ICS ××.×××.××

× ×××

|  |
| --- |
| 备案号： |

团体标准

特种设备无损检测缺陷自动识别系统

202X - XX - XX发布

核电厂无损检测缺陷自动识别系统评价方法

第1部分：射线

Evaluation Method for Automatic Defect Recognition System in Nondestructive Testing of Nuclear Power Plant

Part 1: X-ray

（征求意见稿）

中国核能行业协会

202X - XX - XX实施

发 布

目次

1 范围 2

2 规范性引用文件 2

3 术语和定义 2

4 人员要求 5

5 基本要求 5

6 评价流程 6

7 标准测试集 6

8 性能测试 9

9 评价指标 11

10 评价结果 13

附录A 16

A.1 评价报告模板 16

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。

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。

本文件由中国核能行业协会提出并归口管理，技术支持单位为上海核工程研究设计院股份有限公司、核工业标准化研究所、苏州热工研究院有限公司、华能核能技术研究院有限公司。

本文件起草单位：XXX

本文件主要起草人：XXX

核电厂无损检测缺陷自动识别系统评价方法

第1部分：射线

* 1. 范围

本文件规定了适用于核电厂无损检测（射线）缺陷自动识别系统的评价方法。

本文件规定了核电厂射线检测缺陷自动识别系统评价方法的人员要求、系统基本要求、评价流程和评价方法。

本文件适用于核电厂金属熔化焊焊接接头射线底片缺陷自动识别系统的应用评价，适用的射线检测图像来源为通过X或伽马射线检测获得的物理底片经底片数字化设备扫描后的图像。通过X射线数字成像（DR）检测直接获取的图像和通过X射线计算机辅助成像（CR）检测拍摄并扫描的图像可参照使用。

* 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 25000.51 系统与软件工程 系统与软件质量要求和评价（SQuaRE） 第51部分：就绪可用软件产品（RUSP）的质量要求和测试细则

GB/T 28452 信息安全技术 应用软件系统通用安全技术要求

NB/T 2709 核电厂射线照相底片数字化技术规范

NB/T 47013 承压设备无损检测 第2部分：射线检测

* 1. 术语和定义

下列术语和定义适用于本文件。

3.1

缺陷自动识别系统（ADR） automatic defect recognition

基于计算机视觉算法，对无损检测数字图像缺陷自动识别的计算机应用系统，以下简称系统。

3.2

辅助评片（AssistDR）assisted defect recognition

系统辅助评片人员识别缺陷及评估焊接接头是否合格的过程，包括缺陷定性、定量、定位及评级。

3.3

自动评片（AutoDR）automated defect recognition

系统自主进行缺陷识别及评估焊接接头是否合格的过程，包括缺陷定性、定量、定位及评级。

3.4

标准底片 standard film

用于评价及测试缺陷自动识别系统性能的数字底片。

3.5

重点关注缺陷 key defects

评价缺陷自动识别系统识别性能的危害性缺陷，指裂纹、未熔合及未焊透。

3.6

一般关注缺陷 general defects

除重点关注缺陷外，其余的缺陷类型。

3.7

缺陷标注 defect label

对数字底片上的缺陷进行缺陷性质（cls）和缺陷位置（bbox）的标记，生成标注文件。

3.8

标注样本（Lable）label samples

对标准底片进行缺陷标注生成的标注文件，包含缺陷性质（cls\_L）及缺陷位置（bbox\_L）。

3.9

预测样本（Prediction）prediction samples

由缺陷自动识别系统预测的结果文件，包含缺陷性质（cls\_P）及缺陷位置（bbox\_P）。

3.10

训练集 train set

缺陷自动识别系统中使用人工智能算法训练模型的数字底片集合及对应的标注文件。

3.11

测试集 test set

用于测试缺陷自动识别系统性能的数字底片集合。

3.12

缺陷位置交集 defect location intersection

对于同一张底片上，预测样本缺陷位置（bbox\_P）与标注样本缺陷位置（bbox\_L）相交。



3.13

缺陷定位精度（La）location accuracy

预测样本与标注样本缺陷位置中心距离的偏差。

$$La=\sqrt{(x\_{P}−x\_{L})^{2}+(y\_{P}−y\_{L})^{2}}$$

式中:

$x\_{P}$、$y\_{P}$——预测样本的缺陷位置中心坐标；

$x\_{L}$、$y\_{L}$——标注样本的缺陷位置中心坐标。

3.14

正检（TD） true detection

定性并定位正确的缺陷。

3.15

误检(FD) false detection

定性或定位错误的缺陷。

3.16

漏检(MD) miss detection

未识别到的缺陷。

3.17

 缺陷定量精度（Da）dimension accuracy

 对于正检缺陷，预测样本的缺陷长径占标注样本的缺陷长径的比例。

$$Da=\left(1−\frac{\left|L\_{P}−L\_{L}\right|}{L\_{L}}\right)×100\%$$

式中:

$L\_{P}$——预测样本的缺陷长径；

$L\_{L}$——标注样本的缺陷长径。

3.18

不合格底片检出（ND）no good detection

被正确检出的超过相关标准验收等级的底片。

3.19

合格底片检出（GD）good detection

被正确检出的未超过相关标准验收等级的底片。

3.20

误判（FJ）false judgment

将未超标底片判断为超标底片。

3.21

漏判（MJ）miss judgment

将超标底片判断为未超标底片。

* 1. 人员要求

4.1 测试集标注人员应按《民用核安全设备无损检验人员资格管理规定》规定取得RT Ⅱ级及以上证书或TSG Z8001规定取得RT Ⅱ级及以上证书。

4.2 系统评价人员应按《民用核安全设备无损检验人员资格管理规定》规定取得RT Ⅲ级证书TSG Z8001规定取得RT Ⅲ级证书。

* 1. 系统要求

5.1 识别效率

以系统测试500张数字底片进行缺陷自动识别所需要的平均时间作为系统识别效率的计算指标。

5.2 功能要求

5.2.1 系统应具备打开DICONDE格式数字底片及底片格式转换的功能。

5.2.2 系统应具备焊缝区自动识别功能，可以自动识别核电厂射线检测常见接头类型，包含对接焊缝、小径管焊缝及插套焊缝。

5.2.3 系统应具备缺陷自动识别功能，包括自动识别有无缺陷、缺陷类型、缺陷位置、缺陷尺寸等功能。

5.2.4 系统应具备按照核电厂相关验收标准对缺陷自动识别结果进行质量评级的功能，并自动判断底片是否合格。

5.2.5 系统应具备自动识别结果人工复核功能，包括人工复核评级结果、缺陷增加、删除、修改等功能，并可将人工复核后的结果保存。

5.2.6 系统应具备图像降噪、平移、放大、缩小、旋转、镜像、适应屏幕尺寸、1:1 尺寸、还原、对比度调整、亮度调整、窗位窗宽调整、浮雕、锐化、正反片转换、双片对比等图像操作功能。

5.2.7 系统应具备双片叠加功能。

5.2.8 系统应具备底片黑度、灵敏度、缺陷尺寸等测量功能。

5.3 运行环境

5.3.1 系统部署方式应支持云端部署或本地部署。

5.3.2 系统硬件是自动识别系统软件安装、运行的计算硬件设备,包括CPU模块、内存模块、GPU推理模块、存储模块及显示模块。

5.3.3 CPU和内存模块应根据系统处理的任务规模与复杂程度、系统单位时间访问量以及系统对响应时间等要求进行配置。

5.3.4 GPU推理模块应满足5.1.2识别效率要求。

5.3.5 存储模块的选择应根据系统内文件和数据量、对系统扩展性的要求、对数据读写速度的要求等因素进行配置。

5.3.6 显示器模块应结合系统对图像显示及处理的精度要求进行配置、并符合NB/T 2709《核电厂射线照相底片数字化技术规范》的要求。

5.4 系统可靠性

系统软件质量应符合标准GB/T 25000.51-2016《系统与软件工程 系统与软件质量要求和评价（SQuaRE） 第51部分：就绪可用软件产品（RUSP）的质量要求和测试细则》中的相关要求。

