Enhanced X-Ray Inspection of Solder Joints in SMT Electronics Production using Convolutional Neural Networks

Conference Paper · October 2020

DOI: 10.1109/SIITME50350.2020.9292292

8 authors, including:

Konstantin Schmidt
Friedrich-Alexander-University of Erlangen-Nürnberg
11 PUBLICATIONS **30** CITATIONS
SEE PROFILE

Nils Thielen
Friedrich-Alexander-University of Erlangen-Nürnberg
13 PUBLICATIONS **28** CITATIONS
SEE PROFILE

Christian Voigt
Friedrich-Alexander-University of Erlangen-Nürnberg
9 PUBLICATIONS **15** CITATIONS
SEE PROFILE

Reinhardt Seidel
Friedrich-Alexander-University of Erlangen-Nürnberg
14 PUBLICATIONS **61** CITATIONS
SEE PROFILE

Enhanced X-Ray Inspection of Solder Joints in SMT Electronics Production using Convolutional Neural Networks

Konstantin Schmidt, Nils Thielen, Christian Voigt, Reinhardt Seidel and Jörg Franke
Institute for Factory Automation and Production Systems
Friedrich-Alexander University Erlangen-Nürnberg
Nürnberg, Germany
Konstantin.Schmidt@faps.fau.de

Yannik Milde, Jochen Bönig and Gunter Beitinger
Segment of Factory Automation, Manufacturing
Siemens AG
Amberg, Germany

***Abstract*—The electronics production is prone to a multitude of possible failures along the production process. Therefore, the manufacturing process of surface-mounted electronics devices (SMD) includes visual quality inspection processes for defect detection. The detection of certain error patterns like solder voids and head in pillow defects require radioscopic inspection. These high-end inspection machines, like the X-ray inspection, rely on static checking routines, programmed manually by the expert user of the machine, to verify the quality. The utilization of the implicit knowledge of domain expert(s), based on soldering guidelines, allows the evaluation of the quality. The distinctive dependence on the individual qualification significantly influences false call rates of the inbuilt computer vision routines. In this contribution, we present a novel framework for the automatic solder joint classification based on Convolutional Neural Networks (CNN), flexibly reclassifying insufficient X-ray inspection results. We utilize existing deep learning network architectures for a region of interest detection on 2D grayscale images. The comparison with product-related meta-data ensures the presence of relevant areas and results in a subsequent classification based on a CNN. Subsequent data augmentation ensures sufficient input features. The results indicate a significant reduction of the false call rate compared to commercial X-ray machines, combined with reduced product-related optimization iterations.**

Keywords—electronics production; automated X-ray inspection; machine learning; computer vision; Convolutional Neural Networks

I. Introduction

The process performance in industrial production is commonly specified in occurring defects per million opportunities (DPMO). Trough progressing automation and enhanced process capability, electronics manufacturers nearly achieve single-digit DPMO-rates within their production processes. [1] Insufficient solder connections in the electronics production can result from a multitude of influencing factors so that a direct attribution to a single influencing factor is improbable. [2] The validation of the quality levels is realized with an extensive inspection coverage. Besides electrical test methods, optical inspection processes characterize the outer appearance of the products defined regions of interest (ROI). The increased application of integrated circuits (ICs) with covered ROIs like ball grid arrays (BGAs) justify the usage of X-ray based methods in addition to inspections using light in the visible spectrum.

The automated X-ray inspection (AXI) typically represents the last optical inspection step in the SMD production. Accurate classification is essential and leads to narrow checking routines in the qualification process. As a direct consequence of the static decision rules, two main problems arise. For the setup of these rule-based test routines, extensive product- and process-specific know-how are required. The high overhead of creating and maintaining product-specific checking routines, leads to increased costs, and severe dependency on expert knowledge, impeding the continuous improvement of existing test procedures. Static checking routines lead to reduced sensitivity with varying conditions, as misclassification rises.

