FIQA

Public Report

Authors: Eric Lindahl, Patrik Rydén, Jun Yu, Mattias Silver

Date: 2019-08-20

A project within Hållbar produktion

FFI Fordonsstrategisk Forskning och Innovation

Table of contents

Kort om FFI
FFI är ett samarbete mellan staten och fordonsindustrin om att gemensamt finansiera forsknings- och innovationsaktviteter med fokus på områdena Klimat & Miljö samt Trafiksäkerhet. Satsningen innebär verksamhet för ca 1 miljard kr per år varav de offentliga medlen utgör drygt 400 Mkr.

För närvarande finns fem delprogram; Energi & Miljö, Trafiksäkerhet och automatiserade fordon, Elektronik, mjukvara och kommunikation, Hållbar produktion och Effektiva och uppkopplade transportsystem. Läs mer på www.vinnova.se/ffi.

1 Sammanfattning

Vid tillverkning av fordon (lastbilar, bilar etc.) är ytbehandlingsprocessen i hög grad automatiserad, med undantag av kvalitetskontrollen som i stor utsträckning sker manuellt. Volvo Group Truck Operations (GTO), som är en del av Volvo Group har tillsammans med Umeå Universitet och Volvo Cars identifierat ett behov av att införa ett automatiserat kvalitetssystem för inspektion och rotorsaksanalys av målade karosser. Detta för att minska produktionskostnaden, minska miljöpåverkan samt skapa ett långsiktigt samarbete mellan Umeå Universitet och svensk fordonsindustri.

Projektet FIQA består av två arbetspaket: WP1 - Defektdetektering och klassificering samt WP2 - Alarmsystem och rotorsakanalys.

WP1 har utvecklat ett Computer Vision system baserat på tekniken Phase Measuring Deflectometry (PMD). En stor mängd annoterat bilddata har samlats in i Volvo GTOs produktionsanläggning i Umeå, där hytter till Volvo Lastvagnar produceras. Datat har använts för att träna och testa klassificerare och även för att utvärdera dess prediktionsförmåga. Vi har utvecklat en statistisk inlärningsmetod för att upptäcka defekter, inklusive bildtagning, extraktion och defektklassificering. Den utvecklade nya metoden ger inte bara noggrann och pålitlig klassificering utan ger också en osäkerhetsbedömning av klassificeringsresultaten.

WP2 har bidragit med att bygga ett koncept för ett automatiserat alarm- och rotorsaksanalys, inklusive datainsamling, konstruktion av hyttspecifika variabler, validering av variabler samt modellering av kvalitetsutfallet. Från ett stort antal sensorer från produktionsprocessen mäts olika processvariabler i realtid och detta används sedan för att övervaka kvaliteten på målade hytter och för att analysera orsakssambandet när defekter uppstår.

De utvecklade metoderna är generiska och kan tillämpas i svensk bilindustri.

Ett pilotsystem har byggts inom måleriet i hyttfabriken. Systemet hittar och klassificerar defekter online i produktion med nästan 100% noggrannhet för två defekttyper på en del av hytten. Ett system för alarm- och rotorsakanalys byggdes också genom att ansluta 139 sensorer online till systemet. Detta system är mer av konceptuell art, eftersom produktionssystemet inte tillåter överföring av större datamängder utan några större investeringar. Fortfarande anses resultaten vara mycket lovande, eftersom vi har visat att ett sådant system kan byggas, om än i mindre skala.

FIQA-projektet har potential att minska produktionskostnader för kommersiella fordon genom att sänka tillverkningskostnader, minska behovet av reparationer och justeringar (minskad miljöpåverkan), generera stabilare kvalitetsutfall samt minska antalet falska larm.

FIQA har lett till omfattande samarbeten mellan Volvo GTO, Volvo Cars och Umeå Universitet och resulterat i ett antal vetenskapliga publikationer och examensarbeten av hög kvalitet. Samarbetet mellan den svenska fordonsindustrin och Umeå Universitet har gjort det möjligt för universitetet att utveckla kurser och program som är mycket relevanta för fordons- och tillverkningsindustrin. Att utbilda studenter av hög kvalitet är avgörande för den svenska bilindustrins framtid.

2 Executive summary in English

Surface treatment (production) of commercial (trucks, cars, busses etc.) bodies is highly automated except for the inspection, evaluation of what to repair and root cause analysis to avoid production disturbances.

The FIQA projects aims to develop an automated system to find and classify defects in a paint shop environment and a predictive root cause and alarm system to minimize disturbances. This will increase the quality level, stabilize the quality level and lower cost. This will strengthen the competiveness of the Swedish automotive industry and strengthen then research and innovation climate. This will also lead to an increased cooperation within the academy and the industry.

The project consists of two work packages:

- WP1 - Defect detection and classification
- WP2 – Alarm system and root cause analysis

In WP1, we have developed a statistical learning approach for defect detection on painted cab surfaces, covering image acquisition, feature extraction and defect classification. The aim is to develop an approach that not only produces accurate and reliable classification but also provides uncertainty assessment for the classification results. To accomplish this we build feature descriptors using regression spline ideas applied to grey intensity pixel values of the acquired images. As the inspection of specular surfaces inflicts special challenges, PMD technique using reflected sinusoidal fringe is applied to capture images of the considered surfaces. Classification is then achieved using a probabilistic classification algorithm based on k-NN classifier.

