

团 体 标 准

T/CWAN 0174—2026

焊接接头射线探伤缺陷智能识别模型评价方法

Evaluation method for an intelligent defect-recognition model in radiographic testing
of welding joints

2026-01-07 发布

2026-02-01 实施

中国焊接协会 发布

目 录

前 言	II
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 人员要求	3
6 基本要求	3
7 流程	3
8 标准测试集	4
9 测试	5
10 指标	8
11 结果	10

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。

本文件由中国焊接协会提出并归口。

本文件起草单位：北京理工大学（珠海）、北京博清科技有限公司、中焊科技发展（哈尔滨）有限公司、天津市特种设备监督检验技术研究院、重庆科技大学、南昌航空大学、上海中巽科技股份有限公司、江苏北人智能制造科技股份有限公司、天津大学、唐山松下产业机器有限公司、哈尔滨职业技术大学、中国兵器工业集团航空弹药研究院有限公司、北部湾大学、哈尔滨华德学院、南昌职业大学、黑龙江科技大学、坤智大数据科技（哈尔滨）有限公司、山东交通学院、福建省特种设备检验研究院、内蒙古机电职业技术学院。

本文件主要起草人：于兴华、李琰、冯消冰、武鹏博、杜非、冯伟、陈玉华、王铭秋、林涛、刘晨曦、刘金龙、刘建国、范东辉、马青军、邓军林、李长威、张大林、王永东、于修和、郝亮、李爱民、王滨滨、孟君晟、孙明辉、张丽红。

焊接接头射线探伤缺陷智能识别模型评价方法

1 范围

本文件规定了焊接接头射线探伤缺陷智能识别模型评价的人员要求、基本要求、流程、标准测试集、测试、指标、结果等内容。

本文件适用于承压设备金属熔化焊焊接接头射线检测底片缺陷自动识别系统的评价,其余焊接场景可参照适用。

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 3375 焊接术语

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 26141.1 无损检测 射线照相底片数字化系统的质量鉴定

GB/T 28452 信息安全技术 应用软件系统通用安全技术要求

GB/T 45288 人工智能大模型

NB/T 47013.2 承压设备无损检测 第2部分:射线检测

NB/T 47013.11 承压设备无损检测 第11部分:射线数字成像检测

NB/T 47013.14 承压设备无损检测 第14部分:射线计算机辅助成像检测

TSG Z8001 特种设备无损检测人员考核规则

3 术语和定义

GB/T 3375 和 GB/T 45288 界定的以及下列术语和定义适用于本文件。

3.1

缺陷自动识别系统 automatic defect recognition system

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 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 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

数字底片 digital film

通过射线检测（Radiographic Testing, RT）拍摄的物理底片经底片数字化系统扫描后的数字图像、通过射线数字成像（DR）检测直接获取的数字图像、通过射线计算机辅助成像（CR）检测扫描后的数字图像。

3.5

标准底片 standard film

用于测试及评价系统性能的数字底片。

3.6

标注样本 label samples

对标准底片进行缺陷标注生成的标注文件，包含缺陷类型及缺陷位置。

3.7

预测样本 prediction samples

由缺陷自动识别系统预测的缺陷识别结果文件，包含缺陷类型及缺陷位置。

3.8

训练集 training set

使用人工智能算法训练缺陷识别模型的数字底片集合及对应的标注文件。

3.9

标准测试集 standard test set

用于测试及评价系统性能的数字底片集合及对应的标注样本。

3.10

正检 true detection

标注样本中的缺陷位置及缺陷类型被正确识别。

3.11

误检 false detection

标注样本中的缺陷位置被正确识别，但缺陷类型未被正确识别。

3.12

漏检 miss detection

标注样本中的缺陷位置未被正确识别。

3.13

误报 false report

标注样本中无缺陷，而被系统错误识别为有缺陷。

3.14

交并比 intersection over union

同一缺陷标注区域与预测区域的交集比并集的值。

3.15

风险分析 risk analysis

对系统未正确识别的数字底片进行分析，以明确导致系统预测错误的原因和潜在的风险，在风险分析过程中，需要关注系统未正确识别图像的缺陷类型、缺陷尺寸、缺陷评定级别等。

4 缩略语

下列缩略语适用于本文件。

AssistDR: 辅助评片 (Assisted Defect Recognition)

AutoDR: 自动评片 (Automated Defect Recognition)

