浮游植物监测 人工智能图像识别技术指南

编制说明

《浮游植物监测 人工智能图像识别技术指南》标准编制组

二零二五年一月

目录

[1 项目背景 3](#_Toc182763189)

[1.1 任务来源 3](#_Toc182763190)

[1.2 工作过程 3](#_Toc182763191)

[2 制定标准的必要性 5](#_Toc182763192)

[2.1 标准制定的背景 5](#_Toc182763193)

[2.2 标准制定的必要性 5](#_Toc182763194)

[2.3 标准制定的目的 6](#_Toc182763195)

[2.4 国内外相关技术发展动态 7](#_Toc182763196)

[2.5 与现行有关法律法规、强制性标准和其他有关标准的关系 8](#_Toc182763197)

[2.6 现有工作基础 9](#_Toc182763198)

[3 标准编制依据和原则 9](#_Toc182763199)

[3.1 编制依据 9](#_Toc182763200)

[3.2 编制原则 10](#_Toc182763201)

[4 规范主要技术内容和适用范围 10](#_Toc182763202)

[4.1 标准适用范围 10](#_Toc182763203)

[4.2 总体框架和主要内容 10](#_Toc182763204)

[4.3 规范性引用文件 11](#_Toc182763205)

[4.4 术语和定义 11](#_Toc182763206)

[4.5 试剂器材 13](#_Toc182763207)

[4.5.1 试剂和材料 13](#_Toc182763208)

[4.5.2 仪器与设备 14](#_Toc182763209)

[4.6 浮游植物监测 人工智能图象识别系统构建与使用流程 14](#_Toc182763210)

[4.6.1 样本采集 15](#_Toc182763211)

[4.6.2 图像获取与鉴定 16](#_Toc182763212)

[4.6.3 数据预处理 17](#_Toc182763213)

[4.6.4 技术原理与构建流程 18](#_Toc182763214)

[4.7 质量控制 18](#_Toc182763215)

[5 案例 19](#_Toc182763216)

[5.1 基于卷积神经网络的藻类识别系统 19](#_Toc182763217)

[6 标准实施建议 21](#_Toc182763218)

[附录 22](#_Toc182763219)

[附录A 浮游植物样品采集与检测记录表格 22](#_Toc182763220)

[附录B 浮游植物人工智能监测系统误差统计表 23](#_Toc182763221)

[附录C 浮游植物人工智能监测系统性能评估表 24](#_Toc182763222)

[附录D 人工智能监测系统浮游植物定量检测结果表（密度） 25](#_Toc182763223)

[附录E 人工智能监测系统浮游植物定量检测结果表（生物量） 26](#_Toc182763224)

[参考文献 27](#_Toc182763225)

# 1 项目背景

## 任务来源

本标准由中国环境科学研究院提出，并由中国环境科学学会归口，2023年申请立项，被列入中国环境科学学会 2023 年第二批标准编制计划正式批准立项，由中国环境科学研究院、南京大学、中国环境监测总站、中日友好环境保护中心、重庆市生态环境科学研究院、水利部交通运输部国家能源局南京水利科学研究院、广东工业大学等单位起草。

## 工作过程

按照标准编写要求，项目承担单位组织相关科研人员组成了标准编制组。编制组成员及时查阅国内外相关文献资料，按照 GB/T 1.1--2020 给出的最新规定起草和编制指南。在前期项目研究、文献资料分析以及实际应用的基础上，编制组讨论并确定了开展标准编制工作的原则、程序、步骤和方法，目前形成标准初稿及编制说明。主要工作如下：

1. 研究基础

2019-2023年，编制组通过阅读文献和收集国内外相关资料，确定了构建浮游植物监测 人工智能图像识别技术的工作，并针对相关内容和关键技术开展了大量的科学研究工作，积累了大量的研究基础和实践经验，初步确定了浮游植物 人工智能图像识别技术的基本框架和流程。

1. 编制启动

编制组接到标准制定任务后，立刻组织落实标准制定工作。确定由来自高校、科研机构、企业的相关专家组成起草组，形成标准初稿。

1. 理论研究

2023-2024年：为了按照文件要求，准确完成制定工作，标准起草组通过各种途径，收集并学习了《水质 湖泊和水库采样技术指导》（GB/T 14581），和《海洋微微型光合浮游生物的测定 流式细胞测定法》（GB/T 30737—2014）等生态环境部发表的本领域相关标准，同时也深入学习了《水生态监测技术指南河流水生生物监测与评价（试行）》（HJ1295-2023），《水生态监测技术指南湖泊和水库水生生物监测与评价（试行）》（HJ1296-2023），《地表水和污水监测技术规范》（HJ/T 91），《水质 采样技术指导》（HJ 494），《水质 浮游植物的测定 滤膜-显微镜计数法》（HJ 1215-2021），《水质 浮游植物的测定 0.1ml计数框-显微镜计数法》（HJ 1216-2021），《内陆水域浮游植物监测技术规范》（SL 733）等行业/地方标准，收集和研究了众多国内外浮游植物监测 人工智能图像识别的技术方法和实际案例。记过资料分析和共性总结，初步对浮游植物监测 人工智能图像识别技术指南的构建方案进行梳理和提炼。理顺了标准制定的方向和思路，形成标准编制大纲。