The aforementioned issues justify the investigation of a deep learning approach. A flexible inspection algorithm is expected to increase the classification accuracy, reduce optimization costs, and lower dependency on expert knowledge. A system based on object detection is capable of autonomously detecting relevant ROIs for new inspection tasks derived from historic data. For this study, the solder joints of an electrical connector used in programmable logic controllers are evaluated. A flexible computer vision system is developed from the automated classification of the extracted grayscale images.

II. State of the Art

A. Industrial X-Ray Inspection of Solder Joints

The quality assessment of a covered solder joint is carried out by the utilization of X-ray inspection based on grayscale images. The intensity of certain areas of the image varies depending on the material being irradiated. This results in darkened image areas for metallic material accumulations. Automated evaluation of predefined image areas is carried out based on these differences in intensity. If a certain amount of pixels within this area exceeds a set threshold, the area is marked as insufficient. A confirmation from the operator must be carried out. These product-specific test routines are manually created and adjusted.

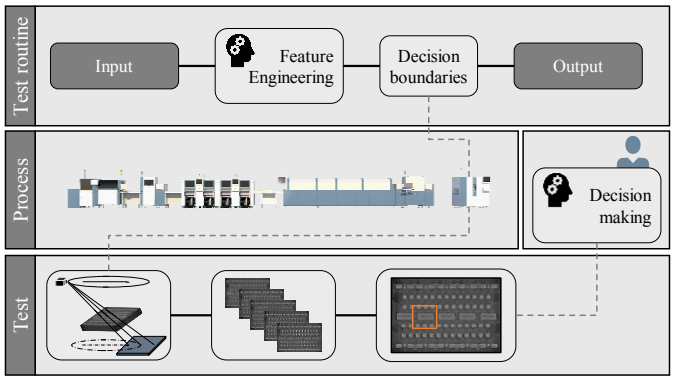

Fig. 1: Schematic overview of the human influence on the test result.

B. Object Detection based on Deep Learning

Machine Learning based Computer Vision applications have gained the highest interest with increasing computing power. Besides the image recognition, the object detection and instance segmentation are in general exercised for knowledge extraction from images. These applications are based on deep learning techniques. They use multiple processing layers to learn representations from the abstracted data. [3] The feature engineering is no longer a manual process, executed by human engineers, but learned in the abstraction layers of the network. With a sufficiently large combination of layers, functions of any complexity can be approximated, using general learning techniques.

One of the most frequently applied types of learning is supervised learning. This is an attempt to learn the relationship between an input vector and an assigned output. It requires a sufficient quantity of labeled inputs to create significant relations. An error function computes the difference between the given and the predicted outcome to indicate the score during the training. With continuous updating of function parameters (weights), the result is being optimized towards a minimal error. This adjustment is based on optimization methods. Stochastic gradient descent is a commonly exploited method, due to its ability to properly generalize on unknown data. It minimizes the loss by computing the gradient of the loss curve for training examples and updating function parameters accordingly. [4]

The general architecture of Neural Networks (NNs) consists of an input layer, directing the information to the output over a generalized linear function (activation function). Intermediate layers, hidden between input and output are applied for the non-linear abstraction of the forwarded information. For updating the weights according to the gradient of the error function, the information has to be propagated back into the network. With the backpropagation algorithm, the loss of every node can be computed and the weights can be optimized subsequently. [5]

A Convolutional Neural Network (CNN) is a specific type of feed-forward network commonly used for grid-structured data (arrays). So that this type of network can be used for image classification or object detection, a mathematical operation called convolution is used instead of the conventional matrix-vector multiplication. The convolution is a cross-correlation between the signal and a given filter (feature map) to detect structural similarities between both functions. [3] With the extension to the two-dimensional space, image data can be covered in height and width. For improved generalization and reduced computational costs, pooling layers are applied. The array is divided into a grid of a set resolution and only certain features are transferred. [6] The conclusive fully connected layer with multiple 1D layers is the successor of the output layer with the predicted classes. [7] Two main reasons, why CNNs show superior performance on image data, are a high local correlation of associated values and the local invariance of the values, both being considered with the specific architecture. [3]

