

文章编号: 2097-1974(2025)04-0045-08

DOI: 10.7654/j.issn.2097-1974.20250406

# 基于深度学习的博弈态势认知不确定度新质测量方法研究

王永光<sup>1,2</sup>, 孙静<sup>1,2</sup>, 张楠<sup>1</sup>, 张修建<sup>1,2</sup>

(1. 北京航天计量测试技术研究所, 北京, 100076;

2. 国家市场监督管理总局重点实验室(人工智能计量测试与标准), 北京, 100076)

**摘要:** 当前, 随着博弈环境复杂多变而引入深度学习模型如深度卷积神经网络, 以辅助提升人员对博弈态势的认知和决策水平。然而, 将深度学习引入博弈态势理解的同时, 也引入了人工智能中的数据不确定度和认知不确定度, 导致人工智能预测和决策结果存在发散性等问题。基于测量不确定度评定方法对博弈态势理解中的测量过程进行不确定度关键要素分解、提取及测量模型构建。试验结果表明, 基于GUM的物理测量方法能有效对博弈态势的认知不确定度进行准确高效的测量和评估。最后, 基于蒙特卡罗方法对提出的博弈态势认知不确定度新质测量方法进行验证, 表明提出方法的准确性和适用性。

**关键词:** 深度学习; 态势理解; 认知不确定度; 测量不确定度; 蒙特卡罗方法

中图分类号: V42; TP183

文献标识码: A

## A New Qualitative Measurement Method of Game Situation Cognitive Uncertainty Based on Deep Learning

WANG Yongguang<sup>1,2</sup>, SUN Jing<sup>1,2</sup>, ZHANG Nan<sup>1</sup>, ZHANG Xiujian<sup>1,2</sup>

(1. Beijing Aerospace Institute for Metrology and Measurement Technology, Beijing, 100076; 2. Key Laboratory of Artificial Intelligence Measurement and Standards for State Market Regulation, Beijing, 100076)

**Abstract:** At present, with the complex and changeable game environment, deep learning models such as deep convolutional neural networks are introduced to assist in improving personnel's cognition and decision-making level of the game situation. However, when deep learning is introduced into game situation understanding, it also introduces data uncertainty and cognitive uncertainty in artificial intelligence, which leads to problems such as divergence of artificial intelligence prediction results. Key elements of uncertainty in the measurement process of game situation understanding are decomposed, extracted and measurement modeling constructed based on the measurement uncertainty evaluation method. The experimental results show that the physical measurement method based on GUM can effectively measure and evaluate the cognitive uncertainty of game situation accurately and efficiently. Finally, based on Monte Carlo method, the proposed new qualitative measurement method of game situation cognition uncertainty is verified, which shows the accuracy and applicability of the proposed method.

**Keywords:** deep learning; situation understanding; cognitive uncertainty; measurement uncertainty; Monte Carlo method

## 0 引言

在博弈对抗领域, 博弈态势理解的认知能力水平是指挥决策的重要先决条件, 对博弈指挥人员作出正确决策具有决定性影响。由于博弈环境复杂, 对抗双方态势瞬息万变, 仅依靠指挥员的认知, 难以满足复杂多变的博弈需求。人工智能技术可用于快速提取和高效分析博弈态势中的复杂非线性因素。在获取博弈

初步信息数据后, 态势理解类智能算法对这些数据进行深度分析与推理, 构建对抗双方的博弈态势模型。该类智能算法运用数据融合、机器学习和推理模型, 预测蓝方行动、资源配置和潜在威胁, 助力决策者建立对博弈全局的认知, 有助于提升指挥决策人员对博弈态势的快速有效理解与认知<sup>[1-2]</sup>。

深度学习算法较强的非线性处理能力、自动分

析、记忆储存和逐层理解能力特性的网络结构,是研究人脑、语音识别、图像处理、视觉处理和自然语音处理的重要智能算法技术,将深度学习算法代替博弈指挥员以快速高效解决复杂态势信息分析任务,是解决博弈态势理解问题行之有效的办法,也是军事现代化的重要发展方向<sup>[3-7]</sup>。

然而,在博弈态势理解中,以认知智能中的卷积神经网络(Convolutional Neural Network, CNN)为代表的深度学习算法面临多种不确定性因素,这些不确定性可能会对军事行动和决策产生一定的影响。对基于深度学习的博弈态势认知不确定度进行测量和评估是非常重要的,可以帮助了解智能算法的可信度和稳定性。

## 1 人工智能不确定度测量技术研究

### 1.1 国内外人工智能不确定度测量技术研究现状

2022年1月,国务院发布的《计量发展规划(2021—2035年)》中指出,加强人工智能计量基础理论、评估方法和技术研究,开发用于评测人工智能系统性能的参考数据集;研究智能基础设施计量测试技术,形成各领域通用的人工智能计量体系框架、接口与方法、标准规范;建立适用于各领域智能水平评价标准和计量测试平台,提升数据和知识协同驱动的计量测试能力。

