Contract No:

This document was prepared in conjunction with work accomplished under Contract No. DE-AC09-08SR22470 with the U.S. Department of Energy (DOE) Office of Environmental Management (EM).

Disclaimer:

This work was prepared under an agreement with and funded by the U.S. Government. Neither the U. S. Government or its employees, nor any of its contractors, subcontractors or their employees, makes any express or implied:

1) warranty or assumes any legal liability for the accuracy, completeness, or for the use or results of such use of any information, product, or process disclosed; or
2) representation that such use or results of such use would not infringe privately owned rights; or
3) endorsement or recommendation of any specifically identified commercial product, process, or service.

Any views and opinions of authors expressed in this work do not necessarily state or reflect those of the United States Government, or its contractors, or subcontractors.

A U.S. DEPARTMENT OF ENERGY NATIONAL LABORATORY • SAVANNAH RIVER SITE • AIKEN, SC

FY20 Progress Report: SRNL Analysis of ICCWR LCM and WAMS data for Corrosion and Cracking

Bruce Hardy (SRNL)

Anna d'Entremont (SRNL)

Jason Bakos (University of South Carolina)

Taylor Clingenpeel (University of South Carolina)

Michael Martínez-Rodríguez (SRNL)

Brenda García-Díaz (SRNL)

September 30, 2020

SRNL-STI-2020-00355, Revision 0

DISCLAIMER

Printed in the United States of America

Prepared for
U.S. Department of Energy

Keywords: *3013 Corrosion, 3013 Cracking, 3013 Surveillance Program, Machine Learning 3013 Corrosion*

Retention: *Varies*

FY20 Progress Report: SRNL Analysis of ICCWR LCM and WAMS data for Corrosion and Cracking

Bruce Hardy (SRNL)

Anna d'Entremont (SRNL)

Jason Bakos (University of South Carolina)

Taylor Clingenpeel (University of South Carolina)

Michael Martínez-Rodríguez (SRNL)

Brenda García-Díaz (SRNL)

September 30, 2020

Prepared for the U.S. Department of Energy under contract number DE-AC09-08SR22470.

Savannah River National Laboratory®

OPERATED BY SAVANNAH RIVER NUCLEAR SOLUTIONS

REVIEWS AND APPROVALS

AUTHORS:

Bruce Hardy, SRNL — Oct. 12, 2020

Anna d'Entremont, SRNL — Oct. 12, 2020

Jason Bakos, University of South Carolina — Oct. 19, 2020

Taylor Clingenpeel, University of South Carolina — Oct. 21, 2020

Michael Martínez-Rodríguez, SRNL — Oct. 16, 2020

Brenda García-Díaz, SRNL — Oct. 20, 2020

TECHNICAL REVIEW:

Emmanuel Perez, SRNL — Oct. 21, 2020

APPROVAL:

Patricia Lee, Manager
SRNL/Advanced Modeling, Simulation and Analytics — Oct. 16, 2020

Marissa Reigel, Manager
SRNL/Actinide Materials Science and Technology — Oct.21, 2020

EXECUTIVE SUMMARY

The development of algorithms for machine learning and data analysis for the 3013 Surveillance Program is a collaborative effort by the Savannah River National Laboratory (SRNL) and the University of South Carolina (USC). For corrosion detection, Laser Confocal Microscope (LCM) or Wide Area 3D Measurement System (WAMS) data is extracted from large binary files, with software written to convert the data to physical attributes (e.g. height, color and grayscale values; all as functions of a location in a plane projection). It is the objective of this project to produce a user-friendly interface that incorporated all operations needed to perform surface examination. For this reason, a Matlab-based Graphical User Interface (GUI) was created to integrate data input with software developed for processing and evaluation. In summary, the GUI permits selective downloading of binary data, interrogation of attributes, data labeling, flagging of significant features, execution of Machine Learning (ML) algorithms, output of parameters for trained ML algorithms, reports of ML model accuracy with respect to labeled data, and generation of graphical representations of various analyses.

Machine learning algorithms for determining the presence of cracks and corrosion are at various stages of development. The algorithms are designed to be input to the GUI as modules that can be easily exchanged as their development progresses and specialized needs arise.

