

استفاده از میدان‌های عصبی پویا در یک الگوریتم بینایی ماشین به منظور ردگیری همزمان اهداف

شیوا کامکار^{۱*}، حمید ابریشمی مقدم^۲، رضا لشگری^۳

چکیده

ردگیری همزمان چند شیء یکی از موضوعات مهم در زمینه بینایی ماشین و پایه کاربردهای بسیاری مانند انواع سامانه‌های نظارتی انسانی و حیوانی است. از اصلی‌ترین چالش‌های الگوریتم‌های این حوزه، مرتبط‌سازی داده‌ها به ویژه هنگام بروز انسداد می‌باشد. از آنجا که انسان‌ها عملکرد مناسبی در مواجهه با این چالش دارند، انتظار می‌رود استفاده از الگوریتم‌های ملهم از مغز در ردگیری همزمان اهداف بتواند به بهبود کارایی این گونه از سامانه‌ها منجر شود. میدان‌های عصبی پویا سازگاری بالایی با عملکرد نوروئی و شناختی مغز انسان دارند. ما در این مقاله از این میدان‌ها بهره بردیم تا مانند حافظه کاری انسان در نگهداری و پردازش اطلاعاتی مانند مکان اشیاء، جهت و سرعت آن‌ها ایفای نقش کنند. استخراج این اطلاعات با استفاده از شگردهای بینایی ماشین صورت می‌گیرد. ارزیابی این روش با مقایسه عملکرد آن با روش‌های اخیراً توسعه داده شده روی مجموعه داده حاوی ویدیوهای ضبط‌شده از حرکات آزادانه لاروهای قزل‌آلا که در دسترس همگان قرار دارد انجام شده است. نتایج نشان داد که روش پیشنهادی نه تنها از روش‌های رقیب عملکرد بهتری دارد، بلکه تقریباً در همه موارد به ویژه بعد از انسداد قادر است مرتبط‌سازی داده‌ها را به درستی انجام دهد.

کلیدواژه‌ها

ردگیری همزمان اشیاء، میدان‌های عصبی پویا، مرتبط‌سازی داده‌ها، انسداد، ویدیوهای میکروسکوپی.

۱ مقدمه

ردگیری همزمان اهداف یکی از کاربردهای مهم بینایی ماشین در دنیای واقعی است [۱]–[۴]. سامانه‌هایی که مسئول نظارت بر

این مقاله در آذرماه ۱۳۹۹ دریافت، در اسفندماه همان سال بازنگری و متعاقباً پذیرفته شد.

این پژوهش با حمایت مالی مرکز مطالعات و همکاری‌های علمی بین‌المللی وزارت علوم، تحقیقات و فناوری بر اساس قرارداد شماره ۱۴۸۳ انجام شده است.

^۱ دانش‌آموخته دکتری مهندسی کامپیوتر، دانشکده مهندسی کامپیوتر، دانشگاه صنعتی خواجه نصیرالدین طوسی.

رایانامه: sh.kamkar@email.kntu.ac.ir

^۲ استاد دانشکده مهندسی برق، دانشگاه صنعتی خواجه نصیرالدین طوسی، مرکز مطالعات و همکاری‌های علمی بین‌المللی وزارت علوم، تحقیقات و فناوری.

رایانامه: moghaddam@kntu.ac.ir

^۳ استادیار پژوهشکده علوم و فناوری‌های پزشکی، دانشگاه شهید بهشتی.

رایانامه: r_lashgari@sbu.ac.ir

انسان‌ها در فرودگاه، بانک یا مترو هستند، باید قادر باشند موقعیت دقیق اهداف مورد ردگیری را در هر لحظه به درستی گزارش کنند. سامانه‌های نظارت ترافیکی نیز بایستی مکان دقیق هر خودرو را در هر لحظه به درستی ثبت کنند تا انواع تخلفات رانندگی به درستی تشخیص داده شود. سامانه‌هایی که به نظارت و تحلیل اهداف میکروسکوپی می‌پردازند یا سامانه‌هایی که وظیفه تعامل با انسان‌ها یا حیوانات را دارند نیز از دیگر مواردی هستند که نیاز به ردگیری همزمان چند هدف دارند.

علی‌رغم پیشرفت‌های این حوزه، هنوز چالش‌هایی مطرح است که باعث شده الگوریتم‌های موجود کارایی مناسبی در رویارویی با برخی از سناریوهای واقعی نداشته باشند. یکی از مهم‌ترین این چالش‌ها، وقوع انسداد^۱ است. انسداد زمانی رخ می‌دهد که قسمتی یا کل یک هدف توسط اهداف یا سایر عوامل محیطی پوشانده شود. در این صورت مشخص نمودن مکان دقیق هدف مشکل است. به خصوص اگر دو یا چند هدف در یک انسداد واحد

¹ Occlusion

حافظه بلند مدت بازیابی می‌گردد. به این ترتیب کارایی الگوریتم ردگیری بهبود می‌یابد. همچنین این ویژگی به ردگیر کمک می‌کند که در مواجهه با انسداد به خوبی شیء هدف را از عامل مزاحم جدا نماید. به این منظور ابتدا نقاط کلیدی و ویژگی‌های تصویر استخراج می‌گردد، سپس با مقایسه آنها با ویژگی‌های مربوط به شیئی که از قبل در حافظه ذخیره شده است آن را از سایر نواحی جدا می‌کند. در مرحله بعد ویژگی‌ها را از شیء که در قاب^۴ جدید مکان آن مشخص شده است استخراج کرده و هر دو حافظه را به روز می‌کند. از آنجا که در این روش، اختلاف ویژگی‌های شیء و پیرامون آن در نظر گرفته نمی‌شود در محیط‌هایی که شیء و زمینه شباهت زیادی با هم دارند دچار مشکل شده و قادر نیست موقعیت دقیق شیء را گزارش کند. این نحوه استفاده از حافظه، الهام گرفته از مدل ASMM^۵ می‌باشد که با ذخیره، به‌روزرسانی و حذف برخی اطلاعات، مشابه مکانیزم حافظه انسان عمل می‌کند. در [۹] نیز از مدل یادشده برای لحاظ کردن حافظه در الگوریتم پیشنهادی استفاده شده است. ویژگی‌های شیء در این الگوریتم با در نظر گرفتن پیرامون آن استخراج می‌شوند و از این جهت می‌توان گفت این روش در شرایطی که شیء و زمینه شباهت زیادی با هم دارند کمتر دچار مشکل می‌شود. هرچند تغییرات سریع زمینه می‌تواند کارایی آن را تحت تاثیر قرار دهد. در مقاله دیگری از مدل حافظه مشابه با فوق استفاده شده است تا اطلاعات ظاهری شیء که از ابتدا تاکنون ثبت شده است به خاطر سپرده و یا در صورت لزوم فراموش شود. این کار باعث می‌شود که ردگیر آسان‌تر با ظاهر جدید شیء سازگار شود و انسداد را مدیریت کند [۱۰]. ردگیر یادشده از فیلتر ذرات برای تکمیل فرآیند ردگیری استفاده می‌کند.

پژوهش ارائه شده در [۱۱] از محدود روش‌هایی است که با در نظر گرفتن حافظه، یک روش ردگیری همزمان چند شیء ارائه کرده است. این روش قادر است چند روبات مثل هم را که ظاهر آنان از قبل شناخته شده است تشخیص داده، ردگیری و بازشناسی کند. به این منظور یک شبکه LSTM واحد طراحی شده که اطلاعات اشیاء را حفظ می‌کند. زیربخش‌هایی از این شبکه مسئول اطلاعات هر شیء می‌باشد و به این ترتیب قادر است با وجود انسدادهای طولانی مدت کارایی مطلوبی داشته باشد. روش پیشنهادی در کاربرد خاصی ارائه شده است و قابل تعمیم به همه مسائل دنیای واقعی نیست. در مجموع، ردگیرهایی که تنها به مدل کردن حافظه پرداختند به نسبت سایر روش‌ها عملکرد بهتری در رویارویی با چالش‌های انسداد و تغییرات ظاهری شیء داشتند. اما برای مواجهه با سایر چالش‌ها مانند تغییرات سریع روشنایی و تغییرات ناگهانی زمینه که در برخی مواقع منجر می‌شود شیء بسیار شبیه به زمینه به نظر برسد، لازم است روش مناسبی برای استخراج شیء از زمینه به کار گرفته شود.

شرکت داشته باشند، در بسیاری از الگوریتم‌ها هویت اهداف با هم اشتباه گرفته می‌شود. این باعث می‌شود پس از انسداد، به اهداف برچسب اشتباه نسبت داده شود حتی اگر مکان آن‌ها به درستی گزارش شده باشد.

از سوی دیگر، مطالعات روان‌فیزیک^۱ نشان داده‌اند که انسان‌ها در ردگیری اهداف حین انسداد به خوبی عمل می‌کنند و چالش یادشده تاثیر اندکی بر عملکرد آن‌ها دارد [۵]–[۷]. بنابراین، می‌توان دریافت که مکانیزم‌های شناختی انسان و نحوه پردازش اطلاعات در مغز به گونه‌ی موثرتری نسبت به الگوریتم‌های موجود با این چالش روبه رو می‌شوند. این می‌تواند راه‌حلی باشد برای بهبود کارایی الگوریتم‌های بینایی ماشین در ردگیری همزمان اشیاء. بر این اساس، ما در این مقاله از میدان‌های عصبی پویا^۲ بهره گرفتیم تا روشی برای ردگیری همزمان اهداف ارائه دهیم. این میدان‌ها منطبق بر نحوه پردازش نورونی و شناختی مغز هستند و به ما کمک می‌کنند تا الگوریتمی سازگارتر با عملکرد مغز برای ردگیری اهداف ارائه دهیم.

ساختار مقاله بدین شرح است. ابتدا در بخش دوم به بیان پیشینه‌ای از الگوریتم‌های ارائه شده در این حوزه می‌پردازیم. سپس، میدان‌های عصبی پویا که اساس الگوریتم پیشنهادی است را در بخش سوم معرفی می‌کنیم. بخش چهارم به جزئیات روش پیشنهادی اختصاص دارد. ارزیابی روش و مقایسه آن با سایر روش‌ها در بخش پنجم صورت گرفته است. بخش ششم حاوی جمع‌بندی از پژوهش انجام شده می‌باشد.

