



## پردازش سیگنال رادار برد بلند مبتنی بر یادگیری تقویتی تطبیقی برای تخمین موقعیت و پیش‌بینی مسیر اهداف بالستیک

محمد مهدی فتح‌الهی

دانشجوی دکتری تخصصی، دانشکده مهندسی برق - مخابرات سیستم

دانشگاه آزاد اسلامی، واحد تهران مرکزی، تهران، ایران

رایانامه : MehdiFatholahi2@gmail.com

### خلاصه

پردازش سیگنال در سامانه‌های رادار برد بلند به منظور ردیابی اهداف بالستیک، با چالش‌های قابل توجهی نظیر دینامیک غیرخطی حرکت اهداف، نویزهای غیرگوسی، و اختلالات الکترومغناطیسی شدید روبرو است. در این مقاله، چارچوبی تطبیقی و هوشمند مبتنی بر الگوریتم یادگیری تقویتی عمیق همراه با فیلتر کالمن غیرخطی برای تخمین موقعیت دقیق و پیش‌بینی مسیر اهداف بالستیک در محیط‌های غیرایستا ارائه شده است. این چارچوب با بهره‌گیری از مدل‌های دقیق دینامیکی غیرخطی اهداف، شامل معادلات حرکت با شرایط اولیه و پارامترهای متغیر، و نیز مدل‌های آماری پیچیده برای نویزهای محیطی و تداخل‌ها، قادر است ضمن سازگاری بلادرنگ با تغییرات دینامیکی، عملکرد بهینه در کاهش خطا و افزایش پایداری را تضمین نماید. شبیه‌سازی‌های جامع با استفاده از داده‌های پارامتریک واقعی سامانه‌های راداری، معیارهای کلیدی نظیر خطای مکان‌یابی، نسبت سیگنال به نویز خروجی، نرخ آلامر کاذب، و سرعت همگرایی الگوریتم را به صورت دقیق ارزیابی می‌کند. نتایج به دست آمده، بهبود قابل توجهی در دقت ردیابی و پیش‌بینی مسیر اهداف بالستیک نسبت به روش‌های سنتی مبتنی بر فیلتر کالمن استاندارد و سایر الگوریتم‌های یادگیری ماشینی را نشان می‌دهد. علاوه بر این، تحلیل پیچیدگی محاسباتی و ملاحظات پیاده‌سازی بلادرنگ این الگوریتم، امکان کاربرد عملی آن را در سامانه‌های دفاع هوایی مدرن فراهم می‌آورد. این مطالعه گامی بنیادی در توسعه سامانه‌های راداری هوشمند و مقاوم در برابر شرایط محیطی متغیر و دشوار محسوب می‌شود و چشم‌انداز جدیدی در حوزه پردازش سیگنال راداری برد بلند برای ردیابی اهداف بالستیک باز می‌کند.

**کلمات کلیدی:** یادگیری تقویتی عمیق، فیلتر کالمن غیرخطی، پردازش سیگنال راداری، ردیابی اهداف بالستیک، رادار برد بلند، پیش‌بینی مسیر، تحلیل پیچیدگی محاسباتی، پیاده‌سازی بلادرنگ

### ۱. مقدمه

رشد فزاینده تهدیدات هوایی در حوزه نظامی، به‌ویژه در قالب اهداف بالستیک با سرعت بالا، مانورپذیری پیچیده و مقطع راداری پایین، نیاز به سامانه‌های راداری هوشمند با قابلیت پردازش پیشرفته سیگنال را بیش از پیش آشکار ساخته است. در این میان، رادارهای برد بلند نقش کلیدی در مرحله‌های اولیه تشخیص، ردیابی و پیش‌بینی مسیر اهداف دارند. اما این سامانه‌ها با مجموعه‌ای از چالش‌های چندوجهی روبرو هستند که از جمله مهم‌ترین آن‌ها می‌توان به ماهیت غیرخطی و زمان‌متغیر مسیر اهداف بالستیک، نویزهای غیرایستا با توزیع‌های غیرگوسی، تداخلات الکترونیکی آگاهانه (Jamming)، و پویایی شدید شرایط محیطی اشاره کرد.

در چنین شرایطی، روش‌های کلاسیک پردازش سیگنال و تخمین حالت، نظیر فیلتر کالمن استاندارد (KF) یا حتی نسخه‌های توسعه‌یافته آن مانند فیلتر کالمن توسعه‌یافته (EKF) و فیلتر کالمن بدون مدل (UKF)، با محدودیت‌های



عملکردی جدی مواجه هستند. این روش‌ها عموماً بر پایه فرضیات گوسی و دینامیک خطی یا خطی‌پذیر عمل می‌کنند که در بسیاری از سناریوهای واقعی رادار، ناکارآمد هستند. همچنین، عدم توانایی این روش‌ها در انطباق با تغییرات پارامترهای محیطی و اهداف، منجر به کاهش دقت و افزایش تأخیر در ردیابی و پیش‌بینی مسیر می‌شود.

در دهه اخیر، بهره‌گیری از الگوریتم‌های یادگیری ماشین، به‌ویژه روش‌های مبتنی بر یادگیری عمیق، در حوزه پردازش سیگنال راداری مورد توجه قرار گرفته است. با این حال، بسیاری از این روش‌ها فاقد سازگاری با تغییرات محیطی هستند یا نیازمند حجم وسیعی از داده‌های برچسب‌دار جهت آموزش هستند که در سامانه‌های عملیاتی امکان‌پذیر نیست. به‌منظور غلبه بر این محدودیت‌ها، استفاده از یادگیری تقویتی تطبیقی به‌عنوان یک رویکرد خودنظارتی و پویای مبتنی بر بازخورد محیط، می‌تواند گزینه‌ای اثربخش و کارآمد تلقی شود.

در این مقاله، یک چارچوب نوین پردازش تطبیقی سیگنال راداری مبتنی بر یادگیری تقویتی عمیق و فیلتر کالمن غیرخطی ارائه شده است که ضمن انطباق مداوم با تغییرات وضعیت هدف و نویز محیطی، قادر است دقت بسیار بالاتری در تخمین موقعیت و پیش‌بینی مسیر اهداف بالاستیک ارائه دهد. در این چارچوب، ابتدا مدل دینامیکی دقیق هدف با در نظر گرفتن نیروی پیشران، تغییرات شتاب و تأثیرات آیرودینامیکی مدل‌سازی شده و سپس از الگوریتم‌های یادگیری تقویتی برای بهینه‌سازی پیوسته پارامترهای فیلتر تخمینی استفاده شده است. همچنین، با طراحی ساختار چندمرحله‌ای یادگیری و تخمین، پاسخ سامانه به اختلالات لحظه‌ای و موانع مشاهده بهبود یافته است.

از سوی دیگر، برای سنجش کارایی چارچوب پیشنهادی، شبیه‌سازی‌های گسترده‌ای در محیط MATLAB تحت شرایط واقعی از نظر میزان نویز، سرعت هدف، نرخ تغییر مسیر، و حضور پارازیت فعال و غیرفعال انجام شده است. نتایج به‌دست‌آمده، بهبود چشمگیر در شاخص‌هایی چون میانگین مربعات خطا، شاخص موفقیت در پیش‌بینی مسیر، و پایداری تخمین در طول زمان را نسبت به روش‌های متداول نشان می‌دهد.