1. 调研/实验研究

2022年6月-2023年12月，编制组完成资料收集、文献调研，以及了解浮游植物智能监测在国内外的最新研究成果及发展趋势，整理过往在国内长江、松花江等流域开展的野外生态调研工作中收集到的浮游植物样本，进行鉴定、质控等工作，开展模型训练与系统构建等准备工作并在此基础上初步拟定标准方法制订的基本原则和技术路线。为了使标准具有科学性和可操作性，标准起草组在已有的实验研究和资料分析的基础上，进一步与相关技术和管理人员进行深入地探讨，调整已有方法。

1. 标准草稿

2023 年 6月-2023 年 12 月：标准起草组召开起草工作研讨会，就标准起草过程中存在的问题进行集中研讨。经过若干次课题组内部研讨会和专家咨询会，编制组根据各大项目的总体规划和已有研究成果，组织优势力量和团队学术骨干编写了《浮游植物监测 人工智能图像识别技术指南》大纲。

1. 标准立项

2023年11月：标准起草组向中国环境科学学会提交制修订立项申请书。

2023年12月：召开标准立项论证会，专家组一致同意标准通过立项论证。

2023年12月：经中国环境科学学会审议进行立项公示。

2024年 1月：经中国环境科学学会审议进行正式立项。

# 2 制定标准的必要性

## 2.1 标准制定的背景

我国水环境质量持续改善，但水生态环境保护不平衡、不协调问题较为突出，部分河湖水生态系统严重失衡，蓝藻水华频发，生物多样性显著下降，水生态保护修复任务艰巨。2023年，国家印发《长江流域水生态考核指标评分细则（试行）》，标志着我国走向统筹水资源、水环境、水生态治理，加强生物多样性保护和水生态修复，提升水生态系统健康水平的新阶段。

浮游植物作为水生生态系统中的重要成分，对于水环境的变化十分敏感，是许多水生生物赖以生存的能量与物质来源，其多样性与丰富度等指标对水质和生态平衡具有深远影响。然而传统的浮游植物形态学鉴定工作需要在显微镜下进行人工分类计数，不仅费时费力，且对鉴定人员具有极高的专业知识要求，上述两种因素限制了水生生态系统中的浮游植物监测的大规模长时间开展。利用机器学习算法为浮游植物识别和计数提供技术支撑能够帮助实现浮游植物监测的智能化和快速化，提高监测质量。因此本标准旨在规定浮游植物智能识别系统的设计、开发、实施和应用标准，以促进水体生态监测技术的创新与升级，提高水生态环境保护和水资源管理的效能。智能识别系统的引入将加强水体监测，实现对水生态灾害的及时预警和干预，从而保护生态环境，促进可持续发展。此标准的实施有望为水生生态监测技术产业带来新的创新和发展机遇，推动技术的进步和应用的扩展。这将有助于加强对水生生态系统的监测和保护，促进环境可持续发展，并为未来的研究和政策制定提供更多参考。智能识别系统的引入将催生新一代环境监测设备，促进相关产业的升级和转型。

## 2.2 标准制定的必要性

随着环境监测技术的发展，浮游植物智能监测已成为水环境管理的关键工具。国内外众多研究团队和企业已经开发出多种浮游植物智能监测技术产品，这些产品在水环境监测、生态评估和预警系统中发挥着重要作用。然而，尽管这些技术产品在实际应用中取得了一定的进展，目前浮游植物智能监测技术的评估体系尚未形成统一标准，这在一定程度上限制了技术的有效性和推广。

我国在浮游植物智能鉴定领域已取得显著研究成果，这些成果主要体现在产品开发、学术文献发表以及项目研究报告中。但现有技术产品在识别精度和分类一致性方面存在不统一的问题。大多数产品仅能达到属级别的识别精度，且缺乏统一的分类标准，这导致不同产品间的识别结果存在差异，难以实现有效比较和协同应用。此外，数据标定和处理的标准化问题同样亟待解决。不同数据集之间的差异，以及缺乏统一的数据标定和处理标准，严重影响了算法的准确性和可重复性。这不仅增加了数据预处理的复杂性，也对监测结果的可靠性提出了挑战。在质控方法方面，目前缺乏引入人工审核、置信度指标等质控措施，这限制了监测技术的准确性和可靠性评估。建立适当的质控指标，如误判率、漏判率等，并规定其合理范围，对于评估识别技术的性能至关重要。算法评估和比较指标的一致性同样面临挑战。目前缺乏统一的评估方法和比较指标，这使得不同算法的性能比较和选择变得复杂且主观。