1993年,国际计量局及国际标准化组织等七个国际组织联合公布了《Guide to the Expression of Uncertainty in Measurement, 1993》,即《测量不确定度表示指南》,其中所用方法简称“GUM”<sup>[8]</sup>,中国对测量结果的表述与国际一致。

2019年1月,中国发布最新的计量技术规范JJF1059—2019《测量不确定度评定与表示》,包括两个部分:a) JJF1059.1—2019《测量不确定度评定与表示》;b) JJF1059.2—2019《用蒙特卡罗法评定测量不确定度》,该规范规定了用蒙特卡罗法(Monte Carlo Method, MCM)评定与表示测量不确定度的方法,其核心内容是在建立测量模型的基础上采用MCM进行概率分布传播,本规范适用于具有多个输入量和单一输出量的测量模型。

人工智能不确定度评估方面,文献[9]提出不确定度量化不仅有助于理解神经网络的行为,而且能增强对其预测结果的可信度。文中描述了一种为深度神经网络中的不确定性量化创建基准测试问题的方法,为了对当前已有的不确定度估计方法的质

量进行评估,还提出了一个新的框架,该框架中可以设计对问题复杂性和输入维度明确控制的通用数据集,以对深度回归中的不确定性量化进行基准测试,并演示了如何利用生成模型在其参数中的线性来创建一个锚模型(Anchor Model),该锚模型有助于判断预测结果的质量和量化深度学习方法的不确定度。

文献[10]自动测试系统ATS主要用于航空航天等系统的使用、诊断、维护等过程,并对该过程中发挥关键作用的电子测量等复杂设备与系统可靠运行提供必要保证。ATS测量结果的准确与否直接影响系统的安全和效能发挥,ATS通过GUM和蒙特卡罗仿真进行各传递单元静态测量、动态测量及测量链静态测量不确定度合成等工作,已经取得不少研究成果。

### 1.2 国内外研究现状面临的挑战

国内外对于通用人工智能的测量不确定度研究均处于起步阶段。现有研究面临的挑战分析如下:

a) 关键要素的分解不清晰。面向复杂态势的认知智能算法的测量不确定度关键要素提取不全面,关键要素之间的关系不清晰,例如测试数据不确定度的测试需求覆盖性等。

b) 测量不确定度的方法不规范。现有测量数据或模型不确定度的方法不统一、不规范、缺乏参考基准,未能形成全面、系统、规范的智能算法测量不确定度评估方法。

c) 测量不确定度的客观有效指标体系不完备。现有测量不确定度指标与态势理解类智能算法的理解逻辑和能力边界的联系不够紧密,测量不确定度的指标难以仅通过获得不确定度的大小值来表征态势理解智能算法的理解能力边界。

d) 测量不确定度缺乏计量技术规范的指导。现有的态势理解智能算法测量不确定度主要以准确性应用需求为导向,缺乏客观的A类评定和基于先验知识的B类评定的测量不确定度方法指导,降低了测试的可信性。

## 2 博弈态势认知算法不确定度影响因素分解与表征

在基于深度学习的博弈态势理解与认知的过程中,博弈态势的理解和认知水平关键受到两个方面的影响:a) 数据的不确定度(主要是由博弈双方的人员部署、装备、地理环境、天气等条件的随机变化与发展引起)<sup>[1]</sup>;b) 模型认知的不确定度(主要指智能

模型的认知水平有限，导致对各个方向上博弈态势的预测概率存在波动和随机性)<sup>[11-12]</sup>。

本文基于物理域的JJF1059.1—2019《测量不确定度评定与表示》所述方法，将基于深度学习的博弈态势认知测量不确定度标准化流程定义为如图1所示。具体步骤：a) 分析基于深度学习的博弈态势智能算法不确定度来源并建立测量模型；b) 进行数据分量方差与灵敏系数计算；c) 计算各分量标准不确定度；d) 计算合成标准不确定度；e) 确定扩展不确定度；f) 报告测量不确定度评估结果。

数据不确定度主要来源于蓝方人员部署、对方的隐蔽行动以及环境的随机变化等。这些不确定因素可

能导致态势感知系统无法获取完整、准确的信息，进而影响对博弈的理解与判断。

态势认知智能算法的不确定度主要来自模型在推理过程中使用的随机采样技术。由于态势理解涉及对复杂动态环境的感知和预测，算法往往依赖概率模型或基于样本的统计方法来应对信息的不完备性和随机性。随机采样通过生成多个可能的情景或未来态势进行分析，但这种方法不可避免地引入了采样偏差和模型不精确性。模型参数的选择、采样规模的限制以及采样过程中的随机波动，都会导致预测结果的不稳定性。这种不确定度使得态势理解系统可能无法充分捕捉环境中的微妙变化，影响决策的精确性。