Label: 标注样本 (Label Samples)

Prediction: 预测样本 (Prediction Sample)

TD: 正检 (True Detection)

FD: 误检 (False Detection)

MD: 漏检 (Miss Detection)

FR: 误报 (False Report)

IOU: 交并比 (Intersection Over Union)

RT: 射线检测 (Radiographic Testing)

DR: 数字射线 (Digital Radiography)

CR: 射线计算机辅助成像检测 (Computed Radiography)

5 人员要求

评价人员应按 TSG Z8001 进行考核并取得 RT (D) II 级及以上资格证书。

标准测试集标注人员应按 TSG Z8001 规定取得 RT II 级及以上证书。

6 基本要求

6.1 功能要求

6.1.1 系统宜具备显示数字底片的功能，数字底片的格式包括但不限于 DICONDE 及 TIF 格式。

6.1.2 系统宜具备缺陷自动识别功能，包括但不限于自动识别有无缺陷、缺陷类型、缺陷位置等功能。

6.1.3 系统宜具备按照 NB/T 47013.2 的规定自动对缺陷识别结果进行评级并输出报告的功能。

6.1.4 系统宜具备自动识别结果人工复核功能，包括但不限于复核评级结果、缺陷增删、缺陷类型、缺陷位置等内容。

6.1.5 系统宜具备但不限于图像降噪、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。

6.1.6 系统宜具备但不限于数字底片的缺陷尺寸、归一化信噪比、空间分辨率、灰度、图像灵敏度等测

量功能。

6.2 运行环境

6.2.1 系统部署方式应支持云端部署或本地部署。

6.2.2 硬件应满足系统软件安装及运行。

6.2.3 CPU 和内存模块应根据系统处理的任务规模、复杂程度、单位时间访问量及系统响应时间等要求进行配置。

6.2.4 存储模块的选择应根据系统内文件、数据量、扩展性及数据读写速度等要求进行配置。

6.2.5 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 47013.11 的要求。

6.3 系统可靠性

系统可靠性应符合 GB/T 25000.51 的要求。

6.4 系统安全性

系统安全性应符合 GB/T 28452 的要求。

6.5 系统更新

系统更新后应按照本文件重新评价。

7 流程

包括标准测试集、评价测试、评价指标及评价结果4个步骤，如图1所示。

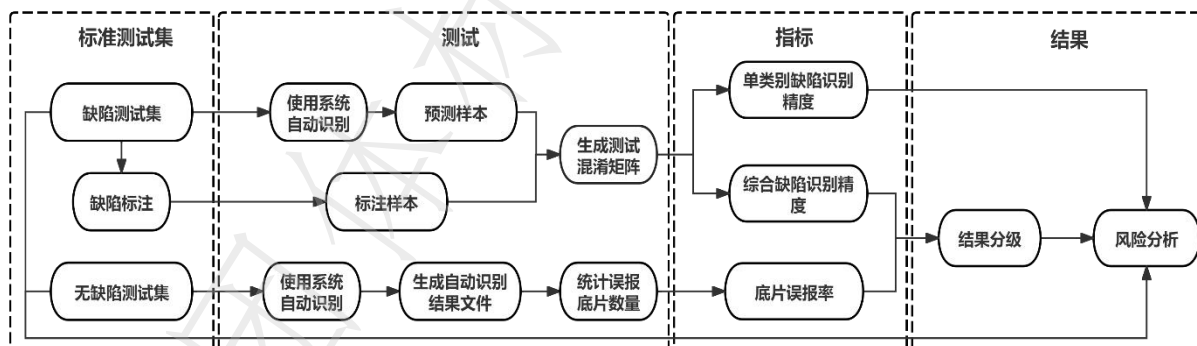


图1 评价流程

8 标准测试集

8.1 底片质量

8.1.1 RT 底片

8.1.1.1 通过 X 射线检测 (RT) 获得的物理底片质量应满足 NB/T 47013.2 要求。

