

VEREIN  
DEUTSCHER  
INGENIEURE  
  
VERBAND DER  
ELEKTROTECHNIK  
ELEKTRONIK  
INFORMATIONSTECHNIK  
  
VERBAND DEUTSCHER  
MASCHINEN- UND  
ANLAGENBAU

Industrielle Bildverarbeitung  
Oberflächeninspektionssysteme in der  
Flachstahlproduktion  
Leistungsbewertung der Klassifikation  
Machine vision/industrial image processing  
Surface inspection systems in flat steel production  
Performance evaluation of classification

VDI/VDE/  
VDMA 2632  
Blatt 4.2 / Part 4.2  
  
Ausg. deutsch/englisch  
Issue German/English

Die deutsche Version dieser Richtlinie ist verbindlich.

The German version of this standard shall be taken as authoritative. No guarantee can be given with respect to the English translation.

Inhalt	Seite	Contents	Page
Vorbemerkung .....	2	Preliminary note.....	2
Einleitung.....	2	Introduction.....	2
<b>1 Anwendungsbereich.....</b>	<b>3</b>	<b>1 Scope.....</b>	<b>3</b>
<b>2 Normative Verweise.....</b>	<b>4</b>	<b>2 Normative references.....</b>	<b>4</b>
<b>3 Begriffe.....</b>	<b>4</b>	<b>3 Terms and definitions.....</b>	<b>4</b>
<b>4 Formelzeichen und Abkürzungen.....</b>	<b>4</b>	<b>4 Symbols and abbreviations.....</b>	<b>4</b>
<b>5 Eindeutige Rahmenbedingungen zur Nachvollziehbarkeit einer Leistungsbewertung.....</b>	<b>6</b>	<b>5 Clear general conditions for the traceability of a performance evaluation.....</b>	<b>6</b>
5.1 Voraussetzungen.....	7	5.1 Requirements.....	7
5.2 Fachkräfte für die Leistungsbewertung.....	8	5.2 Specialists for performance assessment.....	8
5.3 Festlegung des zu bewertenden Klassifikationsprozesses.....	8	5.3 Determination of the classification process to be evaluated.....	8
5.4 Zusammenstellung einer geeigneten Bewertungsmenge.....	9	5.4 Compilation of a suitable evaluation set.....	9
5.5 Bewertung auf Ereignisebene.....	10	5.5 Event level assessment.....	10
<b>6 Methode zur Leistungsbewertung.....</b>	<b>11</b>	<b>6 Performance assessment method.....</b>	<b>11</b>
6.1 Statistische Grundlagen.....	11	6.1 Statistical basics.....	11
6.2 Bestimmung der Precision.....	15	6.2 Determination of the precision.....	15
6.3 Bestimmung des Recalls.....	16	6.3 Determination of the recall.....	16
6.4 NK-Rate.....	17	6.4 nc rate.....	17
<b>7 Darstellung der Ergebnisse.....</b>	<b>17</b>	<b>7 Presentation of the results.....</b>	<b>17</b>
<b>8 Praktische Einordnung der Ergebnisse einer Leistungsbewertung.....</b>	<b>22</b>	<b>8 Practical integration of the results of a performance evaluation.....</b>	<b>22</b>
8.1 Bewertung von OIS-Ergebnissen.....	22	8.1 Evaluation of SIS results.....	22
8.2 Anwendungsfelder für OIS-Ergebnisse.....	24	8.2 Fields of application for SIS results.....	24
8.3 Leistungsanforderungen.....	24	8.3 Performance requirements.....	24
Schrifttum.....	28	Bibliography.....	28

VDI/VDE-Gesellschaft Mess- und Automatisierungstechnik (GMA)

Fachbereich Methodik der Mess- und Sensortechnik

VDI/VDE-Handbuch Optische Technologien  
VDI/VDE-Handbuch Automatisierungstechnik  
VDI/VDE-Handbuch Fertigungsmesstechnik

Frühere Ausgabe: 04.23 Entwurf  
Former edition: 04/23 Draft

Zu beziehen durch / Available at Beuth Verlag GmbH, 10772 Berlin – Alle Rechte vorbehalten / All rights reserved © Verein Deutscher Ingenieure e.V., Düsseldorf 2024

Vervielfältigung – auch für innerbetriebliche Zwecke – nicht gestattet / Reproduction – even for internal use – not permitted

## Vorbemerkung

Der Inhalt dieser Richtlinie ist entstanden unter Beachtung der Vorgaben und Empfehlungen der Richtlinie VDI 1000.