We've focused our efforts on the two most common defect types: dirt and crater. An image acquisition and annotation Graphical User Interface (GUI) was created and implemented in to facilitate an extensive data collection. The data collection consists of images from cabs in production along with annotated defects. Images from 20527 cabs was collected. The dataset was divided in two parts one for training and validation.

Our results shown that almost 100% of defect detection and 0% of false alarms are achieved when applying the proposed approach. Equivalently, nearly all patches with both crater and dirt were correctly classified, and no defect-free patch was predicted as defective.

The developed statistical learning approach is generic and can be applied to similar tasks and therefore it is of strongly relevance to the Swedish automobile industry in general.

Within WP2 we have developed methods and procedures for data collection, construction of cab-specific process variables, modelling of quality and alarm and root cause analysis. We have collected raw data for a period of 65 weeks, including data from more than 80,000 cabs, 700 sensors, 100 categorical variables and 100 quality variables. Altogether close to 1 billion observations were collected.

The collected raw data are unstructured and need to be structured in order to enable modeling of different quality variables (the response variables). We have developed a novel approach that combines senor and tracking data to construct a set cab-specific process variables. Altogether more than 3,000 cab-specific process variables (CPVs) were constructed and observed on more than 80,000 cabs.

The relationships between the CPVs and the quality variables (for example indicators for dust and crater defects) were initially modelled using univariate approaches resulting in further pre-processing of the data as well as some basic understanding on how different sub-processes influence quality. Two interesting results were that sanding by machine increases the risk of crater defects later in the process and that cabs spending long time in some of the buffer zones have an increased risk of having dust defects.

The idea behind the large scale modelling is that the risk of a certain type of defect (e.g. craters) can be described as a function of the CPVs. Several approaches, including both machine learning and statistical learning, have been used for modeling the risk of having defects and the results show that the predictions to high degree mimic the empirical data.

In the Volvo Group Truck Operations (GTO) production plant in Umeå, cabs for Volvo Trucks are produced. At the production plant, a pilot system was built within the paint shop. The system finds and classify defects online in production with nearly 100% accuracy for two defect types on a 0.2 m^2 surface. A concept system for an alarm system and root cause analysis was also built by connecting 139 sensors online into the system. The alarm system setup is more of conceptual art, as the production systems don’t allow big data handling without some bigger investments. Still the results are considered very promising, as we’ve proved that such a system could be built in theory.

The FIQA project has delivered the desired pilot system, meeting most of the specifications, and has in addition generated general knowledge, specific improvements and a platform for future research.

The FIQA-project has the potential to reduce the conversion costs of commercial vehicles by: lowering labor cost, reducing the need of repair and adjustments (reduced environmental impact), generating a more stable quality, reducing false alarm costs, enable a more effective update of the production process.

FIQA has resulted in a successful long-term collaboration between Swedish automotive industry and academia with focus on automated quality control systems based on high-dimensional data and signal processing and statistical learning methods. It provides also a foundation for future collaborations and research applications including grants from EU. The research conducted in FIQA is innovative and internationally competitive.

FIQA has led to extensive collaborations between Volvo GTO, VC and UMU and resulted in a number of high quality scientific publications and master thesis projects. The closer collaboration between the Swedish automotive industry and UMU has enabled and encouraged the University to develop courses and programs that are highly relevant for the automotive and manufacturing industry. To produce high quality master students is crucial for the future of Swedish automotive industry.

3 Background

Surface treatment (production) of commercial (trucks, cars, buses etc.) bodies is highly automated except for the inspection and evaluation of what to repair and adjust. Today quality inspection and evaluations are commonly performed manually by operators (inspectors). The inspector identifies and describes defects (Detection), determines if and how the defects should be repaired (Evaluation and Repair instructions), and alarms if defects occur too frequently (Alarm) in which case an investigation, aiming to identifying the root causes, is initiated (Root cause analysis). Manual inspection and evaluations leads to very high costs connected to inspection, repair of defects and identification of root causes. The main cost drivers are manning (manual inspection and evaluation) and production losses (due to unneeded and extra repair work).

To be able to replace the manual inspection, minimize the consequences of defects, and improve the quality of delivered products to customers with a more efficient and automated quality control system. Volvo Group Truck Operations (GTO) as a part of Volvo Group, has together with Umeå University (UMU) and Volvo Cars (VC) identified a need for a research and development project within the quality inspection and repair of painted vehicle bodies.

Automated statistical process quality control systems based on real time data from a large number of sensors, including 2D and 3D cameras have successfully been implemented in production processes including surface treatment of vehicle bodies, for instance in The Ford Motor Company.

Developing and implementing an automated system for monitoring production quality is a highly interdisciplinary task that requires expertise from several areas. Within the project we have expertise in surface treatment of vehicle bodies (Volvo GTO and VC), statistical process quality control (Volvo GTO, VC and UMU), machine color-based vision and image analysis (UMU), classification and pattern recognition (UMU), and multivariate analysis of high-dimensional spatiotemporal data (UMU). The proposed research project FIQA (Finish Inspection and Quality Analysis) covers the quality work connected to detection and evaluation of surface finish defects within the manufacturing of painted vehicle bodies.