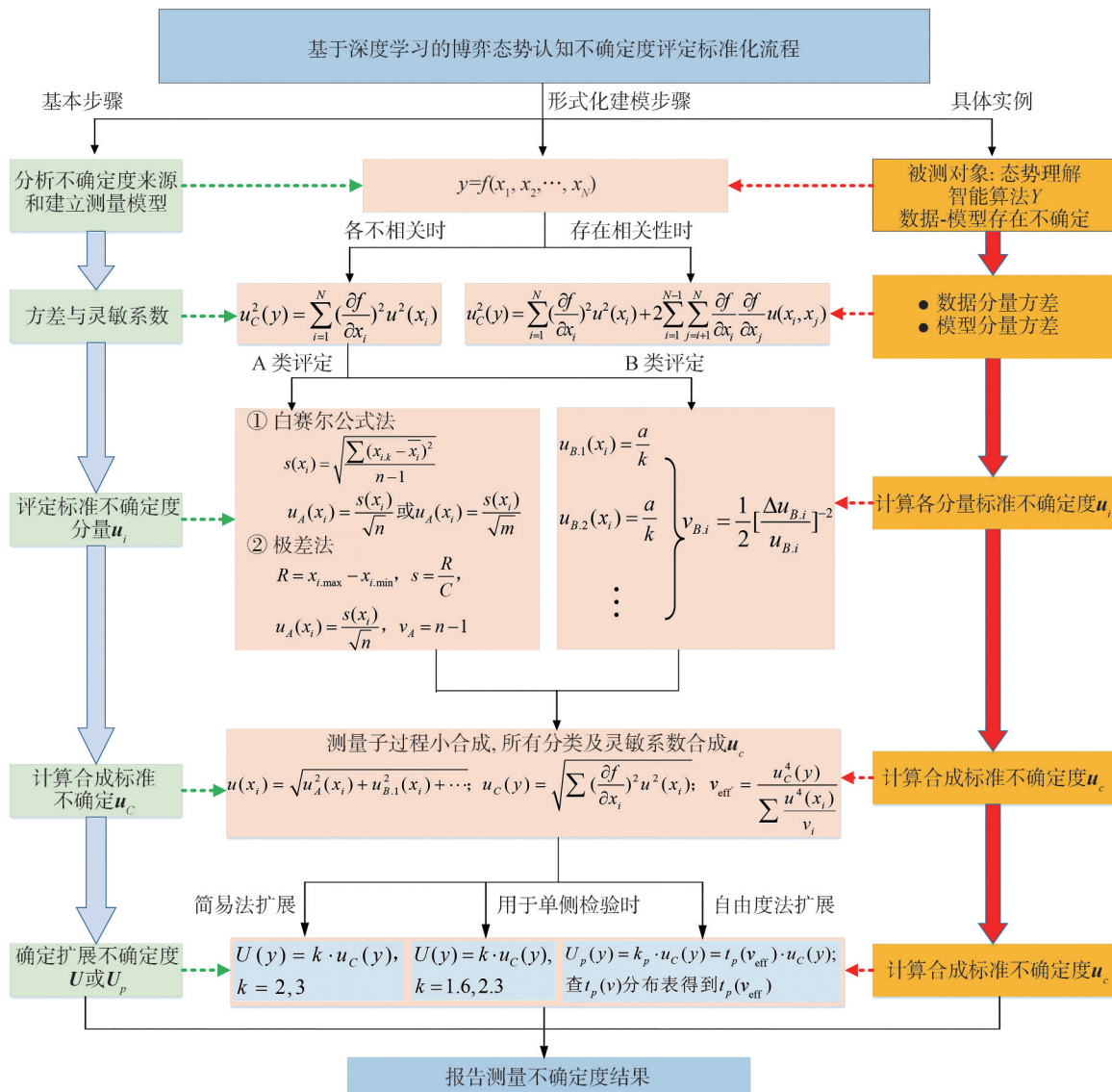


图1 博弈态势认知不确定度测量标准化流程

Fig.1 A standardized process for measuring the uncertainty of game situation cognition

## 2.1 数据的不确定度影响因素分解及表征

在本研究中,通过模拟博弈场景,对红方重要目标位置、蓝方集结区域及蓝方主要博弈方向展开深入分析。模拟场景设定为:对抗双方各博弈单元在南北30 km宽、东西30 km宽的博弈平面上呈地理分布。构建的博弈态势图像矩阵由30×30个元素组成,每个元素作为图像像素,取值范围为[0, 255]内的整数,且每个矩阵元素对应着不同状态。

这些状态通过对博弈尺寸、前线宽度、单位数量以及基准点位置等参数的精确控制来生成。具体而言,博弈中心的前线宽度设定为5,红方与蓝方的单位分别随机分布于左右两侧,双方各拥有30个单位。基准点分别位于博弈区域的左侧和右侧,通过设置相应参数实现对其位置的精准调控,进而为深入研究蓝方的博弈意图及行动企图奠定基础。

