

ارائه یک روش مبتنی بر یادگیری عمیق به منظور تخمین موقعیت مطلق چاله‌ها در خیابان‌های شهری

همایون هادیگل^۱، علی حسینی نوه^{۲*}

^۱ دانشجوی کارشناسی ارشد فتوگرامتری-دانشکده مهندسی نقشه‌برداری-دانشگاه صنعتی خواجه‌نصیرالدین طوسی
homayoonh79@email.kntu.ac.ir

^۲ دانشیار دانشکده مهندسی نقشه‌برداری-دانشگاه صنعتی خواجه‌نصیرالدین طوسی
hosseininaveh@kntu.ac.ir

(دریافت: اسفند ۱۴۰۳، تصویب: شهریور ۱۴۰۴)

چکیده

شناسایی و تعیین دقیق موقعیت چاله‌های سطح آسفالت نقش مهمی در بهبود ایمنی رانندگان، کاهش هزینه‌های نگهداری و بهینه‌سازی مدیریت زیرساخت‌های حمل‌ونقل شهری دارد. در این پژوهش، روشی مبتنی بر یادگیری عمیق ارائه شده است که با استفاده از فناوری‌های در دسترس و مقرون‌به‌صرفه، این اهداف را محقق می‌سازد. روش پیشنهادی از یک شبکه عصبی عمیق برای تشخیص و بخش‌بندی چاله‌ها در تصاویر ثبت‌شده توسط دوربین گوشی‌های هوشمند استفاده می‌کند. سپس با به‌کارگیری یک الگوریتم تلفیق داده، اطلاعات استخراج‌شده از تصاویر با داده‌های سامانه موقعیت‌یاب جهانی (GPS) و واحد اندازه‌گیری اینرسی (IMU) دریافت‌شده از سنسورهای گوشی ترکیب می‌شود تا موقعیت مطلق چاله‌ها در سیستم مختصات جهانی (UTM) با دقت بالا تخمین زده شود. این رویکرد نوآورانه، چالش‌های مرتبط با تبدیل مختصات تصویر به مختصات جهانی و خطاهای ناشی از دقت پایین داده‌های سامانه موقعیت‌یاب جهانی را به‌طور مؤثر برطرف می‌کند. به‌منظور ارزیابی عملکرد روش پیشنهادی، مجموعه داده‌ای جامع از خیابان‌های شهری با استفاده از گوشی هوشمند متداول جمع‌آوری شد. در این فرآیند موقعیت چاله‌ها به دو روش تعیین شد: یک بار با استفاده از روش پیشنهادی ارائه‌شده در این پژوهش و بار دیگر با استفاده از گیرنده مولتی‌فرکانس G1 Plus Sout با دقت ۰.۱۲ متر که به عنوان مرجع دقیق در نظر گرفته شد. سپس نتایج حاصل از این دو روش با یکدیگر مقایسه شدند تا دقت و کارایی روش پیشنهادی ارزیابی گردد. نتایج آزمایش‌ها نشان می‌دهد که این روش قادر است موقعیت مطلق چاله‌ها را با میانگین خطای کمتر از ۲ متر در سیستم مختصات UTM تخمین بزند. این رویکرد امکان نقشه‌برداری دقیق و به‌روز از وضعیت سطح خیابان‌های شهری را با استفاده از ابزارهای روزمره و در دسترس عموم فراهم می‌کند. علاوه بر این، سادگی و مقرون‌به‌صرفه بودن این روش، امکان مشارکت گسترده شهروندان در جمع‌آوری داده‌ها و بهبود زیرساخت‌های شهری را فراهم می‌سازد.

واژگان کلیدی: یادگیری عمیق، شبکه عصبی، شناسایی اشیاء، ویژوال اودومتری، تلفیق داده‌ها

* نویسنده رابط

۱- مقدمه

شناسایی چاله‌های سطح آسفالت با استفاده از تصاویر، یک موضوع مهم و ضروری در حوزه نگهداری و مدیریت جاده‌ها و سطوح آسفالت است. این فناوری می‌تواند به طور موثری در بهبود ایمنی رانندگان، کاهش هزینه‌ها و بهبود کیفیت جاده‌ها نقش اساسی داشته باشد. چاله‌های سطح آسفالت می‌توانند عوامل خطرناکی برای رانندگان و وسایل نقلیه باشند. می‌توان با تشخیص و موقعیت‌یابی دقیق این چاله‌ها، اقدامات مناسبی انجام داد تا از تصادفات و آسیب به وسایل نقلیه جلوگیری شود. با داشتن اطلاعات دقیق درباره موقعیت و تعداد چاله‌ها، نهادهای مسئول می‌توانند برنامه‌ریزی مناسبی برای تعمیرات انجام داده و بهره‌وری بیشتری در هزینه‌های نگهداری داشته باشند. همچنین، این روش امکان تعمیرات به موقع و پیشگیری از آسیب‌های جدی را فراهم می‌کند. استفاده از روش‌های اتوماتیک و تصویربرداری برای شناسایی و موقعیت‌یابی چاله‌ها، هزینه‌های نگهداری و تعمیرات را نیز به شدت کاهش می‌دهد. این روش با دقت بالا می‌تواند چاله‌ها را شناسایی کرده و به نتیجه‌گیری مناسب درباره نیاز به تعمیرات بپردازد که این امر منجر به افزایش بهره‌وری مالی می‌شود. می‌توان با شناسایی و موقعیت‌یابی چاله‌ها، تعمیرات را به موقع را انجام داد و از تخریب بیشتر سطح جاده جلوگیری کرد. این بهبودات در نهایت باعث افزایش رضایتمندی شهروندان و بهبود کیفیت زندگی آنان می‌شود. در دنیای امروز، تکنولوژی و ارتباطات به طور فراگیری در حال رشد هستند و این رشد باعث تحولات چشمگیری در صنایع مختلف و زمینه‌های کاربردی مختلف شده است. یکی از نمونه‌هایی که این تحولات را به خوبی نشان می‌دهد، توانایی شناسایی و موقعیت‌یابی چاله‌های سطح آسفالت با استفاده از تصاویر و ترکیب داده‌های دوربین، داده‌های سامانه موقعیت‌یابی جهانی^۱ (GPS) و واحد اندازه‌گیری اینرسی (IMU)^۲ به منظور موقعیت‌یابی آن‌ها است. این موضوع به عنوان یک عنصر کلیدی در فرآیندهای کنترل و مدیریت زیرساخت‌های حمل و نقل، از جمله جاده‌ها و مسیرهای حمل و نقل عام مهم است. یکی از چالش‌های اصلی در این زمینه، ترکیب و تلفیق

داده‌های حاصل از دوربین‌های دیجیتال، GPS و IMU به منظور بهبود دقت و کیفیت داده‌های موقعیتی چاله‌ها است. این فرآیند معمولاً به نام تلفیق داده‌ها^۳ شناخته می‌شود. واژه اودومتري بينايي (Visual Odometry) ابتدا توسط Nister و همکارانش [۱] معرفی شد. این روش پیشنهاد شد تا با استفاده از تناظریابی مجموعه‌ای از نقاط پراکنده بین فریم‌های متوالی، موقعیت دوربین را تخمین بزند. پس از آن کارهای بسیار دیگری بر اساس این رویکرد مبتنی بر ویژگی انجام شد. ایده اصلی آن‌ها استخراج و ردیابی ویژگی‌ها در تصویر و کمینه کردن خطای تصویر کردن مجدد (Reprojection Error) است. کارهای بعدی مانند MonoSLAM [۲] و PTAM [۳] با وارد کردن سرشکنی دسته اشعه توانایی تولید همزمان یک نقشه پایدار از محیط را به این سیستم افزودند. در زمینه روش‌های مستقیم، Engel و همکاران الگوریتم DSO را معرفی کردند [۱۶] که با استفاده از یک مدل فوتومتریک پیچیده و بهینه‌سازی مشترک همه پارامترها، دقت قابل توجهی را به دست آورد. Mur-Artal و همکاران [۱۹] سیستم ORB-SLAM را توسعه دادند که با ترکیب تشخیص حلقه بسته و بهینه‌سازی گراف، یک سیستم مکان‌یابی و نقشه‌برداری همزمان (SLAM)^۴ را ارائه داد. در زمینه ترکیب حسگرها، Leutenegger و همکاران [۲۰] یک چارچوب بهینه‌سازی یکپارچه برای ترکیب اطلاعات بینایی و اینرسی ارائه کردند که به OKVIS معروف شد. Qin و همکاران [۲۱] سیستم VINS-Mono را معرفی کردند که با استفاده از یک رویکرد اتصال محکم و بهینه‌سازی پنجره لغزان^۵، عملکرد بلادرنگ را با دقت بالا ترکیب کرد. Forster و همکاران [۲۲] یک روش جدید برای پیش‌انتگرال‌گیری IMU معرفی کردند که محاسبات را به طور قابل توجهی کاهش داد. در زمینه یادگیری عمیق، Wang و همکاران [۲۳] یک شبکه عصبی عمیق برای تخمین عمق از تصاویر تک دوربین ارائه دادند که می‌توان با سیستم‌های SLAM کلاسیک ترکیب شود. Zhou و همکاران [۲۴] یک سیستم end-to-end برای تخمین همزمان عمق و حرکت دوربین معرفی کردند که بدون نیاز به داده‌های برچسب‌خورده آموزش می‌بیند. Bloesch و همکاران [۲۵] روشی برای استخراج ویژگی‌های عمیق ارائه دادند که می‌تواند مستقیماً

۴ Simultaneous Localization and Mapping

۵ sliding window

۱ Global Positioning System

۲ Inertial measurement unit

۳ Data Fusion

در یک فیلتر کالمن توسعه‌یافته (EKF)^۱ برای تخمین حالت استفاده شود. در زمینه سیستم‌های مقاوم، Yang و همکاران [۲۶] یک چارچوب تخمین حالت مقاوم در برابر اغتشاشات دینامیکی ارائه دادند که می‌تواند با حضور اشیاء متحرک در صحنه کار کند. Campos و همکاران [۲۷] یک سیستم بینایی-اینرسی برای محیط‌های چالش‌برانگیز مانند تونل‌ها و معادن توسعه دادند که در شرایط نوری ضعیف نیز عملکرد خوبی دارد. روش دیگر برای انجام Visual Odometry روش مستقیم است. در این روش به جای پیدا کردن ویژگی‌ها در تصویر و کمینه کردن خطای Reprojection Error، مستقیماً از پیکسل‌های تصویر و با کمینه کردن خطای فوتومتریک برای به دست آوردن موقعیت دوربین استفاده می‌کند که نتیجه آن کاهش حجم پردازش‌ها و افزایش سرعت و کارایی سیستم می‌باشد. در [۴] برای اولین بار روش مستقیم با استفاده از دوربین دوچشمی ارائه شد که از دقت و سرعت بالاتری نسبت به روش‌های غیرمستقیم مشابه برخوردار بود. روش‌های مستقیم مخصوصاً در کاربرد ردیابی با دوربین‌های RGB-D بسیار موثر بوده که توسط Newcombe و همکاران [۵] و Kerl و همکاران [۶] نمایش داده شده است. اخیراً روش‌های مستقیم برای سیستم‌های تک دوربین نیز در حالات ابرنقطه متراکم [۷]، ابر نقطه نیمه-متراکم [۸] و ابرنقطه پراکنده [۲] ارائه شده است که قدرت و پایداری سیستم‌های تک دوربین مشابه خود را تا حد قابل قبولی افزایش داده است. با توجه به ماهیت مکمل داده‌های IMU نسبت به سایر حسگرها، تلاش بسیاری برای ترکیب آن‌ها با دوربین صورت گرفته است. این سنسور می‌تواند حرکات سریع و ناگهانی دوربین را تخمین بزند و علاوه بر آن زوایای غلتش (Roll) و پیچش (Pitch) را نیز مشاهده کند. بنابراین چندین روش برای ترکیب این سنسور با دوربین وجود دارد. این روش‌ها به دو دسته کلی تقسیم می‌شوند:

دسته اول که معروف به "اتصال محکم" است. در این حالت اندازه‌گیری‌های IMU و دوربین در یک تابع انرژی متشکل از ترم خطاهای دوربین و IMU، به صورت همزمان بهینه می‌شود. حسن این روش آن است که اندازه‌گیری‌های اولیه این دو سنسور در هنگام بهینه‌سازی، در سطح عمیقی با یکدیگر ترکیب شده‌اند و همبستگی‌های بین آن‌ها نیز در نظر گرفته می‌شود و این باعث افزایش دقت و قدرت این دسته روش‌ها می‌شود. در [۱۲] از روش اتصال محکم برای

تلفیق اطلاعات دوربین استریو و IMU جهت تعیین موقعیت یک هواپیما استفاده شده است. روند انجام کار بدین صورت است که بعد از شناسایی عوارض کلیدی در هر تصویر، با توجه به مشاهدات حاصل از IMU، موقعیت قرارگیری آن‌ها در فریم بعدی پیش‌بینی می‌شود که این امر موجب کاهش فضای جستجو در فریم جدید می‌شود. پس از تناظریابی و محاسبه موقعیت، اختلاف بین موقعیت‌های پیش‌بینی شده و محاسبه شده برای تصحیح خطاهای IMU به کار گرفته می‌شود. این دسته [۱۶] و [۱۷] این امکان را فراهم می‌کند که هر زمان که نیاز باشد به صورت تطبیقی تابع انرژی را دوباره خطی‌سازی کنیم تا از خطای سیستماتیک ناشی از خطی‌سازی اجتناب شود.