4 Purpose, research questions and methodology

In this chapter, two work packages will be described

4.1 WP1 - Defect detection and classifications

Purpose:
Develop a pilot Automated Quality control System (AQS) for detection of defects: automated inspection, detection and description of surface finish defects, and determine if the defect needs to be repaired or not.

AQS specification and review of available monitor systems:
Clarify which defects are relevant to be detected with AQS, how fast needs AQS to operate, how AQS will need to interact with other systems. Identification and review of available systems to monitor quality and detect defects (advantages, limitations and costs for implementing the system).

Description of the modelling approach:
Machine color-based vision and image analysis together with advanced statistical modeling of high-dimensional spatio-temporal data can be very useful for image segmentation and pattern recognition, e.g. recognizing defects on painted vehicle bodies. The apparent color of an object depends on various factors and color-based machine vision can be broadly categories in developing computational color constancy algorithms and related classification algorithms. An analysis based on only one image/sensor is often insufficient and therefore combining data from multiple sensors is crucial to obtain the desired information. In this work, we will employ various types of classification procedures and computational color constancy modeling to develop an AQS that enables automated detection and description of defects.

AQS will for specified time (training period and test period) run in parallel with manual inspections, resulting in data collection consisting of manually identified defects (response data), raw data from AQS (explanatory image data) and data collection from the production process (explanatory production data). A classifier will be obtained using state-of-the-art statistical methods and the observed training data (response data, explanatory image data and explanatory production data). The performance of the classifier will be evaluated using the test data, with respect to its ability to predict and describe different types of manually identified defects and not least the uncertainty of these predictions. The modelling approach will for any type of surface abnormality estimate the probability that this deviation corresponds to a certain type of defect. These types of estimations will be based on a continuous test variable that measures quality at a higher resolution than the manual inspection does, which in turn can be used to monitoring quality over time, determining which defects should be repaired and to improve alarm detection.

Pilot system – Detection of defects:
Make a model (pilot) of a system to be used to test its application. This includes setting up AQSs and conducting data collection and modelling as described above.

Implementation costs:
Summarize investment and start-up cost for an implementation

Standardize way of working:
Describe the new way of working including necessary support to run the new system.

4.2 WP2 – Alarm and root cause analysis

Purpose:
To identify sources that contributes to variation in quality (in particular defects) and develop an automated system that alarms if the production quality is drifting.

Production data:
Determine what production, process and product data (explanatory production data) can be collected, together with the AQS-data and the manual inspection data. Determine what type of generic data base should be used. Identify needs to implement new or change existing standards or demands.

Root cause analysis:
AQS measures quality at different scales: it predicts defects and measures quality at a continuous scale with some test variable T (quality response), where T can be a vector addressing different aspects of quality. Statistical techniques based on multivariate regression and correlation will be used to link the quality response T with the explanatory production data and to identify exploratory variables that impact the final quality. Note that the exploratory variables can be both continuous and categorical, and the interaction terms between these variables can also be a part of root cause. Options for variable screening or feature selection enable us to screen exploratory variables for root cause analysis as well as the methods that can be used to find the exploratory variables that are important. In general, predictor statistics can be computed by the respective method, and then predictors can be ranked based on the method-specific measure of predictor importance. There are a number of existing methods that may be appropriate for our purpose, such as generalized linear model, classification and regression trees, boosted trees, multivariate regression splines, neural networks, etc.

To most efficiently find desired improvements and examine the interactions of exploratory variables statistical design of experiments (DOE) is of key importance. DOE can be used to find the optimum process parameters for a defined use environment or adapt it to find a robust design that is well-suited for a range of use environments. Furthermore, our DOE should also take spatiotemporal structure of data into consideration.

Alarm algorithm:
Specify when the system should alarm and what data the system should be based on (e.g. observed defects or some test variable T).

Pilot system – Root cause analysis:
Identify and organize the production data and the AQS-data in a data base. Apply a root cause analysis as described and identify sources contributing to variation in quality and correlated with defect frequency. Make a report model to visualize the root causes. Evaluate existing Key Performance Indexes (KPI) and propose new or changed KPI's

Implementation costs:
Summarize investment and start-up cost for an implementation

Standardize way of working:
Describe the new way of working including necessary support to run the foreseen new system

5 Objectives

The FIQA project will deliver results linked to Swedish automotive industry, method development and evaluation within the area of automated defect detection, automated alarm system and root cause analysis (Development and evaluation of methods), and results linked to the general task of implementing an automated quality control system (AQS implementation).

Results directly linked to Swedish automotive industry:

- Design and develop test systems for defect detection.
- Design and develop models for automated alarm and root cause analysis.
- Develop a prototype and build a pilot system for automated defect detection, automated alarm and root cause analysis that is evaluated in production environment.
- Deliver a specification for implementation of a complete system based on the research results.

