大模型军事应用需求分析与开发流程

雷 震，李立伟，张 龙*，冯轩铭，杨 波
（军事科学院系统工程研究院，北京 100101）

摘 要：随着大模型技术的不断发展，其在军事领域的应用价值日益凸显。为探索适配大模型军事应用的技术架构和开发流程，进一步提升作战效能，在深入分析大模型军事应用需求的基础上，通过研究Transformer和混合专家（MoE）架构技术机理及军事适用性，提出“预训练—监督微调—奖励建模—强化学习—知识蒸馏”的系统架构开发流程，并以海上联合作战智能态势感知系统为例进行验证。研究结果明确了大模型在多源异构数据融合、智能决策支持等任务的潜在应用价值，表明了开发流程的可行性，为推动大模型在军事领域深度应用，加速军事智能化转型提供理论参考。

关键词：大模型；军事应用；需求分析；技术机理；开发流程；案例分析
中图分类号：E919 **文献标识码：**A **文章编号：**1009-1300(2026)01-0049-12
DOI：10. 16358/j. issn. 1009-1300. 20250059

Requirement analysis and development process for militarized application of large-scale models

Lei Zhen, Li Liwei, Zhang Long*, Feng Xuanming, Yang Bo
(System Engineering Research Institute of the Academy of Military Sciences, Beijing 100101, China)

Abstract: With the continuous development of large-scale model technology, its application value in the military field is increasingly prominent. To explore the technical architecture and development process suitable for military applications of large-scale models and further enhance combat effectiveness, based on an in-depth analysis of the military applications requirements for large-scale models, a systematic architecture development process of "pretraining—supervised fine-tuning—reward modeling—reinforcement learning—knowledge distillation" is proposed by studying the technical mechanisms and military applicability of Transformer and Mixture of Experts (MoE) architectures. The feasibility of the development process is verified by taking the intelligent situation awareness system for joint maritime operations as an example. The results clearly demonstrate the potential application value of large-scale models in tasks such as multi-source heterogeneous data fusion and intelligent decision support, indicating the feasibility of the development process. The research provides theoretical references for promoting the in-depth application of large-scale

收稿日期：2025-04-03；**修回日期：**2025-06-09
作者简介：雷震，博士研究生。
通讯作者：张龙，博士研究生。
引用格式：雷震，李立伟，张龙，等. 大模型军事应用需求分析与开发流程[J]. 战术导弹技术，2026（1）：49-60.（Lei Zhen，Li Liwei，Zhang Long，et al. Requirement analysis and development process for militarized application of large-scale models[J]. Tactical Missile Technology，2026（1）：49-60.）

models in the military field and accelerating the intelligent transformation of the military.

Key words: large-scale models; military application; requirement analysis; technical mechanism; development process; case analysis

1 引 言

大模型，起源于大语言模型（Large Language Models，LLM），是在大数据、大算力、强算法基础上衍生的，具有通用的、泛化能力和庞大参数量的深度神经网络模型[1]，现已逐渐扩展出视觉大模型、多模态大模型以及基础科学大模型等新兴概念。

大模型军事应用是将大模型与战场情报分析、指挥决策控制等多个军事领域专业知识融合，通过数据分析挖掘，快速识别威胁、预测战场态势、优化作战方案，辅助军事指挥官快速科学决策，同时构建虚拟战场环境，提升军事训练实战化水平。

当前，大模型军事应用的核心突破点在于需求分析和开发流程的系统性构建。由于大模型技术存在可解释性不足、溯源困难、评估体系不完善等“黑箱”特性，亟需进一步深入考量军事任务特点、数据特性以及场景要求，明确功能、性能指标。从模型架构选择、训练优化至部署应用全环节入手，积极探索研究开发流程，推动和引领大模型军事应用落实落地。

2 大模型军事应用需求分析

随着人工智能（AI）技术的突破性发展，大模型军事应用已成为国防现代化的重要驱动力。开展基于场景的大模型军事应用需求分析，可以精准对接战场实际需求，适配多样化差异化任务，提高大模型军事应用的针对性和适用性，强化人工智能技术与军事行动深度融合，加速推动军事智能化转型发展[2]。

2.1 情报分析与预测预警

借助先进传感器技术，大模型能整合卫星影像、雷达信号等多模态数据，通过智能降噪识别虚假情报与背景噪声，自动标记高价值目标，缩短情报处理链。采用基于自注意力机制的Transformer架构，大模型能将海量多模态数据映射到同一语义空间，解决异构数据时间戳和空间尺度不匹配问题，还可通过自监督预训练和小样本学习补全残缺数据，实现异构数据交叉关联与动态融合。此外，构建军事知识图谱并建立语义关联网络，基于动态知识更新机制，大模型能快速关联历史与新数据，挖掘推理、分析攻击行为和模式，预测态势演变趋势、预警潜在威胁并辅助生成应对策略。

2.2 决策支持与协同控制

在决策支持方面，大模型通过整合多源异构数据生成战场态势全景图，基于历史战例库和强化学习算法生成作战方案，并通过模拟推演评估可行性与风险，还能结合实时数据动态调整方案，助力决策者应对复杂战场环境。

在协同控制方面，大模型赋予无人作战平台自主感知、决策与执行能力，支持无人集群任务分配、编队控制等协同行动；同时通过轻量化技术部署至单兵终端等边缘设备，以多模态交互实现离线对话、威胁识别等功能，避免电磁干扰与路径冲突，提升有人/无人协同作战效能。

2.3 武器研发与综合保障

大模型赋能武器装备全寿命周期管理，有望重塑现代武器装备体系与军事保障模式，推动装备体系从“经验驱动”向“数据智能驱动”转型。在现役装备升级改造方面，大模型通过深度学习与多模态感知技术，自动生成优化算法并完成软件定义升级，提升装备任务规划、威胁规避等智能化水平，强化目标识别跟踪与火力打击精度。在新式装备研发设计方面，大模型与生成式人工智能结合实现代码自动生成与生产流程精准控制，有效降低人力成本，缩短研发周期。在综合保障方面，大模型通过分析历史记录和实时数据，不仅可以优化供应链路线，预测物资运输干扰并动

