WARC
2026 World "AI+Radio" Challenge
"AI+无线电"挑战赛

GDPS
全球开发者先锋大会
Global Developer Pioneers Summit

【CommAndNetwork】汇报路演

基于目标条件关联与迭代位置修正的多站多模态融合的无人机识别与三维定位





2026.10.25　　中国 · 上海

团队介绍

何继光 导师	大湾区大学副教授、研究员、博导；奥卢大学兼职教授（Docent）。 2018年获奥卢大学通信工程博士学位；IEEE高级会员、IEEE期刊编委。 研究方向：6G 无线通信、感知及**通感一体化**。
郑灿 队长、汇报人	西安交通大学信息工程本科；高丽大学电子工程硕士；大湾区大学访问学生。 曾在爱立信、英特尔中国研究院实习。 研究方向：AI/ML 使能无线通信、多模态通信与感知。
王俊杰 队员	郑州大学通信工程本科；郑州大学信息与通信工程专业硕博连读。 曾公派赴加拿大西安大略大学留学，现于大湾区大学开展学术访问。 研究方向：**无人机检测、身份认证及定位**。
强冠 队员	上海大学计算机科学与技术本科；加拿大西安大略大学攻电气与信息工程硕博连读。 研究方向：边缘计算、**智传网**与边缘推理优化。

目录

01 研究背景

02 系统模型与问题表述

03 方案设计

04 性能比较

05 结论与展望

研究背景

面向频谱管理与低空安全的多模态无人机感知

- **频谱智能化管理**

  频谱是重要战略资源和关键发展要素。挑战赛围绕 AI 与无线电融合，推动频谱智能感知、资源高效利用及相关产业应用。[1]

- **低空安全监管**

  检测目标有无、识别机型并确定三维位置，为异常飞行发现、空域核验及后续处置提供依据，保障低空作业安全。[2]

- **白名单目标区分**

  已报备无人机、授权目标或已知通信用户仍会出现在混合观测中。识别白名单对象，减少无效告警，将分类与定位重点放在名单外目标上。

- **6G通感一体化与多源感知**

  6G 将通感一体化列为典型应用场景。本赛题融合无线电、雷达和光电信号，研究多模态低空目标的识别与三维定位。[3]

资料：[1] 主办方启动材料、普陀区政府（2026）；[2]《无人驾驶航空器飞行管理暂行条例》；[3] ITU-R M.2160（2023）。

目录

01 研究背景

02 系统模型与问题表述

03 方案设计

04 性能比较

05 结论与展望

应用场景

多模态输入



多站无线电



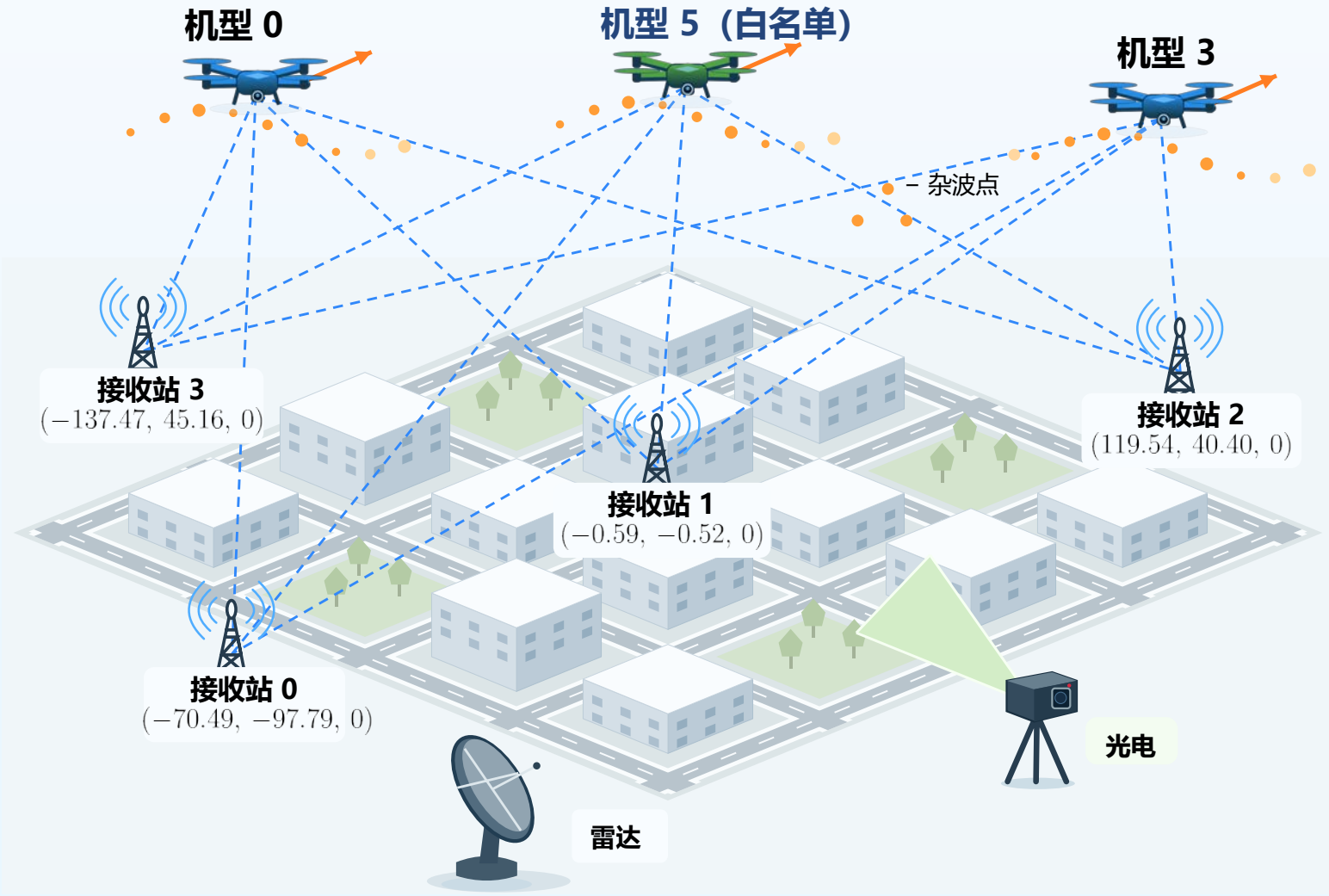
雷达



光电

(单帧RGB)