Results connected to development and evaluation of research methods:

- Evaluate various sensors (e.g. 2D and 3D cameras) and classification methods for identification of manually identified defects. This work should result in research papers that shall be submitted to high quality scientific journals and presented on international scientific conferences.
- Develop and evaluate methods for automated alarm and root cause analysis that utilize the data obtained from the automated quality inspection. This work should result in research papers that shall be submitted to high quality scientific conferences and journals and presented on international scientific conferences.

Results connected to general implementation of an AQS:
Propose general instructions and standards on how to implement AQS for manufacturing industry. This is an ambitious task, but we intend to present what worked and did not work with the FIQA-approaches for setting up an AQS at Volvo Umeå and give qualitative and quantitative results showing how these approaches can be applied in general. This work should result in a case study report of the FIQA project.

The proposed research project will contribute to the prioritized aims of the strategic vehicle research and innovation (FFI): sustainable production technology (sv. Hållbar produktionsteknik), surface treatment and painting as stated below.

FFI-1: *To enhance research and innovation capacity in Sweden and thus secure the competitiveness of the Swedish automotive industry.*

The FIQA-project has the potential to reduce the conversion costs of commercial vehicles by:

a) lowering labor cost
b) reducing the need of repair and adjustments (reduced environmental impact)
c) generating a more stable quality
d) reducing false alarm costs
e) enable a more effective update of the production process

FFI-2*: To develop internationally connected and competitive research and innovation environments in Sweden.*

The Volvo GTO’s production plant in Umeå, cabs are producesd for Volvo Trucks. This production plant is among most modern and technologically advanced production units within the Volvo Group with well trained and highly skilled personnel. At Umeå

University several research groups (from the Department of Mathematics and Mathematical Statistics and the Department of Applied Physics and Electronics) will be part of the FIQA-project and contribute with expertise in areas linked to Automated Quality control Systems. This includes expertise in: machine based vision, image analysis, statistical process control, and analysis of high-dimensional spatio-temporal data. The FIQA-project has the potential to result in a successful long-term collaboration between Swedish automotive industry and academia with focus on automated quality control systems based on high-dimensional data and signal processing methods. We believe that the FIQA project builds a foundation for future collaborations and research applications including grants from EU.

FFI-3*: To promote cooperation between industry and universities, colleges and institutes.*

In the short perspective, the FIQA-project will lead to extensive collaborations between Volvo GTO, VC and UMU, which will result in an increased number of high quality scientific publications and master thesis projects. In the longer perspective a closer collaboration between the Swedish automotive industry and the involved departments at UMU, will enable the University to develop courses and programs that are highly relevant for the automotive and manufacturing industry.

6 Result and target compliance

FIQA has successfully built two pilot systems, for automated defect detection and classification and for alarm systems and root cause analysis, respectively. The developed methods are generic and can be applied to Swedish automotive industry.

WP1 developed a computer vision system based on the phase measuring deflectometry (PMD) technique, including a test system, followed by a pilot system. A large amount of image data and annotation data was collected for training and testing our classifiers and also for evaluation of the prediction ability. We have developed a statistical learning approach for defect detection, including image acquisition, feature extraction and defect classification. The developed novel approach does not only produce accurate and reliable classification but also provide uncertainty assessment for the classification results.

WP2 has contributed to build a proof of concept for Volvo GTO's paint process in Umeå, including data collection, construction of cab specific variables, validation of variables, modelling the quality and developing an automated root cause analysis system. From a large number of sensors on the production line various process variables are measured and in real time and then utilized to monitor the quality of painted cab bodies and to analyze the causality when defects occur.

One of main challenges for coming work is to scale up the established pilot system to the whole cab body and develop a real time system for defect detection and classification with an efficient and automated statistical process quality control system.

FFI Target Compliance

FFI-1*: To enhance research and innovation capacity in Sweden and thus secure the competitiveness of the Swedish automotive industry.*
The FIQA-project has the potential to reduce the conversion costs of commercial vehicles by:

a) lowering labor cost
b) reducing the need of repair and adjustments (reduced environmental impact)
c) generating a more stable quality
d) reducing false alarm costs
e) enable a more effective update of the production process

FFI-2: *To develop internationally connected and competitive research and innovation environments in Sweden.*
FIQA has resulted in a successful long-term collaboration between Swedish automotive industry and academia with focus on automated quality control systems based on high-dimensional data and signal processing and statistical learning methods. It provides also a foundation for future collaborations and research applications including grants from EU. The research conducted in FIQA is innovative and internationally competitive.

FFI-3: *To promote cooperation between industry and universities, colleges and institutes.*
FIQA has led to extensive collaborations between Volvo GTO, VC and UMU and resulted in a number of high quality scientific publications and master thesis projects. The closer collaboration between the Swedish automotive industry and UMU has enabled and encouraged the University to develop courses and programs that are highly relevant for the automotive and manufacturing industry. To produce high quality master students is crucial for the future of Swedish automotive industry.