态调整，还可以实现预测性维护，及时诊断故障原因并提供维修建议。

2.4　舆论布势与认知攻防

在认知对抗方面，大模型利用搜索引擎优化、数据标签操纵等“算法解释”技术，定向传播特定内容，影响受众群体的观念塑造，实现意志直达。例如，在俄乌冲突爆发前，双方利用人工智能、大数据等先进技术展开认知攻防，抢占道义优势，为各自即将展开的军事行动塑形造势。在精准画像方面，大模型依托互联网采集多维多层异构认知数据，分析找准对手认知体系的薄弱点和敏感点，设计精准信息“弹药”。在深度伪造方面，大模型可以生成大量的假视频、假音频、假文件，借助社交网络快速扩散，制造信息茧房。俄乌冲突中出现的乌总统 “投降” 伪造视频，即通过大模型合成并传播，干扰对手士气与舆论走向。

3　大模型军事应用面临的矛盾困难

大模型军事应用虽然前景广阔，但仍受到数据质量不高、算法机理不明、算力资源紧缺等因素的制约和影响，必须要谨慎应对，提前谋划、科学布局，确保大模型在军事领域安全、可靠、高效应用，分析如下。

3.1　算料:预训练数据质量不高

（1）数据采集获取难。常态化战备执勤、演习演训等会产生海量数据，但非结构化数据占比高、聚合能力弱[3]，且受机密性与敏感性限制，数据生成、采集和存储流程严格，导致 “数据量大而可用量小”的矛盾突出。

（2）清洗标注过程难。军事数据涉及时间、空间、数量、状态等多维要素，传统知识图谱以“实体–关系–属性”三元组表示知识的方式，难以完整刻画复杂军事事实，导致知识存储更新滞后、问答逻辑复杂化[4]。此外，不同军种部门的数据标准与格式不统一，进一步加剧标注整合难度。

（3）数据真假辨别难。战场环境不确定性、传感器精度局限性及敌方虚假数据诱饵攻击等因素，导致数据普适性差、易受污染。以坦克目标识别为例，仅基于沙漠环境数据训练的模型，面对雪地伪装的坦克图像时，表现出泛化能力严重不足[5]。

解决真实作战环境与模型训练需求之间的矛盾，需进一步强化数据治理、统一标准体系、增强数据抗干扰能力，为大模型军事应用发展注入高品质“算料”。

3.2　算法:运行机理可解释性不够

（1）复杂性导致“黑箱”效应。深度学习架构的内部结构复杂，参数量巨大，决策过程难以追溯，内部特征提取与逻辑推理路径也无法直观呈现，其计算复杂性进一步加剧了“黑箱”问题[6]。

（2）特殊性导致标准不一。不同作战场景对可解释性的需求差异显著。传统“事后解释”方法难以满足动态战场环境下的即时性要求，而针对 “直觉决策” 等人类指挥艺术的算法模拟仍处于理论阶段，导致解释标准难以统一。

（3）局限性导致信任失衡。现有可解释技术（如模型蒸馏、特征反演）仅能局部解析特征关联，无法全局解释模型涌现能力，而且在受到数据投毒等外部攻击后，大模型还可能会生成虚假决策逻辑，输出错误信息，“解释鸿沟”导致其在军事决策和应用中可信任程度和安全性受到质疑。

解决算法“黑箱”特性与透明性、可追溯性需求之间的矛盾，需持续推动模型架构创新，结合领域知识图谱构建全流程统一的评估标准和多层级信任评估体系，为大模型军事应用发展打造强逻辑“大脑”。

3.3　算力:成本资源算力紧缺

（1）算力需求与资源限制的矛盾突出。大模型训练需处理海量多模态军事数据，并执行复杂深度学习任务，对算力需求极大。但受制于数据来源，以及保密安全等问题，当前主要通过“大模型调用小模型 + 外挂知识库”的轻量化模式实现有限应用，无法满足高复杂度任务的算力需求。

（2）硬件支持和技术发展的瓶颈约束。大模型训练和推理过程不仅依赖大算力服务器支持，而且需要高效的通信架构来同步模型参数和梯度信息，其高昂的部署成本和通信带宽延迟约束限制，一定程度上影响了算力执行效率。因此现阶

段轻量化模型或性价比更高的异构计算平台或许是更现实的选择[7]。

（3）能源保障与资源分配的失衡困境。随着大模型向体量更大、参数更多的方向发展，必然会消耗更多的成本和能源[8]，加剧军事能源资源的紧张局面，影响其合理分配。尤其是在战时，作战装备能源供应的优先保障与大模型训练的高能耗矛盾将进一步凸显。

解决高算力需求与有限资源供给的矛盾，要积极巩固和加强基础设施建设，推动算法创新应用，优化资源配置，为大模型军事应用发展加装高性能“引擎”。

4 大模型架构技术机理

大模型通过创新架构设计和复杂算法组合，形成了强大的通用泛化能力与海量数据处理能力。深入研究大模型架构技术机理，有助于更好地理解大模型在军事应用中的潜力与局限，为其军事应用系统的开发提供坚实的理论基础。

4.1 大模型发展历程

大模型技术的发展是一个不断演进的过程（表1）。20世纪80年代，神经网络开始兴起，但受限于数据量和计算能力，模型规模较小且功能有限。21世纪初，机器学习算法取得一定进展，支持向量机等算法在特定领域得到应用，但处理大规模复杂数据的能力依然不足。

2012年，深度学习的标志性成果AlexNet在图像识别领域取得重大突破，开启了深度学习快速发展的新时代。此后，深度神经网络不断发展，模型结构日益复杂，性能也不断提升。2017年，Transformer架构横空出世，成为大模型技术发展的关键转折点，其基于自注意力机制的序列建模新范式开启了预训练模型时代。

2018年以来，BERT、GPT等模型持续迭代升级，多模态、混合专家等架构不断创新引用推动大模型技术向更高效、更智能方向发展。

Transformer架构作为大模型发展的里程碑，在长距离依赖建模与复杂时序数据并行处理等方面优势显著。一方面，其基于自注意力机制的序列建模范式革新了深度学习，通过全局上下文表征能力直接捕捉序列任意位置依赖关系，突破传统模型局部特征捕捉时梯度消失或爆炸的局限。另一方面，Transformer支持输入序列并行处理，大幅提升训练效率，为大规模模型训练提供可能。此外，该架构通用性强，可通过迁移学习适配情报分析、指挥控制、装备保障等多类军事任务，

