



دانشگاه صنعتی امیرکبیر
(پلی تکنیک تهران)
دانشکده مهندسی کامپیوتر

گزارش درس روش پژوهش و ارائه

سیستم ادراک و تصمیم‌گیری در وسایل نقلیه خودران

نگارش
کیان شکرانه

استاد راهنما
دکتر مهدی صدیقی

خرداد ۱۴۰۲

تقدیر و تشکر

سپاس بیکران پروردگار یکتا را که به من این توفیق را عطا فرمود تا بتوانم قدمی در راه علم و دانش بردارم و این گزارش را به نتیجه برسانم.

تشکر قلبی خودم را از استاد گرانقدر جناب آقای دکتر صدیقی که زحمت راهنمایی این گزارش را عهده‌دار گردیدند و از راهنمایی‌های مدبرانه ایشان استفاده نمودم ابراز می‌دارم و توفیق روزافزون ایشان را همراه با صحت و سعادت خواستارم.

چکیده

امروزه وسایل نقلیه خودران، از مباحث مطرح و مورد بحث به شمار می‌روند. ادراک را می‌توان وظیفه اصلی هر سیستم رانندگی خودران برشمرد. ادراک یک وسیله نقلیه خودران، به معنی جمع‌آوری داده‌های مورد نیاز از محیط پیرامون است. داده‌های ادراک شده، به عنوان ورودی به سیستم تصمیم‌گیری تحویل داده می‌شوند تا تصمیم مناسب با شرایط گرفته شود. معیارهای گوناگونی در تصمیم‌گیری مورد بررسی قرار می‌گیرند که می‌توان مهم‌ترین آنها را حفظ ایمنی سرنشینان دانست.

در سیستم ادراک، از حسگرهای گوناگونی از جمله رادار، لایدار، دوربین، و ... استفاده می‌شود. هر یک از این حسگرها، کاربردهای مختص به خود، انواع مختلف، و نیز مزایا و معایبی دارند. بنابراین، استفاده از حسگرهای مختلف در کنار هم ضرورت دارد، و به شکل‌گیری یک سیستم ادراک مناسب در گرو آن است. همچنین در سیستم ادراک یک خودروی خودمختار، رویکردهای مختلفی مورد استفاده قرار می‌گیرند. دو رویکرد مهم و اساسی برای ادراک وسایل نقلیه خودران، تقسیم‌بندی معنایی و تشخیص جسم هستند. هر یک از این رویکردها در پردازش و تصمیم‌گیری خودرو حائز اهمیت و نیز جداناپذیر هستند. در طول دهه‌های اخیر، مدل‌ها و الگوریتم‌های متعددی برای هر یک از دو رویکرد مذکور، ارائه شده‌اند. یادگیری عمیق، همانطور که امروزه در هر زمینه‌ای کاربردی است، در زمینه سیستم وسایل نقلیه خودمختار نیز کاربرد گسترده پیدا کرده است. انواع مدل‌ها و الگوریتم‌های یادگیری عمیق بطور گسترده در هر دو رویکرد تقسیم‌بندی معنایی و تشخیص جسم مورد استفاده قرار گرفته‌اند. هر یک از این الگوریتم‌ها محدودیت‌های مدل‌های پیش از خود را رفع کرده‌اند. بنابراین یادگیری عمیق را می‌توان ستون فقرات سیستم ادراک خودروهای خودمختار دانست.

واژه‌های کلیدی:

ادراک، حسگر، تقسیم‌بندی معنایی، تشخیص جسم، یادگیری تقویتی

چکیده.....	أ
فصل اول مقدمه.....	۱
فصل دوم حسگرها.....	۴
۲-۱- دوربین.....	۵
۲-۲- لایدار.....	۷
۲-۳- رادار.....	۹
۲-۴- حسگر فراصوتی.....	۹
۲-۵- مرور و مقایسه حسگرها.....	۱۰
فصل سوم رویکردهای ضروری ادراک.....	۱۲
۳-۱- تقسیم‌بندی معنایی.....	۱۳
۳-۱-۱- ویژگی‌های مکانی و معنایی.....	۱۴
۳-۱-۲- پژوهش‌های مرتبط.....	۱۴
۳-۲- تشخیص شیء.....	۱۵
۳-۲-۱- ابر نقاط.....	۱۶
۳-۲-۲- پژوهش‌های مرتبط.....	۱۶
فصل چهارم یادگیری عمیق.....	۱۸
۴-۱- یادگیری عمیق در تقسیم‌بندی معنایی.....	۲۰
۴-۲- یادگیری عمیق در تشخیص اجسام.....	۲۰
۴-۲-۱- یادگیری تحت نظارت.....	۲۱
۴-۲-۲- یادگیری تقویتی.....	۲۱
فصل پنجم مرور و جمع‌بندی.....	۲۳
منابع و مراجع.....	۲۶

شکل ۱-۲ - دوربین تک‌چشمی و دوربین استریو [۵].....	۶
شکل ۲-۲ - لایدار دوبعدی (a)، سه‌بعدی (b)، حالت مستحکم (c) [۶].....	۸
شکل ۳-۲ - محدوده طول موج مورد استفاده در ادراک وسایل نقلیه خودران [۶].....	۱۰
شکل ۴-۲ - مقایسه ویژگی‌های سنسورهای گوناگون [۶].....	۱۱
شکل ۱-۳ - نمونه‌ای از کاربرد رنگ‌ها در تقسیم‌بندی معنایی [۹].....	۱۳
شکل ۲-۳ - نمونه‌ای از تشخیص اجسام توسط مدل YOLOv3 [۱۳].....	۱۶
شکل ۱-۴ - لایه‌های مختلف CNN [۱۷].....	۱۹

فصل اول

مقدمه

مقدمه

مردم آمریکا بطور متوسط سه تریلیون مایل در سال رانندگی می‌کنند، که مدت زیادی از این زمان، صرف انتظار کشیدن در ترافیک می‌شود. سالانه چندین هزار نفر در حوادث رانندگی جان خود را از دست می‌دهند [۱]. حال، پیشرفت خودروهای خودمختار می‌تواند تاثیر شگرفی در صنعت حمل و نقل بگذارد. پیشرفت این خودروها، می‌تواند کیفیت حمل و نقل را افزایش دهد، و شمار مرگ و میر ناشی از خطاهای انسانی کاهش پیدا می‌کند. همچنین، خودمختار شدن خودروها به سرنشینان این امکان را می‌دهد که نیازی به کنترل خودرو نداشته باشند، و بتوانند در زمان سپری شده در مسیر و ترافیک، به کارهای دیگر خود بپردازند. اینها تنها تعدادی از مواردی هستند که منجر به محبوبیت و مورد بحث قرار گرفتن روزافزون وسایل نقلیه خودران شده‌اند. این محبوبیت، موجب انجام پژوهش‌های فراوان در این زمینه شده است. همچنین، پیشرفت‌های سریع در فناوری‌هایی مانند الکترونیک، منجر به توسعه فناوری‌های وسایل نقلیه خودران شده است. اما همچنان چالش‌ها و محدودیت‌هایی در این زمینه وجود دارند، و همچنان، پیشرفت و توسعه باید ادامه پیدا کند.

تعریف‌های مختلفی برای وسایل نقلیه خودران ارائه شده‌اند. یکی از این تعریف‌ها توسط UCSUSA [۲] ارائه شده است که چنین است: "وسایل نقلیه خودران، خودروها یا کامیون‌هایی هستند که انسان‌ها هرگز مجبور نیستند کنترل آنها را برای عملکرد ایمن وسیله نقلیه، به دست بگیرند. آنها همچنین به عنوان خودروهای خودران یا بدون راننده شناخته می‌شوند، که حسگرها و نرم‌افزار را برای کنترل، هدایت و رانندگی وسیله نقلیه ترکیب می‌کنند." همچنین تعریف دیگری که توسط Thrun [۳] ارائه شده است، چنین است: "وسیله نقلیه بدون سرنشینی که قادر است محیط خود را تشخیص دهد و بدون کمک انسان حرکت کند." با این حال، انجمن مهندسين خودرو تعریفی استاندارد و اصولی بر پایه سطوح خودران بودن ارائه کرده است. سطوح مذکور متشکل از شش سطح مختلف، از ساده‌ترین سیستم‌ها، تا سیستم‌هایی با قابلیت خودران بودن بصورت کامل هستند. سطح متناسب با سیستم یک خودرو، بیانگر میزان پیشرفته بودن آن سیستم است.

