

# 高质量演训数据集的构建与评估研究

霍明亮, 黄永亮, 赵程亮, 林李杰, 郑向通

中国人民解放军 63611 部队, 新疆 库尔勒 841000

## 摘要

实战化演训是提升体系作战能力、检验装备效能、优化指挥决策的核心手段, 而演训数据则是支撑复盘评估、战法验证、能力认证与指挥决策的关键数据要素。针对当前演训活动中存在的数据采集零散、标准不统一、多源异构难以融合、全生命周期治理机制缺失、质量评估体系不健全、数据安全与复用水平薄弱等现实问题, 构建了高质量演训数据集全生命周期理论框架, 系统研究了演训数据的需求建模、标准体系、数据预处理和质量管控、精准标注等关键构建技术, 提出了“静态基础质量+动态作战效用”的双层评估模型, 构建了演训数据集评估指标体系, 并结合联合演训数据开展了验证分析。结果表明, 该方法在完整性、准确性、时空一致性、标注一致性、复现率五个维度均实现了显著超越, 且具备更强的自适应能力与鲁棒性, 对联合演训数据质量评估具有有效性和可行性, 可为实战化演训数字化转型、数据要素赋能体系作战提供理论支撑与技术路径。

## 关键词

实战化演训; 高质量数据集; 构建技术; 评估模型; 数字化转型

中图分类号: TP391.9

文献标志码: A

doi:10.11959/j.issn.2096-0271.2026xxx

## Research on the construction and evaluation of high-quality exercise training datasets

Huo Mingliang, Huang Yongliang, Zhao Chengliang, Lin Lijie, Zheng Xiangtong

Unit 63611, PLA, Korla Xinjiang 841000, China

## Abstract

Combat-oriented exercise training serves as a core enabler for enhancing system-of-systems operational capabilities, validating equipment effectiveness, and optimizing command decision-making. As a fundamental data element, exercise training data underpins after-action review, doctrine validation, capability certification, and command decision-making. To address prevailing challenges in current exercise training activities, including fragmented data collection, inconsistent standards, difficulties in fusing multi-source heterogeneous data, the absence of a full-lifecycle governance mechanism, an incomplete quality evaluation system, and inadequate levels of data security and reuse, a full-lifecycle theoretical framework for high-quality exercise training datasets is established. It systematically investigates key construction technologies, such as requirement modeling, standard system formulation, data preprocessing and quality control, and precise annotation. A two-layer evaluation model integrating “static foundational quality” and “dynamic

operational effectiveness” is proposed, along with a corresponding evaluation indicator system for exercise training datasets. Validation is conducted using data from joint exercise training. The results demonstrate that the proposed method achieves significant improvements across five critical dimensions: completeness, accuracy, spatiotemporal consistency, annotation consistency, and reproducibility, while exhibiting stronger adaptability and robustness. It is proven effective and feasible for assessing the quality of joint exercise training data, thereby offering theoretical support and technical pathways for the digital transformation of combat-oriented exercise training and the empowerment of system-of-systems operations through data elements.

### *Key words*

combat-oriented exercise training, high-quality dataset, construction technology, evaluation model, digital transformation, system-of-systems operations

## 1 引言

现代战争呈现信息化、智能化、体系化、联合化特征，实战化演训已从单一科目训练转向联合对抗、跨域协同、全域覆盖的高强度演训模式。演训数据承载兵力行动、装备状态、指挥指令、战场环境、毁伤效果、协同关系等核心信息，是复盘评估、战法推演、能力认证、决策优化的核心依据。

当前演训数据建设面临突出问题<sup>[1]</sup>：一是多军兵种、多平台、多系统数据异构，格式与标准不统一，难以融合共享；二是原始数据存在缺失、异常、时序错位、冗余失真等缺陷，直接影响复盘结论可信度；三是数据治理权责不清，缺乏从采集到销毁的全流程管控机制，安全保密与合规风险突出；四是质量评估停留在完整性、准确性等浅层指标，未与作战任务、战法效能深度绑定；五是数据孤岛严重，跨演训场、跨军兵种、跨任务复用能力弱。在此背景下，开展高质量演训数据集的构建、治理与评估研究，对提升演训质效、缩短复盘周期、优化作战方案、提升体系作战能力具有重要理论价值与军事应用价值。

## 2 高质量演训数据集概述

### 2.1 基本内涵

演训数据集<sup>[2]</sup>指的是为满足实战化演训、复盘评估、战法验证、能力认证、指挥决策需求，按照军用标准与演训大纲规范，经多源采集、预处理、标注、质控、融合形成的，可支撑演训分析、模型验证、决策优化的结构化、时序型、空间地理型、视频图像型、多模态数据集合。

所谓高质量演训数据集，即满足标准合规、原始保真、完整无缺、准确一致、时空对齐、代表性强、安全可控、可追溯、可复现、可作战复用等要求的演训数据集。

### 2.2 演训数据集核心特征

演训数据集作为支撑实战化演训、战法验证与指挥决策的核心数据底座，其核心特征<sup>[3-6]</sup>从数据来源、时空属性、业务特性、安全要求等维度形成了鲜明的技术与应用标签，主要包括以下五方面。