表1 大模型架构发展历程

Table 1 Development history of large-scale model architecture

发展阶段	时间	代表性成果	主要贡献
初始探索期	1986年	循环神经网络(RNN)	处理序列数据,引入时间依赖性
	1989年	卷积神经网络(CNN)	引入卷积层和池化层,适用于图像识别
	1997年	长短期记忆网络(LSTM)	解决RNN中的梯度消失问题,适用于长序列数据
	2006年	深度学习概念	开启深度学习的新高潮
	2012年	AlexNet	标志着深度学习在计算机视觉领域的全面爆发
转折引领期	2017年	Transformer架构	引入自注意力机制,解决长程依赖问题,开启预训练模型时代
	2018年	BERT	双向Transformer编码器,显著提升自然语言处理(NLP)任务性能
	2018年	GPT	单向Transformer解码器,展示语言生成能力
	2019年	GPT-2	扩大模型规模(15亿参数),提升生成能力
	2020年	GPT-3	参数规模达1750亿,展示少样本学习能力
	2020年	Gshard(MoE模型)	引入MoE架构,扩展模型规模,降低计算成本
	2021年	Switch Transformers	简化MoE实现,提高训练效率
融合拓展期	2023年	GPT-4V	结合文本和图像,扩展模型的多模态能力
	2024年	GPT-4o	进一步整合音频和视频,提升多模态交互能力
	2025年	DeepSeek-R1	利用MoE架构,显著降低运营成本,提升性价比

使其成为军事智能系统中实现多源异构数据融合、实时态势推演及动态决策优化的理想基座。

4.2　Transformer架构核心机理

Transformer架构以自注意力机制（Self-attention Mechanism）为核心[9]，通过并行计算获取输入序列全局信息，并通过网络层进行传递。标准架构（图1）由编码器和解码器组成，分别包括一个编码层和若干相同的Transformer模块层。

编码层对输入词序列进行词嵌入与位置编码处理后转化为连续向量。词嵌入将词汇映射为高维语义向量，捕捉词语间语义与语法关系；位置编码弥补自注意力机制的位置感知缺陷，采用正余弦函数构建固定维度向量，为每个词赋予位置信息。具体而言，位置i对应的位置编码向量第j个元素按奇偶规则计算，d为词嵌入维度：如果j是偶数，元素值为$\sin(i/10000^{j/d})$；如果j是奇数，元素值为$\cos(i/10000^{j/d})$。

编码器的每个Transformer模块通过多头注意力层捕捉特征关联，再经全连接前馈网络处理信息，两者通过残差连接与层归一化增强稳定性。解码器额外引入交叉注意力层，融合编码器输出与当前解码状态，适用于文本生成等任务。

（1）自注意力层，是Transformer模型的核心组成部分。它包含一个查询矩阵$Q\in\mathbb{R}^{n\times d_q}$，一个键矩阵$K\in\mathbb{R}^{m\times d_k}$和一个值矩阵$V\in\mathbb{R}^{m\times d_v}$，其中，矩阵中的每一行对应一个词，$\mathbb{R}$为实数集，表示该矩阵中所有元素均为实数；$n$为当前输入序列的词数（token数量），$m$为序列长度，通常$m=n$；$d_q$、$d_k$、$d_v$分别是对应矩阵的值向量维度，通常$d_q=d_k=d_v$。注意力机制的计算方式：

$$\mathrm{Attention}(Q,K,V)=\mathrm{Softmax}\left(\frac{QK^{\mathrm{T}}}{\sqrt{d_k}}\right)V \tag{1}$$

其中，QK^{T}为Q与K的转置相乘，可得到一个$n\times m$的相似度矩阵，$\frac{1}{\sqrt{d_k}}$为缩放因子，用于稳定点积方差，Softmax(·)对缩放后的相似度矩阵的每一行进行归一化操作，得到注意力权重矩阵，再与值矩阵V相乘，得到最终输出。

直观来说，实数域下的矩阵$H\in\mathbb{R}^{n\times d_v}$中的每一行是矩阵$\boldsymbol{V}$中$m$个行向量（每个维度为$d_v$）的加权和，且该加权和的权重由查询序列对应的查询向量与键-值序列对应的键矩阵的点积结果确定。

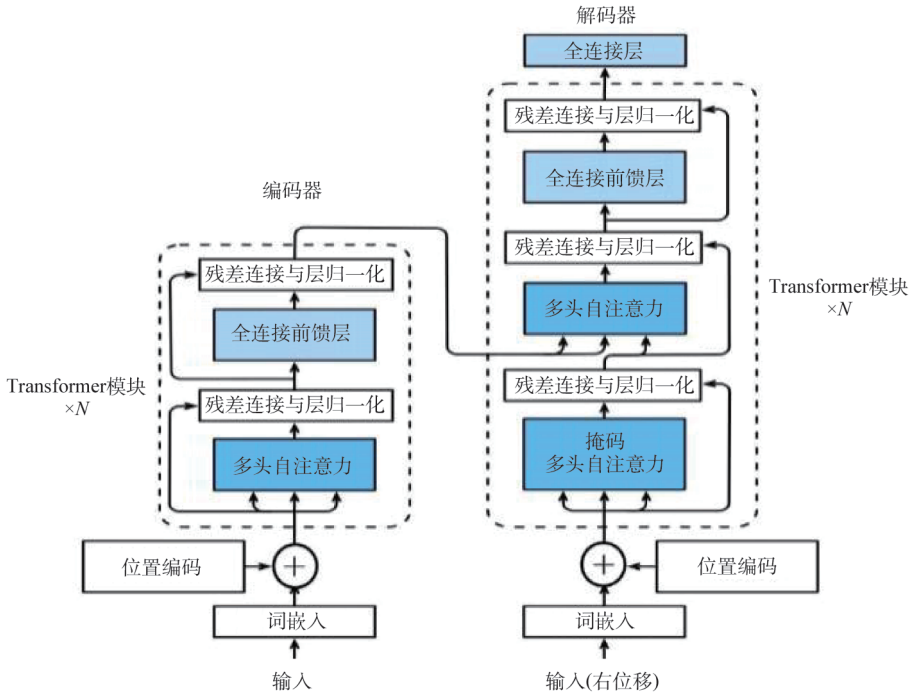

图1　Transformer架构

Fig. 1　Transformer architecture

记具有序列长度 n 的查询序列特征矩阵和具有序列长度 m 的键-值序列特征矩阵分别为 $X_q \in \mathbb{R}^{n\times d}$ 和 $X_{kv} \in \mathbb{R}^{m\times d}$，三个矩阵 Q、K、V 由三个线性变换得到，即通过 W_q、W_k、W_v 三个不同的可学习权重矩阵，分别对 X_q 和 X_{kv} 做线性变换，满足 $Q = X_qW_q$、$K = X_{kv}W_k$、$V = X_{kv}W_v$。