دسته دوم به "اتصال سست" معروف هستند [۱۰] و [۱۱]. در این حالت سیستم بینایی فقط اندازه‌گیری‌های مربوط به موقعیت را انجام می‌دهد و هیچ تداخلی با اندازه‌گیری‌های IMU ندارد و توسط یک فیلتر کالمن توسعه‌یافته باهم ترکیب می‌شوند. یک مزیت این دسته روش‌ها پیاده‌سازی ساده‌تر آنهاست چرا که بهینه‌سازی و تعیین مجهولات هر سنسور به طور جداگانه انجام می‌شود. مزیت دیگر این روش‌ها، فراهم آوردن امکان استفاده از الگوریتم‌های فقط بینایی موجود مثل PTAM، LSD-SLAM و ... است. در [۱۳]، سامانه‌های مکان‌یابی و نقشه‌برداری همزمان تک‌چشمی با رویکرد اتصال سست، با IMU ادغام شده‌اند. این دسته، روش‌های مبتنی بر فیلتر هستند [۱۴] و [۱۵] که در چارچوب فیلتر کالمن با مدل‌های احتمالاتی کار می‌کنند. در زمینه شناسایی و تخمین موقعیت چاله‌ها، تحقیقات متعددی انجام شده است. Koch و Heinrich یک سیستم تشخیص چاله [۲۹] مبتنی بر پردازش تصویر ارائه دادند که از شبکه عصبی کانولوشنی عمیق (CNN) برای طبقه‌بندی مناطق آسیب‌دیده جاده استفاده می‌کند و دقت تشخیص ۹۲ درصد را به دست آورد. Jo و Ryu یک روش ترکیبی [۳۰] برای شناسایی چاله‌ها معرفی کردند که از داده‌های IMU و دوربین استفاده می‌کند. آن‌ها با استفاده از الگوریتم‌های پردازش سیگنال پیشرفته، الگوهای ارتعاشی خاص ناشی از چاله‌ها را شناسایی کرده و سپس با تصاویر دوربین تأیید می‌کنند. Madli و همکاران [۳۱] یک سیستم هوشمند برای نظارت بر وضعیت جاده توسعه دادند که از سنسورهای فراصوت و GPS برای تشخیص و ثبت موقعیت چاله‌ها

۱ Extended Kalman Filter

استفاده می‌کند. این سیستم با دقت ۹۳ درصد قادر به تشخیص چاله‌ها و دست‌اندازها است. Wang و همکاران [۳۲] یک روش مبتنی بر یادگیری عمیق برای تشخیص و طبقه‌بندی خودکار آسیب‌های سطح جاده ارائه دادند که می‌تواند انواع مختلف خرابی از جمله چاله‌ها را با دقت بالا شناسایی کند. Dhiman و Klette یک سیستم بینایی ماشین پیشرفته [۳۳] برای شناسایی چاله‌ها در شرایط نوری مختلف توسعه دادند که از تکنیک‌های پردازش تصویر چندمقیاسی و CNN استفاده می‌کند. این سیستم قادر به تشخیص چاله‌ها در شرایط مختلف آب و هوایی با دقت ۸۹ درصد است. Chen و همکاران [۳۴] یک سیستم نظارت بر جاده مبتنی بر SLAM ارائه دادند که می‌تواند به طور همزمان موقعیت خودرو و نقشه سه‌بعدی جاده را با جزئیات چاله‌ها تخمین بزند. آن‌ها از یک الگوریتم بهینه‌سازی چند هدف برای بهبود دقت تخمین عمق چاله‌ها استفاده کردند. Jahanshahi و همکاران [۳۵] یک روش جدید برای اندازه‌گیری خودکار حجم و عمق چاله‌ها با استفاده از تصاویر استریو معرفی کردند. این سیستم با استفاده از تکنیک‌های بازسازی سه‌بعدی و مدل‌سازی هندسی، می‌تواند ابعاد چاله‌ها را با خطای کمتر از ۵ درصد تخمین بزند. Yu و Salari یک سیستم تشخیص چاله مبتنی بر لیدار و دوربین [۳۶] ارائه دادند که از تلفیق داده‌های دو سنسور برای افزایش دقت تشخیص استفاده می‌کند. آن‌ها با استفاده از الگوریتم‌های پردازش ابر نقاط و تصویر، توانستند دقت تشخیص ۹۵ درصد را به دست آورند. Fox و همکاران [۳۷] یک سیستم نقشه‌برداری خودکار از چاله‌ها با استفاده از پهپاد و تصویربرداری هوایی توسعه دادند. این سیستم از تکنیک‌های فتوگرامتری و یادگیری عمیق برای شناسایی و اندازه‌گیری چاله‌ها استفاده می‌کند و می‌تواند نقشه‌های دقیق از وضعیت خرابی جاده‌ها تهیه کند. Telceken و همکاران [۵۹] یک روش جدید برای تشخیص اشیاء با وضوح بالا ارائه می‌دهند که از ترکیب الگوریتم برش تصویر (Sahi)^۱ [۶۴] و DeepSort [۴۰] و مدل SAM^۲ [۶۵] استفاده می‌کند. هدف اصلی این مطالعه، بهبود دقت و کارایی تشخیص اشیاء در تصاویر با وضوح بالا است. الگوریتم Sahi به تقسیم تصاویر بزرگ به بخش‌های کوچکتر می‌پردازد تا پردازش آنها ساده‌تر شود. مدل SAM برای تقسیم‌بندی دقیق و شناسایی اشیاء در این

بخش‌ها به کار میرود. این ترکیب باعث افزایش دقت تشخیص و کاهش خطا در تصاویر پیچیده میشود. نتایج آزمایش‌ها نشان می‌دهد که این روش در مقایسه با روش‌های سنتی عملکرد بهتری در تشخیص اشیاء با وضوح بالا دارد و می‌تواند در کاربردهایی مانند نظارت تصویری و سیستم‌های تشخیص خودکار مفید باشد. Kaess و همکاران در [۶۰] الگوریتم iSAM2 را معرفی کردند و از آن برای بهبود کارایی و دقت در حل مسائل SLAM استفاده نمودند. iSAM2 یک روش بهینه‌سازی افزایشی است که از ساختار داده‌ای Bayes Tree بهره می‌برد تا به‌روزرسانی‌های کارآمد و دقیقی در سیستم‌های SLAM انجام دهد. با استفاده از iSAM2، دقت موقعیت‌یابی و نقشه‌برداری بهبود یافت، زیرا این روش به‌طور مداوم خطاها را به‌روزرسانی و اصلاح می‌کند. iSAM2 قابلیت ادغام داده‌های سنسورهای مختلف را دارد و می‌تواند در محیط‌های پیچیده و داینامیک به‌خوبی عمل کند. Rajshekar و همکاران [۳۸] به توسعه یک اپلیکیشن هوشمند برای تشخیص چاله‌های خیابانی با استفاده از الگوریتم YOLO^۳ [۳۹] پرداختند. هدف این پروژه ارائه راه‌حلی برای شناسایی و مکان‌یابی گودال‌ها با استفاده از دوربین و GPS گوشی هوشمند است. کاربران می‌توانند تصاویر گودال‌ها و موقعیت آن‌ها را در اپلیکیشن بارگذاری کنند و موقعیت چاله‌ها روی نقشه نمایش داده می‌شود.

با توجه به مرور کارهای گذشته اکثر تحقیقات انجام‌شده بر شناسایی و موقعیت‌یابی چاله‌ها در تصاویر متمرکز بوده‌اند. روش‌هایی نظیر استفاده از الگوریتم‌های یادگیری عمیق و داده‌های سنسوری مختلف، عمدتاً توانسته‌اند دقت بالایی در شناسایی و تحلیل چاله‌ها ارائه دهند. با این حال، در زمینه تخمین موقعیت چاله‌ها در سیستم‌های مختصات جهانی چالش‌های قابل توجهی وجود دارد. به‌ویژه تا کنون سیستم‌هایی که بتوانند به صورت خودکار و بدون نیاز به وارد کردن اطلاعات دستی توسط کاربر، همزمان شناسایی و موقعیت‌یابی چاله‌ها را در سیستم مختصات جهانی انجام دهند، توسعه نیافته‌اند. این ضرورت، لزوم طراحی یک سیستم جامع و هوشمند را برجسته می‌کند.

هدف این پژوهش، ارائه روشی نوین مبتنی بر الگوریتم‌های یادگیری عمیق و استفاده از داده‌های سنسوری دوربین، GPS و IMU است که بتواند به‌طور همزمان قابلیت

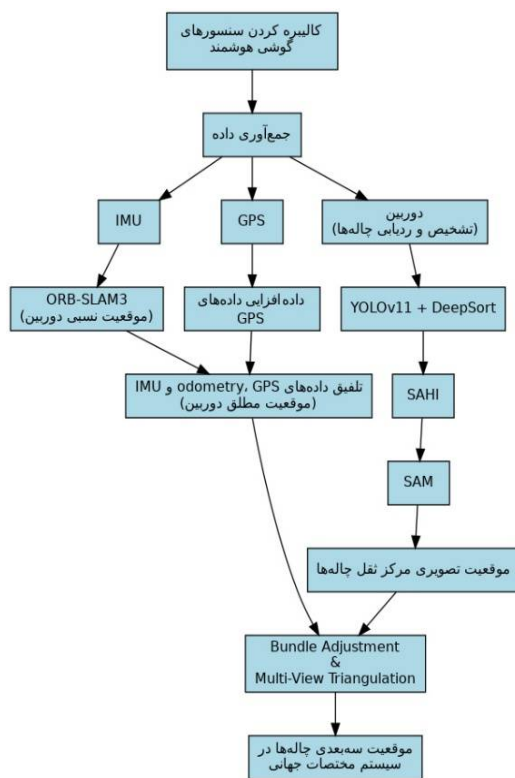
شناسایی چاله‌ها و تخمین موقعیت مطلق آن‌ها را در سیستم مختصات جهانی فراهم کند. این رویکرد با بهره‌گیری از سنسورهای موجود در یک گوشی هوشمند به‌جای تجهیزات پیشرفته و گران‌قیمت، هزینه‌های شناسایی و تحلیل معابر شهری را به‌طور چشمگیری کاهش می‌دهد. دستاورد این پژوهش شامل توسعه سیستمی مقرون‌به‌صرفه، کارآمد و قابل حمل برای بهبود ایمنی معابر و مدیریت شهری است که می‌تواند به‌سادگی در مقیاس گسترده پیاده‌سازی شود و گامی مؤثر در جهت خودکارسازی فرآیندهای مرتبط با زیرساخت‌های شهری باشد.

در بخش دوم این مقاله، روش پیشنهادی برای تخمین موقعیت سه‌بعدی چاله‌ها شرح داده می‌شود. این بخش شامل توضیح الگوریتم‌ها و فرآیندهای استفاده‌شده برای دستیابی به موقعیت دقیق چاله‌ها در سیستم مختصات جهانی است. در بخش سوم، مراحل پیاده‌سازی روش پیشنهادی توضیح داده می‌شود. در بخش چهارم، دقت، کارایی و نتایج حاصل از اجرای روش پیشنهادی مورد بررسی قرار می‌گیرد. نهایتاً در بخش پنجم این پژوهش، نتیجه‌گیری کلی ارائه می‌شود که شامل مرور دستاوردهای اصلی، بررسی محدودیت‌ها و پیشنهادهایی برای تحقیقات آینده است.

