Hanhirova, Jussi; Harjuhahto, Jaakko; Harjuhahto, Janne; Hirvisalo, Vesa

A machine learning based quality control system for power cable manufacturing

Published in:
Proceedings - 2019 IEEE 17th International Conference on Industrial Informatics, INDIN 2019

DOI:
10.1109/INDIN41052.2019.8972281

Published: 01/01/2019

Document Version
Peer reviewed version

Please cite the original version:
Hanhirova, J., Harjuhahto, J., Harjuhahto, J., & Hirvisalo, V. (2019). A machine learning based quality control system for power cable manufacturing. In *Proceedings - 2019 IEEE 17th International Conference on Industrial Informatics, INDIN 2019* (pp. 193-198). Article 8972281 (IEEE International Conference on Industrial Informatics). IEEE. https://doi.org/10.1109/INDIN41052.2019.8972281

A machine learning based quality control system for power cable manufacturing

Jussi Hanhirova*, Jaakko Harjuhahto†, Janne Harjuhahto† and Vesa Hirvisalo*
**Department of Computer Science*, *Aalto University*, Espoo, Finland
{jussi.hanhirova, vesa.hirvisalo} @aalto.fi
†*Maillefer Extrusion Oy*, Vantaa, Finland
jaakko.harjuhahto@mailletech.net janne.harjuhahto@maillefer.net

***Abstract*—We study methods for observing physical defects on the surface of power cables. Quality control is essential for power cable manufacturing and surface defects are an important quality factor. Traditionally power cable manufacturing has relied on manual inspection as automated methods have not been sufficient to be used in industrial production.**

We have designed and implemented a novel defect detection system that applies machine learning methods to detect power cable surface defects. Our system uses laser scanning to map the surface of a cable during production. For the machine learning, we have evaluated different CNN (Convolutional Neural Network) architectures and studied their performance and accuracy. According to our results, CNNs are suitable for the detection of surface defects as they can be trained with large amounts of cable surface data.

***Index Terms*—deep learning, convolutional neural networks, power cables, quality control**

I. Introduction

Using machine learning is an option for automating some parts of the quality control that is needed in power cable manufacturing. We address the quality control for manufacturing of high and extra high voltage power cables, which need rigorous testing after manufacturing. The traditional approaches have mostly relied on human inspection.

There is added value in detecting manufacturing defects as soon as possible, before the produced reel is sent to the time consuming subsequent processes. Visual anomalies at the surface of a cable can be caused by numerous underlying problems inside the insulated cable core. Directly connected to the production process, detecting such defects calls for soft real-time methods. Such approaches are necessary in the path toward, *e.g.*, just in time and zero defect manufacturing [6].

Using machine learning methods, it is possible to distinguish between naturally occurring surface features and features that are related to significant defects. In the recent years, particularly Convolutional Neural Networks (CNN) have proven to be a powerful tool for many similar applications [1] and excel in image classification. In addition to their classification performance, CNNs can automatically extract features from complex data sets without manual programming their recognition. Additionally, pre-trained CNN models can be fine-tuned with limited data sets to meet specific industrial requirements, such as tailoring system performance for a specific production line or product. Finally, CNN model inference can be accelerated to achieve differing real-time requirements.

We have done research to understand what kind of CNNs can be used and how the quality control based on them could be arranged in power cable manufacturing. There exists a number of CNN architectures and a number ways of training them. Further, CNNs can be implemented with various machine learning frameworks. All the frameworks have specific formats to model the structure of a CNN network. Also, arranging for the related sensing and sensor data stream processing is demanding in factory environments.

Our contribution in this paper is to present our findings with a prototype that we have designed and implemented. Our experiments with the prototype show that well-known pre-trained CNN models can effectively be fine-tuned to detect defects. Our system has internal threshold parameters that can be used to control the false positive rate and defect detection thresholds. Our performance evaluations with different GPU hardware show that the computational requirements can be satisfied with currently available hardware.

The structure of this paper is the following. Section II describes power cable manufacturing and discusses using machine learning for quality control in that context. Section III introduces a cable topography scanner that is used for defect detection and Section IV describes the related training of the CNNs used. We continue our presentation by describing our experimentation and our experimental results in Sections V and VI. We end our paper with our conclusions in Section VII.