Transformer 模型采用自注意力机制，因为三个矩阵 Q、K、V 都来自于前一层的相同特征矩阵 $X \in \mathbb{R}^{n\times d}$。通过这种方式，模型能够有效捕捉输入序列中各元素间的依赖关系，从而更好地理解上下文信息。

（2）全连接前馈层，位于注意力层之后，由两个线性变换和一个非线性激活函数组成。将输入矩阵表示为 $X \in \mathbb{R}^{n\times d_i}$，$n$ 为 token 数量，d_i 为每个 token 的输入特征维度。前馈层的输出：

$$\mathrm{FFN}(X) = \sigma\left(XW_1 + b_1\right)W_2 + b_2 \tag{2}$$

其中，$\sigma(\cdot)$ 是激活函数（通常为 ReLU 或 GELU），而 $W_1 \in \mathbb{R}^{d_i\times d_f}$，$b_1 \in \mathbb{R}^{d_f}$，$W_2 \in \mathbb{R}^{d_f\times d_o}$，$b_2 \in \mathbb{R}^{d_o}$ 均为可学习的参数。在实践中，d_i 通常设置为 d_o，d_f 设置为 d_i 的 4 倍。前馈神经网络（FFN）作用包括两个方面：1）通过非线性激活函数增强模型的表达能力；2）在每个位置独立整合信息，同时与自注意力机制协同，使模型既能捕捉全局长距离信息，又能在局部位置进行信息整合。

（3）残差连接和层归一化，在每个注意力层和每个全连接前馈层之后。对于某一层神经网络 $f(\cdot)$，残差连接和归一化层定义为 $\mathrm{LayerNorm}(X+f(X))$。这种设计能够确保模型在加深时能保留信息并维持性能，有助于缓解梯度消失或爆炸问题，使模型训练更加稳定，学习效果更好。

4.3 混合专家架构（MoE）核心机理

2025 年，DeepSeek 开源推理引擎采用 MoE 架构并轻量化设计，实现亿级参数模型实时推理部署，标志着基于 Transformer 的通用智能架构进入规模化应用新阶段，为大模型军事应用架构创新提供新范式。

4.3.1 MoE 架构组成

MoE 架构是一种将多个称为“专家”的子模型组合在一起的机器学习架构，主要由门控网络（Gating Network）和专家网络（Expert Networks）两部分组成（图 2）。

门控网络的主要作用是通过计算输入数据与各个专家之间的相关性，生成权重矩阵，依据专家模型对当前任务的重要度决定其激活程度。例如，路由网络可以计算每个词元对应各个专家的权重，并选择概率最高的几个专家进行激活。

专家网络的主要作用是处理特定的任务或数据子集。每个专家模型作为一个独立的子网络，可以是 FFN，也可以是其他类型的神经网络。

4.3.2 MoE 架构工作原理

在 MoE 架构中，每个混合专家层包含 k 个专家组件，记为 $[E_1, E_2, \cdots, E_k]$，假设其中每个专家组件 E_i 都是一个前馈神经网络。

对于输入的每个词元表示 x_t，模型通过门控网络（或称为门控函数）G 来计算该词元对应于各个专家的权重。

（1）通过线性层 $\boldsymbol{W}^G \in \mathbb{R}^{H\times K}$ 映射为 k 个专家的得分。

（2）这 k 个专家的得分将被送入 Softmax 函数计算出它们的权重矩阵：

$$G(x_t) = \left[G(x_t)_1, G(x_t)_2, \cdots, G(x_t)_k\right] \tag{3}$$

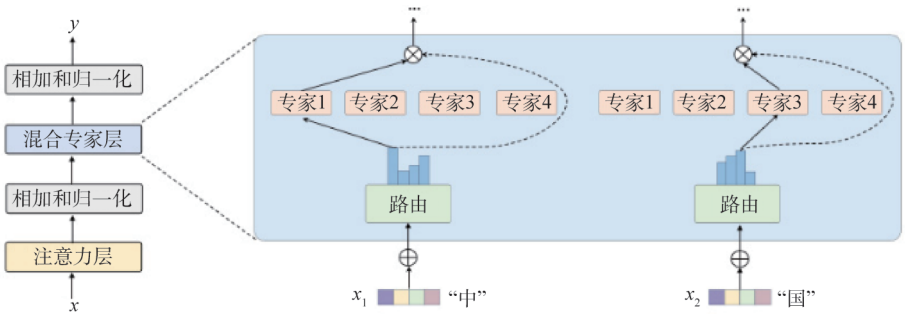

图 2 混合专家架构（MoE）

Fig. 2 Mixed expert architecture (MoE)

（3）基于生成的权重矩阵选择出概率最高的 k 个专家进行激活。没有被选择的专家权重将被置为0。

上述门控网络的计算过程如下：

$$G(x_t)=\mathrm{Softmax}(\mathrm{top}_k(x_t\cdot W^G)) \tag{4}$$

（4）每个被选择的词元的输出的加权和将作为该混合专家网络层的最终输出 O_t：

$$O_t=\mathrm{MoELayer}(x_t)=\sum_{i=1}^{k}G(x_t)_i\cdot E_i(x_t) \tag{5}$$

MoE 架构通过门控网络动态选择专家处理输入，利用多专家优势提升模型性能与泛化能力。每个输入指令（token）可激活不同专家，使模型根据输入特性灵活调用适配专家，在增强表达能力的同时，降低计算成本与参数规模。