(1) 多源异构性：全要素数据的融合基础

演训数据集的构建依托多维度、多类型、多载体的数据源体系，打破了单一数据类型的局限性。其覆盖范围包含指挥终端的指令流数据、武器平台的运行参数与状态数据、侦察感知系统的情报数据、模拟仿真生成的虚拟推演数据、北斗定位系统的时空基准数据、视频监控的可视化图像数据，以及毁伤评估环节的结果判定数据等。这些数据在格式、结构、传输协议上存在显著差异，既包含结构化的数值型数据，也涵盖非结构化的视频、图像、文本数据，形成了异构化的数据生态。这种多源异构特性决定了演训数据集构建必须以标准化适配为前提，通过统一的接口与格式转换，实现不同来源数据的有效汇聚与融合，是保障数据完整无缺、代表性强的核心基础，也为后续“可作战复用”提供了数据广度支撑。

(2) 强时空关联性：战场态势复现的核心前提

演训数据的核心价值在于还原真实的战场时空场景，因此强时空关联性是其不可或缺的关键特征。所有演训数据必须具备统一的时间戳基准与标准化的空间坐标体系，通过高精度的授时系统与地理配准技术，确保不同来源、不同类型的数据在时间维度上严格同步、在空间维度上精准对应。这一特征支撑了战场态势的完整复现，能够精准还原兵力行动的时序脉络、装备部署的空间位置、事件发生的先后逻辑，让演训数据不仅是单纯的数值集合，更是能够还原作战全过程的“时空账本”。这直接对应高质量内涵中的时空对齐、可复现特征，无论是复盘评估时的态势回溯，还是战法验证时的场景还原，都依赖于强时空关联性实现数据的有效关联与解读。

(3) 高实时高并发：演训场景的适配特性

演训过程具有动态性、突发性、连续性的特点，数据产生呈现出高实时、高并发的显著特征。演训现场的传感器、指挥系统、参演平台会持续产生海量数据，数据传输速率高、瞬时数据量极大，对数据采集的响应速度、处理的并行能力与存储的吞吐效率提出了极高要求。这一特征决定了演训数据集的构建必须适配高并发的数据生产环境，需搭建高可用的采集与处理架构，通过实时缓存、分布式处理、流式计算等技术手段，保障海量数据的高效采集、及时处理与完整留存。这是保障数据原始保真、时效性的关键，若无法满足高实时高并发的要求，极易出现数据丢失、时序错位、处理延迟等问题，直接影响演训数据的完整性与时效性，进而削弱数据对实战化演训的支撑能力。

(4) 强作战关联性：数据应用的价值导向

演训数据集并非无的放矢的“数据堆砌”，而是以作战任务为核心、以实战需求为导向的专业化数据集合，强作战关联性贯穿全流程。其构建目标直接服务于战法验证、能力评估、指挥决策、装备效能分析等核心作战场景，数据选取、采集、处理与标注均围绕具体作战任务展开。数据内容紧密贴合实战化演训的核心需求，涵盖作战方案落地、兵力部署执行、装备参数验证、指挥指令响应等关键环节，每一类数据都对应明确的作战应用场景，任务导向清晰、应用价值明确。这种强作战关联性确保了演训数据能够精准支撑作战业务需求，是实现准确一致、可作战复用的核心保障，避免了“为数据而数据”的无效建设，最大化发挥数据对实战化训练的支撑作用。

(5) 高安全密级：数据管控的底线要求

演训数据涉及作战方案、兵力部署、装备核心参数、指挥通信机密等敏感内容，属于国家军事机密范畴，因此具备极高的安全密级，是演训数据管理的刚性约束。不同类型的演训数据根据涉密等级划分，需严格执行分级保密管控要求，从数据采集、传输、存储、处理到销毁的全生命周期，均需遵循保密法规与军队保密制度。在数据管理环节，需通过加密存储、权限分级、访问控制、安全审计等技术与管理手段，严防数据泄露、篡改、窃取与非法访问；在数据应用环节，需严格限定数据使用范围与人员权限，确保演训数据在合规范围内流转与使用。这直接对应高质量内涵中的安全可控、可追溯特征，是演训数据集构建与管理必须坚守的底线，也是保障军事数据安全与国家军事利益的核心要求。

总的来看，多源异构性与强时空关联性，是实现完整无缺、时空对齐、可复现的基础；高实时高并发与强作战关联性，是保障原始保真、准确一致、可作战复用的关键；高安全密级，则是落实安全可控、可追溯、标准合规的刚性保障。

### 2.3 国内外研究现状

美军依托 Joint-Training-and-Training-Technology (J-T3)、OneTESS、Advanced Distributed Learning (ADL) 等系统构建了覆盖全流程的联合演训数据支撑体系，实现了演训数据从采集、标准化治理到全域共享与闭环评估的全链路管理<sup>[7-10]</sup>。在数据采集层面，美军通过多源异构传感器、指挥信息系统和仿真平台实现了对战场态势、指挥决策、作战行动等多维度数据的自动化采集，强调数据的可追溯性与可复现性；在