II. Power cable manufacturing

Quality control is a key factor in power cable manufacturing [5]. Power cables consisting of high and extra high voltage cables, ranging from 66 to 525 kV, are subject to rigorous testing after manufacturing, in order to find any defective production lengths before installation. Routine electrical testing can only be done after all of the production phases. The first production phase where a metal conductor is insulated with three layers of plastic, inner semi-conductive layer, insulating layer and outer semi-conductive layer, is the most critical and prone to manufacturing defects. There is added value in detecting manufacturing defects as soon as possible, before the produced reel is sent to time consuming subsequent processes.

Physical defects on the surface of the insulated core can be caused by numerous underlying problems inside the plastic layers of the insulated core. Any irregularities at the plastic layer interfaces will cause heightened electrical stress through a field enhancement phenomena. Increased local electrical stress will significantly increase the risk of electrical breakdown.

Currently most of the produced high voltage or extra high voltage cables are not examined by any surface quality monitoring system. Current measurement systems are typically aimed towards measuring the layer thicknesses of the three plastic layers with x-rays and are not designed for local surface defect detection.

The surface defects encountered in power cable production can be roughly divided into two categories. First, a piece of non-homogeneous material visible as a roundish bump on the surface of the cable. The bump can be preceded or succeeded by a dent at the surface, due to disrupted flow of the plastic. Second, any mechanical damage at the surface of the cable due to unintended physical contact of the cable with production machinery. These scratches and dents can be of any shape or depth. Not all scratches are critical, but the presence of any mechanical damage is an indication of potentially faulty equipment or a need of maintenance.

Quantitative real-time metrics are valuable for the quality- and process control of an ongoing production process, as well as enablers for long-term data-driven advances towards smarter factories and intelligent manufacturing [6], [7]. The primary challenge with gathering such metrics for this application using automated surface defect detection is accuracy. The surface of power cables during manufacture is not necessarily uniform. Sometimes there is visual discoloration, chemical marks, intentional markings, or residues left over from the manufacturing process, such as cooling water, on the cable surface, which can cause false alarms. However, most of the variations found at cable surfaces are not related to significant quality defects. Automated methods have typically yielded too many false positive detections to be practically usable. Using machine learning methods it is possible to distinguish between naturally occurring features and meaningful surface defects.

Methods based on machine learning are an option for surface defect identification as they can be trained with large amounts of cable surface data. In the recent years, particularly Convolutional Neural Networks (CNN) have proven to be a powerful and efficient way to implement detection [1]. One of their biggest advantages compared to many other approaches is the absence is manually programming the various features that describe an observation that is sufficiently significant to be counted as a quality defect.

CNNs can be implemented with various machine learning frameworks [2]. The frameworks all have specific formats to model the structure of a CNN network. Further, there exists a number of CNN architectures [3], whose properties differ. CNNs are mainly used for supervised learning in which the network is trained for a specific task. Training requires usually a large amount training data and considerable computational time. The inference done by CNNs is usually much faster than their training, but often computational acceleration is needed in order to achieve a desirable inference rate. A large variety of specific accelerators have been developed [4], but general purpose GPUs are often used as they are readily available and relatively cheap.

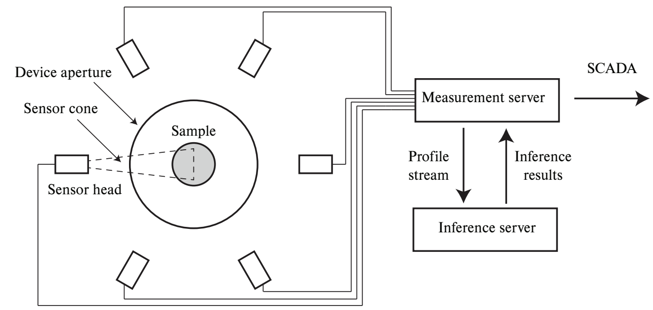


Fig. 1: Logical system overview of the measurement device. The number and arrangement of sensors depends on the range of cable dimensions a particular variant of the scanner supports.