یک وسیله نقلیه خودران، با استفاده از سیستم ادراک خود، اطلاعاتی را از محیط پیرامون خود گردآوری می‌کند. جمع‌آوری این داده‌ها با استفاده از انواع گوناگونی از حسگرها صورت می‌گیرد. حسگرها، هر یک قابلیت‌ها و محدودیت‌هایی دارند. بنابراین، در یک سیستم ادراک مناسب نیاز است که از ترکیب چندین حسگر متفاوت استفاده شود. داده‌های جمع‌آوری شده، در ادامه باید توسط یک مدل پردازش شوند تا یادگیری حاصل شود. دو رویکرد تقسیم‌بندی معنایی و تشخیص شیء، به عنوان رویکردهای اصلی سیستم‌های ادراک خودرو دانسته می‌شوند. تقسیم‌بندی معنایی را می‌توان طبقه‌بندی تصویر در سطح پیکسل دانست. تشخیص

اشیاء، در واقع شناسایی و تشخیص یک شیء در یک تصویر، و مشخص کردن آن با رسم کادر مییاشد. برای اطمینان از ایمنی افراد، لازم است مدل‌های یادگیری بسیار کارآمد و دقیق آموزش‌دیده بر روی طیف وسیعی از سناریوهای رانندگی برای شناسایی دقیق اشیاء اطراف تحت شرایط مختلف آب و هوایی و نور، به کار گرفته شود. این روش یادگیری از طریق آموزش، فرآیند تصمیم‌گیری و مکانیسم کنترل خودرو را برای انجام اقدامات لازم تنظیم می‌کند [۴].

یادگیری عمیق، به عنوان یکی از شاخه‌های یادگیری ماشین، در تقسیم‌بندی معنایی و تشخیص شیء بطور گسترده مورد استفاده قرار گرفته است. مدل‌های گوناگونی مانند SegNet در ادراک خودرو استفاده شده‌اند [۴] و هر یک توانسته‌اند محدودیت‌های و مشکلاتی را رفع کنند. اما همچنان راه درازی برای رسیدن به خودمختاری کامل باید طی شود.

بنابراین، با توجه به مواردی که در این بخش ذکر شد، ابتدا حسگرهای پرکاربرد در زمینه خودروهای خودمختار مورد بررسی و مقایسه قرار خواهند گرفت. سپس، دو رویکرد تقسیم‌بندی معنایی و تشخیص شیء بررسی می‌شوند، و چالش‌های هر یک با بیان راه حل ارائه شده، ذکر می‌شود. در ادامه به بررسی کاربرد مدل‌های یادگیری عمیق در این حوزه پرداخته می‌شود. در نهایت نیز مروری بر موارد مذکور و جمع‌بندی وجود دارد.

فصل دوم

حسگرها

حسگرها

انسان‌ها برای رانندگی به حواس شنوایی و بینایی خود متکی هستند. یک وسیله نقلیه خودران، از حسگرهای متعدد برای ادراک استفاده می‌کند. حسگرها را به دو دسته کلی می‌توان تقسیم کرد. دسته اول، حسگرهای فعال^۱ مانند رادار، لایدار، و سونار هستند که انرژی را به محیط اطراف ساطع می‌کنند، و سپس بازخورد انرژی از اجسام محیط پیرامون را دریافت می‌کنند، تا ادراکی را با استفاده از این داده‌ها حاصل کنند. نوع دیگر، حسگرهای غیرفعال^۲ هستند، که انرژی ساطع شده از محیط اطراف را برای ادراک مورد استفاده قرار می‌دهند. از این نوع حسگرها می‌توان به دوربین‌ها اشاره کرد [۴].

هر حسگر خودرو، قابلیت‌ها و محدودیت‌هایی دارد. بنابراین، استفاده از چندین حسگر متفاوت، امکان غلبه بر محدودیت‌های هر حسگر را فراهم می‌کند. اصلی‌ترین حسگرهایی که در این زمینه مورد استفاده قرار می‌گیرند را می‌توان حسگر فراصوتی، دوربین، رادار، و لایدار برشمرد. در ادامه، قابلیت‌ها و محدودیت‌های هر یک از آنها مورد بررسی قرار گرفته‌اند.

۱-۲- دوربین

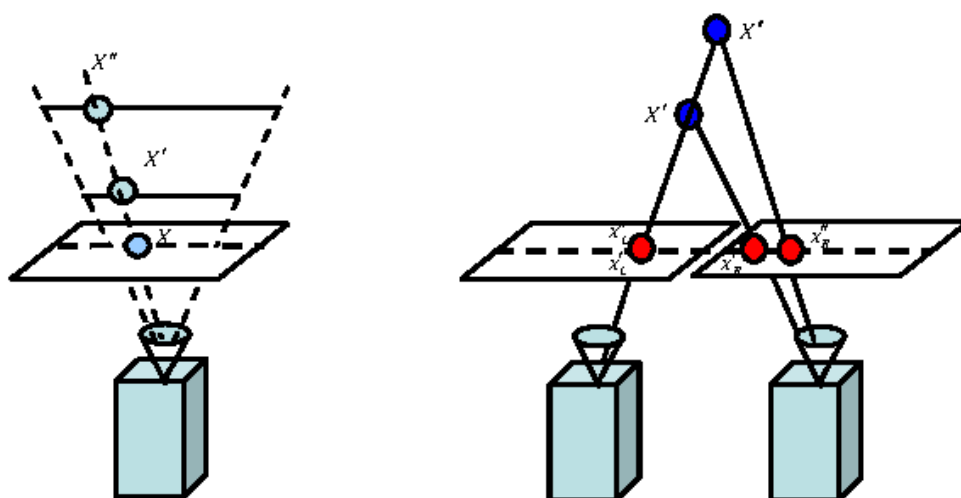
پیش‌تر بیان شد که دوربین‌ها از انواع حسگرهای غیرفعال می‌باشند. این حسگر به طور کلی در دو نوع مختلف استریو^۳ و تک‌چشمی^۴ مورد استفاده قرار می‌گیرد. دوربین‌های تک‌چشمی قابلیت جمع‌آوری اطلاعات بافت و شکل هستند. این اطلاعات، به تشخیص و طبقه‌بندی علائم جاده، چراغ‌های راهنمایی و رانندگی، و نیز انواع خطوط جاده، کمک‌کننده هستند. علاوه بر شکل و بافت محیط پیرامون، نیاز به تعیین موقعیت و فاصله اجسام نیز وجود دارد. این امکان توسط دوربین‌های استریو فراهم می‌شود [۴]. در شکل ۱-۲ نمای از دو نوع دوربین ذکر شده قابل مشاهده است. با توجه به توضیحات داده شده، می‌توان نتیجه گرفت که هر وسیله نقلیه خودران باید مجهز به چندین دوربین باشد، چرا که دوربین‌ها می‌توانند غنی‌ترین اطلاعات را در مورد محیط پیرامون جمع‌آوری کنند. بنابراین می‌توان گفت که استفاده از دوربین‌ها علاوه بر سایر حسگرها ضرورت دارد.

^۱ Active Sensors

^۲ Passive Sensors

^۳ Stereo Cameras

^۴ Monocular Cameras



شکل ۲-۱- دوربین تک‌چشمی و دوربین استریو [۵].