In addition to the suite of LCM data that was initially used, and which represents the majority of the work presented in this report, WAMS image data was also reviewed at a preliminary level. The review included a comparison between image resolution and dynamic range for each method. WAMS (ZON file) image data was found to have a pixel pitch of 3.69μm compared to 1 μm for the LCM (vk4 file) data, which implies a lower resolution for the WAMS images. Conversely, the ratio of dynamic range of the WAMS data to the LCM data was approximately 41:20 for height data, suggesting that information from WAMS should more accurately determine the depth of pits. At this point, the significance of the reduced image resolution and greater dynamic range of the WAMS data relative to the LCM data in the detection of corrosion and cracking is not apparent.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
CAE	Convolutional Autoencoder
CNN	Convolutional Neural Network
CS	Computer Science
DNN	Deep Neural Network
DOE	Department of Energy
GUI	Graphical User Interface
ICCWR	Inner Container Closure Weld Region
LANL	Los Alamos National Laboratory
LCM	Laser Confocal Microscope
ML	Machine Learning
NN	Neural Network
RELU	Rectified Linear Unit
RGB	Red, Green, Blue Colors
SCC	Stress Corrosion Cracking
SRNL	Savannah River National Laboratory
USC	University of South Carolina, Columbia, SC
WAMS	Wide Area 3D Measurement System

1.0 Introduction

Through-wall penetration from stress corrosion cracking (SCC) of the 3013 inner container has been identified as the most credible condition for failure withing the 50-years lifetime[1]. Chlorides contained in Pu-bearing material, together with intra-canister humidity levels, metallurgical conditions, and internal stresses have been found to produce corrosion in the Inner Container Closure Weld Region (ICCWR) of the 3013 canister system, which is used throughout the DOE complex. A Laser Confocal Microscope (LCM) is used as part of the 3013 Surveillance Program protocol to identify the prevalence of corrosion and corrosion-related cracking in the ICCWR[2]. With the LCM, a close visual examination is made of the ICCWR surface along with measurements of corrosion-related features [2]. LCM inspections produce immense amounts of image data: approximately 6000 images per can, having 786,432 pixels per image, with 8 layers of data for each pixel. There is currently an 8-year backlog of images, with approximately 45 canisters that must be evaluated. Simplistic computer-aided image analysis can flag some basic surface characteristics, such as pit depth, to guide manual examinations for corrosion. However, while this approach greatly improves the efficiency of the examination process compared to unaided manual screening, it is still excessively time consuming. A more efficient and sophisticated approach is to assess the data using Machine Learning (ML) algorithms to identify corrosion without manual intervention.

The objective of this project was the development of computer-facilitated methods to facilitate rapid identification of corrosion and corrosion-related damage in 3013 canisters from very large sets of LCM (vk4 files) and (more recently) Wide Area 3D Measurement System (WAMS) data (ZON files). The LCM data includes: RGB, RGB + laser intensity, grayscale + laser intensity and height data. Height and RGB and data from WAMS files were examined at a preliminary level.

Computational methods for identifying surface corrosion and cracking include user-specified thresholds for flagging, covariance, morphological filtering and machine learning. While all of these methods were implemented at some stage of development for this project, the latter, machine learning, was a major focus area due to its potential for rapid interpretation of surface data and direct incorporation into statistical methods. In this study, the development of algorithms for machine learning and data analysis was a collaborative effort by the Savannah River National Laboratory (SRNL) and the University of South Carolina, Columbia (USC).

The process for detection of corrosion and cracking in the 3013 ICCWR consists of first extracting surface data from large binary files. This data is processed by using software that was written to convert it into physical attributes. Data taken by LCM and WAMS measurements consisted of height, color and grayscale values; all as functions of a location in a plane projection. To facilitate this operation, a user-friendly Graphical User Interface (GUI) was developed to selectively download binary data and interrogate its attributes. The GUI is a complete package that:

- Reads binary data files to enable viewing and processing of both height and optical data.
- Stitches individual LCM images into a mosaic with matched edges to give a panoramic view of the surface.
    - The image "view" can include height as well as optical data.
- Can zoom and rotate images and 3D height data for local and regional examination.
- Allows height thresholds, input by the user, to automatically flag surface features of interest, especially pits and other surface irregularities.
- Permits an expert to label features for training ML algorithms.
- Computes histograms that summarize the distribution of features.
- Can incorporate and execute ML algorithms.

  - The ML algorithms are implemented as a module so that they can readily be replaced as they are improved or exchanged with algorithms better suited to a particular data set or application.
- Can apply and visualize statistical operations to surface features.

Currently, ML algorithms are being refined and developed to identify corrosion and cracking. Figure 1 shows a flowchart for the overall development of the methodology, as applied to LCM data.

2.0 FY2020 Objectives

The FY2020 objectives for this project were:

- Develop machine learning methods, based on computer vision, to analyze data (including both optical and height data) for corrosion.
- Identify 3013 data sets, and numerical methods, suitable for near-term development.
- Determine preliminary sets of attributes for training supervised ML algorithms.
- Assemble training sets; train and test ML algorithms.
- Classify features by size, quantity, density, and location.
- Utilize computer vision to reduce amount of data needing manual analysis.