4.3.3　MoE 与 Transformer 的内在联系

MoE 架构本质是对 Transformer 并行计算范式的扩展与动态重构。Transformer 通过自注意力机制实现序列全局建模，但固定权重难以适配动态任务；MoE 引入门控路由策略，以专家网络动态组合替代标准化多头注意力（图3），使模型可根据输入特征自动选择最优子网络路径。这一融合既继承 Transformer 全局上下文建模能力，又通过稀疏激活机制（如仅激活部分专家）显著降低计算冗余。

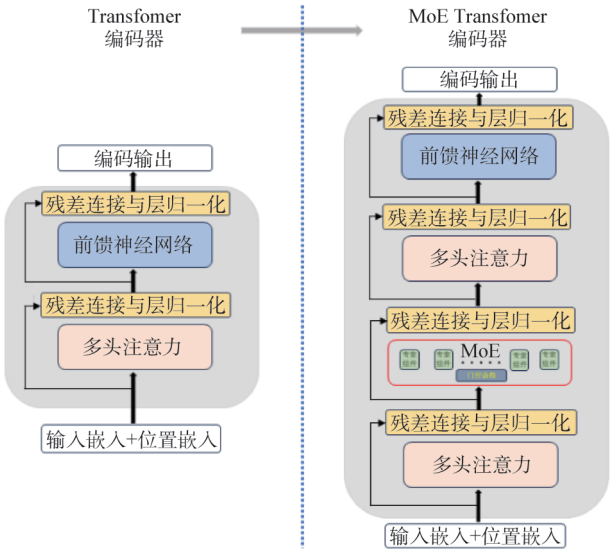

图3　前馈神经网络替换为专家层

Fig. 3　Replacing the feedforward neural network with the expert layer

4.4　大模型架构军事适用性分析

加速军事智能化转型发展，对大模型在军事领域应用的实时性、鲁棒性以及多任务并行处理能力提出了严苛要求。可以将军事任务的核心技术需求抽象为一个四元组：

$$E=(T_{\mathrm{latency}},R_{\mathrm{robustness}},M_{\mathrm{multimodal}},C_{\mathrm{resource}}) \tag{6}$$

其中，T_{latency} 表征实时性阈值，$R_{\mathrm{robustness}}$ 为对抗环境下系统鲁棒性，$M_{\mathrm{multimodal}}$ 定义为多源数据融合处理能力，C_{resource} 为资源约束条件。基于此框架，开展 Transformer 与 MoE 架构在军事场景的应用潜力与局限性分析，为大模型在军事领域应用的系统设计提供思路。

4.4.1　Transformer 架构的军事适用性分析

当前，Transformer 架构已经广泛应用于自然语言处理、计算机视觉等领域，有大量的研究成果和实践经验可供参考，这使得其在军事领域的应用更容易落地和优化。

一方面，Transformer 架构具备强大的多模态时空建模能力，在执行多源情报融合、战场态势感知等任务时，具有巨大的潜在优势。但自注意力矩阵的 $X(n^2)$ 空间复杂度对计算和内存的需求高，导致其在边缘设备上部署相对困难。

另一方面，战场上信息流高度交叉，实时动态变化，Transformer 的自注意力机制能够很好地捕捉信息之间的长距离依赖关系，在标注数据稀缺、样本有限的条件下实现高效微调，不仅能提高决策效率，还能增强系统鲁棒性。但在执行对实时性要求极高的任务时，例如，制导导弹在打击目标过程中，需要快速完成目标识别和轨迹预测，Transformer 计算复杂度高使其难以应对战场突发变量，打击实时性受限。

4.4.2　MoE 架构的军事适用性分析

MoE 采用了专家稀疏激活机制，一方面，可以根据不同的任务分配不同的专家模块，同时支持专家级冗余备份，具有较高的扩展性、灵活性和适应性。另一方面，由于 MoE 只激活部分专家网络，具有较高的计算经济性，在大规模、分布式多任务并行处理时具有一定的优势。

MoE 的适用瓶颈主要体现在其复杂的门控路由机制可能会导致负载不均衡，并引入额外的计

算开销，制约鲁棒性和实时性表现。

综上所述，尽管在实时性表现方面Transformer和MoE都存在一定的缺陷，但仍可以考虑利用知识蒸馏，在边缘设备上部署轻量化Transformer来改善这一情况。在面对大规模情报分析和多任务复杂决策时，则可以考虑采用MoE架构，实现任务-专家的动态匹配。而在强对抗环境下，可以采用Transformer+MoE的混合架构，在加快信息处理、提高决策效率、缩短杀伤链闭合时间的同时，提高作战单元分布式协同能力。总之，要根据任务场景的复杂性和多样性，审慎选择适配策略。

5 大模型军事应用架构开发流程

军事数据特殊、环境威胁多变、决策输出可信、边缘迁移部署等对大模型军事应用提出了新的更高要求，需构建适配的规范化架构开发流程，推动大模型从技术概念向实战能力转化。

5.1 预训练(Pre-Training)

作为基础构建环节，预训练要考虑到军事数据的特殊性、术语密集性以及训练数据不足等现实挑战，基于Transformer架构，采用二次预训练和增量预训练的多节点分布式训练策略，在有限的数据资源下提升模型的适应性。

（1）分词处理。使用定制化的分词器对军事文本数据进行分词操作，将文本序列分解为字符（Character-level）或字词级别（Word-level）的token。考虑到军事术语的复杂性和专业性，分词器需具备对军事专有词汇的识别能力，以确保分词的准确性和语义完整性。

（2）词典映射构建。根据分词结果，构建词典映射（Vocabulary Mapping），将每个token映射为唯一的索引值。若采用预训练词向量（Pre-trained Word Embeddings），则需将词典映射与词向量文件进行对齐处理，确保词向量的语义信息能够准确嵌入模型。

（3）序列转换。基于构建好的词典，将分词后的文本序列转换为数字序列（Numerical Sequences）。将文本中的每个token替换为其对应的词典索引值实现，为模型训练提供结构化的输入数据。

（4）数据包处理。对转换后的数字序列进行标准化处理，长度不足的进行填充（Padding），超出最大长度的进行截断（Truncation），确保每个数据包内的数据维度一致，且长度符合模型的可接受范围。