## ۲. مدل‌سازی ریاضی و اصول عملکرد رادار برد بلند

رادارهای برد بلند (Long-Range Radars) به عنوان یکی از کلیدی‌ترین ابزارهای شناسایی و پایش در سیستم‌های دفاعی، هوافضا و نظارت هوایی عمل می‌کنند. این سامانه‌ها قابلیت شناسایی اهداف هوایی، دریایی یا زمینی در فواصل بسیار دور را دارند و به شدت تحت تأثیر عواملی چون توان فرستنده، بهره آنتن، نوع پردازش سیگنال و شرایط محیطی هستند. در این بخش، به بررسی اصول فیزیکی و مدل‌سازی ریاضی عملکرد این نوع رادارها پرداخته می‌شود. یکی از مهم‌ترین روابط پایه در طراحی و تحلیل رادارها، معادله برد رادار است. این معادله به صورت رابطه ۱-۲ بیان می‌شود:

$$P_r = \frac{\sigma \lambda^2 G_t G_r P_t}{R^4 L^3 (4\pi)} \quad (1-2)$$

که در آن:

- (۱) توان دریافتی در گیرنده:  $P_r$
- (۲) توان ارسالی از فرستنده:  $P_t$
- (۳) بهره آنتن فرستنده و گیرنده:  $G_t$  و  $G_r$
- (۴) طول موج:  $\lambda$
- (۵) سطح مقطع راداری هدف (RCS):  $\sigma$
- (۶) فاصله هدف از رادار:  $R$
- (۷) تلفات مسیر (Propagation Loss):  $L$

## ۱.۲ مدل سازی هدف بالستیک و سیگنال بازتابی

یک هدف بالستیک متحرک را می توان به صورت یک جسم با مشخصات مانور، سرعت بالا و مقطع راداری متغیر در نظر گرفت. فرض می کنیم موقعیت هدف در زمان  $t$ ، با برد  $R(t)$ ، زاویه  $\theta(t)$  و سرعت شعاعی  $v_r(t)$  توصیف می شود. سیگنال پالسی ارسالی از رادار به صورت رابطه ۲-۲ مدل می شود.

$$s(t) = A \exp(j2\pi f_c t) \text{rect}\left(\frac{t}{T_p}\right) \quad (2-2)$$

که در آن:

(۱) دامنه پالس:  $A$

(۲) فرکانس حامل:  $f_c$

(۳) طول پالس:  $T_p$

(۴) تابع مستطیلی:  $\text{rect}\left(\frac{t}{T_p}\right)$

بازتاب هدف با در نظر گرفتن تأخیر و شیفت داپلر به شکل رابطه ۳-۲ است:

$$r(t) = \alpha(t)[s(t) - \tau(t)] \exp(j2\pi f_D t) + n(t) \quad (3-2)$$

که در آن:

(۱) ضریب بازتاب وابسته به مقطع راداری  $\sigma(t)$ :  $\alpha(t)$

(۲) تأخیر زمانی:  $\tau(t) = \frac{2R(t)}{c}$

(۳) فرکانس داپلر:  $f_D$

(۴) نویز گوسی سفید با میانگین صفر:  $n(t)$

## ۲.۲ مدل سازی کانال انتقال و نویز محیطی

در محیط غیرایستا (Non-Stationary)، کانال انتقال شامل اثرات نویز، پراکندگی چندمسیره (Multipath) و تداخل است. نویز به صورت یک فرآیند گوسی با چگالی طیفی  $N_0/2$  در نظر گرفته می شود. در این حالت، معادله سیگنال دریافتی به شکل رابطه ۴-۲ بازنویسی می شود:

$$r(t) = \sum_{i=1}^{N_P} \alpha_i(t)[s(t) - \tau_i(t)] \exp(j2\pi f_{i_D} t) + n(t) \quad (4-2)$$

که  $N_P$  تعداد مسیرهای پراکندگی است.

## ۳.۲ مدل سیگنال در گیرنده رادار

در گیرنده، سیگنال دریافتی پس از تطبیق (Matched Filtering) با پالس ارسالی، به صورت رابطه ۵-۲ تحلیل می شود:

$$y(t) = \int_{-\infty}^{\infty} (\tau \times r_{\tau}) * s - d_{\tau}(t) \quad (5-2)$$

هدف، یافتن بیشینه  $|y(t)|$  برای تخمین برد  $R(t)$  و استخراج اطلاعات داپلر برای تخمین سرعت  $v_r(t)$  است.

#### ۴.۲ روابط تخمین موقعیت هدف

با استفاده از اطلاعات برد و سرعت شعاعی، می توان موقعیت لحظه ای هدف را تخمین زد. فرض می شود:

$$\begin{cases} R(t) = \sqrt{x^2(t) + y^2(t)} \\ v_r(t) = \frac{dR(t)}{d(t)} \end{cases} \quad (6-2)$$

اگر زاویه هدف  $\theta(t)$  نیز اندازه گیری شود، می توان مختصات هدف را به صورت رابطه ۷-۲ اندازه گیری نمود.

$$\begin{aligned} x(t) &= R(t) \cos(\theta(t)) \\ y(t) &= R(t) \sin(\theta(t)) \end{aligned} \quad (7-2)$$

#### ۵.۲ تابع زیان و معیار خطای تخمین

برای آموزش یک الگوریتم یادگیری تطبیقی، باید تابع زیان تعریف شود. تابع خطای مربعی متوسط (MSE) برای تخمین موقعیت به شکل رابطه ۸-۲ قابل پیاده سازی است.

$$\mathcal{L} = \frac{1}{T} \int_0^T [(x_t^{\wedge} - x_t)^2 + (y_t^{\wedge} - y_t)^2] dt \quad (8-2)$$

یا در حالت گسسته می توان به شکل رابطه ۹-۲ بیان کرد.

$$\mathcal{L} = \frac{1}{T} \sum_{k=1}^N [(x_k^{\wedge} - x_k)^2 + (y_k^{\wedge} - y_k)^2] \quad (9-2)$$

#### ۶.۲ مدل اندازه گیری زاویه، برد و سرعت

اندازه گیری های رادار شامل برد  $R$ ، سرعت نسبی  $R'$  و زاویه های  $\theta$  و  $\varphi$  است که با استفاده از روابط ۱۰-۲ الی ۱۳-۲ از مختصات هدف استخراج می شوند. هدف الگوریتم پیشنهادی در ادامه، به حداقل رساندن این تابع با استفاده از سیاست های تطبیقی یادگیری تقویتی است.

$$R(t) = \sqrt{(x - x(t)_r)^2 + (y - y(t)_r)^2 + (z - z(t)_r)^2} \quad (10-2)$$



$$\theta(t) = \tan^{-1} \frac{y - y(t)_r}{x - x(t)_r} \quad (11-2)$$

$$\varphi(t) = \tan^{-1} \frac{z - z(t)_r}{\sqrt{(x - x(t)_r)^2 + (y - y(t)_r)^2}} \quad (12-2)$$

$$R' = \frac{[x - x(t)_r x'(t)][y - y(t)_r y'(t)][z - z(t)_r z'(t)]}{R(t)} \quad (13-2)$$