3.0 Approach

Corrosion is strongly, but not exclusively, associated with surface pitting and cracking, coloration, along with shapes and patterns of surface features, see Figure 2. Conversely, not all pits and surface lesions are the result of corrosion: some are artifacts of fabrication, impact, scoring or other non-corrosion events. Corrosion is identified via the combined properties of pit depth, area, edge contour, color and clustering. Software was developed to extract these features from large binary files generated by the LCM, and analogously will apply to images from WAMS data. The individual images, which collectively span the ICCWR were stitched together and corrected to eliminate the effect of the canister surface curvature on measurement of the local height. Various image processing methods were tested for identifying the presence of cracks and corrosion. The methods included Deep Neural Networks (DNNs), gradient methods, statistical characterization, correlations and filters[3,4]. The process is shown schematically in Figure 1.

Samples of LCM image data taken for the 3013 Surveillance Program containing identified cracks, pits and other features characteristic of corrosion were used as training data for ML algorithms. The low incidence of corrosion and cracking in the actual ICCWR samples made it necessary to incorporate data augmentation schemes for proper training of the ML algorithms developed for this application. Images containing cracks were augmented by generating vertical and horizontal translations of the original labeled image (Figure 3). Capability for labeled image rotation was also developed (Figure 4). To provide an efficient means for handling large amounts of binary image data a GUI was developed to serve as an interface with the data files, manipulate and group images, label features for training the ML algorithm, group features with user-defined thresholds, correct for sample tilt and curvature, stitch images, train ML algorithms, and apply the algorithms for crack and corrosion identification. Further, methods were developed to read binary WAMS data, which has recently been adopted for 3013 image interrogation. After preprocessing, images obtained from LCM and WAMS data were partitioned into tiles (rectangular blocks of pixels). Image data used for training and testing ML algorithms was labeled on a per-tile basis. Studies conducted during FY 20 emphasized that larger views, represented by image tiles containing a larger number of pixels, improve the accuracy of crack detection by the ML algorithms.

4.0 Results/Discussion

Cracks, pits and color patterns are all associated in various forms with corrosion. Pits can readily be detected using height data thresholds. However, cracks, particularly "hairline" cracks, do not always have a definitive height signature. Rather, crack identification is a combination of grayscale image intensity (pixel value) and height data. Initially, it was hoped that standard edge detection methods could be used with pixel values to extract crack edges. Methods considered included: erosion and dilation, blurring, Fourier and Gaussian filters, and gradient methods. Unfortunately, other surface features combined to create background noise that was similar in frequency to that associated with crack edges. To overcome this problem, crack identification was attempted through the use of DNNs.

Generally, the training of the DNN algorithms suffered due to the small amount of crack data available and the severe imbalance between the amount of crack data and the amount of non-crack surface. This was particularly true in the early stages of algorithm development. The lack of images containing cracks was offset somewhat by using augmented image data to synthesize additional labeled crack data.

Studies conducted during FY 20 indicated that larger views, represented by image tiles containing a larger number of pixels, improved the accuracy of crack detection by the ML algorithms. Increasing the amount of labeled training data through augmentation, adjusting the DNN algorithms, and increasing the image tile size from 64x64 to 112x112 pixels improved the precision and recall for crack identification. Examples of labeled training data for cracks, taken from LCM images are shown in Figure 5. It was found that the greatest accuracy was obtained by using a consensus drawn from an ensemble of randomly generated Convolutional Neural Network (CNN) algorithms, having the following characteristics:

- 2 channel input – grayscale contrast-adjusted peak intensity and tilt/curvature-corrected height.
- A CNN depth of 1 to 5 layers, with each layer consisting of a convolution layer using a Rectified Linear Unit (RELU) activation function and a maximum value pooling layer.
    - The convolution layers had kernel sizes ranging from 1x1 to 3x3, with 1 to 128 feature maps.
    - The pooling kernel size ranged from 1x1 to 3x3, with strides that ranged from 1x1 to 3x3.
- The final layer of the CNN consisted of a fully connected layer of 1 to 128 neurons.

The consensus was a vote by all models in the ensemble on whether or not an image contained a crack. Results from this classification method are shown in Figure 6, which compares consensus accuracy with that of a single model and shows that the ensemble reached an accuracy greater than that of any individual algorithm within it. It was found that that accuracy was improved when training sets contained an approximately an equal number of tiles with and without cracks, achieved via data augmentation of the crack tiles.

5.0 FY2020 Accomplishments

The FY2020 accomplishments were the result of a collaborative effort by USC and SRNL.

- The functions developed for image analysis are now available through user-friendly GUIs developed specifically for this application. Functions executed by the GUI include:
    - Data input.
    - Data labeling, flagging of features of interest, and other diagnostics. Includes zoom capability.
    - 3D surface display of height data.
    - Implementation of ML algorithms.
        - ML training methods can be implemented through the GUI.
        - Trained ML algorithms can be applied using the GUI.