场景示意



输出结果

非白名单无人机
识别与三维定位



$(0, \hat{p}_0)$

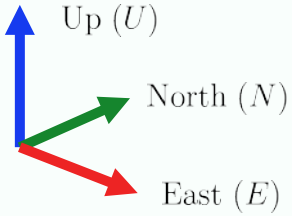


$(3, \hat{p}_3)$



机型 5
白名单

ENU 坐标系



系统模型与评价指标

01 系统模型

输入观测： 多站 RF 观测 $\mathcal{X} = \{(X_k, m_k)\}_{k=0}^{K-1}$

雷达历史点集 $\mathcal{R} = \{(\boldsymbol{q}_\ell, \Delta t_\ell)\}_{\ell=1}^{N_r}$

光电图像 $I,\ m_{\mathrm{EO}}$

模型输出： 机型与当前三维位置

$$\widehat{\mathcal{Y}} = f_\theta(\mathcal{X}, \mathcal{R}, I, \boldsymbol{m}, \mathcal{A}) = \{(\hat{c}_j, \hat{\boldsymbol{p}}_j)\}_{j=1}^{\hat{n}}, \quad \hat{c}_j \notin \mathcal{A},\ \hat{\boldsymbol{p}}_j \in \mathbb{R}^{d_{\mathrm{pos}}}$$

坐标模型： $\boldsymbol{q}_\ell \approx R_z(\varphi_b)\boldsymbol{p}_j(\Delta t_\ell) + \boldsymbol{b}_b + \boldsymbol{\varepsilon}_\ell$

02 评价指标

分类 F1 $\mathrm{F1}_s = 2|P_s \cap G_s|/(|P_s| + |G_s|)$

定位 LocMass $\mathrm{LocMass}_s = |G_s|^{-1} \sum_{i \in M_s} \exp(-d_i/(100\,\mathrm{m}))$

综合得分 $\mathrm{Score} = 100\ \mathrm{mean}_s[0.5\,\mathrm{F1}_s + 0.5\,\mathrm{LocMass}_s]$

符号定义

- X_k, m_k, k, K
  站点 I/Q、可用掩码、站点索引及总数
- $\boldsymbol{q}_\ell, \Delta t_\ell, N_r, \ell$
  雷达点坐标、相对时间、点数与索引
- $I, m_{\mathrm{EO}}, \mathcal{A}$
  可选图像、图像可用标记与白名单
- $f_\theta, \theta, \boldsymbol{m}$
  融合模型、参数与模态可用掩码
- $\widehat{\mathcal{Y}}, \hat{n}, j$
  输出集合、目标数与目标索引
- $\hat{c}_j, \hat{\boldsymbol{p}}_j, d_{\mathrm{pos}}$
  机型、当前 ENU 坐标及坐标维数
- $b, R_z, \varphi_b, \boldsymbol{b}_b$
  视图、竖轴旋转、偏航角及平移量
- $\boldsymbol{p}_j(\Delta t_\ell), \varepsilon_\ell$
  历史目标位置与观测噪声
- s, P_s, G_s
  样本索引、上报机型集合与真实机型集合
- i, M_s, d_i
  匹配对索引、同机型匹配集合与三维误差

数据集与观测难点

针对观测扰动、站点缺失与多模态不完整，分别设计处理方法

训练样本 **7,362**　目标实例 **12,352**　公开测试样本 **487**　含白名单样本约 **28.0%**

观测	观测条件与难点	本方案处理
无线电 I/Q 最多 4 站 10 ms 观测	多目标信号叠加、观测存在扰动； 站间非严格同步，采样率可不同； 部分站点可能缺失。	归一化与频谱扰动增强； 共享 CNN 分站提取频谱特征， 结合站点身份与可用标记融合。
雷达点列 历史观测 坐标 + 时间	杂波多、雷达点不带目标身份； 存在坐标偏差，历史点也不能 直接作为目标的当前位置。	标定后，按机型、位置和时间 关联雷达点，结合迭代修正与 运动补偿估计当前坐标。
光电图像 单帧 RGB 图像	图像并非每个样本都有， 训练样本覆盖率仅 16.5%； 小目标细节模糊，视觉证据有限。	预训练视觉模型提取视觉特征， 结合可用标记作为辅助信息； 缺图时仍由射频与雷达推断。

目录

01 研究背景

02 系统模型与问题表述

03 方案设计

04 性能比较

05 结论与展望

方案思路：从 DETR 到目标条件定位

DETR [1] 目标检测	Query：可学习的目标查询 Key/Value：图像空间特征	Transformer 交叉注意力	类别与检测框集合
前期工作 [2] CSI 引导多用户定位	Query：信道状态信息（CSI）特征 Key/Value：图像空间特征	跨模态 Transformer 用户信息交互	逐用户三维坐标 保留导频用户索引
本挑战方案 多模态无人机定位	条件：机型 + 位置初估 + 射频/视觉特征 观测：雷达点的相对位置与时间	目标条件关联 与位置迭代修正	目标当前三维坐标 白名单外目标列表

迁移到本赛题：以目标条件引导多模态信息关联，再逐轮修正目标位置。

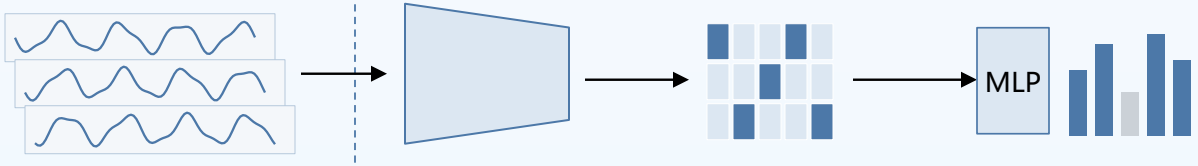
[1] Carion et al., "End-to-End Object Detection with Transformers," ECCV 2020.
[2] Zheng et al., "Vision–Wireless Fusion for Multi-User Localization: A Cross-Modal Transformer Approach," arXiv:2609.23372, 2026.

