

Arquitetura Híbrida para Rastreamento e Predição de Trajetória de VANTs Baseada em Visão Computacional e Filtro de Kalman

Ighor de Souza Ribeiro¹, Rigel Procópio Fernandes² e Gabriel de Sapienza Luna³

¹Centro de Instrução Almirante Alexandrino (CIAA), Rio de Janeiro/RJ - Brasil

²Diretoria de Sistemas de Armas da Marinha (DSAM), Rio de Janeiro/RJ - Brasil

³Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro/RJ - Brasil

Resumo— Este trabalho propõe uma arquitetura híbrida para rastreamento e predição de trajetória de veículos aéreos não tripulados baseada na integração entre o detector YOLO26, o rastreador local Channel and Spatial Reliability Tracking (CSRT) e um filtro de Kalman com amortecimento de velocidade. O detector é utilizado para inicialização e redetecção periódica do alvo, enquanto o rastreador local mantém a continuidade da trajetória entre sucessivas correções. Os experimentos foram realizados utilizando as 20 sequências do dataset DUT Anti-UAV Tracking, totalizando 24.804 quadros. Os resultados demonstraram desempenho de rastreamento próximo aos obtidos por rastreadores consagrados, com redução de até 38% no erro de predição para horizontes de 30 quadros à frente, indicando que a combinação entre redetecção periódica, rastreamento local e estimação de estado constitui uma alternativa promissora para sistemas anti-VANT de baixo custo computacional.

Palavras-Chave— Anti-VANT, Rastreamento Visual, Predição de Trajetória

I. INTRODUÇÃO

A rápida disseminação dos Veículos Aéreos Não Tripulados (VANTs) em aplicações civis e militares ampliou as preocupações relacionadas à proteção de infraestruturas críticas e áreas sensíveis de defesa, como as Organizações Militares das Forças Armadas. Embora sistemas baseados em radar e radiofrequência continuem desempenhando papel relevante nesse contexto, sua utilização pode ser limitada pelo custo de implantação ou pela dificuldade em detectar drones de pequeno porte operando em baixas altitudes, cuja seção reta radar é reduzida [1], [2].

Nesse cenário, a visão computacional tem se consolidado como uma alternativa complementar para aplicações anti-VANT [2]–[5]. Além de permitir operação passiva, essa abordagem pode aproveitar infraestruturas de segurança já existentes, como sistemas de Circuito Fechado de TV (CFTV), reduzindo custos associados à instalação de sensores dedicados. Mesmo assim, detectar e rastrear drones em imagens do espectro visível permanece uma tarefa desafiadora. O pequeno tamanho aparente dos alvos, a presença de fundos complexos, as mudanças de iluminação, as oclusões temporárias e as manobras bruscas frequentemente observadas em voo afetam diretamente o desempenho dos algoritmos de detecção e rastreamento [1], [3]–[9].

I. S. Ribeiro, ighoribeiro@gmail.com; R. P. Fernandes, rigelfernandes@gmail.com; G. S. Luna, gabriellunacos.ufrj.br. Código e materiais experimentais disponíveis em: <https://github.com/IghorRibeiro/hybrid-uav-tracking-prediction>.

Nesse contexto, este trabalho propõe uma arquitetura híbrida baseada em detecção periódica, rastreamento local e estimação de estado para rastreamento e predição de trajetória de VANTs, com o fito de reduzir o custo computacional sem sacrificar a eficácia do rastreamento ou a qualidade da predição, aspecto primordial para futuras aplicações que visam a neutralização do alvo.

As principais contribuições deste trabalho são:

- 1) A validação experimental de uma arquitetura híbrida YOLO+CSRT+Kalman em base pública do espectro visível, demonstrando que é possível manter desempenho competitivo de rastreamento, mesmo reduzindo significativamente a frequência de execução do detector.
- 2) A análise do impacto da qualidade do rastreamento na predição de trajetória, incluindo a avaliação de um modelo de Kalman com amortecimento de velocidade (*Kalman Damped*), capaz de reduzir o erro de predição em horizontes mais longos e melhorar a consistência das estimativas futuras.

O restante deste artigo está organizado da seguinte forma. A Seção II apresenta os principais trabalhos relacionados ao tema. A Seção III descreve a arquitetura proposta e os métodos empregados. Os resultados experimentais são apresentados na Seção IV, enquanto as conclusões e perspectivas de continuidade são discutidas na Seção V.

II. TRABALHOS RELACIONADOS

O desenvolvimento de sistemas anti-VANT baseados em visão computacional tem sido impulsionado principalmente pelos avanços recentes em detecção de objetos, rastreamento visual e predição de trajetória. Embora essas linhas de pesquisa sejam frequentemente investigadas de forma independente, a integração dos três componentes em uma única arquitetura pode gerar ganhos expressivos.