C. Industrial applications of Machine Learning

The growing number of industrial applications strengthen the suitability and even the superiority over rule-based approaches in versatile production processes. Among others, Cia et al. [8] present a sophisticated approach for solder joint inspection using a cascaded CNN. For the region proposal of the ROIs, a sliding window approach was conducted. For the proposed regions, CNN-based quality predictions are executed, surpassing conventional SMT solder joint inspection. [9, 10] Metzner et al. [11] use a sophisticated approach to challenge the false call rate of commercial Automated Optical Inspection (AOI) systems via a Neural Network classifier. Within a supervised learning approach, static ROIs are defined for the relevant inspection areas, by manually labeling the associated areas of the images. In a benchmark, the precision of the proposed model surpassed commercial AOI systems.

D. Need for Action in Research

State of the art X-ray inspection systems demand extensive domain expert knowledge for the setup and optimization of testing routines. The test programs suffer from insufficient classification accuracy, particularly under varying conditions. Subsequent false call reclassification demands high personnel retention. Current research efforts have mainly focused on training Neural Networks with predefined regions of interest.

This contribution is to research if deep learning-based algorithms improve the generation of testing routines regarding flexibility and classification accuracy of SMD solder joints. Grayscale images of the solder joints of an electrical connector used in programmable logic controllers are evaluated. The classification is trained and tested on multiple types of solder joints and error types. In a subsequent approach, an object detection is trained for the autonomous detection of solder joints with no predefined regions of interest. It is evaluated if the developed concept applies to the existing manufacturing infrastructure, regarding prediction accuracy and computation time. The used training images are generated as a by-product in a real production environment.

III. System Design

For the realization of the tests, the concept shown in Fig. 2 is proposed. After the inspection process of the X-ray machine, the result images are fed into the inspection algorithm, running on an industrial computer. After the detection, the result is transferred to the test station, where it can be further processed, depending on the outcome. The setup allows a direct inspection of components marked as defective by the test routine. Only marked components are evaluated and considered in the learning

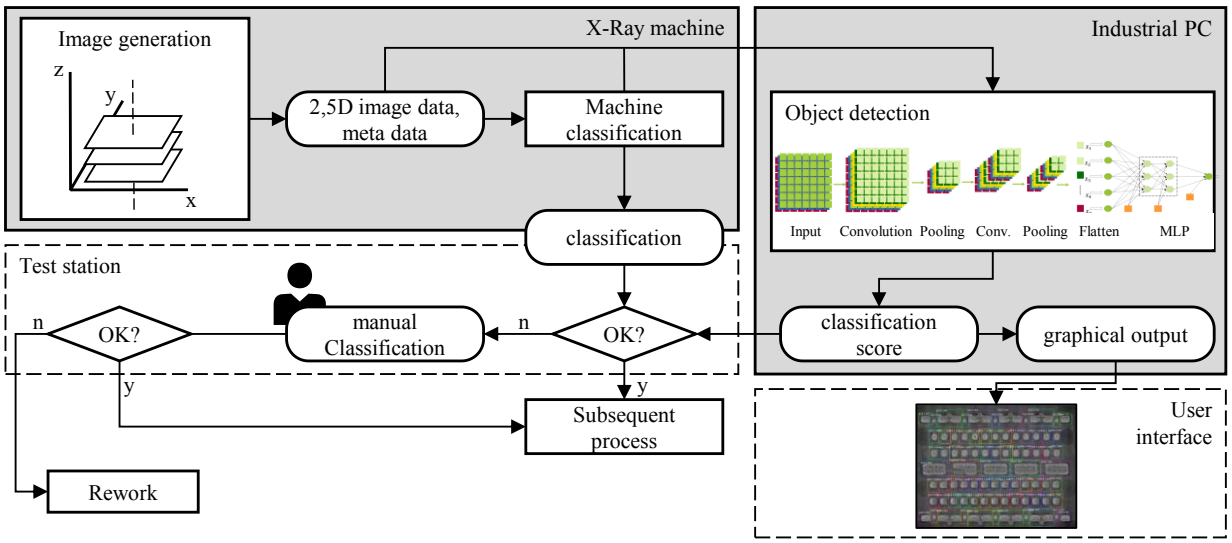

Fig. 2: Data flow in the proposed, CNN-based system design

process so that the algorithm has to distinguish between false calls and true failures.