鉴于上述挑战，建立浮游植物智能监测相关的指南和规范显得尤为迫切。通过制定统一的技术标准和评估方法，可以提高浮游植物智能监测技术的一致性和可比性，促进技术的标准化发展。本标准将引入浮游植物智能检测技术相关标准，有助于实现藻类检测的高效和智能化，从而为深入了解水生生态系统做出更大的贡献，进一步推动水生生物物种多样性的保护工作。此举将有助于建立水生生物指示种与智能检测系统的紧密联系，为更好地监测和评估水体健康状况提供支持。

## 2.3 标准制定的目的

1.规范化技术流程：通过制定明确的技术流程规范，确保浮游植物识别技术的标准化操作，从而提高研究的一致性和复现性。

2.促进数据共享与合作：建立统一的数据集格式和共享机制，鼓励数据的开放获取和跨团队、跨机构的合作，加速浮游植物研究的进展。

3.提升技术准确性与可靠性：通过优化算法和模型，提高浮游植物识别的精度，确保监测数据的科学性和实用性，为生态研究和环境管理提供坚实的数据支撑。

4.推动技术创新与发展：鼓励算法、模型和数据集的研究创新，提高识别技术的精度和效率，推广先进技术方法，促进技术在环境监测和生态研究中的实际应用。

5.建立评估与反馈机制：创建一套综合的评估体系，包括性能评估和用户反馈，以确保技术的持续改进和适应性。

6.培养专业人才：通过教育合作和专业培训，培养浮游植物识别和环境监测领域的专业人才，为技术发展和应用提供人力资源支持。

7.知识普及与公众参与：通过科普教育和宣传活动，提高公众对浮游植物生态价值的认识，促进社会对环境保护的参与和支持。

8.确保技术伦理与可持续性：在技术发展和应用过程中，确保遵循伦理原则，考虑环境和社会的可持续性，为实现绿色发展目标做出贡献。

## 2.4 国内外相关技术发展动态

我国在浮游植物识别上迄今为止多数仍采用传统的人工识别方法，没有充分利用计算机与图像处理技术相结合的方式。这导致了对浮游植物的分析与研究受到限制，使得我国在这一领域稍逊于国际先进水平。至今，国内的研究几乎都是基于专业人员在显微镜下镜检的结果开展的。2007年，汪振兴等人运用遗传算法对藻类的多种形态学特征进行提取，并对引发赤潮的藻类进行了识别[1]。2008年，苏荣国等研制了使用单色发光二极管为光源的浮游植物荧光自动监测仪，对我国近海常见的12种浮游植物进行了识别[2]。

随着深度学习技术的发展，市面上也出现了许多与深度学习技术相结合的浮游生物识别系统，目前大致可分为两种技术方向，分别是目标检测和图像检索，其主要区别在于图像检索需要在已构建的图像数据库内进行“以图搜图”以确定浮游植物的种类，而提升软件的检测能力，即增加可识别种类以及识别准确率是主要升级方向之一。在目标检测方面，胡圣等人研发的基于神经网络的藻类智能监测系统能在属水平上对绝大多数浮游藻类进行准确鉴定，平均准确率可达到87%[3]；郑州英视江河生态环境科技有限公司的Algae Blue智能藻类监测系统利用深度神经网络算法，能够对生物细胞进行3D建模，实现了对浮游藻类的细胞级监测。而杭州迅数科技有限公司的Algacount藻类智能鉴定计数仪和杭州万深检测科技有限公司的Algae-C浮游动物藻类智能鉴定计数仪都是利用图像检索在数据库中搜索相似度最高的结果输出。目前市面上的产品所给出的检测报表一般只能到分类单元的属层面。

国外对于浮游植物的智能识别发展较早，早在上个世纪90年代Pech-Pacheco等利用一种名为光学图像捕捉的技术识别并计量了墨西哥的Todos Santos湾中5种角藻（*Ceratium*），实现了90%准确率的效果[4]。2000年，Boddy等将人工神经网络与流式细胞仪结合，实现了72种浮游植物的识别，准确率达到70%[5]。2006年，美国Fluid Imaging Technology公司开发了一种被称为Flow CAM微藻识别系统的技术系统。该系统通过光学成像和图像处理技术，能够有效地识别微藻的种类和数量。相比于传统的方法，Flow CAM微藻识别系统具备了更高的识别准确率，达到了80%~90%的水平，并且具有低成本、低耗时的优点。2007年，Yao等人运用一种名为DGANN的分布式遗传算法与神经网络相结合的方法对水体中的蓝绿藻进行识别[6]，虽然准确率较高，但该方法受训练数据量和时间的影响较大。2012年，哥伦比亚大学研发出Phytoplankton Key系统，它可以运用二分类法根据浮游植物特征逐级确定到属。