۲ پیشینه پژوهش

الگوریتم‌های متعددی تا کنون برای ردگیری همزمان چند هدف ارائه شده است. اما تمرکز ما در این بخش بیشتر روی الگوریتم‌هایی است که سعی کردند از شیوه عملکرد مغز در روش پیشنهادی خود الهام بگیرند. در برخی مواقع به الگوریتم‌های ردگیری یک شیء نیز اشاره نمودیم چراکه اغلب قابل تعمیم به منظور ردگیری چند شیء هستند.

مطالعات رفتاری نشان داده است که انسان‌ها در هنگام ردگیری همزمان چند هدف از حافظه کاری^۳ استفاده می‌کنند. این حافظه مسئول نگه داشتن و به‌روز کردن موقتی اطلاعاتی مانند مکان و ویژگی‌های ظاهری اهداف است و به فرد کمک می‌کند تا هر هدف را به درستی دنبال کرده و بازشناسی نماید [۲].

در [۸] از حافظه کوتاه‌مدت و بلندمدت انسان الهام گرفته شده است تا اطلاعات تغییرات ظاهری شیء در طول زمان حفظ شود. وقتی که شیء از نظر ظاهری تغییرات کمی دارد پردازش در حافظه کوتاه مدت به سرعت انجام می‌شود و زمانی که در ظاهر شیء تغییرات ناگهانی اتفاق بیفتد، اطلاعات به خاطر سپرده شده در

^۴ Frame

^۵ Atkinson-Shiffrin Memory Model (ASMM)

^۱ Psychophysics

^۲ Dynamic Neural Fields

^۳ Working Memory

برای ردگیری همزمان اشیاء استفاده نموده است. این سامانه در واقع از کنار هم قرار دادن چند ردگیر که هر کدام مسئول ردگیری یک شیء مجزا هستند تشکیل شده است. از توجه زمانی برای یادگیری مدل از روی نمونه‌های صحیح از شیء در طول زمان استفاده شده و به این ترتیب از نظر مواجهه با انسداد به خوبی عمل کرده است. شبکه عصبی پیشنهادی شامل تعدادی لایه مشترک است که حاوی اطلاعات کل قاب هستند. به ازای هر شیء هدف، یک شاخه از شبکه مسئول کد کردن اطلاعات آن شیء و جدا کردنش از زمینه است و با استفاده از نمونه‌های درست و غلط به شکل برخط آموزش می‌بیند. مرجع [۱۸] نمونه‌ای دیگر از استفاده از توجه مکانی و زمانی در قالب یک شبکه عصبی مصنوعی است.

برخی از پژوهشگران حوزه علوم شناختی بر این باورند که دو سامانه مجزا هنگام ردگیری همزمان چند شیء در مغز درگیر می‌شوند. یک سامانه موازی برای ردگیری موقعیت و مکان اشیاء متحرک و حفظ آن‌ها در حافظه کاری بصری و یک سامانه متوالی برای متمایز کردن اشیاء از هم از روی ویژگی‌هایشان و پیوند دادن آن‌ها به موقعیت مکانی‌شان [۱۹]. این با ایده‌ای که سامانه‌های کانونی و محیطی بصری^۷ را از هم متمایز می‌کند همخوانی دارد. بر طبق این ایده، سامانه کانونی به پردازش اطلاعات دریافت شده از گودی مرکزی چشم^۸ می‌پردازد که نیازمند دقت و جزئیات بالا است مانند شناسایی اشیاء یا جستجوی بصری. سامانه محیطی مسئول کنترل و جهت‌گیری‌های فضایی هنگام حرکت است. به این ترتیب ردگیری همزمان اشیاء یکسان معیار ارزیابی عملکرد سامانه محیطی و ردگیری همزمان اشیاء متمایز معیار ارزیابی هر دو سامانه کانونی و محیطی می‌باشد. این ایده مشابه تمایز دو مسیر اصلی در بینایی نیز هست که یکی مسئول بازشناسی اشیاء و دیگری مسئول تعیین موقعیت اشیاء است.

راهبرد استفاده شده در برخی از الگوریتم‌های ردگیری نیز سازگار با این یافته‌هاست. در زمینه ردگیری همزمان چند شیء، در [۲۰] الگوریتمی مبتنی بر شبکه‌های پیچشی ارائه داده شده است که دو مسیر تحتانی^۹ و فوقانی^{۱۰} در آن به شکل جداگانه مدل شده اند. بخشی از این شبکه از مکانیزم توجه نیز بهره برده و قادر است با در نظر گرفتن ویژگی‌های اهداف با چالش‌های تغییر شکل و حرکت اشیاء مقابله کند. هرچند این روش اشیائی که بسیار نزدیک هم به شکل مشابه حرکت می‌کنند را یک شیء در نظر می‌گیرد و به دلیل نداشتن حافظه امکان ردگیری مجدد شیئی که موقتاً از دست رفته است در آن وجود ندارد.

بهره‌گیری از اطلاعاتی مانند ساختار نورون‌ها در مغز و نحوه ارتباط آن‌ها در قالب شبکه‌های عصبی مصنوعی در الگوریتم‌های ردگیری بسیار معمول است. در [۲۱] محققان با استفاده از مدل

توجه^۱، دیگر مکانیزم شناختی است که نقش موثری در ردگیری همزمان اهداف توسط انسان بازی می‌کند. مرسوم است که این مکانیزم را به دو نوع توجه پایین به بالا^۲ و بالا به پایین^۳ تقسیم می‌کنند. توجه پایین به بالا محرک محور است به این معنا که توجه بیننده به سمت بخشی از صحنه جلب می‌شود که از نظر ویژگی‌های سطح پایینی مانند رنگ و شدت روشنایی نسبت به نواحی همسایه‌اش برجسته باشد. اما توجه بالا به پایین هدف محور است به این معنا که توجه فرد به سمتی هدایت می‌شود که متناسب با هدف از پیش تعریف شده‌ای در ذهن وی باشد.

استفاده از مدل‌های محاسباتی توجه در الگوریتم‌های ردگیری اخیراً مورد استقبال فراوانی قرار گرفته است [۱۲]. در این پژوهش‌ها اغلب برای تمییز شیء از زمینه از مدل‌های تشخیص برجستگی اشیاء^۴ استفاده می‌شود و از آنجا که این مدل‌ها اساساً بر پایه مشاهدات و آزمایش‌های حوزه علوم اعصاب و نظریات مطرح بودند می‌توان آنها را نمونه‌ای از روش‌های ملهم از بیولوژی دانست.

در برخی از پژوهش‌ها روشی دو مرحله‌ای ملهم از توجه بصری برای ردگیری یک شیء در زمینه بینایی ماشین مطرح شده است. در این روش ابتدا از توجه پایین به بالا برای ایجاد تمایز بین شیء مورد نظر و زمینه استفاده شده است و در مرحله بعد روشی ملهم از توجه بالا به پایین برای پیدا کردن شیء مورد نظر در قاب‌های بعدی به کار گرفته شده است [۱۳]–[۱۵]. استفاده از این شیوه به تنهایی نمی‌تواند چالش انسداد جزئی یا کلی را حل نماید و امکان گم کردن شیء برای آن وجود دارد. همچنین قادر به ردگیری مجدد شیئی که موقتاً از دست رفته است نمی‌باشد. در [۱۶] یک شبکه عصبی نوسانی^۵ برای مدل کردن ردگیری همزمان چند شیء پیشنهاد شده است. معماری این شبکه اولین بار برای مدل کردن توجه پیشنهاد شد. محققان آن را به گونه‌ای تعمیم دادند که بتواند ردگیری یک شیء را با در نظر گرفتن بازخوردها و تعاملاتی که طبق یافته‌های علوم اعصاب، نورون‌های دریافت‌کننده هدف با نورون‌های دریافت‌کننده پیرامون آن دارند انجام دهد. در این روش روی ویژگی‌های پیرامون شیء نیز تمرکز شده و از آن استفاده شده تا پردازش ویژگی در شیء را بهبود بخشد. به همین دلیل با چالش انسداد بهتر مواجه شده است اما همچنان امکان ردگیری مجدد شیئی که به طور موقت از دست رفته است برای آن وجود ندارد.

روش [۱۷] را می‌توان جزو معدود پژوهش‌هایی دانست که در زمینه ردگیری همزمان چند شیء، توجه را در الگوریتم پیشنهادی لحاظ کرده است. در این مقاله یک چارچوب مبتنی بر شبکه‌های عصبی پیچشی^۶ ارائه شده است که از مکانیزم توجه مکانی-زمانی

¹ Attention

² Bottom-up

³ Top-down

⁴ Salient Object Detection

⁵ Oscillatory Neural Network

⁶ Convolutional Neural Network

⁷ Foveal and Peripheral Vision

⁸ Fovea

⁹ Ventral

¹⁰ Dorsal

شرکت کننده های انسانی باید تعدادی از آن ها را به عنوان هدف در میان سایرین دنبال کنند. این مدل، تنها از اطلاعات مکانی دایره ها استفاده می کند و قادر است نتایج به دست آمده حاصل از تحلیل سیگنال های مغزی را نیز توجیه کند [۲۸].

همچنین در [۲۹]، الگوریتمی برای نمایش و ردگیری اشیاء هندسی ساده روی یک میز توسط روبات پیشنهاد شده است. بخشی از این الگوریتم به ردگیری همزمان حداکثر دو شیء موجود می پردازد. در صورتی که شیء سوم به میز اضافه شود ظرفیت حافظه سامانه به حدی می رسد که احتمال فراموش کردن یکی از آنها زیاد می شود. این مشابه فراموشی حافظه کاری است که در انسان اتفاق می افتد. اشیاء تعریف شده در این سامانه ساده هستند و تنها سه ویژگی طول، عرض و رنگ برای تعریف اشیاء در صحنه کافی است. همچنین مسیر حرکت متمایزی نسبت به هم ندارند، از هم فاصله دارند و مسئله انسداد تعریف نشده است. اشیاء صلب هستند و امکان ردگیری اشیاء در حالی که ظاهر آن و محیط زمینه تغییر می کند لحاظ نشده است و سامانه قادر به یادگیری و تطبیق پذیری با گذر زمان نبوده است.