Na etapa de detecção, arquiteturas da família YOLO têm sido amplamente empregadas devido à combinação entre elevada precisão e baixa latência de inferência. Trabalhos recentes demonstraram a viabilidade da detecção de drones em tempo real utilizando diferentes variantes desses detectores, inclusive em plataformas embarcadas e dispositivos de borda [2], [4], [5], [10]. Apesar dos resultados obtidos, a execução contínua do detector em todos os quadros representa uma parcela significativa do custo computacional desses sistemas.

Como alternativa, diversos autores passaram a explorar arquiteturas híbridas que combinam detecção global e rastreamento local. Nesse tipo de abordagem, o detector é utilizado para inicialização e correção periódica da posição do alvo e o rastreador mantém sua continuidade temporal entre detecções [3], [9], [11]. Estratégias semelhantes também têm sido empregadas em sistemas de perseguição e interceptação de drones, evidenciando o potencial dessa colaboração para reduzir o processamento sem comprometer significativamente a qualidade do rastreamento [12]–[14].

Paralelamente, o rastreamento visual de VANTs tem evoluído de métodos baseados em associação por aparência, como DeepSORT e BoT-SORT [15], [16], para abordagens mais sofisticadas que incorporam mecanismos de atenção, memória temporal e arquiteturas baseadas em *Transformers* [6], [17]–[19]. Tais métodos frequentemente apresentam ganhos expressivos em benchmarks especializados, ao passo que demandam modelos complexos e recursos computacionais incompatíveis com aplicações de baixo custo ou sistemas embarcados.

Além do rastreamento, observa-se um interesse crescente na previsão de trajetória de VANTs. Redes recorrentes, arquiteturas baseadas em *Transformers* e métodos multimodais têm sido empregados para antecipar o movimento futuro dos alvos em cenários dinâmicos [20]–[23], em geral concentrando esforços na maximização da acurácia preditiva.

Nesse contexto, permanece em aberto a investigação de arquiteturas capazes de combinar detecção, rastreamento e previsão em uma solução computacionalmente eficiente. Em particular, ainda são limitados os estudos que analisam como a qualidade e a consistência temporal do rastreamento influenciam o desempenho da previsão de trajetória em sistemas Anti-VANT. Este trabalho busca contribuir para essa lacuna por meio da integração entre um detector leve, um rastreador local e um filtro de Kalman voltado à estimação e previsão do movimento do alvo.

III. MÉTODO PROPOSTO

Esta seção apresenta a arquitetura proposta para detecção, rastreamento e previsão de trajetória de Veículos Aéreos não Tripulados (VANTs). O sistema foi desenvolvido para operar em cenários de vigilância visual utilizando câmeras fixas, buscando equilibrar precisão e custo computacional.

A. Definição do Problema

O problema abordado consiste na detecção e no rastreamento visual de um único VANT do tipo quadcopter em sequências de vídeo. Considera-se um problema de *Single Object Tracking* (SOT), no qual existe no máximo um alvo de interesse por sequência.

O objetivo do sistema é estimar continuamente a posição do drone no plano da imagem, manter a consistência temporal da trajetória observada e fornecer estimativas futuras da posição do alvo para horizontes de curto prazo. Essa abordagem busca aproximar o problema de aplicações de monitoramento perimetral e apoio à decisão em sistemas anti-VANT no tocante à solução de tiro.

B. Visão Geral do Pipeline

Com o objetivo de atender ao problema definido na seção anterior, foi desenvolvido um pipeline que combina detecção, rastreamento e previsão em uma arquitetura única. A estratégia adotada busca reduzir a frequência de execução do detector, etapa computacionalmente mais custosa do sistema, sem comprometer a capacidade de acompanhar o movimento do alvo ao longo da sequência.

A operação é iniciada por uma detecção global realizada pelo YOLO26. Nos quadros subsequentes, o acompanhamento do alvo é executado pelo CSRT, enquanto redetecções periódicas ($N_{red} = 15$) corrigem possíveis desvios acumulados. Paralelamente, um filtro de Kalman estima o estado cinemático do alvo e produz previsões de curto prazo para sua trajetória futura.

A Fig. 1 ilustra a arquitetura proposta, composta pelo detector YOLO26n, rastreador CSRT e filtro de Kalman (Seção III-E), operando em regime de redetecção periódica.

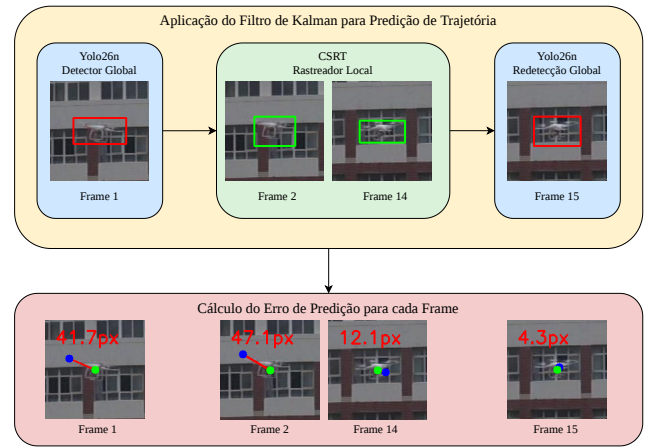


Fig. 1: Fluxo operacional do pipeline proposto, integrando detecção global, rastreamento local e previsão de trajetória.

