدانی سنتی می‌پردازد (Brownson و همکاران، ۲۰۰۹). دوم، جهش در مقیاس‌پذیری تحول‌آفرین است: یک تیم ارزیابان آموزش‌دیده چندین هفته برای ارزیابی ۵۰۰ موقعیت نیاز داشتند، در حالی که خط لوله آموزش‌دیده ۹۲٬۱۲۵ موقعیت را در کمتر از ۱۲ ساعت پردازش کرد. سوم، مقایسه بین مدل کامل GWPR و مدل پایه صرفاً بخش‌بندی نشان می‌دهد که مدل تنها-ویژگی‌های-صریح ۶۱/۸٪ از انحراف تصادفات را تبیین کرد، در حالی که APSEI ترکیبی به ۷۳/۲٪ رسید.

پیامدهای سیاستی و برنامه‌ریزی: عملیاتی‌سازی APSEI

خروجی‌های این چارچوب به‌عنوان ابزارهای پشتیبانی تصمیم قابل‌اقدام با سه کاربرد سیاستی فوری طراحی شده‌اند. نخست، نقشه APSEI در مقیاس شهری (شکل ۵) به‌عنوان یک ابزار

اولویت‌بندی فضایی قدرتمند برای سرمایه‌گذاری‌های ایمنی عمل می‌کند. دوم، عینیت و قابلیت تفسیر شاخص راه را برای توسعه یک پورتال مشارکت شهروندی عمومی و آنلاین هموار می‌کند. سوم، ماهیت SVI-محور APSEI پایش طولی و دینامیک ایمنی را ممکن می‌سازد.

محدودیت‌ها و مسیرهای پژوهش آینده

محدودیت نخست به حالت حسگر مربوط می‌شود: پانور اما SVI استاتیک و تک‌چشمی نمی‌تواند عوامل خطر دینامیک حیاتی مانند سرعت وسایل نقلیه را ثبت کند. دوم، دامنه جغرافیایی مطالعه محدود به اصفهان است. سوم، مسیر علّی موضوعی برای مدل‌سازی معادلات ساختاری باقی می‌ماند. چهارم، ممیزی متخصص یک حقیقت زمینی ناقص است. در نهایت، چارچوب محاسباتی به‌سوی یک معماری یکپارچه و چندوجهی اشاره می‌کند.

٦- نتیجه‌گیری

این پژوهش با یک چالش جهانی پایدار و مرگبار انگیزه یافت: فقدان ابزارهای مقیاس‌پذیر، عینی و به‌طور تجربی اعتبارسنجی‌شده برای تشخیص تعیین‌کننده‌های محیطی مقیاس خرد خطر تصادفات عابر پیاده. در پاسخ، ما یک چارچوب محاسباتی نوین پیشنهاد، توسعه و به‌طور دقیق اعتبارسنجی کردیم. شاخص خودکار ایمنی عابر پیاده و محیط (APSEI)، که به‌طور کامل از تصاویر نمای خیابان مشتق شده، از طریق یک مطالعه تجربی در مقیاس شهری در اصفهان به‌عنوان یک جانشین قابل‌اعتماد، دقیق و قابل‌اقدام اثبات شده است. این مطالعه پاسخ‌های مستقیم و قطعی به سه پرسش راهنمای پژوهش خود ارائه می‌دهد. مشارکت فکری مرکزی این کار در نشان دادن این است که ادغام یادگیری بصری صریح و خودنظارتی نه صرفاً افزایشی بلکه هم‌افزا برای وظیفه تشخیص ایمنی عابر پیاده است.

۷-سپاسگزاری

نویسندگان از پلیس راهور اصفهان و شهرداری اصفهان برای فراهم‌آوردن داده‌های تصادفات و تصاویر نمای خیابان قدردانی می‌کنند.

۸- پی‌نوشت‌ها

1.APSEI: Automated Pedestrian Safety and Environment Index

2. SVI: Street-View Imagery

3. SSL: Self-Supervised Learning

4. DINO: self-DIstillation with NO labels

5. ViT: Vision Transformer

6.GWPR: Geographically Weighted Poisson Regression

7. ICC: Intraclass Correlation Coefficient

8. UMAP: Uniform Manifold Approximation and Projection

9. mIoU: mean Intersection over Union

10. AADT: Average Annual Daily Traffic

۹- مراجع

- Arellana J, Saltarin M, Larranaga AM, et al. (2020). Walkability indicators and their association with pedestrian crashes. *Transp Res Part C.* 111:156-175.

-Bugs G, Granell C, Fonts O, et al. (2010). An assessment of Public Participation GIS. *Cities.* 27(3):172-181.

- Caron M, Touvron H, Misra I, et al. (2021). Emerging properties in self-supervised vision transformers. *IEEE ICCV;* 9630-9640.

- Chen S, Kuhn M, Prettner K, Bloom DE. (2019). The global economic burden of road traffic injuries. *Lancet Glob Health.* 7(12):e1663-e1674.

-Hadjidemetriou GM, Vela PA, (2020). Christodoulou SE. Vision-based deep learning for automated detection of pedestrian-vehicle conflicts. *Accid Anal Prev.* 135:105355.

- Hall CM, Ram Y. (2018). Walk Score and its potential contribution. *Transp Res Part D.* 61:310-324.

- Jacobsen PL. (2003). Safety in numbers. *Inj Prev.* 9(3):205-209.

- Millington C, Ward Thompson C, Rowe DA, et al. (2009). Development of the Scottish

-He K, Fan H, Wu Y, et al. (2020). Momentum contrast for unsupervised visual representation learning. *IEEE CVPR,* 9729-9738.

-Ibrahim MR, Haworth J, Cheng T. (2020). Understanding cities with machine eyes. Cities. 96:102481.