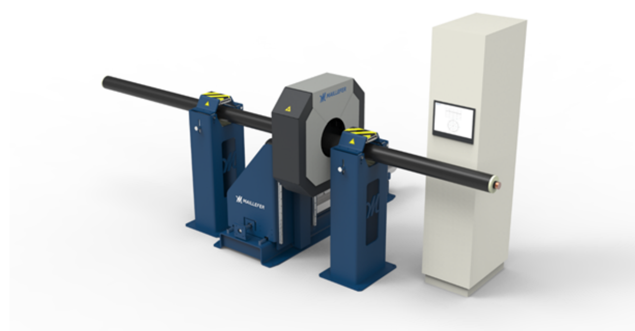


Fig. 2: Render of the topography scanner. The measured cable sample passes through the aperture of the device.

III. Cable topography scanner

Our experimental setup is built upon a power cable surface scanning instrument developed by Maillefer Extrusion. This device was designed for measuring the surface geometry of power cable products in real-time in order to establish online quality metrics for the product and the production process. The device generates a detailed 3D mesh of a cable's surface geometry: a topographic map of the surface. This surface map is then used to compute analytic quality metrics for the product and for detecting the presence and type of surface defects, such as incisions, scratches, extrusion residues, scorched raw material etc.

This topography scanner is composed of several optical sensors, a control and computing unit and an inference unit. An overview of the system is presented in Figure 1. Figure 2 presents a 3D render of the topography scanner installed as a part of a power cable production line.

A. Sensor subsystem

An array of laser displacement sensors is placed around the continuously moving cable product in an arrangement to cover the full 360° of the cylindrical cable. Each sensor measures the distance from itself to an arc of the cable's circumference. Merging these measurements into a single cross section yields a single profile and a sequence of profiles forms the surface mesh. For a typical high-voltage power cable application, the sensors operate at a sampling rate of 200 Hz to 800 Hz while the cable moves at a speed of 0.5 - 2.0 meters per minute.

B. Measurement subsystem

The sensor system is attached to a server running a custom data acquisition and processing application. Measurements from individual sensors are assembled into profiles and the the data is filtered and conditioned as necessary.

The number of raw data points per profile depends on the number of sensors, the size and position of the measured sample. During regularization, profile size is reduced to a fixed number of points (*e.g.*, 3600), equally spaced around the sample to facilitate further processing and analysis, such as computing quality metrics for the product.

The computing server offers an interface for publishing the full profile data and any derived metrics to subscribers over a local network. The three most significant subscribers are: a visualization client for real-time monitoring of metrics and the cable surface, a SCADA plant automation systems for logging and alerting, and the deep learning inference service used for identifying abnormalities on the cable products surface.

C. Inference subsystem

The inference subsystem uses CNN model inference to detect surface defects on the measurement subsystem data. Overview of the inference subsystem is presented in Figure 3.

The inference subsystem retrieves profile measurement data from the measurement subsystem and maintains a batch buffer for the profile data. Based on overlap configuration, it transforms the profile data batches into overlapping tiles that are sent for inference. Results of the CNN inference are transmitted back to the measurement subsystem.

We evaluate the inference subsystem functionality using three metrics. First, does the system detect all of the defects. Second, how much it generates false positives. Third, does the system function in real-time.

IV. TRAINING

For prototyping purposes, it is difficult to obtain necessary amount of real life defects in order to train a machine learning system. We overcome this by artificially generating the majority of the defects in laboratory conditions. Reports and pictures from cable factories were used as guidelines to model the most common defect types.

The cable samples used to generate measurement data for this paper were realistic samples having outer diameter between 60 mm and 130 mm from Maillefer's Pilot Vertical Line as well as cable samples from commercial cable factories.