C. Detector Global YOLO26

Como detector global, foi adotada a variante YOLO26 nano (YOLO26n) [24], especializada para a classe *drone* por meio de *fine-tuning* no conjunto de detecção do DUT Anti-UAV [25]. A escolha baseia-se no compromisso entre precisão e velocidade de inferência, adequado ao monitoramento visual em tempo real. Em relação às versões anteriores da família, o YOLO26 introduz uma arquitetura *end-to-end* sem NMS (*Non-Maximum Suppression*), reportada como até 43% mais rápida em inferência via CPU, característica relevante para o regime de redetecção periódica em CPU/borda adotado neste trabalho.

No pipeline proposto, o YOLO26n é acionado para inicialização e redetecção periódica do alvo. A cada acionamento, somente a detecção de maior confiança da classe *drone* é utilizada para atualizar o sistema, enquanto os demais frames são processados pelo rastreador.

D. Rastreador Local CSRT

Entre as execuções do detector global, o acompanhamento contínuo do alvo é realizado pelo rastreador CSRT [26]. O

algoritmo é inicializado a partir da caixa delimitadora fornecida pelo YOLO26 e atualizado nos quadros subsequentes. Para limitar erros acumulados, o rastreador é reinicializado a cada ciclo de redetecção utilizando a posição fornecida pelo detector.

Além disso, o CSRT opera de forma eficiente em CPU [3], o que representa uma vantagem importante em plataformas que não dispõem de GPU dedicada ou cuja GPU é compartilhada com outras tarefas críticas, cenário comum em dispositivos de borda. Essa escolha foi validada experimentalmente frente a rastreadores locais alternativos (Seção IV-C).

E. Predição com Filtro de Kalman

Além da localização instantânea da posição do alvo, o sistema busca estimar sua trajetória futura. Para tal, foi empregado um filtro de Kalman baseado em um modelo cinemático de velocidade constante [27], cujo vetor de estado é definido por

$$\mathbf{x} = [x, y, \dot{x}, \dot{y}]^T, \quad (1)$$

onde x e y representam as coordenadas, em pixels, do centroide do alvo no plano da imagem, enquanto $\dot{x}$ e $\dot{y}$, correspondentes às derivadas temporais de x e y , representam as velocidades instantâneas do drone nos eixos horizontal e vertical da imagem, respectivamente.

A etapa de predição segue a formulação clássica do filtro de Kalman,

$$\mathbf{x}_{k|k-1} = \mathbf{F}\mathbf{x}_{k-1|k-1}, \quad (2)$$

$$\mathbf{P}_{k|k-1} = \mathbf{F}\mathbf{P}_{k-1|k-1}\mathbf{F}^T + \mathbf{Q}, \quad (3)$$

em que $\mathbf{P}$ representa a matriz de covariância do erro de estimação e $\mathbf{Q}$ a covariância associada ao ruído de processo.

Considerando a hipótese de velocidade constante entre observações consecutivas e intervalo temporal unitário entre quadros ($\Delta t = 1$), a matriz de transição de estados é dada por

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (4)$$

Por sua vez, as observações são compostas apenas pelas coordenadas do centro do alvo, sendo modeladas pela matriz de observação

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}. \quad (5)$$

As observações utilizadas na etapa de correção correspondem ao centro da caixa delimitadora estimada pelo módulo de rastreamento. Dessa forma, a atualização do filtro pode ser realizada tanto a partir das redetecções periódicas fornecidas pelo YOLO26 quanto das posições estimadas pelo rastreador CSRT nos quadros intermediários.

Embora a hipótese de velocidade constante represente uma simplificação da dinâmica real de voo, ela fornece uma aproximação adequada para horizontes curtos de predição e estabelece uma base consistente para a avaliação das modificações propostas neste trabalho.

1) *Kalman Persistente*: Na configuração persistente, o estado e a matriz de covariância são mantidos ao longo de toda a sequência, sem reinicialização e sem amortecimento das componentes de velocidade. Dessa forma, as estimativas de velocidade acumuladas pelo filtro são preservadas entre detecções, contribuindo para maior estabilidade da trajetória estimada e reduzindo oscilações decorrentes de pequenas variações nas observações.

2) *Kalman Damped*: Embora o modelo de velocidade constante apresente bom desempenho em movimentos suaves, drones frequentemente executam manobras que violam essa hipótese. Com o intuito de reduzir erros de extrapolação provocados por acelerações abruptas, foi introduzido um fator de amortecimento $\gamma = 0,95$ aplicado às componentes de velocidade antes da etapa de predição:

$$\dot{x}_k \leftarrow \gamma \dot{x}_k, \quad \dot{y}_k \leftarrow \gamma \dot{y}_k. \quad (6)$$

Esse procedimento reduz a influência de velocidades elevadas acumuladas em instantes anteriores, produzindo previsões mais conservadoras para horizontes temporais mais longos.