        - ML algorithms are not yet in final form, but GUI access is modular so that updated algorithms can readily be imported, replacing the current ones.
- Different forms of ML algorithms were tested.
    - The algorithms included well regarded CNN's: such as ResNet 50, AlexNet, and custom CNN's, which gave the best performance so far.
    - It was determined that overfitting (excessive response when applied to test data) and recall (a measure of the rate of false negative predictions) need to be improved in CNN algorithm development.
    - It was found that deeper neural networks don't necessarily result in better performance.
    - Hyperparameter optimization was performed, including genetic algorithm methods.
        - Due to the number of hyperparameters, internal gradients (a custom method), may be used for optimization.
- The fine scale detail associated with cracks (in terms of grayscale and color intensity, along with height variation) was found to complicate the process of distinguishing them from scratches, tooling marks, and sequences of pits/protrusions on the surface.
    - The horizontal machining marks in particular are visually similar to linear cracks, and all of the currently identified cracks overlap the machined region, posing challenges for the algorithm.
- Developed Convolutional Autoencoder (CAE) to obtain transfer learning for CNN's used for image analysis.
    - The CAE supports transfer learning for accelerating the training of the CNN's, to help compensate for a relatively small amount of training data.
- Developed data labeling and augmentation capability. Labeled data is necessary for training the CNN's and a means for rapid labeling is necessary due to the volume of data required to train the algorithms.
    - Developed software to enable rapid labeling of corrosion training data.
        - Completed software for mouse click labeling of corrosion training data for supervised learning (Figure 7). This method associated pixel tagging with tile labeling.
        - Labeled pixels are retained during geometric transformation for data augmentation (Figure 4).
        - Automated pit flagging functionality and k-means clustering on optical data facilitates rapid labeling of multi-tile regions (which can be fine-tuned during final labeling).
    - Developed baseline software for data augmentation via simulated geometric variations (Figures 3, 4, 7).
- Developed software that enables reading of WAMS binary data files.
    - USC has provided the software to the Los Alamos National Laboratory (LANL) statistics group.
    - SRNL will compare surface contour data from WAMS with that from the LCM files.
        - Tests will include proportionality between data sets for areas and heights of pits and protrusions.
- SRNL is collaborating with the LANL statistics directorate on image analysis methods, possibly coupling ML based corrosion/defect identification with statistical evaluation.

6.0 Conclusions

A user-friendly Matlab GUI that reads data from either LCM or WAMS files was developed. The GUI can label features for reference, further examination, archival storage, or for development of a training set for machine learning. Features can be called out by user-specified thresholds, manual labeling or machine learning algorithms when they have been completed. The ability to rapidly label data is important because of the volume of data required for training machine learning algorithms.

The GUI has the flexibility to allow addition of improved ML algorithms, methods for data visualization, and statistical computations. Available data statistics include areas of pits within a defined range of pit depths, correlations between RGB or grayscale intensity and relative surface height, covariances between values associated with features, and feature histograms.

Overall, the development of supervised learning algorithms is hindered by a lack of training data. The machine learning algorithms for crack identification are in a state of partial development and require improvements to the true positive rate. This shortcoming is an artifact of the limited training data currently used, perhaps more so than the structure of the neural networks. The best results are had from an ensemble consensus of CNN models that differ in their hyperparameters. The training accuracy of the ensemble consensus attained true positive and true negative rates of 100% for ensembles containing more than 10 networks. Upon testing, the consensus accuracy leveled off for ensembles consisting of 25-35 networks with a true negative rate of approximately 88%, but a true positive rate of only around 14%.

Data used to train the neural networks used for the machine learning algorithms consisted of unaltered LCM images augmented by translated images as shown in Figure 3. The capability for augmentation via rotated images has also been developed. Although not used for training the algorithms discussed in this report, image rotation will be used for future efforts.

WAMS (ZON file) image data was found to have a pixel pitch of 3.69µm compared to 1 µm for the LCM (vk4 file) data. The significance of the lower resolution for the WAMS data with regards to corrosion and crack identification is not clear at this point and will require further evaluation. Further, the ratio of dynamic range of the height data from WAMS to that from the LCM was approximately 41:20. This implies that the WAMS data should more accurately determine height than the LCM data. Again, the significance of the greater dynamic range of the WAMS data relative to the LCM data has not yet been evaluated.