5.2 监督微调(Supervised Fine-Tuning,SFT)

作为定向校准环节，监督微调是在预训练基础上继续注入军事领域特定数据和任务指令，提升涉密数据处理能力，在不暴露原始数据的前提下，优化模型性能，提升军事任务需求匹配度。

（1）指令数据构建。结合文本、图像、视频、语音以及传感器信号等多模态数据，可以通过人工标注或者大模型自动生成与少量人工标注相结合的半自动生成方式，构建多样化指令数据。

（2）整合上下文与输入文本。采用“指令+上下文+输入文本”的模式构建提示（Prompt），以实现军事领域语义理解的简单提示和少样本（Few-shot）提示，在少量样本的情况下快速适应特定任务[10]。

（3）构建多层级的智能化军事专用语料库。结合向量知识库，拓展大模型支持的军事专用语料长度，完成外挂知识的加载、读取、分割与向量化、匹配、Top-K推荐以及语义提示与回答生成等过程[11]。

5.3 奖励建模(Reward Modeling,RM)

作为价值对齐环节，奖励模型能够为后续的强化学习阶段提供可靠的奖励信号，使其在军事任务中生成更高质量、更符合人类偏好的响应。

（1）定义奖励模型的输入输出。将监督微调（SFT）模型的最终非嵌入层（通常为Softmax层）删除，并替换为一个线性输出层。定义输入为提示和响应（Response），输出为标量奖励值（Scalar Reward），用于衡量生成响应的质量。

（2）数据的采集与标注。针对同一输入提示采样多条不同的输出响应，进行响应对比，并给出倾向性标签。通过计算每条样本中两个回复的奖励值之差，拟合标注的倾向性概率[12]。

（3）应用K折交叉验证（K-Fold Cross-Valida-

tion）优化损失函数。奖励模型的训练一般选用交叉熵损失算法，为提高模型泛化能力，避免对特定子集的过度拟合，考虑引入K-Fold交叉验证，损失函数则定义为[13]

$$\mathrm{loss}(\theta)=-\frac{1}{|D|}E_{(x,y_w,y_l)\sim D}\cdot\left[\log\left(\sigma\left(C\cdot\left(r_\theta(x,y_w)-r_\theta(x,y_l)\right)\right)\right)\right] \tag{7}$$

在K-fold交叉验证中，首先将数据集随机分成K个大小相等、互不重叠的子集，其中$r_\theta(x,y)$表示奖励模型在参数为θ时，针对输入提示x和对应的完成文本y输出的标量奖励值；y_w代表人类偏好的“首选完成文本”，y_l为对比的“非首选完成文本”，$|D|$表示数据集样本数，包含所有人类标注的比较样本。$\frac{1}{|D|}E_{(x,\ y_w,\ y_l)\sim D}$表示遍历数据集$D$中所有人类标注的比较样本三元组（$x$，$y_w$，$y_l$），并对后续损失项取平均值。在每次迭代中，选择一个子集作为验证集，其余K-1个子集作为训练集，同时引入缩放因子C（通过交叉验证或其它技术选择的一个常数），以减少损失函数的波动。统计K次验证结果的均值和标准差，作为模型泛化性能的最终评估指标。

（4）归一化处理。通过计算奖励值的均值和标准差，对奖励模型的输出进行标准化处理，使其在强化学习训练前达到平均分数为0。

5.4　强化学习(Reinforcementlearning,RL)

作为策略优化环节，通过持续迭代优化，大模型能够不断适应动态变化的军事任务需求，有效提高其垂直领域的业务能力，为后续能交互、能决策、能学习和自成长的演化奠定基础[14]。

（1）初始化生成策略，根据输入的查询采样回复，并计算奖励值。

（2）根据反馈的奖励值，使用近端策略优化（Proximal Policy Optimization，PPO）算法更新当前策略，将每个时间步视为一个字节对编码（Byte Pair Encoding，BPE）令牌，把文本分解为更小的、可管理的单元，从而简化模型输入。同时引入词级别惩罚项，用于惩罚学习策略之间的KL散度（Kullback-Leibler，也叫相对熵，用于量化信息损失程度）差异π_ϕ^{RL}，防止策略过度优化。由此可将全部奖励R写为[15]

$$R(x,y)=r_\theta(x,y)-\beta\log\left[\pi_\phi^{\mathrm{RL}}(y\,|\,x)/\pi^{\mathrm{SFT}}(y\,|\,x)\right] \tag{8}$$

其中，π^{SFT}为原始的监督模型。

（3）通过评估回复质量、策略稳定性等指标，调整奖励函数、惩罚项等参数，不断优化生成策略。

5.5　知识蒸馏(Knowledge Distillation,KD)

作为部署应用环节，该过程将大型、复杂模型中的知识迁移至更轻量、高效的学生模型中，在保持性能的同时，显著降低计算和存储需求，使大模型的军事应用兼顾效率和效能。

（1）在选定军事领域大模型作为教师模型的前提下，考虑战术边缘设备的硬件平台、计算能力、存储容量，以及响应时延等约束条件，遴选结构相对简单、参数量更少的学生模型。

（2）动态调整蒸馏温度、损失函数等参数，实现教师模型和学生模型结构对齐。通过计算最小化KL散度或均方误差，对教师模型的软硬标签输出以及中间层特征表进行迁移训练（图4）。

（3）通过性能对比，评估学生模型的准确性、鲁棒性和泛化能力等指标，并进一步迭代优化，为轻量化部署提供可靠性支撑。

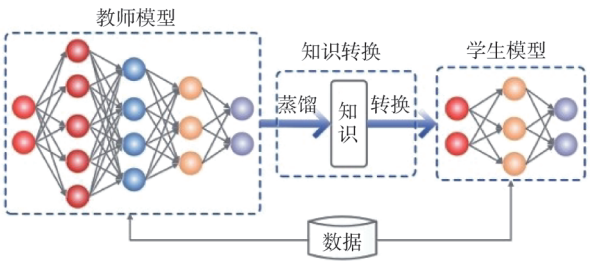

图4　教师-学生模型知识蒸馏

Fig. 4　Teacher-student model knowledge distillation

6　大模型军事应用架构案例分析

在现代海上联合作战中，态势感知能力是决定作战成败的关键因素之一。近年来，大模型技术在自然语言处理、图像识别和多模态数据融合等领域取得了显著进展，为提升海上联合作战态势感知能力提供了新的思路和技术手段。本文以“基于大模型的海上联合作战智能态势感知系统”