Alle Rechte, insbesondere die des Nachdrucks, der Fotokopie, der elektronischen Verwendung und der Übersetzung, jeweils auszugsweise oder vollständig, sind vorbehalten.

Die Nutzung dieser Richtlinie ist unter Wahrung des Urheberrechts und unter Beachtung der Lizenzbedingungen ([www.vdi.de/richtlinien](http://www.vdi.de/richtlinien)), die in den VDI-Merkblättern geregelt sind, möglich.

Allen, die ehrenamtlich an der Erarbeitung dieser Richtlinie mitgewirkt haben, sei gedankt.

Eine Liste der aktuell verfügbaren und in Bearbeitung befindlichen Blätter dieser Richtlinienreihe sowie gegebenenfalls zusätzliche Informationen sind im Internet abrufbar unter [www.vdi.de/2632](http://www.vdi.de/2632).

## Einleitung

Ein wichtiger Aspekt der Industrie 4.0 ist es, Daten von Sensoren, Mess- und Prüfsystemen nicht nur zur Automatisierung des Fertigungsprozesses oder zur Qualitätskontrolle einzusetzen, sondern diese Daten auch zur kontinuierlichen Prozessoptimierung und damit auch für strategische Unternehmensentscheidungen zu nutzen. Dadurch steigt der Anspruch an die Zuverlässigkeit der Ergebnisse. Das gilt insbesondere dann, wenn bisher die finale Qualitätsentscheidung vom Personal der Qualitätssicherung und nicht automatisch von den Systemen getroffen wurde. Auch wenn Rückschlüsse auf die Zuverlässigkeit des Fertigungsprozesses getroffen werden sollen, müssen die dazu ausgewerteten Daten der Inspektionssysteme zuverlässig sein.

Für die Kontrolle der Bandoberflächen stehen in der Flachstahlproduktion Oberflächeninspektionssysteme zur Verfügung, die nicht nur die Klassifikationsergebnisse, sondern auch die Position der klassifizierten Ereignisse auf dem Band, einschließlich eines Bilds davon, ausgeben. Damit steht grundsätzlich eine sehr große Datenmenge für übergeordnete Analysen zur Verfügung.

Nicht nur vor dem Hintergrund einer erweiterten Datenanalyse im Industrie-4.0-Kontext, sondern auch im Rahmen der eigentlichen Inspektionsaufgabe ist es wichtig, die Leistungsfähigkeit eines Oberflächeninspektionssystems mit angemessenem Aufwand quantifizieren zu können. Mit diesen Aussagen zur Leistungsfähigkeit kann beispielsweise die Spezifikation des Systems im Lasten- und Pflichtenheft vereinfacht, der Abnahmeprozess

## Preliminary note

The content of this standard has been developed in strict accordance with the requirements and recommendations of the standard VDI 1000.

All rights are reserved, including those of reprinting, reproduction (photocopying, micro copying), storage in data processing systems and translation, either of the full text or of extracts.

The use of this standard without infringement of copyright is permitted subject to the licensing conditions ([www.vdi.de/richtlinien](http://www.vdi.de/richtlinien)) specified in the VDI Notices.

We wish to express our gratitude to all honorary contributors to this standard.

A catalogue of all available parts of this series of standards and those in preparation as well as further information, if applicable, can be accessed on the Internet at [www.vdi.de/2632](http://www.vdi.de/2632).

## Introduction

An important aspect of Industry 4.0 is to use data from sensors, measuring and testing systems not only to automate the manufacturing process or for quality control, but also to use this data for continuous process optimisation and thus also for strategic corporate decisions. This increases the demand on the reliability of the results. This is especially true if the final quality decision was previously made by quality management staff and not automatically by the systems. The inspection system data shall be also reliable if they are evaluated to draw conclusions on the reliability of the production process.

For the inspection of coil surfaces, surface inspection systems are available in flat steel production that not only provide the classification results, but also the position of the classified events on the coil, including an image of them. This basically means that a very large amount of data is available for higher-level analyses.

Not only against the background of extended data analysis in the Industry 4.0 context, but also within the framework of the actual inspection task, it is important to be able to quantify the performance of a surface inspection system with reasonable effort. These statements on performance can be used, for example, to simplify the specifications of the systems, standardise the acceptance process or monitor the performance of the systems during operation.

standardisiert oder die Leistungsfähigkeit der Systeme im laufenden Betrieb überwacht werden.

