|  |  |
| --- | --- |
| ICS  |   |
| CCS  |       |

|  |
| --- |
|  11 |

北京市地方标准

DB11/T XXXX—XXXX

浮游植物人工智能识别技术规范

多微流道成像法

Technical regualtions for phytoplankton identification using artificial intelligence multi-microchannel imaging

XXXX - XX - XX发布

XXXX - XX - XX实施

北京市市场监督管理局  发布

目次

[前言 II](#_Toc204957298)

[1 范围 1](#_Toc204957299)

[2 规范性引用文件 1](#_Toc204957300)

[3 术语和定义 1](#_Toc204957301)

[4 方法原理 2](#_Toc204957302)

[5 检测流程 2](#_Toc204957303)

[6 样品准备 3](#_Toc204957304)

[7 样品检测 3](#_Toc204957305)

[8 群落特征参数计算 5](#_Toc204957306)

[9 成果形式 6](#_Toc204957307)

[10 质量保证与质量控制 6](#_Toc204957308)

[附录A（资料性） 浮游植物优势属名录 8](#_Toc204957309)

[附录B（资料性） 浮游植物分类计数汇总表 9](#_Toc204957310)

[参考文献 10](#_Toc204957311)

1. 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。

本文件由北京市水务局、北京市生态环境局提出。

本文件由北京市水务局、北京市生态环境局归口。

本文件起草单位：北京市水文总站（北京市水务局水质水生态监测中心）、北京市生态环境监测中心、武汉大江锐视生态科技有限公司。

本文件主要起草人：

浮游植物人工智能识别技术规范

多微流道成像法

* 1. 范围

本文件规定了使用微流道法开展浮游植物人工智能识别的检测流程、数据分析、质量保证和质量控制等技术要求。

本文件适用于河流、湖泊、水库和渠道等水体的浮游植物检测。

* 1. 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

HJ 1216 水质 浮游植物的测定 0.1 mL计数框-显微镜计数法

HJ 1295 水生态监测技术指南 河流水生生物监测与评价（试行）

HJ 1296 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）

SC/T 9402 淡水浮游生物调查技术规范

SL 733 内陆水域浮游植物监测技术规程

DB11/T 1721 水生生物调查技术规范

* 1. 术语和定义

下列术语和定义适用于本文件。

微流道 microchannel

用于处理或操纵微小流体、尺寸从几微米到几毫米的微管道。

浮游植物 phytoplankton

悬浮于水中生活的微小藻类，亦称浮游藻类。通常包括蓝藻门、绿藻门、硅藻门、隐藻门、甲藻门、金藻门、黄藻门、裸藻门等，不包括细菌和其他植物。

[来源：DB11/T 1721，3.1]

人工智能检测 artificial intelligence detection

使用人工智能算法对采集到的图像或视频数据进行自动化处理、识别分类、特征提取及量化分析的综合技术。

人工智能检测系统 artificial intelligent detection system

基于人工智能检测技术实现自动化检测的系统，包含硬件设备、算法模型与软件平台等模块。

浮游植物细胞密度 phytoplankton density

单位体积水体中全部或某种（类）浮游植物的细胞数，单位为cells/L。

浮游植物生物量 phytoplankton biomass

单位体积水体中全部或某种（类）浮游植物的鲜重，单位为mg/L。

浮游植物优势种 dominant species of phytoplankton

浮游植物样品中占比超过总细胞密度或总生物量15%的种类。

召回率 recall

被正确预测的正样本占全部正样本的比率。

[来源：GB/T 41864，3.9.20]

精度 precision

预测类别为正样本的集合中真实类别为正样本的比率。

[来源：GB/T 41864，3.9.8]

重叠度 intersection over union

预测范围和标记范围的交叠率，即两个区域重叠的部分除以两个区域的集合部分得出的结果。

[来源：GB/T 41864，3.9.4，有修改]