-Koo B, Guhathakurta S, Botchwey N. (2022). Development and validation of a street-level pedestrian safety index. *J Transp Geogr.* 99:103300.

-Koo TK, Li MY. (2016). A guideline of selecting and reporting intraclass correlation coefficients. *J Chiropr Med.* 15(2):155-163.

-Li X, Ratti C, Seiferling I. (2018). Assessing street-level urban greenery. *Urban For Urban Green.* 34:52-59.

-Mooney SJ, DiMaggio CJ, Lovasi GS, et al. (2016). Use of Google Street View. *Am J Public Health.* 106(5):876-82.

-Naci H, Chisholm D, Baker TD. (2009). Distribution of road traffic deaths by road user group. *Inj Prev.* 15(1):55-9.

-Naik N, Philipoom J, Raskar R, Hidalgo C. (2014). Streetscore. *IEEE CVPR Workshops,* 779-786.

-Rundle AG, Bader MDM, Richards CA, et al. (2011). Using Google Street View to audit neighborhood environments. *Am J Prev Med.* 40(1):94-100.

-Sleet DA, Pollack Porter KM, Ewing R, et al. (2020). Transportation and Public Health. *Elsevier.*

-Stoker P, Garfinkel-Castro A, Khayesi M, et al. (2015). Pedestrian safety and the built environment. *J Transp Land Use.* 8(2):81-105.

-World Health Organization (2018). Global status report on road safety Geneva.

-Zhao H, Shi J, Qi X, et al. (2017). Pyramid scene parsing network. *IEEE CVPR*, 2881-2890.

-Zuniga-Teran AA, Orr BJ, Gimblett RH, et al. (2017). Designing healthy communities. *Front Archit Res.* 6(1):63-73.

Walkability Assessment Tool. *Health Place.* 15(2):474-481.

-Nasar JL. (2008). Assessing perceptions of environments for active living. *Am J Prev Med.* 34(4):357-363.

- Yin L, Cheng Q. Wang Z, Shao Z. (2015). 'Big data' for pedestrian volume. *Comput Environ Urban Syst.* 53:96-108.

- Yin L, Wang Z. (2016). Measuring visual enclosure for street walkability. *Comput Environ Urban Syst.* 59:164-175.

- Zhang F, Zhou B, Liu L, et al. (2018). Measuring human perceptions. Landsc Urban Plan. 180:148-160.

-Aghaabbasi M, Moeinaddini M, Zaly Shah M, Asadi-Shekari Z. (2020). A new assessment model. *Accid Anal Prev.* 134:105336.

-Ayush K, Uzkent B, Meng C, et al. (2021). Geography-aware self-supervised learning. *IEEE ICCV;* 10181-10190.

-Biljecki F, Ito K. (2021). Street view imagery in urban analytics and GIS: A review. *Landsc Urban Plan.* 215:104217.

-Brownson RC, Hoehner CM, Day K, et al. (2009). Measuring the built environment for physical activity. *Am J Prev Med.* 36(4 Suppl):S99-S123.

-Carr LJ, Dunsiger SI, Marcus BH. (2010). Walk Score™ as a global estimate. *Am J Prev Med.* 39(5):460-463.

-Chen LC, Zhu Y, Papandreou G, et al. (2018). Encoder-decoder with atrous separable convolution. *ECCV,* 801-818.

-Chen T, Kornblith S, Norouzi M, Hinton G. A (2020). simple framework for contrastive learning. *ICML,* 1597-1607.

-Day K, Boarnet M, Alfonzo M, Forsyth A. (2006). The Irvine-Minnesota inventory. *Am J Prev Med.* 30(2):144-152.

-Ewing R, Cervero R. (2010). Travel and the built environment: A meta-analysis. *J Am Plann Assoc.* 76(3):265-294.

-Harvey C, Aultman-Hall L. (2016). Measuring urban streetscapes for livability. *Prof Geogr.* 68(1):149-158.

Automated Assessment of Pedestrian-Safety Features from Street-View Imagery Using Semantic Segmentation and Self-Supervised Representation Learning

(A Citywide Study in Isfahan)

Rasoul Zabihian, Department of Civil Engineering, Ki.C., Islamic Azad University, Kish, Iran.
***Mohammad Reza Ahadi,** Transportation Research Institute, Road, Housing and Urban Development Research Center, Tehran, Iran*
Ali Naderan, Department of Civil Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran.

E-mail: m.ahadi@bhrc.ac.ir

Received: May 2026- Accepted: November 2026

ABSTRACT
Microscale characteristics of the built environment substantially influence pedestrian safety; however, conventional field audits are costly, subjective, and difficult to implement at the city scale. This study develops and validates a hybrid computer-vision framework for the automated assessment of pedestrian safety using street-view imagery. The framework integrates semantic segmentation for extracting explicit and interpretable environmental attributes with self-supervised representation learning for capturing latent visual characteristics, including pavement condition, spatial enclosure, material quality, and visual disorder. The empirical analysis covers 368,500 street-view images collected from 92,125 georeferenced locations in Isfahan, Iran. A fine-tuned SegFormer-B4 model was used to quantify eight safety-related features, while a DINO-trained Vision Transformer generated contextual visual representations. These feature streams were fused within an XGBoost regression model to construct the Automated Pedestrian Safety and Environment Index (APSEI), using expert field-audit scores from 500 stratified locations as ground truth. External validity was evaluated against 3,212 geocoded pedestrian–vehicle crashes through geographically weighted Poisson regression.

Keywords: Pedestrian Safety, Street-View Imagery, Semantic Segmentation, Self-Supervised Learning, Geographically Weighted Poisson Regression