7.0 Recommendations for Future Work

- Improvement to ML for crack detection.
  - Include hyperparameter optimization and other forms of classification algorithms, e.g. decision trees, support vector machines, etc.
  - Further development of AI for corrosion detection based on surface image and height data.
- Testing of ML training based on multiclass labeling.
  - Separately labeling non-crack features that are visually similar to cracks, such as machining marks, may help to drive the learning process to learn distinguishing features between them and thus improve crack identification over binary labeling.
  - Multiclass training/classification allows for detection of separate classes of corrosion features, e.g., cracks versus pitting.
- Processing of WAMS image data using the GUI interfaces. ML training using WAMS data (which requires sufficient WAMS files containing identified cracks/corrosion).
- Provide a more complete assessment of the impact of reduced image resolution for WAMS images on ML capability for identification of cracks and corrosion.
- Evaluate parity between LCM and WAMS measurements for areas associated with depressions, and for pit areal density (individual pits per area).
- Interface ML crack and corrosion assessments with data statistics.
  - Collaborate with LANL statistics group.
  - Incorporate advanced statistical methods developed by LANL into the GUI.
- Acquisition of additional training data.
  - Augmentation via geometric manipulation.
  - Using data from corrosion coupons.

    - Make use of baseline canister images as a large collection of relatively easily labeled data, i.e. images demonstrating canister features in the absence of corrosion. In conjunction with multiclass labeling, this data could help to refine classification of non-corroded surfaces. (Note: Training on baseline canister images solely as "non-crack"/"non-corrosion" data in binary-class training would further imbalance the training set and thus be detrimental to crack identification.)
- Investigate potential for using AI methods to relate surface features to diagnostic data for subsurface corrosion, voiding and deterioration.
    - Investigate whether there are surface features associated with subsurface voids and cracks observed in tomography images.

8.0 References

1 Berg, J.M., D.K. Veirs, E.J. Kelly, Juan G. Duque, S.A. Joyce, J.E. Narlesky, J.M. Duffey, J.I. Mickalonis, and K.A. Dunn, Test Plan for Assessing Potential for Stress Corrosion Cracking in the 3013 Inner Container Closure Weld Region (FY 2014). 2014, LA-UR-14-20785.

2 Martínez-Rodríguez, M.J., Laser Confocal Microscope for Analysis of 3013 Inner Container Closure Weld Region. 2017, SRNL-STI-2017-00589.

3 Ian Goodfellow, Yoshua Bengio, and Aaron Courville. Deep Learning. MIT Press, Cambridge MA (2016)

4 Aurélien Géron. Hands-On Machine Learning with Scikit-Learn &Tensorflow. O'Reilly, Sebastapol, CA (10/12/2018).

9.0 Figures

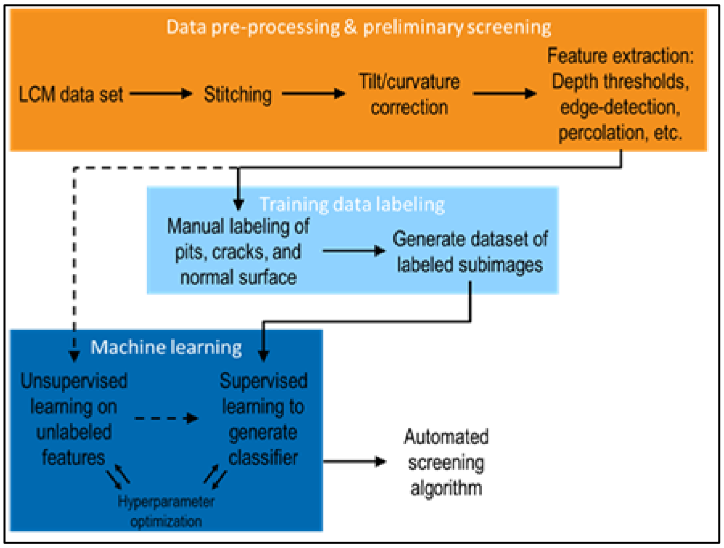

Figure 1. Flowchart for data processing and application of machine learning and other analysis methods for identification of corrosion and cracking in LCM and WAMS data.

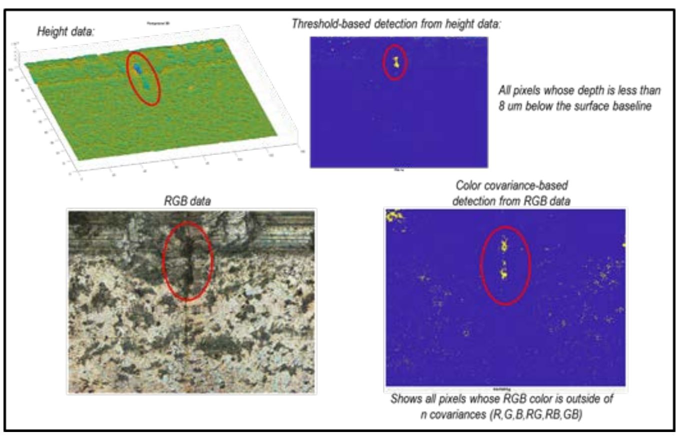

Figure 2. Crack and corrosion data from LCM images. Data channels include RGB, grayscale and height.