方案总览：多站感知与多模态定位

A 多站射频感知

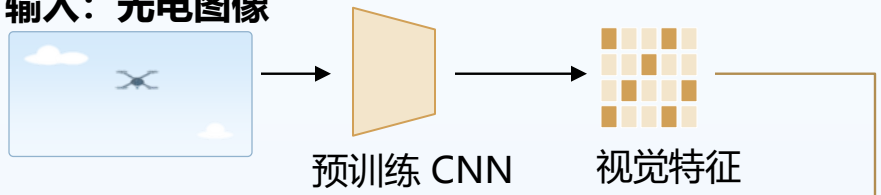
作用：融合多站无线电信息，识别目标机型

输入：多站 RF 观测



输出：机型 + 射频特征

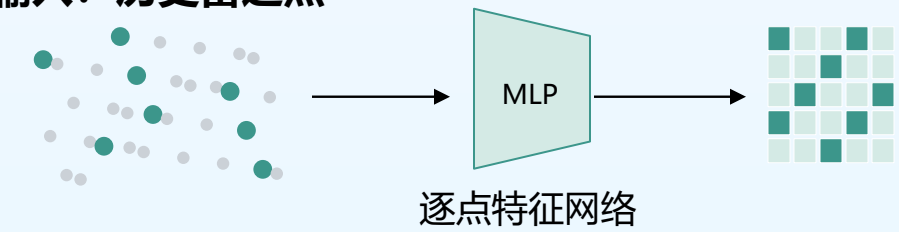
输入：光电图像

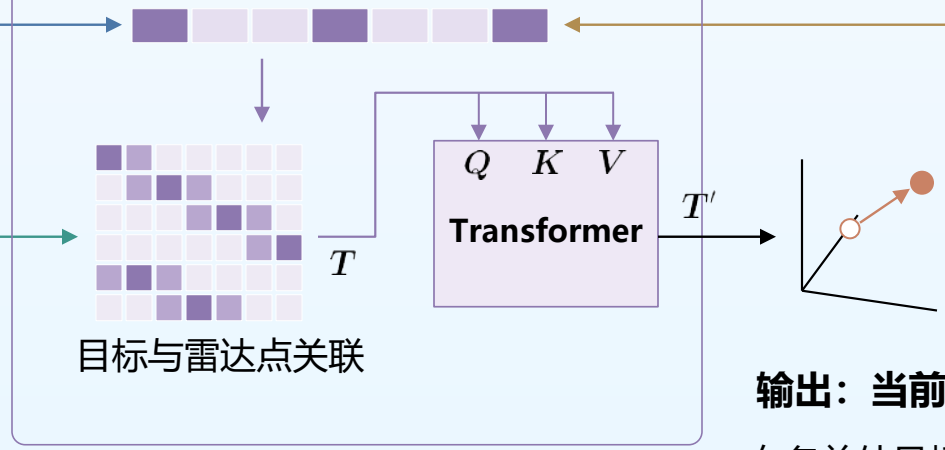


B 雷达关联与位置修正

输入：机型、位置初估及环境特征

输入：历史雷达点





输出：当前坐标

白名单外目标列表

$\hat{c}_j$，$\hat{p}_j$

C 数据增强

作用：模拟位置初估偏差，让 B 学会纠正错误初估

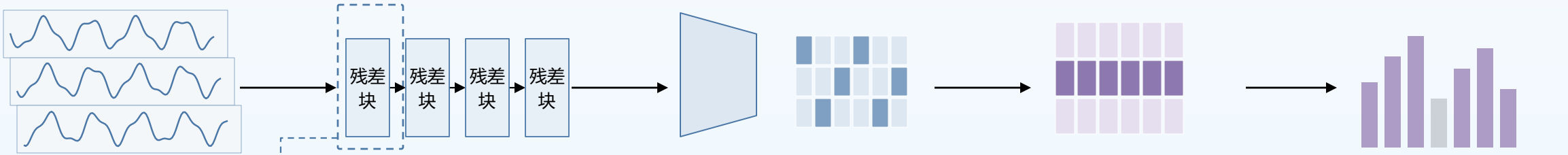
多站射频感知：特征提取与融合

从各站 I/Q 提取频谱与功率特征，再融合站点信息，完成机型判别

A

输入：多站 RF 观测

站点特征融合与机型判别

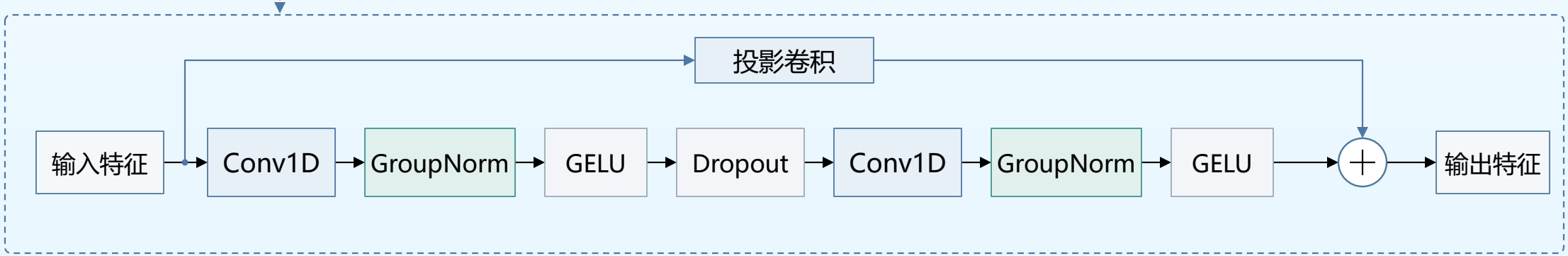


共享残差卷积　频谱池化 + 功率统计　站点标识 + 可用站点融合　分类头 / 数量头

局部放大

输出：白名单外机型与融合射频特征

残差块结构



$\boldsymbol{H}_{\mathrm{RF}} \in \mathbb{R}^{B\times K\times d_h}$

各站特征张量；结合站点标识与可用标记进行融合

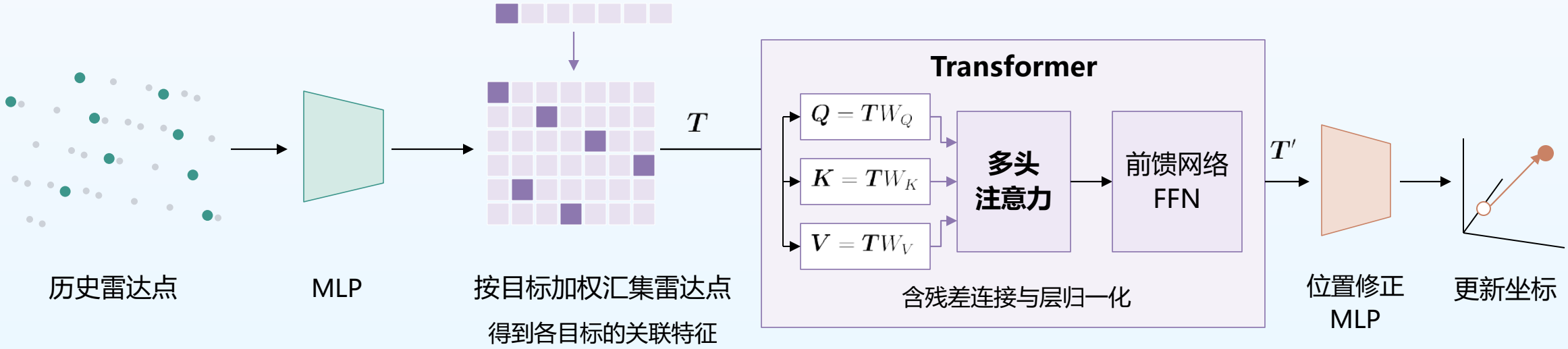
雷达点关联与位置迭代修正

按目标机型与当前估计位置关联雷达点，预测修正量，并逐轮更新坐标

B

输入：机型、位置初估、射频 / 视觉特征，以及历史雷达点

目标条件：机型 + 当前估计位置 + 射频 / 视觉特征



T, T' 交互前、后的目标特征；　Q, K, V 由同一组目标特征分别投影得到（自注意力）。

W_Q, W_K, W_V 可学习的投影矩阵。

位移尺度

本轮位置 $\boldsymbol{p}_j^{(t)} = \boldsymbol{p}_j^{(t-1)} + s_r \tanh(\boldsymbol{r}_j^{(t)})$ 网络预测残差