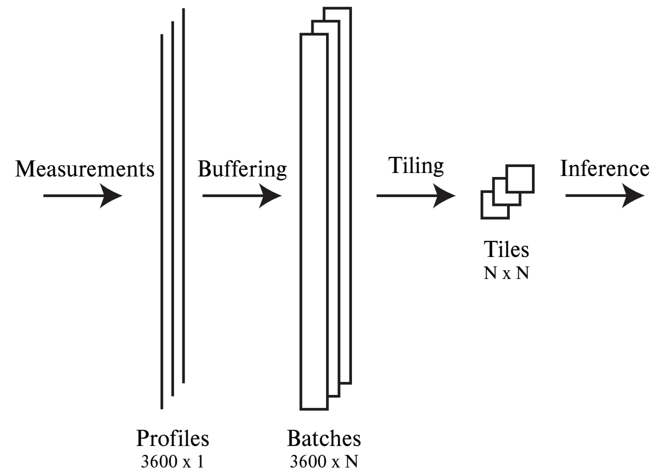


Fig. 3: Overview of the inference subsystem. The inference subsystem pulls profile measurement data from the measurement subsystem and stores the profile data in a buffer. Based on the tile size NxN and overlap configuration, it generates multiple tiles from the buffered data. Tiles are sent for inference and the results are passed back to the measurement subsystem.

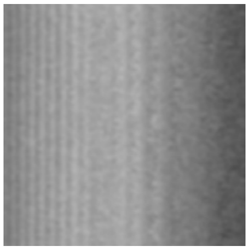
The Pilot Vertical Line is an R&D platform designed for manufacturing realistic high and extra high voltage cable samples at pilot scale. Most of the machinery is identical to that of a real vertical continuous vulcanization line and representative samples can be created with identical dimensions using similar process parameters. The Pilot Vertical Line was also used to purposefully generate defective cable samples representing real production defects.

When this quality control system is employed at a power cable manufacturing line it will be necessary to obtain more training data from that particular line and fine-tune the system for better line specific usage. In a real production environment the cable surface has a certain amount of natural texture and possibly some line specific surface marks. From a cable manufacturers point of view the system has to be adjustable to minimize the amount false positives while retaining losing detection accuracy of true positives.

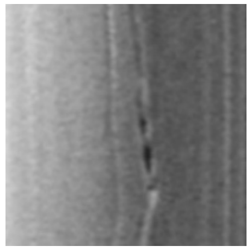
To train and evaluate the CNN models for defect recognition, we develop a training data library. Development of the training data library is based on scanning the sample cables using different measurement configurations. The generated measurement data is then annotated by hand to create masks representing the defect locations on the surface mesh. From this library, we are able to generate different training and validation data sets using adjustable augmentation parameters.

Due to relatively small amount of cable surface measurement data, we map the surface mesh data to grayscale images. This allows us to use pre-trained image classification CNNs as a base to fine-tune the defect detection CNN models.

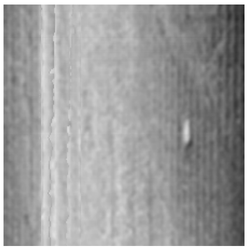
For performance evaluation and accuracy-speed trade-off characterization we have selected three well-known CNN



(a) A clean sample.



(b) An incision on the outer semi-conductive layer.



(c) A lump of non-homogeneous material embedded onto the sample's surface.

Fig. 4: Grayscale renders of surface mesh tiles. The tiles represent areas roughly 12mm by 8mm on the surface of a cable sample. The defects shown in 4b and 4c are less than 0.2mm in depth and height, respectively.

models that have been pre-trained using the ImageNet data set [8]. We fine-tune the selected CNN models using transfer learning [13] with our custom training data.

Figure 4 presents examples of different types of samples in the data. By using a configurable overlap parameter in the inference subsystem tile generation, we attempt to catch all defects in multiple tiles for maximizing the possibility of detection.

In this paper, we use a balanced training data set consisting of 28000 clean and 28000 non-clean tiles. The test data set is generated from cable sample not used in the training data set. The test data set contains a total of 27 defect area annotations. By generating the test data sets for different tile overlap values, we acquire test data sets representing industrial usage scenarios, where the ratio between clean and non-clean tiles is realistic. With an overlap value of 0%, approximately 200 tiles per defect class and 7000 clean tiles are generated. With an overlap value of 80%, 6000 tiles per defect class and 190k clean tiles are generated. Using such a unbalanced test data sets allows us to measure the defect hit-rate and the false positive rate, and estimate how the system would perform in a real industrial usage scenario.