## 2.5 与现行有关法律法规、强制性标准和其他有关标准的关系

在我国，淡水浮游植物监测标准主要集中于形态学鉴定领域。为了确保藻类智能监测技术的准确性和可比性，本标准在样本采集方法上，严格遵循了现行的相关规定。这些规定包括但不限于《水生态监测技术指南河流水生生物监测与评价（试行）》(HJ1295-2023)、《水生态监测技术指南湖泊和水库水生生物监测与评价（试行）》(HJ1296-2023)、《水质 湖泊和水库采样技术指导》(GB/T 14581)等，它们为样本的采集、保存、分析提供了详尽的技术指导，确保了监测工作的标准化和规范化。

在样品的鉴定和计数方法上，本标准借鉴了《水质 浮游植物的测定 0.1 ml 计数框-显微镜计数法》(HJ 1216-2021)和《水质 浮游植物的测定 滤膜-显微镜计数法》(HJ 1215-2021)等标准，这些方法涵盖了从样品采集到分析步骤、各指数计算方法以及质量控制标准的全面流程，为藻类的准确鉴定和计数提供了科学依据。

通过整合和执行这些规定和方法，本标准旨在构建一个完善的藻类智能监测流程。这不仅为水生生态系统的评价与管理提供了坚实的科学基础，而且通过智能监测技术的应用，能够更全面、更准确地评估水体健康状况，实现对水环境的及时管理和保护。此举将有助于提高我国水体监测工作的技术水平，并对保护并改善我国的水环境质量做出积极贡献。

本标准的制定，是在现有法规和标准的基础上，对藻类智能监测技术进行的一次系统性整合与创新。它不仅体现了技术进步对环境保护的推动作用，也展示了我国在水环境监测领域不断追求卓越的坚定决心。通过这些努力，我们期望为水生生态系统的可持续管理和水环境的长期健康提供坚实的支撑。

## 2.6 现有工作基础

目前，编制组开展了大量我国淡水水生生物的形态学鉴定以及浮游植物人工智能识别系统的开发相关工作，积累了大量野外调查和实验室分析的工作经验，对于规范的合理可操作性提供了很好的保障。同时，整理收集了浮游生物智能识别的经验案例，也为规范的编制提供了很好基础支撑。

# 3 标准编制依据和原则

## 3.1 编制依据

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 14581 水质 湖泊和水库采样技术指导

HJ1295-2023水生态监测技术指南河流水生生物监测与评价（试行）

HJ1296-2023水生态监测技术指南湖泊和水库水生生物监测与评价（试行）

HJ/T 91 地表水和污水监测技术规范

HJ 494 水质 采样技术指导

HJ 1215-2021 水质 浮游植物的测定 滤膜-显微镜计数法

HJ 1216-2021 水质 浮游植物的测定 0.1ml计数框-显微镜计数法

SL 733 内陆水域浮游植物监测技术规范

## 3.2 编制原则

本标准的编制主要遵循以下原则：

（1）严格遵守我国相关法律、法规和标准。以《中华人民共和国环境保护法》《中华人民共和国水污染防治法》《环境监测管理办法》以及我国现行的环境保护法律法规、政策、条例、标准的相关规定和要求为主要依据，体现标准的科学性、规范性和一致性，标准要和国家现有相关规范性文件相协调，避免冲突。

（2）技术规范应突出浮游植物人工智能监测技术的适用性和可操作性，能够提供清晰明确的操作指导，具备充分的理论科学基础和实际应用推广价值。技术规范需要在整个浮游植物智能监测系统领域内保持统一和一致。这包括统一的识别和分类标准、数据处理和分析方法、质控及质量评估等方面，确保不同监测系统之间的数据比对和共享的可行性。

（3）本标准基于可获得的最新科学证据进行编写，并承诺随着新的科学认识和科学证据的出现，对评估结果进行更新，以适应新的科学发现和技术进步。通过持续的评估和修订，确保标准的长期有效性和相关性，为环境保护和水环境监测提供持续的支持。

# 4 规范主要技术内容和适用范围

## 4.1 标准适用范围

根据建立技术规范的目的和专家讨论意见，本文件规定了浮游植物监测人工智能识别技术使用时的一般性原则、内容、程序、方法和技术要求，主要适用于我国淡水浮游植物的鉴定。

## 4.2 总体框架和主要内容

本标准规定了浮游植物监测 人工智能图像识别技术指南的相关术语、构建方法及质量控制等。

（1）范围

（2）规范性引用文件

（3）术语和定义

（4）仪器和设备

（5）技术使用原则

（6）系统构建与使用流程

（7）结果分析与计算

（8）质量控制和质量保证

## 4.3 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 14581 水质 湖泊和水库采样技术指导

HJ1295-2023水生态监测技术指南河流水生生物监测与评价（试行）

HJ1296-2023水生态监测技术指南湖泊和水库水生生物监测与评价（试行）

HJ/T 91 地表水和污水监测技术规范

HJ 494 水质 采样技术指导

HJ 1215-2021 水质 浮游植物的测定 滤膜-显微镜计数法

HJ 1216-2021 水质 浮游植物的测定 0.1ml计数框-显微镜计数法

SL 733 内陆水域浮游植物监测技术规范

## 4.4 术语和定义

（1）分类单元 taxon

分类单元是物种分类工作中的客观操作单位，有特定的名称和分类特征，主要包括门（Phylum）、纲（Class）、目（Order）、科（Family）、属（Genus）、种（Species）等。