A. Data acquisition and preparation

On the investigated PCB, two functionally different types of solder connections are defined. These connections are subdivided into six categories for the detection, according to their geometry. Each side of the board contains 79 or 52 solder joints, depending on the current component.

For the X-ray inspection, an Omron VT-750 X-ray machine is conducted. The imaging system is based on a slice imaging method using transversal computer tomography (CT). The assembled PCB is positioned between the X-ray tube and the detector, which are horizontally circulating the PCB. [12] The generated image data consists of 65 8-bit grayscale images vertically varying over the inspected product. One image layer consists of 623x483 pixels for the inspected component.

The image data is extracted as *.rec-files, containing the multilayer pixel information. In combination with the provided meta-data file, containing information about layers and resolution, the information can be transferred into an array structure of the form 623x483x65. In the second step of the data preparation, the array is reduced to one layer, selected as sufficient by the inspection system after the quality assessment. It is ensured that the same database is used for the training as seen and evaluated by the operator. The feature space is reduced to a two-dimensional vector.

Subsequently, the positioning of the solder connections on the board is defined. A self-developed graphical user interface (GUI) was conducted for the position-labeling and type categorization of the solder connections. This step is performed each time a new connector is trained.

For the labeling of the individual solder connections, the database of the in-plant quality system is utilized: The true label, assigned by the operator in the confirmation process, is directed to the system, overruling the X-ray inspection result if required. The pin-based evaluation result is merged with the 2D image array by a unique identifier. For the training 6387 images are available.

To counteract potential overfitting, data augmentation is executed, to synthetically create a variance within the image data. This includes rotation of the image (90° - 270°), relocation of the ROIs, and insertion of interfering elements. For the training of the image classification, the data is split in the proportions 70 % (training), 20 % (validation), and 10 % (testing). Due to the high yield (> 99.99%), a class imbalance is created. To prevent a model, biased towards one class, class weights are introduced. According to the proportion of a class in the dataset, a weight is added to relativize the effect of this class during training. The minority classes are rated > 1, whereas the majority class is rated < 1. The application is written and executed in Python (v3.7). The used TensorFlow-GPU (v1.15) backend is addressed via Keras (v2.0). [13, 14]

B. Test Setup

In Table I. the selected test setup investigated in this work is shown. A distinction is made between the detection method and the image section used (cut out, full image). Consequently, the required classes and the available data for the training are adapted. The first training and testing (A) are done on cut out areas of the images. The defined ROIs are derived from the labeling process. Further tests (B-D) were carried out on the complete images. The second dataset (B) includes two classes, functional pins, and open solder connections. For the third dataset (C) functional pins and solder bridges are trained. In the last image classification (D) the tests are merged to distinguish functional connections from general defects.

In the training of the object detection, the different pin types occurring on the component are classified. Besides, the error types ‘open solder’ and ‘solder bridge’ are included for the detection.

TABLE I. EVALUATED TEST SETUPS

Test Set	Test Setups			
	Detection method	*Image type*	*Classes*	*Class distribution*
A	Image classification	Cut out solder connections	2	5560/827
B	Image classification	Complete image	2	562/728
C	Image classification	Complete image	2	562/326
D	Image classification	Complete image	2	562/1054
--	Object detection	Complete image	8	526/1054

C. Image Classification

As a suitable CNN architecture for the image classification, the VGG-16 is chosen. [7] Compared to AlexNet, the kernel size is reduced to 3x3 over multiple layers. With the increased number of non-linear layers, more complex structures can be learned. [15] The pre-trained convolutional block of the VGG-16 is utilized. As described earlier, the local correlation of associated values within images and the local invariance of the values allow the training of a generic feature extractor. It is suitable for edge and structure detection. [3, 15] After the convolutional block, the compressed data is then transformed into a one-dimensional vector (flatten layer) and transferred to a multi-layer perceptron (MLP) for classification. In three fully connected layers, the output of the convolutional part is converted into a 1x1x4096 vector to be interpreted by the last dense layer for the classification. Two output neurons are defined for the final classification.