在模拟博弈的构建中,单位分布基于正态分布随机生成,以此确保单位能够沿博弈中线呈现合理分布态势。对于基准点布局,通过精心设计的算法避免与其他单位发生重叠,同时在大本营附近促成更多单位的战略集中,这种布局方式有利于形成具有战略意义的力量集聚点。在确定基准点之间的角度关系时,通过精确计算两点的坐标差,得出能够反映空间位置关系的角度值,该角度值在很大程度上决定了双方基准点的相对位置,为后续的态势分析提供了关键的空间位置信息。

动作概率的生成借助训练有素的算法,模型会对博弈场景中的诸多因素进行综合考量,进而计算出每个动作的概率分布。这些概率数值不仅体现了模型对于各个方向决策的信心程度,更是后续深入分析不确定性的关键因变量,通过对动作概率的持续跟踪与分析,能够为博弈策略的制定与调整提供有力支持。当完成对各个博弈目标在位置、身份以及关联关系等方面的估计,并实现数据的一致性融合后,获取到的丰富博弈目标信息需要进一步处理。结合态势理解领域的专业知识以及指挥员积累的宝贵经验,经过深入思考,对各个博弈目标进行抽象化处理,将其聚合成具有特定意义的博弈单元。从空间分布角度出发,对这些博弈单元展开系统研究,并将其提炼为图像矩阵形式。在该图像矩阵中,每个元素均对应于图像的像素,通过对不同像素值的明确区分,得以清晰辨别各方的博弈单元以及背景图像,为从可视化角度洞察博弈态势提供了直观且有效的途径。

在图像矩阵模拟中,构建蓝方主力集结区域具有重要意义。通常,蓝方以某一特定点作为集结中心开展行动,该集结中心与各个博弈实体间的距离应维持在较小水平,如此方能为通信与指挥工作提供便利。通过对图像像素点密度的精细调控,可间接实现蓝方主力集结区域的生成。此过程中,密度较高的像素点区域即代表蓝方主力的主要分布范围,这种模拟方式能够有效反映蓝方力量集中的态势。

红方重要区域的生成亦遵循类似原理。红方重要区域作为高价值目标,往往围绕某核心点布局,在该中心点周围,红方人员呈现出较大密度的分布状态,以此保障红方重要区域的安全。同样通过控制像素点密度的手段,能够模拟出红方重要区域的范围与特征,凸显其关键地位。前线区域作为红方与蓝方冲突和博弈的核心地带,位于双方区域的中间位置。在该区域中,需要有较为密集的人员排布以体现双方的对峙状态。通过对前线区域像素点密度的灵活控制,可以直观地模拟博弈的激烈程度以及双方势力的分布情况。像素点密度高的区域,表明该部分博弈激烈,双方投入的力量较多;而像素点密度较低的区域,则反映出相对缓和的局势。

数据集包含15 000个训练样本和5 000个测试样本,每个数据样本包含一个30×30的博弈矩阵、基准点位置、角度差以及模型输出的动作概率。每个矩阵对应一个博弈场景。矩阵中的元素分别表示不同的博弈对象,其中:

- a) 数值0被设定为博弈背景标识,用以构建场景的基础框架;
- b) 数值1代表红方小队,表示红方基层单元;
- c) 数值2代表红方大本营,象征红方核心战略据点;
- d) 数值-1代表蓝方小队,表示蓝方基层单元;
- e) 数值-2代表蓝方大本营,象征蓝方核心战略据点。

依据上述规则所生成的博弈图像矩阵如图2所示,其中红色像素代表一个单位的红方人员,红色圆圈代表红方重要区域;蓝色像素代表一个单位的蓝方人员,蓝色圆圈代表蓝方人员的集结中心;黄色箭头代表蓝方主力的进攻方向。图2展示了不同的博弈模拟布局,其中红方和蓝方的基层单元以及战略据点位置随场景变化。

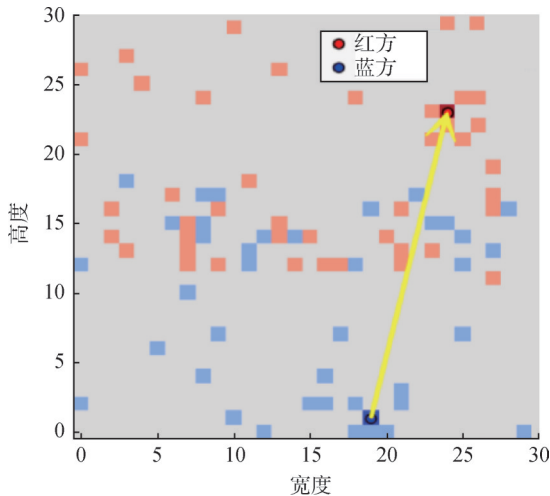


图2 博弈态势图像矩阵

Fig.2 Game situation image matrix diagram

在这个场景中，矩阵中的红方战略据点位于博弈的右上角，蓝方战略据点位于下方中央区域。蓝方的进攻方向由黄色箭头标示，从下方右上进攻。博弈算法需要根据这一进攻方向调整防守策略，重点防守从下方进攻的蓝方单位。