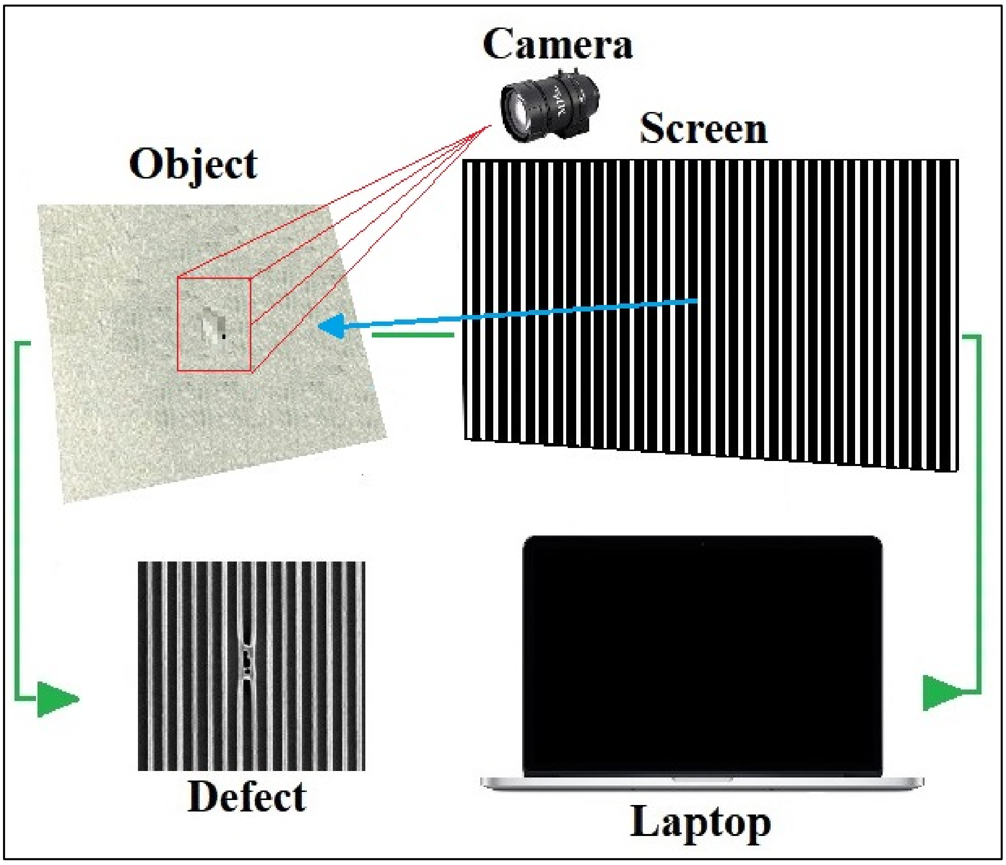


Figure 1: A sketch of the deflectometry-based image acquisition process. At the bottom left, the distortions of the sinusoidal patterns show that there is a defect on the surface.

6.1 WP1 – Find and classify defects

A summary of the results from WP1 is presented below, together with coming challenges for the continuing work. Details are referred to our technical reports and publications [1-4].

A. Computer vision systems

The computer vision algorithm we have developed during FIQA-WP1 is called Phase Measuring Deflectometry (PMD) as shown in Figure 1.

The PMD is a technique which utilizes a reflected sinusoidal pattern to obtain surface local slope, which can be further processed to calculate surface 3D shape. PMD is the well-known technique for analyzing specular surfaces like, lenses, mirror, reflectors and car-body. The truck surface at Volvo facility, Umeå possess similar specular quality. Therefore, we have used this technique to find defects in painted truck body.

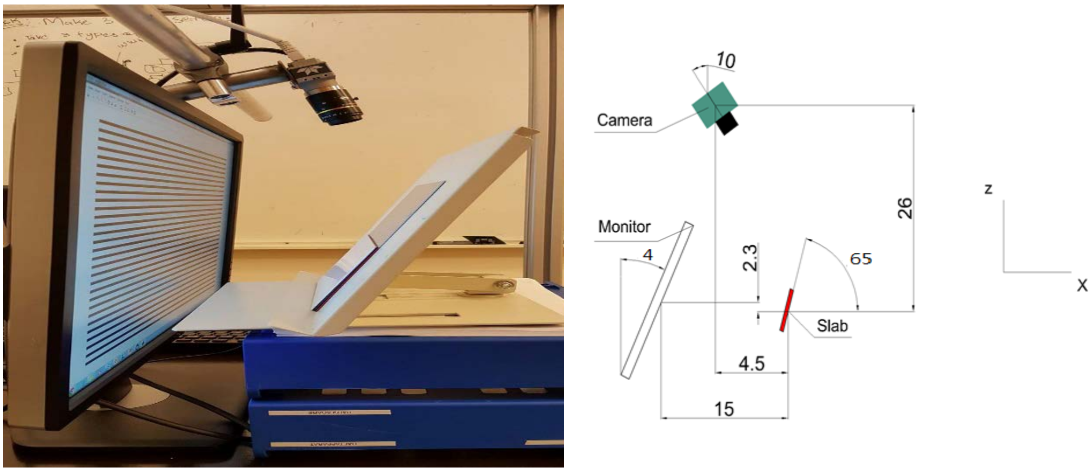
In PMD technique, sinusoidal patterns generated by computer are displayed on the screen and these patterns are then projected on the truck-body and then reflected patterns from the truck body are recorded by the camera.