群落特征参数 community characteristic parameters

用于描述生物群落组成、结构及功能的特征参数，常包含物种密度、生物量、丰富度、多样性、均匀性等指标。

* 1. 方法原理

将前处理后的浮游植物样品导入微流道中，由显微镜多视野多液层采集图像，利用人工智能检测系统，经过图像预处理、目标识别、特征提取等步骤，智能识别浮游植物种类并计数，计算浮游植物种类数、细胞密度、生物量、优势种等信息。

* 1. 检测流程

浮游植物人工智能检测流程见图1。



1. 浮游植物人工智能检测流程
	1. 样品准备
		1. 样品采集

点位布设及采样频次参考GB/T 14581、HJ 1295、HJ 1296和SL 733的相关规定执行。也可根据调查研究目的确定。

定性和定量样品的采集和保存按照HJ 1216相关要求执行。

* + 1. 前处理

参考HJ 1216、SC/T 9402和DB11/T 1721进行浮游植物样品固定、沉淀和细胞密度调整，前处理后加入微流道中的样品浮游植物细胞密度宜为1 ×107 ~ 3×109 cells/L。

若样品泥沙等杂质含量高，应稀释水样使杂质图像与浮游植物图像之间空隙距离多数大于50微米。

* 1. 样品检测
		1. 进样

将样品放至室温，取适量混匀后的样品注入浮游植物微流道板中，使微流道被完全充满，不留气泡。宜选用多通路浮游植物微流道板。

* + 1. 显微成像
			1. 硬件要求

系统应满足以下要求：

1. 应配备具有平场消色差功能且物镜数值孔径不低于0.6的生物级显微镜，满足系统光学分辨率达到0.5微米以上；
2. 应配备图像帧率至少为25帧/秒的工业相机，相机感光单元不低于600万像素，宜2000万像素及以上；
3. 应配备相对于显微物镜至少三轴方向微米级运动自动控制功能的载物台，其中沿物镜轴向（Z轴方向）的运动控制精度应优于1微米；
4. 应采用多液层扫描拍摄方式，对待测样品每个视野的不同液层自动对焦拍照。
	* + 1. 显微扫描拍摄
				1. X、Y轴方向计数视野分布

进样完成后，样品在显微镜载物台的待测位置进行自动显微扫描拍摄，计数视野宜在整个样品观测范围内均匀分布。

参考SL 733，可根据视野平均藻细胞数确定合适的视野数。

* + - * 1. Z轴方向多液层扫描拍摄

在Z轴方向应每隔不高于5微米拍摄一幅图像，所拍摄图像的焦平面范围宜覆盖微流道内底部到顶部所有液层。

* + - 1. 图像质量

显微扫描拍摄图像应满足每微米内至少包含8个像素、平均色差应小于2。为保证拍摄图像质量、相机的曝光时间应小于10ms、感光度ISO数值应小于400、增益应小于200%。

* + 1. 智能识别

将各视野不同液层的显微拍摄图像融合，利用人工智能检测技术，根据合成的高清样品图像自动进行其中浮游植物的种属识别。

* + - 1. 训练数据库要求

标注图像应覆盖常见浮游植物90%以上属，其中包括附录A中所有优势属。

浮游植物图像特征清晰且在角度、形态、大小、颜色、背景等方面具有一定差异性，各属图片数量应能满足7.3.2要求。

训练数据库应定期扩建，增加新发现或需优化的物种图像。

* + - 1. 人工智能检测算法要求

算法对于本地数据库中所有种类的检测召回率不低于95%。

检测平均精度不低于95%，当重叠度阈值在0.5～0.95之间时，所有类别平均精度的均值不低于90%。

算法能够在普通工作站上（中低端中央处理器、主流图形处理器及32GB及以上内存）顺畅运行，对于单张图像检测时间不高于500ms。

算法支持通过不断收集反馈数据优化检测模型权重参数等方式提升性能。

检测模型具备数据适应性，能够通过数据增强、迁移学习等技术，保持在不同数据分布下的准确性。

* + 1. 智能计数

在智能识别的同时进行细胞计数。若多细胞浮游植物个体细胞间隙不明显，宜根据群体的排列特征，测量面积或长度估算细胞数。计数完成后自动计算样品的浮游植物细胞密度、相对丰度、生物量和优势种属。