در ادامه به معرفی روش هایی می پردازیم که عملکرد آن ها با عملکرد روش پیشنهادی در بخش ارزیابی مقایسه شده است. در [۳۰] به ارائه روشی برای ناحیه بندی و ردگیری همزمان لاروهای موجود در ویدیو استفاده شده است. برای تخمین پس زمینه ویدیوها از مدل ترکیب گاوسی^۵ با در نظر گرفتن ۵۰۰ قاب اول ویدیو استفاده شده است. برای ویدیوهایی که کوتاه تر هستند یک قاب را برای چند بار در نظر می گیرند. ناحیه بندی تصاویر با استفاده از آستانه گذاری آتسوف^۶ و عملیات ریخت شناسی انجام می گیرد و لاروهای هر قاب معرفی می شوند. در مرحله بعد به متناظر کردن لاروهای به دست آمده با تعریف یک ماتریس هزینه می پردازند. هزینه تناظر هر دو شیء، به اندازه فاصله اقلیدسی بین مراکز ثقل شان تعریف شده است. در نهایت، با استفاده از الگوریتم مجارستانی^۷ که هر شیء را تنها به یک شیء دیگر مرتبط می کند، بهترین تناظر بین اشیاء مشخص می شود به شکلی که مجموع مربعات فاصله ها کمینه باشد. گاهی اتفاق می افتد که لاروی در چند قاب به درستی تشخیص داده یا ناحیه بندی نمی شود و در قاب های آتی دوباره مشخص می گردد. در این صورت بین اطلاعات ردگیری شیء فاصله می افتد. در این مواقع با استفاده از الگوریتم نزدیک ترین همسایه مسیریایی که در برخی قاب ها بینشان فاصله افتاده به هم متصل می شوند. هر چند این اتصال در صورتی که فاصله اتفاق افتاده از حدی بیشتر باشد انجام نمی گیرد. در روشی دیگر برای هر شیء به شکل خودکار به دنبال اثر انگشت هایی می گردند تا بتوانند آن ها را در طول ویدیو به درستی ردگیری کند [۳۱]. در مرحله اول، الگوریتم از بخش هایی از ویدیو که در

HMAX به عنوان روشی که لایه های سلول های ساده^۱ و سلول های پیچیده^۲ در مسیر تحتانی مغز را شبیه سازی کرده است به استخراج ویژگی های بیولوژیکی سطح پایین پرداختند و روشی برای یادگیری معرفی کردند که مکان شیء را در هر قاب مشخص نماید. الگوریتم پیشنهاد شده در [۲۲]، نمونه ای دیگر از استفاده از مدل HMAX برای بیان ظاهر اشیاء در یک الگوریتم ردگیری است. در [۲۳] نیز با استفاده از یک شبکه پیش خور^۳ پیچشی دو لایه (لایه اول سلول های ساده و لایه دوم سلول های پیچیده)، ردگیری پیشنهاد شد که قادر است اطلاعات ظاهری اشیاء هدف را به شکل برخط به روز کند. هر چند در این روش از وصله های زمینه هم فیلترهایی استخراج می شود اما عدم مقایسه زمینه پیرامون شیء با شیء باعث می شود که متمایز کننده ترین فیلترها و ویژگی ها برای جدا کردن شیء از زمینه مورد استفاده قرار نگیرند. همچنین انسداد و گم کردن شیء چالش های مهمی برای این روش محسوب می شوند. استفاده ابزاری از شبکه های عصبی مصنوعی بدون تلاش برای سازگاری بیشتر روش با عملکرد مغز به ویژه در الگوریتم های اخیر متداول است [۲۴]–[۲۶].

در بین الگوریتم های ارائه شده تا کنون پژوهش های محدودی هستند که از میدان های عصبی پویا به عنوان ابزاری که نه تنها از لحاظ ساختاری با پردازش نورونی سازگار است بلکه از لحاظ شناختی نیز قابلیت مدل کردن بسیاری از مسائل شناختی را دارد استفاده نمودند [۲۷]. این میدان ها دو مفهوم پویایی و پایداری^۴ که در سامانه عصبی انسان وجود داشته و در رفتارهای شناختی نیز مشهود هستند را به خوبی پوشش می دهد. نتایج شبیه سازی با این میدان ها نشان داده است که به خوبی رفتار سوژه های انسانی را مدل می کنند. داده های رفتاری حاصل از بررسی تاثیر سرعت و طول دوره ردگیری حین ردگیری همزمان چند شیء توسط سوژه های انسانی مطابقت دارد. در واقع این نظریه، هم با داده های رفتاری و هم با داده های نورونی مطابقت دارد.

یک مدل پیشنهادی بر این اساس از سه میدان عصبی تشکیل شده است: میدان دریافت ورودی، میدان بازدارنده و میدان حافظه کوتاه مدت. میدان ورودی به دریافت اطلاعات مکانی اهداف می پردازد. موقعیت هر هدف در میدان حافظه کوتاه مدت حفظ می شود. میدان بازدارنده نیز اطلاعات موقعیت هدف را دارد و با تعامل بازدارندگی کمک می کند که در گذر زمان و تغییر مکان شیء، مکان قبلی شیء نادیده گرفته شده و مدل به دنبال کردن هدف در مکان های جدید بپردازد. این گونه تعاملات برانگیزانندگی و بازدارندگی در سامانه عصبی انسان متداول است. مدل مذکور قادر است تا حدی رفتار انسان ها را هنگام انجام آزمایشات ساده روان-فیزیک شبیه سازی کند. این آزمایشات اغلب شامل دایره هایی هستند که در زمینه یکنواخت به مدت چند ثانیه حرکت می کنند.

^۵ Gaussian Mixture Model

^۶ Otsu

^۷ Hungarian

^۱ Simple Cells

^۲ Complex Cells

^۳ Feed-forward

^۴ Stability

سازی مجموعه‌های نورونی می‌پردازد و آن‌ها را به شکلی موقعیت می‌بخشد که نورون‌های دارای مقدار مشابه در یک ویژگی، در میدان در کنار هم قرار بگیرند. هر جا که تحریکی اتفاق بیفتد یک قله^۴ به شکل بی‌درنگ رخ می‌دهد و منجر به تصمیم رفتاری می‌شود. میدان‌های مختلف می‌توانند به روش‌های متفاوت با هم ترکیب شوند و مدل‌های با معماری پیچیده برای شبیه‌سازی عملیات نورونی-شناختی مختلف مانند توجه و حافظه ایجاد کنند. از آنجا که میدان‌های عصبی پویا سامانه‌هایی پویا هستند، یک ویژگی کلیدی آنها این است که تحت تاثیر پیشینه خود هستند. برای مثال، در یک مکانیزم یادگیری ساده، برخی از نورون‌هایی می‌توانند روی حافظه اثری از خود بر جا بگذارند و منجر شوند که در آینده آن نورونها به تحریک‌های مشابه پاسخ قوی‌تری بدهند [۳۲].

چگونگی پردازش مغز با استفاده از اطلاعات حسگرها و موتورها نیازمند یک زبان نظری است که بتواند همه این زمینه‌ها را به هم مرتبط کند. این، فرآیندی است که نیاز به پردازش‌های نورونی دارد که به شکل پیوسته در طول زمان تکامل می‌یابند و به شکل برخط با هم و با سامانه اطلاعاتی حسگرها مرتبط هستند. این پیوستگی در حالات و زمان نیازمند مدل‌سازی به زبان سامانه‌های پویا است.

نظریه میدان‌های پویا ریشه در نظریه شبکه عصبی دارد و توصیف ریاضی چگونگی فعالیت نورون‌ها در شبکه عصبی بازگشتی در زمان پیوسته می‌باشد. در حقیقت، اصل پیوستگی مکان-زمان مفهومی است که میدان‌های عصبی پویا اضافه کرده است. همچنین روی پایداری حالات شناختی تاکید دارد و ایجاد عملکرد شناختی جدید را حاصل ناپایداری دینامیک نورونی می‌داند. به دلیل وجود پایداری در میدان‌های عصبی پویا، آنها با وجود تغییرات کم و نطفه^۵ دار در ورودی همچنان پایداری خود را حفظ می‌کنند و این خلاف چیزی است که در شبکه‌های عصبی مصنوعی وجود دارد. چون در ذات این شبکه‌ها آستانه‌گذاری‌های متعدد وجود دارد.

در مدل‌های شبکه عصبی مانند مدل هاپفیلد، جاذب‌های زیادی همزمان وجود دارند. سامانه از یک شرایط اولیه (به واسطه تحریک ورودی) شروع به کار می‌کند و به نزدیک‌ترین جاذب همگرا می‌شود. همچنین هیچ عنصر درونی در سامانه وجود ندارد که بیان کند شبکه به برون داد^۶ رسیده و محاسبات آن به اتمام رسیده رسیده است. بلکه باید توسط سامانه دیگری بررسی شود که خارج از شبکه نورونی قرار می‌گیرد. در حالی که در سامانه‌های عصبی پویا رفتار در زمان پیوسته ظاهر می‌شود. دینامیک‌های نورونی پیوسته تکامل می‌یابند و ورودی‌هایی دریافت می‌کنند که ممکن است متغیر باشند. در این شرایط نقطه جاذب حرکت می‌کند. از آنجا که دینامیک‌های نورونی همواره در یک نقطه جاذب پایدار

آن شیء مذکور دچار انسداد نشده است تعدادی قاب مرجع استخراج می‌کند. سپس، به تجمیع قاب‌های استخراج شده‌ای که با احتمال زیاد مربوط به یک شیء مشخص هستند می‌پردازد. در نهایت، هر شیء با مجموعه‌ای از قاب‌های مرجع که آن را در حالات مختلف نشان می‌دهند شناسایی می‌گردد. برای به دست آوردن یک اثر انگشت برای هر شیء، آن را در هر تصویر از مجموعه ناحیه‌بندی می‌کند. سپس یک نقشه سطح روشنایی و یک نقشه تباین^۱ استخراج می‌شود که مستقل از تبدیلات انتقال و چرخش هستند. سایر قاب‌های ویدیو با مقایسه با نقشه‌های روشنایی به دست آمده مشخص می‌شوند. مسیر حرکت اشیاء با متصل کردن مرکز اشیاء تشخیص داده شده به هم تشکیل می‌گردد. در مواقع بروز انسداد بین مسیرها وقفه می‌افتد که در این صورت به تخمین فاصله خالی از روی مکان شیء قبل و بعد از وقفه می‌پردازند. استفاده از این رویکرد تخمین مسیر در هنگام بروز انسداد باعث می‌شود تا روش‌های مذکور به شکل بی‌درنگ قادر به تعیین موقعیت اهداف حین انسداد نباشند. همچنین الگوریتم در مرتبط-سازی داده‌ها^۲ در بسیاری از مواقع که سرعت حرکت اهداف یکنواخت نیست و از توزیع گاوسی پیروی نمی‌کند دچار مشکل می‌شود.