Set up	Kowa lens+ White ring light	Kowa lens + IR ring light	Deflectometry
Stability to color	✓	✓	✓✓
Stability to surface category	✓	✓	✓✓
Stability to curved surface	✗✗	✗✗	✓
Stability to ambient light condition	✗	✗	✓
Visibility of dirt on the image	✓✓	✓✓	✓✓
Visibility of crater on the image	✗✗	✗✗	✓✓
Conclusion	Good for dirt	Good for dirt	Good for dirt and crater

Figure 2: Various setup evaluation table. One can see that PMD method is the most suitable method when it comes to 'dirt' and/or 'crater' detection in painted surfaces.

In this project, we evaluated various imaging methods for specular surface defect detection epically painted surfaces. The results are shown below in Figure 2 from our various setup for detection [21].

From these investigations, we have concluded that a PMS setup based on three parts (namely- 1) Monochrome camera 2) Kowa lens 3) Sinusoidal pattern display) is good enough for our purpose of defect detection in painted surfaces. Figure 3 shows our PMD setup.



(a) (b)

Figure 3: Our PMD setup based 1) Monochrome camera 2) Kowa lens 3) Sinusoidal pattern display. a) Prototype design b) schematic measurements.

The dataset from our PMD setup is used for both 1) statistical learning methods and 2) direct deflectometry machine vision method. In first part, we have developed various statistical learning methods for efficient classification of defects. The second can be used to assist the statistical learning methods and AR/VR based visualization method to assist the operators for better identification of detect.

One test system followed by one pilot system based on PMD technique have been built during the project.

B. Data collection

a) Defect type
There are many different types of defects that can occur on painted vehicle bodies. The most common defect types are dirt and crater, thereby this report will focus only on these two defects. Dirt can be described as a small bump deposited in, on, or under the painted surface, whereas crater looks like a circular low spot or bowl-shaped cavity on the painted surface. It should be noted that dirt is more common than crater on the cabs. It is almost impossible to distinguish dirt from crater just by observing the captured images, therefore, human eyes and touch are still vital for this task.

b) Image acquisition and annotation
An image acquisition and annotation Graphical User Interface (GUI) was created and implemented in MatLab to facilitate the shooting and annotation parts. When a new object is detected in the system, sixteen images are automatically taken by the image acquisition and annotation GUI. After the shooting, the object’s surface is inspected manually with care and under good lighting conditions. Together with the defect annotation, the whole process is accomplished in less than fifteen seconds.

c) Data collected
The total number of cabs collected during the whole period (week 12 – 48, 2018, except the holiday period of weeks 29-32) is 20527. The dataset was divided in two parts, Data II and Data III. Note that Data I represents the data collected in the test system, which was used for analysis during the testing stage. The patches close to the edges of the images were excluded from the whole dataset to withdraw the edge’s effect. For Data II, each subset contains 18433 patches of eight different window sizes, which include 4234 patches of dirt, 372 patches of craters and 13827 defect-free patches. 1170 patches close to the edges were removed from each subset. For Data III, each subset is formed by 16170 patches of eight different window sizes, including 3376 patches of dirt, 515 patches of craters and 12279 defect-free patches. 836 patches close to the edges were removed from each subset.

C. Statistical learning for classification

In WP1, we have developed a statistical learning approach for defect detection on painted cab surfaces, covering image acquisition, feature extraction and defect classification. The aim is to develop an approach that not only produces accurate and reliable classification but also provides uncertainty assessment for the classification results. To accomplish this we build feature descriptors using regression spline ideas applied to grey intensity pixel values of the acquired images. As the inspection of specular surfaces inflicts special challenges, PMD technique using reflected sinusoidal fringe is applied to capture images of the considered surfaces. Classification is then achieved using a probabilistic classification algorithm based on k-NN classifier.

The scientific contributions of our work include the following novel elements necessary to succeed in defect detection and classification.

1) We propose novel feature descriptors based solely on the smoothness degree of the fitted splines. These allow avoiding the usage of classification algorithms with computational expensive training phases that are most commonly used for surface defect detection. Instead we apply the probabilistic classifier based on k-NN algorithm that results in highly accurate and reliable classification. A number of feature descriptors and classifiers has been evaluated and compared, including HOG (histogram of oriented gradients) [22,23], LBP (local binary patterns) [24], 2D DWT (2D discrete wavelet transform) [25], smooth features based on P-splines [26], variabilities [27, 28], EDF (effective degrees of freedom) [4], RF (random forest) [10], CNN (convolutional neural network) [29], SVM (support vector machine) [9], probabilistic k-NN (k-nearest neighbors) [30].

2) We provide a nonparametric patch-wise probabilistic classification approach built upon the one nearest neighbor rule with Euclidean distance. The estimates of the proper probabilities for each class are obtained using the concept of NN-balls in the feature space [30].

3) The probability based performance evaluation metrics are presented as alternatives to such conventional metrics as misclassification error rates, false positive and false negative rates. Moreover, the vectors of posterior probabilities of the considered classifier allow for classification quality assessment in terms of the uncertainty measure, Entropy [30].

The preliminary analysis of WP1 has been carried out to compare the performances of different classifiers such as CNN, RF, SVM and k-NN. Regardless of the feature vector chosen, SVM revealed itself to be the best classifier in terms of prediction accuracy whereas k-NN came out as the fastest one in those previous studies. These two classifiers, SVM and probabilistic k-NN, were then selected for the further investigation with more collected new data.