تقسیم‌بندی دیگری نیز برای دوربین‌ها وجود دارد، که آنها را با توجه به طول موج قابل دریافت، به دو دسته مرئی^۱ و فروسرخ^۲ تقسیم می‌کند. دوربین‌های مرئی مانند چشم‌های انسان عمل می‌کنند، و می‌توانند طول موج‌های بین ۴۰۰ نانومتر تا ۷۰۰ نانومتر را دریافت کنند. این دوربین‌ها به دلیل هزینه مناسب، و نیز وضوح و کیفیت خوب، کاربرد نسبتاً زیادی برای جمع‌آوری داده از محیط اطراف برای وسایل نقلیه خودران دارند. اما این دوربین‌ها بدون ایراد نیستند. شرایط آب و هوایی نامناسب مانند مه یا برف، و یا نور بسیار زیاد یا کم، بر عملکرد این دوربین‌ها تأثیرگذار هستند. همچنین، وضوح بالا به معنی حجم عظیم داده‌هایی است که باید پردازش شوند، که سربار زیادی را ایجاد می‌کنند. تشخیص نقطه کور، کنترل نمای جانبی، و ضبط تصادف را می‌توان از کاربردهای دوربین‌های مرئی برشمرد [۶].

یک دوربین فروسرخ، همان‌طور که نام آن نشان می‌دهد، در بازه طول موج‌های فروسرخ، یعنی ۷۸۰ نانومتر تا ۱۰۰۰ نانومتر فعالیت می‌کند. فعالیت در این طول موج‌ها، مزیت کمتر بودن تداخل‌های نوری را به همراه دارد. این مزیت باعث شده است که حسگرهای دیگری مانند لایدار نیز در این محدوده فعالیت کنند. سیستم‌های ادراکی شامل این دوربین‌ها، معمولاً در یکی از بازه‌های نزدیک فروسرخ (۷۸۰ نانومتر تا ۳۰۰۰ نانومتر) یا فروسرخ میانی (۳۰۰۰ نانومتر تا ۵۰۰۰۰ نانومتر) فعالیت می‌کنند [۷، ۸]. در سیستم‌های ادراک معمولاً برای

^۱ Visible (VIS)

^۲ Infrared (IR)

^۳ Near-Infrared (NIR)

^۴ Mid-Infrared (MIR)

تکمیل دوربین‌های مرئی از دوربین‌های نزدیک فروسرخ استفاده می‌شود. با توجه به توضیحات قبل، این دوربین‌ها در هر شرایط آب و هوایی و هر شرایط روشنایی به خوبی عمل می‌کنند. بنابراین در شرایطی مانند خروج از تونل یا وجود نور شدید، کاربرد دارند. علاوه بر اینها، از دوربین‌های فروسرخ می‌توان برای شناسایی عابرین استفاده کرد.

دوربین‌هایی مرئی و فروسرخ بیان شده، همگی از نوع حسگرهای غیر فعال هستند. با این حال، دوربین‌هایی مانند دوربین‌های زمان پرواز^۱ وجود دارند که فعال هستند. یک دوربین زمان پرواز، پالس‌های فروسرخ میانی منتشر می‌کند و با استفاده از اختلاف فاز سیگنال منتشر شده و سیگنال دریافتی، فاصله اجسام را تعیین می‌کند. طی این فرآیند، یک نمای سه‌بعدی از اجسام پیرامون تهیه می‌شود [۶].

۲-۲- لایدار^۲

سنسورهای لایدار، پالس‌های لیزری در بازه امواج فروسرخ (معمولاً ۹۵۰ یا ۱۵۵۰ نانومتر) منتشر می‌کنند. همچنین دریافت‌کننده‌هایی وجود دارند که پالس‌های برگشتی را دریافت می‌کنند، و آنها را برای تولید خروجی ادراک مورد استفاده قرار می‌دهند. کاربرد اصلی لایدار در وسایل نقلیه خودران، شناسایی و تشخیص دسته‌بندی اجسام، و نیز تعیین موقعیت آن با اندازه‌گیری دقیق فاصله جسم است. تعیین فاصله اجسام، با محاسبه زمان رفت و برگشت پالس‌ها انجام می‌شود. تعیین موقعیت به روش بیان شده، باید برای همه اجسام پیرامون انجام شود. بنابراین، بطور کلی لایدار هزاران پالس در ثانیه ارسال می‌کند تا یک نمای جامع ۳۶۰ درجه از محیط اطراف حاصل شود. این نمای حاصل شده را نقشه ابر نقطه‌ای^۳ می‌نامند [۴].

استفاده از پالس‌های لیزر در این حسگر، قابلیت عملکرد دقیق و مناسب در هر شرایط آب و هوایی، و نیز در نور کم و زیاد را می‌دهد. اما با توجه به توضیحات، لایدار مانند دوربین‌های استریو، مبتنی بر عمق است. بنابراین قابلیت تشخیص شکل و بافت اجسام را ندارد. این محدودیت نشان می‌دهد که نمی‌توان از لایدار بصورت مستقل استفاده کرد. بنابراین، استفاده از حسگرهایی مانند دوربین‌ها در کنار لایدار ضروری است.

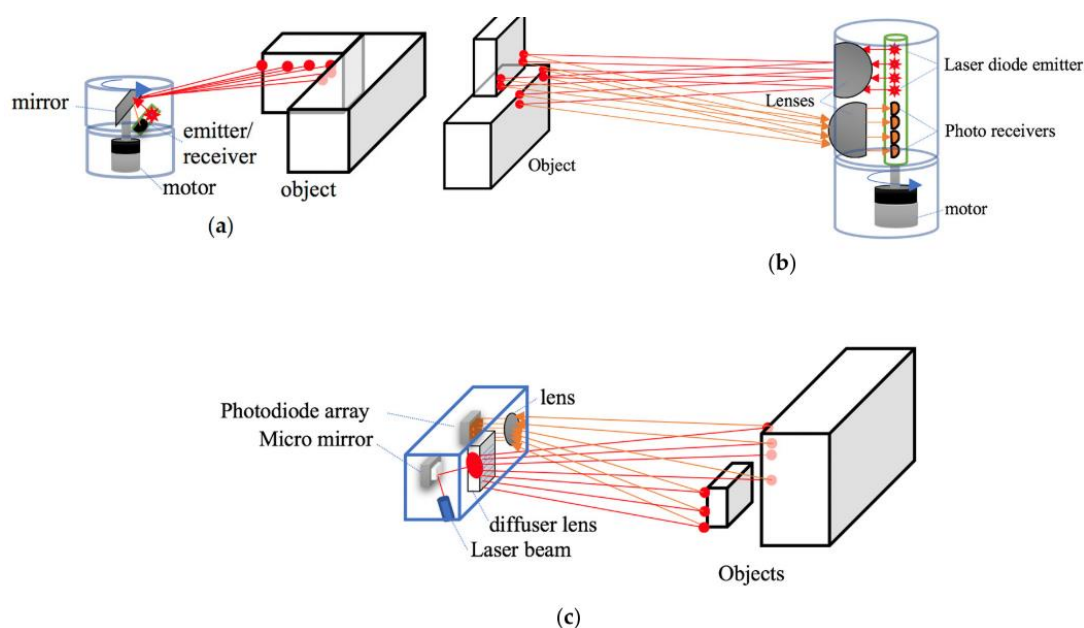
^۱ Time of Flight (ToF)

^۲ LiDAR (Light Detection and Ranging)

^۳ Point Cloud Map

می‌توان لایدارها را بر اساس نوع اطلاعاتی که از محیط دریافت می‌کنند، به دو دسته دوبعدی و سه‌بعدی تقسیم کرد. لایدار دوبعدی با انتشار پرتو لیزر بر روی یک آینه دوار عمود بر محور چرخش، اطلاعات محیط را دریافت می‌کند. یک رادار سه‌بعدی از تعدادی لیزر دیودی نصب شده بر روی یک غلاف که با سرعت بالا می‌چرخد، استفاده می‌کند. هر چه تعداد دیودهای استفاده شده بیشتر باشد، دقت نمای خروجی نیز بیشتر خواهد بود. با این روش، یک نمای سه‌بعدی دقیق از محیط حاصل می‌شود. از لایدارهای سه‌بعدی در کروز کنترل تطبیقی، اجتناب از برخورد، شناسایی اشیاء و ... استفاده می‌شود.

دسته دیگر، لایدارهای حالت مستحکم^۱ هستند. در این حسگرها، یک آینه میکرو پرتو را روی یک عدسی پخش‌کننده منعکس می‌کند. طی این فرآیند، خطوطی ایجاد می‌شوند که به اجسام پیرامون برخورد می‌کنند و طی این فرآیند، میدان دید به خوبی شناخته می‌شود. نحوه عملکرد انواع لایدار ذکر شده در شکل ۲-۲ قابل مشاهده است [۶].