5.5 系统安全性

系统安全应符合标准GB/T 28452-2012《信息安全技术 应用软件系统通用安全技术要求》中的相关要求。

5.6 系统升级

缺陷自动识别系统升级后按照本文件重新评价，评价结果自动更新。

* 1. 评价流程

本章规定了核电厂射线检测缺陷自动识别系统性能评价的流程步骤，包括标准测试集、性能测试、评价指标及评价结果四个步骤，如图1所示。



1. 评价流程
	1. 标准测试集

标准测试集用于测试缺陷自动识别系统性能,本章规定了标准底片的质量要求、测试集类型、测试集构建及测试集标注要求。

7.1 标准底片质量

7.1.1 底片扫描质量按照NB/T 2709《核电厂射线照相底片数字化技术规范》要求执行。

7.2 测试集分类

7.2.1 缺陷测试集

7.2.1.1 缺陷测试集用于测试系统的缺陷识别能力，包括缺陷的定性、定位及定量。

7.2.1.2 缺陷测试集中的每张底片都应存在缺陷，缺陷底片总数不少于500张，其中小径管焊接接头底片占比约20%，插套焊缝底片占比约10%。

7.2.1.3 缺陷测试集中应包含裂纹、未熔合、未焊透、圆形缺陷、条形缺陷、内凹、咬边等类型缺陷。

7.2.1.4 缺陷测试集中裂纹、未熔合、未焊透、内凹、咬边每类缺陷的数量基本相当且每类缺陷数量不少于50个，圆形缺陷和条形缺陷的缺陷数量不受限制。

7.2.2 评级测试集

7.2.2.1 评级测试集用于测试系统的缺陷定级能力，判断缺陷是否合格。

7.2.2.2 评级测试集由缺陷测试集和合格底片测试集组成，缺陷测试集使用7.2.1中规定的测试集，合格底片测试集由按照核电厂相关验收标准评级后为合格的底片组成。

7.2.2.3 合格底片测试集中应包含圆形缺陷和无任何缺陷底片。

7.2.2.4 合格底片测试集的数量与缺陷测试集的数量一致，其中小径管焊接接头底片占比约20%。

7.3 测试集构建

7.3.1 测试集图像与训练集图像在焊缝形式、焊接方法、射线源、射线检测透照布置、图像扫描参数保持一致，确保测试集数据与训练集数据一致性。

7.3.2 测试集与训练集的互斥性：测试数据集中不包含训练数据集中的数据。

7.3.3 测试集的隐私性，测试集需考虑企业隐私泄露，采取数据脱敏处理。

7.4 测试集标注

7.4.1 缺陷标注样本

7.4.1.1 标注人员对缺陷测试集中的标准底片依次进行缺陷标注，生成缺陷标注样本(Label)。缺陷标注样本包含缺陷性质和缺陷位置信息，缺陷标注样本为用于测试系统缺陷检出精度的对比样本。

7.4.1.2 缺陷标注样本加工流程如图2所示，由A、B、C三名标注人员分别独立标注，对比三人独立标注的结果，当三人标注缺陷性质一致时，标注样本的缺陷位置为bbox\_A、bbox\_B、bbox\_C的交集（bbox\_A∩bbox\_B∩bbox\_C）。当三人标注缺陷性质不一致时，该缺陷不计入标注样本。



1. 缺陷标注样本加工流程

7.4.2 评级标注样本

7.4.2.1 标注人员对评级测试集中的缺陷测试集和合格底片测试集依次进行定级（底片评级按照核电厂相关验收标准执行），生成评级标注样本。评级标注样本包含底片级别与是否合格信息，评级标注样本为用于测试系统评级精度的对比样本。