* 1. 群落特征参数计算
		1. 细胞密度

按照公式（1）计算样品中浮游植物细胞密度：

$$N=\frac{n}{A}×\frac{1}{h}×\frac{V\_{1}}{V\_{0}}×1000000…………………………………………\left(1\right)$$

式中：

*N*——浮游植物细胞密度，单位为 （cells/L）；

*n*——检测视野浮游植物累计细胞数，单位为 （cells）；

*A*——检测视野累计计数面积，单位为 （mm2）；

*h*——微流道板内径高度，单位为毫米（mm）；

*V1*——稀释或浓缩后的试样体积，单位为毫升（mL）；

*V0*——稀释或浓缩前的水样体积，单位为毫升（mL）；

1000000——体积换算系数，单位为 （μL/L）。

* + 1. 相对丰度

按照公式（2）计算样品中第i属（种）浮游植物的相对丰度：

$$P\_{i}=\frac{n\_{i}}{n}×100\%…………………………………………………\left(2\right)$$

式中：

*Pi*——第i属（种）浮游植物的相对丰度；

$n\_{i}$——检测视野第i属（种）浮游植物的累计细胞数，单位为 （cells）；

n——检测视野浮游植物累计细胞数，单位为 （cells）。

* + 1. 生物量

浮游植物生物量计算优先采用体积测量法、然后换算成生物量。生物量具体检测方法按照SL 733 6.5执行。也可根据已有资料查得浮游植物细胞平均湿重值计算生物量。

* + 1. Margalef丰富度指数

按照公式（3）计算样品的Margalef丰富度指数：

$$d=\frac{S-1}{lnM}…………………………………………………\left(3\right)$$

式中：

$\overbar{d}$——Margalef丰富度指数；

*S*——浮游植物种类总数；

*M*——浮游植物个体总数。

* + 1. Shannon-wiener多样性指数

按照公式（4）计算样品的Shannon-wiener多样性指数：

$$H^{'}=-\sum\_{i=1}^{S}P\_{i}×lnP\_{i}……………………………………………(4)$$

式中：

*H´*——Shannon-wiener多样性指数；

*S* ——浮游植物种类总数；

*Pi*——第i属（种）浮游植物的相对丰度。

* + 1. Pielou均匀度指数

按照公式（5）计算样品的Pielou均匀度指数：

$$ J=\frac{H'}{lnS}……………………………………………………\left(5\right)$$

式中：

*J*——Pielou均匀度指数；

*H’*—— Shannon-wiener多样性指数；

*S*——浮游植物种类总数。

* + 1. Simpson优势度指数

按照公式（6）计算样品的Simpson优势度指数：

$$D=1-\sum\_{i=1}^{S}P\_{i}^{2}……………………………………………\left(6\right)$$

式中：

*D*——Simpson优势度指数；

*S* ——浮游植物种类总数；

*Pi*——第i属（种）浮游植物的相对丰度。

* 1. 成果形式
		1. 分类计数汇总表

分类计数汇总表由浮游植物的分类单元、细胞密度、种类相对丰度、生物量、优势种、门类细胞占比等组成，浮游植物分类计数汇总表参见附录B。

* + 1. 浮游植物图像识别结果

合成图上应以种属的中文名称或拉丁名标注识别结果。

* 1. 质量保证与质量控制
		1. 人工比对

应在人工智能检测系统投入使用时和算法更新后进行人工比对测试。选择前处理后细胞密度在1 ×107 cells/L ~ 3×109 cells/L之间、分属不同数量级的3个样品，同时使用本方法和HJ 1216、SL 733等规定的显微镜计数法开展检测。检测结果浮游植物细胞密度差异百分比应低于30%，优势种属保持一致。否则应重新调试算法至满足要求为止。