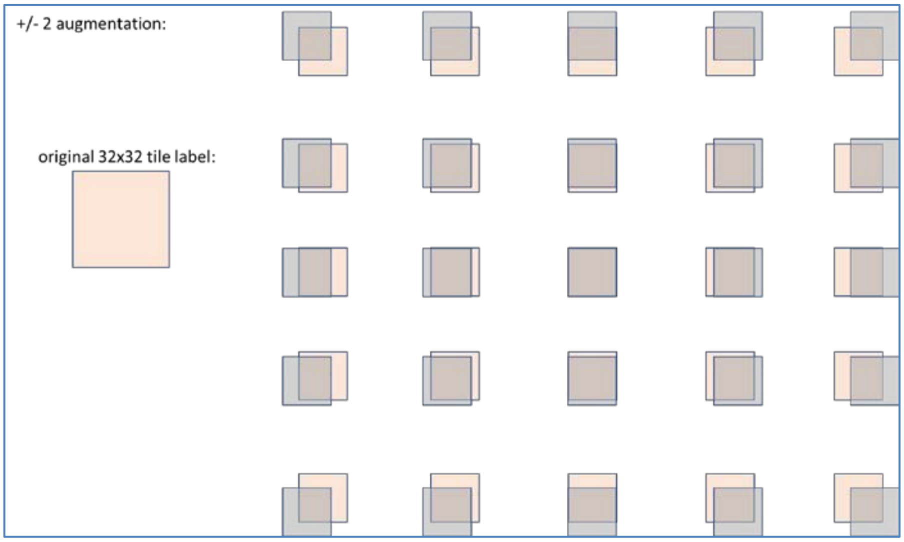

Figure 3. Data augmentation by translation of image. Translated offsets are shown in gray.

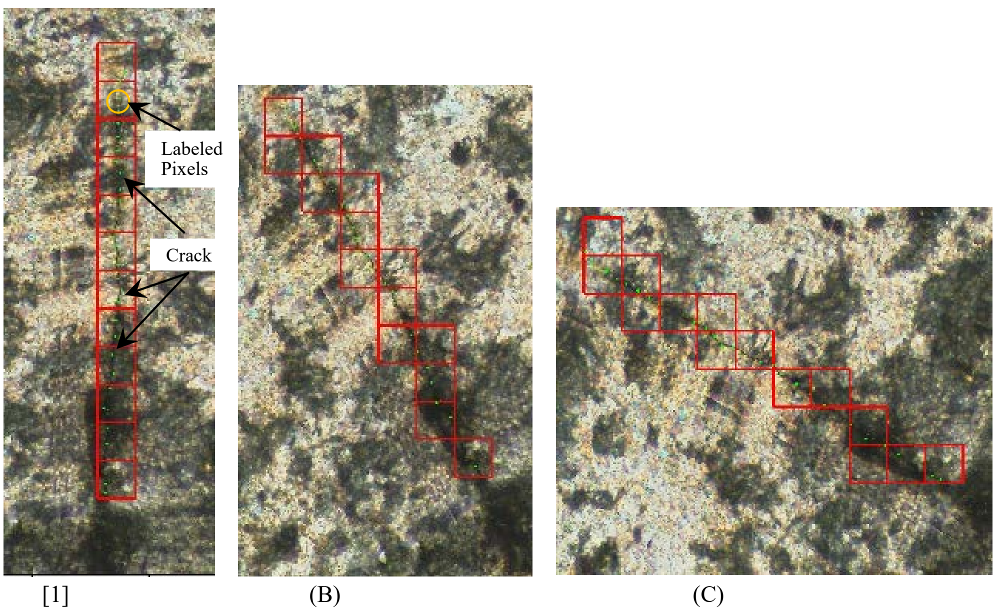

Figure 4. Data augmentation by rotation of crack image. The images are cropped show the highlighted region at a larger scale. Green dots (barely visible) along crack are pixels manually selected by the analyst by mouse click. The baseline orientation is shown in (A). When the image is divided into tiles, those containing a crack pixel are highlighted with a red boundary and labeled as a crack containing tile. Tiles that do not contain a crack pixel are defined and labeled as not containing a crack and are not highlighted. In (B) the image and crack pixels are rotated 30 degrees in the counterclockwise direction and the image is divided into tiles. Again, those tiles containing a crack pixel are highlighted with a red boundary and labeled as a crack containing tile. Tiles that do not contain crack pixels are labeled as such, and not highlighted. Similarly, (C) shows a rotation 60 degrees in the counterclockwise direction. The tile labeling based on the marked pixels updates automatically in each new orientation, resulting in a multiple distinct sets of crack tiles from the same baseline image.