数据治理层面，建立了统一的数据模型、元数据规范和质量控制流程，通过数据清洗、标注、融合等手段提升数据可用性，支撑多军兵种联合作战演训；在评估应用层面，形成了“数据驱动-效能评估-方案优化”的闭环机制，可对作战方案、指挥效能、装备性能等进行量化分析，为体系作战能力提升提供支撑。此外，NATO、美国防部先后发布《联合训练数据标准》《演训数据质量规范》等多项政策文件，明确了演训数据的分类、格式、接口和安全要求，推动演训数据在军种间、盟国间的高效流转与复用，构建了“采集-治理-评估-应用”的成熟生态。

国内围绕演训数据采集、复盘分析、效能评估等方向开展了大量研究与实践<sup>[11]</sup>，但与美军相比仍存在明显短板。在数据标准层面，缺乏面向实战化演训的专用数据集构建标准与流程，不同演训场景、军兵种间的数据格式、采集规范不统一，导致数据难以跨域共享与复用；在治理体系层面，现有治理机制多聚焦于演训后的数据处理，未覆盖“筹划-实施-复盘-改进”的全链条，无法实现对演训全生命周期数据的动态管控；在评估应用层面，评估指标体系与作战任务耦合度不足，多侧重技术指标而非作战效能，实用性与针对性不强，难以有效支撑作战方案优化；在平台支撑层面，工程化演训数据处理平台建设相对薄弱，智能化标注、异常检测、关联分析等能力不足，无法满足大规模、高复杂度演训数据的快速处理需求。尽管国内部分单位在演训数据可视化分析、AI辅助评估等领域取得了阶段性成果，但整体仍处于“局部突破、体系不足”的阶段，与美军成熟的演训数据支撑体系存在较大差距。

## 3 高质量演训数据集构建技术体系

### 3.1 总体流程构建

高质量演训数据集遵循八步闭环构建流程，如图1所示。

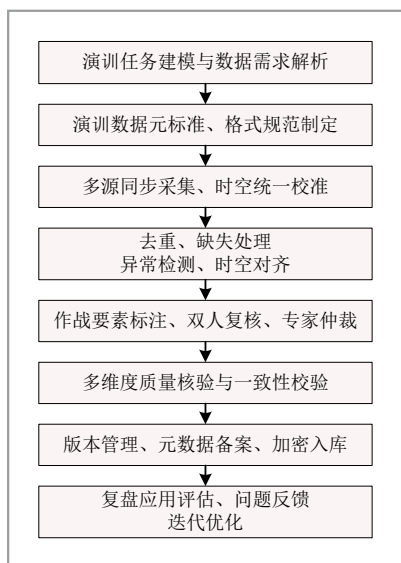


图1 高质量演训数据集全生命周期构建流程

### 3.2 需求建模与标准体系

#### 3.2.1 需求建模

需求建模是演训数据集构建的起点与核心依据，旨在通过多维度拆解演训任务与应用目标，形成精准、可落地的数据需求框架，为后续数据采集、治理与评估提供明确指引。本文从演训类型、演训科目、评估指标、复盘要求、输出形式五个维度开展需求建模，具体建模逻辑如下：

(1) 演训类型维度：针对联合演训、专项对抗演训、装备验证演训、指挥决策推演等不同类型，明确数据采集的粒度与范围。例如，联合演训需覆盖多军兵种跨域协同数据，而装备验证演训则重点聚焦装备性能参数与作战效能数据，确保数据需求与演训目标高度匹配。

(2) 演训科目维度：围绕侦察预警、指挥控制、火力打击、综合保障等典型科目，拆解各科目所需的核心数据要素。如火力打击科目需采集目标识别信息、射击诸元、毁伤评估数据等，保障数据能够完整支撑作战行动的全流程复盘与分析。

(3) 评估指标维度：结合作战效能评估、指挥效能评估、训练水平评估等需求，提取对应的量化指标数据需求。例如，评估指挥效能时，需采集决策响应时间、指令下达准确率、态势感知完整性等指标相关的原始数据，为效能量化计算提供支撑。

(4) 复盘要求维度：根据演训后复盘分析的深度与场景，明确数据的可追溯性、可复现性要求。如针对复杂对抗场景的复盘，需采集完整的时间序列数据、交互日志与环境数据，支持事后复现作战过程、定位问题节点，优化战术战法。

(5) 输出形式维度：结合后续数据应用场景（如模型训练、可视化分析、作战方案推演），明确数据的输出格式、接口规范与存储要求。例如，面向 AI 模型训练的需求，需输出结构化标注数据集；面向指挥辅助决策的需求，需提供可直接接入指挥信息系统的实时数据接口。

通过5个维度的需求建模，可将抽象的演训任务转化为具象化、可执行的数据需求清单，从而进一步明确数据采集范围、质量要求、应用场景等关键信息，避免数据采集的盲目性与冗余性，为后续数据集构建与治理提供精准指引。