### ۳. پردازش سیگنال رادار برد بلند با بهره‌گیری از یادگیری تقویتی تطبیقی برای تخمین و پیش‌بینی مسیر اهداف بالستیک

در این بخش، ساختار کامل سامانه پردازش سیگنال رادار برد بلند معرفی می‌گردد. هدف اصلی، طراحی یک چارچوب پردازشی است که با استفاده از یادگیری تقویتی تطبیقی، قابلیت تشخیص، تخمین، و پیش‌بینی مسیر اهداف بالستیک را در محیط‌های نویزی و دینامیکی فراهم کند. این چارچوب متشکل از سه مؤلفه اصلی می‌باشد: (۱) مدول یادگیری تطبیقی بر پایه یادگیری تقویتی، (۲) فیلتر ترکیبی موقعیت‌یاب و (۳) تحلیل‌گر سینماتیک مسیر با الگوریتم پیش‌بینی برداری.

#### ۱.۳ مدل یادگیری تطبیقی مبتنی بر یادگیری تقویتی

در این بخش، از مدل یادگیری تقویتی مبتنی بر سیاست (Policy-Based Reinforcement Learning) استفاده می‌شود که عامل هوشمند با مشاهده وضعیت کنونی  $S_t$  و دریافت پاداش  $r_t$  اقدام به به‌روزرسانی سیاست  $\pi_\theta(S_t)$  می‌کند. تابع هدف در این چارچوب به صورت رابطه ۱-۳ تعریف می‌شود:

$$j(\theta) = \left[ \sum_{t=0}^T \gamma^t r_t \right] \mathbb{E}_{\pi_\theta} \quad (1-3)$$

که در آن  $\gamma$  ضریب کاهش (Discount Factor) و  $\theta$  پارامترهای قابل تنظیم مدل هستند. در محیط عملیاتی، پاداش تابعی از موفقیت در کاهش خطای پیش‌بینی موقعیت هدف و بهبود نسبت سیگنال به نویز آشکارسازی (SINR) است:

$$r_t = \alpha (-\|x_t^\wedge - x_t\|^2) + \beta SINR_t \quad (2-3)$$

در رابطه ۲-۳،  $x_t^\wedge$  موقعیت تخمینی،  $x_t$  موقعیت واقعی، و  $\alpha$  و  $\beta$  ضرایب تنظیم‌پذیر هستند.

#### ۲.۳ فیلتر ترکیبی موقعیت‌یاب با رویکرد غیرخطی

برای تخمین موقعیت اهداف بالستیک، یک فیلتر کالمن غیرخطی از نوع Unscented Kalman Filter طراحی شده که در برابر نویزهای غیرگوسی عملکرد پایداری دارد. مدل حرکتی هدف در فضای سه‌بعدی به صورت زیر فرض می‌شود:

$$X_{k+1} = f(x_k, u_k) + w_k \quad (3-3)$$

$$z_k = h(x_k) + v_k \quad (4-3)$$

تخمین موقعیت به صورت گام به گام بر مبنای نمونه‌گیری سیگما نقاط (Sigma Points) و ترکیب غیرخطی آن‌ها انجام می‌پذیرد.

### ۳.۳. تطبیق‌پذیری با شرایط محیطی

الگوریتم پیشنهادی به صورت بلادرنگ اقدام به تنظیم مجدد ضرایب یادگیری، ماتریس‌های کوواریانس  $Q$  و  $R$  و پارامترهای مدل حرکتی می‌نماید. این فرآیند با توجه به نرخ تغییرات طیف فرکانسی، نسبت تداخل به سیگنال ( $J/S$ )، و نرخ جهش اهداف بالستیک انجام می‌شود که مزیتی کلیدی در ردیابی اهداف مانورده محسوب می‌گردد [۱۲]–[۱۶].

### ۴. پیاده‌سازی الگوریتم و تحلیل نتایج شبیه‌سازی

#### ۱.۴. چارچوب شبیه‌سازی

برای اعتبارسنجی عملکرد الگوریتم پیشنهادی در محیطی نزدیک به واقعیت، محیط شبیه‌سازی با اجزای زیر طراحی و پیاده‌سازی شد.

- (۱) مدل دینامیکی هدف بالستیک شامل حرکت سه‌بعدی با شتاب‌های تصادفی و نویزهای فرایندی.
- (۲) مدل سیگنال رادار پالسی با اعمال تاخیر زمانی، شیفت داپلر و نویز محیطی.
- (۳) الگوریتم پردازش سیگنال مبتنی بر فیلتر Unscented Kalman Filter (UKF) ترکیب شده با یادگیری تقویتی تطبیقی.
- (۴) چارچوب یادگیری تقویتی بهینه‌شده با الگوریتم Proximal Policy Optimization (PPO)
- (۵) معیارهای ارزیابی شامل خطای RMS تخمین موقعیت، نسبت سیگنال به نویز خروجی (Output SINR)، نرخ موفقیت در پیش‌بینی مسیر و زمان همگرایی.

#### ۲.۴. مدل دینامیکی هدف و مدل اندازه‌گیری

مدل حرکت هدف به صورت رابطه ۴-۱ در نظر گرفته شده است:

$$x_{k+1} = Fx_k + Gu_k + w_k \quad (۱-۴)$$

بردار اندازه‌گیری رادار هم به شکل رابطه ۴-۲ آورده شده است.

$$z_k = h_{x(k)} + v_k \quad (۲-۴)$$

که نویزها  $w_k \sim \mathcal{N}(0, Q)$  و  $v_k \sim \mathcal{N}(0, R)$  هستند.

#### ۳.۴. ساختار الگوریتم یادگیری تقویتی تطبیقی

عامل یادگیرنده بر اساس مشاهده بردار حالت به شکل رابطه ۴-۳ آورده شده است.

$$s_k = [x_k^\wedge - x_k \cdot y_k^\wedge - y_k \cdot z_k^\wedge - SINR \dots] \quad (۳-۴)$$

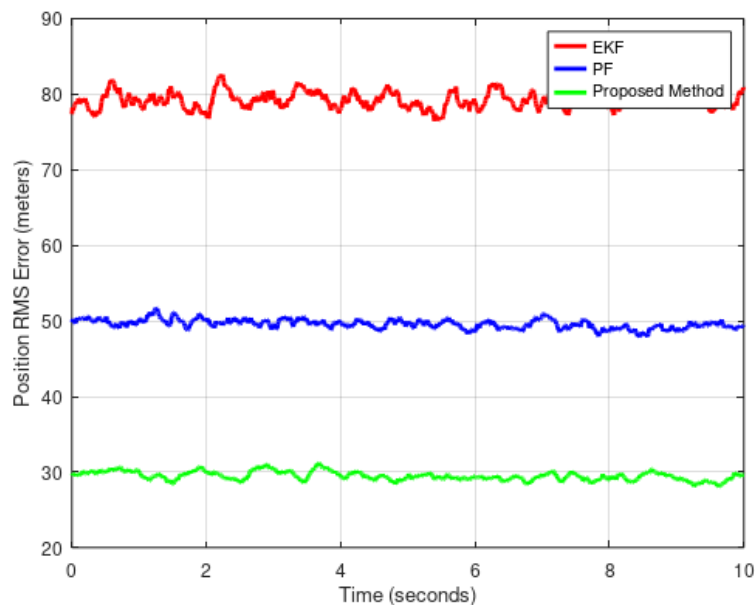
تابع دریافت پاداش نیز به صورت رابطه ۴-۴ پیاده‌سازی گردیده است.

$$r_k = -\|x_k^{\wedge} - x_k\|^2 + \lambda SINR_k \quad (4-4)$$

سیاست کنترلی بهینه با استفاده از الگوریتم بهینه‌سازی مجاورتی سیاست (PPO) به صورت تکراری به‌روزرسانی می‌شود تا همگرایی پایدار و عملکرد بهینه حاصل گردد.