7.4.2.2 评级标注样本加工流程如图3所示，评级测试集由A、B、C三名标注人员分别独立评级，对比三人独立评级的结果，仅当三人评级结果一致时，纳入评级标注样本，否则舍弃该底片。



1. 评级标注样本加工流程
	1. 性能测试

自动识别系统的性能测试过程包括生成预测样本、样本偏离程度及生成混淆矩阵三部分。

8.1 生成预测样本

8.1.1 将缺陷测试集中的数字底片输入待测系统进行缺陷检测，自动生成缺陷预测样本（Prediction），缺陷预测样本中必须包含缺陷性质（cls）和缺陷位置(bbox)信息。

8.1.2 将评级测试集中的数字底片输入待测系统进行评级检测（底片评级按照核电厂相关验收标准执行），自动生成评级预测样本，评级预测样本中必须包含底片评定级别与是否合格信息。

8.2 样本偏离程度

将标注样本（Label）与预测样本(Prediction）进行缺陷定性定位对比，表征两个样本的缺陷偏离程度有以下三种情况：缺陷正检（TD）、缺陷误检(FD)和缺陷漏检(MD）。

（a）缺陷位置相交：bbox\_P∩bbox\_L＞0的前提下，当同时满足以下条件时，判断缺陷正检。

缺陷性质一致：cls\_P=cls\_L;

缺陷定位正确：圆形缺陷La≤1mm，其余缺陷La≤5mm，其余缺陷包括裂纹、未熔合、未焊透、咬边、内凹及条形缺陷。

(b)缺陷位置相交：bbox\_P∩bbox\_L＞0的前提下，满足以下条件之一时，判断缺陷误检。

缺陷性质不一致：cls\_P≠cls\_L;

缺陷性质一致：cls\_P＝cls\_L且缺陷定位错误：圆形缺陷La＞1mm，其余缺陷La＞5mm，其余缺陷包括裂纹、未熔合、未焊透、咬边、内凹及条形缺陷。

(c）缺陷位置不相交时：bbox\_P∩bbox\_L=0，判断缺陷漏检。

自动计算标注样本与预测样本缺陷偏离程度的流程如图4所示。



1. 样本偏离程度计算流程

8.3 生成混淆矩阵

8.3.1 统计缺陷测试集中每类缺陷的正检（TD）、误检（FD）和漏检（MD）数量并生成缺陷检测混淆矩阵，如图5所示。缺陷检测混淆矩阵中，使用TDn、FDn和MDn分别表示每一类缺陷的正检数量、误检数量和漏检数量，n为缺陷序号。



1. 缺陷检测混淆矩阵

8.3.2 统计评级测试集中合格底片检出（GD）、不合格底片检出（ND）、漏判（MJ）和误判（FJ）数量并生成评级检测混淆矩阵，如图6所示。



图7评级检测混淆矩阵

* 1. 评价指标

性能测试的评价指标包括缺陷检出精度和质量评级精度。

9.1 缺陷检出精度

缺陷检出精度表征系统对单类别缺陷的检出能力和所有缺陷的综合检出能力。

9.1.1 缺陷正检率按公式（1）计算。

$$TDRn= \frac{TDn}{TDn+FDn+MDn} \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \left(1\right)$$

式中:

$TDRn$——序号为n的缺陷正检率；

*n*——缺陷序号（圆形缺陷1、条形缺陷2、裂纹3、未熔合4、未焊透5、内凹6、咬边7）。

9.1.2 缺陷误检率按公式（2）计算。

$$FDRn= \frac{FDn}{TDn+FDn+MDn}\cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \left(2\right)$$

式中:

$FDRn$——序号为n的缺陷误检率；

*n*——缺陷序号（圆形缺陷1、条形缺陷2、裂纹3、未熔合4、未焊透5、内凹6、咬边7）。

9.1.3 缺陷漏检率按公式（3）计算。

$$MDRn= \frac{MDn}{TDn+FDn+MDn}\cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \left(3\right)$$

式中:

$MDRn$——序号为n的缺陷漏检率；

*n*——缺陷序号（圆形缺陷1、条形缺陷2、裂纹3、未熔合4、未焊透5、内凹6、咬边7）。

9.1.4 重点关注缺陷检出率按公式（4）计算。

$$KDR= \frac{\sum\_{n=3}^{5}(TDn+FDn)}{\sum\_{n=3}^{5}(TDn+FDn+MDn)}\cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \left(4\right)$$

式中:

$KDR$——重点关注缺陷检出率；

*n*——缺陷序号（圆形缺陷1、条形缺陷2、裂纹3、未熔合4、未焊透5、内凹6、咬边7）。

9.1.5 综合检出率按公式（5）计算。

$$WDR= \frac{\sum\_{n=1}^{7}(TDn+FDn)}{\sum\_{n=1}^{7}(TDn+FDn+MDn)}\cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \left(5\right)$$

式中:

$WDR$——综合检出率；

*n*——缺陷序号（圆形缺陷1、条形缺陷2、裂纹3、未熔合4、未焊透5、内凹6、咬边7）。

9.1.6 综合正检率按公式（6）计算。

$$TDR= \frac{\sum\_{n=1}^{7}TDn}{\sum\_{n=1}^{7}(TDn+FDn+MDn)}\cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \left(6\right)$$

式中:

$TDR$——综合正检率；

*n*——缺陷序号（圆形缺陷1、条形缺陷2、裂纹3、未熔合4、未焊透5、内凹6、咬边7）。

9.1.7 对于所有的正检缺陷，平均定量精度按公式（7）计算。

$$\overbar{Da}= \frac{1}{\sum\_{n=1}^{7}TDn}\sum\_{\sum\_{n=1}^{7}TDn}^{}Da\cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \left(7\right)$$

式中:

$\overbar{Da}$——平均定量精度；

*TD*——缺陷正检数；

*n*——缺陷序号（圆形缺陷1、条形缺陷2、裂纹3、未熔合4、未焊透5、内凹6、咬边7）。

9.2 质量评级精度

质量评级精度表征系统对焊缝底片质量评级的综合能力，用于判断系统是否会发生漏判及误判。

9.2.1 合格底片检出率按公式（8）计算。

$$GDR= \frac{GD}{GD+FJ}\cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \left(8\right)$$

式中:

$GDR$——合格底片检出率；

$GD$——标注样本为合格底片，预测样本判断为合格底片的数量；

$FJ$——标注样本为合格底片，预测样本判断为不合格底片的数量。

9.2.2 不合格底片检出率按公式（9）计算。

$$NDR= \frac{ND}{ND+MJ}\cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \cdots \left(9\right)$$

式中:

$NDR$——不合格底片检出率；

$ND$——标注样本为不合格底片，预测样本判断为不合格底片的数量；

$MJ$——标注样本为不合格底片，预测样本判断为为合格底片的数量。

* 1. 评价结果

10.1 系统分级

按照表1将缺陷自动识别系统分为L1~L4四个等级，其中L1、L2、L3为辅助评片系统，L4为自动评片系统。

表1 系统分级表

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 精度指标系统级别 | L1 | L2 | L3 | L4 |
| 重点关注缺陷检出率(KDR) | ≥95% | ≥98% | 100% | 100% |
| 综合检出率(WDR) | ≥92% | ≥95% | ≥98% | 100% |
| 综合正检率(TDR) | ≥85% | ≥92% | ≥96% | ≥98% |
| 平均定量精度（$\overline{Da}$） | 85% | 90% | 92% | 95% |
| 不合格底片检出率（NDR） | ≥90% | ≥95% | 100% | 100% |
| 合格底片检出率（GDR） | ≥85% | ≥90% | ≥95% | 100% |