۲- روش پیشنهادی

در این پژوهش، یک روش نوین مبتنی بر یادگیری عمیق برای تخمین موقعیت مطلق چاله‌ها در خیابان‌های شهری ارائه شده است. این روش با ترکیب داده‌های چندحسگری شامل ویدئو، GPS و IMU، و بهره‌گیری از الگوریتم‌های پیشرفته پردازش تصویر و یادگیری عمیق عمل می‌کند. در ابتدا داده‌های مورد نیاز با استفاده از سیستم جمع‌آوری داده مجهز به دوربین و حسگرهای مورد نیاز برداشت می‌شوند و با دقت زمانی بالا همگام‌سازی می‌گردند. سپس با استفاده از الگوریتم ORB-SLAM، تخمین موقعیت محلی دوربین انجام شده و اطلاعات مسیر حرکت استخراج می‌شود. در ادامه مدل YOLO11 [39] به منظور شناسایی چاله آموزش داده شد. تصاویر مورد استفاده از منابع موجود در اینترنت گردآوری شدند. تعداد تصاویر آموزش ۶۲۳۵، ۴۰۰ تصویر ولیدیشن و ۲۵۰ تصویر تست بوده است که با استفاده از داده افزایی (تغییر روشنایی تصاویر، چرخش تصاویر و آینه کردن تصاویر در دو جهت عمودی و افقی) به ۱۸۷۰۵ تصویر

آموزش، ۱۲۰۰ تصویر ولیدیشن و ۷۵۰ تصویر تست افزایش پیدا کرد سپس با استفاده از پایتون مدل آموزش دیده به‌دست آمد. مدل YOLO11 [39] برای شناسایی چاله‌ها در تصاویر ویدئویی به کار گرفته شده و با استفاده از روش‌های ردیابی نظیر DeepSORT [40]، چاله‌ها در طول زمان دنبال می‌شوند. از طرفی دیگر با استفاده از الگوریتم‌های Odometry، موقعیت‌های دوربین هنگام اخذ فریم‌های کلیدی به صورت محلی تخمین زده می‌شوند و سپس با استفاده از داده‌های GPS و IMU به موقعیت مطلق تبدیل می‌شوند. به منظور تلفیق داده‌های IMU، GPS و odometry از ISAM2 استفاده شد. در نهایت با استفاده از مثلث‌بندی (Multi-View Triangulation) [41]، موقعیت سه‌بعدی مطلق چاله‌ها تخمین زده می‌شوند.



شکل ۱ - فلوجارت روش پیشنهادی

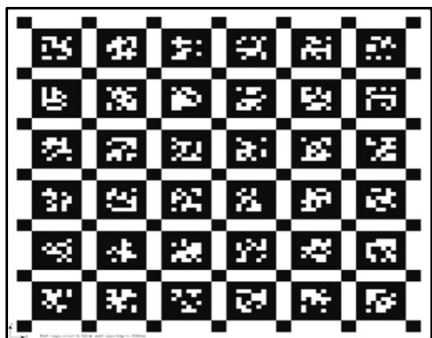
هسته اصلی روش پیشنهادی این پژوهش بر مبنای ادغام اطلاعات از منابع چندگانه حسگری طراحی شده است؛ منابعی شامل تصاویر ویدئویی، داده‌های موقعیت جغرافیایی (GPS) و اطلاعات حسگرهای اینرسی (IMU). هدف از این رویکرد، دستیابی به مکان‌یابی دقیق چاله‌های خیابانی با استفاده از ترکیب هوشمند داده‌های مختلف و کاهش خطاهای سیستماتیک در فرآیند تشخیص است.

ابتدا تصاویر ویدیویی ثبت شده از مسیرهای متنوع تحت شرایط مختلف محیطی، با بهره‌گیری از الگوریتم‌های پردازش تصویر و بینایی ماشین مورد تحلیل قرار می‌گیرند تا ویژگی‌های بصری مرتبط با چاله‌ها، از جمله شکل و مرکز ثقل چاله‌ها استخراج شوند. سپس داده‌های موقعیت‌یاب GPS و IMU شامل شتاب و زاویه چرخش وسیله، به منظور بهبود فرآیند مکان‌یابی و افزایش دقت تشخیص با نتایج تحلیل تصویری ادغام می‌شوند. این ادغام در چارچوب یک مدل بهینه‌سازی چندمنبعی انجام می‌گیرد که با بهره‌گیری از مزایای نسبی هر منبع داده، ضمن کاهش عدم قطعیت و خطای مکان‌یابی، دقت نهایی در تعیین موقعیت مطلق چاله‌های خیابانی را به طور چشمگیری افزایش می‌دهد. در پایان فلوچارتی از مراحل گام‌به‌گام این فرآیند ارائه شده است که نحوه‌ی ترکیب داده‌ها و دست‌یابی به خروجی نهایی را به صورت شفاف نمایش می‌دهد.

۲-۱- کالیبراسیون سنسورهای گوشی هوشمند

کالیبراسیون سنسورهای گوشی هوشمند، از جمله دوربین و IMU، یکی از مراحل اساسی در به‌کارگیری این دستگاه‌ها برای کاربردهای دقیق در سیستم‌های بینایی ماشین و ناوبری است. سنسورهای موجود در گوشی‌های هوشمند به دلیل طراحی تجاری و تنوع در ساختار سخت‌افزاری، معمولاً از دقت کمتری نسبت به تجهیزات تخصصی برخوردارند [۴۶]. برای غلبه بر این چالش و بهبود دقت داده‌های سنسورها، فرآیند کالیبراسیون ضروری است. کالیبراسیون دوربین به تنظیم پارامترهای داخلی و خارجی دوربین می‌پردازد. پارامترهای داخلی شامل فاصله کانونی، نقاط اصلی^۱ و اعوجاج لنز است. پارامترهای خارجی شامل موقعیت و جهت‌گیری دوربین در فضای سه‌بعدی است. ابزارهایی مانند Kalibr [۴۲] و الگوریتم‌های مبتنی بر تخمین ماتریس‌های تبدیل سه‌بعدی [۴۳]، رایج‌ترین روش‌های کالیبراسیون دوربین محسوب می‌شوند. کالیبراسیون IMU شامل تعیین بایاس‌ها و مقیاس‌پذیری داده‌های شتاب‌سنج وژیروسکوپ است. این فرآیند به کاهش خطاهای ناشی از نویز، انحراف طولانی‌مدت و تغییرات حرارتی کمک می‌کند. برای دستیابی به دقت بالاتر، داده‌های خام IMU در شرایط استاندارد و حرکت‌های مشخص جمع‌آوری شده و تحلیل

می‌شوند. سپس پارامترهای اصلاحی از طریق الگوریتم‌هایی مانند تحلیل واریانس آلان (Allan Variance) [۴۳] محاسبه و اعمال می‌شوند. در سیستم‌هایی که داده‌های دوربین و IMU با یکدیگر ترکیب می‌شوند، کالیبراسیون متقابل^۲ بین این دو سنسور نیز اهمیت پیدا می‌کند. این فرآیند شامل تعیین موقعیت نسبی و جهت‌گیری دوربین و IMU نسبت به یکدیگر است. کالیبراسیون دقیق متقابل، به هم‌ترازسازی داده‌های حرکتی و تصویری کمک می‌کند و دقت در تخمین موقعیت و حرکت را بهبود می‌بخشد [۴۵]. در این تحقیق، فرآیند کالیبراسیون سنسورهای گوشی هوشمند شامل IMU و دوربین به صورت دقیق انجام شده است. ابتدا گوشی هوشمند بر روی یک سطح صاف و بدون حرکت قرار داده شد. با استفاده از ابزار SensorServer [۵۳]، داده‌های سنسورهای گوشی از طریق وب‌سوکت روی IP لوکال در دسترس قرار گرفتند. سپس برنامه‌ای در پایتون نوشته شد که اطلاعات IMU را دریافت کرده و در قالب فایل اکسل ذخیره کند. این داده‌ها با استفاده از ابزار Allan Variance در ROS [۵۴] مورد تحلیل قرار گرفتند و پارامترهای Noise Density و Random Walk برای هر دو سنسور شتاب‌سنج وژیروسکوپ تخمین زده شدند. برای کالیبراسیون دوربین، از ابزار Kalibr و تست فیلد Aprilgrid [۵۵] استفاده شد.



شکل ۲ - تست فیلد Aprilgrid

در نهایت پارامترهای دوربین، شامل پارامترهای داخلی و ضرایب اعوجاج تخمین زده شدند. به منظور تخمین ماتریس توجیه نسبی IMU نسبت به دوربین، به داده‌های همگام‌شده IMU و دوربین نیاز بود. بدین منظور یک برنامه اندرویدی به نام Virec توسعه داده شد که می‌تواند داده‌های دوربین و IMU را به طور همزمان و همگام ذخیره کند. با استفاده

^۲ Extrinsic Calibration

^۱ Principal Point

مجدد از ابزار Kalibr، ماتریس توجیه نسبی IMU نسبت به دوربین به دست آمد.

۲-۲- شناسایی اشیا و الگوریتم‌های تشخیص اشیا

شناسایی اشیا یکی از مهم‌ترین چالش‌های پردازش تصویر در سیستم‌های هوش مصنوعی است. این فرآیند شامل تشخیص، طبقه‌بندی و ردیابی اشیا در تصاویر و ویدیوها می‌باشد. در حوزه شناسایی چاله‌های خیابانی، دقت و سرعت این فرآیند از اهمیت بسیار بالایی برخوردار است.

الگوریتم YOLO یکی از پیشرفته‌ترین روش‌های شناسایی اشیا در زمان واقعی است که با رویکردی متفاوت از سایر الگوریتم‌های شناسایی عمل می‌کند. برخلاف روش‌های سنتی که تصویر را به صورت جداگانه تحلیل می‌کنند، YOLO کل تصویر را در یک واحد پردازش می‌نماید. این رویکرد منجر به سرعت بسیار بالا و دقت قابل قبول در شناسایی اشیا می‌شود. در مورد چاله‌های خیابانی، YOLO قادر است با دقت بالایی محدوده و موقعیت چاله‌ها را در تصاویر متحرک شناسایی کند. الگوریتم DeepSort مکمل الگوریتم YOLO در فرآیند ردیابی اشیا است. پس از شناسایی اولیه توسط YOLO، DeepSort وظیفه ردیابی پیوسته و تعقیب شیء در فریم‌های متوالی را بر عهده دارد. این الگوریتم از تکنیک‌های یادگیری عمیق برای تشخیص هویت و حرکت شیء استفاده می‌کند، که در مورد چاله‌ها امکان تعقیب مسیر و تغییرات آنها را فراهم می‌کند. علاوه بر این، ابزار Saha نیز در این سیستم به کار گرفته شده است. Saha به عنوان یک ابزار کمکی، قابلیت‌های اضافی در پردازش و تحلیل داده‌ها ارائه می‌دهد و باعث افزایش کارایی و دقت سیستم در شناسایی و ردیابی اشیا می‌شود.