上一轮位置

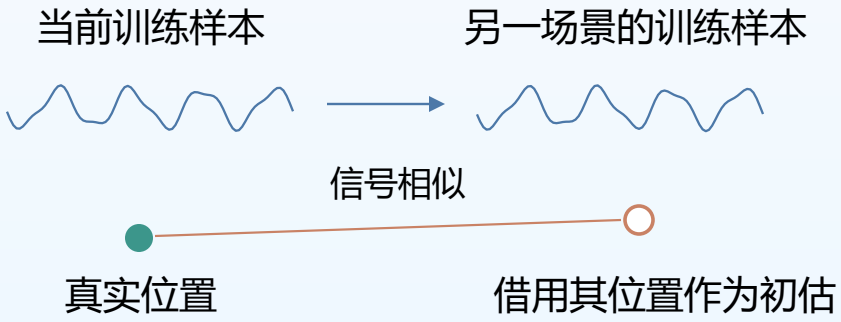
数据增强：粗定位误差模拟训练

训练时模拟粗定位偏差，让定位模型学会将错误初估修正到真实位置

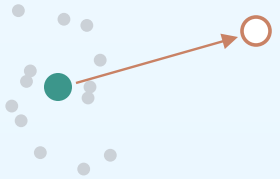
C

方式一：用信号相似样本的位置模拟误定位

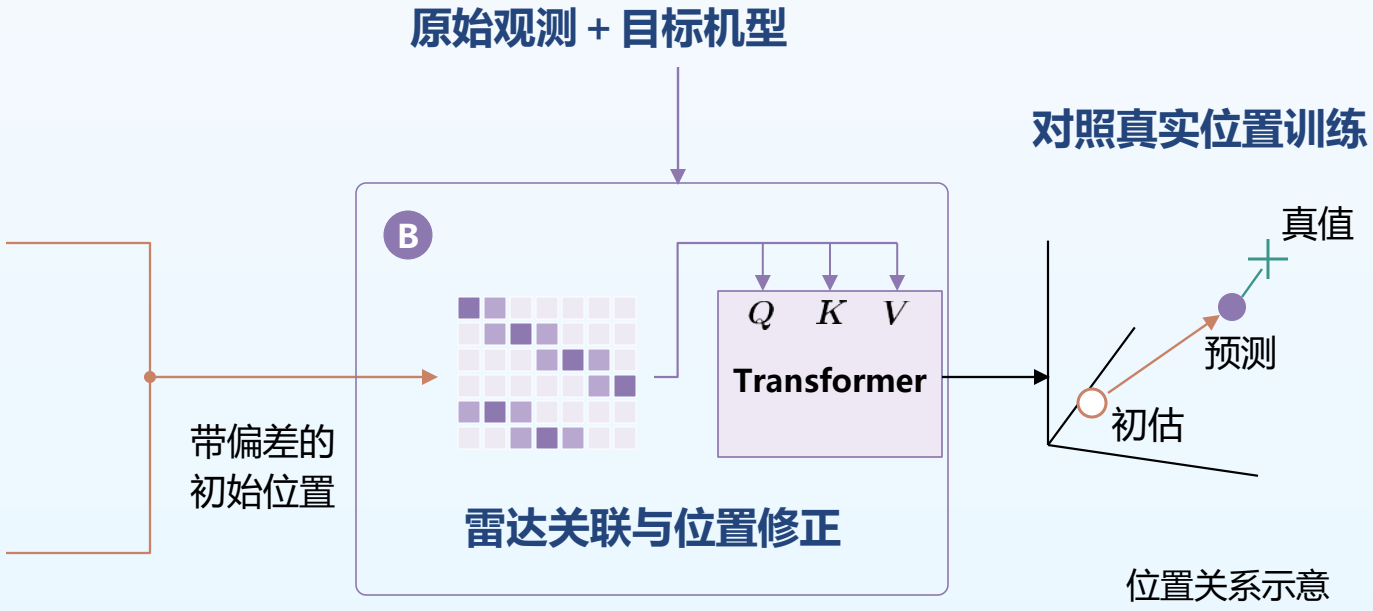
同机型、不同场景；射频特征相近，但目标位置不同



方式二：随机偏移真实位置



补充不同类型的位置偏差



位置替换： $p_{\text{init}} = p_{\text{sim}}$

相似训练样本的位置

随机偏移： $p_{\text{init}} = p_{\text{gt}} + \delta$