هدف اصلی این مقاله، یافتن روشی برای بهبود چالش‌های انسداد و مرتبط‌سازی داده‌ها در مسئله ردگیری همزمان اهداف است. انگیزه فرعی، آن است که ببینیم آیا می‌توان میدان‌های عصبی پویا را به عنوان ابزاری که سازگاری بالایی با پردازش نورونی و شناختی مغز دارد به گونه‌ای تعمیم داد که در الگوریتم‌های خودکار مورد استفاده قرار گیرد و موجب بهبود کارایی آن‌ها شود؟ از آنجا که انسان عملکرد قابل قبولی در هنگام رویارویی با دو چالش یادشده دارد انتظار داریم استفاده از میدان‌های عصبی پویا نیز موثر واقع شود. به این منظور برای اولین بار این میدان‌ها را به گونه‌ای تعمیم دادیم تا در یک کاربرد واقعی حوزه بینایی ماشین (و نه در محیط‌های بسیار ساده و شبیه‌سازی شده) مانند ردگیری لاروها مورد استفاده قرار بگیرند. از آنجا که میدان‌های عصبی پویا اساس روش پیشنهادی را تشکیل می‌دهد و به ندرت از سوی پژوهشگران حوزه بینایی ماشین مورد توجه قرار گرفته‌اند، در ادامه به معرفی اجمالی این مفهوم می‌پردازیم.

۳ مقدمه‌ای بر میدان‌های عصبی پویا

نظریه میدان‌های عصبی پویا، مجموعه‌ای از مفاهیمی است که فرآیندهای رفتاری، نورونی و شناختی را به شکل محصولی از سامانه‌های خودسازمانده^۳ که در طول زمان تکامل می‌یابند تعریف می‌کند. تاثیر همه سامانه‌ها به یک اندازه است و برای ایجاد یک رفتار لازم است همه آنها همگرا شوند. میدان عصبی پویا به شبیه-

⁴ Peak

⁵ Noise

⁶ Output

¹ Contrast

² Data Association

³ Self-organized

شود، نرخ منفی منجر به کاهش تا رسیدن به جاذب می شود. به علاوه، میدان می تواند ورودی هایی داشته باشد که از حسگرها می آیند و یا حاصل تعاملات نوروئی هستند که منجر به نسخه کامل-تری از معادله می گردد [۳۳].

دینامیک های نوروئی به شکل پیوسته با دریافت ورودی هایی که ممکن است متغیر باشند (به این دلیل که رفتار منجر به حرکت حسگرها می شود) تکامل می یابند. ورودی های متغیر با زمان ممکن است نقاط جاذب را حرکت دهند اما از آنجا که دینامیک های نوروئی همواره در یک نقطه جاذب هستند، پس این تغییرات و نقاط جاذب را دنبال می کنند. این فرآیند دائما ادامه می یابد تا یک ناپایداری اتفاق بیفتد. در این صورت دینامیک های نوروئی به سمت جاذب جدید می روند و رفتار تغییر می کند. در جاذب نوروئی جدید، سامانه دوباره آماده دنبال کردن تغییرات بعدی ورودی است. تمام این فرآیند به شکل خودکار انجام می شود.

بیان نوروئی که تاکنون در مورد آن صحبت کردیم باید در یک فضای پیوسته بیان شود که منجر به قرارگیری زنجیره ای از متغیرهای فعالیت و ایجاد میدان فعالیت می شود. قله های فعالیت پایه های تئوری میدان های پویا هستند. این قله ها دو مفهوم را به دنبال دارند. اول اینکه بیانگر این حقیقت است که موردی وجود دارد که قادر است شبکه های نوروئی دیگر را در میدان تحت تاثیر قرار دهد. دوم اینکه مکان قله در میدان نشان می دهد که پارامترها، ویژگی ها و موقعیت ادراکی چگونه هستند.

در بسیاری از مواقع از جمله در پژوهش فعلی، از میدان های با ابعاد بالای یک استفاده می شود. با فرض اینکه میدانی دو بعدی داریم که دو بعد مکانی x و y را داراست، می توان تابع فعالیت دو بعدی را به شکل معادله زیر نوشت.

$$\tau u(x,y) = -u(x,y) + h + s(x,y) + \iint k(x-x',y-y')g(u(x',y'))dx'dy' \quad (4)$$

که در آن، k تابع تعاملی است که مشخص می کند چطور همسایه های دور و نزدیک بر یکدیگر اثر می گذارند. این تابع تعاملی نیز دو بعد خواهد داشت. در شرایطی که هر دو بعد میدان از لحاظ کیفی شبیه باشند مانند ابعاد فضا، میزان تعامل را می توان تابعی از فاصله در نظر گرفت. همچنین تابع تعاملی می تواند حاصل تفاضل دو تابع گاوسی به شکل کلاه مکزیکی باشد:

$$k(x,y) = c_{exc} \cdot \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_{x,exc}^2} + \frac{y^2}{\sigma_{y,exc}^2}\right)\right] - c_{inh} \cdot \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_{x,inh}^2} + \frac{y^2}{\sigma_{y,inh}^2}\right)\right] - c_{glob} \quad (5)$$

چنین تابعی به معنای تحریک در فواصل کوتاه و بازدارندگی در فواصل طولانی تر است. این تابع مانند یک کرنل روی برون داد میدان (که در واقع برون داد تابع سیگموئید در هر نقطه است) کانوالو می شود. حاصل در نهایت با محرک های ورودی (s)، سطح استراحت (h) و فعالیت میدان در همان لحظه جمع می

هستند، سامانه با دریافت ورودی هایی که طی زمان تغییر می کنند، نقطه جاذب را دنبال می کند. این فرآیند دائما ادامه می یابد، هر زمان سامانه ناپایدار شد به جاذب جدید همگرا می شود و در آن نقطه هم آمادگی تغییرات جدید را دارد. بنابراین، تمام نقاط در فضایی که پیوسته تعریف شده است می توانند پاسخی باشند که سامانه در آنجا پایدار شده است. همه این وقایع به شکل خودکار رخ می دهد و هیچ سامانه دیگری نیاز نیست که وضعیت آن را بررسی کند که آیا محاسبات آن به اتمام رسیده است یا خیر.

در نظریه میدان های پویا فرض می شود که فعالیت جمعیتی از نوروئی ها در نواحی مشخصی از مغز است که منجر به الگوهای رفتاری می شود. متغیر فعالیت نوروئی، یک عدد حقیقی برون داد یک تابع سیگموئید^۱ است (g) با شیب (β) که اگر بالاتر از آستانه صفر (یعنی مقادیر مثبت) باشد به سایر متغیرهای فعالیت منتقل می شود و به این ترتیب به ایجاد ارتباطات مختلف بین مجموعه های نوروئی کمک می کند [۳۳]:

$$g(u) = \frac{1}{1 + \exp(-\beta u)} \quad (1)$$

وضعیت درونی سامانه عصبی مرکزی دائما به شکل برخط در طول زمان تغییر می کند. به همین دلیل متغیرهای فعالیت را با تابعی پیوسته از زمان نشان می دهند ($u(t)$). دینامیک نوروئی، پاسخ معادله دیفرانسیلی زیر است [۳۳]:

$$\tau \dot{u} = f(u) \quad (2)$$

که در آن τ نرخ تغییرات u است و τ یک ثابت مثبت است که واحد زمان را تعریف می کند (مانند ثانیه یا میلی ثانیه). $f(u)$ تابع همواری از فعالیت است و ما باید تابعی را که دوره زمانی درستی از فعالیت نشان می دهد بیابیم. معادله فوق، یک سامانه پویا برای متغیر فعالیت $u(t)$ است و پاسخ های آن مسیرهای زمان پیوسته ای (یعنی مشتق پذیر) از فعالیت هستند که نرخ تغییرات آنها تابعی از فعالیت فعلی است. لازم است این تابع را تعیین کنیم. از آنجا که سامانه عصبی مرکزی از اتصالات زیادی برخوردار است. یک متغیر فعالیت تحت تاثیر بسیاری از سایر متغیرهای فعالیت است. بیشتر اوقات بخشی از اثرهایی که دریافت می کند ناسازگار هستند و فعالیت را از وضعیت کنونی دور می کنند. بدون پایداری، سامانه عصبی قادر به حفظ یک شرایط خاص از همه تاثیرات ممکن نیست. بنابر آنچه گفته شد معادله را به شکل معادله زیر بازنویسی می کنیم.

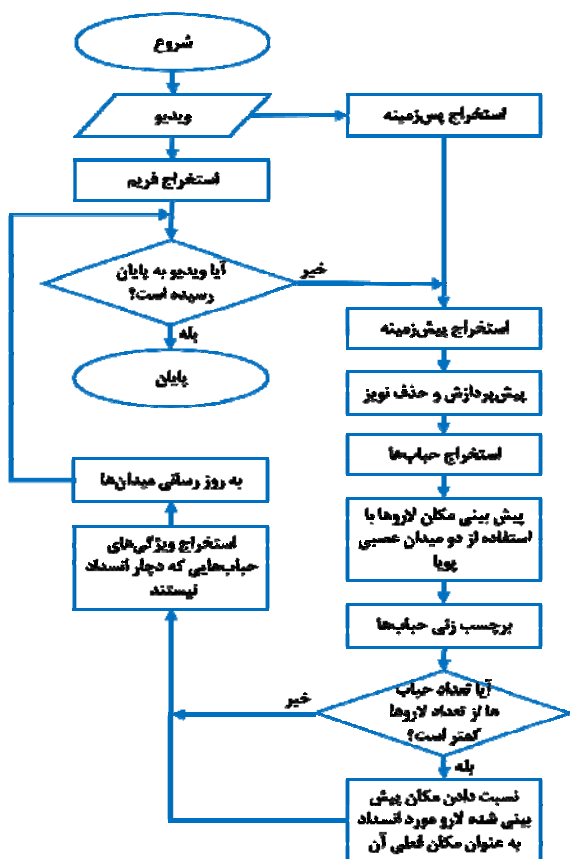
$$\tau \dot{u} = -u + h \quad (3)$$

h بیانگر فعالیت سطح استراحت است و یک جاذب^۲. چراکه فعالیت با گذر زمان از هر سطح فعالیت اولیه در مجاورت این نقطه ثابت به آن همگرا می شود. اگر فعالیت از سطوح پایین تر جاذب شروع شود نرخ مثبت تغییرات منجر به افزایش آن تا رسیدن به نقطه ثابت می شود و اگر از سطحی بالاتر از آن آغاز

¹ Sigmoid

² Attractor

لاروها و موقعیت اولیه آن‌ها را استخراج می‌نماید. در مرحله بعد نوبت به بازنمایی و ردگیری لاروها می‌رسد.