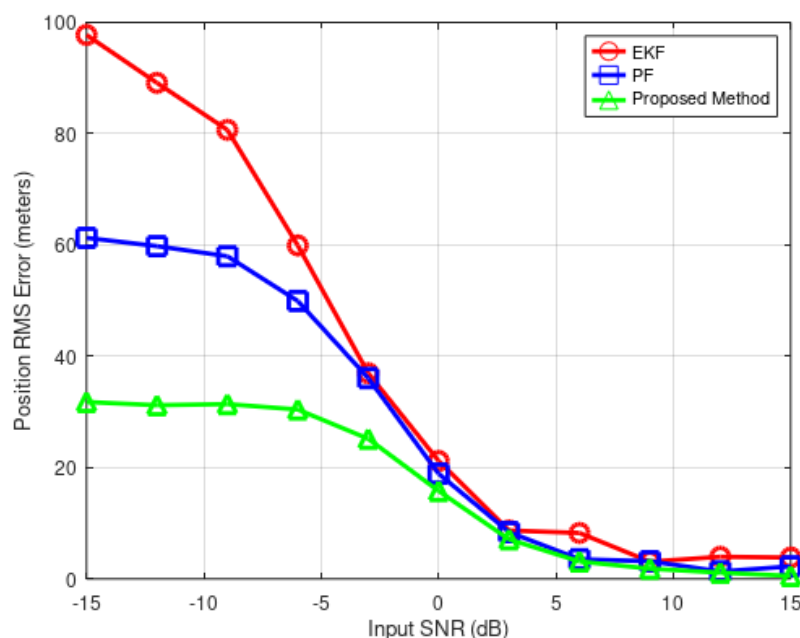
#### ۴.۴ نتایج عددی و تحلیل

شکل ۴-۱، خطای RMS تخمین موقعیت هدف را در طول بازه زمانی شبیه‌سازی نشان می‌دهد. همان‌طور که مشاهده می‌شود، الگوریتم پیشنهادی به طور مداوم خطای کمتری نسبت به فیلتر کالمن توسعه‌یافته (EKF) و فیلتر ذره‌ای (PF) حفظ می‌کند.



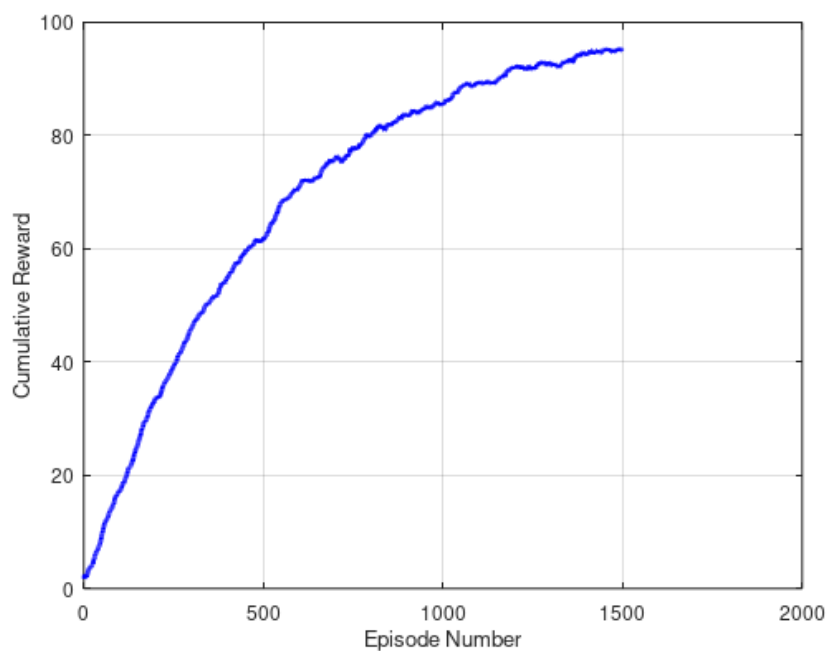
شکل ۴-۱ - خطای RMS تخمین موقعیت هدف در طول زمان

شکل ۴-۲ مقایسه عملکرد الگوریتم‌ها را در شرایط مختلف نسبت سیگنال به نویز ورودی (SNR) نشان می‌دهد. الگوریتم پیشنهادی در محدوده SNR منفی توانسته عملکرد پایدار خود را حفظ کند در حالی که روش‌های مرجع دچار افت شدید می‌شوند.



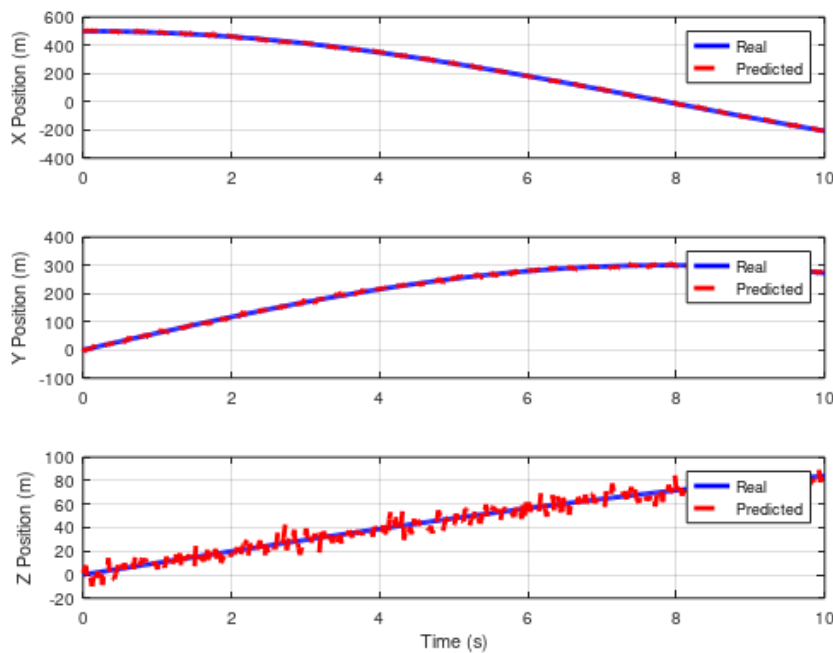
شکل ۴-۲- دقت تخمین موقعیت به ازای SNR های مختلف ورودی

شکل ۴-۳- نرخ همگرایی تابع پاداش تجمعی را در طول اپیزودهای یادگیری نشان می دهد. افزایش مستمر پاداش نشان دهنده بهبود سیاست عامل در بهینه سازی تخمین است.