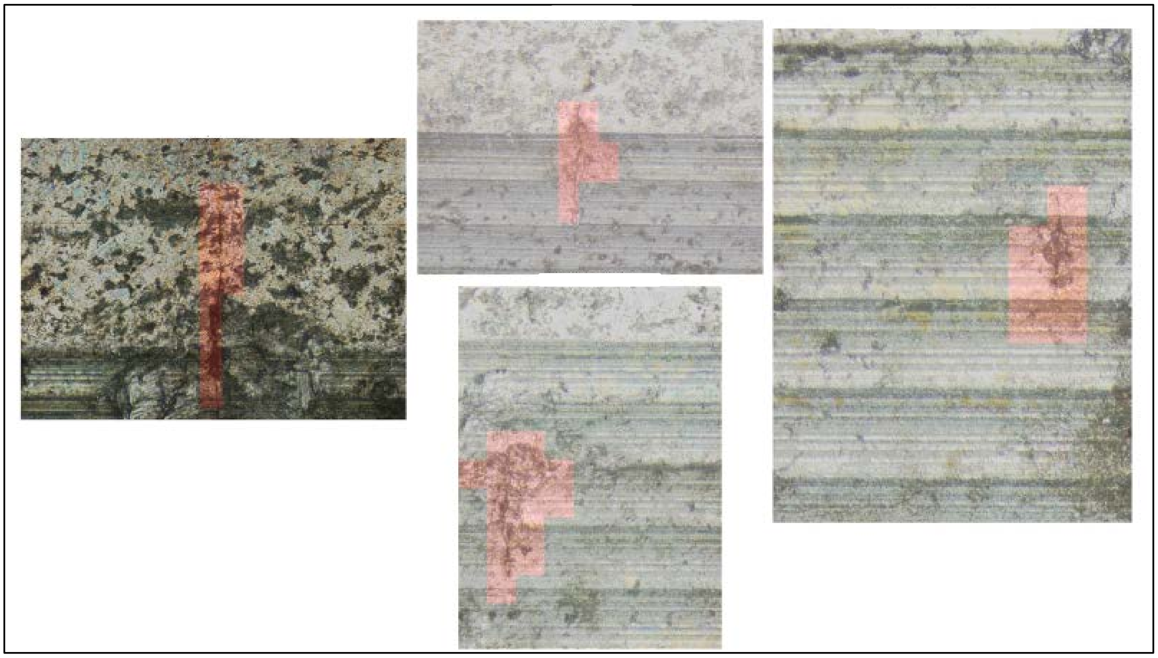

Figure 5. Examples of labeled data consisting of 112x112 pixel tiles (blocks). The training process used bootstrapped data samples that were augmented using horizontal and vertical translation. Highlighted regions are tiles that contain cracks.