#### 3.2.2 标准体系构建

针对上述需求，构建覆盖基础通用、数据元标准、采集处理、标注规范、质量控制、安全共享六类标准体系，如表1所示。

表1 演训数据集标准体系

标准类型	核心内容	适用范围
基础通用	术语定义、分类编码、元数据框架	全流程
数据元标准	字段名称、单位、精度、值域、枚举	结构化数据
采集处理	采样率、时统、定位、格式转换	原始数据
标注规范	作战要素标签、流程、一致性要求	复盘与验证集
质量控制	缺陷分级、抽检规则、返工机制	全流程质控
安全共享	分级分类、脱敏、水印、审计	存储与使用

### 3.3 数据预处理与质量管控

#### 3.3.1 数据清洗与质量管控数学模型构建

围绕演训数据的完整性评估与多维度异常检测，构建量化评价与识别的数学模型，为数据清洗与质量管控提供理论支撑。

##### (1) 缺失率与完整性指标

数据完整性是保障演训分析评估的基础<sup>[12]</sup>，通过构建缺失率（MR）与完整性指标（CI）数学模型，实现对演训数据完整性的量化评估：

$$MR = \frac{N_{null}}{N_{total}}, CI = 1 - MR \quad (1)$$

其中，Nnull为缺失值数量，Ntotal为总记录数；MR反映数据缺失程度，CI表征数据完整度。

该数学模型主要用于评估单字段或全数据集缺失情况，指导均值填充、时序插值等补全策略，或删除高缺失样本，确保满足后续建模的最低完整性要求。

##### (2) 时序异常检测

针对演训数据的时间序列特性，采用3σ原则识别时序异常情况<sup>[13]</sup>，即：

$$\Delta t = \left| t_i - \bar{t} \right| > 3\sigma_t \quad (2)$$

其中， $t_i$ 为第*i*条记录时间戳， $\bar{t}$ 为时间序列均值， $\sigma_t$ 为标准差；当时间偏差 $\Delta t > 3\sigma_t$

时，判定为异常。

该数学模型主要基于正态分布统计特性，主要用于定位时间同步或数据传输问题。

##### (3) 空间异常检测

针对演训数据的空间位置信息，采用欧氏距离阈值法识别空间异常情况<sup>[14]</sup>，即：

$$\Delta s = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2 + (z_i - z_0)^2} > S_{max} \quad (3)$$

其中， $(x_i, y_i, z_i)$ 为当前坐标， $(x_0, y_0, z_0)$ 为参考坐标， $\Delta s$ 为欧氏距离， $S_{max}$ 为预设最大允许距离阈值。

该数学模型主要用于当目标/装备位置与参考点距离超出阈值时，判定为空间异常（如坐标漂移、越界机动），进而保障演训空间态势的合理性。

##### (4) 联合异常判定

为提升异常检测鲁棒性，融合时序、空间、数值三类特征进行联合判定，即：

$$Anomaly(x_i) = Time(x_i) \vee Space(x_i) \vee Value(x_i) \quad (4)$$

其中， $Anomaly(x_i)$ 为异常判定结果， $Time(x_i)$ 、 $Space(x_i)$ 、 $Value(x_i)$ 分别为时序、空间、数值异常标记， $\vee$ 为逻辑或运算。