8.1.1.2 底片数字化系统的性能应不低于 GB/T 26141.1 中规定的 DS 级别要求。

8.1.1.3 扫描后的数字底片每英寸长度内的像素点数 (dpi) 应不低于 600, 数字分辨率应不低于 16 bit, 像质计丝号与原 X 射线底片一致。

8.1.1.4 底片评定区范围内应无底片数字化过程中造成的划痕、灰尘杂质、光学污染等伪影像。

8.1.2 DR 及 CR 底片

DR 数字底片质量应符合 NB/T 47013.11 的相关规定，CR 数字底片质量应符合 NB/T 47013.14 的相关规定。

8.2 底片类型

8.2.1 承压设备熔化焊（双面焊）焊接接头测试集。

8.2.1.1 测试集包括缺陷测试集与无缺陷测试集。

8.2.1.2 缺陷测试集中的每张底片都应有缺陷，以 150 mm 为单张底片标准长度。缺陷测试集的缺陷总数不少于 500 个。

8.2.1.3 缺陷测试集中应至少包含裂纹、未熔合、未焊透、圆形缺陷、条形缺陷等类型缺陷，其中裂纹、未熔合、未焊透为重点关注缺陷，圆形缺陷和条形缺陷为一般关注缺陷。

8.2.1.4 缺陷测试集中每类缺陷的数量应基本相当。

8.2.1.5 无缺陷测试集中的每张底片都应无缺陷，无缺陷测试集底片数量与缺陷测试集底片数量一致。

8.2.2 承压设备熔化焊（单面焊）焊接接头测试集

8.2.2.1 测试集包括缺陷测试集与无缺陷测试集。

8.2.2.2 缺陷测试集中的每张底片都应有缺陷，以 150 mm 为单张底片标准长度（除小径管接头外）。缺陷测试集的缺陷总数不少于 500 个。

8.2.2.3 缺陷测试集中应至少包含裂纹、未熔合、未焊透、圆形缺陷、条形缺陷、内凹、咬边等类型缺陷。其中裂纹、未熔合、未焊透、内凹、咬边为重点关注缺陷，圆形缺陷和条形缺陷为一般关注缺陷。

8.2.2.4 缺陷测试集中每类缺陷的数量应基本相当。

8.2.2.5 无缺陷测试集中的每张底片都应无缺陷，无缺陷测试集底片数量与缺陷测试集底片数量一致。

8.3 构建

8.3.1 测试集与训练集的数据互斥性：测试数据集中不应包含训练数据集中的数据。

8.3.2 测试集的隐私性：测试集应考虑企业隐私泄露，采取数据脱敏处理措施。

8.4 标注样本

8.4.1 由缺陷标注人员对缺陷测试集中的标准底片依次进行缺陷标注，生成标注样本。标注样本为用于系统精度的对比样本。

8.4.2 标注样本的标注流程如图 2 所示，由 A、B、C 三名标注人员独立标注，对比三人标注的结果，当标注缺陷类型一致且缺陷范围的交并比 $IOU(A \& B \& C) \geq 0.5$ 时，标注样本为 A、B、C 的并集 $(A \cup B \cup C)$ 。当标注缺陷类型不一致或缺陷范围的交并比 $IOU(A \& B \& C) < 0.5$ 时，引入更高级别的专家（如 RT III 级）进行最终裁定。