（2）浮游植物 phytoplankton

浮游植物是指在水体中悬浮分布的微小植物生物，其主要由单细胞或多细胞的藻类组成。这些微生物具有光合作用能力，通过吸收阳光和水中的养分，进行光合作用从而为自身生长和繁殖提供能量和养分。通常浮游植物就是浮游藻类，包括原核的蓝藻和其它各类真核藻类。

[来源：SL733—2016，3.1，有修改]

（3）浮游植物密度 density of phytoplankton

单位体积的浮游植物的细胞数，cells/L。

[来源：SL733—2016，3.3，有修改]

（4）浮游植物物种组成 Phytoplankton species composition

浮游植物物种组成是指在特定水生生态系统中，浮游植物群体中不同物种的种类和比例。它反映了该生态系统中浮游植物多样性及其相对丰度，是评估水域健康状况和生态平衡的重要指标。

（5）人工智能 artificial intelligence; AI

人工智能是一项技术和科学，针对人类定义的给定目标，产生诸如内容、预测、推荐或决策等输出的一类工程系统的研究、开发和应用。人工智能旨在通过模拟和实现人类智能的某些任务，从而使计算机系统具备感知、学习、推理、决策和交互等智能能力。本文件中主要指通过机器学习或深度学习方法实现对藻类生物的自动识别分类统计。

（6）机器学习 machine learning

机器学习是一门研究如何让计算机系统基于过去的经验和数据，自动改善和完善执行任务能力的科学和技术。通过机器学习，计算机系统可以自动进行模式识别、决策制定和预测等任务，使模型的行为反映数据或经验。

（7）深度学习 deep learning

深度学习是一种机器学习方法，旨在通过构建和训练多层神经网络模型来实现对数据的高级抽象和表达。它能够模仿人脑神经网络的结构和功能，通过多层次的非线性变换对数据进行表征学习和特征提取，自动学习到任务相关的特征和模式，从而实现对复杂问题的有效解决。

（8）数据标注 data labeling

数据标注是指将数据集中的特定信息、特征或属性进行识别、注解、标记或描述的过程。在浮游植物形态学鉴定中，数据标注就是指对浮游植物图像或数据进行观察和描述，以识别出不同的植物种类或特征，然后记录在相应的数据库中。这一过程需要由经验丰富的专业人员进行，以确保数据标注的准确性和可靠性。

（9）卷积神经网络 convolutional neural network （CNN）

卷积神经网络是一种前馈神经网络，在其中至少一层网络中使用卷积。核心思想是利用卷积操作对输入数据进行特征提取，随后通过层级连接的方式进行深入学习，从而自动识别和提炼出高级特征与模式。在本文件中，卷积神经网络可用来提取图像特征并进行分类鉴定。

（10）神经网络 neural network; neural net

神经网络是一种模仿生物神经系统结构和功能的计算模型，它由多个节点（神经元）以及它们之间的连接组成。神经网络通过学习大量的数据来建立模型，通过输入数据的传递和处理来模拟复杂的非线性函数映射或解决各种任务。

（11）图像检测/识别 image detection / recognition

图像检测/识别是一种计算机视觉技术，旨在通过使用计算机算法和模型，自动分析图像内容并确定其中的对象、物体或特征。它通常涉及从输入图像中提取期望目标（浮游植物）的信息，并用合适的标签或类别进行分类或定位。

（12）训练数据 training data

用于训练机器学习模型的输入数据样本子集。

（13）验证数据 validation data

用于评估单个或多个候选机器学习模型性能的数据样本。

## 4.5 试剂器材

### 4.5.1 试剂和材料

（1）25 号浮游生物网：网孔直径为 0.064 mm。

（2）定性采样瓶：30 ml～100 ml 广口聚乙烯瓶。

（3）采水器：不锈钢或有机玻璃材质，圆柱形。容量和深度规格要满足采样要求。

（4）定量采样瓶：1 L～2 L 广口聚乙烯瓶。

（5）鲁哥氏液：称取 60 g 碘化钾溶解在 100 ml 蒸馏水中，加入 40 g 碘，充分搅拌使其完全溶解，加水定容至 1000 ml，转移至棕色磨口玻璃瓶。鲁哥氏液在室温避光条件下保存。