A Fig. 2 apresenta exemplos das principais etapas executadas pelo sistema: a atuação do detector global na inicialização e redetecção do alvo, o acompanhamento contínuo realizado pelo CSRT e as estimativas produzidas pelo filtro de Kalman para diferentes horizontes de predição.

IV. RESULTADOS

Esta seção apresenta os resultados experimentais obtidos para os detectores, métodos de rastreamento e estratégias de predição avaliadas. São discutidos os resultados dos detectores YOLO26 utilizados no pipeline, comparados os métodos de rastreamento e, por fim, são analisadas as diferentes estratégias de predição de trajetória baseadas em filtro de Kalman.

A. Configuração Experimental

Os experimentos foram realizados utilizando as 20 sequências do dataset DUT Anti-UAV Tracking, totalizando 24.804 quadros. O DUT foi adotado por apresentar predominância de câmeras estáticas, além de contemplar diferentes condições de iluminação, escala do alvo, complexidade do fundo e padrões de movimento.

As métricas de acurácia (SA, SR@0.5, IoU médio, AUC) e de erro de predição (H5, H10, H15, H30) foram computadas a partir das anotações do dataset. O tempo de processamento por quadro e a taxa de quadros por segundo (FPS) foram medidos primariamente em um dispositivo de borda sem GPU dedicada (Apple M2, MPS), cenário representativo de sistemas de vigilância de baixo custo. Adicionalmente, comparações complementares em uma GPU de datacenter (Tesla T4) são discutidas na Seção IV-D.

Para o rastreamento, adotaram-se as métricas *State Accuracy* (SA), *Success Rate* (SR@0.5), calculadas a partir da sobreposição entre a caixa estimada e a anotação de referência (*ground truth*). A sobreposição é medida pela *Intersection over Union* (IoU), definida como

$$IoU = \frac{B_{pred} \cap B_{gt}}{B_{pred} \cup B_{gt}}, \quad (7)$$

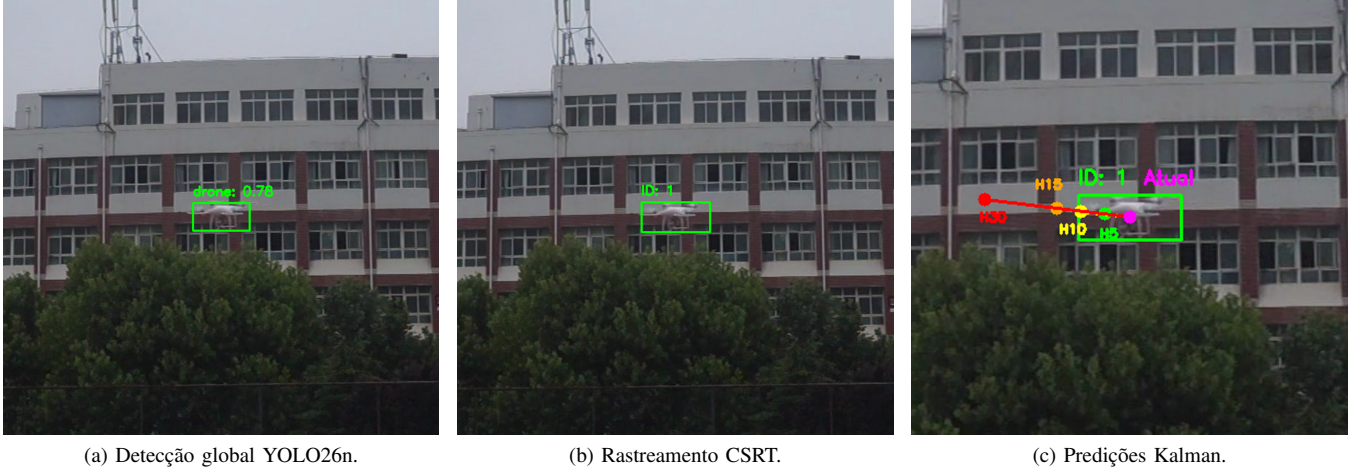


Fig. 2: Exemplos das principais etapas do pipeline proposto: (a) detecção global, (b) rastreamento local e (c) predição de trajetória com filtro de Kalman.

onde B_{pred} e B_{gt} são, respectivamente, a caixa predita e a caixa real. A métrica SA corresponde à média da IoU ao longo da sequência, enquanto $SR@0.5$ é a proporção de quadros com IoU superior a 0,5.

A capacidade de predição foi avaliada pelo erro de posição, definido como a distância euclidiana entre o centro previsto e o centro real do alvo, em pixels:

$$e_H = \sqrt{(x_H - x_{gt})^2 + (y_H - y_{gt})^2}, \quad (8)$$

em que (x_H, y_H) é a posição estimada para um horizonte H e (x_{gt}, y_{gt}) é a posição real. Foram considerados os horizontes H5, H10, H15 e H30, correspondentes a 5, 10, 15 e 30 quadros à frente do instante atual.