该数学模型主要用于多维度融合时，

避免单一维度检测的漏检/误检，进而实现全面异常筛查，为演训数据清洗（删除/修正异常样本）提供精准依据。

### 3.3.2 数据预处理流程

数据预处理流程共包括8个步骤，如图2所示。

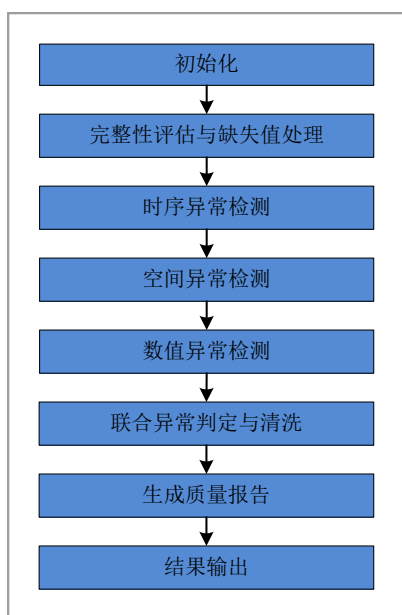


图2 数据预处理流程

(1) 第一步时初始化阶段，主要是创建空的清洗后数据集，并统计原始数据总记录数；

(2) 第二步是完整性评估与缺失值处理，主要是遍历每条记录，统计缺失值数量，超过阈值则丢弃，否则用均值/插值填充；

(3) 第三步是时序异常检测，主要是基于时间戳的 $3\sigma$ 原则，识别偏离时间均值过远的异常记录；

(4) 第四步是空间异常检测，主要是计算空间坐标与参考点的欧氏距离，超出

$S_{max}$  阈值则标记为空间异常；

(5) 第五步是数值异常检测，主要是对速度和海拔高度字段应用 $3\sigma$ 原则，识别数值偏离均值过大的异常；

(6) 第六步是联合异常判定与清洗，只要记录存在时序/空间/数值任一异常，就标记为总异常并丢弃；无异常则加入清洗后数据集；

(7) 第七步是质量报告生成，主要是计算缺失率和完整性指标，输出数据质量评估结果；

(8) 第八步是结果输出，主要是返回清洗后的数据集和完整的质量报告。

## 3.4 作战要素精准标注

作战要素精标准注是保障演训数据集可用性的核心环节，旨在通过规范化标注流程与质量控制机制，实现对战场所态势、作战行动、装备状态等关键要素的精准标记，为后续模型训练、效能评估与复盘分析提供高质量标注数据。

### 3.4.1 作战要素标注数学模型构建

为量化评估标注结果的可靠性，采用标注一致率（Agreement Rate, Agr）作为核心评价指标，其数学模型如下：

$$Agr = \frac{N_{agree}}{N_{total}} \times 100\% \quad (5)$$

其中， $N_{agree}$  表示双标注结果中一致的样本数量， $N_{total}$  表示参与标注的总样本数量， $Agr$  表示标注一致率，取值范围为 0~100%。

该数学模型存在一定的质量约束，即：明确双人标注一致率需 $\geq 85\%$ ，若低于该阈值则进入专家仲裁环节，避免因标注偏差影响后续数据应用效果，保障标注结果的

准确性与一致性。

### 3.4.2 标注质量控制流程

为实现标注一致性的自动化校验与结果闭环管理，设计一套标注质量控制算法，具体流程如图3所示：

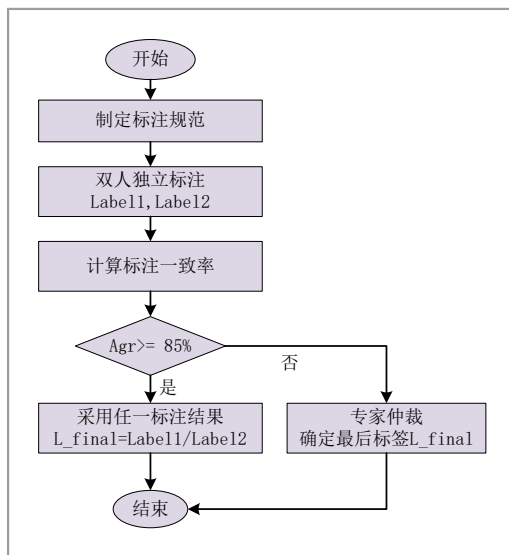


图3 标注质量控制算法流程

(1) 第一步是先制定统一的《演训数据 XX 要素标注规范》，确保标注口径一致；

(2) 第二步是双人标注，每份样本由两名标注人员独立完成标注，生成 Label1 和 Label2；

(3) 第三步是进行一致率计算，通过逐样本比对双标注结果，计算标注一致率 Agr；

(4) 第四步是质量判定，若  $Agr \geq 85\%$ ：直接采用任一标注结果作为最终标签  $L_{final}$ ；若  $Agr < 85\%$ ，则提交领域专家进行仲裁，确定最终标签；

(5) 最后是流程闭环，完成质量校验后，输出最终标注结果。

该算法主要是通过逐样本比对两名标注人员的结果，自动计算一致率；若一致率满足阈值要求，则直接采用标注结果；若不满足，则引入专家仲裁机制，修正标注偏差，保障最终标签的权威性与准确性。该算法实现了标注质量的自动化管控，既提升了标注效率，又避免了人工主观判断的误差，为构建高可信度演训数据集提供支撑。

## 3.5 典型演训数据集构建

以某次联合对抗演训为例，运用上述方法构建作战要素数据集，如表2所示。

## 4 高质量演训数据集评估模型

### 4.1 双层评估框架

采用静态基础质量+动态作战效用构建双层评估模型，实现从数据校验到版本迭

表2 演训作战要素与数据类型

作战要素	数据类型	采集方式	标注内容
兵力行动	时序定位	北斗/定位终端	位置、速度、航向、行动类型
装备状态	结构化参数	装备终端	工况、能耗、通信、故障
指挥指令	文本/指令流	指挥系统	指令内容、发送时间、执行结果
战场环境	栅格/矢量	环境模拟	气象、地形、电磁、水文
毁伤效果	图像/评估值	侦察/评估	目标类型、毁伤等级、命中概率