（6）浓缩装置：1 L～2 L 筒形分液漏斗或量筒。

（7）样品瓶：50 ml 具塞棕色玻璃广口瓶。

（8）0.1 ml 浮游植物计数框：面积 20 mm×20 mm，框内划分横竖各 10 行格，共 100 个小方格。

（9）目镜分划板：又称目镜测微尺，5 mm/50 DIV，分划值为 0.1 mm。

（10）计数器：用于浮游植物计数。

### 4.5.2 仪器与设备

（1）超声波发生装置：工作频率 40 kHz，水浴方式。

（2）微量移液器：100 μl。

（3）盖玻片：面积 22 mm×22 mm，厚度小于 0.2 mm。

（4）载物台测微计：又称镜台测微计，1 mm/100 DIV，分划值为 0.01 mm。

注：DIV 指等分格，1 mm/100 DIV 即 1 mm 等分成 100 格。

（5）集成自动对焦拍照等功能的数控一体化成像系统：终端放大倍数为100～400倍，单张像素不低于500万。

## 4.6 浮游植物监测 人工智能图象识别系统构建与使用流程

浮游植物监测 人工智能图象识别系统构建流程主要包括：样本采集、图像获取与鉴定、数据预处理、系统构建等过程（图1）。



图1 浮游植物监测 人工智能图象识别系统构建基本流程

### 4.6.1 样本采集

#### 4.6.1.1 点位布设

监测点根据GB/T 14581、HJ/T 91和HJ 494标准，结合水体特征和环境条件在河流和湖库的不同区域布设，数量依据监测任务确定。

#### 4.6.1.2 监测频率与采样时间

常规监测按季度或水文节律进行，水华期间增加频次。采样时间建议在8:00-17:00之间，特殊项目根据需求调整。

#### 4.6.1.3 定性样品采集

使用25号浮游生物网在水面下0～0.5m处采集浮游植物，采集分层样品时过滤特定水层。

#### 4.6.1.4 定量样品采集

浮游植物定量样品通过采水器采集不少于500ml，透明度高或浮游植物稀少时增加体积。

#### 4.6.1.5 样品固定

样品采集后立即加入鲁哥氏碘液固定，用量为水样体积的 1.0%～1.5%，水华时增加用量。样品可在室温下保存3周或冷藏保存12个月。

#### 4.6.1.6 沉淀浓缩

固定样品摇匀后静置24-48小时，用虹吸管吸取上清液至接近50ml标记线，保持吸液口与沉淀物间距离大于3cm。

#### 4.6.1.7 样品运输

确保样品数量准确，运输过程中避免污染和破损。

### 4.6.2 图像获取与鉴定

#### 4.6.2.1 载物台对焦调整

使用电脑或手柄控制载物台，将目标区域置于镜头中心，并调整焦距以清晰成像，确保图像稳定和对焦准确。

#### 4.6.2.2 载物台观察与焦距微调

移动载物台观察不同区域，微调焦距以获得清晰成像。对自动对焦系统，设置分辨率、对比度、锐度和对焦精度的量化指标。

#### 4.6.2.3 设置扫描参数

根据HJ 1216-2021标准，在200～400倍放大倍数下设置扫描起点和终点，确保扫描面积覆盖样品总面积的至少30%。

#### 4.6.2.4 扫描拍照

启动设备进行载物台移动，用显微镜顶端相机对每个视野单元拍照，程序存储图像用于后续分析。

#### 4.6.2.5 构建图像数据库

在浮游植物人工智能图像识别技术标准中，数据集构建应确保图像分辨率不低于300 dpi，格式推荐使用JPEG或PNG。图像应清晰无模糊或过度曝光，确保浮游植物特征明显便于识别。数据标注工作应由具有副高职称或从业两年以上鉴定工作的专业人员完成，以减少种类误差，并在有分歧时通过讨论或引入第三方专家达成共识。数据标注建议参考《中国淡水藻类:系统、生态及分类》[7]与《藻类名词及名称》[8]等专著。图像数据库应至少以属为单位构建，每属至少包含200个比对数据，同时确保数据集包含不同生长阶段和形态变异的浮游植物，以提高模型的泛化能力。

### 4.6.3 数据预处理

#### 4.6.3.1　数据清洗

识别并处理数据中的缺失值、异常值和重复值，以确保数据质量和一致性，减少对后续分析和模型建立的干扰。

#### 4.6.3.2　数据增强

利用常见的数据增强方法对处理后的数据进行旋转、剪切、缩放、添加噪点等操作改善训练样本过少、类不平衡等问题。数据的增强应根据实际的应用自行调整，建议增强后不同类别比应大致等于1：1。