真实位置　随机位置偏移

目录

01 研究背景

02 系统模型与问题表述

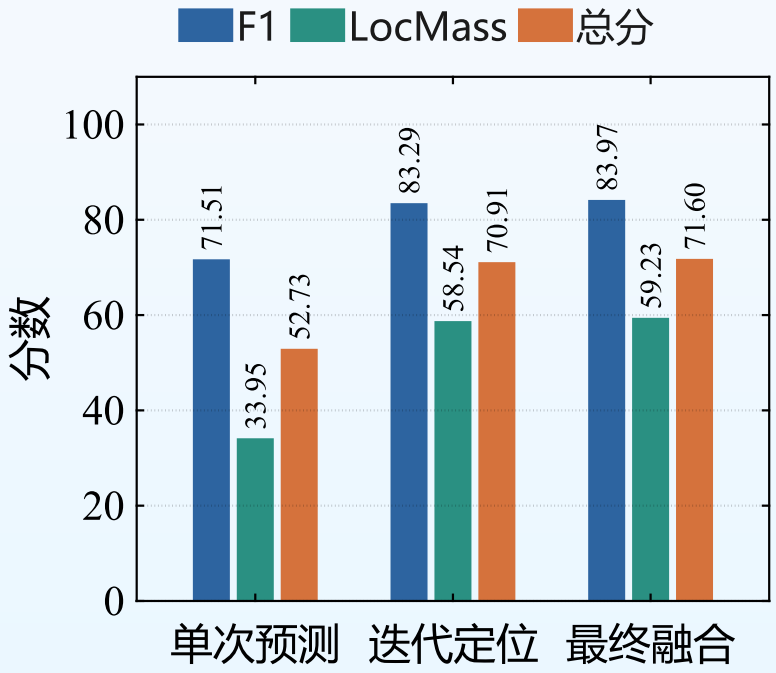
03 方案设计

04 性能比较

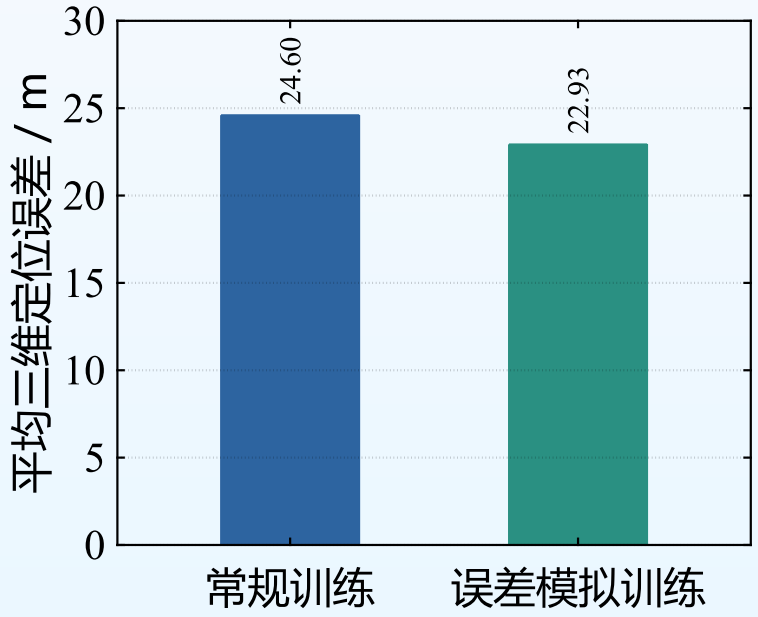
05 结论与展望

测试结果

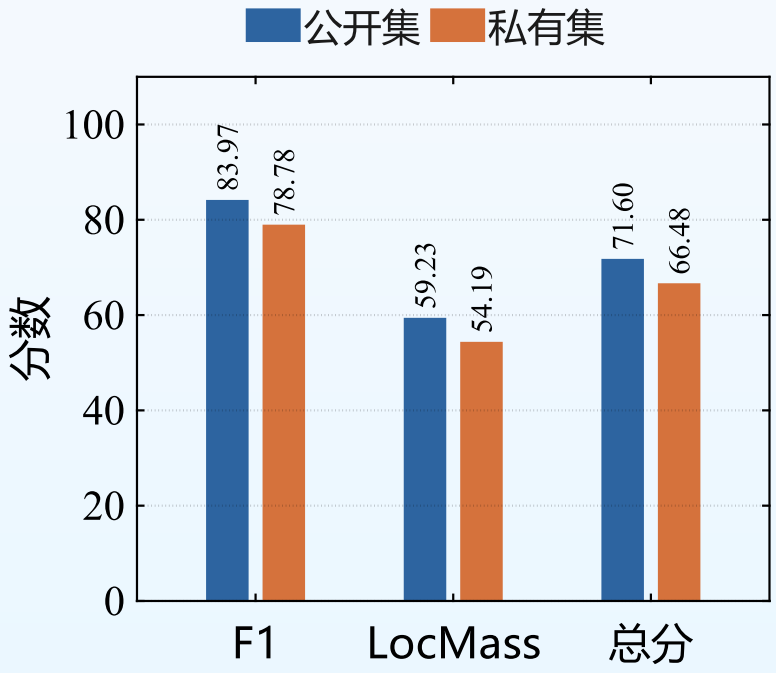
公开集方案演进



误差模拟训练消融



最终公开 / 私有集结果



单次预测：射频 I/Q 与雷达联合输入，一次前向输出机型和坐标。
迭代定位：采用位置迭代修正与粗定位误差模拟训练；三阶段为历史方案对比。
消融实验：比较是否采用误差模拟训练，平均三维定位误差降低 **1.67 m**。