代的闭环管理，如图4所示。

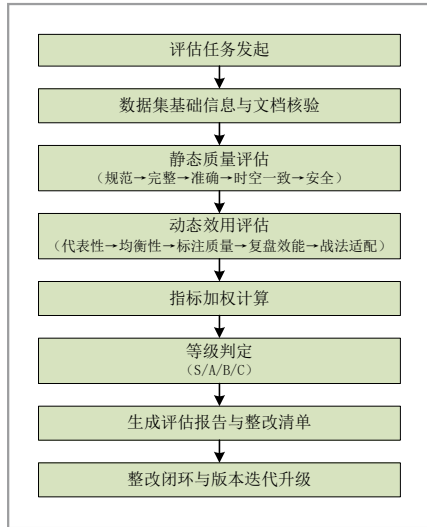


图4 演训数据集双层评估实施流程

评估实施流程起始于评估任务发起，随后进入数据基础信息与文档校验环节，重点核查数据集元数据、标注规范文档、版本说明等基础资料的完整性和合规性，为后续评估奠定基础。核验通过后，依次开展静态质量评估和动态效用评估。其中，静态质量评估遵循“规范→完整→准确→时空一致→安全”的递进逻辑，从数据格式合规性、字段缺失率、真值偏差度、时空坐标匹配度及数据加密防护等维度，量化校验数据集底层质量水平；动态效用评估则围绕“代表性→均衡性→标注质量→复盘效能→战法适配”展开，验证数据样本对真实战场场景的覆盖度、类别分布均衡性、标注精准度，以及支撑作战复盘、战法验证等实战任务的效能表现。完成两类评估后，通过指标加权计算整合各维度得分，依据阈值完成S/A/B/C等级判定。最终生成评估报告与整改清单，明确问题项与优化方向，推动整改闭环与版本迭代

升级，实现演训数据集的持续优化与实战适配能力提升。

## 4.2 评估指标体系

根据演训数据集评估模型，围绕“文档合规性、数据质量、作战效用”构建演训数据集评估指标体系，包括2个一级指标、10个二级指标，并根据需要，可继续向下拓展，直至可量化为止。根据每个指标的重要程度对其赋权，如表3所示。

表3 演训数据集评估指标体系

序号	一级指标	二级指标	权重	量化方式
1		规范性	0.12	标准符合度
2		完整性	0.16	1-缺失率
3	数据质量	准确性	0.16	真值符合率
4		时空一致性	0.08	对齐合格率
5		安全性	0.08	安全合规项占比
6		代表性	0.08	场景覆盖度
7		均衡性	0.08	类别均衡指数
8	作战效用	标注质量	0.08	一致率+准确率
9		复盘效能	0.08	复盘效率提升
10		战法适配度	0.08	战法复现成功率

## 4.3 综合评估计算

根据上述评估指标及相应的权重，得到综合得分S，即：

$$S = 0.6 \times S_{static} + 0.4 \times S_{dynamic} \quad (6)$$

其中，综合得分按百分制计算，采取S级、A级、B级、C级四级制评定，百分制与四级制的换算关系为：90分（含）以上为S级，表明该数据集为高质量，可直接用于战法验证与指挥决策；80分（含）-89分为A级，表明该数据集质量为合格，可用于演训的复盘评估；70分（含）-79分为

B级，表明该数据集质量一般，限制使用，需要整改；不足70分为C级，表明该数据集质量不合格，禁止使用。

## 5 验证分析

### 5.1 验证方法

为深度验证本文所提方法在联合对抗演训数据质量评估中的有效性与优越性，选取某联合对抗演训数据集作为核心研究样本。该数据集聚焦实战化联合演训场景，涵盖兵力行动、装备状态、指挥指令、战场环境四大类战场动态数据，共计16800条结构化样本。数据覆盖演训全周期关键节点，真实反映了复杂电磁环境、强对抗态势下联合作战的数据特征，具备极强的实战代表性。本文采用传统方法和本文设计的方法进行对比分析，以确保结论的客观性与严谨性。

传统方法是基于规则引擎与阈值判定的经典评估方法，该方法依赖人工预设的静态规则（如固定格式校验、单一数值阈值），通过逻辑匹配完成数据质量判定，包括规则配置、批量校验、结果统计等阶段，是目前演训数据评估中应用广泛的基础方法，但其对复杂战场数据的动态适应性与噪声鲁棒性较弱，仅能识别规则覆盖的显性问题，难以挖掘隐含的时序关联与语义冲突。

本文方法是一种融合多模态特征提取、时序依赖建模与深度学习的智能评估方法，通过嵌入作战条令与演训业务规则的知识图谱，实现对兵力行动、指挥指令等数据的语义理解；利用时序卷积网络（TCN）与注意力机制，挖掘战场数据的时空关联规律，可自适应识别缺失、冲突、噪声等多类型质量问题，适配联合作战数据的动态特性。核心实现步骤包括数据预处理、构建三类特征集（基础统计特征、时序特征、语义特征）、模型训练与推理、指标计算等内容。

为验证两种方法评估结果差异的显著性，本文对结果进行了统计检验与误差分析，采用配对t检验（Paired t-test）对两种方法在五大维度上的得分差异进行分析，结果显示p值均小于0.05，说明本文方法与传统方法的评估效果差异具有统计学意义，并非随机误差导致。