#### 4.6.3.3　数据归一化

通过调整数据的尺度和范围进行数据归一化操作，以确保不同尺度的特征能够更好地适应模型的训练和优化过程。

#### 4.6.3.4　特征提取

利用计算机视觉技术提取浮游植物形状、颜色、纹理等特征，并将这些特征作为图像中浮游植物的关键视觉信息。

### 4.6.4 技术原理与构建流程

#### 4.6.4.1　选择深度学习算法

为了学习从特征到浮游植物类别的映射关系，浮游植物监测系统宜采用多种深度学习算法。这样做的目的是为了允许不同的系统根据它们特定的需求和监测的浮游植物种类，选取最适合的算法来优化识别效果，从而提高整体的监测准确性和可靠性。在训练过程中，模型通过调整参数来提升对浮游植物分类的准确性。随着对图像特征的学习和这些特征与浮游植物类别的关联，模型逐渐增强其自动识别和分类的能力。

#### 4.6.4.2　模型性能评估

模型训练完成后，需评估模型的关键性能参数，包括准确率、召回率和F1分数，以确保其综合性能处于预定的标准范围内。这些指标可以量化模型在识别和分类浮游植物时的效果。验证数据在这一过程中扮演关键角色。使用验证数据对模型进行测试，可以在实际投入使用前检测和调整模型的泛化能力，从而在实际应用中达到更高的准确性和可靠性。

#### 4.6.4.3　图像识别与生态指标分析

完成模型训练后，系统可用于实际水体图像识别应用。在获取新的水体图像时，系统会提取图像特征，并将其输入经过训练的模型，自动进行分类推理预测，从而确定图像中的浮游植物种类。在应用此系统并对样品检测完毕后，应产生相应检测结果及其数据报表，其中应包括不同藻种在样品种的分布情况和密度，以及香农-威纳指数、均匀度指数、丰富度指数等基本生态指标，以反映水体的生态健康程度。

## 4.7 质量控制

（1）严格按标准要求进行信息采集，填写各项观测数据。记录表格一般要编页装订成册，内容齐全，填写翔实，字迹工整、清晰。所有原始数据记录表、图片、样品和分类凭证标本应及时保存归档，并及时填写和归档电子数据（包括数据记录表和数码图片等）。

（2）采样完成后，将所有样品运回实验室，与实验室人员交接，填写实验室样品记录表。将样品瓶上的所有信息抄写在实验室样品登记表上，按照采样区域或样点对样品登记表进行统一编号。

（3）为防止外源污染，实验前应将实验用具进行高压灭菌，并用 75%乙醇擦洗药材表面。

（4）同样品可进行两次及以上的平行检测，对于浮游植物样品：控制PDE和PTD分别在10%和15%以内。对于平行样的检测，务必采用相同的拍摄参数。

# 5 案例

## 5.1 基于卷积神经网络的藻类识别系统

编制组结合近年来在中国内陆各流域的工作，构建了基于卷积神经网络的40余种（属）藻类识别系统。覆盖硅藻门、绿藻门、蓝藻门、隐藻门、裸藻门的常见种类，包括布纹藻、颤藻、集星藻、茧形藻、空星藻、菱形藻、卵囊藻、裸藻、盘星藻、平裂藻、十字藻、四角藻、四星藻、微囊藻、纤维藻、小环藻、小球藻、新月藻、衣藻、隐藻、栅藻、长孢藻、直链藻、舟形藻、针杆藻、脆杆藻、星杆藻、拟多甲藻、角甲藻等。

1. 图像数据库构建

编制组搜集中国南北各流域的浮游植物样品，交与具有副高职称或从业两年以上鉴定工作的专业人员完成。使用专业显微成像系统对样品进行拍照，根据鉴定结果进行标注，并对数据进行预处理，平衡类别间差异，使不同类别样本量约等于1：1，构建图像数据库。

1. 藻类识别系统

编制组建立了一个深度卷积神经网络模型，用于处理和分析显微成像系统捕获的浮游植物图像。模型采用先进的数据增强技术，包括旋转、缩放、剪切和颜色变换，以提高模型的泛化能力。通过大量图像数据的训练，网络学习到藻类的特征表示，并通过多轮迭代优化，目前已实现了对相关种（属）藻类的高精度识别。系统的性能通过一系列验证实验进行评估，包括准确率、召回率和混淆矩阵等指标（图2）。使用交叉验证方法来确保评估结果的可靠性。针对识别精度不高的藻类，编制组进一步分析了原因，可能是由于样本量不足或特征相似性造成的混淆。系统识别效果如图3所示。



图2 藻类识别系统性能评估

图3 系统识别效果图

# 6 标准实施建议

本标准为首次制订，在编制过程中，有关条款直接引用了现有国家标准或行业标准的内容，尽量避免重复，力求简化。内容上力求突出浮游植物监测人工智能图像识别的技术要求，层次上尽量体现与各标准之间的衔接。建议在本标准实施过程中，继续广泛听取和收集各方面的意见与建议，并根据实际应用情况，对本标准进行不断地修订与完善，使其实用性和可操作性与时俱进，为规范化开展浮游植物监测人工智能图像识别工作提供依据。