شکل ۱ روندنمای الگوریتم پیشنهادی

الگوریتم ۱ شبه کد روش پیشنهادی

پس‌زمینه را با انتخاب پرتکرار ترین مقدار برای هر پیکسل در ۵۰ فریم اول تخمین بزن

میدان عصبی DNF1 را برای ردگیری مکان اشیاء بر اساس (۴) و (۵) تعریف کن

میدان عصبی DNF2 را برای ردگیری سرعت و راستای اشیاء بر اساس (۴) و (۵) تعریف کن

برای هر قاب از ویدیو ورودی {

پیش‌زمینه را با روش تفاضل زمینه محاسبه کن.

با اعمال ریخت‌شناسی نوفه‌ها را حذف و قاب را پیش‌پردازش کن

اجزای پیوسته تصویر (حباب‌ها) را استخراج کن

ویژگی‌های مکان و راستا را از هر یک استخراج کن

اگر در فریم اول هستیم، {اشیاء را متناظر با حباب‌ها قرار بده}

در غیر این صورت

{مکان هر لارو را با استفاده از DNF1 و DNF2 تخمین بزن}

حباب‌ها را برچسب بزن

اگر تعداد حباب‌ها از تعداد اشیاء کمتر بود،

{برای اشیائی که دچار انسداد شدند، مکان تخمینی را به DNF1 بده}

برای اشیائی که دچار انسداد نشدند {

مکان حباب‌ها را به DNF1 بده

ویژگی سرعت اشیاء را استخراج کن

ویژگی‌های سرعت و راستا را به DNF2 بده}

روش‌های متعددی برای بازنمایی لاروهای قزل آلا به عنوان اشیائی غیر صلب وجود دارد. ما در این روش، از اطلاعات مکان سر لاروها، راستا و سرعت آن‌ها برای بازنمایی، ردگیری و

شود تا روند تغییر میدان را با گذر زمان نشان دهد. میدان‌هایی که ما در الگوریتم پیشنهادی از آن بهره جستیم نیز به طریق مشابه عمل می‌کنند.

یکی از مسائل مهم در پیاده‌سازی میدان‌های عصبی پویا این است که اساساً مغز مانند یک رایانه آنالوگ عمل می‌کند. حال آنکه کامپیوترهای معمول امروزی دیجیتال هستند و مفهوم زمان در آن‌ها گسسته است. هر چند ساعت^۱ رایانه در مقایسه با مقیاس زمانی مورد استفاده در الگوریتم قادر است بازه زمانی بسیار کوچکی را تعریف کند، اما لازم است تا معادلات پویا مطرح شده تاکنون را به شکل گسسته تقریب بزنیم. به این منظور در پیاده‌سازی، با در نظر گرفتن بازه زمانی بسیار کوچک Δt ($\Delta t \ll \tau$) بین هر دو گام متوالی، مشتق تابع فعالیت را با استفاده از روش اولیئر^۲ به شکل معادله زیر تقریب می‌زنیم:

$$\ddot{u}(t_i) \approx \frac{u(t_i) - u(t_{i-1})}{\Delta t}, t_i = i\Delta t, i = 0, 1, 2, 3, \dots \quad (6)$$

۴ الگوریتم پیشنهادی

الگوریتم پیشنهادی ما برای ردگیری همزمان اشیاء، مبتنی بر میدان‌های عصبی پویا است که مفاهیم پایه آن در بخش قبل بیان شد. هر چند ما از این روش ملهم از پردازش مغز به عنوان هسته پردازشی الگوریتم پیشنهادی استفاده می‌کنیم اما همچنان از شگردهای بنیایی ماشین نیز بهره می‌بریم. روند نمای روش پیشنهادی در شکل ۱ و شبه کد آن در الگوریتم ۱ آورده شده است. در اولین مرحله، استخراج پس‌زمینه صورت می‌گیرد. به این منظور از روش تفاضل پس‌زمینه استفاده نمودیم. از آنجا که ویدیو-های مجموعه داده لاروها با دوربین ثابت و در شرایط آزمایشگاهی ضبط شده‌اند به منظور یافتن پس‌زمینه، حدود ۵۰ قاب اول هر ویدیو را در نظر گرفتیم و از میان ۵۰ مقداری که برای هر پیکسل به دست می‌آید پرتکرارترین مقدار را به آن پیکسل نسبت دادیم. تصویری که پیکسل‌های آن به شکل مذکور محاسبه شدند تصویر پس‌زمینه در نظر گرفته شده است. تفاضل این تصویر از هر قاب به تصویر پیش‌زمینه منجر می‌شود که حاوی لاروها می‌باشد. با انجام عمل ریخت‌شناسی^۳ فرسایش به حذف نوفه‌های موجود در تصویر پرداختیم و با انجام عمل ریخت‌شناسی بستن به یکنواخت-سازی بخش‌های باقیمانده پرداختیم. به منظور استخراج لاروها، تمام اجزای پیوسته موجود در تصویر را استخراج نمودیم. شایان ذکر است که در فریم اول ویدیوها هیچ انسدادی وجود ندارد. بنابراین تعداد حباب^۴‌های استخراج شده در قاب اول برابر تعداد لاروها هستند. به همین دلیل روش پیشنهادی ما در قاب اول هیچ اطلاعاتی از کاربر نمی‌گیرد و به شکل بدون ناظر و خودکار،

¹ Clock

² Euler

³ Morphology

⁴ Blob

گیرد که با در نظر گرفتن معیار تعریف شده، در مجموع کمترین هزینه را داشته باشیم.

در صورتی که انسداد رخ داده باشد، تعداد حباب‌ها از تعداد لاروها کمتر است و در هنگام برچسب زنی، حداقل یک حباب بیش از یک برچسب می‌گیرد که هر یک بیانگر لاروی است که در انسداد شرکت داشته است.

در مرحله بعد، به استخراج اطلاعات لاروها در فریم جدید می‌پردازیم. چنانچه تعداد حباب‌های تصویر از تعداد اهداف کمتر باشد (بروز انسداد)، برای اهداف شرکت کننده در انسداد، مکان پیش‌بینی شده را در میدانی که مسئول حفظ و دنبال کردن مختصات سر لارو است به‌روز می‌کنیم. برای سایر لاروهایی که دچار انسداد نبودند مکان هر لارو در قاب جدید گزارش می‌شود و از طرف دیگر اطلاعات سرعت و راستای آن‌ها نیز در قاب جدید استخراج شده و در میدان‌ها به‌روز می‌گردد. الگوریتم به همین ترتیب به تشخیص مکان لاروها در قاب بعدی می‌پردازد.

برای پارامترهای تعریف‌شده در این روش هنگام پیاده‌سازی، به شکل تجربی و با در نظر گرفتن معنی و مفهومی که هر پارامتر دارد مقادیری در نظر گرفته شده که لیست آن‌ها در جدول ۱ آمده است.

شایان ذکر است که روش پیشنهادی با انتخاب مجموعه دیگری از مقادیر برای پارامترها نیز می‌تواند همچنان کارایی مناسبی را ارائه دهد. کافی است مفاهیم اصلی ذکر شده در بخش ۳ در

انتخاب پارامترها لحاظ شود. برای مثال لازم است مقادیر $\sigma_{x,inh}$

و $\sigma_{y,inh}$ نسبت به $\sigma_{x,exc}$ و $\sigma_{y,exc}$ بزرگتر باشد. چراکه نورون‌های با ویژگی مشابه که در میدان در نزدیکی هم قرار می‌گیرند همدیگر را تحریک و نورون‌های حساس به ویژگی‌های متفاوت تر که در میدان از هم دورترند یکدیگر را مهار می‌کنند. نمونه دیگر

بزرگتر بودن a_s نسبت به h است که باعث می‌شود تابع فعالیت به حدی برسد که خروجی تابع سیگموئید مثبت شده و بتواند تعامل ایجاد کند. اطلاعات بیشتر در مورد مفاهیم پارامترها و نحوه صحیح انتخاب آن‌ها در [۳۳] به تفصیل بیان شده است.

جدول ۱ پارامترهای مورد استفاده در تنظیمات میدان‌های پویا

پارامترهای تابع تعاملی		پارامترهای میدان‌ها	
۵	C_{exc}	۲۰	τ
۲	C_{inh}	-۵	h
۰	C_{glob}	۴	β
۱۰	$\sigma_{x,exc}$	پارامترهای ورودی‌ها	
۱۰	$\sigma_{y,exc}$	۲۰	a_s
۱۵	$\sigma_{x,inh}$	۱۰	$\sigma_{s,x}$
۱۵	$\sigma_{y,inh}$	۱۰	$\sigma_{s,y}$

بازشناسی استفاده می‌کنیم. مکان سر لارو با انجام عمل ریخت شناسی فرسایش^۱ روی حباب استخراج می‌شود. راستای لارو زاویه خطی است که مرکز حباب را به سر آن متصل می‌کند. سرعت لارو میانگین سرعت آن در سه فریم متوالی پیشین است. دو میدان عصبی پویا به اندازه ۴۰۰ در ۴۰۰ مسئول حفظ و دنبال کردن این اطلاعات هستند. یکی از آن‌ها معرف مختصات طول و عرض سر لاروها و دیگری معرف سرعت و جهت حرکت آن‌ها می‌باشد. هر دو میدان طبق معادله (۴) تکامل می‌یابند و تابع تعامل مطابق رابطه (۵) برای آن‌ها تعریف شده است. هر میدان قله‌ای متناظر با هر هدف دارد که بیانگر اطلاعات آن هدف است. به طور مثال وجود قله در مختصات (۱۰۰ و ۲۰۰) میدان اول به این معناست که سر یک لارو در این مختصات واقع شده است. اطلاعات لاروها به عنوان ورودی در هر دو میدان به ازای هر قاب به‌روز می‌شود. ورودی میدان‌ها به شکل یک تابع گاوسی به مرکز مقدار تعیین شده p تعریف می‌شود:

$$s(x,y) = a_s \cdot \exp\left[-\frac{1}{2}\left(\frac{(x-p_x)^2}{\sigma_{s,x}^2} + \frac{(y-p_y)^2}{\sigma_{s,y}^2}\right)\right] \quad (۷)$$

در این رابطه، a_s بیانگر ضریب تابع گاوسی ورودی به مرکز (p_x, p_y) و پراکندگی $\sigma_{s,x}^2$ در جهت x و $\sigma_{s,y}^2$ در جهت y است.