为例，全流程探索实践“预训练–监督微调–奖励建模–强化学习–知识蒸馏”技术范式。

6.1 案例背景

海洋空间广阔，环境复杂多变，多军兵种联合作战，态势瞬息万变，海上威胁日益多样化和复杂化。在高强度对抗的海战场景下，作战双方投入大量先进武器装备。航母战斗群作为我方海上作战的核心力量，面临着来自多维空间的威胁。空中层面，敌方先进战机凭借隐身与超视距打击能力构成重大威胁，无人机则凭借其小型化、低成本、零伤亡的优势实施集群作战和灵活侦察袭扰，进一步加剧威胁复杂性[16]；水面上，敌方舰艇先进的反舰导弹系统随时在防御范围外发动饱和攻击，无人舰艇以隐蔽机动作战方式突破我舰艇编队防御，实施近距离侦察攻击；水下领域，静音性能、续航能力、打击能力进一步升级，敌方潜艇可长时间潜伏在深海，发动远程打击或突然袭击，无人潜航器不仅破坏海底通信光缆、水下基础设施等，还具备情报收集和水下攻击能力，增加我方水下防御难度。

传统的雷达、声呐等传感器系统收集的海量数据，存在格式多样、噪声干扰严重，关联性差等问题，人工处理效率低下，无法及时形成完整准确的战场态势图。因此，开发一套基于大模型的智能战场态势感知系统，对海量多源异构数据进行深度挖掘和融合处理，实现对战场态势的精准感知和快速响应，成为提升航母战斗群作战效能的关键。

6.2 开发流程

在基于大模型的海上联合作战智能态势感知系统开发中，“预训练—监督微调—奖励建模—强化学习—知识蒸馏”五个核心环节构成了严密完整的技术链条，彼此间既分工明确又深度协同，形成从数据理解到模型部署的递进式能力跃升路径（图5）。

6.2.1 预训练阶段

预训练作为全流程的基础节点，通过构建多域异构数据特征工程，搭建基础认知能力，让模型能够“读懂”海上战场数据。

（1）数据处理。对包括卫星图像、雷达信号、声呐数据、情报文本等来源广泛且格式多样的作战数据进行预处理，转化为模型能够识别的数字信号。例如，将卫星拍摄的舰艇图像、雷达回波信号以及声呐探测到的水下目标信号转化为数字向量。同时，针对军事领域的专业术语，如“航母战斗群”“水雷区”和无人水下航行器（UUV）等，建立包含海军战术术语在内的军事词汇“专属词典”，通过位置编码融合时间戳信息，使模型能够识别并理解其含义和上下文关系。

（2）分布式预训练。可以采用 MoE 混合专家架构，在对互联网公开情报数据通用特征提取的基础上，引入海上联合作战军事脱敏数据进行领域适配，通过“二次预训练 + 增量更新”的方式[17]，提高大模型“跨域精准识别和目标关联”

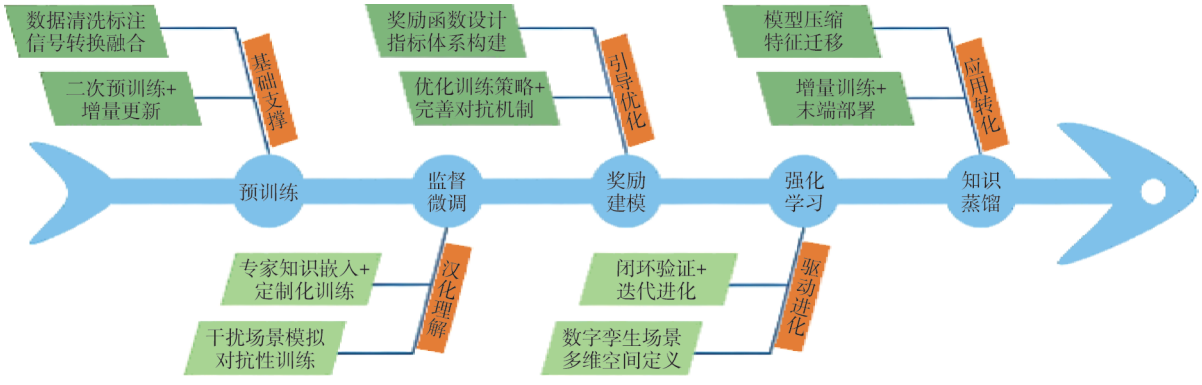

图5 开发流程及相互关系

Fig. 5 Development process and interrelationship

能力。例如，在海上航道中识别伪装成商船的敌方舰艇，或将雷达发现的“高速目标”与声呐检测的“无噪声尾流”关联识别为敌方隐身快艇。

6.2.2　监督微调阶段

监督微调作为深化理解的关键节点，通过构建“多传感器数据–战术意图标签”的细粒度训练集[18]，让模型能够“理解”海上作战决策。

（1）嵌入专家知识，开展定制化训练。引入海上联合作战规则库，采用提示工程构建“规则–数据”映射关系，并对Transformer 编码器进行稀疏化微调，通过人工设计大量“作战问题–答案”对，模拟实际作战场景中的问题和解决方案，为模型深度注入海战规则理解能力。

（2）模拟干扰场景，增强对抗性训练。模拟极端作战环境，如大风浪、低能见度、强对流等复杂海况，或者是雷达信号干扰、传感器故障等干扰场景，通过快速梯度符号法（Fast Gradient Sign Method，FGSM）生成对抗样本，训练模型在数据不完整或缺失情况下的鲁棒推理能力。

6.2.3　奖励建模阶段

奖励建模作为引导优化的核心节点，通过构建科学合理的评分规则库，实现作战效能导向价值对齐，让模型能够“判断”作战决策优劣。

（1）设计奖励函数。按照“战略–战役–战术”等级划分，以及“战前筹划预置–战中执行调控–战后评估规划”阶段划分，构建多层级多阶段奖励函数指标体系[19]，并设置评分规则。