۲-۳- داده‌افزایی داده‌های GPS

در سیستم‌های چندحسگره، چالش اصلی عدم تقارن در نرخ داده‌برداری حسگرهای مختلف است [۵۱]. به عنوان مثال، حسگرهای IMU با فرکانس بسیار بالا (معمولاً ۱۰۰-۲۰۰ هرتز) داده تولید می‌کنند، در حالی که سنسورهای GPS با نرخ پایین‌تر (حدود ۱ هرتز) عمل می‌کنند. این عدم تقارن می‌تواند منجر به خطاهای محاسباتی در تخمین موقعیت

شود. راهکار اصلی برای مقابله با این چالش، استفاده از تکنیک‌های داده‌افزایی است. در مورد داده‌های GPS، راهبردهای مختلفی برای افزایش حجم و کیفیت داده‌ها وجود دارد:

۱- درون‌یابی خطی میان دو داده متوالی GPS

۲- درون‌یابی مکعبی میان دو داده متوالی GPS

هدف از داده‌افزایی، ایجاد تنوع در مجموعه داده و کاهش وابستگی به داده‌های محدود است. این رویکرد به مدل کمک می‌کند تا قابلیت تعمیم بهتری داشته باشد و در شرایط مختلف محیطی عملکرد مناسبی ارائه دهد [۵۲]. این رویکرد نشان می‌دهد که چگونه می‌توان با استفاده از تکنیک‌های هوشمندانه، محدودیت‌های سنتی در جمع‌آوری داده را پشت سر گذاشت و سیستم‌های مکان‌یابی را به مراتب دقیق‌تر و کارآمدتر طراحی کرد. روش پیشنهادی از چندین مزیت کلیدی برخوردار است. اول، دقت بالا در شناسایی و موقعیت‌یابی چاله‌ها و دوم، قابلیت عملکرد در شرایط مختلف محیطی و سوم، بهره‌گیری حداکثری از اطلاعات چندحسگره. در این پژوهش به منظور داده‌افزایی داده‌های GPS ابتدا داده‌های جغرافیایی (عرض جغرافیایی و طول جغرافیایی) به مختصات در سیستم مختصات جهانی (UTM)^۱ تبدیل می‌شوند، چرا که درون‌یابی در سیستم مختصات کارترین UTM دقیق‌تر است [۵۶]. سپس داده‌های GPS با فرکانس ۱ هرتز به داده‌هایی با فرکانس ۱۰ هرتز تبدیل می‌شوند. این فرآیند از طریق درون‌یابی مکعبی انجام می‌گیرد. درون‌یابی مکعبی یک روش پیشرفته برای تخمین مقادیر میانی یک تابع بین نقاط داده مشخص است. این روش از چندجمله‌ای‌های مکعبی استفاده می‌کند و نسبت به روش‌های ساده‌تری مانند درون‌یابی خطی دقت بیشتری دارد، زیرا می‌تواند انحناهای تابع اصلی را بهتر مدل کند. فرمول عمومی برای درون‌یابی مکعبی در مختصات $x(t)$ و $y(t)$ به صورت زیر است [۶۱]:

$$x(t) = a(t - t_0)^3 + b(t - t_0)^2 + c(t - t_0) + d \quad (1)$$

$$y(t) = a'(t - t_0)^3 + b'(t - t_0)^2 + c'(t - t_0) + d' \quad (2)$$

t_0 زمان اولیه است و t زمان ثانویه است و همچنین ضرایب $a, b, c, d, a', b', c', d'$ از داده‌های مختصات UTM در دو زمان متوالی محاسبه می‌شوند. برای محاسبه این

^۱ Universal Transverse Mercator

ضرایب، باید از داده‌های مختصات UTM در دو زمان متوالی t_1 و t_2 استفاده کنیم. این داده‌ها شامل مختصات x_1 و y_1 در زمان t_1 و مختصات x_2 و y_2 در زمان t_2 هستند. همچنین نیاز است که مشتقات مختلف این توابع نیز محاسبه شوند تا تابع مکعبی به درستی بین این دو نقطه تطبیق داده شود.

۲-۴- تخمین موقعیت بصری

تخمین موقعیت بصری یکی از چالش‌برانگیزترین و در عین حال پرکاربردترین مسائل در حوزه رباتیک و سیستم‌های مکان‌یابی است. این روش به دنبال تعیین موقعیت و حرکت یک سیستم متحرک با استفاده از پردازش تصاویر متوالی است که توسط دوربین‌های نصب شده بر روی سیستم ثبت می‌شوند. در گذشته سیستم‌های مکان‌یابی عمدتاً بر پایه حسگرهای مکانیکی مانند انکودرها، ژيروسکوپ‌ها و یا سیستم‌های موقعیت‌یابی جهانی طراحی می‌شدند. هرچند این روش‌ها مزایای خاص خود را دارند، اما محدودیت‌های قابل توجهی نیز به همراه دارند. به عنوان مثال GPS در محیط‌های بسته مانند تونل‌ها، ساختمان‌ها و یا مناطقی با پوشش ضعیف سیگنال، عملکرد مطلوبی ندارد [۴۷]. همچنین، حسگرهای مکانیکی ممکن است در برابر عوامل محیطی همچون لرزش‌ها و ضربات مقاومت کافی نداشته باشند و معمولاً هزینه‌های بالایی نیز دارند. تخمین موقعیت بصری به عنوان یک جایگزین نوآورانه به این محدودیت‌ها پاسخ داده و روشی مبتنی بر اطلاعات بصری ارائه می‌دهد. در این روش، تصاویر متوالی ثبت‌شده توسط دوربین‌ها پردازش شده و نقاط کلیدی یا ویژگی‌های مهم (مانند لبه‌ها، گوشه‌ها یا الگوهای خاص) از هر تصویر استخراج می‌شوند. سپس این نقاط در تصاویر متوالی ردیابی می‌شوند و تغییرات موقعیت آن‌ها تحلیل می‌شود. اصل اساسی این رویکرد بر مبنای مقایسه تغییرات نقاط کلیدی در تصاویر متوالی استوار است [۴۸]. با تحلیل جایجایی این نقاط و استفاده از الگوریتم‌های هندسی و ریاضی پیشرفته مانند الگوریتم‌های مبتنی بر ماتریس‌های اساسی^۱ یا تطبیق‌های سه‌بعدی، حرکت نسبی دوربین محاسبه می‌شود. در نتیجه، می‌توان مسیر حرکت و موقعیت مکانی سیستم را با دقت بسیار بالایی تخمین زد. علاوه بر این، تخمین موقعیت بصری می‌تواند با دیگر منابع اطلاعاتی مانند داده‌های IMU ترکیب

شود تا دقت بیشتری در موقعیت‌یابی فراهم شود. این روش که به عنوان ادغام داده‌ها یا فیوژن سنسورها شناخته می‌شود [۴۹]، به ویژه در سیستم‌های ناوبری پیشرفته و ربات‌های خودمختار اهمیت ویژه‌ای دارد.

در این پژوهش، به منظور تخمین موقعیت بصری، از ORB-SLAM به عنوان یکی از پیشرفته‌ترین و پرکاربردترین الگوریتم‌های تخمین موقعیت بصری استفاده شده است. انتخاب این مدل به دلیل مزایای برجسته آن در پردازش دقیق و کارآمد تصاویر و سازگاری با انواع محیط‌های داخلی و خارجی صورت گرفته است. این الگوریتم به دلیل استفاده از ویژگی‌های کلیدی مقاوم مانند ORB^2 ، توانایی شناسایی و ردیابی نقاط کلیدی را حتی در شرایطی با نویز زیاد و تغییرات نور حفظ می‌کند. همچنین این الگوریتم از روش‌هایی مانند مدیریت مجدد نقاط کلیدی و حذف نقاط متحرک برای افزایش دقت تخمین موقعیت در محیط‌های پویا بهره می‌برد و همچنین قابلیت استفاده از داده‌های تک‌دوربین^۳ را دارد.

در این تحقیق از نسخه ORB-SLAM3 استفاده شده است که داده‌های دوربین در ROS [Δλ] منتشر شده و ORB-SLAM از آن استفاده می‌کند. این الگوریتم نقاط کلیدی را از تصاویر استخراج کرده و مسیر حرکت را تخمین می‌زند. خروجی این الگوریتم حاوی اطلاعات موقعیت نسبی و چرخش دوربین در طول مسیر حرکت است. این داده‌ها به منظور تبدیل به مختصات UTM و استفاده در تخمین سه‌بعدی موقعیت چاله‌ها مورد استفاده قرار گرفته‌اند.

۲-۵- تلفیق داده‌های Odometry، GPS و IMU

در سیستم‌های مکان‌یابی، هر حسگر داده‌های خود را با خطاها و نویزهای خاصی تولید می‌کند. استفاده از یک منبع داده به تنهایی می‌تواند منجر به تخمین‌های نادقیق و غیرقابل اطمینان شود. تلفیق داده‌ها راهکاری است که به منظور کاهش خطاها و افزایش دقت تخمین موقعیت طراحی شده است. هدف اصلی از تلفیق داده‌ها، ترکیب اطلاعات از منابع مختلف حسگری مانند GPS، IMU و تصاویر ویدیویی است [۴۹]. با استفاده از الگوریتم‌های پیشرفته مانند فیلتر کالمن [۵۰]، می‌توان خطاهای تک‌تک حسگرها را کاهش داد و یک تخمین دقیق‌تر از موقعیت ارائه داد. این رویکرد به سیستم اجازه می‌دهد تا نقاط ضعف هر حسگر را با استفاده از نقاط

قوت سایر حسگرها جبران کند. انتخاب تخمین موقعیت بصری و تلفیق داده‌ها برای شناسایی چاله‌های خیابانی بر اساس چندین دلیل کلیدی صورت گرفته است:

- ۱) دقت بالا: این روش‌ها قادر به ارائه تخمین موقعیت با دقت بسیار بالاتر نسبت به روش‌های سنتی هستند.
- ۲) انعطاف‌پذیری: قابلیت عملکرد در شرایط مختلف محیطی و با استفاده از حداقل تجهیزات سخت‌افزاری.
- ۳) هزینه پایین: استفاده از گوشی هوشمند به عنوان پلتفرم اصلی، هزینه‌های سیستم را به شدت کاهش می‌دهد.
- ۴) قابلیت مقیاس‌پذیری: امکان توسعه و تعمیم روش به سایر کاربردهای مکان‌یابی و پردازش تصویر.

در این پژوهش از iSAM2 [۵۷] به عنوان روش تلفیق داده‌های odometry، IMU و GPS استفاده شده است. iSAM2 تلفیق داده‌ها را با استفاده از مدل‌های احتمال و بهینه‌سازی گرافی انجام می‌دهد. داده‌های IMU برای پیش‌بینی وضعیت، داده‌های GPS برای تصحیح مطلق و داده‌های Odometry برای تصحیح نسبی استفاده می‌شوند. این الگوریتم به صورت تدریجی و بهینه اطلاعات را ترکیب می‌کند و تخمین دقیقی از وضعیت دوربین در فریم‌های کلیدی می‌دهد. هر یک از این منابع داده نقش کلیدی در تخمین موقعیت دارند، اما دقت و اطمینان آن‌ها متفاوت است. داده‌های Odometry که از الگوریتم ORB-SLAM3 استخراج شده‌اند، در محیط‌های شهری دقت بالایی دارند و به همین دلیل وزنی قابل توجه در فرآیند تلفیق داده‌ها دریافت می‌کنند. این داده‌ها اطلاعات نسبی حرکت سیستم را به‌طور دقیق ارائه می‌دهند و در کوتاه‌مدت خطای کمی دارند. داده‌های IMU که از حسگرهای داخلی گوشی هوشمند جمع‌آوری شده‌اند، به دلیل طراحی غیرصنعتی حسگر و وابستگی زیاد به شرایط محیطی، ممکن است پارامترهای کالیبراسیون آن‌ها تغییر کرده و در نتیجه نویز و خطای جمعی ایجاد شود. با این حال، اطلاعات لحظه‌ای این داده‌ها (مانند شتاب و تغییرات زاویه) می‌تواند به پر کردن فواصل میان داده‌های Odometry کمک کند و وزن آن‌ها در فرآیند تلفیق با توجه به این ویژگی تعیین شده است. داده‌های GPS نیز در این تحقیق دقتی در حدود چند متر دارند که در محیط‌های شهری به دلیل بازتاب سیگنال‌ها از ساختمان‌ها و موانع دیگر با خطا همراه هستند. وزن این داده‌ها به دلیل عدم دقت مطلق کمتر از Odometry و IMU

تعیین شده است، اما به‌عنوان یک منبع موقعیت مطلق برای کاهش خطای تجمعی Odometry و IMU، نقش حیاتی ایفا می‌کنند. در فرآیند تلفیق، وزن‌دهی به داده‌ها از طریق تنظیم ماتریس‌های کوواریانس انجام می‌شود. داده‌های Odometry با توجه به دقت بالای خود، کوواریانس کمتری دارند. داده‌های IMU به دلیل حساسیت به شرایط محیطی و نویز حسگر گوشی، کوواریانس بالاتری دارند و داده‌های GPS با توجه به خطای چند متری، وزن متوسطی دریافت می‌کنند. این تنظیمات باعث می‌شود که iSAM2 بتواند با ترکیب داده‌های نسبی و مطلق، تخمینی دقیق و پایدار از موقعیت سیستم در محیط‌های شهری ارائه دهد.

