ICS 35.240
CCS L 70

中华人民共和国国家标准

GB/T 45288.4—2026

人工智能　大模型
第4部分：计算机视觉大模型

Artificial intelligence—Large-scale model—
Part 4: Computer vision large-scale model

2026-09-28 发布　　2027-01-01 实施

国家市场监督管理总局
国家标准化管理委员会　发布

目　　次

前言……Ⅲ
引言……Ⅳ
1　范围……1
2　规范性引用文件……1
3　术语和定义……1
4　缩略语……2
5　参考框架……2
6　技术要求……3
6.1　视觉理解……3
6.2　视觉分析……4
7　测评方法……5
7.1　通则……5
7.2　测评流程……6
7.3　测评数据……6
7.4　测评组……6
7.5　指标计算方法……7
附录A（资料性）　多模态视觉大模型……8
A.1　视觉交互……8
A.2　视觉生成……9
A.3　测评结果计算方法……11
附录B（资料性）　专用视觉模型……12
B.1　专用测量……12
B.2　专用识别……12
B.3　测评结果计算方法……13
附录C（规范性）　测评指标计算方法……14
C.1　客观测评指标……14
C.2　主观测评指标……17
参考文献……19

前　言

本文件按照GB/T 1.1—2020《标准化工作导则　第1部分：标准化文件的结构和起草规则》的规定起草。

本文件是GB/T(Z) 45288《人工智能　大模型》的第4部分。GB/T(Z) 45288已经发布了以下部分：

——第1部分：通用要求；

——第2部分：评测指标与方法；

——第3部分：服务能力成熟度评估；

——第4部分：计算机视觉大模型；

——第5部分：多模态大模型；

——第7部分：语音大模型。

请注意本文件的某些内容可能涉及专利。本文件的发布机构不承担识别专利的责任。

本文件由全国信息技术标准化技术委员会(SAC/TC 28)提出并归口。

本文件起草单位：中国电子技术标准化研究院、华为云计算技术有限公司、浙江大学、北京智源人工智能研究院、浦江国家实验室、北京航空航天大学、科大讯飞股份有限公司、蚂蚁科技集团股份有限公司、北京智谱华章科技股份有限公司、上海商汤智能科技有限公司、阿里云计算有限公司、北京百度网讯科技有限公司、清华大学、启元实验室、北京交通大学、北京邮电大学、中国移动通信集团有限公司研究院、中国电信股份有限公司北京研究院、中国电力科学研究院有限公司、中国电信股份有限公司上海研究院、国能数智科技开发(北京)有限公司、南方电网数字电网集团有限公司、海信视像科技股份有限公司、杭州海康威视数字技术股份有限公司、杭州电子科技大学、浙江大华技术股份有限公司、中国移动通信集团有限公司、上海计算机软件技术开发中心、中电信数智科技有限公司、深圳思谋信息科技有限公司、中国工商银行股份有限公司、哈尔滨工业大学、北京市公安局人工智能安全研究中心、浙江理工大学、国投智能信息科技股份有限公司、同方知网数字科技有限公司、北京格灵深瞳信息技术股份有限公司、浪潮通信信息系统有限公司、青岛港国际股份有限公司、中国科学院软件研究所、中国科学院沈阳自动化研究所、中国南方电网有限责任公司超高压输电公司、中关村智联软件服务业质量创新联盟、安徽大学、中兴通讯股份有限公司、中央民族大学、中移(杭州)信息技术有限公司、中移(苏州)软件技术有限公司、中移雄安信息通信科技有限公司、中国航天科技体系与创新研究院。

本文件主要起草人：范科峰、郑子木、郑若琳、吴飞、刘颖、蒋慧、徐洋、刘祥龙、李建欣、陶建华、王志栋、鲍薇、王浩然、况琨、董建、殷兵、孙曦、冯涛、吴庚、吴涛、于佃海、张圣宇、王永威、杨熙、石羡、刘艾杉、王嘉凯、汪群博、左兴权、杨乐、曹晓琦、刘伟东、王峰、陈晰、宋光磊、阎岩、韩韬、杨旭、李晋伟、张学聪、李晨、黄现翠、李亚玲、卢顺、高东辉、焦若丹、刘卫卫、梁寿愚、朱春波、任文奇、李孟择、付建海、曹汐、陈敏刚、陈利明、杨晓诚、李哲欣、李海峰、郭影起、佟威、阙锦龙、李达、冯子勇、梁秉豪、吴宇震、郭乙运、孟令中、王鑫、王宁、吕飞燕、范存航、赵孝武、赵小兵、谢小燕、李旭东、郑庆国。

引　言

大模型已成为人工智能发展的重要技术手段，在引领产业变革方面发挥着重要作用，国内外各企业和研究机构相继开发百余种大模型产品和评测榜单，导致用户难以有效评价业内产品的技术水平和服务能力。GB/T(Z) 45288《人工智能　大模型》旨在规定通用大模型的技术要求、评测指标和服务能力，拟由九个部分构成。

——第1部分：通用要求。目的在于确立大模型的参考架构，规定通用技术要求。

——第2部分：评测指标与方法。目的在于确立大模型的评测指标，描述评测方法。

——第3部分：服务能力成熟度评估。目的在于给出大模型服务能力成熟度等级及评估方法。

——第4部分：计算机视觉大模型。目的在于确立计算机视觉大模型的参考框架，规定技术要求，描述测试方法。

——第5部分：多模态大模型。目的在于规定多模态大模型的参考架构，规定技术要求。

——第6部分：代码大模型。目的在于定义代码大模型的概念和功能，规定技术要求和测试方法。

——第7部分：语音大模型。目的在于定义语音大模型的概念和功能，规定技术要求和测试方法。

——第8部分：时间序列大模型。目的在于定义时间序列大模型的概念和功能，规定技术要求和测试方法。

——第9部分：图大模型。目的在于定义图大模型的概念和功能，规定技术要求和测试方法。

人工智能　大模型
第4部分:计算机视觉大模型

1　范围

本文件确立了计算机视觉大模型的参考框架,规定了相关技术要求,描述了对应的测试指标和方法。

本文件适用于计算机视觉大模型的设计和测试。

2　规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中,注日期的引用文件,仅该日期对应的版本适用于本文件;不注日期的引用文件,其最新版本(包括所有的修改单)适用于本文件。

GB/T 41864—2022　信息技术　计算机视觉　术语

GB/T 41867—2022　信息技术　人工智能　术语

GB/T 45087—2024　人工智能　服务器系统性能测试方法

GB/T 45288.1—2025　人工智能　大模型　第1部分:通用要求

GB/T 45288.2—2025　人工智能　大模型　第2部分:评测指标与方法

GB/T 45288.5—2026　人工智能　大模型　第5部分:多模态大模型

3　术语和定义

GB/T 41864—2022、GB/T 41867—2022、GB/T 45288.1—2025界定的以及下列术语和定义适用于本文件。

3.1

标签　label

分配到样本的目标变量。

[来源:ISO/IEC 22989:2022,3.2.10]

3.2

异常　anomaly

偏离预期的分布。

注1:异常通常体现为离群数据。

注2:异常类型包括小区域异常、大范围区域异常、序列异常。其中序列异常指视频中时间维度的连续异常。

3.3

计算机视觉大模型　computer vision large-scale model

以视觉数据作为一种输入,用于实现视觉任务的大模型。

注1:视觉数据包含图像(含3D图像、点云、红外、高光谱数据等)或视频(含3D视频)。

注2:按技术机制与应用范畴分为通用视觉大模型、多模态视觉大模型和专用视觉大模型。

注3:计算机视觉大模型简称视觉大模型。