（2）优化训练策略。建立“专家标注–对抗训练”双阶段训练策略，设计并优化“红蓝军博弈对抗”机制，对模型方案输出进行两两对比，构建偏好数据集，训练模型理解“最优决策”的内在逻辑，减少无效或错误判断。

6.2.4　强化学习阶段

强化学习作为驱动进化的引擎节点，通过构建数字孪生虚拟场景，实现模型在“虚拟战场”中自主进化。在面对复杂海上作战环境时，让模型能够“调整”作战行动策略。

（1）构建数字孪生训练框架。构建包含直升机、舰载机、舰艇、潜艇等有人作战单元，以及无人机、无人船、无人潜航器等无人作战平台的数字孪生场景，定义“传感数据+己方态势+敌方威胁”三维状态空间和“威胁等级调整+资源动态分配+作战进程转换”三类动作空间。

（2）优化策略算法。建立“模型建议–人工校正–奖励回溯”闭环验证机制[20]，设置正向奖励和负向惩罚边界约束条件，例如，反潜无人机成功避开敌方防空火力并发现目标时，给予正向奖励，反之则给予负向惩罚。通过多次模拟推演和迭代进化，训练模型能够自主优化调整侦察路径和目标分配策略。

6.2.5　知识蒸馏阶段

知识蒸馏作为技术落地的应用节点，通过简化中央模型，实现装备平台终端高效轻量化部署，让模型能够“瘦身”下沉战术末端。

（1）模型压缩。通过多头注意力头数裁剪、编码器层数深度压缩、模型参数量权重优化等方式，保留目标识别、数据分析等核心能力，去除冗余结构。

（2）特征迁移。在保留大模型目标分类概率分布的基础上，根据装备平台不同任务载荷分别迁移相应跨域多模融合特征，进一步约束学生模型学习。

（3）增量训练。结合实时采集的天气、海况、电磁环境、敌方目标信息等战场数据，在末端部署装备上进一步微调训练，确保在通信降级的复杂条件下轻量化模型也能运行流畅，实现快速响应，为战术终端提供决策支持。

7　结　论

随着人工智能、大数据技术的迅猛发展，大模型在军事领域的应用呈现出颠覆性潜力。以国产大模型（如DeepSeek 开源引擎）持续迭代优化为牵引，通过“预训练–监督微调–奖励建模–强化学习–知识蒸馏”的军事化开发流程，未来大模型必将深度融入全域作战链路。要进一步加强军事数据治理，建立完善动态更新机制，持续强化技术创新应用，努力突破数据质量低、算法“黑箱”、算力紧缺等瓶颈问题，实现完全自主可控，方能在新一轮军事变革中抢占战略主动权。

[参 考 文 献]

[1] 郭旺，杨雨森，吴华瑞，等. 农业大模型：关键技术、应用分析与发展方向[J]. 智慧农业（中英文），2024，6（2）：1-13.

[2] 崔翛龙，高志强，姬纬通，等. “艾武大模型+”：一种大模型军事化应用系统的开发与实证[J]. 数据采集与处理，2024，39（3）：588-597.

[3] 蔡磊，孟宪波，韩冬梅，等. 大模型在军事垂直领域的应用[J]. 舰船科学技术，2024，46（5）：171-175.

[4] 徐皮克. 面向军事知识问答的问题语义解析关键技术研究[D]. 长沙：国防科技大学，2021.

[5] Pfaff C A，Lowrance C J，Washburn B M，et al. Trusting AI：Integrating artificial intelligence into the army’s professional expert knowledge[J]. Journal of Defense Analytics，2023，15（3）：45-60.

[6] 刘涛，蒋国权，丁鲲，等. 基于大模型的事件抽取技术及军事应用思考[J]. 网络安全与数据治理，2023，42（S1）：163-168.

[7] 周中元，刘小毅，李清伟，等. ChatGPT技术及其对军事安全影响[J]. 指挥信息系统与技术，2023，14（2）：7-16.

[8] Liu H，Wang Y，Fan W，et al. Trustw-orthy AI：A computational perspective [J/OL]. 2021-07-12. http://arxiv.org/abs/2107.06641.

[9] Wei J，Bosma M，Zhao V Y，et al. Finetuned language models are zero-shot learners [J/OL]. 2021-09-03. http://arxiv.org/abs/2109.01652.

[10] 岳增营，叶霞，刘睿珩. 基于语言模型的预训练技术研究综述[J]. 中文信息学报，2021，35（9）：15-29.

[11] Jhong K Y. Evaluating artificial intelligence for operations in the information environment[J]. Defense Science Review，2023，12（4）：78-92.

[12] Liu P，Yuan W，Fu J，et al. Pre-train，prompt，and predict：A systematic survey of prompting methods in natural language processing[J/OL]. 2021-07-29. http://arxiv.org/abs/2107.13586.

[13] Khoshnoodi M，Jain V，GAO M，et al. A comprehensive survey of accelerated generation techniques in large language models [J/OL]. 2024-05-15. http://arxiv.org/abs/2405.13019.

[14] Christiano P，Leike J，Brown T B，et al. Deep reinforcement learning from human preferences [J/OL]. 2023-02-17. http://arxiv.org/abs/1706.03741.

[15] Schulman J，Wolski F，Dhariwal P，et al. Proximal policy optimization algorithms [J/OL]. 2017-07-20. http://arxiv.org/abs/1707.06347.

[16] 李铮，方琼，刘津玮，等. 2024年美军无人装备领域发展分析[J]. 战术导弹技术，2025（2）：1-12.

[17] 张龙，雷震，冯轩铭，等. 军事大模型：应用分析、关键技术和评估体系框架[C]. 第六届体系工程学术会议论文集——体系工程与高质量发展. 长沙：国防科技大学系统工程学院，2024：1149-1162.

[18] 赵日，赵鹏飞，程运江，等. 人工智能技术在反舰作战中的应用研究[J]. 战术导弹技术，2019（5）：86-91.

[19] 潘长鹏，韩玉龙，庄益夫. 舰载无人机编队协同对海突击作战效能评估指标体系研究[J]. 战术导弹技术，2019（2）：25-32.

[20] 张龙，王数，雷震，等. AIGC军事大模型评估体系框架研究[J]. 战术导弹技术，2025（1）：42-52.

（编辑：沈玉芃）