۲-۶- دستیابی به موقعیت سه‌بعدی چاله‌ها

برای دستیابی به موقعیت سه‌بعدی اشیاء در سیستم مختصات UTM، زمانی که مختصات تصویری شی و موقعیت سه‌بعدی دوربین در دسترس است، از مدل‌سازی هندسی دوربین و روش‌های بازسازی سه‌بعدی استفاده می‌شود. بنابراین به‌منظور تعیین مختصات سه‌بعدی شی، نیاز به مشاهدات تصویری از چندین موقعیت (موقعیت‌های مختلف دوربین) است. این مشاهدات منجر به ایجاد مجموعه‌ای از معادلات می‌شود که با حل آن‌ها می‌توان مختصات سه‌بعدی شی را محاسبه کرد. یکی از روش‌های رایج برای حل این معادلات، روش Multiview Triangulation [۴۱] است. در این روش، از روابط هندسی بین موقعیت‌های تصویری مربوط به یک شی و موقعیت دوربین‌ها استفاده می‌شود تا نقطه تقاطع مسیرهای تصویری در فضای سه‌بعدی محاسبه شود. این تقاطع نشان‌دهنده مختصات سه‌بعدی شی در فضای مورد نظر است. ابتدا با استفاده از پارامترهای کالیبراسیون دوربین و همچنین مقادیر چرخش و انتقال داده‌های Odometry برای هر فریمی که چاله در آن وجود دارد، ماتریس پروجکشن ساخته می‌شود. ماتریس P_i نقاط را از فضای سه‌بعدی به فضای دوبعدی تصویر می‌برد [۶۲].

$$P_i = K \cdot [R_i | t_i] \quad (3)$$

که در آن K ماتریس پارامترهای داخلی دوربین شامل فاصله کانونی و مختصات نقطه اصلی، R_i ماتریس چرخش و t_i بردار انتقال مربوط به فریم مربوطه است [۶۳].

برای بازسازی مختصات سه‌بعدی یک نقطه از مختصات دوبعدی آن در تصاویر، ابتدا با استفاده از مختصات تصویری و ماتریس پروژکشن دوربین‌ها (که شامل پارامترهای داخلی، چرخش و انتقال است)، یک معادله خطی ساخته می‌شود. این معادله خطی (معادله ۵) برای تمامی فریم‌هایی که نقطه در آن مشاهده شده است، جمع شده و یک سیستم معادلات تشکیل می‌دهد. سپس با حل این سیستم با استفاده از SVD، مختصات سه‌بعدی نقطه به دست می‌آید. این مختصات به صورت همگن محاسبه شده و در نهایت برای تبدیل به مختصات کارتزین نرمال‌سازی می‌شود. این فرآیند برای تمامی نقاط موردنظر تکرار می‌شود.

در این تحقیق ابتدا با استفاده از روش Multiview Triangulation، مختصات اولیه سه‌بعدی چاله‌ها محاسبه شد. به دلیل آنکه داده‌های تصویری و موقعیت‌های دوربین برای هر چاله از چندین موقعیت در فریم‌های مختلف جمع‌آوری شده بودند، تعداد معادلات موجود برای هر چاله قابل توجه بود. این موضوع امکان استفاده از روش Bundle Adjustment را فراهم کرد تا با بهینه‌سازی مقادیر اولیه، خطاها به حداقل برسند و موقعیت سه‌بعدی چاله‌ها در سیستم مختصات UTM با دقت بالاتری به دست آید.

۳- پیاده‌سازی

در این پژوهش از ترکیب الگوریتم‌های پیشرفته یادگیری عمیق و روش‌های نوین پردازش داده‌های چندحسگری برای افزایش دقت و کارایی استفاده شده است. فرآیند کار با کالیبراسیون دقیق سنسورهای گوشی هوشمند شامل دوربین و IMU آغاز گردید تا داده‌های جمع‌آوری شده از کیفیت و دقت مطلوب برخوردار باشند. پس از کالیبراسیون، داده‌ها با نرخ‌های زمانی مختلف توسط سنسورها جمع‌آوری و برای تحلیل آماده شدند. مراحل پردازش شامل تقسیم‌بندی دقیق تصاویر با استفاده از الگوریتم‌های یادگیری عمیق، استخراج مختصات مرکز چاله‌ها، و تطبیق این مختصات در فریم‌های مختلف است. همچنین از تلفیق داده‌های چندحسگری برای تخمین موقعیت سه‌بعدی نهایی چاله‌ها استفاده شد. در این تحقیق از گوشی هوشمند Samsung S7 به‌عنوان منبع اصلی داده‌های حسگری استفاده شده است.

$$K = \begin{bmatrix} fx & 0 & cx \\ 0 & fy & cy \\ 0 & 0 & 1 \end{bmatrix} \quad (۴)$$

سپس برای یافتن مختصات سه‌بعدی چاله (X) ، از رابطه زیر استفاده می‌شود [۶۲]:

$$A \cdot X = 0 \quad (۵)$$

که در آن ماتریس A از اطلاعات ماتریس پروژکشن و به صورت زیر محاسبه می‌شود [۶۲]:

$$\begin{aligned} A_{2i} &= x_i P_{i3} - P_{i1} \\ A_{2i+1} &= y_i P_{i3} - P_{i2} \end{aligned} \quad (۶)$$

x_i و y_i مختصات تصویری نقطه در فریم i و P_{i1} ، P_{i2} ، P_{i3} ردیف‌های ماتریس پروژکشن متناظر با آن فریم هستند. معادله ۵ با استفاده از روش SVD^۱ حل می‌شود و یک تخمین اولیه از مختصات سه‌بعدی نقطه X ارائه می‌دهد. زمانی که R (ماتریس چرخش) و t (بردار انتقال) از تلفیق داده‌های GPS و Odometry تخمین زده می‌شوند، دقت این پارامترها به دقت GPS وابسته است. داده‌های GPS ممکن است جابجایی‌هایی در حد چندین متر ایجاد کنند که می‌تواند منجر به چرخش و انتقال اشتباه در ماتریس پروژکشن شود. برای کاهش این خطاها و بهبود دقت بازسازی سه‌بعدی، از روش Bundle Adjustment [۶۲] استفاده می‌شود. این روش بهینه‌سازی غیرخطی است که تمامی نقاط تصویری، موقعیت دوربین‌ها و پارامترهای داخلی دوربین را به صورت همزمان در نظر می‌گیرد و خطای بازپروژکشن^۲ را کمینه می‌کند. استفاده از این روش باعث می‌شود مختصات سه‌بعدی اشیاء با دقت بالاتری تعیین شود. هدف این است که خطای بازپروژکشن کمینه شود [۶۲].

$$x'_i = P_i \cdot X = K \cdot [R_i | t_i] \cdot X \quad (۷)$$

$$\text{reprojection error} = \sum_{i=1}^n \| x_i - x'_i \|^2 \quad (۸)$$

که در آن x_i مختصات تصویری واقعی نقطه در فریم i و x'_i مختصات تصویری پیش‌بینی شده نقطه در آن فریم است. پس از بهینه‌سازی خطای بازپروژکشن با استفاده از Bundle Adjustment، مختصات سه‌بعدی نقاط و همچنین پارامترهای ماتریس چرخش و بردار انتقال بهبود می‌یابند.

^۲ reprojection error

^۱ Singular Value Decomposition

۳-۱- تجهیزات و داده‌های مورد استفاده

در این روش تنها به یک گوشی هوشمند نیاز است که از سنسورهای داخلی آن شامل دوربین، GPS و IMU استفاده می‌شود. تمامی داده‌های مورد نیاز این تحقیق با استفاده از گوشی هوشمندی که بالای یک خودرو در حال حرکت نصب شده بود، جمع‌آوری شدند. در جریان برداشت داده‌ها، خودرو با سرعت ۵۰ کیلومتر بر ساعت مسیری به طول حدود ۵۰ متر را طی کرد. در این فرآیند، دوربین با فریم‌ریت ۶۰ فریم بر ثانیه موفق به ثبت ۲۱۶ فریم از مسیر شد که برای پردازش مورد استفاده قرار گرفت. دوربین گوشی تصاویر متوالی از محیط ثبت کرده، سنسور IMU اطلاعات مربوط به شتاب و تغییرات زاویه را فراهم آورده و GPS گوشی داده‌های موقعیت مکانی را ارائه داده است.



شکل ۳- نحوه قرارگیری دوربین در خودرو

به منظور جمع‌آوری داده‌های واقعی و ارزیابی دقت نتایج، از دستگاه GPS مولتی‌فرکانس مدل G1 Plus Sout با دقت ۰.۰۱۲ متر استفاده شده است. این دستگاه برای ثبت موقعیت دقیق چاله‌ها به کار رفته و با قرار دادن آن روی هر چاله، موقعیت دقیق آن با دقت سانتی‌متری تعیین شده است. این داده‌ها به عنوان مقیاس مرجع برای ارزیابی عملکرد سیستم به کار رفته‌اند.

۳-۲- کالیبراسیون دقیق سنسورهای گوشی هوشمند

کالیبراسیون سنسورهای گوشی هوشمند به منظور بهبود دقت داده‌های ثبت شده انجام می‌شود. بدین منظور از ابزارهای تخصصی استفاده گردید.

۳-۲-۱- کالیبراسیون پارامترهای دوربین

با استفاده از Kalibr، پارامترهای دوربین (پارامترهای داخلی و اعوجاج) به صورت دقیق کالیبره شدند. این مرحله به

بهبود کیفیت داده‌های بصری و افزایش دقت پردازش‌های بعدی کمک می‌کند. در جدول زیر، پارامترهای کالیبراسیون دوربین این گوشی ارائه گردیده‌اند تا دقت فرایند تخمین تضمین شود. پارامترهای دوربین شامل پارامترهای داخلی و ضرایب اعوجاج، مطابق جدول زیر تخمین زده شدند:

جدول ۱- پارامترهای کالیبراسیون دوربین

مدل دوربین	Pinhole
فاصله کانونی (fx, fy)	۱۴۹۲.۹۳ / ۱۴۹۶.۱۸
نقطه اصلی (cx, cy)	۹۷۲.۶۹ / ۵۷۱.۱۶
مدل اعوجاج	Radtan
ضرایب اعوجاج	-۰.۰۰۰۴۷ / ۰.۰۰۰۲۶ / -۰.۰۰۲۱۴ / -۰.۰۰۰۲۶

۳-۲-۲- کالیبراسیون پارامترهای IMU

به منظور کالیبراسیون پارامترهای اینرسی، از ابزار Allan Variance استفاده شد. این ابزار، نویزها و خطاهای سیستماتیک موجود در داده‌های شتاب‌سنج وژیروسکوپ را تحلیل کرده و مدل‌های نویز مناسب را ارائه می‌دهد. این کالیبراسیون، دقت داده‌های حرکتی را به شکل قابل توجهی افزایش می‌دهد. در ادامه به صورت کامل در جدول آورده شده‌اند.

جدول ۲- پارامترهای کالیبراسیون imu

شتاب سنج	Noise Density	Random Walk
ژیروسکوپ <td>۰.۰۰۰۲۰۳۲</td> <td>۰.۰۰۰۰۰۰۱</td>	۰.۰۰۰۲۰۳۲	۰.۰۰۰۰۰۰۱
شتاب سنج <td>۰.۰۰۰۵۲۴۲۲</td> <td>۰.۰۰۰۰۲۸۵۵</td>	۰.۰۰۰۵۲۴۲۲	۰.۰۰۰۰۲۸۵۵

۳-۲-۳- تعیین موقعیت نسبی IMU نسبت به دوربین

در مرحله نهایی، مجدداً از Kalibr برای تخمین موقعیت و جهت نسبی IMU نسبت به دوربین استفاده شد. این مرحله با استفاده از ماتریس T_cam_imu انجام گرفت و برای هماهنگی کامل داده‌های بصری و اینرسی به کار گرفته شد.