V. Measurements

For evaluating the trade-offs in the inference subsystem, we evaluated three different CNN models. The selected models are MobileNet [9], InceptionV3 [10], and ResNet50 [11]. These three models represent well-known CNNs that have been proven to work well for image classification problems and are lightweight enough for soft real-time usage. All three models have been pre-trained with the ImageNet dataset. We evaluate the models using the TensorFlow framework [12].

We evaluated their real-time performance on two different hardware platforms: NVIDIA Tesla V100 and Tesla P4. We measured the average inference latency for different batch sizes corresponding to different tile overlap values. Relationship of the tile overlap values and batch sizes are presented in Tables I and II.

We also measured the CNN model detection performance using a separate test data set. We measured the false positive rate of the different models and a custom defect hit rate metric.

In the surface scan measurement data, a defect can span across multiple tiles. Using overlap in the tile formation we attempt to capture all possible defects onto multiple tiles. We consider a defect to be found if at least one of the tiles where the defect is present is classified as defect. To measure this, we define the defect hit rate metric as the number of found defects divided by the total number of defects.

VI. Results

Figures 5 and 6 present the average inference times measured from the three models using different tile overlap values. With the Tesla V100 GPU the overlap value of 10% CNN models attain the lowest inference times. This is due to the Tesla V100 GPU being underutilized with overlap value of 0%. Going to overlap values of 20% and beyond, the inference times start growing steadily as there is an increased amount of processing load.

With the Tesla P4 GPU, there is a smaller decline with 10% overlap. Tesla P4 is less powerful when compared to the Tesla V100 GPU, and thus it attains relatively good utilization also with smaller batch sizes (overlap of 10%).

Overall the curves in Figures 5 and 6 follow the same rising pattern. The more there is computational load the more it takes time to compute the inference results.

With the Tesla V100 GPU the most time consuming CNN model is the ResNet50, while on Tesla P4 it is the InceptionV3 model. The two GPUs are based on different internal architectures and have different internal designs and capabilities. Different CNN models compose of different computations, which map differently to different GPU hardware.

In general, there is a magnitude of difference between the inference performance of the two GPUs. Based on this observation, an inference subsystem based on the Tesla V100 could be used to serve multiple topography scanner systems simultaneously.

Figure 7 presents the false positive rate measured from the three CNN models using different tile overlap values. The InceptionV3 model attains the false positive rate minimum at the overlap value of 40%. The ResNet50 model attains the false positive rate minimum at the overlap value of 60%. The MobileNet model does not generate any false positive classifications with the test data set.

TABLE I: Inference performance requirements at sampling rate of 400 Hz for 224x224 tiles.

Overlap %	0	10	20	30	40	50	60	70	80
Inferences per batch	17	18	21	23	27	33	41	54	81
Batches per second	1.8	2.0	2.2	2.6	3.0	3.6	4.5	6.0	8.9
Inferences per second	28.7	35.4	44.8	58.6	79.7	114.8	179.4	318.9	717.5
Milliseconds per inference	34.8	28.2	22.3	17.1	12.5	8.7	5.6	3.1	1.4

TABLE II: Inference performance requirements at sampling rate 400 Hz for 299x299 tiles.

Overlap %	0	10	20	30	40	50	60	70	80
Inferences per batch	12	14	16	18	21	25	31	41	61
Batches per second	1.3	1.5	1.7	1.9	2.2	2.7	3.3	4.5	6.7
Inferences per second	16.1	19.9	25.2	32.9	44.7	64.4	100.7	179.0	402.7
Milliseconds per inference	62.1	50.3	39.7	30.4	22.4	15.5	9.9	5.6	2.5