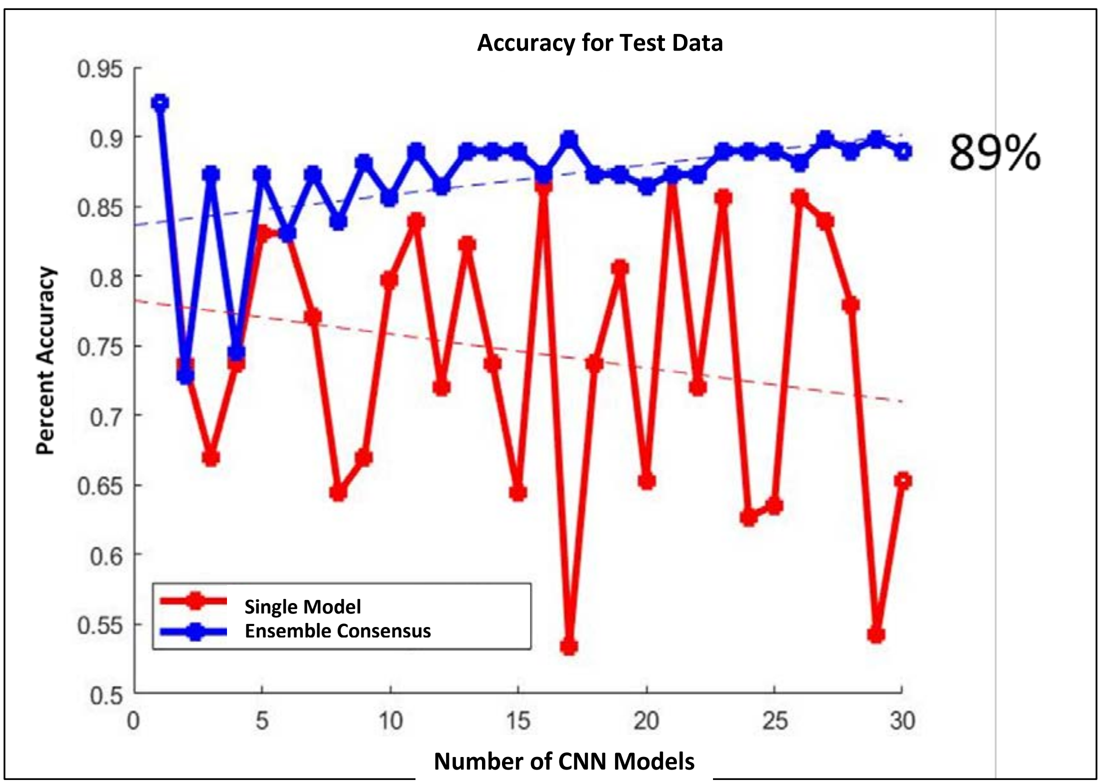

Figure 6. Comparison of the accuracy of crack identification from an ensemble consensus of CNN algorithms with the accuracy of a single model.

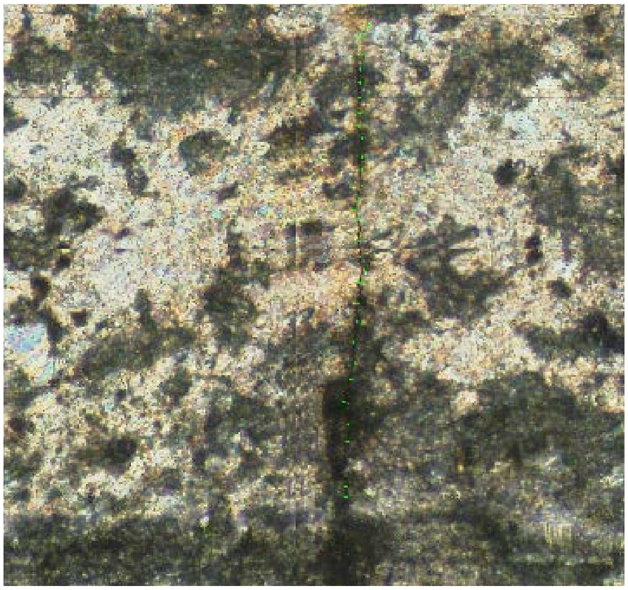

Figure 7. Pixels selected by mouse click are highlighted by green dots. The pixels are used to label the crack in rotational transformations that augment available data.

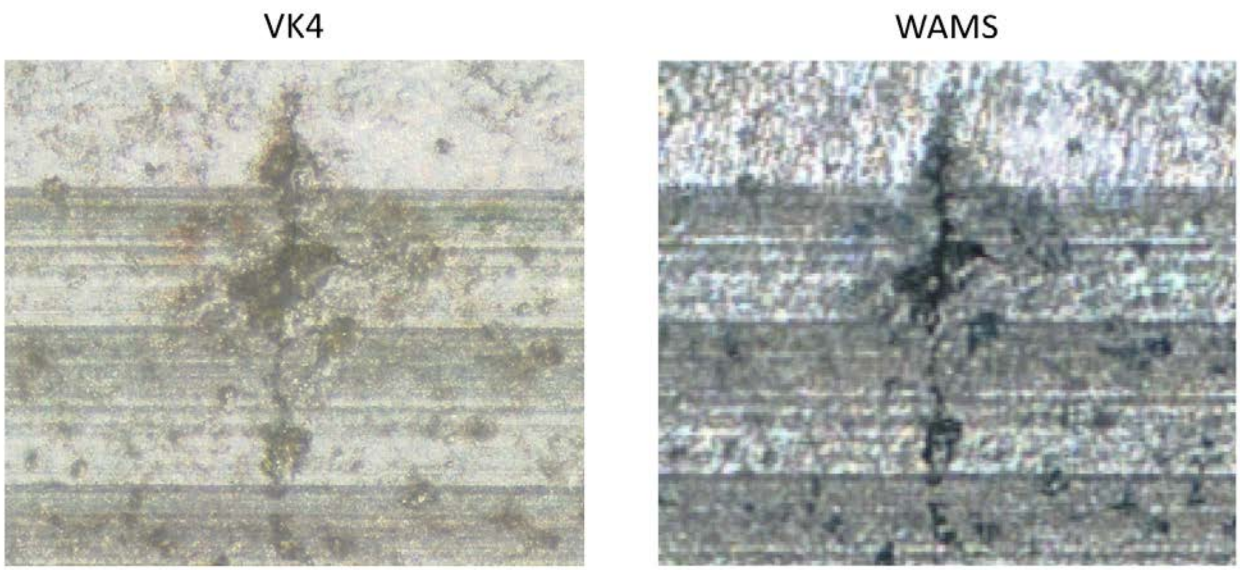

Figure 8. Close up of RGB data for the “Acrux” training feature from VK4 (left) and WAMS (right).