جدول ۳- ماتریس موقعیت نسبی imu نسبت به دوربین

-۰.۰۰۲۴۳	-۰.۹۹۹۹۸	۰.۰۰۴۹۸	-۰.۰۲۴۳۶
-۰.۹۹۹۸۲	۰.۰۰۲۵۲	۰.۰۱۸۷۳	۰.۰۱۷۸۲
-۰.۰۱۸۷۴	-۰.۰۰۴۹۳	-۰.۹۹۹۸۱	۰.۱۲۸۹۴
۰.۰	۰.۰	۰.۰	۱.۰



شکل ۵- شناسایی مرکز ثقل چاله با استفاده از YOLO و SAM

۳-۵- تخمین موقعیت سه‌بعدی چاله‌ها با استفاده از داده‌های چندحسگری

در ادامه، الگوریتم ORB-SLAM برای استخراج داده‌های موقعیتی دوربین از ویدئو به کار گرفته شد. این الگوریتم با استفاده از اطلاعات تصویری و حرکتی دوربین، مسیر حرکت دوربین را تخمین زده و ذخیره می‌کند. داده ای GPS نیز پس از جمع‌آوری، به کمک روش‌های داده‌افزایی بهبود داده شدند تا خطاهای احتمالی در آن‌ها کاهش یابد و همچنین داده‌های کافی در دسترس باشد. برای افزایش دقت و پایداری این داده‌ها و همچنین تبدیل موقعیتی نسبی دوربین در فریم‌های متوالی به موقعیت در سیستم مختصات UTM، الگوریتم iSAM2 به‌منظور تلفیق داده‌های ORB-SLAM و اطلاعات GPS و IMU مورد استفاده قرار گرفت. این ترکیب، داده‌های موقعیتی دقیق‌تر و سازگارتری را برای مراحل بعدی در سیستم مختصات UTM فراهم آورد. سپس با استفاده از مختصات تصویری چاله‌ها و داده‌های تلفیقی حاصل شده از الگوریتم iSAM2، فرآیند تخمین موقعیت سه‌بعدی چاله‌ها آغاز شد. در این مرحله، روش Multiview Triangulation به‌منظور انتقال مختصات تصویری چاله‌ها به فضای سه‌بعدی به کار گرفته شدند.

جدول ۴- شاخص‌های ارزیابی مدل

YOLO + Sahi	YOLO	
۰.۹۲	۰.۸۰	دقت
۰.۶۸	۰.۵۳	بازخوانی
۰.۸۲	۰.۶۲	mAP50
۰.۶۶	۰.۳۵	mAP50-95

۳-۶- بهینه‌سازی موقعیت سه‌بعدی چاله‌ها

برای کاهش خطاهای احتمالی و بهبود دقت مختصات تخمین‌زده‌شده، از روش بهینه‌سازی Bundle Adjustment

۳-۳- شناسایی و ردیابی چاله‌ها در فریم‌های متوالی

در مرحله نخست، از الگوریتم YOLO برای شناسایی چاله‌ها در ویدئوها استفاده شد. این الگوریتم با دقت بالا توانست موقعیت تصویری چاله‌ها را شناسایی کرده و مختصات آن‌ها را به‌صورت دویبعدی ذخیره کند. برای بهبود دقت شناسایی، الگوریتم Sahi نیز به کار گرفته شد که توانایی تشخیص جزئیات کوچک‌تر چاله‌ها در تصاویر با رزولوشن بالا را فراهم می‌سازد. علاوه بر این، برای ردیابی چاله‌ها در طول ویدئو و اختصاص یک شناسه منحصربه‌فرد به هر چاله، از الگوریتم DeepSORT استفاده گردید. این الگوریتم با بهره‌گیری از ترکیب اطلاعات حرکتی و ویژگی‌های ظاهری، توانست چاله‌ها را در چندین فریم ردیابی کرده و موقعیت آن‌ها را در طول زمان مشخص کند.



شکل ۴- شناسایی چاله با استفاده از YOLO

۳-۴- تشخیص مرکز ثقل چاله‌ها و تعیین موقعیت تصویری آن‌ها

به‌منظور تعیین دقیق‌تر موقعیت تصویری چاله‌ها از الگوریتم SAM استفاده شد. این الگوریتم با تقسیم‌بندی دقیق چاله‌ها در تصاویر و استخراج مرکز ثقل هر چاله، مختصات مرکز آن را به‌عنوان موقعیت تصویری نهایی در نظر گرفت. این رویکرد نه‌تنها باعث افزایش دقت موقعیتیابی می‌شود، بلکه امکان ارائه یک مبنای قابل اعتماد برای انتقال مختصات تصویری به فضای سه‌بعدی را فراهم می‌کند. همچنین لازم است هر چاله به‌عنوان یک نقطه‌ی واحد در نظر گرفته شود که این نقطه در تمامی فریم‌هایی که چاله در آن‌ها شناسایی می‌شود، به‌طور دقیق و یکنواخت انتخاب گردد. این امر تضمین می‌کند که مختصات تصویری هر چاله در فریم‌های مختلف سازگار باشد و به درستی برای پردازش‌های بعدی استفاده شود.

استفاده گردید. این روش با تلاش بر کمینه کردن خطای Reprojection Error و تصحیح مختصات سه‌بعدی، مختصات نهایی هر چاله را با دقت بسیار بالا ارائه می‌دهد. ترکیب این روش با الگوریتم‌های قبلی باعث شد تا مختصات سه‌بعدی تخمین‌زده‌شده به واقعیت نزدیک‌تر باشد و از اثرات نویز و خطاهای سنسوری و به خصوص سنسور GPS کاسته شود.

۴- نتایج و ارزیابی

روش پیشنهادی توانست چاله‌ها را با دقت و صحت بالا در محیط‌های شهری شناسایی و موقعیت تصویری آن‌ها را به طور دقیق مشخص کند. مدل YOLO در آزمایش‌های انجام‌شده عملکرد بسیار قابل قبولی داشت و توانست در شرایط مختلف محیطی و نوری، چاله‌ها را با دقت بالا شناسایی کند. شاخص‌های ارزیابی این مدل، شامل دقت، فراخوانی نشان‌دهنده بهبود قابل توجهی نسبت به روش‌های قبلی بودند. این نتایج بیانگر توانایی مدل در شناسایی صحیح و کاهش هشدارهای نادرست است که اهمیت بالایی در کاربردهای عملی دارد. استفاده از الگوریتم DeepSORT، امکان ردیابی دقیق چاله‌های شناسایی‌شده در چندین فریم ویدئویی را فراهم کرد. این الگوریتم با اختصاص شناسه‌های منحصر‌به‌فرد به هر چاله، توانست از ادغام یا اشتباه گرفتن چاله‌های مختلف در فریم‌های متوالی جلوگیری کند. این ویژگی به‌ویژه در شرایطی که چاله‌ها در نزدیکی یکدیگر قرار دارند یا دوربین در حال حرکت سریع است، نقش حیاتی ایفا کرد و باعث افزایش دقت و انسجام نتایج شد. علاوه بر این، استفاده از ابزار Sahi به طور مؤثری توانست چاله‌های کوچک‌تر یا چاله‌هایی را که در مرز دید دوربین قرار داشتند را شناسایی کند. این ابزار با بهره‌گیری از تقسیم تصاویر به زیرنواحی و تمرکز بیشتر بر جزئیات، موفق به شناسایی مناطقی شد که ممکن است در روش‌های معمول نادیده گرفته شوند. در جدول زیر مقایسه هر دو روش یعنی تشخیص فقط با استفاده از YOLO و تشخیص با استفاده از YOLO و استفاده همزمان از Sahi را نشان می‌دهد.

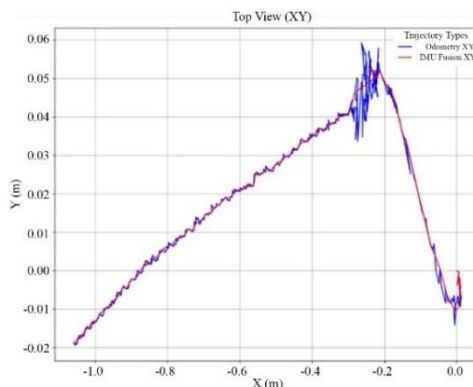
استفاده همزمان Sahi در کنار YOLO باعث افزایش قابل توجه دقت و کارایی در شناسایی چاله‌ها شده است. یکی از دلایل اصلی این بهبود، توانایی Sahi در تقسیم تصاویر به بخش‌های کوچک‌تر و پردازش دقیق‌تر هر بخش است که منجر به شناسایی بهتر جزئیات می‌شود. الگوریتم YOLO

به‌تنهایی در تشخیص اشیاء کوچک ضعف دارد. از آنجا که چاله‌ها اغلب به‌عنوان اشیاء کوچک در تصاویر ظاهر می‌شوند، با استفاده از Sahi، این محدودیت برطرف شده و عملکرد مدل بهبود یافته است. نتایج نشان می‌دهد که با ترکیب Sahi، دقت از ۰.۸۰ به ۰.۹۲ و mAP50-95 از ۰.۳۵ به ۰.۶۶ افزایش یافته است که نشان‌دهنده بهینه‌سازی قابل توجه این روش برای شناسایی اشیاء کوچک مانند چاله‌ها است.

ترکیب داده‌های Odometry با داده‌های IMU در این پژوهش نقشی کلیدی در بهبود دقت و جامعیت نتایج سیستم تشخیص و ردیابی چاله‌ها ایفا کرد. داده‌های Odometry که از الگوریتم ORB-SLAM استخراج شده‌اند، مسیر حرکتی دوربین را شبیه‌سازی می‌کنند اما به دلیل محدودیت‌های ذاتی ممکن است در شرایط خاص مانند تغییرات ناگهانی حرکت یا از دست دادن فریم‌های کلیدی دچار خطا شوند. داده‌های IMU، شامل اطلاعات دقیق شتاب‌سنج و ژيروسکوپ، به‌عنوان مکملی حیاتی برای کاهش این خطاها به‌کار گرفته شده‌اند. با ترکیب این دو منبع داده، دقت موقعیت‌یابی سیستم به طور چشمگیری افزایش یافت. داده‌های IMU امکان ثبت تغییرات ناگهانی دوربین را فراهم کردند که در داده‌های Odometry ممکن است به درستی بازتاب داده نشود. این ویژگی به‌ویژه در شرایطی که دوربین در معرض حرکات سریع یا لرزش‌های ناگهانی قرار دارد، اهمیت بیشتری پیدا می‌کند. به‌علاوه استفاده از داده‌های IMU باعث شد که خطاهای تجمعی در داده‌های Odometry به‌طور قابل توجهی کاهش یابد. یکی از نتایج قابل توجه این ترکیب، افزایش اطمینان‌پذیری کلی سیستم است. سیستم توانست در محیط‌های پیچیده‌تر و با موانع بیشتر، چاله‌ها را با دقت بالاتری شناسایی و ردیابی کند. این ویژگی از آنجا ناشی می‌شود که داده‌های IMU به سیستم امکان داد تا حرکات سه‌بعدی دوربین را بهتر مدل‌سازی کند و اطلاعات اضافی درمورد چرخش‌ها و شتاب‌ها را به داده‌های Odometry بیفزاید. علاوه بر این، این ترکیب به سیستم اجازه داد تا در طولانی‌مدت پایداری بیشتری داشته باشد و خطاهای ناشی از انباشت داده‌ها را به حداقل برساند.

در هنگام دوران، داده‌های odometry (آبی) معمولاً دچار خطای بیشتری می‌شوند، زیرا در چرخش‌های شدید یا تغییرات سریع جهت به دلیل محدودیت در شناسایی نقاط متناظر کافی، باعث انحراف و ناپایداری در مسیر محاسبه‌شده می‌شود.