The classification results show that the proposed approach is very promising. In comparison with all considered feature descriptors, the proposed EDF based feature vector performs the best in terms of both conventional and probability based performance metrics. The value of the frequency parameter of the sinusoidal pattern has only marginal influence on the detection performance of the EDF. In contrast, the performance of the alternative features decreases for higher values of the frequency, which is even more profound when using the probabilistic k-NN. Furthermore, the proposed features showed equally good classification results in terms of false positive and false negative rates, whereas the competitive approaches appear less capable of detecting defect, erroneously yielding its absence. This indicates that the EDF features achieve not only higher classification performance but also can be viewed as an optimal approach to resolve the sensitivity-specificity trade-off task. Moreover, the introduced probability based evaluation metrics suggest that the approach is certain its decisions.

To investigate the defect detection capability of the proposed approach for finding the most common types of defects such as crater and dirt, we examined a 3-class classification problem. Among the alternative features considered in the study, the feature vector based on HOG produced the best results when solving the binary classification task. So, the performance results of the HOG are shown here together with those of the EDF. The SVM classifier using HOG features failed to distinguish between three classes, at the same time SVM on the EDF performed as good as k-NN on EDF. So, the results on SVM performance are not included here.

Our results shown that almost 100% of defect detection and 0% of false alarms are achieved when applying the proposed approach. Equivalently, nearly all patches with both crater and dirt were correctly classified, and no defect-free patch was predicted as defective. Contrary to the EDF, the HOG descriptors are capable of defect detection to a much lower degree, accounting for more than 10% of false alarms and higher than 50% of non-detections. The performance results of the HOG are worsening for the datasets of patches that correspond to channels with larger values of the frequency parameter, whereas the EDF performs equally well for all presented datasets.

The developed statistical learning approach is generic and can be applied to similar tasks and therefore it is of strongly relevance to the Swedish automobile industry in general. Besides the success, there are challenges left for further investigation:

a) It was time-consuming and expensive to collect the data and many challenges emerged during this process. Hundreds of data were thrown away due to unavailable information caused by some reported bugs from the image acquisition and annotation GUI. According to the size of the mouse click error from the operators, some defects could be hidden. Sometimes when the click error was big, some patches (often with smaller window sizes) did not even include these defects. The click error tended to be bigger when dealing with smaller defects because it was easier to miss them.

b) The defect's hiding could as well be triggered by their micrometer size or by darker colors. Another vital and previously known issue was to separate craters from dirt during annotation. Touching the surface was still crucial to tell if it is a bump or a hole. Only a small part of the cab is currently investigated, the luggage lid. The luggage lid is flat while many other areas of painted vehicles might be curved, uneven, or contain pressings or gaps. With the luggage lid being one of the easier areas to investigate, the current results might not be applicable to the whole cab.

c) The coming challenges include certainly scaling up the pilot system and developing a real time pilot system for defect detection and classification with an efficient and automated statistical process quality control system. The self-driven pilot system is currently running at Umeå Volvo GTO cab plant and the results look promising despite the limited time for each cab.

d) Developing this system was complex, hence different areas of expertise are requested. Due to the short time available for prediction in real time, it was preferable to use the HOG features. The prediction time was reduced by utilization of

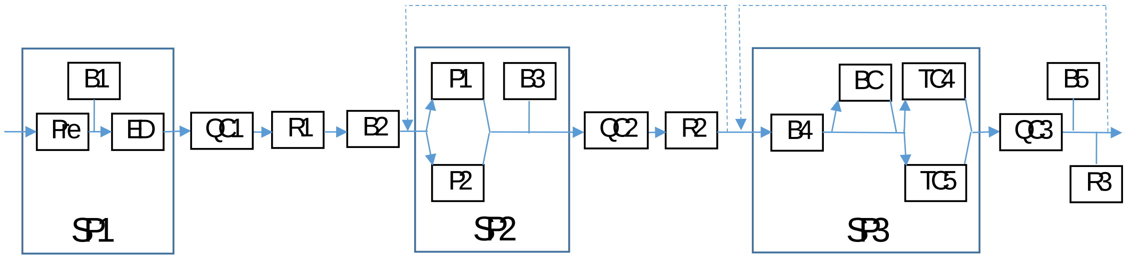


Figure 4: A map of the paint shop with the main sub-processes (SP1, SP2, SP3), quality control stations (QC1, QC2, QC3), repair stations (R1, R2, R3) and selected buffers (B1, B2, B3, B4, B5). Solid arrows indicate the normal movement of the cabs, and dashed arrows the possible routes for send backs.

patches with larger window sizes. The task of estimating the size defects could not be accomplished in this report due to data complexity and restricted time in the project.

6.2 WP2 – Alarm system and root cause analysis

The general problem addresses a production line where a large number of sensors are collecting data in real time and where the objective is to use these data to monitor variation in quality. As a proof of concept Volvo Trucks paint process in Umeå was considered, see Figure 4. The work involved data collection, construction of cab-specific process variables, validation of variables, modelling the quality and developing an automated root cause analysis system. A summary of the results from WP2 is presented below, for further details see [5-8].