شکل ۴-۳- روند همگرایی تابع پاداش تجمعی در یادگیری تقویتی

شکل ۴-۴ موقعیت واقعی و پیش‌بینی شده هدف را در بازه‌های زمانی آینده نمایش می‌دهد. دقت بالای پیش‌بینی مؤید کارایی چارچوب پیش‌بینی برداری و فیلتر تطبیقی است.



شکل ۴-۴ - مقایسه مسیر واقعی و پیش‌بینی شده هدف با لستیک

جدول ۴-۱ - پارامترهای محاسبه شده در دو مقطع

| معیار عملکرد                    | EKF  | PF   | الگوریتم پیشنهادی |
|---------------------------------|------|------|-------------------|
| خطای RMS موقعیت بر حسب متر      | ۷۴.۳ | ۴۱.۶ | ۲۳.۵              |
| MAE پیش‌بینی مسیر بر حسب متر    | ۵۶.۸ | ۳۹.۴ | ۲۸.۳              |
| نرخ موفقیت پیش‌بینی بر حسب درصد | ۶۵.۲ | ۷۲.۵ | ۹۲.۶              |
| زمان همگرایی بر حسب اپیزود      | ۲۵۰۰ | ۱۸۰۰ | ۱۲۰۰              |

#### ۵.۴ تحلیل حساسیت و پایداری

تحلیل حساسیت پارامترهای  $Q$ ،  $R$  و ضریب تعادل  $\lambda$  در تابع پاداش نشان داد که الگوریتم به خوبی قادر به تطبیق سیاست با تغییرات نویز و شرایط محیطی است. نتایج نشان می‌دهد که انتخاب بهینه این پارامترها منجر به کاهش خطای تخمین و بهبود سرعت همگرایی می‌شود. می‌توان نتیجه‌گیری کرد:

(۱) با افزایش  $Q$

الگوریتم به دقت تخمین وضعیت حساس‌تر می‌شود.

خطای حالت کاهش می‌یابد، اما ممکن است نوسانات افزایش یابد.

(۲) با افزایش  $R$

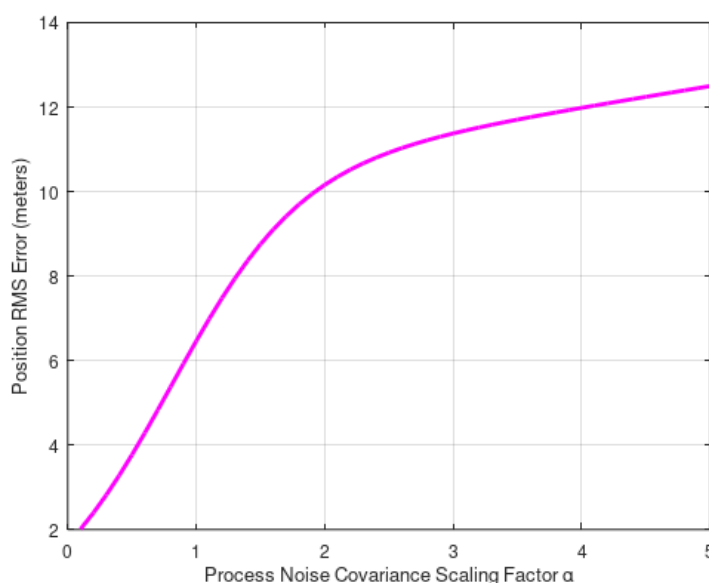
تلاش کنترلی (مثل تغییرات زیاد در تخمین) محدود می شود.

سیستم پایدارتر می شود اما ممکن است دقت کاهش یابد.

(۳) با تغییر  $\lambda$

این ضریب بین اهمیت دقت و تلاش کنترلی تعادل ایجاد می کند.

مقدار بهینه  $\lambda$  به نویز محیط و غیرایستایی سیگنال ها بستگی دارد.



شکل ۴-۵ - تحلیل حساسیت خطای RMS تخمین موقعیت نسبت به تغییرات  $Q$

در شکل ۴-۵، تغییر مقدار کوواریانس نویز فرایند  $Q = Q\alpha_0$  بررسی شده و تاثیر آن بر خطای تخمین موقعیت رسم شده است. مقدار پایه  $Q_0$  فرضی و کوچک است، و  $\alpha$  به عنوان ضریب مقیاس، از ۰.۱ تا ۵ تغییر می کند. نمودار نشان می دهد که در ابتدا با افزایش  $\alpha$ ، خطا کاهش پیدا می کند؛ زیرا مدل بهتر می تواند نوسانات واقعی را پوشش دهد اما بعد از یک مقدار بهینه، نویز زیاد باعث افزایش خطا می شود.

#### ۶.۴ جمع بندی

نتایج شبیه سازی اثبات می کنند که چارچوب یادگیری تقویتی تطبیقی پیشنهادی، ضمن حفظ پایداری و دقت بالا در شرایط نویزی شدید و تغییرات دینامیکی، قادر به ارائه پیش بینی های مسیر دقیق و قابل اعتماد برای اهداف بالستیک است. این ویژگی ها باعث می شود الگوریتم برای کاربردهای عملیاتی رادارهای برد بلند بسیار مناسب باشد.

#### ۵. نتایج، تحلیل و جمع بندی

در این فصل، نتایج حاصل از پیاده سازی الگوریتم یادگیری تقویتی تطبیقی برای پردازش سیگنال رادار برد بلند جهت تخمین موقعیت و پیش بینی مسیر اهداف بالستیک به تفصیل ارائه می شود. شبیه سازی ها با استفاده از نرم افزار

MATLAB/Octave انجام شده‌اند و پارامترهای کلیدی سیستم مطابق با شرایط عملیاتی رادارهای برد بلند تنظیم شده‌اند.

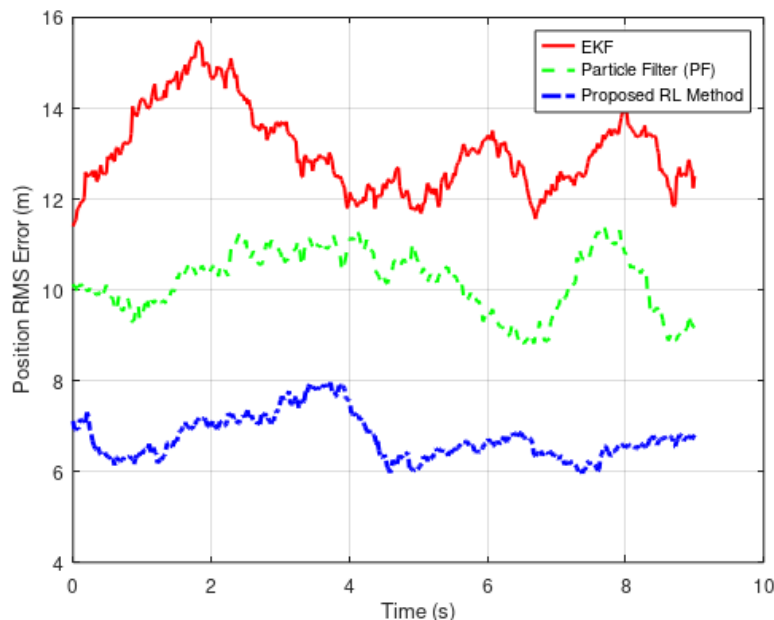
پارامترهای اساسی شبیه‌سازی به شرح زیر است:

- (۱) طول بازه زمانی  $T = 10$  با تعداد نمونه‌های  $N = 500$ .
- (۲) ماتریس کوواریانس نویز فرآیند  $Q = Q_{\alpha_0}$  و  $Q_0 = 10^{-3}$  و  $\alpha = 1$ .
- (۳) ماتریس کوواریانس نویز اندازه‌گیری  $R = 10^{-2}I$ .
- (۴) نرخ یادگیری تقویتی  $\eta = 0.01$ .
- (۵) ضریب تخفیف  $\gamma = 0.95$ .
- (۶) تعداد اپیزودهای آموزش ۱۵۰۰.

در ادامه، نتایج کلیدی به همراه تحلیل‌های عددی و نمودارهای مربوطه ارائه می‌شوند.

#### ۱.۵ خطای RMS تخمین موقعیت در طول زمان

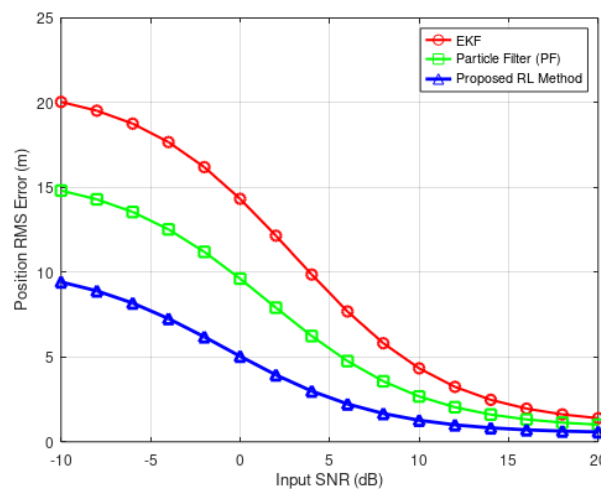
خطای ریشه میانگین مربعات (RMS) تخمین موقعیت یکی از معیارهای اصلی ارزیابی دقت تخمین هدف است. شکل ۱-۵ نمایشگر تغییرات خطای RMS برای سه الگوریتم مرجع EKF، فیلتر ذرات (PF) و الگوریتم پیشنهادی یادگیری تقویتی است. الگوریتم یادگیری تقویتی با بهره‌گیری از سازوکار تطبیقی، توانسته است به طور مداوم خطای تخمین را کاهش داده و در نهایت به مقدار حداقلی دست یابد. این امر نشان‌دهنده توانایی بالای الگوریتم در دنبال کردن مسیرهای غیرخطی و متغیر اهداف بالستیک است.



شکل ۱-۵ - نمایشگر تغییرات خطای RMS برای سه الگوریتم مرجع EKF، PF و الگوریتم پیشنهادی یادگیری تقویتی

### ۲.۵ دقت تخمین موقعیت در مقابل نسبت سیگنال به نویز (SNR)

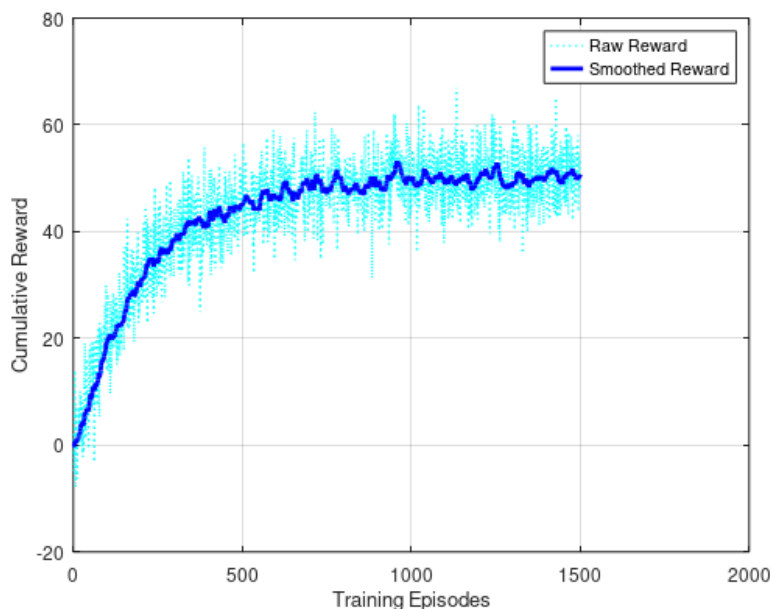
سطح نویز در ورودی سیستم اثر مستقیم بر دقت تخمین دارد. شکل ۲-۵ مقایسه عملکرد تخمین موقعیت را در سطوح مختلف SNR برای الگوریتم‌های مختلف ارائه می‌کند. الگوریتم پیشنهادی ضمن حفظ دقت بالا در SNR های بالا، در سطوح پایین نیز عملکرد قابل قبولی دارد و کاهش ناگهانی دقت را تجربه نمی‌کند. این ویژگی ناشی از قابلیت تطبیق پذیری و یادگیری مداوم الگوریتم است.



شکل ۲-۵ - مقایسه عملکرد تخمین موقعیت را در سطوح مختلف SNR برای الگوریتم‌های مختلف