با ورود قاب جدید، اطلاعات مربوط به لاروها در قاب(های) قبل در میدان‌ها موجود است. با استفاده از این اطلاعات، مکان هر لارو در قاب جدید به صورت زیر تخمین زده می‌شود.

$$X_i = X_{i-1} + vD \quad (۸)$$

در این معادله، X_i, X_{i-1}, v و D به ترتیب مکان لارو در قاب قبلی، مکان آن در قاب فعلی، سرعت لارو و جهت لارو در قاب قبل هستند.

اکنون که برای فریم جدید، هم حباب‌ها و هم مکان‌های پیش بینی شده لاروها را بر اساس اطلاعات قبلی آن‌ها داریم به برچسب زنی حباب‌ها می‌پردازیم. به این منظور، دو مورد را به عنوان معیار در نظر می‌گیریم. اول اینکه مکان حباب متناظر با یک لارو باید نزدیک به مکان پیش‌بینی شده آن باشد (از نظر فاصله اقلیدسی) و دوم اینکه میزان تغییر جهت لارو نسبت به قاب قبلی نمی‌تواند زیاد باشد. وزن این دو مورد در معیار، وابسته به شرایط تغییر میکند. به این ترتیب که در شرایط عادی که انسدادی رخ نداده است هر دو مورد وزن یکسان (۰,۵) دارند. اما وقتی که انسداد رخ دهد تشابه جهت لارو با وزن ۰,۶ نقش قوی تری نسبت به نزدیکی مکانی با وزن ۰,۴ پیدا می‌کند. چراکه وقتی اهداف روی هم می‌افتند راستای آن‌ها اطلاعات مفیدتری برای تمییز و برچسب زنی در اختیار قرار می‌دهد. برچسب‌زنی حباب‌ها به شکلی صورت می‌-

¹ Erosion

۵ ارزیابی

سوم و چهارم به ترتیب نتیجه روش‌های [۳۱]، [۳۴] و [۳۰] هستند. در ستون دوم ردگیری مطابق روش [۳۱] انجام شده اما برای ناحیه‌بندی از روش [۳۰] استفاده شده است. ستون پنجم نتایج الگوریتم پیشنهادی ما را نشان می‌دهد.

همانطور که در جدول دیده می‌شود روش پیشنهادی برای همه ویدیوها نسبت به سایر روش‌ها بهترین کارایی را داشته است. در حقیقت، راهبردهای مورد استفاده در این روش نسبت به سایرین برتری‌هایی دارد. در ادامه به تحلیل روش و مقایسه با همتایان آن می‌پردازیم. اول: روش پیشنهادی ما با در نظر گرفتن این حقیقت که دوربین ثابت بوده است و ویدیوها در شرایط آزمایشگاهی ضبط شده‌اند، پس‌زمینه را تخمین زده است. بدیهی است که برای تخمین روش به سایر کاربردهایی که چالشی چون تغییرات روشنایی در آن‌ها مطرح است، نیاز به استفاده از روش‌های تخمین پس‌زمینه به شکل پویا خواهیم داشت.

دوم: در هنگام بروز انسداد، حباب‌های اهداف مورد انسداد با هم همپوشانی دارند و یک حباب واحد را تشکیل می‌دهند. در این صورت، حباب مذکور توسط روش پیشنهادی ما برچسب همه اهدافی که در آن انسداد شرکت داشتند را می‌گیرد. این موضوع هرچند اندکی دقت مکان‌یابی را کاهش می‌دهد (نسبت به زمانی که مکان دقیق اهداف گزارش شوند). اما باعث می‌شود اهداف از دست نروند و بعد از اتمام انسداد، هویت آن‌ها به درستی تشخیص داده شود. به همین دلیل نرخ تشخیص اشتباه الگوریتم از لحاظ تعداد دفعاتی که یک هدف با هدف دیگری اشتباه گرفته شود صفر است (جدول ۳). این در حالی است که در سایر روش‌های رقیب، در لحظه بروز انسداد اطلاعاتی از شیء مورد انسداد در اختیار نیست. در بیشتر موارد، به ردگیری هدف تنها قبل و بعد از انسداد می‌پردازند. این روش‌ها مکان هدفی که دچار انسداد است را در لحظه انسداد، با درونیابی موقعیت آن به شکل برون‌خط در قبل و بعد از انسداد گزارش می‌کنند. در این رویکرد گاهی اشتباهاتی رخ می‌دهد به این دلیل که تنها معیار متناظر کردن مسیر هدف قبل و بعد از انسداد نزدیکی مکانی آن‌هاست.

سوم: روش پیشنهادی ما همچنان به تخمین مکان اشیاء می‌پردازد، حتی زمانی که دچار انسداد هستند. این باعث می‌شود مرتبط‌سازی داده‌ها بعد از انسداد به درستی انجام گیرد. در این شرایط، به دلیل اینکه دو هدف متمایز که با هم انسداد دارند تنها با یک حباب متناظر هستند نمی‌توان از اطلاعات همان قاب برای تخمین مکان آن‌ها استفاده نمود. به همین دلیل از اطلاعاتی که بلافاصله قبل از شروع انسداد در میدان‌ها ذخیره شده است بهره می‌بریم. این تخمین مکان حین ردگیری همزمان چند شیء با یافته‌های برخی پژوهشگران علوم شناختی که معتقدند انسان در ردگیری از اطلاعات سرعت و برون‌یابی بهره می‌برد نیز سازگار است [۳۵]. بر این اساس، ما از این روش تخمین مکان نه تنها هنگام بروز انسداد، بلکه در سایر مواقع نیز استفاده می‌کنیم.

به منظور ارزیابی روش پیشنهادی، عملکرد آن را روی مجموعه داده ارائه شده در [۳۰] بررسی و با سایر روش‌های ارائه شده روی این مجموعه داده مقایسه نمودیم. مجموعه داده یادشده حاوی ویدیوهای ضبط شده از حرکات طبیعی لاروهای ماهی قزل‌آلا می‌باشد که در شرایط آزمایشگاهی و ظرفی بسته قرار دارند. بنابراین تعداد لاروها با گذر زمان تغییر نمی‌کند. پیاده‌سازی با استفاده از نرم‌افزار MATLAB2019b و با کمک جعبه‌ابزار Cosivina^۱ که مشخصاً برای کار با میدان‌های عصبی پویا طراحی شده است در رایانه‌ای با پردازنده ۲٫۱۰ گیگاهرتز و ۱۲ گیگابایت حافظه با دسترسی تصادفی و سیستم عامل ویندوز ۸ صورت گرفت. دو معیار اصلی برای ارزیابی یک روش برای ردگیری همزمان اشیاء صحت (MOTA) و دقت (MOTP) است که مطابق روابط ۹ و ۱۰ تعریف می‌گردد:

$$MOTA = 1 - \frac{\sum_i (m_i + fp_i + mme_i)}{\sum_i g_i} \quad (9)$$

$$MOTP = \frac{\sum_{i,t} d_i^t}{\sum_i c_i} \quad (10)$$

در معادلات فوق، m_i ، fp_i ، mme_i و g_i به ترتیب تعداد اهدافی که به عنوان هدف گزارش نشده‌اند^۲، تعداد اشیائی که به اشتباه هدف گزارش شده‌اند^۳، تعداد تطابق‌های اشتباه^۴ و تعداد اشیاء موجود در قاب t می‌باشند. c_i تعداد تطابق‌ها^۵ در قاب t است و d_i^t فاصله شیء i ام از مکان حقیقی آن است. داشتن مقدار بالاتر برای صحت به معنای داشتن خطای کمتر است و داشتن مقدار پایین‌تر برای دقت به معنای موقعیت‌یابی دقیق‌تر اهداف توسط الگوریتم است.

به منظور ارزیابی روش پیشنهادی، برون‌داد آن را با برون‌داد روش‌های [۳۰] و [۲۶] مقایسه نمودیم. همچنین، روش دیگری که نتایج را با آن مقایسه نمودیم برون‌داد نرم‌افزار تجاری LoliTrack^۶ است [۳۴] که مشخصاً برای تحلیل رفتار حیوانات به خصوص ماهی‌ها که شکل لوکوموتیووار دارند طراحی شده است. نتایج ردگیری این نرم‌افزار روی مجموعه داده لاروهای قزل‌آلا نیز در [۳۰] گزارش شده است.

روش‌هایی که بیان شد مشخصاً برای ردگیری اهداف زیستی به ویژه ماهی‌ها طراحی شده‌اند. به همین دلیل ما آن‌ها را به عنوان روش‌های رقیب انتخاب نمودیم. مقایسه روش پیشنهادی با روش‌های رقیب با استفاده از دو معیار دقت و صحت که در بالا بیان شد در جدول ۲ آورده شده است. از چپ به راست ستون‌های اول،

^۱ <https://dynamicfieldtheory.org/cosivina/>

^۲ Miss

^۳ False positive

^۴ Mismatches

^۵ Matches

^۶ Loligo Systems ApS, Viborg, Denmark

می دهد. الگوریتم پیشنهادی به درستی بر هر دو این چالش ها غلبه نموده است.

پنجم: روش پیشنهادی سازگاری بیشتری با نحوه پردازش مغز دارد که این موضوع منجر به بهبود کارایی آن نیز شده است. میدان ها در این الگوریتم نقشی مشابه حافظه کاری انسان دارند و مانند انسان که اطلاعات مکان، جهت و سرعت لاروها را در طول ردگیری در حافظه کاری نگه داشته و از آن استفاده می کند، ما نیز این اطلاعات را در میدان ها حفظ کرده و از آن ها کمک می گیریم. در مورد مکانیزم توجه، باید گفت که در ویدیوهای لاروها و ماهی های قزل آلا و ویدیوهای مشابه پیچیدگی پس زمینه زیاد نیست و شرایط محیطی هم تشخیص لاروها را نسبت به پس زمینه نسبتا آسان می کند. به همین دلیل مکانیزم ساده تفاضل پس زمینه منجر به استخراج کامل لاروها می گردد. مکانیزمی که ما در این الگوریتم از آن استفاده نمودیم معادل این است که انسان تنها با استفاده از ویژگی شدت روشنی به تمییز لاروها از زمینه پردازد.