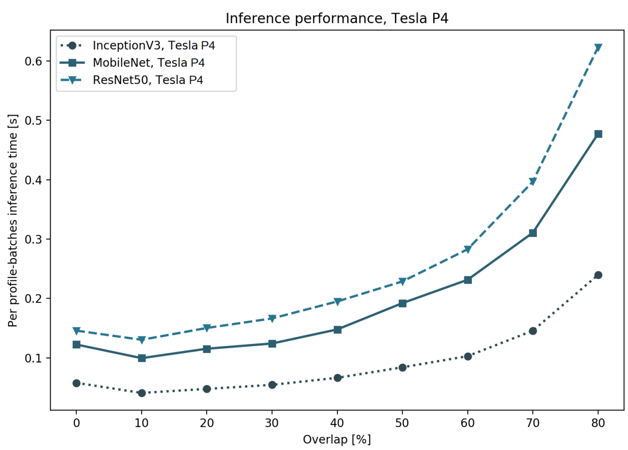


Fig. 5: Mean inference times for different amount of profile batches (overlap values) using the three CNN models on NVidia Tesla P4 GPU.

Fig. 6: Mean inference times for different amount of profile batches (overlap values) using the three CNN models on NVidia Tesla V100 GPU.

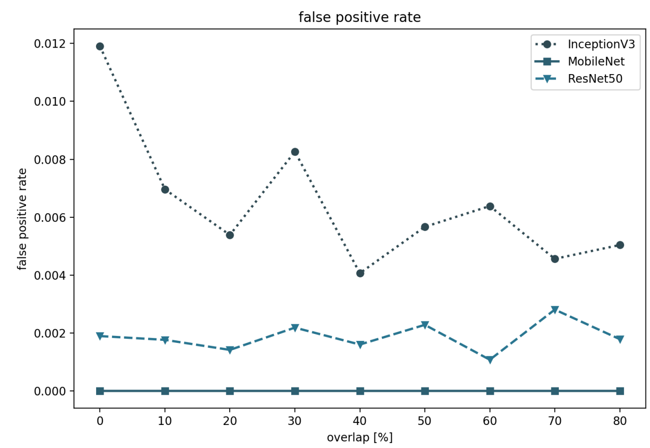


Fig. 7: False positive rate of three CNN models presented with increasing the tile overlap values.

The false positive rate results depend on the test data set. There is a random factor influencing how the tiles are formed based on the overlap factor, which affects the false positive rate results. With more data more accurate evaluations could be made.

Figure 8 presents the custom defect hit rate measurement results for the three CNN models under evaluation. There are 27 defect areas in the test data set. When the overlap value is larger, more test data is generated. The low number of defects causes the hit ratios to get fixed values.

The MobileNet and ResNet50 models perform better with overlap value of 0% than with overlap of 10%. This is likely due to random factor of placement of defects in the 224x224 pixel tiles (for Inception and ResNet50) as compared to 299x299 tiles for the InceptionV3 model.

From Figure 8, it can be seen that the ResNet50 model achieves the highest defect hit rate. With the test data, it is in practise able to find all defects. But as seen from Figure 7, it also generates false positive classifications.

While the MobileNet model in Figure 8 is not able to detect all the defects it does not generate any false positives, as seen in Figure 7.

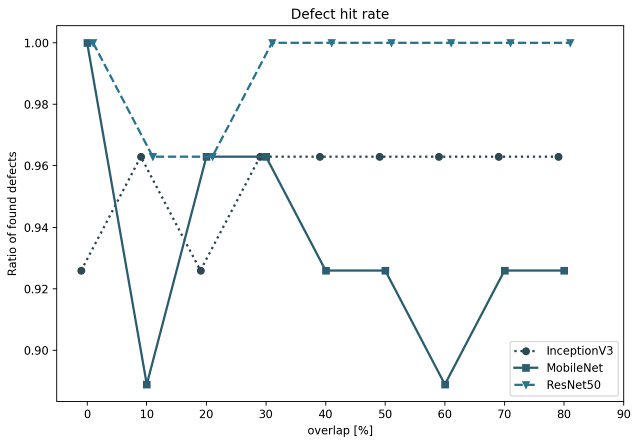


Fig. 8: Defect hit rate shows how many of the defects in the test data are found. A defect is considered found if at least one of its tiles is classified as non-clean.

The InceptionV3 model performs worst of the three models on both the false positive rate and mean defect detection metrics. This is most likely due to the different input size of the InceptionV3 model (299x299 vs. 224x224). The cable defect data seems to be better captured using smaller tile sizes.