目录

01 研究背景

02 系统模型与问题表述

03 方案设计

04 性能比较

05 结论与展望

研究总结与应用展望

分类效果较好，后续重点减少极端定位误差

最终测试结果	分类 F1	定位 LocMass	总分
公开测试集	83.97	59.23	71.60
私有测试集	78.78	54.19	66.48

主要结论与展望

- **总体性能**：机型识别表现较好，定位仍是主要短板，需要降低少数大误差样本的影响。
- **性能局限**：内部验证表明，定位模块误差中位数约 7.0 m，但**约 4.9% 的目标误差超过 100 m**。
- **后续改进**：面向雷达观测稀疏、观测陈旧与初始位置大偏差，增强异常识别、可信度判断和自适应修正，降低极端误定位风险。
- **应用展望**：探索无线电 / 通感组网覆盖、雷达区域定位跟踪、视觉聚焦局部目标核验的互补感知，用于军事基地周界、机场与关键设施的低空警戒。[1–3]

参考：[1] 英国交通部（2024）；[2] 英国皇家空军（2026）；[3] ETSI GR ISC 003（2026）。

致谢

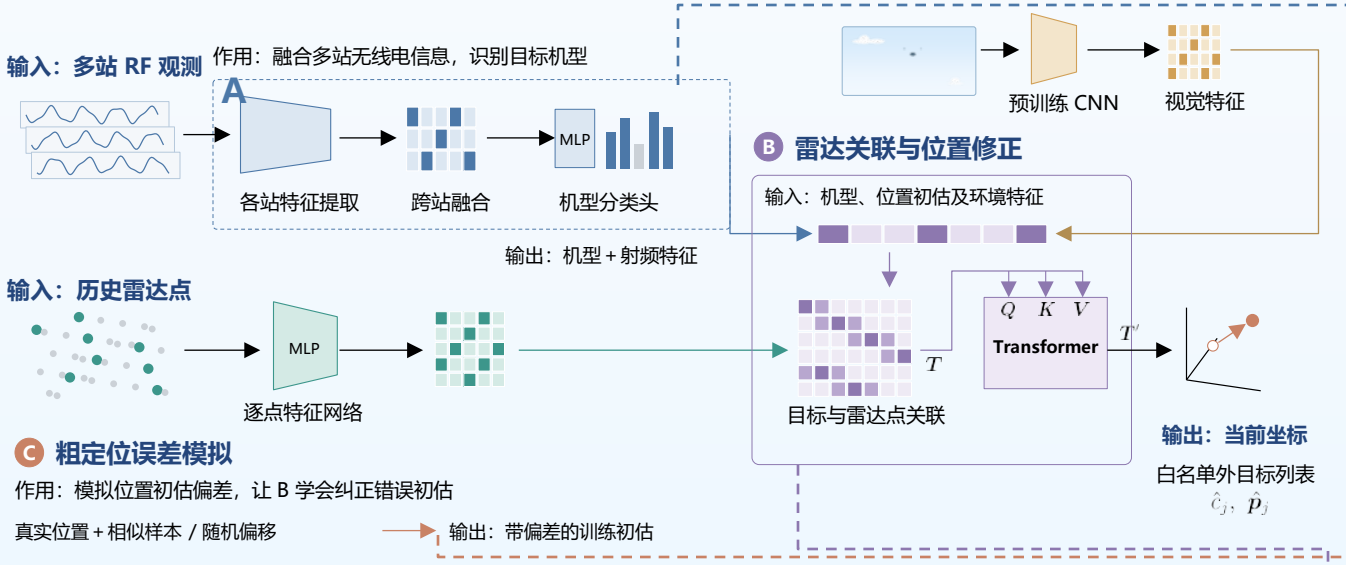
感谢各位聆听
敬请批评指正

感谢赛事主办方、评审专家及各位同仁的支持。

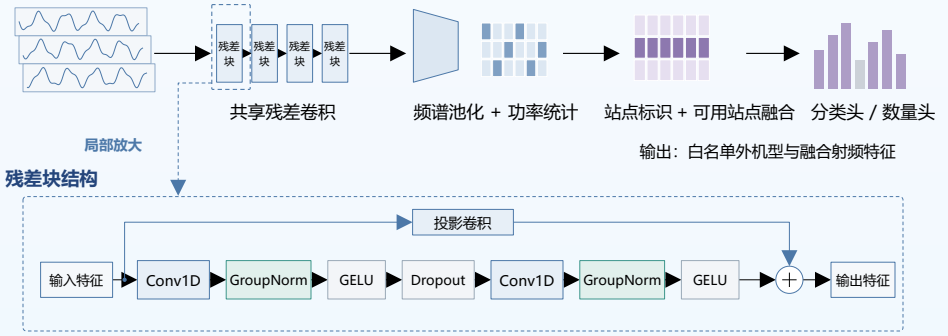
附录

附录：整体架构与模块展开

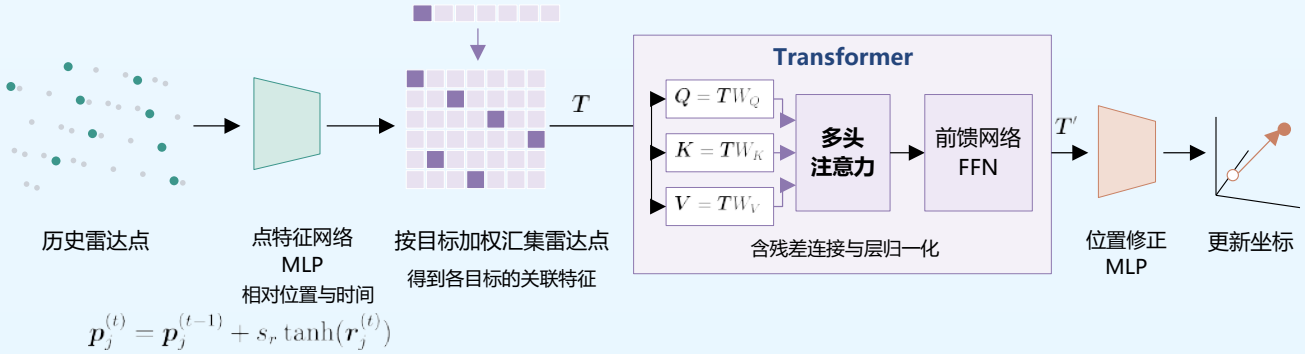
整体处理流程



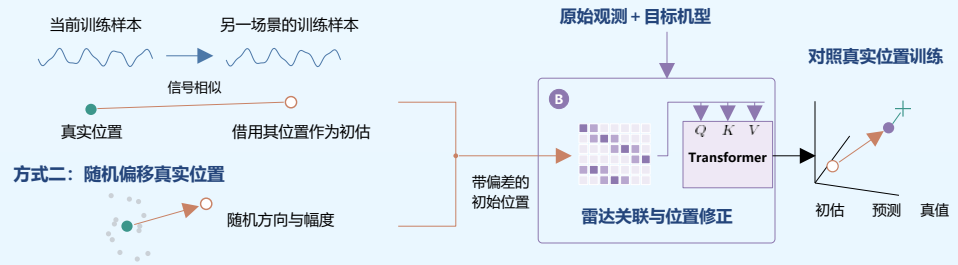
A 多站射频感知与残差结构



B 雷达点关联与位置迭代修正



C 粗定位误差模拟训练



相同标记对应同一模块；虚线引出局部结构

附录：监测平台的输入与输出

统一输出名单外机型与ENU坐标，衔接低空监测平台

观测接入

多站I/Q、雷达点列
可选图像与白名单

缺失站或图像时，
显式传入可用掩码。

按白名单过滤已知目标，
减少无效上报。

目标处理

机型判决 → 雷达关联
迭代定位 → 运动校正

利用观测时间偏移，
补偿历史点中的目标运动。

仅更新可信关联的坐标，
保持机型和目标数量。

结果输出

名单外机型集合
当前ENU三维坐标

输出统一的目标列表，
对接平台展示与后续核验。

面向重点区域，
监测名单外目标。

应用关注点：观测可缺失、名单可配置、结果可对接。

附录：位置修正网络结构与维度

融合射频、雷达与光电信息，逐轮修正位置估计

模块	结构与作用	维度符号
① **环境特征**	提取射频与光电特征，结合可用标记形成环境特征。	$\boldsymbol{z} \in \mathbb{R}^{B\times d_z}$
② **机型与位置初估**	将机型、当前估计位置及误差尺度组成各目标的条件向量。	$\mathcal{D}_q = (u_q, h_q, d_q)$
③ **点—目标关联**	提取雷达点的空间与时间特征，按目标计算权重并汇集相关点。	$\mathcal{D}_p = (u_p, h_p)$ $\mathcal{D}_a = (u_a, d_a), \quad \mathcal{D}_m = (u_m, d_m)$
④ **目标交互与输出**	融合目标条件与雷达关联结果，通过 Transformer 和修正网络更新坐标。	$\mathcal{D}_f = (u_f, h_f, d_f)$ $\mathcal{D}_r = (u_r, h_r, d_{\mathrm{pos}})$