此外，为了提高数据多样性，生成的博弈矩阵通过一定角度的旋转进行数据增强，确保模型能够适应不同视角下的战斗场景。

下一步标注生成的矩阵，以蓝方人员集结中心为原点，建立极坐标系如图3所示。

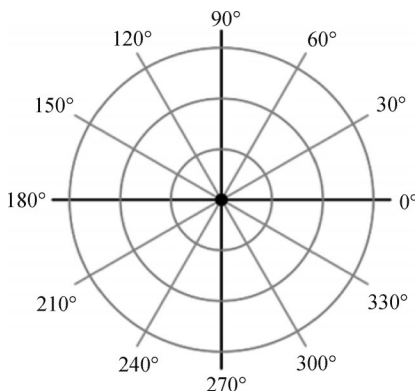


图3 极坐标系

Fig.3 Polar coordinate diagram

将蓝方的进攻方向，转化成东、南、西、北、东北、西北、东南、西南八个方向，八个方向对应八个分类，以数字0到7表示。标签用以表示蓝方的进攻方向，进攻方向由基准点之间的相对位置决定。参数定义如表1所示。

表1 博弈态势图像矩阵标签

Tab.1 Game situation picture matrix label

极坐标角度 $a/(^\circ)$	方向	标签
$-22.5 \leq a < 22.5$	东	0
$22.5 \leq a < 67.5$	东北	1
$67.5 \leq a < 112.5$	北	2
$112.5 \leq a < 157.5$	西北	3
$157.5 \leq a < 202.5$	西	4
$202.5 \leq a < 247.5$	西南	5
$247.5 \leq a < 292.5$	南	6
$292.5 \leq a < 337.5$	东南	7

根据上述规则生成态势理解类算法的数据，其不确定度主要来源于测试数据生成的随机性。由于博弈图像矩阵会以一定的概率密度随机分布于博弈上，博弈每一个位置存在博弈单位的概率并不相同，红方重要区域与蓝方集结区域的半径也不相同，蓝方的主攻方向受到以上因素的影响，会有一些的随机性，最终导致不确定度。使用GUM法统计测试数据中的不同标签下的概率，进而通过计算均值和标准差，得到数据不确定度。

为了分析数据不确定度来源，通过热力图展示了各个测试场景中不同方向的预测概率分布，如图4所示。其中，颜色的深浅反映了模型对各个方向的预测置信度，图中显示出模型在某些场景下对特定方向的预测较为可信，而在其他场景中则存在一定的不确定性。红色区域表示模型对这些方向的预测置信度较高，而蓝色区域则表示置信度较低。通过分析这些高置信度区域和低置信度区域，可以进一步理解模型的行为，并找出可能导致不确定性的场景特征。

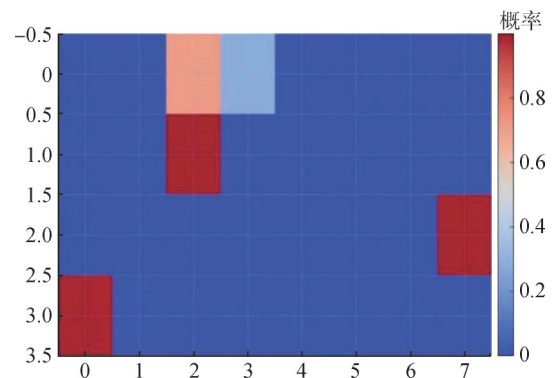


图4 不同方向上预测概率的热力图

Fig.4 Heat maps of predicted probabilities in different directions

通过对态势理解类智能算法进行不同角度差异和预测概率的分析，探索了模型预测不确定性的来源，

并通过试验验证不同方向和角度下的模型表现。

试验验证表明,模型在较小角度差异时表现更为稳定,较大角度差异时预测不确定性有所增加。模型对某些方向的预测存在较大波动,这是由训练数据中这些方向样本的不均衡或其他因素造成的。模型在不同测试场景下的预测置信度呈现出一定的波动,表明不确定性来源于输入数据中的特定特征,尤其是角度差异的大小。