In the input layer, the images are resized to a vector of the size 224x224x3. Since X-ray inspection generates grayscale images, which corresponds to a width of one, two channels of the images used are left empty. The MLP is replaced with a two-class soft matrix, which means that the predefined weights are lost. Binary cross-entropy is chosen as the cost function, as a two-class categorization is required. Weight adjustment is carried out by the SGD optimizer, according to the loss calculated by backpropagation. For the training a learning rate of $\eta = 0.0001$ is chosen. The learning rate regulates the convergence of the weights over the epochs. The model is trained over 100 epochs.

For Training A an accuracy of 98 % could be achieved after 4 epochs on the validation set (Fig. 3). The similar learning curve for both sets indicates minimal overfitting in the process. During Training B similar indicators were observed (accuracy = 98 %). In Training C, the increase in accuracy of the training and validation set is significantly reduced in comparison to Training A and B. After 60 epochs the accuracy exceeds 90 %. This is considered to be a result of the higher variance in the occurrence of errors. Training D showed improved performance compared to Training C (40 epochs > 90 % accuracy), indicating an enhanced learning effect due to the higher volume of training data.

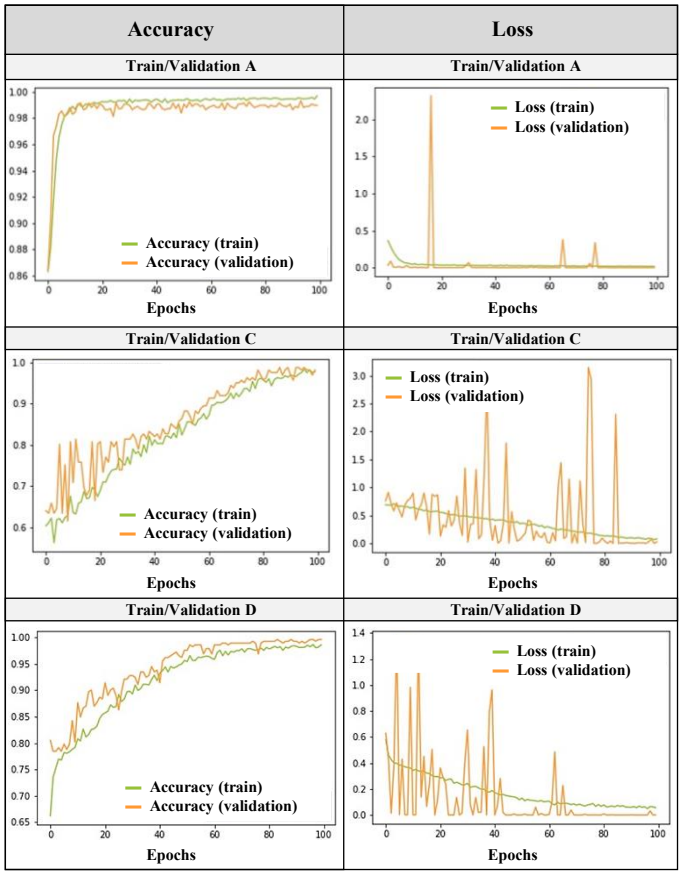

Fig. 3: Accuracy and loss for the training and the validation of the image classification (A, C, D).