为直观展示数据集形态，采用分层随机抽样策略，按照“数据类型+时间阶段”选取部分演训数据片段如表4所示。

从样本片段可见，数据集包含四类核心要素：

- （1）兵力行动类：记录作战单元的位置、机动方向与行为状态，用于刻画战场态势变化；
- （2）装备状态类：反映武器装备的运行工况、健康度与可用性，支撑装备保障决策；

表4 部分演训数据片段

样本ID	要素类型	时间戳	经度	纬度	状态	指令	合格
T001	兵力行动	TT:TT:02.33	XX.2345	YY.8765	机动	前进	1
T002	装备状态	TT:TT:02.35	XX.2346	YY.8764	正常	无	1
T003	指挥指令	TT:TT:02.40	-	-	已执行	压制	1
T004	战场环境	TT:TT:02.42	XX.2347	YY.8763	晴	无	1
T005	兵力行动	TT:TT:02.45	XX.2348	YY.8762	隐蔽	待命	1

(3) 指挥指令类：承载上级作战意图与行动命令，部分指令因保密或数据采集限制存在经纬度缺失；

(4) 战场环境类：记录气象、地形等环境信息，为作战行动提供环境约束参考。

为确保标签的权威性与准确性，本次评估采用的真实标签基于“人工校验+领域专家会审”的方式构建，最终构建的标签集包含完整的质量问题标注。此次验证从完整性、准确性、时空一致性、标注一致性、复现率五大核心维度，对两种方法的评估效果进行量化对比，结果如表5所示。

## 5.2 结果分析

表5 两种方法评估效果对比结果

质量维度	评估指标	传统方法得分	本文方法得分	提升幅度	指标内涵
完整性	关键字段覆盖率	88.7%	98.5%	+9.8%	衡量兵力行动、指挥指令等关键字段无缺失的比例，直接反映数据可用基础。
准确性	信息匹配准确率	90.6%	99.2%	+8.6%	验证数据内容与战场真实情况的一致性，剔除虚假、错误数据的能力。
时空一致性	时空逻辑符合率	81.3%	97.8%	+16.5%	校验时间戳与空间位置的逻辑合理性(如机动轨迹不冲突、指令与位置匹配)。
标注一致性	标注一致性系数	0.73	0.93	+0.20	衡量多源数据标注规则的统一度，消除歧义与标注冲突的效果。
复现率	质量问题复现率	80.5%	99.6%	+19.1%	评估方法对缺失、异常等质量问题的检出与复现能力，反映方法的鲁棒性。

从验证结果看，在完整性方面，传统方法仅能识别显性缺失字段，对隐性缺失与语义缺失的检出能力不足，导致关键字段覆盖率仅为88.7%；本文设计的方法通过作战语义补全与多源数据关联，可基于兵力行动轨迹、上级指令逻辑等上下文信息，推断并补全指挥指令类数据的缺失位置，使覆盖率提升至98.5%，解决了演训数据“字段缺失不可用”的核心痛点。在准确性方面，传统方法依赖固定阈值，难以识别战场环境中的细微误差与标注歧义，准确率为90.6%；本文设计的方法引入战场知识图谱与真值比对机制，通过比对权威演训数据库，可精准修正位置偏差、状

态误标等问题，将准确率提升至99.2%，为作战决策提供可靠的数据支撑。在时空一致性方面，传统方法缺乏时序与空间的关联建模，时空一致性符合率仅81.3%；本文设计的方法通过TCN时序建模与空间拓扑分析，可有效识别“时间错位、位置冲突”等逻辑矛盾，将符合率提升至97.8%，保障战场数据的逻辑自洽。在标注一致性与复现率方面，标注一致性系数从0.73提升至0.93，表明本文设计的方法统一了多军兵种、多传感器数据的标注规范，消除了“一词多义”的歧义；质量问题复现率大幅提升19.1%，证明本文设计的方法对噪声、异常值的检出与还原能力

显著强于传统方法，具备更强的工程应用价值。

验证结果表明，本文设计的方法在完整性、准确性、时空一致性、标注一致性、复现率五个维度均实现了显著超越，且具备更强的自适应能力与鲁棒性，对联合演训数据质量评估具有有效性和可行性，可为后续实战化演训数字化转型、数据要素赋能体系作战提供理论支撑与技术路径。

## 6 结束语

本文系统研究了高质量演训数据集的构建技术和评估方法，提出了面向实战化演训的数据集“构建-治理-评估”一体化框架，构建了“静态基础质量+动态作战效用”双层评估模型，实现了数据质量与作战效用的统一度量。实证表明，所设计的方法能够有效提升演训质效，可为实战化演训数字化、智能化、体系化发展提供技术支撑。后续，将在强化标准互认、平台互通、资源共享方面进一步加强研究，充分释放数据要素在军队现代化建设与体系作战能力提升中的核心价值，并结合数字孪生、大模型智能标注、区块链存证、联邦学习共享、深度学习、时序模型等先进技术<sup>[15-16]</sup>，持续构建联合演训数据共享生态。