شکل ۲-۲- لایدار دوبعدی (a)، سه‌بعدی (b)، حالت مستحکم (c) [۶].

^۱ Solid State LiDAR

۳-۲- رادار^۱

اساس کار حسگرهای رادار، مشابه با لایدارها می‌باشد. تفاوت آنها این است که با توجه به توضیحات داده شده، لایدار از پالس‌های لیزری استفاده می‌کند، اما رادارها آنتنی دارند که سیگنال‌های رادیویی را منتشر می‌کند. محدوده فعالیت رادارها در طول موج‌هایی از مرتبه میلیمتر می‌باشد. مشابه با لایدار، رادار گیرنده‌ای دارد که سیگنال رادیویی را که از اجسام در محیط اطراف منعکس شده است، تشخیص می‌دهد. همچنین رادار از مدت زمان رفت و برگشت این سیگنال‌ها به آنتن، می‌تواند فاصله و موقعیت اجسام پیرامون را تعیین کند.

این نوع از حسگرها، در شرایط آب و هوایی مختلف، از جمله مه یا برف، عملکرد مناسبی دارند و نسبت به تعداد کثیری از حسگرهای دیگر، برتری دارند. از معایب رادار می‌توان به این اشاره کرد که نسبت به لایدار و دوربین، دقت کمتری دارد و محدوده قابل پوشش داده شدن توسط یک رادار، معمولاً در حدود ۵۰ متر می‌باشد. این محدودیت باعث می‌شود که رادار در تعیین جزییات و تشخیص و طبقه‌بندی اجسام، توانایی نسبتاً کمتری داشته باشد. بنابراین استفاده از حسگرهای دیگر در کنار رادار ضروری است. از کاربردهای رادار می‌توان به تشخیص فاصله از خودروهای مجاور، فعال کردن سیستم ترمز اضطراری در صورت احتمال برخورد، سیستم‌های کمک کننده برای پارک کردن، و ... اشاره کرد [۴،۶].

۴-۲- حسگر فراصوتی^۲

یک حسگر فراصوتی شامل یک غشای مغناطیسی است، که امواج صوتی در بازه ۲۰ کیلوهرتز تا ۴۰ کیلوهرتز منتشر می‌کند. با استفاده از اختلاف زمان انتشار امواج و دریافت بازخورد آنها، فاصله اجسام پیرامون محاسبه می‌شود. این حسگرها نسبتاً ارزان هستند، و برای محاسبه فاصله اجسام نزدیک در سرعت کم کاربرد دارند. مزیت دیگر حسگرهای فراصوتی این است که در شرایط آب و هوایی مختلف به خوبی عمل می‌کنند. با این حال، این حسگرها بدون ایراد نیستند. در پالس‌ها ممکن است جهش ایجاد شود، و یا نقطه کور در اندازه‌گیری‌ها وجود داشته باشد. این موارد ممکن است منجر به محاسبات و تشخیص نادرست شوند [۶].

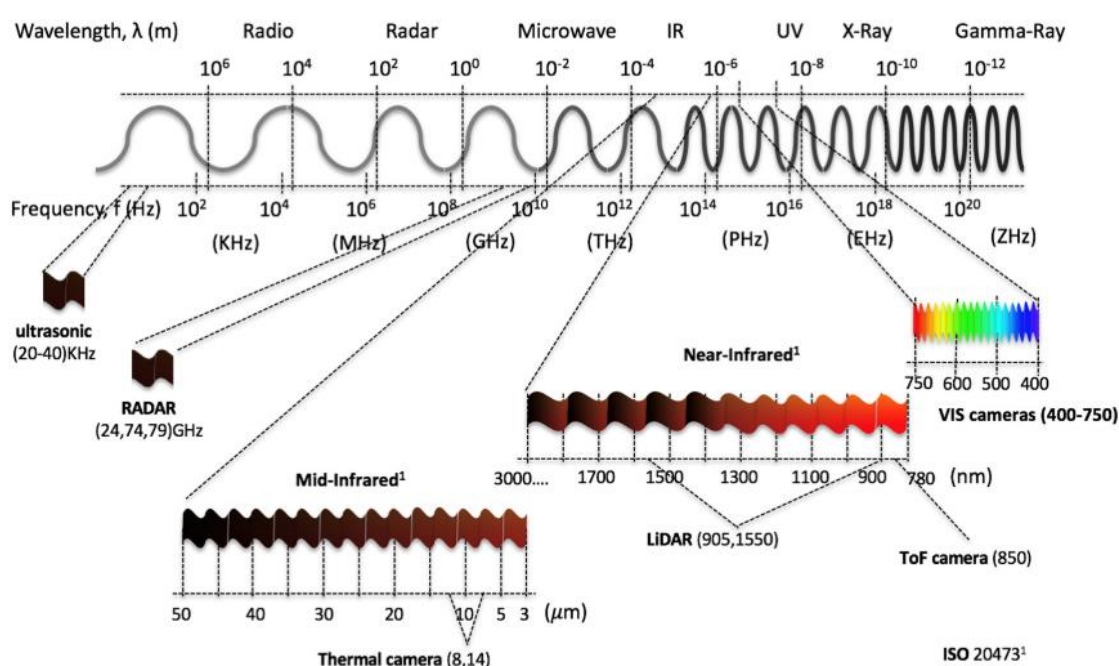
^۱ RADAR (Radio Detection and Ranging)