10.2 风险分析

 对缺陷自动识别系统进行风险分析，分析导致系统预测缺陷错误的原因和潜在的使用风险。分析数据包括评价指标计算结果和图像。

10.2.1 缺陷漏判风险

漏判风险级别评定按照表2进行，其中风险级别由高到低依次为Ⅰ类、Ⅱ类风险。

表2 漏判风险级别评定表

|  |  |  |  |
| --- | --- | --- | --- |
| 风险级别 | 系统级别 | 精度指标 | 潜在漏检风险 |
| I类风险 | L1 | 重点关注缺陷检出率(KDR)≥95%综合检出率(WDR)≥92%不合格底片检出率（NDR）≥90% | 不合格的底片漏判。 |
| L2 | 重点关注缺陷检出率(KDR)≥98%综合检出率(WDR)≥95%不合格底片检出率（NDR）≥95% |
| Ⅱ类风险 | L3 | 重点关注缺陷检出率(KDR)=100%综合检出率(WDR)≥98%不合格底片检出率（NDR）=100% | 存在缺陷漏检，但不会发生不合格底片漏判。 |

10.2.2 缺陷误判风险

误判风险级别评定按照表3进行，其中风险级别由高到低依次为Ⅰ类、Ⅱ类风险。

表3 误判风险级别评定表

|  |  |  |  |
| --- | --- | --- | --- |
| 风险级别 | 系统级别 | 精度指标 | 潜在误判风险 |
| I类风险 | L1 | 合格底片检出率（GDR）≥85% | 最多15%的合格底片误判为不合格底片，需要人工复核。 |
| L2 | 合格底片检出率（GDR）≥90% |
| Ⅱ类风险 | L3 | 合格底片检出率（GDR）≥95% | 最多5%的合格底片误判为不合格底片，需要人工复核。 |

10.3 出具评价报告

按照本文件评价的缺陷自动识别系统需出具评价报告。

10.3.1 评价报告至少包含评价人员、基本要求达标情况、标准测试集数量、缺陷类型及占比、焊缝形式、单类别缺陷识别精度、综合缺陷识别精度、系统分级结果、风险分析结果等内容。

10.3.2 评价报告模板参考见附录A。

1.

（资料性）

A.1 评价报告模板

* 1. 缺陷自动识别系统评价报告模板

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 系统名称 |  | 系统版本号 |  | 评价日期 |  |
| 开发单位 |  | 系统负责人 |  | 联系方式 |  |
| 单位地址 |  |
| 评价人员 | 姓名 |  | 姓名 |  |
| 持证类型 |  | 持证类型 |  |
| 标准底片类型 | 检测方法 |  | 透照布置 |  | 焊缝形式 |  | 焊接方法 |  |
| 缺陷测试集集缺陷数量（张） |  | 各类缺陷占比 | 裂纹： %；未熔合： %；未焊透： %；圆形缺陷： %；条形缺陷： %；咬边： %；内凹： % |
| 评级测试集底片数量（张） |  | 各级底片占比 | Ⅰ级片： %；Ⅱ级片： %；Ⅲ级片： %；Ⅳ级片： %。 |
| 标注样本人员 | 姓名 |  | 姓名 |  | 姓名 |  |
| 持证类型 |  | 持证类型 |  | 持证类型 |  |
| 缺陷检出精度 | *TDRn* | 裂纹： %；未熔合： %；未焊透： %；圆形缺陷： %；条形缺陷： %；咬边： %；内凹： % |
| *FDRn* | 裂纹： %；未熔合： %；未焊透： %；圆形缺陷： %；条形缺陷： %；咬边： %；内凹： % |
| *MDRn* | 裂纹： %；未熔合： %；未焊透： %；圆形缺陷： %；条形缺陷： %；咬边： %；内凹： % |
| *KDR* |  |
| *WDR* |  |
| *TDR* |  |
| $$\overbar{Da}$$ |  |
| 质量评级精度 | *GDR* |  |
| *NDR* |  |
| 系统分级结果 | □ L1；□ L2；□ L3；□ L4 |
| 风险分析结果 | 缺陷漏判风险：□Ⅰ类；□Ⅱ类缺陷误判风险：□Ⅰ类；□Ⅱ类 |
| 评价人签字 |  | 评价日期 |  |