$$\hat{\boldsymbol{p}} = \boldsymbol{p}_{\mathrm{anchor}} + s_r \tanh(\boldsymbol{r}), \qquad \boldsymbol{r} \in \mathbb{R}^{d_{\mathrm{pos}}}$$

$\boldsymbol{z}$, B, d_z 全局条件、批量大小、特征维度；映射维度按输入、隐层、输出排列。

$\boldsymbol{p}_{\mathrm{anchor}}$, s_r, $\boldsymbol{r}$ 当前估计位置、位移尺度与预测修正量；取值见末页。

附录：观测、输出与评价符号

输入观测、坐标模型和评价指标的符号定义

符号	含义
$k,\ K$	接收站索引、接收站总数
$\mathcal{X},\ X_k,\ m_k$	多站输入集合、对应站I/Q观测及可用掩码
$\mathcal{R},\ N_r$	雷达历史点集及观测点总数
$\boldsymbol{q}_\ell,\ \Delta t_\ell,\ \ell$	雷达点坐标、相对当前的时间偏移及点索引
$I,\ m_{\mathrm{EO}}$	可选光电单帧及其可用标记
$\boldsymbol{m},\ \mathcal{A}$	模态可用掩码、给定的白名单集合
$f_\theta,\ \theta$	多模态识别定位模型及模型参数
$\hat{\mathcal{Y}},\ \hat{n}$	预测目标集合、预测目标数量
$j,\ \hat{c}_j,\ \hat{\boldsymbol{p}}_j$	目标索引、预测机型及当前三维坐标
$\boldsymbol{p}_j(\Delta t)$	目标在相对时间偏移处的三维位置

符号	含义
$b,\ R_z(\varphi_b)$	标定视图索引、绕竖直轴的旋转矩阵
$\varphi_b,\ \boldsymbol{b}_b,\ \boldsymbol{\varepsilon}_\ell$	偏航角、平移偏差及观测噪声项
$s,\ P_s,\ G_s$	样本索引、上报机型集合及真实机型集合
$\mathrm{F1}_s$	该样本的机型集合F1得分
$i,\ M_s,\ d_i$	匹配对索引、同机型匹配集合及三维误差（m）
$\mathrm{LocMass}_s$	该样本的定位质量得分，漏检贡献为零
Score	最终百分制综合得分，分类与定位等权汇总
$E,\ N,\ U$	ENU坐标轴：东、北、天；位置单位为m
$t=0$	场景示意图中的当前时刻

附录：网络结构与训练符号

特征维度、映射结构、迭代条件与扰动变量

符号	含义
$B,\ C,\ V$	批量大小、机型类别数、雷达标定视图数
d_{pos}	位置向量的维度
$d_{\mathrm{sp}},\ d_{\mathrm{pw}},\ d_{\mathrm{aux}}$	频谱、功率统计、辅助特征的维度
$d_{\mathrm{RF}},\ \boldsymbol{X}_{\mathrm{RF}}$	站级总特征维度、批量多站输入张量
$\boldsymbol{H}_{\mathrm{RF}},\ E_{\mathrm{RF}}$	射频特征、各站共享的射频特征网络
$L_{\mathrm{RF}},\ d_z,\ d_I$	射频网络层数、环境特征及光电特征维度
$\boldsymbol{z},\ N$	全局条件特征、每个视图保留的雷达点数
$\mathcal{D}_q,\ \mathcal{D}_p$	目标条件向量与逐点特征网络的维度配置
$\mathcal{D}_a,\ \mathcal{D}_m$	关联映射、运动统计映射的维度配置
$\mathcal{D}_f,\ \mathcal{D}_r$	融合映射、残差预测映射的维度配置

