

UDC 004.932:620.179.152

DIGITAL RADIOGRAPHIC IMAGES PRELIMINARY PROCESSING FOR
GENERATING TRAINING SAMPLESE. E. KOVSHOV¹, V. S. KUVSHINNIKOV¹, A. V. SUDAREV^{1,2}¹Joint-stock company «Research and Development Institute of Construction
Technology – Atomstroy»²Dmitry Mendeleev University of Chemical Technology of Russia
Moscow, Russia

For any modern industrial facility or technological machines and equipment, regardless of its industry affiliation, safety and reliability are critical factors. Various types and methods of non-destructive testing (NDT) are used to prevent structural failure, especially in the field of welded joints. Among the existing NDT methods, special attention is paid to main three: visual and measuring, ultrasonic, and radiographic. The latter, representing the radiation type of NDT, is the most accurate and objective, allowing not only to localize the defect but also to accurately determine its size and type.

However, the complexity, high cost and high volume of data requiring processing and analysis, together, actualize the problem of digitalization of the process of obtaining, processing and decoding radiographic images. A significant breakthrough has been made by ASTM, which in 1999 developed the specialized DICONDE (Digital Imaging and Communication in Non-Destructive Evaluation) format, allowed solving the tasks of digitalization related to obtaining, storing and transmitting the results of NDT. In turn, the active development and integration of artificial neural networks (ANN) into various industries creates a foundation for solving the remaining tasks for automating the processing and interpretation of radiographic NDT results.

It should be noted that the use of ANN in the field of radiation-based NDT is currently hindered by a severe lack of relevant and high-quality training data. Additionally, most scientific research in this area doesn't explore the emerging correlations between the test results and the conditions under which the radiographic images were acquired and reflected in the metadata attached to the images in the form of tags in the DICONDE file structure. Performing radiographic image quality assessment, checking the integrity and completeness of tags, and labeling images are necessary preprocessing steps before full-fledged neural network training. Using the PACS (Picture Archiving and Communication System) and its integrated database allows the automation for generating training samples of high-quality and relevant data.

Based on the results researches, quantitative mathematical models, algorithms, and application software for an automated radiographic image quality assessment system have been developed, as well as a structure for storing their metadata. The developed and verified mathematical model strictly formalizes the algorithms' requirements for assessing the quality of digital radiographic images in accordance with national and international standards, ensuring the accuracy and reproducibility of the results obtained.

Considering the DICONDE file structure, the relational database entities are the DICONDE hierarchical units: Component, Study, Series, and Instance. This database structure allows for the inclusion of attributes that are not searchable in PACS. After generating a selection of records that meet the query criteria, the database generates a path to each record in PACS using unique identification numbers (UID). The effectiveness of the developed data structure has been confirmed by a number of computational experiments, but it is worth noting the need for testing on large data sets or in real industrial settings.

The main four stages of automatic preliminary processing of radiographic DICONDE images and formation of training samples based on them have been determined: 1 – Data generation; 2 – Validation of images and metadata; 3 – Storage in the database; 4 – Formation of the training sample.

Training sample is formed based on the specified parameters using semantic search. This allows ANN model to be trained using not only radiographic images, but also factual information that characterizes both the test object and the conditions of radiographic technics itself.

The proposed approach ensures that ANN is invariant to the material of the test object, the exposure parameters, and the radiographic technics, which helps to reduce the dimensionality of the parameter space and improve the generalization ability of the model.

Quality assessment of a radiographic image requires the localization and analysis of key elements such as the groove, wire, and step sensitivity indicators, the two-wire sensitivity standard, and other auxiliary elements. The task of segmenting key image objects has been solved, and the information about their location is stored in the corresponding DICONDE tags and markup file for subsequent training of the ANN. Software implementation of the automated system involves its application for collecting, processing, and storing data from radiation NDT in the industrial sector in order to generate representative datasets necessary for the effective training of neural networks.

The key feature of this system is the integration of automated image quality assessment and metadata completeness verification, which allows for the creation of high-quality and reliable training datasets suitable for the effective training of machine learning models and reduces the risk of artifacts in the datasets.

Proposed automated approach to the formation of training samples based on synthetic data in the DICONDE format, generated using the industrial radiography simulator, allows to effectively solve the problems of the limited and unbalanced nature of real data of radiation NDT. The creation of representative samples with specified characteristics opens the way to the training of neural network models with greater generalization ability. An algorithmic approach to object segmentation in an image (without considering defects) and its quality assessment eliminates the probabilistic factor of results inherent in ANN and returns a strictly deterministic result.

The conducted research confirms the scientific validity and practical significance of the proposed approach to automatic preliminary preprocessing of digital radiographic DICONDE images, which creates a solid foundation for the further development of machine learning technologies in the field of NDT.