赵程亮（1991-），男，硕士，中国人民解放军63611部队工程师，主要研究方向为训练任务规划。



林李杰（2000-），男，学士，中国人民解放军63611部队助理工程师，主要研究方向为训练任务规划。



郑向通（1985-），男，学士，中国人民解放军63611部队高级技师，主要研究方向为训练任务保障。

## 参考文献：

- [1] 葛承堃, 贾晨星, 明月伟. 联合作战教学演训数据治理基础问题研究[J]. 指挥控制与仿真, 2024, 46(5): 139-146.  
Ge C L, Jia C X, Ming Y W. Research on fundamental issues of data governance in joint operations teaching and training [J]. Command Control & Simulation, 2024, 46(5): 139-146.
- [2] 张大永, 杨镜宇, 马骏, 等. 面向兵棋推演复盘分析的机器学习数据集构建[J]. 系统仿真学报, 2024, 36(3): 652-661.  
Zhang D Y, Yang J Y, Ma J, et al. Construction of a machine learning dataset

- for post-game analysis of military board games[J]. Journal of System Simulation, 2024, 36(3): 652-661.
- [3] 吴信东, 董丙冰, 堵新政, 等. 数据治理技术[J]. 软件学报, 2019, 30(9): 2830-2856.  
Wu X D, Dong B B, Du X Z, et al. Data governance technology[J]. Journal of Software, 2019, 30(9): 2830-2856.
- [4] 陈军, 赵仁凯, 王涛. 战场时空数据关联与一致性校验方法[J]. 指挥控制与仿真, 2022, 44(2): 56-61.  
Chen J, Zhao R K, Wang T. Method for correlating battlefield time-space data and verifying consistency[J]. Command Control & Simulation, 2022, 44(2): 56-61.
- [5] 张军, 王磊, 刘芳. 面向演训场景的高并发数据采集架构设计[J]. 计算机工程与应用, 2023, 59(10): 215-223.  
Zhang J, Wang L, Liu F. Design of high-concurrency data acquisition architecture for training and simulation scenarios[J]. Computer Engineering and Applications, 2023, 59(10): 215-223.
- [6] 王婷婷, 程远国, 王磊. 军事大数据环境下的数据安全访问控制模型[J]. 指挥控制与仿真, 2020, 42(4): 16-21.  
Wang T T, Cheng Y G, Wang L. Data security access control model in the military big data environment[J]. Command Control & Simulation, 2020, 42(4): 16-21.
- [7] U. S. Department of Defense. Joint Training and Training Technology (J-T<sup>3</sup>) System Overview[R]. Washington D. C.: U.S. DoD, 2020.
- [8] NATO. Allied Joint Training Data Standard (AJTDS)[S]. Brussels: NATO Standardization Office, 2019.
- [9] Advanced Distributed Learning (ADL) Initiative. Sharable Content Object Reference Model (SCORM) 2004 4th Edition [EB/OL]. (2018-06-15). <https://adlnet.gov>.
- [10] U. S. Department of Defense. OneTESS: Training Evaluation and Support System [R]. Washington D.C.: U.S. DoD, 2021.
- [11] 康红宴, 赵阳, 张鹏. 作战数据工程建设与发展[J]. 国防科技, 2018, 39(1): 120-124.  
Kang H Y, Zhao Y, Zhang P. Combat data engineering and development[J]. National Defense Science & Technology, 2018, 39(1): 120-124.
- [12] Han J, Kamber M, Pei J. Data Mining: Concepts and Techniques[M]. 3rd ed. Morgan Kaufmann, 2011.
- [13] 张阳, 郭刚, 刘思伟. 面向联合作战的时序数据质量校验技术[J]. 火力与指挥控制, 2021, 46(8): 8-14.  
Zhang Y, Guo G, Liu S W. Time-series data quality verification technology for joint operations[J]. Fire Control & Command Control, 2021, 46(8): 8-14.
- [14] 刘思伟, 郭刚, 张阳. 基于时空基准的联合作战数据融合技术[J]. 火力与指挥控制, 2021, 46(8): 1-7.  
Liu S W, Guo G, Zhang Y. Joint combat data fusion technology based on spatio-temporal reference[J]. Fire Control & Command Control, 2021, 46(8): 1-7.
- [15] 张军. 基于数字孪生的战场环境仿真与数据孪生技术研究[J]. 国防科技, 2024, 45(2): 33-40.  
Zhang J. Research on battlefield environment simulation based on digital twin and data twin technology[J]. National Defense Science & Technology, 2024, 45(2): 33-40.
- [16] 李磊. 区块链在军事数据安全共享与存证中的应用[J]. 指挥与控制学报, 2023, 9(3): 289-297.  
Li L. The application of blockchain in military data security sharing and evidence preservation[J]. Journal of Command and Control, 2023, 9(3): 289-297.

## 作者简介



霍明亮（1989-），男，硕士，中国人民解放军63611部队工程师，主要研究方向为训练任务规划。

收稿日期: 2026-05-19

通信作者: 霍明亮, 554296566@qq.com

基金项目: 黄永亮(1994-), 男, 学士, 中国人民解放军63611部队工程师, 主要研究方向为训练任务规划。

**Foundation Items:**