符号	含义
$u_*,\ h_*,\ d_*$	对应映射的输入、隐层、输出宽度
$L_{\mathrm{att}},\ H_{\mathrm{att}}$	类别注意力的层数、注意力头数
$D_\theta,\ \theta$	各迭代步共享的条件去噪网络及参数
$t,\ T,\ \sigma_t$	迭代步索引、总步数、当前去噪尺度
$\boldsymbol{p}^{(0)},\ \boldsymbol{p}_{\mathrm{coarse}}$	初始候选坐标，即上游粗定位坐标
$\boldsymbol{p}^{(t)},\ \boldsymbol{p}^{(t-1)}$	当前迭代结果、上一步候选坐标
$\boldsymbol{p}_{\mathrm{anchor}},\ s_r,\ \boldsymbol{r}$	当前估计位置、位移尺度、预测修正向量
$\sigma,\ \sigma_{\min},\ \sigma_{\max}$	训练扰动标准差、采样下界及上界
$\boldsymbol{\delta}_{\mathrm{KNN}},\ R_{\mathrm{KNN}},\ \rho$	近邻方向扰动、截断半径及近邻采样概率
$\delta_z,\ \gamma_z$	扰动的高度分量、高度方向缩放系数

附录：融合与超参数符号

分类校准、位置融合及训练推理设置

符号	含义
$j,\ p^{(j)},\ J,\ p_e$	折索引、专家概率、折数及融合后的概率
$\mathcal{F}_\omega,\ \omega$	可学习的专家融合器及其参数
$\ell_{\text{old}},\ \ell_{\text{new}}$	分类校准前、校准后的logit
$\eta_\ell,\ c_\ell$	分类校准权重、logit差值截断上限
$\boldsymbol{p}_{\text{base}},\ \boldsymbol{p}_{\text{assoc}}$	基线定位坐标、雷达关联候选坐标
$\boldsymbol{\Delta},\ \alpha_{\text{pair}}$	关联候选相对位移、关联融合权重
$\Pi_R,\ R,\ R_{\text{pair}}$	欧氏球投影、球半径、关联候选的位移半径
$\boldsymbol{p}_b$	有界关联融合后的坐标
$\boldsymbol{p}^{(j)}_{\text{parent}},\ \boldsymbol{p}^{(j)}_{\text{restore}}$	该折中配对的原分支与恢复分支预测

符号	含义
$\overline{\boldsymbol{\Delta}}_{\text{rec}},\ \boldsymbol{p}_{\text{rec}}$	平均配对恢复残差、恢复后的坐标
$\boldsymbol{p}_{\text{forest}},\ \boldsymbol{p}_{\text{safe}}$	森林候选坐标、位移限幅后的坐标
R_F	森林候选对坐标的最大新增位移
$\boldsymbol{p}_{\text{corrected}},\ \boldsymbol{p}_{\text{final}}$	运动修正候选坐标、最终输出坐标
$\alpha_m,\ \tau_{\text{match}},\ N_{\text{tree}}$	运动融合权重、拒配代价阈值、回归树数
$T_{\text{RF}},\ T_{\text{obs}},\ n_{\text{IQ}}$	射频窗口、历史观测时长、I/Q采样点数
$s,\ E_s,\ E_{\text{total}}$	阶段索引、各阶段的轮数及总训练轮数
$\mathcal{L},\ \mathcal{L}_t,\ w_t$	总损失、该步监督损失及其权重
KNN	射频特征空间中的近邻采样

参数设置汇总

输入配置、网络规模与训练推理设置

输入与网络结构

输入与输出	$K = 4, \quad C = 8, \quad V = 2, \quad d_{\mathrm{pos}} = 3$
观测窗口	$T_{\mathrm{RF}} = 10\,\mathrm{ms}, \quad T_{\mathrm{obs}} \approx 3\,\mathrm{s}, \quad n_{\mathrm{IQ}} = 1.25 \times 10^{6}$
射频特征	$(d_{\mathrm{sp}}, d_{\mathrm{pw}}, d_{\mathrm{aux}}) = (512, 32, 10), \quad d_{\mathrm{RF}} = 554$
射频与光电	$L_{\mathrm{RF}} = 4, \quad d_z = 128, \quad d_I = 2560$
雷达点数	$N = 64$
目标条件映射	$\mathcal{D}_q = (164, 192, 128)$
点列与关联	$\mathcal{D}_p = (11, 128), \quad \mathcal{D}_a = (384, 128), \quad \mathcal{D}_m = (26, 128)$
融合与残差	$\mathcal{D}_f = (512, 256, 128), \quad \mathcal{D}_r = (128, 128, 3)$
类别注意力	$L_{\mathrm{att}} = 1, \quad H_{\mathrm{att}} = 4$
跨折融合	$J = 5$

训练与推理超参数

分类校准	$\eta_\ell = 0.5, \quad c_\ell = 2$
去噪迭代	$T = 2, \quad (\sigma_1, \sigma_2) = (10, 5)\,\mathrm{m}, \quad s_r = 160\,\mathrm{m}$
误差样本混合	$\rho = 0.25, \quad 1 - \rho = 0.75$
扰动范围	$\sigma_{\min} = 5\,\mathrm{m}, \quad \sigma_{\max} = 55\,\mathrm{m}, \quad R_{\mathrm{KNN}} = 40\,\mathrm{m}$
高度扰动	$\gamma_z = 0.55$
训练与监督	$(E_s) = (20, 20), \quad E_{\mathrm{total}} = 40, \quad (w_t) = (0.35, 0.65)$
关联融合	$R_{\mathrm{pair}} = 25\,\mathrm{m}, \quad \alpha_{\mathrm{pair}} = 0.75$
森林限幅	$R_F = 2\,\mathrm{m}$
轨迹回归	$N_{\mathrm{tree}} = 160, \quad \tau_{\mathrm{match}} = 40$
运动融合	$\alpha_m = 0.5$