When using smaller tile sizes the inference performance requirements grow rapidly. In addition, small tiles do not cover as much of cable surface, making some of the defects difficult to distinguish from the clean surface. Further research is needed to find out optimal tile size and CNN model configurations.

Regarding the results presented in Figures 7 and 8 it can be seen that there is a trade-off between minimizing the false positive rate and maximizing the defect detection rate. For the goal of maximizing the defect detection probability the ResNet50 model performs most accurately. When minimization of the false positives is wanted, the MobileNet model gives the lowest scores.

It is possible to combine the capabilities of the MobileNet and ResNet50 models. The ResNet50 model can be used to find the defects and the MobileNet model can be used to validate the findings.

Based on the initial system evaluations presented in this paper, detecting cable surface defects using CNNs appears viable. More real-world data and measurements are needed to further evaluate the real-world performance of the approach.

VII. Conclusions

Quality control is vital element in power cable manufacturing. Traditional manufacturing methods have relied on manual inspection as automated methods have yielded too many false positive detections to be practically usable. In this paper, we present an automatic machine learning based system for quality control of power cable manufacturing.

Our experiment results show that well-known pre-trained CNN models can effectively be fine-tuned to detect typical power cable surface defects. Our measurements show that with suitable system configuration the false positive rate can be limited near zero while defects can still be detected with high confidence. More research is required to find an optimal CNN model for this application, in terms of detection accuracy as well as computational requirements.

Our performance evaluations with different GPU hardware show that the computational requirements can be handled even with high manufacturing line speeds. Furthermore, the measurements suggest that with more capable hardware a shared inference platform could be used to serve multiple topography scanner systems.

Future work is required to evaluate different custom CNN models with varying input sizes and internal complexity, in order to find even better balance with false positive rate minimization, detection accuracy maximization and optimal computational performance requirements. This evaluation work is mostly manual, while the search space is huge, which is one of core challenges for deep learning system development.

Our topography scanner system can extract detailed quality information from cable manufacturing process, which has not been possible earlier. Linking the detailed information to line control data opens up new means for smart process control and future smart manufacturing.

Acknowledgment

We would like to thank Dr. Mikko Lahti at Maillefer Extrusion for his support.

This work has been supported by a Technology Industries of Finland Centennial Foundation grant.

References

[1] Y. LeCun Y. Bengio and G. Hinton. "Deep learning," Nature, Vol. 521, May 2015

[2] S. Pouyanfar et. al. "A survey on deep learning: algorithms, techniques, and applications," ACM Computing Surveys, Vol. 51, No. 5, September 2018

[3] Open Neural Network eXchange (ONNX) Model Zoo. 2019. https://github.com/onnx/models

[4] V. Sze, Y. Chen, T. Ya and J.S. Emer. "Efficient processing of deep neural networks: a tutorial and survey," Proceedings of the IEEE Vol. 105, No. 12, December 2017

[5] Harry Ernest Orton and Rick A. Hartlein. "Long-life XLPE-insulated Power Cables," 2006

[6] R. Y Zhong, X. Xu, E. Klotz and S.T Newman. "Intelligent manufacturing in the context of Industry 4.0: a review," Engineering, Vol 3, No. 5, October 2017

[7] E. Westkämper. "Strategic development of factories under the influence of emergent technologies," CIRP Annals, Volume 56, Issue 1, 2007

[8] ImageNet. 2019. http://www.image-net.org/

[9] Andrew G. Howard, et. al. "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision". 2017. http://arxiv.org/abs/1704.04861.

[10] Szegedy, Christian, et. al. "Rethinking the Inception architecture for computer vision," CVPR. 2016.

[11] K. He, X. Zhang, S. Ren, and J. Sun. "Deep residual learning for image recognition,". In CVPR, pages 770–778, 2016.

[12] TensorFlow. 2019. https://www.tensorflow.org/

[13] K. Weiss, T.M. Khoshgoftaar, D.D. Wang. "A survey of transfer learning," Journal of Big Data Vol. 3, Issue 1, December 2016