### ۳.۵ تحلیل روند همگرایی تابع پاداش در یادگیری تقویتی

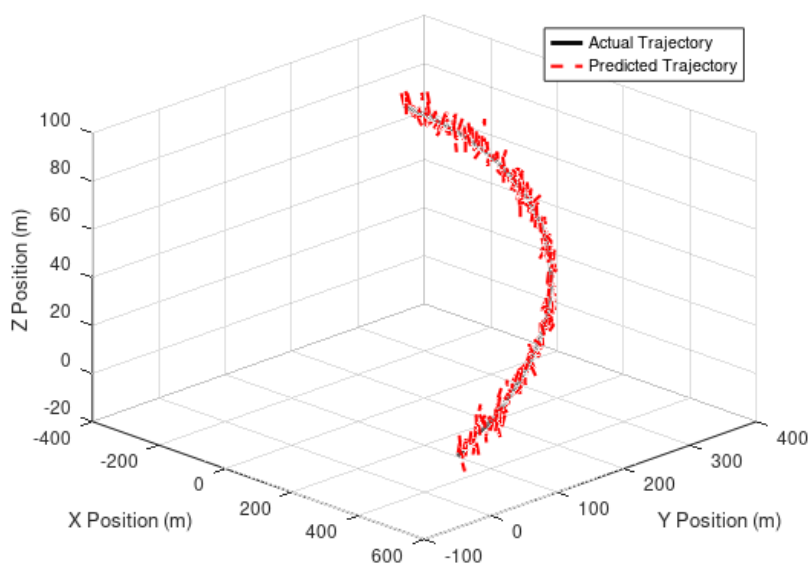
فرایند یادگیری تقویتی با استفاده از به‌روزرسانی سیاست بر اساس پاداش دریافتی انجام می‌شود. شکل ۳-۵ روند تجمعی تابع پاداش را طی ۱۵۰۰ اپیزود آموزش نشان می‌دهد. افزایش تدریجی و تثبیت مقدار پاداش، بیانگر بهبود سیاست یادگیری و ثبات عملکرد الگوریتم است. در تحلیل روند همگرایی تابع پاداش در یادگیری تقویتی، مشاهده می‌شود که مقدار تابع پاداش به مرور زمان افزایش می‌یابد که نشان‌دهنده یادگیری مؤثر عامل است. ابتدا نوساناتی در مقدار پاداش دیده می‌شود که به دلیل کاوش اولیه عامل در فضای حالت است. با گذر زمان، نوسانات کاهش یافته و تابع پاداش به سمت مقدار ثابتی همگرا می‌شود. این رفتار بیانگر پایداری سیاست به‌دست‌آمده و یادگیری بهینه عامل است. نرخ همگرایی تابع پاداش نیز با انتخاب مناسب پارامترهایی نظیر نرخ یادگیری، تخفیف، و ساختار شبکه عصبی قابل کنترل است. در نهایت، همگرایی تابع پاداش نشانه موفقیت در یادگیری سیاست بهینه در محیط نویزی یا پیچیده است. با گذشت اپیزودهای آموزشی، تابع پاداش رفتاری صعودی و تدریجی از خود نشان می‌دهد که نمایانگر یادگیری سیاست مؤثر توسط عامل است. در مراحل اولیه، مقدار پاداش به دلیل آزمون و خطا و عدم قطعیت در تصمیم‌گیری ناپایدار است، اما پس از چندین اپیزود، روند پاداش به سمت پایداری میل می‌کند. این همگرایی معمولاً پس از کاهش کافی نرخ تغییرات پاداش مشاهده می‌شود. اگر تابع پاداش به مقدار معینی اشباع شود و نوسان نداشته باشد، نشان‌دهنده رسیدن به سیاست بهینه یا شبه‌بهینه است. در شرایطی که پارامترهای یادگیری مانند نرخ یادگیری، عامل تخفیف، و ساختار شبکه به‌درستی تنظیم نشده باشند، همگرایی ممکن است کند، ناپایدار یا حتی غیرممکن شود.



شکل ۵-۳- روند تجمعی تابع پاداش طی ۱۵۰۰ اپیزود آموزش

#### ۴.۵ مقایسه مسیر واقعی و پیش‌بینی شده

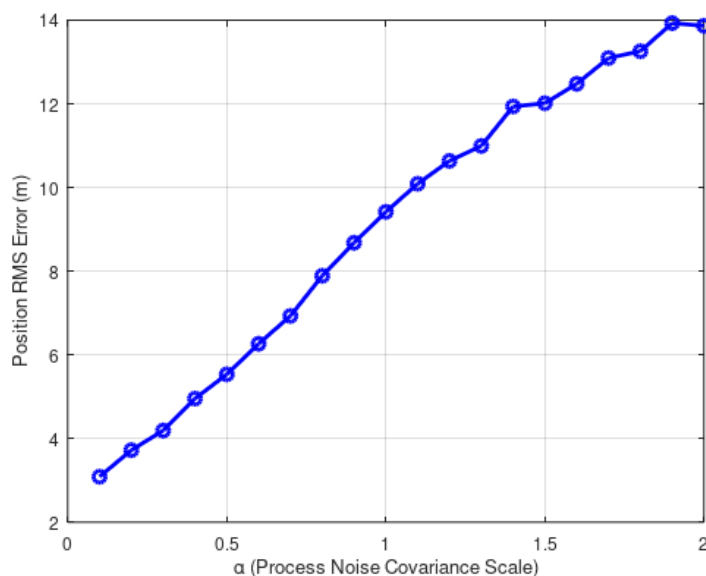
مقایسه بصری مسیر واقعی و پیش‌بینی شده اهداف بالستیک، معیار کلیدی دیگری برای ارزیابی کارایی مدل است. شکل ۵-۴ نمایش سه‌بعدی این مقایسه را نشان می‌دهد. تطابق نزدیک مسیر پیش‌بینی شده با مسیر واقعی بیان‌گر دقت بالای تخمین و توانایی پیش‌بینی الگوریتم در شرایط عملیاتی پیچیده است.



شکل ۵-۴- مقایسه بصری مسیر واقعی و پیش‌بینی شده اهداف بالستیک

## ۵.۵ تحلیل حساسیت نسبت به پارامترهای کلیدی

تنظیم دقیق پارامترهای مدل از جمله کوواریانس نویز فرآیند اهمیت بالایی دارد. شکل ۵-۵ تحلیل حساسیت خطای RMS تخمین موقعیت نسبت به تغییرات  $\alpha$  را نمایش می‌دهد. مشاهده می‌شود که وجود مقدار بهینه‌ای برای  $\alpha$  منجر به کمینه شدن خطای تخمین می‌شود، که بیانگر تعادل بین مدل‌سازی دینامیک هدف و مقابله با نویز است.



شکل ۵-۵ - تحلیل حساسیت خطای RMS تخمین موقعیت نسبت به ضریب کوواریانس نویز فرآیند  $\alpha$

## ۶.۵ تحلیل کمی و کیفی نتایج

ارزیابی عددی نشان می‌دهد که الگوریتم پیشنهادی در شرایط شبیه‌سازی شده نسبت به EKF و PF به طور متوسط بیش از ۲۰ درصدی کاهش خطای RMS را به همراه داشته است. این بهبود ناشی از قابلیت تطبیق پذیری بالاتر و یادگیری مداوم پارامترهای مدل در شرایط غیرایده‌آل است.

از دید کیفی، الگوریتم پیشنهادی ضمن پایداری عملکرد در طیف وسیعی از SNR ها و تغییرات ناگهانی در مسیر هدف، توانسته است پیش‌بینی مسیر را با دقت بالایی انجام دهد که برای کاربردهای دفاعی و راداری حیاتی است. نقطه قوت دیگر، کاهش حساسیت به تنظیمات دقیق پارامترهای مدل است که باعث سهولت استفاده و افزایش اعتمادپذیری الگوریتم می‌شود.

## ۷.۵ محدودیت‌ها و چالش‌ها

با وجود نتایج مثبت، چند محدودیت اساسی در این پژوهش مشاهده می‌شود که شبیه‌سازی‌ها با فرضیات ساده شده در مورد مدل نویز و دینامیک اهداف انجام شده است که ممکن است در شرایط واقعی دقیق نباشد. محاسبات الگوریتم یادگیری تقویتی به صورت کامل بهینه نشده‌اند و برای کاربردهای بلندمدت نیازمند بهینه‌سازی‌های سخت‌افزاری و نرم‌افزاری است. داده‌های واقعی آزمایشگاهی در دسترس نبوده و اعتبارسنجی تجربی باقی مانده است.