D. Object Detection

As a modification to image classification, an object recognition, taking the location under consideration is evaluated. A Mask-RCNN architecture is conducted. [16] With this architecture, the possibility of an instance segmentation is given but is not considered in this work. [17] The according layers are removed. A ResNet101 architecture is used as a backbone for the feature extraction. [18]

As eight different objects are targeted with the detection, categorical cross-entropy is chosen as the cost function. The SGD optimizer is regulated by a learning rate of $\eta = 0.0001$. Due to the high computing complexity, the batch size is reduced to one, and the epochs for the training set to 5. The previous tests with the VGG-16 architecture have shown, that after only a few epochs an accuracy of over 90 % could be achieved. Consequently, the reduced number of epochs is considered sufficient. For the evaluation, the mean average precision (mAP) is used. The loss curves of the training and validation data fall from 1 to 0.5 after the first epoch and converge towards zero with a negative gradient in the following epochs, as shown in Fig. 4.

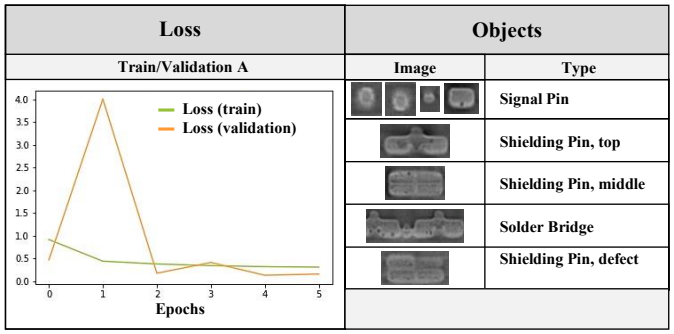

Fig. 4: Loss for the training and the validation of the object detection and the according object classes.

E. Evaluation of the Results

For the evaluation of the trained model, the confusion matrices are shown in Fig. 5. Misclassification by the model occurred in five of 647 cases for Test A (cut out pin images). The misclassification of true failures commonly represents the most cost-intensive decision in production. With the adaption of the discrimination threshold, the False Positive Rate (FPR) can be reduced to zero, resulting in a decreased True Positive Rate (TPR) of 0.95, as shown in the receiver operating characteristic (ROC) curve (see Fig. 6). The classification of Test B to D show a similar proportion but result in a stronger decay of the TPR for FPR of zero (Test C).

The object detection models each achieve an average mAP score of 0.99, with an intersection over union (IoU) threshold of 0.5. For the classification of the functional solder connections, the confidence score varies between 0.98 – 1.00 for functional solder connections and 0.96 – 0.99 for bad solder connections. The slight difference in favor of the functional pins could be attributed to the increased variance of errors (open solder vs. solder bridge). To ensure that an unexecuted classification is not resulting in an unclassified solder connection, a product-related quantity of objects must be identified for a positive test result.

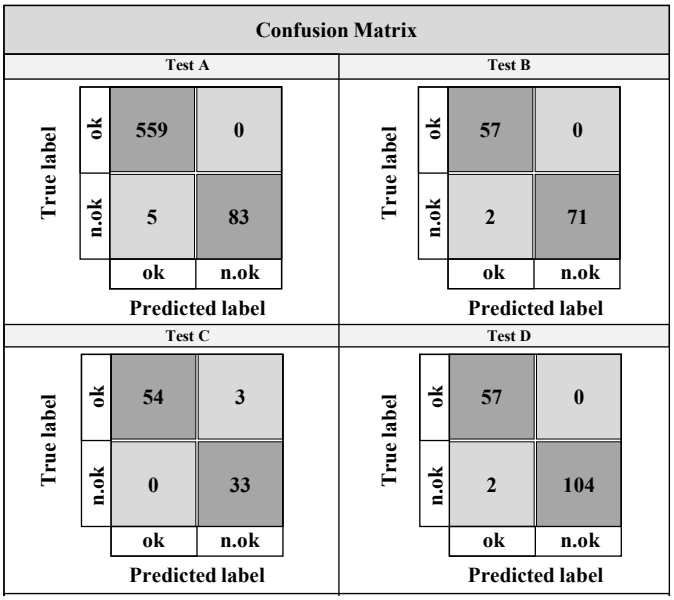

Fig. 5: Confusion matrices of the testing of image classification (Test A-D)