# 附录

1. 浮游植物样品采集与检测记录表格

|  |  |  |  |
| --- | --- | --- | --- |
| 样本编号 |  | 样本照片 |  |
| 凭证标本编号 |  | 凭证标本保存地点及保存方式 |  |
| 采样位点 |  | 采样人/单位 |  | 采样时间 |  |
| 经纬度 |  | 海拔 |  | 气温 |  | 水温 |  |
| 检测人/单位 |  | 检测时间 |  |
| 浓缩后体积（mL） |  | 稀释倍数 |  |

1. 浮游植物人工智能监测系统误差统计表

|  |  |  |  |
| --- | --- | --- | --- |
| 样本编号 |  | 样本照片 |  |
| 凭证样本编号 |  | 监测系统名称及编号 |  |
| 浓缩倍数 |  | 稀释倍数 |  |
| 取样量 |  | 计数方法 |  |
| 系统检测数量/（104cells/L） | 实际浮游植物数量/（104cells/L） |
| 蓝藻门 |  | 蓝藻门 |  |
| 绿藻门 |  | 绿藻门 |  |
| 硅藻门 |  | 硅藻门 |  |
| 裸藻门 |  | 裸藻门 |  |
| 甲藻门 |  | 甲藻门 |  |
| 金藻门 |  | 金藻门 |  |
| 黄藻门 |  | 黄藻门 |  |
| 隐藻门 |  | 隐藻门 |  |
| 计数差异百分比（PDE） |  |
| 系统检测种类数 |  | 实际浮游植物种类数 |  |
| 分类差异百分比（PTD） |  |

分析人： 分析日期：

1. 浮游植物人工智能监测系统性能评估表

|  |  |
| --- | --- |
| 监测系统名称及编号 |  |
| 成像系统像素 |  | 算法类型 |  |
| 监测内容 | 🞎浮游植物密度; 🞎物种组成; 🞎生物量 |
| 评价指数 | 🞎香农-维纳多样性指数; 🞎辛普森多样性指数; 🞎均匀度指数 |
| 系统性能指标 |
| 自动化程度 | 精确度 | 召回率 | F1分数 |
| 🞎自动数据采集🞎浮游植物检测🞎报表生成  |  |  |  |

1. 人工智能监测系统浮游植物定量检测结果表（密度）

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 样本编号 | 浓缩/稀释倍数 | 取样量 | 计数方法 | 计数面积 | 浮游植物总密度/（104cells/L） | 各门浮游植物密度/（104cells/L） | 优势种密度/（104cells/L） | 多样性指数 |
| 蓝藻门 | 绿藻门 | 硅藻门 | 甲藻门 | 裸藻门 | 金藻门 | 隐藻门 | 黄藻门 | 优势种1 | 香农-维纳多样性指数（H） | 辛普森多样性指数（D） | 均匀度指数（J） |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |   |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

1. 人工智能监测系统浮游植物定量检测结果表（生物量）

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 样本编号 | 浓缩/稀释倍数 | 取样量 | 计数方法 | 计数面积 | 浮游植物总生物量/（g/L） | 各门浮游植物生物量/（g/L） | 优势种生物量/（g/L） |
| 蓝藻门 | 绿藻门 | 硅藻门 | 甲藻门 | 裸藻门 | 金藻门 | 隐藻门 | 黄藻门 | 优势种1 |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

# 参考文献

[1] 汪振兴, 佘焱, 姜建国. 赤潮藻类图像自动识别的研究 [J]. 海洋环境科学, 2007, (01): 42-44.

[2] 苏荣国, 胡序朋, 张传松，等. 基于活体荧光的藻类识别测定技术 [J]. 热带海洋学报, 2008, 27(05): 24-29.

[3] 胡圣, 刘浩兵, 刘辉，等. 基于深度学习技术的藻类智能监测系统开发 [J]. 中国环境监测, 2022, 38(01): 200-210.

[4] Pech-Pacheco J, Alvarez-Borrego J. Optical–digital system applied to the identification of five phytoplankton species [J]. Marine Biology, 1998, 132: 357-365.

[5] Boddy L, Morris C, Wilkins M, et al. Identification of 72 phytoplankton species by radial basis function neural network analysis of flow cytometric data [J]. Marine Ecology Progress Series, 2000, 195: 47-59.

[6] Yao Z, Fei M, Li K, et al. Recognition of blue-green algae in lakes using distributive genetic algorithm-based neural networks [J]. Neurocomputing, 2007, 70(4-6): 641-647.

[7] 胡鸿钧, 魏印心. 中国淡水藻类--系统、分类及生态 [M]. 北京: 科学出版社, 2006.

[8] 曾呈奎, 毕列爵. 藻类名词及名称 [M]. 北京: 科学出版社, 1979.