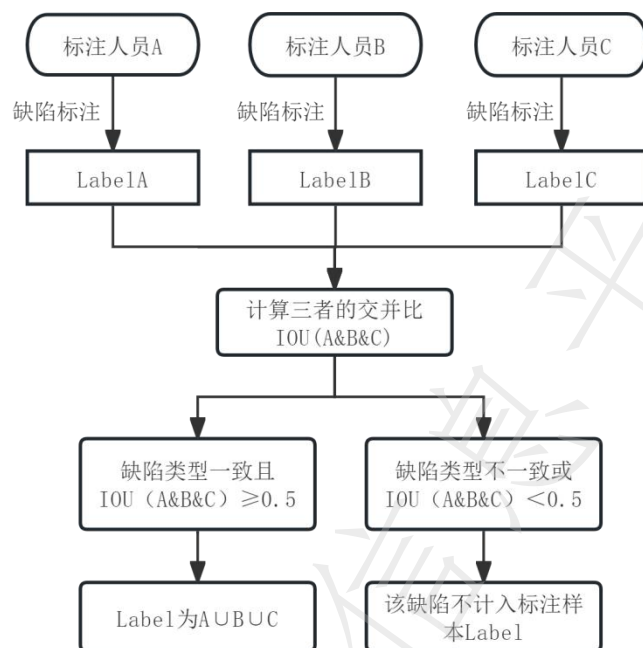


图2 标注样本标注流程

9 测试

9.1 预测样本

9.1.1 将缺陷测试集中的数字底片输入系统进行缺陷识别，生成预测样本。

9.1.2 将无缺陷测试集中的数字底片输入系统进行识别，分别统计系统识别为有缺陷的底片数量 a 和无缺陷的底片数量 b 。

9.2 偏离程度

将标注样本与预测样本进行对比，表征两个样本的偏离程度有以下四种情况：正检、误检、漏检及误报。

9.2.1 当 $IOU \geq 0.1$ ，标注样本与预测样本缺陷类型一致时，样本偏离程度表征为正检。

9.2.2 当 $IOU \geq 0.1$ ，标注样本与预测样本缺陷类型不一致时，样本偏离程度表征为误检。

9.2.3 当 $IOU < 0.1$ ，样本偏离程度表征为漏检。

9.2.4 在无缺陷测试集上测试，识别出缺陷的情况，样本偏离程度表征为误报。

9.3 混淆矩阵

根据 9.2 中的判定条件，统计每类缺陷的正检、误检、漏检及误报数量，绘制测试结果混淆矩阵，如图 3 和图 4 所示。

预测样本	1	圆形缺陷	TD1	FD2	FD3	FD4	FD5	FD6	FD7
	2	条形缺陷	FD1	TD2	FD3	FD4	FD5	FD6	FD7
	3	裂纹	FD1	FD2	TD3	FD4	FD5	FD6	FD7
	4	未熔合	FD1	FD2	FD3	TD4	FD5	FD6	FD7
	5	未焊透	FD1	FD2	FD3	FD4	TD5	FD6	FD7
	6	内凹	FD1	FD2	FD3	FD4	FD5	TD6	FD7
	7	咬边	FD1	FD2	FD3	FD4	FD5	FD6	TD7
	8	无缺陷	MD1	MD2	MD3	MD4	MD5	MD6	MD7
		圆形缺陷	条形缺陷	裂纹	未熔合	未焊透	内凹	咬边	
		1	2	3	4	5	6	7	
		标注样本							

图3 混淆矩阵表（缺陷测试集）

标注样本	无缺陷	FR1	FR2	FR3	FR4	FR5	FR6	FR7
	圆形缺陷	条形缺陷	裂纹	未熔合	未焊透	内凹	咬边	
	1	2	3	4	5	6	7	
		预测样本						

图4 混淆矩阵表（无缺陷测试集）

测试结果混淆矩阵表中，使用 TD_n、FD_n、MD_n、FR_n 分别表示每一类缺陷的正检数量、误检数量、漏检数量及误报数量，n 为缺陷序号。

10 指标

10.1 单类别缺陷识别精度

10.1.1 单类别缺陷正检率按公式（1）计算。

$$TDR_n = \frac{TD_n}{TD_n + FD_n + MD_n} \times 100\% \dots\dots\dots (1)$$

式中：

TDR_n——序号为第n类的缺陷正检率；

n——缺陷序号（1代表圆形缺陷、2代表条形缺陷、3代表裂纹、4代表未熔合、5代表未焊透、6代表内凹、7代表咬边）。

10.1.2 单类别缺陷误检率按公式（2）计算。

$$FDRn = \frac{FDn}{TDn+FDn+MDn} \times 100\% \quad (2)$$

式中：

$FDRn$ ——序号为第 n 类的缺陷误检率；

n ——缺陷序号（1代表圆形缺陷、2代表条形缺陷、3代表裂纹、4代表未熔合、5代表未焊透、6代表内凹、7代表咬边）。

10.1.3 单类别缺陷漏检率按公式（3）计算。

$$MDRn = \frac{MDn}{TDn+FDn+MDn} \times 100\% \quad (3)$$

式中：

$MDRn$ ——序号为第 n 类的缺陷漏检率；

n ——缺陷序号（1代表圆形缺陷、2代表条形缺陷、3代表裂纹、4代表未熔合、5代表未焊透、6代表内凹、7代表咬边）。

10.1.4 单类别缺陷误报率按公式（4）计算。

$$FRRn = \frac{FRn}{\sum_{n=1}^7 FRn} \times 100\% \quad (4)$$

式中：

$FRRn$ ——序号为第 n 类的缺陷误报率；

n ——缺陷序号（1代表圆形缺陷、2代表条形缺陷、3代表裂纹、4代表未熔合、5代表未焊透、6代表内凹、7代表咬边）。

10.2 综合缺陷识别精度

10.2.1 重点关注缺陷识别率按公式（5）计算。

$$KDR = \frac{\sum_{n=3}^7 (TDn+FDn)}{\sum_{n=3}^7 (TDn+FDn+MDn)} \times 100\% \quad (1)$$