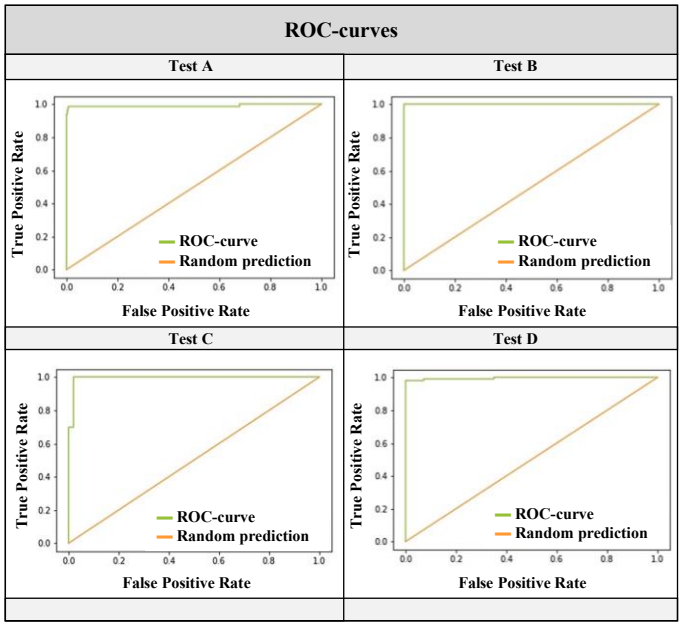

Fig. 6: ROC curves for the testing of the image classification (Test A-D)

F. Transferability of the model

The evaluation of various images showed promising results, regarding the transferability to further components. The model was able to identify objects as solder connections of components, which have not been included in the training set. The fully connected layer of the model can learn a generalized concept of a solder connection, as shown in Fig. 7. On 50 test images of an unknown component, the model reached an average mAP score of 0.98. The application of a generalized model, based on a similar training, evaluating multiple products, could enable a more time-efficient and flexible generation of test programs for X-ray detection.

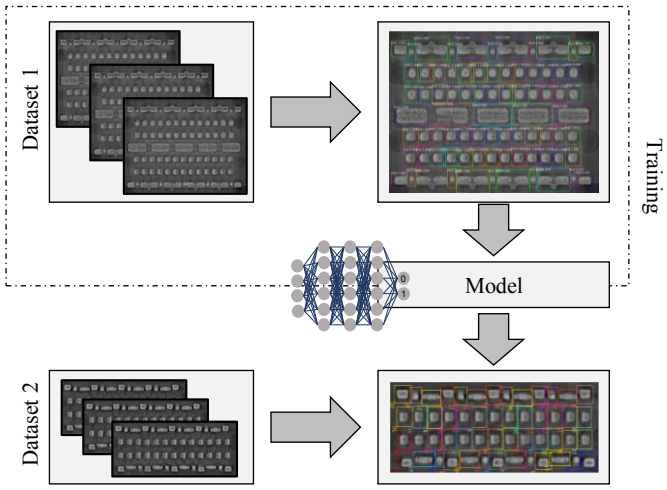

Fig. 7: Concept for the transferability of the evaluated model.

IV. CONCLUSIONS AND FUTURE RESEARCH

The application of different CNN architectures, for the detection and classification of different solder connections and error types in grayscale X-ray images, could successfully be evaluated. The proposed system shows superior precision compared to the commercial X-ray inspection test routines. The dependence on domain-specific know-how for the generation and optimization of the test-routines could be reduced with the CNN-based classification approach. The application of an object detection algorithm shows promising results, regarding the autonomous generation of test routines for the X-ray inspection. The model was able to perform cross-component detections of solder connections. For a reliable integration of the system, improved classification accuracy is still demanded. The sufficiently large dataset required for this purpose is one of the main challenges to be addressed. Further research activities will investigate the application of the proposed supervised techniques in combination with unsupervised techniques, to focus on the recognition of unknown failure types and eliminate the prerequisite of a comprehensive dataset.

Also, there is still a dependency on the selection of the appropriate focus level of the image files made by the machine, which requires further development of the proposed solution. A promising approach could be the integration of a multilayer feature extractor. Moreover, the image data could be extended by a further dimension, increasing the utilized information, as 3D information is provided by the image data generated by the X-ray inspection.