B. Desempenho dos Detectores

A Tabela I apresenta os resultados obtidos pelas variantes YOLO26 avaliadas neste trabalho. Observa-se que a variante *medium* apresentou desempenho superior em todas as métricas de detecção avaliadas. Entretanto, os ganhos observados foram relativamente modestos quando comparados ao aumento do custo computacional associado ao modelo mais complexo.

TABELA I: Desempenho dos detectores YOLO26 no dataset DUT Anti-UAV.

Modelo	Precisão	Recall	mAP@0.5	mAP@0.5:0.95
YOLO26n	0.956	0.899	0.944	0.616
YOLO26m	0.969	0.923	0.961	0.662

Considerando o compromisso custo-benefício, o YOLO26n foi selecionado para os experimentos seguintes.

C. Comparação entre Rastreadores Locais

Para validar a escolha do CSRT como rastreador local, o componente foi substituído por três rastreadores alternativos: KCF [28], MOSSE [29] e MedianFlow [30], mantendo o restante do pipeline inalterado. A Tabela II apresenta os resultados.

TABELA II: Comparação de rastreadores locais no pipeline proposto

Rastreador	SA	SR@0.5	FPS	H30 (px)
CSRT (adotado)	0.7858	0.8108	66.4	139.92 ± 97.44
KCF	0.7394	0.7409	227.2	179.47 ± 189.35
MOSSE	0.8050	0.8108	406.4	191.80 ± 228.68
MedianFlow	0.7226	0.7356	258.8	168.99 ± 121.08

O CSRT apresentou o menor erro em H30, embora o MOSSE tenha obtido maior SA e FPS, com $SR@0.5$ equivalente. Apesar de operar a 406,4 FPS, o MOSSE apresentou erro 37% maior em H30 (191,8 vs. 139,9 px) e maior dispersão entre sequências (228,7 vs. 97,4 px). Esses resultados sustentam a escolha do CSRT.

D. Comparação entre Métodos de Rastreamento e Predição

A Tabela III reúne o desempenho de rastreamento (SA, $SR@0.5$, IoU médio, AUC) e de predição de trajetória para todos os métodos comparados neste trabalho, incluindo dois rastreadores baseados em *Transformer* (OSTrack [32] e ODTrack [33], inicializados a partir da anotação do primeiro quadro de cada sequência), além de DeepSORT [15] e ByteTrack [31]. O ODTrack foi selecionado por ser apontado por Qiao *et al.* [34] como o método de maior precisão entre os avaliados especificamente no dataset DUT Anti-UAV, tornando-o um baseline de estado da arte relevante para este trabalho. Para padronizar a avaliação da predição, as trajetórias obtidas por todos os métodos foram submetidas ao mesmo filtro de Kalman persistente, sem amortecimento.

Observa-se que o ByteTrack obteve a maior SA, enquanto o ODTrack apresentou os maiores valores de $SR@0.5$, IoU médio e AUC. O pipeline híbrido manteve desempenho próximo aos métodos avaliados, mesmo executando o detector apenas em intervalos periódicos.

Em uma GPU de *datacenter* (Tesla T4), o BoT-SORT (YOLO26n) atinge 23,7 FPS contra 12,0 FPS do Híbrido. No Apple M2/MPS, sem GPU dedicada, a relação se inverte: 42,7 FPS para o BoT-SORT contra 66,4 FPS para o Híbrido.

Embora o ODTrack tenha apresentado a maior acurácia instantânea, seu erro em H30 (167,3 px) foi superior ao do

TABELA III: Comparação dos métodos de rastreamento e previsão de trajetória no dataset DUT Anti-UAV Tracking (Erro em H30 em pixels, média $\pm$ desvio-padrão entre sequências e FPS medido em Apple M2/MPS).

Método	SA	SR@0.5	IoU médio	AUC	H30 (px)	Ganho (%)	FPS (M2)
BoT-SORT (YOLO26n)	0.828	0.832	0.659	0.662	226.2 $\pm$ 303.6	–	42.7
DeepSORT [15]	0.830	0.834	0.662	0.664	261.9 $\pm$ 446.0	-15.8	4.5
ByteTrack [31]	0.833	0.836	0.660	0.663	269.1 $\pm$ 463.3	-19.0	58.7
OSTrack [32]	0.803	0.831	0.668	0.672	204.3 $\pm$ 206.4	+9.7	13.7
ODTrack [33]	0.819	0.860	0.707	0.709	167.3 $\pm$ 160.8	+26.0	7.3
Híbrido (proposto)	0.789	0.811	0.645	0.647	139.9 $\pm$ 97.4	+38.2	66.4

Híbrido (139,9 px), assim como o do OSTRack (204,3 px). O resultado sugere que métricas instantâneas de rastreamento, isoladamente, não são suficientes para caracterizar a qualidade da previsão de trajetória.

E. Impacto das Modificações do Filtro de Kalman

Com o objetivo de avaliar individualmente a contribuição das modificações propostas para o filtro de Kalman, foram realizados experimentos comparando a configuração persistente (sem amortecimento) e a versão com amortecimento (*Kalman Damped*), mantendo fixos o detector, o rastreador CSRT e o regime de detecção periódica.