在标签0~7中,为每个标签建立一个数列,记为 $r_i(i=0, 1, \dots, 7)$ ,设某一个矩阵的预测结果中,标签 $i$ 对应的分类概率为 $p_i(i=0, 1, \dots, 7)$ 。一般而言, $p_i$ 中最大的结果即为最终分类预测结果,在数值上会显著高于其余7个分类。将 $p_i$ 的最大值加入到数列 $r_i$ 中。假设测试数据集中有 $n$ 个样本,每个标签有 $n$ 个样本,则每个数列也包含有 $n$ 个样本数据。统计每个数列 $r_i$ 的均值,记为 $m_i(i=0, 1, \dots, 7)$ ,则数据的不确定度可表示为

$$u_d = \frac{s_{id}(m)}{\sqrt{n}} \quad (1)$$

式中  $s_{id}(m)$ 为均值 $m$ 的标准差,计算公式为

$$s_{id}(m) = \sqrt{\frac{1}{n} \sum_{i=1}^n (r_i - \bar{r})^2} \quad (2)$$

## 2.2 模型的认知不确定度影响因素分解及量化表征

作为一个分类问题,态势理解类智能算法的神经网络在预测时会给出每个分类下的概率。对数据集中的矩阵进行预测,将每一次预测结果的最高概率进行数据收集,构建概率数列,对数列进行统计学验证,以获得不同分类的方差和均值,进而借助GUM法求得模型的认知不确定度。

模型使用Softmax函数作为多分类模型的输出层,以生成每个类的预测概率。对于一个模型的输出,Softmax函数将每个输出转换为一个概率值,其公式为

$$p_i = \frac{e^{z_i}}{\sum_{j=1}^n e^{z_j}} \quad (3)$$

在不确定度分析中,Softmax输出概率值反映了模型对每个类别(动作方向或博弈意图)的预测置信度。

箱线图展示了模型在不同方向上的认知预测概率波动情况,如图5所示。模型在所有方向上的预测概率中位数较为一致,但在每个方向上仍然存在一定程度的波动,特别是在方向0、2和5处,置信区间较

大,表明这些方向上模型的认知不确定性较高。与此同时,部分方向(如方向1和3)的预测概率集中在较高的区间,表明模型对这些方向的认知更为稳定和可靠。由此可以得出结论,模型在应对某些方向时存在更大的认知不确定性,而在其他方向上则表现较为稳定。

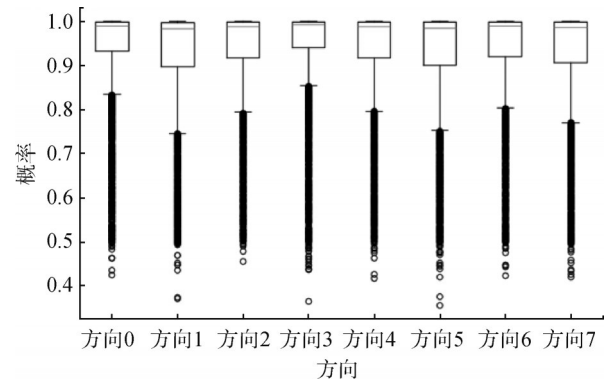


图5 博弈态势的认知不确定度

Fig.5 Cognitive uncertainty analysis diagram of game situation

对于一个博弈图像矩阵,神经网络会给出每个标签的分类概率,将标签0~7的分类概率,记为 $z_i(i=0, 1, \dots, 7)$ 。假设数据集中共有 $n$ 个图像矩阵,提取出一次预测下分类概率最高的值,将其设置为一个长度 $n$ 的数列 $r$ ,则数据的不确定度可以表示为

$$u_m = \frac{s_{id}(r)}{\sqrt{n}} \quad (4)$$

式中  $s_{id}(r)$ 为数列 $r$ 的标准差,计算公式为

$$s_{id}(r) = \sqrt{\frac{1}{n} \sum_{i=1}^n (r_i - \bar{r})^2} \quad (5)$$

## 2.3 不确定度测量模型函数关系构建

对基于深度学习的博弈态势认知不确定度进行函数建模,使用线性函数和非线性函数两种方式进行建模。

基于物理域的GUM法A类评定如式(6)所示,其中 $u_d$ 为数据的不确定度、 $u_p$ 为平台的不确定度(试验结果表明平台的不确定度对预测结果影响很小,这里加入是考虑试验的充分性)。 $u_p$ 通过大量试验使用极差法获得不同平台(Windows和Ubuntu系统)之间的统计学结果差异,获得平台不确定度的分量。

### a) 线性函数关系

使用线性函数关系式,对不确定度影响要素进行建模:

$$u = a_1 \cdot u_d + a_2 \cdot u_m + a_3 \cdot u_p \quad (6)$$

式中  $a_1, a_2, a_3$  分别为各分量不确定度对合成不确定度的影响程度。由于3个分量线性无关且对合成不确定度的影响程度相同，因此函数系数设置为常数1。

b) 非线性函数关系

对于态势理解类智能算法，其数据遵循特定的生成规则，基于 Python 批量模拟生成数据，数据的原始格式为矩阵格式，矩阵中的数值为0~7的整数，数据的值与模型无关。对于算法模型，其深度卷积神经网络在训练完成后参数保持固定，因此不会受到数据的影响。因此，态势理解类的数据和模型相互独立。