^۲ Ultrasonic

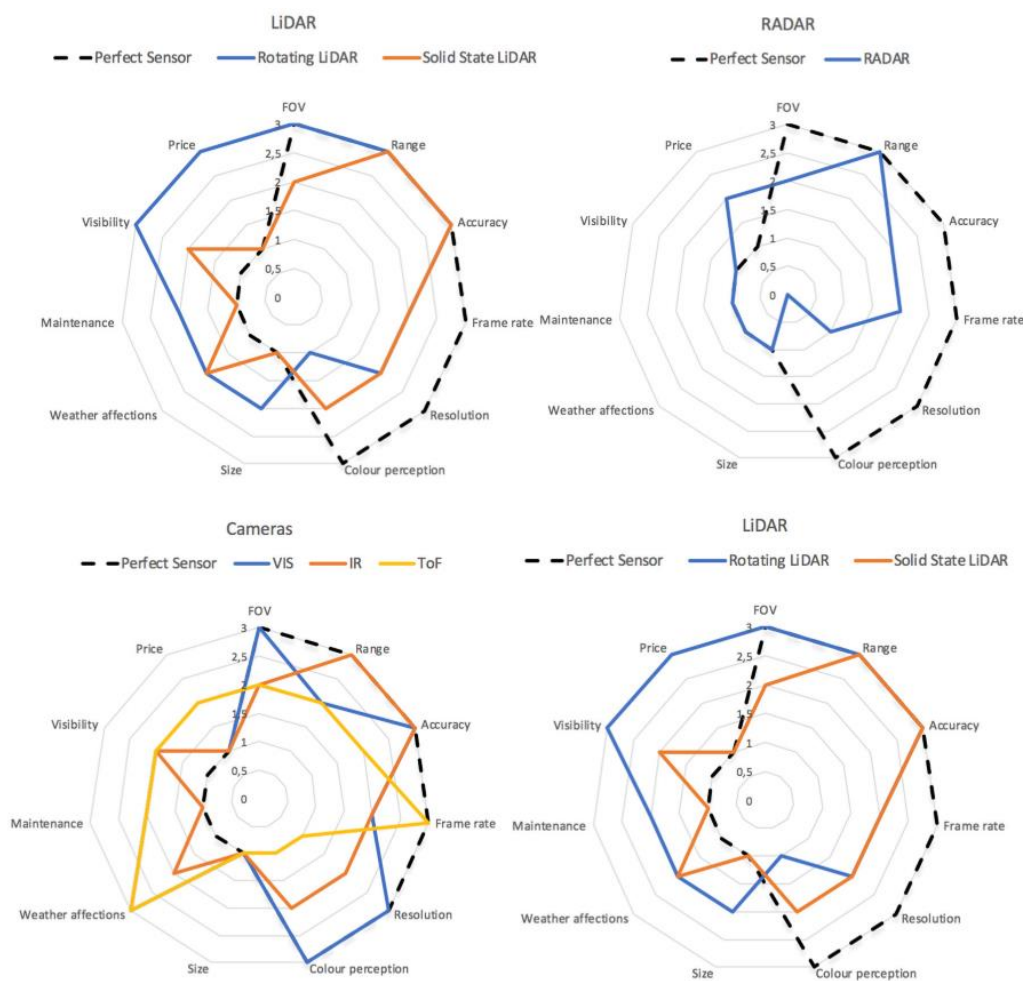
۵-۲- مرور و مقایسه حسگرها

بنابراین با تعدادی از حسگرهای کاربردی در سیستم ادراک وسایل نقلیه خودران آشنا شدیم. همانطور که توضیح داده شد، هر یک از حسگرها در محدوده ای از طول موج‌ها فعالیت می‌کند، که مشابه یا متفاوت با تعدادی از سایر حسگرها می‌تواند باشد. هر یک از محدوده‌ها، کاربرد و مزایای مختص به خود را دارند. در شکل ۲-۳، نمایی کلی از محدوده طول موج مرتبط با هر یک از حسگرهای بررسی شده، نمایش داده شده است.

در شکل ۲-۴ نیز مقایسه ای از ویژگی‌ها و عملکرد سه حسگر اصلی و پرکاربرد دوربین، رادار، و لایدار آورده شده است. معیارهای مورد استفاده شامل هزینه، میدان دید، تاثیرپذیر بودن تحت شرایط آب و هوایی مختلف، دقت، و ... است. همانطور که قابل مشاهده است، هر یک از حسگرها مزایا و محدودیت‌هایی دارند. مزایای هر یک از آنها، محدودیت‌های حاصل از سایر حسگرها را می‌تواند رفع کند. بنابراین، استفاده از ترکیب چندین حسگر از گونه‌های مختلف در یک سیستم ادراک ضروری و لازم است.



شکل ۲-۳- محدوده طول موج مورد استفاده در ادراک وسایل نقلیه خودران [۶].



شکل ۲-۴- مقایسه ویژگی‌های سنسورهای گوناگون [۶].

فصل سوم

رویکردهای ضروری ادراک

رویکردهای ضروری ادراک

برای تجربه رانندگی ایمن در وسایل نقلیه خودران دو رویکرد ضروری وجود دارد که عبارتند از: تقسیم بندی معنایی^۱ و تشخیص شیء^۲.

۱-۳- تقسیم بندی معنایی

خودروهای خودران برای یافتن مسیر خود به صورت اساسی بر تقسیم بندی معنایی تکیه می کنند؛ به این صورت که به هر پیکسل در تصویر یک کلاس اختصاص می دهد و تمام پیکسل هایی که مربوط به یک کلاس خاص هستند، با رنگ یکسان مشخص می شوند. برای مثال، همانطور که در شکل ۱-۳ قابل مشاهده است، وسایل نقلیه با رنگ قرمز، گیاهان و درختان با رنگ سبز و ساختمان ها با رنگ خاکستری نمایش داده می شوند.



شکل ۱-۳- نمونه ای از کاربرد رنگ ها در تقسیم بندی معنایی [۹].

^۱ Semantic Segmentation

^۲ Object Detection

۱-۳- ویژگی‌های مکانی و معنایی

ویژگی‌های مکانی می‌توانند به صورت تصویر یا بردار شامل اطلاعات موقعیت ارائه شوند؛ این ویژگی‌ها در تصاویر به صورت سلول‌های همسایه با نام نواحی^۱ و در بردارها به صورت خط یا نقطه تعریف می‌شوند. علاوه بر حالت‌های تصویر و بردار، ویژگی‌های مکانی در داده‌های لایدار که توسط وسایل نقلیه، ماهواره‌ها، پهپادها و ... نیز یافت می‌شوند. این داده‌ها توسط سیستم اطلاعات جغرافیایی^۲ که ترکیبی از نرم‌افزارهاست تحلیل می‌شوند تا یک تصویر بصری برای درک داده‌های مکانی ایجاد شود. ویژگی‌های مکانی محتوای بصری یک تصویر را با ارتباط دادن آن به خواص دارای سطح پایین‌تر مثل رنگ توصیف می‌کنند. این ویژگی‌ها در خودروهای خودران شامل وسایل نقلیه، علائم جاده، چراغ‌های راهنمایی و ... می‌شوند که تعیین می‌کنند خودرو برای حرکت از یک نقطه و رسیدن به نقطه دیگر، چه مسیری را انتخاب کند، چه زمانی بایستد و چه زمانی جهت حرکت خود را تغییر دهد [۴].

۲-۱-۳- پژوهش‌های مرتبط

یکی از مشکلات در زمینه تقسیم‌بندی معنایی، ارزیابی عملکرد مدل‌های مختلف تقسیم‌بندی معنایی در شرایط گوناگون است. اهمیت ارزیابی به این دلیل است که عملکرد مدل‌ها در صورت بارش و یا روشنایی زیاد، کاهش پیدا می‌کند. راه حل ارائه شده این است که با استفاده از الگوریتم خط لوله شامل لایدار، عملکرد سیستم تقسیم‌بندی معنایی در شرایط گوناگون آزموده می‌شود. یک مجموعه داده بصورت هفتگی به مدت شش ماه جمع‌آوری شده است، که در این داده‌ها، کلاس‌های شناسایی شده برچسب‌گذاری شده‌اند. در ادامه، با استفاده از این روش ارزیابی، دو مدل Enet و Bonnet در شرایط گوناگون بررسی و مقایسه شدند. نتایج حاصل شده نشان داد که هر یک از مدل‌ها با توجه به شرایط محیطی، در برخی شرایط از مدل دیگر عملکرد بهتری دارد [۱۰].

یکی دیگر از مسائل مورد بحث در این حوزه، عدم تشخیص گوشه اجسام در مدل‌ها می‌باشد. راه حل ارائه شده برای رفع این مشکل، استفاده از یک شبکه کارآمد^۳ به عنوان شبکه اصلی و ستون فقرات، و نیز افزودن اطلاعات مکانی غنی‌تر به داده‌ها است. مدل مذکور، عملکرد بهتری از بیشتر مدل‌های دیگر ارائه کرد، و مشکل

^۱ Regions

^۲ Geographic Information System

^۳ EfficientNet

تشخیص گوشه اجسام را برطرف کرد. این افزایش عملکرد، در مورد تشخیص افراد و دوچرخه‌ها بسیار مشهود بوده است [۱۱].

مسئله مطرح دیگر، نیاز به منابع محاسباتی عظیم برای رویکردهای مکانی-زمانی در رهگیری اجسام متحرک بوده است. یک رویکرد زمانی-مکانی سازگار شونده با سرعت خودرو توانست این مشکل را رفع کند. در الگوریتم ارائه شده، از یک خط یک پیکسلی در هر قاب نمونه برداری می‌کند. سپس، نمایه جاده با استفاده از اتصال زمانی خطوط متوالی حاصل می‌شود. در نتیجه استفاده از این روش، داده‌های پردازشی به کسری از ویدئو کاهش پیدا کردند. علاوه بر آن، سازگاری مناسبی با سرعت حرکت خودرو نیز حاصل شد [۱۲].

۲-۳- تشخیص شیء

تشخیص شیء یک رویکرد بنیادین در سیستم تمام وسایل نقلیه خودران است که اشیاء مورد نظر را شناسایی و موقعیت‌یابی می‌کند و یک جعبه محدودکننده دور آنها ایجاد می‌کند. YOLOv2^۲ و YOLOv3^۳ (شکل ۲-۳) از الگوریتم‌های مشهور تعیین‌کننده موقعیت به شمار می‌روند. سایر الگوریتم‌ها از مدل‌هایی بر مبنای یادگیری عمیق^۴ استفاده می‌کنند و مدل‌هایی پیچیده هستند [۴].

^۱ Bounding Box

^۲ You Only Look Once

^۳ Deep Learning



شکل ۳-۲- نمونه‌ای از تشخیص اجسام توسط مدل YOLOv3 [۱۳].