式中：

KDR ——重点关注缺陷识别率；

n ——缺陷序号（3代表裂纹、4代表未熔合、5代表未焊透、6代表内凹、7代表咬边）。

10.2.2 综合缺陷识别率按公式（6）计算。

$$WDR = \frac{\sum_{n=1}^7 (TDn+FDn)}{\sum_{n=1}^7 (TDn+FDn+MDn)} \times 100\% \quad (2)$$

式中：

WDR ——综合缺陷识别率；

n ——缺陷序号（1代表圆形缺陷、2代表条形缺陷、3代表裂纹、4代表未熔合、5代表未焊透、6代表内凹、7代表咬边）。

10.2.3 综合正检率（TDR）按公式（7）计算。

$$TDR = \frac{\sum_{n=1}^7 TDn}{\sum_{n=1}^7 (TDn+FDn+MDn)} \times 100\% \quad (3)$$

式中：

TDR ——综合正检率；

n ——缺陷序号（1代表圆形缺陷、2代表条形缺陷、3代表裂纹、4代表未熔合、5代表未焊透、6代表内凹、7代表咬边）。

10.3 底片误报率

底片误报率（FRR）按公式（8）计算。

$$FRR = \frac{a}{a+b} \times 100\% \quad (1)$$

式中：

FRR ——底片误报率；

a ——在无缺陷测试集上，系统识别为有缺陷的底片数量；

b ——在无缺陷测试集上，系统识别为无缺陷的底片数量。

11 结果

11.1 系统分级

按照表 1 将系统分为 L1~L4 四个等级。

表 1 系统分级表

精度指标		系统级别			
		L1	L2	L3	L4
重点关注缺陷识别率(KDR)		≥95%	≥98%	99.9%	99.9%
综合缺陷识别率(WDR)		≥92%	≥95%	≥98%	99.9%
综合正检率(TDR)		≥85%	≥92%	≥96%	≥98%
底片误报率 (FRR)	自动焊	≤8%			≤3%
	手工焊	≤10%			≤4%

11.2 风险分析

对系统进行风险分析，分析导致系统预测缺陷错误的原因和潜在的使用风险。分析数据包括第 10 章中评价指标计算结果、人工标注缺陷原始图像和系统自动识别缺陷图像。

11.2.1 漏检风险

系统漏检风险级别按照表 2 评定，缺陷的评级参考 NB/T 47013.2 规定。如果存在多种风险级别，结果按较高风险级别评定，其中风险级别由高到低依次为 I 类、II 类风险。

表 2 漏检风险级别评定表

风险级别	内容	潜在风险
I 类风险	重点关注缺陷和按照 NB/T 47013.2 规定评定为 II 级及以上的一般关注缺陷漏检。	缺陷漏检并影响焊接接头合格与否的评定。
II 类风险	按照 NBT/47013.2 规定评定为 I 级的圆形缺陷漏检。	缺陷漏检但不影响焊接接头合格与否的评定。

11.2.2 误检风险

误检风险级别评定按照表 3 评定，如果评定存在多种风险级别，结果按较高风险级别评定，其中风险级别由高到低依次为 I 类、II 类风险。

表 3 误检风险级别评定表

风险级别	内容	潜在风险
I类风险	重点关注缺陷误检为一般关注缺陷。	不合格的焊缝漏检。
II类风险	一般关注缺陷误检为重点关注缺陷。	合格焊缝误判为不合格焊缝。

11.2.3 误报风险

误报风险级别按照表 4 评定，如果评定存在多种风险级别，结果按较高风险级别评定，其中风险级别由高到低依次为I类、II类风险。

表 4 误报风险级别评定表

风险级别	焊接方法	内容	潜在风险
I类风险	自动焊	$FRR > 8\%$	人工复核工作量大。
	手工焊	$FRR > 10\%$	
II类风险	自动焊	$FRR \leq 8\%$	
	手工焊	$FRR \leq 10\%$	

11.3 出具评价记录表

评价记录表至少包含评价人员、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。