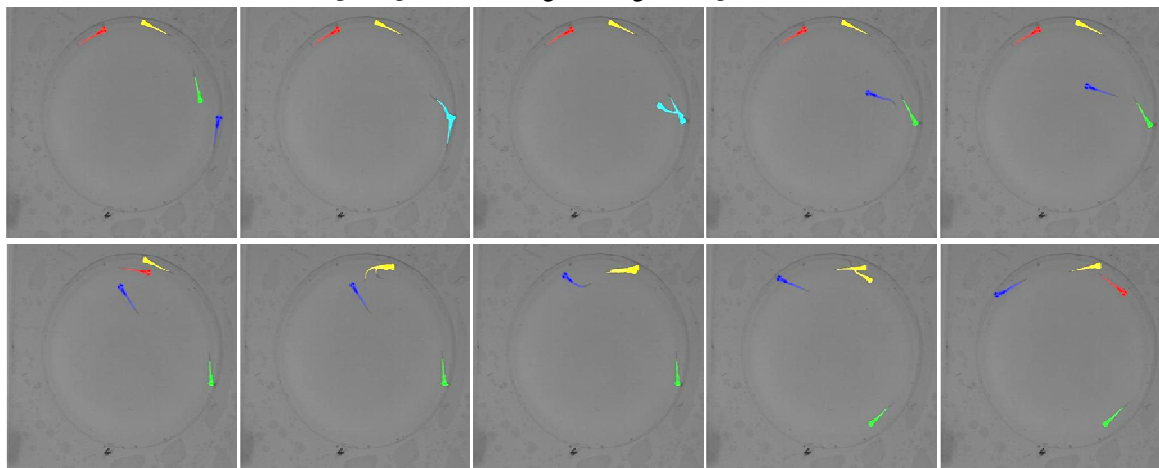
چهارم: روش پیشنهادی، در معیاری که مرتبط سازی داده ها در هر لحظه بر اساس آن انجام می شود نزدیکی مکانی و تغییر جهت اندک برای هر لارو را در نظر می گیرد. استفاده از اطلاعات راستای لاروها در کنار مکان و سرعت آن ها کمک کرده است تا لاروها را بهتر بتوان متمایز نمود و به ندرت در برچسب زنی آن ها دچار اشتباه شد. حال آنکه سایر روش ها تنها به استفاده از اطلاعات مکان یا شدت روشنیایی بسنده کرده اند. شکل ۲ دو نمونه از برونداد الگوریتم را هنگام انسداد نشان می دهد. ویدیوهای مجموعه داده رنگی نیستند. بلکه برونداد الگوریتم به شکل رنگی نمایش داده شده است. به این ترتیب که هر لارو تشخیص داده شده توسط روش پیشنهادی، با یک رنگ متمایز مشخص شده است. توالی قاب ها در این شکل از چپ به راست است. ردیف اول یک سناریو انسداد بین لاروهایی است که به رنگ های سبز و آبی برچسب خورده اند و ردیف دوم یک سناریو انسداد دیگر بین لاروهایی که به رنگ های زرد و قرمز برچسب زده شده اند را نشان

جدول ۲ ارزیابی صحت و دقت

شماره ویدیو	دقت (پیکسل)					صحت				
	[۳۱]	[۳۰] و [۳۱]	[۳۴]	[۳۰]	روش پیشنهادی	[۳۱]	[۳۰] و [۳۱]	[۳۴]	[۳۰]	روش پیشنهادی
۱	۱۱,۳۸۸	-	۱۱,۶۶۲	۶,۳۴۶	۴,۵۵۸	۱	-	۰,۹۹۰	۰,۹۸۸	۱
۲	۲۱,۴۳۴	۱۲,۱۷۶	۱۸,۳۹۵	۱۵,۰۲۴	۱۳,۳۰۱	۰,۵۴۱	۰,۹۲۲	۰,۹۲۱	۰,۸۵۳	۰,۹۷۹
۳	۲۰,۶۴۸	۲۰,۸۷۹	۱۸,۸۵۴	۸,۱۱۳	۶,۷۱۱	۱	۰,۹۷۵	۰,۹۸۵	۰,۹۹۸	۱
۴	۱۶,۷۲۸	۱۷,۶۶۵	۲۳,۲۰۸	۱۰,۶۶۹	۸,۳۷۴	۰,۹۸۷	۰,۸۴۰	۰,۹۸۷	۰,۹۹۳	۰,۹۹۶
۵	۲۱,۵۴۵	۲۱,۷۲۷	۲۱,۸۹۰	۱۵,۵۲۵	۵,۵۲۵	-۰,۰۱	۰,۸۹۵	۰,۱۳۹	۰,۹۹۴	۱
۶	۱۳,۱۵۱	۱۲,۵۶۲	۱۵,۰۲۰	۱۲,۷۸۶	۵,۳۷۴	-۰,۲۷	-۰,۰۸۶	۰,۹۱۴	۰,۹۳۶	۱
۷	۲۵,۲۳۰	۴۳,۷۴۶	۸۰,۶۳۰	۳۰,۰۸۲	۱۰,۷۱۶	۰,۰۰۵	۰,۷۲۵	۰,۵۰۴	۰,۹۵۴	۰,۹۸۹
۸	۵۳,۰۹۶	۵۹,۶۶۶	۹۸,۹۳۶	۳۶,۹۰۱	۸,۰۶۵	-۰,۳۹۰	۰,۷۳۹	۰,۲۰۹	۰,۹۵۶	۱
۹	۲۹,۹۲۱	۴۸,۷۳۹	۱۴۲,۸۳۴	۱۵,۹۶۰	۸,۴۴۰	۰,۶۷۳	-۰,۱۰۷	۰,۳۲۷	۰,۹۸۹	۱
۱۰	۲۱۹,۳۲۹	۲۷,۵۳۲	۱۸۹,۹۷۵	۲۵,۱۲۱	۶,۶۱۲	-۰,۱۸۰	۰,۱۳۳	۰,۹۰۶	۰,۹۲۰	۱
میانگین	۴۳,۲۴۷	۲۹,۴۱۰	۶۲,۱۴۰	۱۷,۶۵۰	۷,۷۶۷	۰,۳۳۰	۰,۵۶۰	۰,۶۹	۰,۹۶	۰,۹۹۶

میدان ها که دائما اتفاق می افتد نیازمند به روز رسانی مداوم توابع فعالیت و تعامل است که از نظر اجرا نسبتا زمان بر هستند. به همین دلیل استفاده از این روش در کاربردهای بلا درنگ نیازمند بهبود

ششم: علی رغم آنچه که مطرح شد، روش پیشنهادی هنوز نیاز به بهبودهایی دارد تا بتواند در کاربردهایی که از نظر محیط بصری پیچیده تر هستند نیز مورد استفاده قرار بگیرد. همچنین، تکامل



شکل ۲ نمونه ای از برونداد الگوریتم پیشنهادی

- [2] S. Kamkar, F. Ghezloo, H. A. Moghaddam, A. Borji, and R. Lashgari, "Multiple-target tracking in human and machine vision," *PLOS Comput. Biol.*, vol. 16, no. 4, p. e1007698, Apr. 2020.
- [3] S. Kamkar and R. Safabakhsh, "Vehicle detection, counting and classification in various conditions," *IET Intell. Transp. Syst.*, vol. 10, no. 6, pp. 406–413, Aug. 2016.
- [4] E. Soroush, A. Mirzaei, S. Kamkar, and K. Nasir, "Near Real-Time Vehicle Detection and Tracking in Highways." 15th International Conference on Traffic and Transportation Engineering, Tehran, 2016.
- [5] T. S. Horowitz, R. S. Birnkrant, D. E. Fencsik, L. Tran, and J. M. Wolfe, "How do we track invisible objects?," *Psychon. Bull. Rev.*, vol. 13, no. 3, pp. 516–523, 2006.
- [6] M. J. Lochner and L. M. Trick, "Multiple-object tracking while driving: the multiple-vehicle tracking task," *Attention, Perception, Psychophys.*, vol. 76, no. 8, pp. 2326–2345, Jun. 2014.
- [7] G. J. Zelinsky and M. B. Neider, "An eye movement analysis of multiple object tracking in a realistic environment," *Vis. cogn.*, vol. 16, no. 5, pp. 553–566, Jul. 2008.
- [8] C. C. Wu and J. M. Wolfe, "Comparing eye movements during position tracking and identity tracking: No evidence for separate systems," *Attention, Perception, Psychophys.*, vol. 80, no. 2, pp. 453–460, Feb. 2018.
- [9] B. Du, Y. Sun, C. Wu, L. Zhang, and L. Zhang, "Real-time tracking based on weighted compressive tracking and a cognitive memory model," *Signal Processing*, vol. 139, pp. 173–181, Oct. 2017.
- [10] Y. Qi, Y. Wang, and T. Xue, "Brain memory inspired template updating modeling for robust moving object tracking using particle filter," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2012, vol. 7366 LNAI, pp. 112–119.
- [11] H. Farazi and S. Behnke, "Online visual robot tracking and identification using deep LSTM networks," in *IEEE International Conference on Intelligent Robots and Systems*, Dec. 2017, vol. 2017–September, pp. 6118–6125.
- [12] X. Li, W. Hu, C. Shen, Z. Zhang, A. Dick, and A. Van Den Hengel, "A survey of appearance models in visual object tracking," *ACM Transactions on Intelligent Systems and Technology*, vol. 4, no. 4. 2013.
- [13] V. Mahadevan and N. Vasconcelos, "Biologically inspired object tracking using center-surround saliency mechanisms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 3, pp. 541–554, 2013.

زمان اجرا است. استفاده از برنامه‌نویسی موازی و یا محاسبات عصبی^۱ می‌تواند به حل این مسئله کمک کند.