۳-۲-۱- ابر نقاط

ابره‌ای نقاط مجموعه‌ای از نقاط در یک فضای سه‌بعدی هستند که در واقع یک شکل ساده شده از مدل‌های سه‌بعدی را نمایش می‌دهند. هر نقطه نمایانگر کمیت‌های متعددی از جمله مولفه‌های X ، Y ، Z ، مقدار رنگ در قالب RGB و مقدار درخشندگی^۱ است. ابرهای نقاط با اسکن یک شیء یا ساختار به وسیله اسکنرهای لیزری ساخته می‌شوند. کارکرد این اسکنرها به این صورت است که یک پالس لیزر را به سطح شیء ارسال می‌کنند و زمان بازگشت پالس را اندازه‌گیری می‌کنند تا موقعیت دقیق و شکل شیء را بدست آورند و سپس این نقاط برای ایجاد ابر نقاط مورد استفاده قرار می‌گیرند. ابرهای نقاط در وسایل نقلیه خودران به وسیله سنسورهای لایدار جمع‌آوری می‌شوند [۴].

۳-۲-۲- پژوهش‌های مرتبط

یکی از مسائل مهم در تشخیص اجسام، انجام این فرآیند بصورت بلادرنگ است. الگوریتمی بر پایه الگوریتم Real AdaBoost ارائه شده است، که با استفاده از لایدار و تولید ابرهای نقطه‌ای، تشخیص و

^۱ Luminance Value

دسته‌بندی بلادرنگی را ارائه کرده است. آزمایش‌های انجام شده بر روی این الگوریتم، دقت بالغ بر ۹۰ درصدی آن در محدوده ۵۰ متری را گزارش کرده‌اند [۱۴].

نقشه‌های سه‌بعدی حاصل از پردازش لایدها ممکن است دارای نویز باشند، و نتوان بطور کامل به داده‌های آنها اعتماد کرد. راه حل این مشکل، استفاده از دوربین‌ها و سیستم موقعیت‌یاب جهانی در کنار لایدار است. سیستم ارائه شده متشکل از یک شبکه تشخیص خودرو YOLOv2^۱ است. YOLOv2 شامل چندین لایه از مدل‌های شبکه عصبی مصنوعی است، و کمک به تعیین محدوده‌های مرزی هر یک از وسایل نقلیه پیرامون می‌کند. علاوه بر اینها، الگوریتم k قاب پیش رو و گذشته‌انیز مورد استفاده قرار می‌گیرد. این الگوریتم کمک می‌کند که محدوده‌های مرزی از دست رفته را بتوان بازیابی کرد. همچنین بر اساس قاب‌های قبلی و بعدی می‌توان پیش‌بینی مناسبی از موقعیت اجسام در لحظات بعدی بدست آورد [۱۵].

جلوگیری از برخورد، ترمز اضطراری، و نیز برنامه‌ریزی مسیر را می‌توان از مواردی برشمرد که بر تشخیص و دسته‌بندی اجسام متکی هستند. برای انجام مناسب تشخیص شی در مسیر حصول موارد مذکور، دو الگوریتم تشخیص شی با نام‌های YOLOv3 و Viola-Jones ارائه شده‌اند. Viola-Jones با استفاده از تعدادی تشخیص دهنده، اجسام پیرامون را در چهار دسته علائم راهنمایی و رانندگی، عابرین پیاده، چراغ‌های راهنمایی، و خودروها دسته‌بندی می‌کند. مقایسه‌های انجام شده میان این الگوریتم‌ها نشان داد که YOLOv3 از دقت و سرعت پردازش بهتری برخوردار است. اما بطور کلی هر دو الگوریتم، بهبود شایانی نسبت به سایر مدل‌های پیشین کسب کرده‌اند. علاوه بر مزایای مذکور، این دو الگوریتم قابلیت تشخیص و رهگیری چندین جسم در یک لحظه را دارند [۱۶].

^۱ YOLOv2 Vehicle Detection Network

^۲ K-Frames Forward-Backward

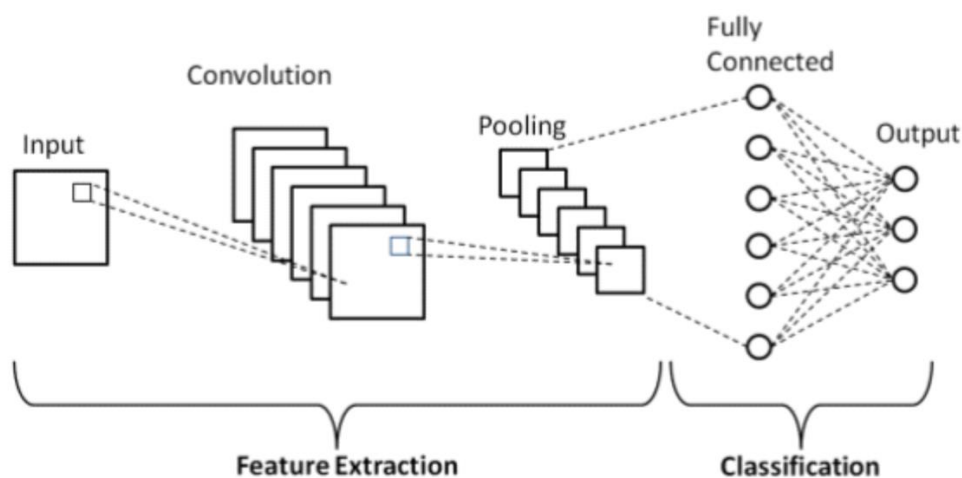
فصل چهارم

یادگیری عمیق

یادگیری عمیق

پیش از این، دو رویکرد اصلی در ادراک وسایل نقلیه خودران را بررسی کردیم. الگوریتم‌ها و مدل‌های مختلف یادگیری عمیق در تشخیص و رهگیری اجسام، و نیز طبقه‌بندی معنایی کاربرد دارد، و در غالب مدل‌ها و الگوریتم‌های این زمینه مورد استفاده قرار می‌گیرد. بنابراین، یادگیری عمیق را می‌توان ستون فقرات سیستم‌های ادراک و تصمیم‌گیری وسایل نقلیه خودران برشمرد.

یکی از مدل‌هایی که کاربرد بسیار گسترده‌ای در زمینه ادراک وسایل نقلیه خودران دارد، مدل CNN^۱ است. این مدل، علاوه بر خودروها، در بسیاری از زمینه‌های دیگر نیز کاربردی است. CNN از سه لایه اصلی تشکیل شده است. این سه لایه به ترتیب لایه کانولوشنال^۲، لایه ادغام^۳، و لایه کاملاً متصل^۴ هستند. لایه کانولوشنال برای استخراج ویژگی‌ها از تصویر ورودی مورد استفاده قرار می‌گیرد. لایه ادغام، نقشه ویژگی‌های تولید شده توسط لایه قبلی را دریافت می‌کند و اندازه آن را کاهش می‌دهد. این کاهش اندازه، به منظور کاهش بار محاسبات انجام می‌شود. در نهایت، لایه کاملاً متصل تصویر را به شکل یک بردار می‌گیرد و خروجی را بصورت تقسیم‌بندی شده به کلاس‌ها ارائه می‌دهد [۴]. لایه‌های CNN در شکل ۴-۱ نیز قابل مشاهده هستند.



شکل ۴-۱ - لایه‌های مختلف CNN [۱۷].

^۱ Convolutional Neural Network

^۲ Convolutional Layer

^۳ A Pooling Layer

^۴ Fully Connected Layer

بسیاری از مدل‌های دیگر شبکه‌های یادگیری عمیق، ساختار مشابهی با CNN دارند که بیان ساختار آنها در اینجا نیازی نیست. پس، با اکتفا به دید کلی ارائه شده به یکی از شبکه‌های یادگیری عمیق، در ادامه به بررسی کاربرد یادگیری عمیق در هر یک از دو رویکرد ادراک وسیله نقلیه خودران پرداخته خواهد شد.