A Fig. 3 compara o erro médio de previsão das duas configurações. O amortecimento reduziu o erro em todos os horizontes avaliados: em H30, a redução foi de 139,9 para 109,1 px, correspondente a 22%, além de reduzir o desvio-padrão de 97,4 para 68,8 px. No dataset avaliado, o ganho foi mais expressivo nos horizontes longos.

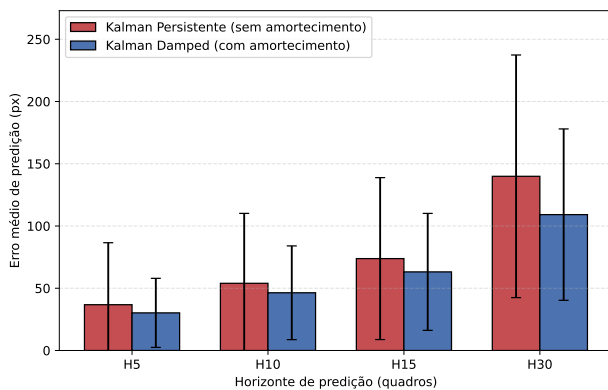


Fig. 3: Erro médio de previsão do filtro de Kalman com e sem amortecimento (*damping*) em função do horizonte temporal, com barras de erro indicando o desvio-padrão entre as 20 sequências avaliadas.

F. Síntese: Rastreamento vs. Previsão

A Tabela III mostra que o pipeline Híbrido apresentou erro de 139,9 px em H30 sob a configuração persistente, uma redução de 38,2% em relação ao BoT-SORT. Com a introdução do amortecimento, avaliada especificamente no pipeline Híbrido, o erro foi reduzido para 109,1 px, o menor valor obtido pelo sistema proposto, com desvio-padrão de 68,8 px. Para referência, a largura média dos drones no dataset é de ≈ 70 px, de modo que o erro alcançado em H30 corresponde a cerca de $1,5\times$ a dimensão típica do alvo.

Os intervalos de confiança de 95% para H30 foram [94,3; 185,5] px para o Híbrido e [84,1; 368,3] px para o BoT-SORT. A menor dispersão observada no Híbrido indica maior consistência entre sequências. Assim, os resultados sugerem que métricas tradicionais de rastreamento, isoladamente, não são suficientes para caracterizar a qualidade da previsão de trajetória.

V. CONCLUSÕES

Este trabalho apresentou uma arquitetura híbrida para rastreamento e previsão de trajetória de VANTs baseada na integração entre um detector YOLO26, um rastreador local CSRT e um filtro de Kalman. A proposta reduz a dependência da execução contínua do detector por meio de redetecções periódicas, mantendo o acompanhamento do alvo entre elas.

Os resultados demonstraram que o pipeline híbrido manteve desempenho competitivo de rastreamento em relação aos métodos avaliados e apresentou menor erro de previsão em H30 sob o protocolo comum de Kalman persistente. O erro foi de 139,9 px, contra 167,3 px do ODTrack e 226,2 px do BoT-SORT, enquanto o processamento atingiu 66,4 FPS no Apple M2/MPS.

No pipeline Híbrido, a introdução do amortecimento de velocidade reduziu o erro em H30 de 139,9 para 109,1 px, uma redução de 22%, com desvio-padrão de 68,8 px. Esse foi o melhor resultado obtido pelo pipeline proposto no dataset avaliado, indicando que a atenuação da velocidade estimada pode melhorar a previsão, particularmente em horizontes mais longos.

De forma geral, os resultados sugerem que arquiteturas híbridas constituem uma alternativa viável para sistemas anti-VANT baseados em visão computacional, oferecendo um compromisso entre rastreamento, previsão e custo computacional. Os resultados também indicam que métricas tradicionais de rastreamento, quando analisadas isoladamente, não são suficientes para caracterizar sistemas cujo objetivo inclui a previsão da trajetória.

A. Trabalhos Futuros

Como continuidade desta pesquisa, busca-se investigar a extensão da arquitetura proposta para cenários de *Multiple Object Tracking* (MOT), permitindo o acompanhamento simultâneo de múltiplos VANTs. Também constitui uma perspectiva relevante a avaliação de outros métodos de rastreamento e de técnicas mais avançadas de previsão de trajetória, incluindo abordagens baseadas em aprendizado profundo e arquiteturas fundamentadas em mecanismos de atenção.

Outra etapa importante consiste na validação da arquitetura em condições reais. Embora os experimentos realizados sobre o dataset DUT Anti-UAV tenham permitido comparar os métodos em um ambiente controlado e reproduzível, a aplicação do pipeline em sistemas reais de Circuito Fechado de TV (CFTV) ainda permanece em aberto.

Por fim, a integração da arquitetura com sistemas de apoio ao engajamento, capazes de utilizar as estimativas de trajetória para auxiliar mecanismos de apontamento, acompanhamento automático ou futuras contramedidas, representa uma direção promissora para ampliar a aplicabilidade prática da solução proposta.