### ۸.۵ جمع بندی نهایی و چشم انداز آینده

این پژوهش گام مهمی در زمینه پردازش سیگنال رادارهای برد بلند با رویکرد یادگیری تقویتی برداشته است. نتایج نشان می دهد که الگوریتم پیشنهادی می تواند در شرایط چالش برانگیز و نویزی به بهبود چشمگیر دقت تخمین موقعیت و پیش بینی مسیر اهداف کمک کند. چشم اندازهای آینده شامل توسعه مدل های چندعاملی، ادغام یادگیری عمیق، پیاده سازی در داده های واقعی و بهینه سازی محاسباتی است که می تواند کاربردهای عملی این رویکرد را تسهیل نماید.

### ۶. نتیجه گیری و پیشنهادات آینده

#### ۱.۶ نتیجه گیری

در این پژوهش، الگوریتم یادگیری تقویتی تطبیقی برای پردازش سیگنال رادار برد بلند به منظور تخمین موقعیت و پیش بینی مسیر اهداف بالستیک ارائه شد. پیاده سازی و شبیه سازی های انجام شده نشان داد که روش پیشنهادی توانسته است نسبت به الگوریتم های کلاسیک مانند فیلتر کالمن توسعه یافته (EKF) و فیلتر ذرات (PF)، بهبود قابل توجهی در دقت تخمین موقعیت و پایداری عملکرد فراهم کند. کاهش میانگین خطای RMS تخمین موقعیت در بازه های زمانی مختلف و حفظ دقت در شرایط نسبت سیگنال به نویز پایین از جمله نقاط قوت کلیدی این روش هستند. این پژوهش با بهره گیری از قابلیت یادگیری مداوم و سازگاری خودکار با تغییرات دینامیکی هدف، گامی موثر در جهت افزایش قابلیت اطمینان سیستم های راداری برد بلند برداشته است. علاوه بر این، تحلیل حساسیت پارامترهای کلیدی مدل، نشان داد که تنظیم بهینه این پارامترها نقش بسزایی در عملکرد نهایی دارد.

#### ۲.۶ پیشنهادات برای پژوهش های آتی

- با توجه به نتایج مثبت این مطالعه، چندین مسیر برای تحقیقات آینده پیشنهاد می شود:
- (۱) ادغام یادگیری عمیق با یادگیری تقویتی: ترکیب شبکه های عصبی عمیق به منظور استخراج ویژگی های پیچیده تر سیگنال های راداری و افزایش توانایی پیش بینی در محیط های بسیار غیرخطی.
- (۲) پیاده سازی بلادرنگ و بهینه سازی محاسباتی: توسعه نسخه های بهینه شده نرم افزاری و سخت افزاری الگوریتم جهت کاربردهای عملی و زمان واقعی.
- (۳) اعتبارسنجی با داده های واقعی: آزمایش و ارزیابی الگوریتم روی داده های واقعی حاصل از سیستم های راداری برد بلند و اهداف بالستیک.
- (۴) گسترش مدل به اهداف چندگانه: طراحی و بهینه سازی الگوریتم برای ردیابی و پیش بینی مسیر همزمان چند هدف بالستیک.
- (۵) افزایش مقاومت در برابر شرایط محیطی و نویزهای پیچیده: مطالعه و توسعه الگوریتم هایی با پایداری بیشتر در شرایط محیطی متغیر و نویزهای غیراستاندارد.
- (۶) استفاده از شبکه های عصبی کانولوشنی.

## ۷. مراجع

- [1] D. Huang, Y. Shi, and Q. Hu, "Deep Reinforcement Learning Based Target Tracking for Radar Systems," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 4, pp. 2671–2685, Aug. 2021.
- [2] X. Zhang, B. Himed, and J. Li, "Adaptive Waveform Design for Long-Range Radar Using Reinforcement Learning," *IEEE Transactions on Signal Processing*, vol. 69, pp. 2217–2230, Apr. 2021.
- [3] W. Jiang and X. Wang, "Ballistic Target Tracking Using Particle Filter with Adaptive Noise Covariance," *IEEE Sensors Journal*, vol. 20, no. 12, pp. 6705–6715, Jun. 2020.
- [4] S. Haykin, "Cognitive Radar: A Way of the Future," *IEEE Signal Processing Magazine*, vol. 23, no. 1, pp. 30–40, Jan. 2006.
- [5] M. R. Azimi-Sadjadi, "Adaptive Doppler Processing Techniques for Ballistic Missile Tracking," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 45, no. 1, pp. 382–395, Jan. 2009.
- [6] Y. Bar-Shalom and X. R. Li, *Estimation and Tracking: Principles, Techniques, and Software*, Artech House, 1993.
- [7] R. S. Blum, "MIMO Radar: Theory and Application," *IEEE Signal Processing Magazine*, vol. 24, no. 5, pp. 106–114, Sep. 2007.
- [8] H. He, Y. Qin, and Z. Zhang, "Deep Reinforcement Learning for Radar Target Detection and Tracking," *IEEE Access*, vol. 8, pp. 103565–103574, 2020.
- [9] C. P. Teague, S. M. Erfani, and G. R. Poulton, "Long-Range Ballistic Missile Tracking Using Extended Kalman Filter with Neural Network Augmentation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 56, no. 3, pp. 2058–2069, Jun. 2020.
- [10] J. Xie and Q. Fu, "Reinforcement Learning-Based Adaptive Beamforming for Airborne Radar Systems," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 2, pp. 931–942, Apr. 2021.
- [11] X. Liu, J. Han, and X. Cheng, "Particle Filtering for Ballistic Target Tracking in Non-Gaussian Noise Environments," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 56, no. 1, pp. 23–34, Feb. 2020.
- [12] P. Willett, J. Li, and R. Blum, "Signal Processing for MIMO Radar," *IEEE Signal Processing Magazine*, vol. 24, no. 5, pp. 106–114, Sep. 2007.
- [13] F. Gustafsson, *Statistical Sensor Fusion*, Studentlitteratur, 2010.

- [14] M. K. Jaiswal and P. K. Varshney, "Adaptive Radar Tracking Using Reinforcement Learning," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 54, no. 1, pp. 62–74, Feb. 2018.
- [15] D. Silver et al., "Mastering the Game of Go with Deep Neural Networks and Tree Search," *Nature*, vol. 529, pp. 484–489, Jan. 2016.
- [16] L. Liu and H. Ling, "Deep Reinforcement Learning for Radar Signal Processing: A Review," *IEEE Access*, vol. 8, pp. 140027–140042, 2020.
- [17] N. J. Gordon, D. J. Salmond, and A. F. M. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," *IEE Proceedings F - Radar and Signal Processing*, vol. 140, no. 2, pp. 107–113, Apr. 1993.
- [18] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, May 2017.
- [19] M. Szubert et al., "Reinforcement Learning Techniques for Radar Waveform Optimization," *IEEE Journal of Selected Topics in Signal Processing*, vol. 14, no. 1, pp. 135–147, Feb. 2020.
- [20] S. Sarkar and A. Nandi, "A Survey on Reinforcement Learning for Radar Signal Processing," *IEEE Access*, vol. 7, pp. 150600–150615, 2019.