۴-۱- یادگیری عمیق در تقسیم‌بندی معنایی

یکی از موارد حائز اهمیت در این رویکرد این است که کلاس‌های اجسام، دارای اهمیت‌های متفاوتی هستند. مثلاً برای یک خودروی متحرک، عابرین پیاده یا خودروهای دیگر، اهمیت بیشتری از ساختمان‌ها یا آسمان دارند. بنابراین در تقسیم‌بندی و بررسی نیاز است که برای موارد ضروری‌تر، اولویت بالاتری قرار داده شود، و بررسی و پردازش بهتری روی آنها اعمال شود. یادگیری عمیق در اینجا به کمک می‌آید و روش IAL^۱ را ارائه می‌دهد. روش IAL دسته‌بندی‌ها را، و با استفاده از چهار شبکه FCN، ENet، ERFNet، و SegNet، کلاس‌های مختلف را بر اساس میزان اهمیت، به حالت سلسله مراتبی در می‌آورد. نتایج آزمایش‌های انجام شده بر روی این مدل، حصول تقسیم‌بندی بهتر بر روی کلاس‌های دارای اهمیت را نشان می‌دهند [۱۸].

تشخیص و طبقه‌بندی خطوط و لبه‌های جاده در بسیاری از مدل‌ها، پیش از استفاده از یادگیری عمیق، از دقت و اطمینان مناسبی برخوردار نبود. اما بعدها مدلی ارائه شد، که با استفاده از دوربین تک‌چشمی، لایدار، و یک شبکه یادگیری عمیق SegNet توانست در تشخیص خطوط و لبه‌های جاده پیشرفت به سزایی حاصل کند [۱۹]. موارد ذکر شده، تنها گوشه‌ای از دستاوردهای استفاده از یادگیری عمیق در زمینه تقسیم‌بندی معنایی هستند. ده‌ها کاربرد و پیشرفت دیگر نیز با استفاده از مدل‌های یادگیری عمیق در این حوزه حاصل شده‌اند.

۴-۲- یادگیری عمیق در تشخیص اجسام

انواع مختلفی از مدل‌های یادگیری عمیق در تشخیص و رهگیری اجسام مورد استفاده قرار گرفته‌اند. در ادامه دو دسته از یادگیری‌های عمیق استفاده شده در این زمینه معرفی خواهند شد، و کاربردهایی از هر یک برای تشخیص و رهگیری اجسام ذکر می‌شوند.

^۱ 'Importance-Aware Loss' IAL

۱-۲-۴- یادگیری تحت نظارت^۱

این شاخه از یادگیری عمیق، کاربرد گسترده‌ای در زمینه وسایل نقلیه خودران دارد. در یادگیری تحت نظارت، تعدادی داده برچسب‌گذاری شده به مدل داده می‌شوند. برچسب‌گذاری اجسام در سیستم ادراک خودرو، بصورت مشخص کردن محدوده جعبه مانند هر جسم، در کنار یک رنگ که نشان‌دهنده کلاس آن است، صورت می‌گیرد. مدل، با استفاده از این داده‌ها آموزش داده می‌شود. پس از آموزش، مدل توانایی شناسایی و دسته‌بندی اجسام گوناگون را خواهد داشت. در ادامه دو مدل کاربردی از شبکه‌های تحت نظارت، بیان خواهند شد [۴].

سیستم‌های ردیابی مبتنی بر بینایی محدودیت‌هایی دارند. مثلاً، این سیستم‌ها قابلیت رهگیری مجدد یک جسم را پس از گم کردن آن ندارند. روشی ارائه شده است، که این مشکل را حل می‌کند. در این روش از لایدار در کنار الگوریتم شبکه تقسیم بندی سیامی^۲ که شامل دوربین تک‌چشمی است، استفاده می‌شود. این شبکه یادگیری عمیق بصورت تحت نظارت آموزش داده می‌شود. نتایج آزمایش‌ها، بهبود عملکرد در رهگیری مجدد اجسام را نشان دادند [۲۰].

الگوریتم دیگری نیز با استفاده از شبکه‌های LaneNet و Faster-RCNN که از یادگیری تحت نظارت استفاده می‌کنند، ارائه شده است. این الگوریتم برای تشخیص همزمان افراد، وسایل نقلیه، خطوط و وسایل نقلیه غیر موتوری، و نیز تشخیص خطوط برای تقسیم‌بندی آنها می‌باشد. شبکه LaneNet برای دسته‌بندی خطوط جاده، و شبکه Faster-RCNN برای تشخیص افراد و سایر موارد مورد استفاده قرار گرفته‌اند [۲۱]. این الگوریتم علاوه بر یک مجموعه داده شخصی‌سازی شده، در خیابان‌های چین نیز آزموده شده است. نتایج بیانگر توانایی این مدل در مدیریت کردن دقیق سناریوهای بلادرنگ بوده‌اند.

۲-۲-۴- یادگیری تقویتی^۳

این دسته از یادگیری‌های نیز در سیستم ادراک وسایل نقلیه خودران، در جهت تشخیص و رهگیری اجسام کاربرد دارد. در یک مدل یادگیری تقویتی، عامل بصورت مستقل و بدون داشتن داده اولیه آموزشی، با تعامل با محیط خود تلاش می‌کند که تجربه کسب کند و یاد بگیرد. در واقع، عامل تصمیماتی را در هر مرحله

^۱ Supervised Learning

^۲ Siamese Segmentation Network

^۳ Reinforcement Learning

اتخاذ می‌کند و بر اساس سودمندی عمل انجام شده، جایزه دریافت می‌کند. هدف عامل، بیشینه کردن مقدار جایزه دریافت شده است.

یکی از شبکه‌های یادگیری تقویتی عمیق مبتنی بر تشخیص موانع و ناوبری مستقل با نام DQN^۱ [۲۲]، امروزه بسیار مورد توجه واقع شده است. این مدل، ناوبری و تشخیص موانع را با استفاده از یک دوربین و یک لایدار که در سپر جلوی ماشین قرار گرفته‌اند، انجام می‌دهد. این مدل بر روی یک خودروی پرسرعت آزموده شده است. نتایج، نشان دهنده عملکرد مناسب مدل بصورت بلادرنگ بوده‌اند.

یکی دیگر از دستاوردهای یادگیری تقویتی در این زمینه، مدلی بوده است که با استفاده از نتایج شبیه‌سازی، توانسته است در محیط حقیقی شامل موانع، به درستی تصمیم‌گیری و ناوبری کند. در این مدل، پاداش متناسب با هر اقدام عامل با توجه به مقادیر ورودی محاسبه می‌شوند. این پاداش‌ها تحت عنوان مقادیر Q ذخیره می‌شوند. در ادامه با توجه به تعریفی که از یادگیری تقویتی ارائه شد، مدل سعی می‌کند اقداماتی را اتخاذ کند که پاداش دریافت شده بیشینه شود. طبق نتایج آزمایش‌ها، اجتناب از موانع و دنبال کردن خطوط با دقت بسیار خوب با استفاده از این الگوریتم، انجام شده است [۲۳].

^۱ Deep Q Network

فصل پنجم

مرور و جمع بندی

مرور و جمع‌بندی

در این گزارش، اجزای مختلف تشکیل‌دهنده سیستم ادراک یک وسیله نقلیه خودران مورد بررسی قرار گرفت. ابتدا پرکاربردترین حسگرهای استفاده شده در سیستم ادراک بررسی شدند. دسته‌بندی‌های فعال و غیرفعال حسگرها نیز بیان شدند. در ادامه با بررسی حسگرها نتیجه گرفته شد که هر یک از این حسگرها قابلیت و کاربرد خاص خود را دارند. همچنین، هر یک ضعف‌ها و محدودیت‌هایی دارند. مثلاً، رادارها برای فواصل نسبتاً نزدیک کاربردی هستند و می‌توان از آنها برای کاربردهایی مانند سیستم کمک‌کننده پارک استفاده کرد. همچنین، عملکرد آنها در شرایط مختلف آب و هوایی تحت تاثیر قرار نمی‌گیرد. با این حال، رادارها نسبت به سایر حسگرها از دقت کمتری برخوردارند. همچنین، پالسهای مورد استفاده ممکن است دچار جهش شوند. برای سایر حسگرها نیز معایب و مزایا به همین شکل بررسی شدند.