REFERÊNCIAS

- [1] H. Lee, D. Lim, S.-H. Han, J. Lee, S.-h. Lee, J. Park, "An empirical study of ptz camera-based drone tracking under latency and detector challenges," *Machine Vision and Applications*, vol. 36, no. 4, p. 84, 2025. [Online]. Available: <https://doi.org/10.1007/s00138-025-01708-6>
- [2] N. Lafta S. Hreshree, "Design and implementation of anti-drone system based on computer vision with you only look once (yolo) algorithm," *Journal of Internet Services and Information Security*, vol. 15, pp. 324–334, 11 2025. [Online]. Available: <https://doi.org/10.58346/JISIS.2025.14.024>
- [3] E. Smith Y. Diskin, "A modular algorithmic architecture for small uav detection and tracking," in *NAECON 2025 - IEEE National Aerospace and Electronics Conference*, 2025, pp. 1–6. [Online]. Available: <https://doi.org/10.1109/NAECON65708.2025.11235331>
- [4] X. Zhou, G. Yang, Y. Chen, C. Gao, B. Zhao, L. Li, B. M. Chen, "Admnet: Anti-drone real-time detection and monitoring," in *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2023, pp. 3009–3016. [Online]. Available: <https://doi.org/10.1109/IROS55552.2023.10341901>
- [5] B. Wang, J. Li, M. Zhou, Q. Lu, "Drone detection and tracking: An edge-deployable efficient algorithm based on vision sensor," *IEEE Sensors Journal*, vol. 25, no. 17, pp. 34 126–34 140, 2025. [Online]. Available: <https://doi.org/10.1109/JSEN.2025.3588414>
- [6] Y. Wang, T. Xu, J. Li, "Focustrack: A self-adaptive local sampling algorithm for efficient anti-uav tracking," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 63, pp. 1–14, 2025. [Online]. Available: <https://doi.org/10.1109/TGRS.2025.3562958>
- [7] Y. W. Jun S.-C. Kim, "Segment-based uav tracking via shape-aware kalman filter," *KSII Transactions on Internet and Information Systems*, vol. 20, no. 3, pp. 1429–1450, March 2026. [Online]. Available: <https://doi.org/10.3837/tiis.2026.03.016>
- [8] S. Yang, W. Zhou, S. Qu, W. Chu, Q. Yang, S. Sigrarom, B. C. Khoo, "Fast and high-accuracy state estimator for some unknown dynamic objects with a stereo camera in aerial tracking," *Aerospace Science and Technology*, vol. 169, p. 111454, 2026. [Online]. Available: <https://doi.org/10.1016/j.ast.2025.111454>
- [9] S. Liu, T. Xu, X.-F. Zhu, X.-J. Wu, J. Kittler, "Learning adaptive detection and tracking collaborations with augmented uav synthesis for accurate anti-uav system," *Expert Systems with Applications*, vol. 282, p. 127679, 04 2025. [Online]. Available: <https://doi.org/10.1016/j.eswa.2025.127679>
- [10] J. You, Z. Ye, J. Gu, J. Pu, "Uav-pose: A dual capture network algorithm for low altitude uav attitude detection and tracking," *IEEE Access*, vol. 11, pp. 129 144–129 155, 2023. [Online]. Available: <https://doi.org/10.1109/ACCESS.2023.3333394>
- [11] Y. Luo, H. Cha, T. Fu, H. Tang, B. Tian, "Adaptive multi-feature fusion-based uav detection and tracking collaboration framework trigger mechanism," in *2025 IEEE 7th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, vol. 7, 2025, pp. 1103–1107. [Online]. Available: <https://doi.org/10.1109/IMCEC66174.2025.11331589>
- [12] K. Kim, J. Kim, H.-G. Lee, J. Choi, J. Fan, J. Joong, "Uav chasing based on yolov3 and object tracker for counter uav systems," *IEEE Access*, vol. 11, pp. 34 659–34 673, 2023. [Online]. Available: <https://doi.org/10.1109/ACCESS.2023.3264603>
- [13] X. Cui, L. Wan, L. Kong, J. Li, C. Zhang, R. Zhao, P. Wu, S. He, "Strongsiatracker: A siamese tracker with dynamic global detection for robust anti-uav tracking," in *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2025, pp. 6685–6695. [Online]. Available: <https://doi.org/10.1109/CVPRW67362.2025.00654>
- [14] T. Zhang, R. Lu, X. Yang, X. Xie, J. Fan, B. Tang, "Uav hunter: A net-capturing uav system with improved detection and tracking methods for anti-uav defense," *Drones*, vol. 8, no. 10, p. 573, 2024. [Online]. Available: <https://doi.org/10.3390/drones8100573>
- [15] Y. Ghazlane, A. E. H. Alaoui, H. Medomi, H. Bnouachir, "Real-time airborne target tracking using deepsort algorithm and yolov7 model," *International Journal of Advanced Computer Science and Applications*, vol. 15, no. 2, 2024. [Online]. Available: <http://dx.doi.org/10.14569/IJACSA.2024.0150248>