In VDI/VDE/VDMA 2632 Blatt 3.1 werden bereits verschiedene Methoden zur Prüfung von klassifizierenden Bildverarbeitungssystemen beschrieben. Diese Methoden sind aber mit den Rahmenbedingungen, wie sie in der Flachstahlproduktion typisch sind, nicht praktikabel. Zu diesen Rahmenbedingungen zählen insbesondere:

- Es existiert eine sehr große Anzahl zu klassifizierender Ereignisse. Praktisch ist es daher unmöglich, jedem Ereignis eine „wahre“ Klasse zuzuordnen.
- Es fehlen trennende Merkmale für eine eindeutige Klassenzuweisung.
- Es gibt einen hohen Anteil von nicht qualitätsrelevanten Ereignissen.
- Die Bewertung des Produkts als „gut“ kann zulässig sein, auch wenn detektierte Einzelereignisse als Fehler klassifiziert wurden.
- Produkte und/oder Produktbereiche können abhängig von den Klassifikationsergebnissen in verschiedene Qualitätsklassen eingeteilt oder für unterschiedliche Einsatzzwecke freigegeben werden. Es gibt nicht nur binäre „gut/schlecht“-Entscheidungen.
- Eine Zurückstellung von Proben für spätere Vergleichsmessungen ist nicht möglich.

Die vorliegende Richtlinie stellt ein Verfahren für die Leistungsbewertung von Inspektionssystemen vor, das unter den Rahmenbedingungen der Flachstahlproduktion praktikabel ist.

## 1 Anwendungsbereich

Die Richtlinie wendet sich an alle, die als Anbieter oder Anwender von Oberflächeninspektionssystemen in der Flachstahlproduktion die Klassifikationsleistung erfassen wollen, um damit eine Leistungsaussage über das Oberflächeninspektionssystem zu treffen. Diese Aussage kann zum Nachweis einer vereinbarten Leistung, zum Vergleich unterschiedlicher Systemkonfigurationen oder zum Funktionsnachweis im laufenden Betrieb eingesetzt werden.

Das in dieser Richtlinie vorgestellte Verfahren kann auf andere klassifizierende Inspektionssysteme, bei denen die inspizierten Produkte ähnliche Randbedingungen haben, übertragen werden.

Verallgemeinert kann das Verfahren zur Leistungsbewertung auf alle Klassifikationsaufgaben angewandt werden, bei denen die „Grundwahrheit“, also eine Aussage über die wahre Verteilung der Ereignis-

In VDI/VDE/VDMA 2632 Part 3.1, various methods for testing classifying machine vision systems are already described. However, these methods are not practicable with the general conditions that are typical in flat steel production. These general conditions include in particular:

- There is a very large number of events to be classified. Practically, it is therefore impossible to assign a “true” class to every event.
- There is a lack of separating features for a unique class assignment.
- There is a high ratio of events that are not quality relevant.
- The assessment of the product as “good” can be acceptable even if detected individual events were classified as defects.
- Products and/or product areas can be classified into different quality classes or released for different purposes depending on the classification results. Not only binary “good/bad” decisions are made.
- It is not possible to defer samples for later reproducibility measurements.

This standard presents a procedure for the performance evaluation of inspection systems that is practicable under the general conditions of flat steel production.

## 1 Scope

The standard is intended for anyone who, as a supplier or user of surface inspection systems in flat steel production, wishes to acquire the classification performance to make a performance statement about the surface inspection system. This statement can be used to prove an agreed performance, to compare different system configurations or to prove functionality during operation.

The procedure presented in this standard can be transferred to other classifying inspection systems where the inspected products have similar boundary conditions.

Generalised, the performance evaluation procedure can be applied to all classification tasks where the “ground truth”, i.e., a statement about the true distribution of the events among the different classes,

nisse auf die verschiedenen Klassen, nicht möglich ist. Diese Situation trifft in der Praxis immer dann zu, wenn die Anzahl der zu berücksichtigenden Ereignisse sehr groß und eine Überprüfung der Klassifikation in jedem Einzelfall daher nicht möglich ist.

is not possible. This situation always applies in practice when the number of events to be considered is very large and it is therefore not possible to check the classification in each individual case.