جدول ۳ نرخ تطبیق‌های اشتباه

شماره ویدیو	[۳۱]	[۳۰] و [۳۱]	[۳۴]	[۳۰]	روش پیشنهادی
۱	۶	–	۲	۰	۰
۲	۲	۳	۴	۳	۰
۳	۰	۰	۰	۰	۰
۴	۴	۰	۰	۰	۰
۵	۳	۲	۶	۰	۰
۶	۱	۱	۱	۲	۰
۷	۲	۱	۳	۱	۰
۸	۱۱	۴۱	۲۴	۱۴	۰
۹	۴	۱۷	۳	۰	۰
۱۰	۴	۸	۰	۸	۰
میانگین	۳,۷	۸,۱	۴,۳	۲,۸	۰

۶ نتیجه‌گیری

در این مقاله، روشی برای ردگیری همزمان اشیاء ارائه شد که قادر است به خوبی با یکی از مهمترین چالش‌های این موضوع یعنی مسأله انسداد و مرتبط‌سازی داده‌ها مقابله نماید. از آنجا که انسان‌ها با این چالش همه روزه مواجه هستند و در زندگی روزمره بارها بر آن غلبه می‌کنند، ما از روشی ملهم از عملکرد مغز به نام میدان‌های عصبی پویا به عنوان پایه الگوریتم استفاده نمودیم. این میدان‌ها سازگاری بیشتری با پردازش مغز انسان دارند و در مدل کردن انواع حالات شناختی که پایه و اساس بسیاری از پردازش‌ها در مغز هستند مانند توجه و حافظه موفق بوده‌اند. همچنین، استفاده از روش‌های بینایی ماشین به ما کمک کرد تا ویژگی‌های مناسبی را حین ردگیری و مرتبط‌سازی داده‌ها در نظر بگیریم. ارزیابی‌ها نشان داد که الگوریتم پیشنهادی نسبت به سایر روش‌های مرز دانش عملکرد بهتری دارد. به ویژه در چالش‌های انسداد و مرتبط‌سازی داده‌ها به خوبی عمل می‌کند.

۷ سپاسگزاری

از آقای دکتر Wolfram Erlhagen، دانشیار دانشکده ریاضیات دانشگاه Minho پرتغال که ما را در رابطه با روش پیشنهادی یاری دادند تشکر و قدردانی می‌شود.

مراجع

- [1] W. Luo, J. Xing, A. Milan, X. Zhang, W. Liu, and T.-K. Kim, "Multiple object tracking: A literature review," *Artif. Intell.*, vol. 293, p. 103448, Dec. 2020.

¹ Neuromorphic

- [27] J. P. Spencer, K. Barich, J. Goldberg, and S. Perone, "Behavioral dynamics and neural grounding of a dynamic field theory of multi-object tracking," *J. Integr. Neurosci.*, vol. 11, no. 3, pp. 339–362, Sep. 2012.
- [28] J. Spencer and S. Perone, "A dynamic neural field model of multi-object tracking," *J. Vis.*, vol. 8, no. 6, pp. 508–508, Mar. 2010.
- [29] S. K. U. Zibner, C. Faubel, I. Iossifidis, G. Schöner, and J. P. Spencer, "Scenes and tracking with dynamic neural fields: How to update a robotic scene representation," in *2010 IEEE 9th International Conference on Development and Learning, ICDL-2010 - Conference Program*, 2010, pp. 244–250.
- [30] X. Wang, E. Cheng, I. S. Burnett, Y. Huang, and D. Wlodkowic, "Automatic multiple zebrafish larvae tracking in unconstrained microscopic video conditions," *Sci. Rep.*, vol. 7, no. 1, pp. 1–8, Dec. 2017.
- [31] A. Pérez-Escudero, J. Vicente-Page, R. C. Hinz, S. Arganda, and G. G. De Polavieja, "IdTracker: Tracking individuals in a group by automatic identification of unmarked animals," *Nat. Methods*, vol. 11, no. 7, pp. 743–748, Jun. 2014.
- [32] S. Perone and V. R. Simmering, "Applications of Dynamic Systems Theory to Cognition and Development: New Frontiers," in *Advances in Child Development and Behavior*, vol. 52, Academic Press Inc., 2017, pp. 43–80.
- [33] G. Schöner, J. Spencer, and D. Research Group, *Dynamic Thinking*. Oxford University Press, 2016.
- [34] L. Závorka *et al.*, "Co-existence with non-native brook trout breaks down the integration of phenotypic traits in brown trout parr," *Funct. Ecol.*, vol. 31, no. 8, pp. 1582–1591, Aug. 2017.
- [35] B. P. Keane and Z. W. Pylyshyn, "Is motion extrapolation employed in multiple object tracking? Tracking as a low-level, non-predictive function," *Cogn. Psychol.*, vol. 52, no. 4, pp. 346–368, Jun. 2006.
- [14] V. Mahadevan and N. Vasconcelos, "Saliency-based discriminant tracking," Mar. 2010, pp. 1007–1013.
- [15] L. Ma, J. Cheng, J. Liu, J. Wang, and H. Lu, "Visual attention model based object tracking," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2010, vol. 6298 LNCS, no. PART 2, pp. 483–493.
- [16] O. Yilmaz, "Oscillatory synchronization model of attention to moving objects," *Neural Networks*, vol. 29–30, pp. 20–36, May 2012.
- [17] Q. Chu, W. Ouyang, H. Li, X. Wang, B. Liu, and N. Yu, "Online Multi-Object Tracking Using CNN-based Single Object Tracker with Spatial-Temporal Attention Mechanism," 2017.
- [18] J. Zhu, H. Yang, N. Liu, M. Kim, W. Zhang, and M.-H. Yang, "Online Multi-Object Tracking with Dual Matching Attention Networks," 2018.
- [19] L. Oksama and J. Hyönä, "Position tracking and identity tracking are separate systems: Evidence from eye movements," *Cognition*, vol. 146, pp. 393–409, Jan. 2016.
- [20] S. C. Wong, V. Stamatescu, A. Gatt, D. Kearney, I. Lee, and M. D. McDonnell, "Track Everything: Limiting Prior Knowledge in Online Multi-Object Recognition," *IEEE Trans. Image Process.*, vol. 26, no. 10, pp. 4669–4683, Oct. 2017.
- [21] B. Cai, X. Xu, X. Xing, K. Jia, J. Miao, and D. Tao, "BIT: Biologically Inspired Tracker," *IEEE Trans. Image Process.*, vol. 25, no. 3, pp. 1327–1339, Mar. 2016.
- [22] S. Zhang, X. Lan, H. Yao, H. Zhou, D. Tao, and X. Li, "A Biologically Inspired Appearance Model for Robust Visual Tracking," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 28, no. 10, pp. 2357–2370, Oct. 2017.
- [23] K. Zhang, Q. Liu, Y. Wu, and M. H. Yang, "Robust visual tracking via convolutional networks without training," *IEEE Trans. Image Process.*, vol. 25, no. 4, pp. 1779–1792, Apr. 2016.
- [24] K. Fang, Y. Xiang, X. Li, and S. Savarese, "Recurrent Autoregressive Networks for Online Multi-object Tracking," in *Proceedings - 2018 IEEE Winter Conference on Applications of Computer Vision, WACV 2018*, May 2018, vol. 2018-January, pp. 466–475.
- [25] S. Sun, N. Akhtar, H. Song, A. Mian, and M. Shah, "Deep Affinity Network for Multiple Object Tracking," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 1, pp. 104–119, Jan. 2021.
- [26] A. Rangesh, P. Maheshwari, M. Gebre, S. Mhatre, V. Ramezani, and M. M. Trivedi, "TrackMPNN: A Message Passing Graph Neural Architecture for Multi-Object Tracking," Jan. 2021, Available: <http://arxiv.org/abs/2101.04206>.



شیوا کامکار تحصیلات خود را در مقطع کارشناسی مهندسی کامپیوتر گرایش نرم افزار در دانشگاه الزهرا (س) در سال ۱۳۹۱ به اتمام رساند. وی فارغ التحصیل دوره کارشناسی ارشد مهندسی کامپیوتر گرایش هوش مصنوعی از دانشگاه صنعتی امیرکبیر در سال ۱۳۹۳ است. او همچنین دوره دکتری

مهندسی کامپیوتر گرایش هوش مصنوعی را در دانشگاه صنعتی خواجه نصیرالدین طوسی در سال ۱۳۹۹ به پایان برد. زمینه های پژوهشی مورد علاقه ایشان، بینایی ماشین، پردازش تصویر، علوم شناختی و الگوریتم های ملهم از مغز است.



حمید ابریشمی مقدم دارای مدرک کارشناسی مهندسی برق از دانشگاه صنعتی امیرکبیر و کارشناسی ارشد مهندسی پزشکی بیوالکتریک از دانشگاه صنعتی شریف و دکترای مهندسی پزشکی گرایش پردازش تصاویر پزشکی از دانشگاه صنعتی کمپین فرانسه است. ایشان در سال ۱۳۷۷ همکاری خود را به عنوان استادیار گروه الکترونیک با دانشکده مهندسی برق

دانشگاه صنعتی خواجه نصیرالدین طوسی آغاز کرد و سپس با تاسیس گروه مهندسی پزشکی در سال ۱۳۸۱ به این گروه ملحق گردید. در سال ۱۳۷۹ با تاسیس آزمایشگاه پژوهشی بینایی ماشینی و پردازش تصاویر پزشکی به سمت سرپرست آزمایشگاه منصوب گردید و تاکنون طرحهای پژوهشی متعددی با حمایت صنایع و مراکز تحقیقاتی داخل و خارج از کشور در این آزمایشگاه و تحت نظر ایشان به انجام رسیده است. وی عضو هیات موسس انجمن ماشین بینایی و پردازش تصویر بوده و از بدو تاسیس بمدت شش سال به عنوان نایب رییس هیات مدیره انجمن فعالیت نموده است. او همچنین از بدو تاسیس انجمن مهندسی پزشکی ایران بمدت شش سال به عنوان عضو هیات مدیره انجمن فعالیت داشته است. ایشان عضو هیات موسس انجمن بازرسی غیر مخرب ایران نیز بوده است. دکتر ابریشمی مقدم دبیر دومین کنفرانس ماشین بینایی و پردازش تصویر در سال ۱۳۸۱ در دانشگاه صنعتی خواجه نصیرالدین طوسی بوده است. او هم اکنون استاد گروه مهندسی پزشکی دانشگاه صنعتی خواجه نصیرالدین طوسی بوده و عضو هیات تحریریه مجلات علمی پژوهشی مهندسی پزشکی زیستی و ماشین بینایی و پردازش تصویر می باشد. زمینه های پژوهشی مورد علاقه ایشان بینایی ماشینی، پردازش تصاویر و بازشناسی الگو است و تاکنون بیش از یکصد مقاله در مجلات معتبر داخلی و بین المللی از ایشان به چاپ رسیده است.



رضا لشگری تحصیلات خود را در حوزه نوروفیزیولوژی از دانشگاه علوم پزشکی شهید بهشتی و در حوزه علوم اعصاب بینایی از دانشگاه های کلمبیا و ایالتی نیویورک اخذ نموده است. ایشان در حال حاضر استاد دانشگاه شهید بهشتی می باشد و فعالیت پژوهشی ایشان معطوف به درک مکانیسم مدارهای عصبی سیستم بینایی، فعالیت های شناختی در مغز و علوم فناوری های هوشمند پزشکی می باشد.