- [16] Y.-H. Chen, "Strong baseline: Multi-uav tracking via yolov12 with bot-sort-reid," in *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2025, pp. 6575–6584. [Online]. Available: <https://doi.org/10.1109/CVPRW67362.2025.00654>
- [17] X. Cui, X. Li, P. Wu, X. Liu, S. He, "Taptrack: An efficient temporal-aware prompt tracker for infrared anti-uav tracking," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 64, pp. 1–14, 2026. [Online]. Available: <https://doi.org/10.1109/TGRS.2026.3682081>
- [18] J. Zhang, Y. Lin, X. Zhou, P. Shi, X. Zhu, D. Zeng, "Precision in pursuit: a multi-consistency joint approach for infrared anti-uav tracking," *The Visual Computer*, vol. 41, no. 4, pp. 2187–2202, 2025. [Online]. Available: <https://doi.org/10.1007/s00371-024-03525-w>
- [19] J. K. Park J.-H. Han, "Memlotrack: Enhancing tir anti-uav tracking with memory-integrated low-rank adaptation," *Sensors*, vol. 25, no. 23, p. 7359, 2025. [Online]. Available: <https://doi.org/10.3390/s25237359>
- [20] W. Wang, R. Gao, C. Du, Y. Li, "Enhanced transformer-based uav-to-uav tracking using distractor elimination and trajectory prediction," in *2024 IEEE 17th International Conference on Signal Processing (ICSP)*, 2024. [Online]. Available: <https://doi.org/10.1109/ICSP.2024.00000>
- [21] H. Lei J. Su, "Robust uav trajectory prediction under diverse disturbances via teacher-student framework," *Expert Systems with Applications*, vol. 308, p. 131185, 2026. [Online]. Available: <https://doi.org/10.1016/j.eswa.2026.131185>
- [22] Z. Dang, B. Sun, C. Li, S. Yuan, X. Huang, Z. Zuo, "Ca-lstm: An improved lstm trajectory prediction method based on infrared uav target detection," *Electronics*, vol. 12, no. 19, p. 4081, 2023. [Online]. Available: <https://doi.org/10.3390/electronics12194081>
- [23] X. Nian, H. Liu, X. Dai, "Audio-visual fusion sim2real platform for anti-uav detection and tracking," *Drones*, vol. 10, no. 3, p. 190, 2026. [Online]. Available: <https://doi.org/10.3390/drones10030190>
- [24] Ultralytics, "Yolo26: Real-time object detection models," <https://docs.ultralytics.com/models/yolo26>, 2025, accessed: Jun. 2026.
- [25] J. Zhao, J. Zhang, D. Li, D. Wang, "Vision-based anti-uav detection and tracking," 2022. [Online]. Available: <https://doi.org/10.48550/arXiv.2205.10851>
- [26] A. Lukežić, T. Vojir, L. C. Zajc, J. Matas, M. Kristan, "Discriminative correlation filter with channel and spatial reliability," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 4847–4856. [Online]. Available: <https://doi.org/10.1109/CVPR.2017.515>
- [27] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME-Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960. [Online]. Available: <https://doi.org/10.1115/1.3662552>
- [28] J. F. Henriques, R. Caseiro, P. Martins, J. Batista, "High-speed tracking with kernelized correlation filters," *IEEE transactions on pattern analysis and machine intelligence*, vol. 37, no. 3, pp. 583–596, 2014.
- [29] D. S. Bolme, J. R. Beveridge, B. A. Draper, Y. M. Lui, "Visual object tracking using adaptive correlation filters," in *2010 IEEE computer society conference on computer vision and pattern recognition*. IEEE, 2010, pp. 2544–2550.
- [30] Z. Kalal J. Mikolajczyk, Krystian Matas, "Forward-backward error: automatic detection of tracking failures," in *2010 20th International Conference on Pattern Recognition*. IEEE, 2010, pp. 2752–2755.
- [31] Y. Zhang, P. Sun, Y. Jiang, D. Yu, F. Weng, Z. Yuan, P. Luo, W. Liu, X. Wang, "Bytetrack: Multi-object tracking by associating every detection box," in *Proceedings of the European Conference on Computer Vision (ECCV)*. Springer, 2022, pp. 1–21.
- [32] B. Ye, H. Chang, B. Ma, S. Shan, "Joint feature learning and relation modeling for tracking: A one-stream framework," in *Proceedings of the European Conference on Computer Vision (ECCV)*. Springer, 2022, pp. 341–357.
- [33] Y. Zheng, B. Zhong, Q. Liang, Z. Mo, S. Zhang, X. Li, "Odtrack: Online dense temporal token learning for visual tracking," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 38, no. 7, 2024, pp. 7150–7158.
- [34] X. Qiao, D. Zhang, Y. Guo, J. Gao, Z. Zhao, X. Li, "Semantic-aware temporal adaptation for UAV anti-UAV tracking," *arXiv preprint arXiv:2607.26511*, 2026.