منظور دستیابی به مختصات سه‌بعدی دقیق چاله‌ها در سیستم مختصات جهانی طراحی و اجرا شد. برای هر چاله، تعدادی موقعیت تصویری در فریم‌های مختلف ثبت شد که این نقاط با استفاده از الگوریتم‌های بینایی ماشین و ردیابی اشیاء استخراج شدند. علاوه بر این برای هر فریم تصویری، مختصات دوربین که پیش‌تر با استفاده از Odometry محاسبه شده بود، در دسترس بود. اطلاعات GPS نیز به‌عنوان داده مرجع افزوده، امکان نگاشت مختصات محلی دوربین‌ها به مختصات UTM را فراهم کرد. جهت تخمین موقعیت سه‌بعدی از روش Multiview Triangulation استفاده شد. این روش با بهره‌گیری از مختصات تصویری چاله‌ها و مختصات دوربین‌ها در سیستم مختصات UTM، مجموعه‌ای از معادلات خطی را برای هر نقطه تصویری ایجاد کرد. برای حل این معادلات و تخمین مختصات سه‌بعدی دقیق، تمامی داده‌های تصویری و موقعیت دوربین‌ها به‌طور همزمان در فرآیند Bundle Adjustment وارد شدند. این روش با اصلاح و بهینه‌سازی کلیه نقاط و روابط هندسی، دقت تخمین را به میزان قابل توجهی افزایش داد و اثر نویز یا عدم تطابق داده‌ها را کاهش داد. استفاده همزمان از Multiview Triangulation و Bundle Adjustment در این پژوهش باعث شد که موقعیت سه‌بعدی هر چاله با دقت بالایی در سیستم مختصات UTM محاسبه شود. این فرآیند نه‌تنها مختصات سه‌بعدی دقیق‌تر و قابل اعتمادتری ارائه داد، بلکه از بروز خطاهای ناهمخوانی و کاهش کیفیت در نتیجه نهایی جلوگیری کرد. استفاده از این روش‌ها به‌ویژه در محیط‌های شهری که تداخل داده‌ها و تغییرات محیطی می‌توانند بر نتایج تأثیرگذار باشند، اهمیت ویژه‌ای دارد. در مرحله نهایی، مختصات تخمین‌زده‌شده با داده‌های GPS مرجع که پیش‌تر با استفاده از دستگاه دقیق مولتی‌فرکانس برداشت شده بودند، مقایسه شدند.



شکل ۶- مقایسه trajectory قبل و بعد از استفاده از IMU

در مقابل، داده‌های IMU (قرمز) با استفاده از ژيروسکوپ، سرعت زاویه‌ای دقیق‌تری ارائه می‌دهند و از این طریق مسیر دوران را با دقت بالاتری مشخص می‌کنند. به همین دلیل در غیاب سنسور IMU، ممکن است داده‌های odometry به‌تنهایی در هنگام دوران با شکست مواجه شود. این شکست باعث می‌شود که تلفیق داده‌های GPS و odometry با استفاده از SAM2 غیرممکن شود، زیرا SAM2 برای تخمین موقعیت دقیق به داده‌های پیوسته و پایدار نیاز دارد. نبود IMU در چنین شرایطی منجر به از دست رفتن پایداری و قطع روند تخمین موقعیت می‌شود. همچنین ترکیب داده‌های odometry و IMU باعث جبران خطاهای ناشی از انباشت نویز و انحراف می‌شود. به همین دلیل در نمودار، مسیر IMU Fusion (مسیر قرمز) در ناحیه دوران صاف‌تر و دقیق‌تر از مسیر odometry (مسیر آبی) است که نشان‌دهنده عملکرد بهینه IMU در سناریوهای چرخش و تغییر جهت است. به‌منظور تخمین دقیق موقعیت سه‌بعدی چاله‌ها، داده‌های GPS، اطلاعات به دست آمده از تحلیل Odometry و موقعیت تصویری چاله‌ها به‌صورت ترکیبی مورد استفاده قرار گرفتند. این فرآیند شامل چندین گام کلیدی بود که به

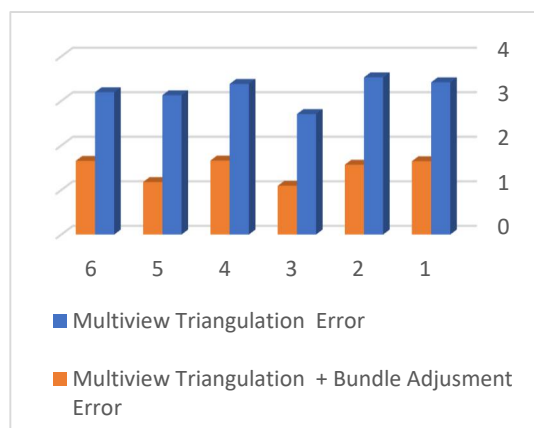
جدول ۵- نتایج تخمین موقعیت چاله‌ها هنگام استفاده از Multiview Triangulation و همچنین استفاده همزمان Bundle Adjustment

مختصات‌های بدست آمده حاصل این پژوهش استفاده همزمان از Multiview Triangulation و Bundle Adjustment		مختصات‌های بدست آمده با استفاده از Multiview Triangulation		مختصات‌های برداشت شده با استفاده از دستگاه مولتی فرکانس به عنوان داده‌های دقیق		شماره چاله
Northing	Easting	Northing	Easting	Northing	Easting	
۴۰۲۲۳۶۱.۲۴	۶۴۹۲۷۳.۳۲۶۲	۴۰۲۲۳۶۳.۳۴۲	۶۴۹۲۷۳.۱۰۲۴	۴۰۲۲۳۶۰.۱۲۴	۶۴۹۲۷۱.۹۵۱۷	۱
۴۰۲۲۳۵۹.۳۵۱	۶۴۹۲۷۳.۹۹۷۲	۴۰۲۲۳۶۲.۱۲۷	۶۴۹۲۷۳.۷۱۲۳	۴۰۲۲۳۵۹.۰۰۵	۶۴۹۲۷۲.۰۶۵۳	۲
۴۰۲۲۳۵۸.۴۴۹	۶۴۹۲۷۴.۰۹۶۲	۴۰۲۲۳۵۸.۶۰۳	۶۴۹۲۷۵.۳۲۴۴	۴۰۲۲۳۵۷.۳۲۷	۶۴۹۲۷۲.۹۴۰۲	۳
۴۰۲۲۳۵۸.۴۵۶	۶۴۹۲۷۵.۱۸۳	۴۰۲۲۳۵۹.۹۴۵	۶۴۹۲۷۵.۹۵۱	۴۰۲۲۳۵۷.۰۹۹	۶۴۹۲۷۴.۱۲۵۰	۴
۴۰۲۲۳۵۱.۲۴۸	۶۴۹۲۸۲.۱۵۳۹	۴۰۲۲۳۵۲.۶۳۵	۶۴۹۲۸۲.۳۴۵	۴۰۲۲۳۵۰.۰۰۵	۶۴۹۲۸۰.۶۵۲۸	۵
۴۰۲۲۳۴۶.۴۲۳	۶۴۹۲۸۶.۹۳۱	۴۰۲۲۳۴۸.۱۹۶	۶۴۹۲۸۷.۱۹۲	۴۰۲۲۳۴۵.۲۷۰	۶۴۹۲۸۵.۹۱۰۵	۶

جدول ۶- اختلاف نتایج تخمین موقعیت چاله‌ها با داده‌های دستگاه مولتی فرکانس (متر)

شماره چاله	هنگام استفاده از Multiview Triangulation	استفاده همزمان از Multiview Triangulation و Bundle Adjustment
۱	۳.۴۱	۱.۶۴
۲	۳.۵۲	۱.۵۶
۳	۲.۷۰	۱.۰۹
۴	۳.۳۸	۱.۶۶
۵	۳.۱۲	۱.۱۷
۶	۳.۱۹	۱.۶۵

این ارزیابی‌ها نشان می‌دهند که مختصات سه‌بعدی تخمین‌زده‌شده دارای خطایی کمتر از ۲ متر هستند، اما این خطا به طور مستقیم با دقت GPS در آن محل مرتبط است. در مناطقی با پوشش ضعیف GPS یا شرایطی که سیگنال‌های GPS دچار انحراف می‌شوند، احتمال افزایش خطا وجود دارد. همچنین عواملی مانند خطای GPS در محیط‌های شهری، حساسیت مدل به کیفیت ویدئوها و یا حرکت ناگهانی دوربین می‌توانند موجب افزایش خطا شوند. روش پیشنهادی توانست دقت مناسبی را برای کاربردهای عملی ارائه دهد. همچنین این مقایسه نتایج بین مختصات برداشت‌شده و مختصات محاسبه‌شده نشان می‌دهد که اختلاف‌ها عمدتاً به دلیل محدودیت‌های اندازه‌گیری و دقت



شکل ۷- مقایسه نموداری خطاهای تخمین موقعیت چاله‌ها با استفاده و بدون استفاده از Bundle Adjustment (متر)

کرده است. این امر سبب شده که مختصات نهایی از خطاهای تصادفی و سیستماتیک تأثیر کمتری بپذیرند. روش پیشنهادی این قابلیت را دارد که در پروژه‌های مختلفی همچون مدیریت زیرساخت‌های شهری، سیستم‌های حمل‌ونقل هوشمند و نقشه‌برداری پیشرفته به کار گرفته شود و به عنوان یک راهکار موثر برای تحلیل و ارزیابی چاله‌های شهری مورد استفاده قرار گیرد.

۵- بحث و نتیجه‌گیری

روش ارائه‌شده در این پژوهش، با بهره‌گیری از تکنیک‌های پیشرفته یادگیری عمیق و تلفیق داده‌های چندحسگره، رویکردی نوین برای تخمین موقعیت مطلق چاله‌ها در خیابان‌های شهری ارائه داده است. این روش با استفاده از سنسورهای داخلی گوشی‌های هوشمند، از جمله دوربین، GPS و IMU، امکان شناسایی و موقعیت‌یابی دقیق چاله‌ها را فراهم کرده و در عین حال هزینه‌های مرتبط با استفاده از تجهیزات پیچیده و گران‌قیمت را کاهش داده است. یکی از نقاط قوت روش پیشنهادی، استفاده از شبکه عصبی YOLO به منظور شناسایی چاله‌ها و الگوریتم DeepSORT به منظور ردیابی دقیق آن‌ها است. این ترکیب به‌ویژه در شرایط چالش‌برانگیز مانند تغییرات نوری یا حرکت سریع دوربین، عملکرد بسیار قابل‌قبولی ارائه داده است. علاوه بر این، ابزار Sahi نیز توانست دقت شناسایی چاله‌های کوچک و جزئیات را افزایش دهد که این مسئله به‌ویژه در محیط‌های پیچیده شهری بسیار حائز اهمیت است.

تلفیق داده‌های چندحسگری شامل داده‌های GPS، IMU و تصاویر ویدئویی، از طریق الگوریتم ORB-SLAM و ISAM2 و همچنین استفاده از تکنیک‌های بهینه‌سازی نظیر Bundle Adjustment، دقت تخمین مختصات سه‌بعدی چاله‌ها را به میزان قابل‌توجهی افزایش داده است. این ترکیب هوشمندانه باعث کاهش خطاهای ناشی از هر یک از حسگرها به صورت مستقل شده و سیستم را قادر ساخته است تا با دقت بالاتری مختصات چاله‌ها را در سیستم مختصات UTM محاسبه کند.

در نتیجه، این پژوهش راهکاری مقرون‌به‌صرفه، کارآمد و قابل حمل برای مدیریت زیرساخت‌های شهری و برنامه‌ریزی تعمیر و نگهداری معابر ارائه داده است. سادگی استفاده از این سیستم روی گوشی‌های هوشمند، امکان مشارکت شهروندان

GPS در محیط بوده و در بازه قابل‌قبولی قرار دارند. نتایج نشان‌دهنده کارایی و دقت روش پیشنهادی در تخمین موقعیت سه‌بعدی چاله‌ها در محیط‌های شهری است. به علاوه، استفاده از Bundle Adjustment امکان بهبود کلیه پارامترهای سیستم را با بهینه‌سازی همزمان داده‌ها فراهم

استفاده شود. همچنین، بهره‌گیری از سیستم‌های GPS با دقت بالاتر می‌تواند منجر به کاهش بیشتر خطاهای مکانی شود. ترکیب این تجهیزات پیشرفته با الگوریتم‌های تلفیق داده می‌تواند عملکرد سیستم را در محیط‌های پیچیده‌تر بهبود بخشد و دامنه کاربرد آن را گسترش دهد.