当数据、模型、平台相互独立时，非线性函数关系会退化为线性函数关系。

### 3 博弈态势认知测量不确定度评估与验证

#### 3.1 测量不确定度评估

对于数据不确定度，将输出概率使用 Softmax 归一化，模拟计算数据不确定度为

$$u_d = \frac{s(x)}{\sqrt{n}} = 0.0062 \quad (7)$$

对于模型不确定度，使用 A 类评定法，设计取样次数为6 000次，计算结果如下：

$$u_m = \frac{s(x)}{\sqrt{n}} = 0.0015 \quad (8)$$

对于平台不确定度，使用 A 类评定法，通过计算大量样本在不同平台之间的相对不确定度，推导出平台绝对不确定度为

$$u_p = z_u \cdot u_R(D) = 0.0007 \quad (9)$$

式中  $z_u$  为1； $u_R$  为0.0007。

基于式 (7) ~ (9) 得到的数据不确定度  $u_d$ 、模型不确定度  $u_m$  和平台不确定度  $u_p$  以及式 (6) 得到合成不确定度结果为

$$u_c^2 = u_d^2 + u_m^2 + u_p^2 = 0.0064 \quad (10)$$

因此，态势理解类算法的合成不确定度为0.0064。

#### 3.2 测量不确定度评估验证

构建博弈态势理解认知不确定度的验证指标及初始值如表2所示。

表2 测量不确定度验证指标及初始值

Tab.2 Measurement uncertainty verification index and initial value

验证指标	数值
概率 $Y$ 的估计值 $y$	0.9164
标准不确定度 $u(y)$	0.0065
扩展不确定度 $U$	0.0129
样本量大小 $M$	$10^6$
数值容差 $\delta$	$5 \times 10^{-4}$

基于蒙特卡罗方法对 GUM 法的不确定度进行验证<sup>[13-15]</sup>，具体步骤如下：

a) 定义被评估验证对象，被评估验证对象为博弈态势理解的认知不确定度。

b) 确定输入量及其输入量概率密度函数，各个输入量及其概率分布如表3所示。

表3 博弈态势算法蒙特卡罗输入量及其概率密度函数

Tab.3 Monte Carlo input of game situation algorithm and its probability density function

输入量 $X$	分布类型	标准不确定度
样本数量差	高斯分布	0.0022
样本标准差	高斯分布	0.0036
样本均值	高斯分布	0.0048
模型采样标准差	高斯分布	0.0012

c) 建立被测量  $Y$  与  $X$  的测量模型，建立系数为1的线性测量模型。

d) 选择蒙特卡罗试验样本量的数  $M$ ，即测量模型计算的次数。 $M=10^6$  通常会为输出量提供包含概率  $p$  为95%的包含区间。

e) 输入量概率密度函数 PDF 的传递。在 MCM 的实现过程中，从  $N$  个输入量  $X$  的 PDF 中抽取  $M$  个样本值，代入测量模型，计算模型在每个样本输出的值。将 MCM 得到的模型值按递增次序排列，得到输出量的 PDF。输出量服从高斯分布：

$$y_{low} = \mu - \sigma = 0.9032$$

$$y_{high} = \mu + \sigma = 0.9290$$

f) 验证 GUM 法测量结果。对 GUM 法和 MCM 获得的包含区间进行比较，确定是否能获得 GUM 法提供的包含区间中正确十进制数字的所需位数，确定：

$$d_{low} = |(y - U) - y_{low}| = 0.0003$$

$$d_{high} = |(y + U) - y_{high}| = 0.0003$$

即两个包含区间的各自端点差的绝对偏差  $d_{low} \leq \delta$ ， $d_{high} \leq \delta$ ，通过验证。

## 4 结论

本文分析了面向博弈态势理解的深度卷积神经网络中存在的认知不确定度, 并对其进行了测量评估与验证。通过物理域GUM法的不确定度关键要素分解及新质测量模型构建, 进行认知不确定度测量, 并使用蒙特卡罗法对测量结果进行验证。

由上述试验及不确定度测量结果可知, 通过经典MCM法的数理统计分析验证了基于物理域GUM法对博弈态势的认知不确定度新质测量建模方法的准确性; 同时, 相比于传统极其耗时的MCM方法, 本文提出的测量模型极大地提高了不确定度评估效率, 本文基于计量学理论对认知模型存在的不确定度进行评估建模分析, 为其在智能装备中的安全可靠性能试验鉴定提供理论支撑和实际部署应用发挥重要指导作用。