* + 1. 重现性

应在人工智能检测系统使用过程中按以下要求进行重现性测试：

1. 随机选择10%的样品，对加样完成的同一微流道板重复检测2次，2次细胞密度检测结果的相对误差应小于15%，优势种属一致；
2. 样品加样计数2次，2次浮游植物细胞总计数量结果相对偏差应在±15%以内，否则应增加计数一次，直至某两次计数结果符合这一要求为止。测定结果为相对偏差在±15%以内的两次计数结果的平均值。
3.
4. （资料性）
浮游植物优势属名录

浮游植物优势属名录参见表A.1。

* 1. 浮游植物优势属名录

| 舟形藻属 | 桥弯藻属 | 曲壳藻属 | 蓝隐藻属 | 栅藻属 |
| --- | --- | --- | --- | --- |
| 沟链藻属 | 针杆藻属 | 小环藻属 | 脆杆藻属 | 锥囊藻属 |
| 平裂藻属 | 微囊藻属 | 假鱼腥藻属 | 隐球藻属 | 盘星藻属 |
| 泽丝藻属 | 空星藻属 | 颤藻属 | 纤维藻属 | 隐藻属 |
| 衣藻属 | 金杯藻属 | 细鞘丝藻属 | 四角藻属 | 裸甲藻属 |
| 直链藻属 | 星杆藻属 | 蹄形藻属 | 卵囊藻属 | 束丝藻属 |
| 鱼鳞藻属 | 鱼腥藻属 | 游丝藻属 | 尖头藻属 | 浮鞘丝藻属 |
| 多甲藻属 | 羽纹藻属 | 转板藻属 | 鼓藻属 | 丝藻属 |
| 菱形藻属 | 拟柱孢藻属 | 纺锤藻属 | 浮丝藻属 | 小球藻属 |
| 拟多甲藻属 | 空球藻属 |  |  |  |

1. （资料性）
浮游植物分类计数汇总表

浮游植物分类计数汇总表参见表B.1。

* 1. 浮游植物分类计数汇总表

|  |  |  |  |
| --- | --- | --- | --- |
| 分析日期 |  | 样品编号 |  |
| 物镜倍数 |  | 设备名称及编号 |  |
| 浓缩倍数 |  | 稀释倍数 |  |
| 取样量(mL) |  | 计数方式 |  |
| 方法依据 |  |
| 门 | 纲 | 目 | 科 | 属 | 种 | 拉丁名 | 细胞数(cells) | 细胞密度(cells/L) | 相对丰度(%) | 生物量(mg/L) |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |
| 总计 |  |  |  |  |  |  |  |  |  |  |
|  | 蓝藻门 | 绿藻门 | 硅藻门 | 隐藻门 | 裸藻门 | 黄藻门 | 甲藻门 | 金藻门 |
| 细胞密度(cells/L) |  |  |  |  |  |  |  |  |
| 细胞密度占比(%) |  |  |  |  |  |  |  |  |
| 总细胞密度(cells/L) |  |
| 优势种（属）1 |  | 密度(cells/L) |  |
| 优势种（属）2 |  | 密度(cells/L) |  |
| 优势种（属）3 |  | 密度(cells/L) |  |
| 备注： |

参考文献

[1] GB/T 41867 信息技术 人工智能 术语

[2] GB/T 42886 数字成像显示显微镜 提供给用户的成像性能信息

[3] GB/T 45225 人工智能 深度学习算法评估

[4] HJ/T 91.2 地表水环境质量监测技术规范

[5] HJ 494 水质 采样技术指导

[6] DB14/T 2527 云平台 人工智能建模系统框架及功能要求

[7] 胡鸿均，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] Greenberg J H. The measurement of linguistic diversity[J]. Language, 1956, 32: 105-119.