را در بهبود وضعیت خیابان‌ها افزایش داده و می‌تواند به بهبود کیفیت زندگی و ایمنی رانندگی کمک کند. به منظور گسترش این پژوهش، پیشنهاد می‌شود در تحقیقات آینده از دوربین‌های استریو یا سامانه‌های چنددوربین به منظور تسهیل فرایند و افزایش دقت در تشخیص موقعیت سه‌بعدی

مراجع

- [۱] Nistér, D., O. Naroditsky, and J. Bergen. Visual odometry. in Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition .۲۰۰۴ ,CVPR 2004. 2004. Ieee
- [۲] Davison, A.J., et al., MonoSLAM: Real-time single camera SLAM. IEEE transactions on pattern analysis and machine intelligence, 2007. 29(6): p. 1052-1067.
- [۳] Klein, G .and D. Murray. Parallel tracking and mapping for small AR workspaces. in 2007 6th IEEE and ACM international symposium on mixed and augmented reality. 2007. IEEE.
- [۴] Comport, A.I., E. Malis, and P. Rives. Accurate quadrifocal tracking for robust 3d visual odometry. in Proceedings 2007 IEEE International Conference on Robotics and Automation. 2007. IEEE.
- [۵] Newcombe, R.A., et al. KinectFusion: Real-time dense surface mapping and tracking. in 2011 10th IEEE International Symposium on Mixed and Augmented Reality. 2011. IEEE.
- [۶] Kerl, C., J. Sturm, and D. Cremers. Robust odometry estimation for RGB-D cameras. in 2013 IEEE international conference on robotics and automation. 2013. IEEE.
- [۷] Newcombe, R.A., S.J. Lovegrove, and A.J. Davison. DTAM: Dense tracking and mapping in real-time. in 2011 international conference on computer vision. 2011. IEEE.
- [۸] Engel, J., T. Schöps, and D. Cremers. LSD-SLAM: Large-scale direct monocular SLAM. in European conference on computer vision. 2014. Springer.
- [۹] Forster, C., M. Pizzoli, and D. Scaramuzza. SVO: Fast semi-direct monocular visual odometry. in 2014 IEEE international conference on robotics and automation (ICRA). 2014. IEEE.
- [۱۰] Meier, L., et al .Pixhawk: A system for autonomous flight using onboard computer vision. in 2011 IEEE International Conference on Robotics and Automation. 2011. IEEE.
- [۱۱] Engel, J., J. Sturm, and D. Cremers. Camera-based navigation of a low-cost quadcopter. in 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems. 2012. IEEE.
- [۱۲] Comport, A.I., E. Malis, and P. Rives. Accurate quadrifocal tracking for robust 3d visual odometry. in Proceedings 2007 IEEE International Conference on Robotics and Automation. 2007. IEEE.
- [۱۳] Zhang, W. and J. Kosecka. Image based localization in urban environments. in Third international symposium on 3D data processing, visualization, and transmission (3DPVT'06) .۲۰۰۶ .IEEE.
- [۱۴] Bloesch, M., et al. Robust visual inertial odometry using a direct EKF-based approach. in 2015 IEEE/RSJ international conference on intelligent robots and systems (IROS). 2015. IEEE.
- [۱۵] Li, M. and A.I. Mourikis, High-precision, consistent EKF-based visual-inertial odometry. The International Journal of Robotics Research, 20 : (۶)۳۲ .۱۳p. 690-711.
- [۱۶] Leutenegger, S., et al .,Keyframe-based visual–inertial odometry using nonlinear optimization. The International Journal of Robotics Research, 2015. 34(3): p. 314-334.
- [۱۷] Usenko, V., et al. Direct visual-inertial odometry with stereo cameras. in 2016 IEEE International Conference on Robotics and Automation (ICRA). 2016. IEEE.
- [۱۸] Engel, J., V. Koltun, and D. Cremers, Direct sparse odometry. IEEE transactions on pattern analysis and machine intelligence, 2017. 40(3): p. 611-625.

- [۱۹] Mur-Artal, R. and J.D. Tardós, Visual-inertial monocular SLAM with map reuse. *IEEE Robotics and Automation Letters*, 2017. 2(2): p. 796-803.
- [۲۰] Leutenegger, S., et al. Keyframe-based visual-inertial odometry using nonlinear optimization. *The International Journal of Robotics Research*, 2015. 34(3): p. 314-334.
- [۲۱] Qin, T., P. Li, and S. Shen, Vins-mono: A robust and versatile monocular visual-inertial state estimator. *IEEE Transactions on Robotics*, 2018. 34(4): p. 1004-1020
- [۲۲] Forster, C., M. Pizzoli, and D. Scaramuzza. SVO: Fast semi-direct monocular visual odometry. in *2014 IEEE international conference on robotics and automation (ICRA)*. 2014. IEEE.
- [۲۳] Wang, X., et al. DeepVO: Towards End-to-End Visual Odometry with Deep Recurrent Convolutional Neural Networks. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2017.
- [۲۴] Zhou, T., et al. Unsupervised Learning of Depth and Ego-Motion from Video. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.
- [۲۵] Bloesch, M., et al. CodeSLAM - Learning a Compact, Optimisable Representation for Dense Visual SLAM. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.
- [۲۶] Yang, Z., et al. CubeSLAM: Monocular 3D Object Detection and SLAM without Prior Knowledge. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019
- [۲۷] Campos, R., et al. ORB-SLAM3: An Accurate Open-Source Library for Visual, Visual-Inertial and Multi-Map SLAM. *IEEE Transactions on Robotics*, vol. 37, no. 6, pp. 1460-1476, Dec. 2021.
- [۲۸] Engel, J., V. Koltun, and D. Cremers. Direct Sparse Odometry. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 3, pp. 611-625, March 2018.
- [۲۹] Koch, C., and M. P. Heinrich. "Deep learning-based pothole detection using smartphone images." *Journal of Transportation Engineering, Part A: Systems*, vol. 145, no. 11, 2019.
- [۳۰] Jo, Y., and S. Ryu. "Pothole detection system using vehicle vibration signals and image processing." *Sensors*, vol. 19, no. 14, 2019.
- [۳۱] Madli, R., et al. "Intelligent road surface monitoring system using ultrasonic sensors and GPS." *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 4, pp. 1046-1055, April 2019.
- [۳۲] Wang, W., et al. "Road surface defect detection using deep convolutional neural networks." *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 5, pp. 1231-1242, May 2020.
- [۳۳] Dhiman, C., and R. Klette. "Multiscale image processing and deep learning for pothole detection." *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 1, pp. 241-252, January 2021.
- [۳۴] Chen, Y., et al. "Road surface monitoring system using SLAM and multi-objective optimization." *IEEE Robotics & Automation Letters*, vol. 6, no. 2, pp. 2530-2537, April 2021.
- [۳۵] Jahanshahi, M. R., et al. "Automated measurement of pothole depth and volume using stereo images." *Journal of Computing in Civil Engineering*, vol. 35, no. 4, 2021.
- [۳۶] Yu, X., and E. Salari. "Pothole detection using LiDAR and camera fusion." *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 11, pp. 7342-7353, November 2021.
- [۳۷] Fox, D., et al. "Automated pothole mapping using drones and aerial imagery." *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1-13, 2022.
- [۳۸] Hiremath, R., et al. (2021). A Smart App for Pothole Detection Using Yolo Model. In *ICT Analysis and Applications* (pp. 1-13). Springer.
- [۳۹] Diwan, T., Anirudh, G., & Tembhurne, J. V. (2023). Object detection using YOLO: Challenges, architectural successors, datasets and applications. *Multimedia Tools and Applications*, 82(1), 1-23.
- [۴۰] Zhang, Y., Chen, Z., & Wei, B. (2020). A Sport Athlete Object Tracking Based on Deep Sort and Yolo V4 in Case of Camera Movement. In *2020 IEEE 6th International Conference on Computer and Communications (ICCC)*.

- [۴۱] J. Chen et al., "Multi-View Triangulation: Systematic Comparison and an Improved Method," in *IEEE Access*, vol. 8, pp. 21017-21027, 2020.
- [۴۲] Rehder, J., Nikolic, J., Schneider, T., Hinzmann, T., & Siegwart, R. (2016). Extending kalibr: Calibrating the extrinsics of multiple IMUs and of individual axes. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pp. 4304-4311.
- [۴۳] Xu, G., et al. (2021). Position and orientation measurement adopting camera calibrated by projection geometry of Plücker matrices of three-dimensional lines. *Scientific Reports*, 7, 44092.
- [۴۴] Lam, S. K., et al. (2016). Improved Allan Variance Analysis for Inertial Measurement Units. *Journal of Navigation*, 75(4), 747-764.
- [۴۵] Madry, M., et al. (2018). Visual-Inertial Odometry: A Comparative Study. *IEEE Robotics and Automation Letters*, 3(4), 3383-3390.
- [۴۶] Poulou, A., & Han, D. S. (2020). Hybrid Indoor Localization Using IMU Sensors and Smartphone Camera. *Sensors*, 19(23), 5084.
- [۴۷] Kaplan, E. D., & Hegarty, C. J. (2018). *Understanding GPS: Principles and Applications (Third Edition)*. Artech House.
- [۴۸] Wang, X., et al. (2023). Temporal Feature Aggregation for Video Understanding: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(11), 28345-28364.
- [۴۹] Yan, Y., Zhang, B., Zhou, J., Zhang, Y., & Liu, X. (2022). Real-time localization and mapping utilizing multi-sensor fusion and visual-IMU-wheel odometry for agricultural robots in unstructured, dynamic and GPS-denied environments. *Agronomy*, 12(8), 1740.
- [۵۰] Ma, H., Yan, L., Xia, Y., & Fu, M. (2020). *Kalman filtering and information fusion*. Springer.
- [۵۱] Chen, J., Li, W., & Chen, G. (2023). Multi-sensor data fusion with heterogeneous sampling rates: Challenges and solutions. *IEEE Access*, 11, 10501-10515.
- [۵۲] Xu, B., Wang, D., Chen, H., & Li, M. (2023). Data augmentation for deep learning: A survey. *IEEE Access*, 11, 11353-11373.
- [۵۳] <https://github.com/umer0586/SensorServer>
- [۵۴] Oliveira, M., Castro, A., Madeira, T., Pedrosa, E., et al. (2020). A ROS framework for the extrinsic calibration of intelligent vehicles: A multi-sensor, multi-modal approach. *Robotics and Autonomous Systems*, 134, 103509.
- [۵۵] Huai, J., Zhuang, Y., Shao, Y., Jozkow, G., & Wang, B. (2023). A Review and Comparative Study of Close-Range Geometric Camera Calibration Tools. *arXiv preprint arXiv:2306.09014*.
- [۵۶] Zhang, Q., Li, Z., & Zhang, X. (2022). GPS data augmentation using UTM coordinate transformation and linear interpolation. *Journal of Applied Geoinformatics*, 18(2), 147-158.
- [۵۷] Schneider, J., Förstner, W. (2015). Real-Time Accurate Geo-Localization of a MAV with Omnidirectional Visual Odometry and GPS. In L. Agapito, M. M. Bronstein, & C. Rother (Eds.), *Computer Vision - ECCV 2014 Workshops* (pp. 233-248). Springer International Publishing.
- [۵۸] Zhang, G., Liu, B., & Liang, Z. (2023). Implementation and optimization of ORB-SLAM2 algorithm based on ROS on mobile robots. *Third International Conference on Robotics and Artificial Intelligence*, 1-9.
- [۵۹] Telceken, M., Akgun, D., Kacar, S., & Bingol, B. (2024). A New Approach for Super Resolution Object Detection Using an Image Slicing Algorithm and the Segment Anything Model. *Sensors (Basel, Switzerland)*.
- [۶۰] Kaess, M., Johannsson, H., Roberts, R., Ila, V., Leonard, J. J., & Dellaert, F. (2012). iSAM2: Incremental Smoothing and Mapping Using the Bayes Tree. *The International Journal of Robotics Research*, 31(2), 216-235.
- [۶۱] Zhang, B., Li, Y., Liu, S., Pang, Z. and Zhao, H., 2022, July. Trajectory planning of upper limb exoskeleton rehabilitation Robot based on polynomial interpolation. In *2022 IEEE International Conference on Real-time Computing and Robotics (RCAR)* (pp. 727-732). IEEE.

- [۶۲] Hakim, I.M., Zakaria, H., Muslim, K. and Ihsani, S.I., 2023, February. 3D human pose estimation using blazepose and direct linear transform (DLT) for joint angle measurement. In 2023 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC) (pp. 236-241). IEEE.
- [۶۳] Luhmann, T., Fraser, C. and Maas, H.G., 2016. Sensor modelling and camera calibration for close-range photogrammetry. ISPRS Journal of Photogrammetry and Remote Sensing, 115, pp.37-46.
- [۶۴] Akyon, F.C., Altinuc, S.O. and Temizel, A., 2022. Slicing aided hyper inference and fine-tuning for small object detection. In 2022 IEEE International Conference on Image Processing (ICIP) (pp. 966-970). IEEE.
- [۶۵] Anonymous, 2023. Custom Object Detection Using YOLO Integrated with a... Segment Anything Model (SAM). International Research Journal of Engineering and Technology (IRJET), 10(10), pp.1