## 参考文献

- [1] 朱丰, 胡晓峰, 吴琳, 等. “深度学习”在军事领域及态势感知评估中的应用研究与展望[C]. 北京: 第七届全军“战争复杂性与信息化战争模拟”高层学术研讨会, 2016.  
ZHU Feng, HU Xiaofeng, WU Lin, et al. Research and prospects on the application of "deep learning" in military domain and situation awareness evaluation[C]. Beijing: Proceedings of the 7th Whole-Army High-Level Academic Symposium on War Complexity and Informationized War Simulation, 2016.
- [2] 朱丰, 胡晓峰, 吴琳, 等. 基于深度学习的战场态势高级理解模拟方法[J]. 火力与指挥控制, 2018, 43(8): 25-30.  
ZHU Feng, HU Xiaofeng, WU Lin, et al. Simulation method of battlefields situation senior comprehension based on deep learning[J]. Fire Control & Command Control, 2018, 43(8): 25-30.
- [3] MCCULLOCH W S, PITTS W. A logical calculus of the ideas immanent in nervous activity[J]. The Bulletin of Mathematical Biophysics, 1943, 5(4): 115-133.
- [4] HEBB D O. The organization of behavior[M]. Cambridge: MIT Press, 1988.
- [5] RUMELHART D E, HINTON G E, WILLIAMS R J. Learning internal representations by error propagation, parallel distributed processing, explorations in the microstructure of cognition[J]. Biometrika, 1986, 71(599-607): 6.
- [6] HINTON G E, OSINDERO S, TEH Y W. A fast learning algorithm for deep belief nets[J]. Neural Computation, 2006, 18(7): 1527-1554.
- [7] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[C]. Lake Tahoe: Advances in Neural Information Processing Systems, 2012.
- [8] ISO I, OIML B. Guide to the expression of uncertainty in measurement[M]. Madrid: Aenor, 1993.
- [9] SCHMÄHLING F, MARTIN J, ELSTER C. A framework for benchmarking uncertainty in deep regression[J]. Applied Intelligence, 2023, 53(8): 9499-9512.
- [10] 丁畅. 军用ATS测量不确定度评定研究[D]. 长沙: 国防科学技术大学, 2009.  
DING Chang. Research on measurement uncertainty evaluation for military automatic test system[D]. Changsha: National University of Defense Technology, 2009.
- [11] 曹兰英, 郭明明, 罗美方. 雷达与电子战的认知博弈[J]. 雷达科学与技术, 2021, 19(5): 552-557.  
CAO Lanying, GUO Mingming, LUO Meifang. Cognitive game between radar and electronic countermeasure[J]. Radar Science and Technology, 2021, 19(5): 552-557.
- [12] 曾丽娇. 基于攻防演化博弈的网络安全态势研究[D]. 西安: 西安电子科技大学, 2018.  
ZENG Lijiao. Research on network security situation based on attack-defense evolutionary game[D]. Xi'an: Xidian University, 2018.
- [13] 曹芸, 陈怀艳, 韩洁. 采用MCM对GUM法测量不确定度评定的验证方法研究[J]. 宇航计测技术, 2012, 32(2): 75-78.  
CAO Yun, CHEN Huaiyan, HAN Jie. Verification method for measurement uncertainty evaluation of GUM method using MCM[J]. Journal of Astronautic Metrology and Measurement, 2012, 32(2): 75-78.
- [14] 刘园园, 杨健, 赵希勇, 等. GUM法和MCM法评定测量不确定度对比分析[J]. 计量学报, 2018, 39(1): 135-139.  
LIU Yuanyuan, YANG Jian, ZHAO Xiyong, et al. Comparative analysis of measurement uncertainty evaluation by GUM and MCM methods[J]. Acta Metrologica Sinica, 2018, 39(1): 135-139.
- [15] 马福强, 刘彦森, 杨学猛, 等. 基于GUM法测试不确定度评价方法研究[C]. 大连: 2016年中国造船工程学会水中目标特性学组学术交流大会, 2016.  
MA Fuqiang, LIU Yansen, YANG Xuemeng, et al. Research on test uncertainty evaluation method based on GUM method[C]. Dalian: Proceedings of 2016 Academic Exchange Conference of Underwater Target Characteristics Group of Chinese Society of Naval Architects and Marine Engineers, 2016.

## 作者简介

王永光 (1987—), 男, 博士, 工程师, 主要研究方向为人工智能计量与测试技术、深度学习不确定性估计技术、人工智能安全可信测评技术等。

孙静 (1993—), 女, 工程师, 主要研究方向为人工智能计量与测试技术、智能特性与安全可信测评技术、人工智能测评体系建设与标准规范等。

张楠 (1990—), 女, 工程师, 主要研究方向为软件工程、人工智能与机器人、人工智能计量与测试技术等。

张修建 (1984—), 男, 研究员, 主要研究方向为人工智能计量与测试技术、对抗安全性测评技术、人工智能测评规范等。