در بخش بعد، دو رویکرد اساسی در سیستم‌های ادراک، یعنی تقسیم‌بندی معنایی و تشخیص جسم، مورد بررسی قرار گرفتند. با بررسی‌ها نتیجه گرفته شد که تقسیم‌بندی معنایی برای شناخت محیط پیرامون و طبقه‌بندی اجسام انجام می‌شود. همچنین، تشخیص اجسام برای شناسایی و رهگیری اجسام در یک محیط کاربرد دارد. بنابراین، همانطور که پیش‌تر هم ذکر شد، این دو رویکرد وابسته به یکدیگر هستند، و عدم استفاده از هر یک از آنها، می‌تواند درک سیستم را دچار ضعف و محدودیت کند. برای هر یک از این دو رویکرد، الگوریتم‌های مختلف ارائه شده بررسی شدند و مزیت هر یک، مورد بحث قرار گرفت. مثلاً نتیجه گرفته شد که الگوریتم بر پایه Real AdaBoost توانسته است تشخیص و دسته‌بندی بصورت بلادرنگ را با دقت بسیار بالا ارائه دهد.

در بخش بعدی به یادگیری عمیق پرداخته شد و گستردگی استفاده از آن در زمینه ادراک خودروهای خودمختار نشان داده شد. مدل‌ها و الگوریتم‌های گوناگون یادگیری عمیق در زمینه تقسیم‌بندی معنایی و تشخیص جسم بیان شدند، و پیشرفت و رفع محدودیت‌های حاصل از هر یک، به عنوان گواهی برا کاربرد گسترده و اهمیت یادگیری عمیق، ذکر شد. همچنین انواع مختلفی از یادگیری‌های عمیق مانند یادگیری تقویتی و یادگیری تحت نظارت مورد بررسی قرار گرفتند. تفاوت اصلی این دو نوع از یادگیری در این است که یادگیری تحت نظارت با استفاده از آموزش توسط داده‌های برچسب‌گذاری شده صورت می‌گیرد. درحالی که یادگیری تقویتی، مدل بدون استفاده از داده‌های کمکی، با استفاده از کسب تجربه و پاداش در محیط پیرامون، بصورت مستقل یادگیری می‌کند.

در این گزارش، تعدادی پژوهش و الگوریتم ارائه شده در حوزه ادراک خودروهای خودمختار بیان شدند. همانطور که مشهود است، مشکلات زیادی وجود داشته‌اند که توسط این راهکارها رفع شدند. اما همچنان ضعف‌های زیادی در این زمینه وجود دارد که نیازمند پژوهش و پیشرفت بیشتری است. بنابراین، این موضوع همچنان پرچالش و مطرح است و راه برای ارائه راه‌حل‌های جدیدتر وجود دارد. بنابراین با وجود اینکه همچنان وسایل نقلیه خودمختار به سطحی نرسیده‌اند که بطور کامل مستقل شوند، اما به دلیل مزایای کثیری که به همراه خواهند داشت، محبوبیت و توجه به آنها روزافزون است و بدون شک، الگوریتم‌ها و روش‌های جدیدتری بصورت گسترده و در آینده‌ای نزدیک مطرح می‌شوند، تا یک پله به خودمختاری کامل خودروها نزدیک‌تر شویم.

منابع و مراجع

- [1] W. Schwarting, J. Alonso-Mora, and D. Rus, "Planning and decision-making for autonomous vehicles," *Annual Review of Control, Robotics, and Autonomous Systems*, vol. 1, pp. 187-210, 2018.
- [2] Self-Driving Cars Explained. Available online: <https://www.ucsusa.org/clean-vehicles/how-self-drivingcars-work> (accessed on 29 October 2018).
- [3] Thrun, S. Toward Robotic Cars. *Commun. ACM* 2010, 53, 99–106.
- [4] H.-H. Jebamikyous and R. Kashef, "Autonomous vehicles perception (avp) using deep learning: Modeling, assessment, and challenges," *IEEE Access*, vol. 10, pp. 10523, 2022.
- [5] www.semanticscholar.org/paper/Stereo-Vision-Based-Advanced-Driver-Assistance-Jung-Lee/Od6d203201400f28f542964cb05432d5279a775a/figure/0
- [6] F. Rosique, P. J. Navarro, C. Fernández, and A. Padilla, "A systematic review of perception system and simulators for autonomous vehicles research," *Sensors*, vol. 19, p.2019.
- [7] Olmeda, D.; de la Escalera, A.; Armingol, J.M. Far infrared pedestrian detection and tracking for night driving. *Robotica* 2011, 29, 495–505.
- [8] Gade, R.; Moeslund, T.B. Thermal cameras and applications: A survey. *Mach. Vis. Appl.* 2014, 25, 245–262.
- [9] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele, "The cityscapes dataset for semantic urban scene understanding," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun.2016
- [10] W. Zhou, J. S. Berrio, S. Worrall, and E. Nebot, "Automated evaluation of semantic segmentation robustness for autonomous driving," *IEEE Trans. Intell. Transp. Syst.*, vol.21, no. 5, pp. 1951–1963, May 2020.
- [11] J. S. Lee and T. H. Park, "Semantic segmentation with improved edge detail for autonomous vehicles," in *Proc. IEEE 16th Int. Conf. Autom. Sci. Eng. (CASE)*, Aug. 2020, pp. 520–525.
- [12] G. Cheng, J. Y. Zheng, and M. Kilicarslan, "Semantic segmentation of road profiles for efficient sensing in autonomous driving," in *Proc. IEEE Intell. Vehicles Symp. (IV)*, Jun. 2019, pp. 564–569.

- [13] M. Gluhakovic, M. Herceg, M. Popovic, and J. Kovacevic, "Vehicle detection in the autonomous vehicle environment for potential collision warning," in Proc. Zooming Innov. Consum. Technol. Conf. (ZINC), May 2020, pp. 178–183.
- [14] M. Yoshioka, N. Suganuma, K. Yoneda, and M. Aldibaja, "Real-time object classification for autonomous vehicle using LIDAR," in Proc. Int. Conf. Intell. Informat. Biomed. Sci. (ICIIBMS), Nov. 2017, pp. 210–211.
- [15] M. Feng, S. Hu, G. Lee, and M. Ang, "Towards precise vehicle-free point cloud mapping: An on-vehicle system with deep vehicle detection and tracking," in Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC), Oct. 2018, pp. 1288–1293.
- [16] J. Ciberlin, R. Grbic, N. Teslić, and M. Pilipović, "Object detection and object tracking in front of the vehicle using front view camera," in Proc. Zooming Innov. Consum. Technol. Conf. (ZINC), May 2019, pp. 27–32.
- [17] www.upgrad.com/blog/basic-cnn-architecture/
- [18] B. Chen, C. Gong, and J. Yang, "Importance-aware semantic segmentation for autonomous vehicles," IEEE Trans. Intell. Transp. Syst., vol. 20, no. 1, pp. 137, 2019.
- [19] K. L. Lim, T. Drage, and T. Braunl, "Implementation of semantic segmentation for road and lane detection on an autonomous ground vehicle with LIDAR," in Proc. IEEE Int. Conf. Multisensor Fusion Integr. Intell. Syst. (MFI), Nov. 2017, pp. 429–434.
- [20] L. Zhao, M. Wang, S. Su, T. Liu, and Y. Yang, "Dynamic object tracking for self-driving cars using monocular camera and LIDAR," in Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS), Oct. 2020, pp. 10865–10872.
- [21] J. Yang, C. Wang, H. Wang, and Q. Li, "A RGB-D based real-time multiple object detection and ranging system for autonomous driving," IEEE Sensors J., vol. 20, no. 20, pp. 11959–11966, Oct. 2020.
- [22] A. R. Fayjie, S. Hossain, D. Oualid, and D.-J. Lee, "Driverless car: Autonomous driving using deep reinforcement learning in urban environment," in Proc. 15th Int. Conf. Ubiquitous Robots (UR), Jun. 2018, pp. 896–901.
- [23] T. Okuyama, T. Gonsalves, and J. Upadhyay, "Autonomous driving system based on deep Q learning," in Proc. Int. Conf. Intell. Auto. Syst. (ICoIAS), Mar. 2018, pp. 201–205.