ACKNOWLEDGMENT

The provision of the measurement, process, and quality data used, as well as the IT infrastructure, was made possible by Siemens AG's electronics plant Amberg. The results were obtained within the scope of research activities at the Institute for Factory Automation and Production Systems (FAPS) at the Friedrich-Alexander University Erlangen-Nuremberg (FAU).

REFERENCES

[1] G. Reinhart, Ed., *Handbuch Industrie 4.0: Geschäftsmodelle, Prozesse, Technik*. München: Hanser, 2017.

[2] K. Schmidt, J. Bönig, G. Beitinger, N. Thielen, and J. Frank, "An approach to quality prediction using intelligent SMT solder joint inspection through the use of Machine Learning," in *GMM-Fachbericht*, vol. 94, *EBL 2020 - Elektronische Baugruppen und Leiterplatten*, M. Nowottnick, Ed., Berlin, Offenbach: VDE VERLAG GMBH, 2020.

[3] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539.

[4] Léon Bottou and Olivier Bousquet, "The Tradeoffs of Large Scale Learning," in *Advances in Neural Information Processing Systems*, 2008, pp. 161–168. [Online]. Available: http://leon.bottou.org/publications/pdf/nips-2007.pdf

[5] T. Hastie, R. Tibshirani, and J. Friedman, *The elements of statistical learning, second edition: Data mining, inference, and prediction,* 2nd ed. New York: Springer, 2009.

[6] V. Christlein, L. Spranger, M. Seuret, A. Nicolaou, P. Kral, and A. Maier, "Deep Generalized Max Pooling," in *2019 International Conference on Document Analysis and Recognition (ICDAR)*, Sydney, Australia, Sep. 2019 - Sep. 2019, pp. 1090–1096.

[7] A. Z. K. Simonyan, "Very Deep Convolutional Networks for Large-Scale Image Recognition," [Online]. Available: http://arxiv.org/pdf/1409.1556v6

[8] N. Cai, G. Cen, J. Wu, F. Li, H. Wang, and X. Chen, "SMT Solder Joint Inspection via a Novel Cascaded Convolutional Neural Network," *IEEE Trans. Compon., Packag. Manufact. Technol.*, vol. 8, no. 4, pp. 670–677, 2018, doi: 10.1109/TCPMT.2018.2789453.

[9] G. Acciani, G. Brunetti, and G. Fornarelli, "Application of Neural Networks in Optical Inspection and Classification of Solder Joints in Surface Mount Technology," *IEEE Trans. Ind. Inf.*, vol. 2, no. 3, pp. 200–209, 2006, doi: 10.1109/TII.2006.877265.

[10] W. Hao, Z. Xianmin, K. Yongcong, O. Gaofei, and X. Hongwei, "Solder joint inspection based on neural network combined with genetic algorithm," *Optik*, vol. 124, no. 20, pp. 4110–4116, 2013, doi: 10.1016/j.ijleo.2012.12.030.

[11] Maximilian Metzner, Daniel Fiebag, Andreas Mayr, and Jorg Franke, "Automated Optical Inspection of Soldering Connections in Power Electronics Production Using Convolutional Neural Networks,"

[12] Omron Corporation, *VT-X750: High-speed automated X-ray CT inspection system Handbook*.

[13] Abadi et al., "TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems," [Online]. Available: https://arxiv.org/pdf/1603.04467.pdf

[14] F. Chollet and others, *Keras*.

[15] Alex Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in *Advances in Neural Information Processing Systems 25*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds.: Curran Associates, Inc, 2012, pp. 1097–1105. [Online]. Available: http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf

[16] Waleed Abdulla, *Mask R-CNN for object detection and instance segmentation on Keras and TensorFlow*: Github, *GitHub repository*.

[17] E. Shelhamer, J. Long, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 4, pp. 640–651, 2017, doi: 10.1109/TPAMI.2016.2572683.

[18] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.