A. Data collection

The raw data consisted of four types of data: sensor data, tracking data, categorical data (e.g. color, cab type and cab specific repair data) and quality data. We included data from several sensors, including both frequently and sparsely measured variables. For example oven temperatures measured every third second and chemical concentrations measured once a day. Tracking data consisted of position and time measurements of each cab throughout the paint process. The quality data included manually observed defects. For example, craters, dust, orange peel and color disagreement observed at the different quality control stations, see Figure 4. Raw data were collected for the time period April 2017 to June 2018, i.e., 65 weeks excluding vacation. Data from 700 sensors, 100 categorical variables, 100 quality variables and more than 80,000 cabs were collected. Altogether close to 1 billion observations were collected. The data collection procedure is described in a technical report [5].

B. Construction of cab-specific process variables

The collected raw data are large and complex. These unstructured data need to be structured in order to enable modeling of different quality target variables. We have developed a novel approach that combines senor and tracking data to construct a set of cab-specific process variables (CPVs). A sensor can be physically related to a specific sub-process or to the entire paint process. For example, a sensor measuring oven temperature is physically linked to the oven and a senor measuring outside air-temperature can be relevant for the entire paint process. For each senor and sub-process two related tracking positions, t1 and t2, were identified, where t1 and t2 represents the time when the cab entered and left the sub-processes respectively. Cab-

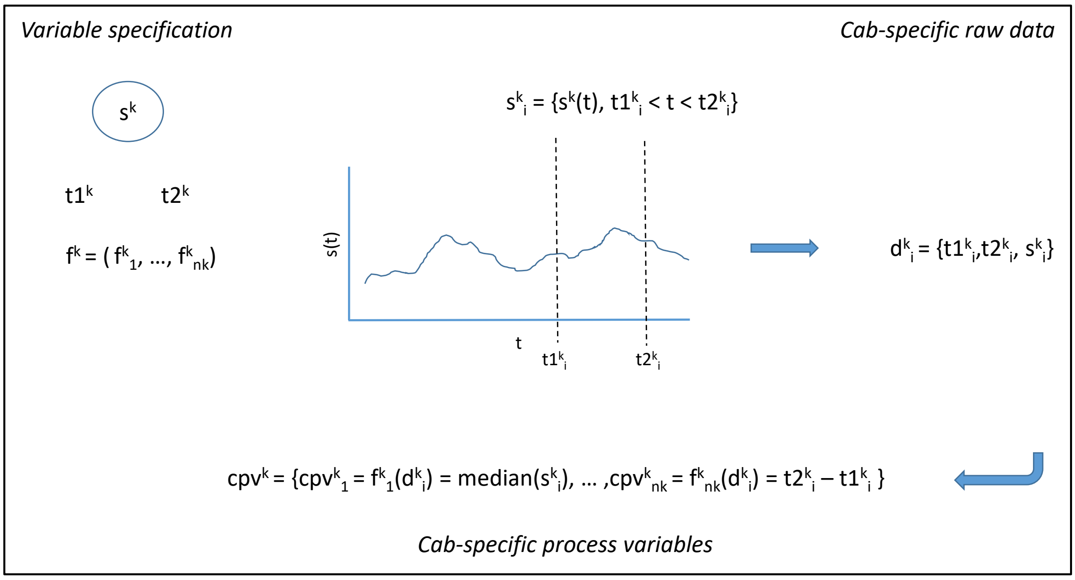


Figure 5: The cab-specific process variables (CPVs) were constructed in a three stage procedure: 1) Variable specification where a sensor s and two corresponding tracing positions t1 and t2 were identified together with a set of relevant functions f. 2) Cab-specific raw data (d) were collected for each cab (indexed with *i*). 3) CPVs were constructed using the raw data and the specified functions. The CPVs can be time specific, sensor specific or depend on both time and sensor values.

specific data on the tracking positions, as well as sensor data observed within the timespan t1-t2, were collected and used to construct several CPVs including for example time within the sub-process (a cab-specific time variable) and the median sensor-value while the cab was in the sub-process (a cab-specific sensor variable), see Figure 5. Interpolation was used to derive CPVs in the case the sensor was sparsely measured.

A second type of CPVs were constructed from cab specific repair data, for example if sanding was applied or not at a repair station. The variable specification of the CPVs were made by personal with deep knowledge of the local paint process. Altogether more than 3,000 CPVs were constructed and observed on more than 80,000 cabs. Henceforth the CPV-data are called the explanatory data.

The quality data are by its nature cab-specific and approximately 20 univariate response variables were constructed and considered. The quality variables included binary response variables (for example cabs with or without crater defects observed at a quality station), count variables (for example the number of observed dust defects) and continues variables (for example the quantified measurements of orange peel). Prior to the modelling the variables were pre-processed and filtered. Variables with a high degree of missing values were removed from further analyses and the remaining missing values were imputed, commonly by replacing the missing data with the variable´s mean value. Similarly, cabs with a high degree of missing values were removed. The data collection procedure is described in a technical report [5].

C. Univariate investigation of the cab-specific process variables

Henceforth the matrix with the collected explanatory variables are denoted X and the response variables are denoted Y. The aim is to model the relationships between each

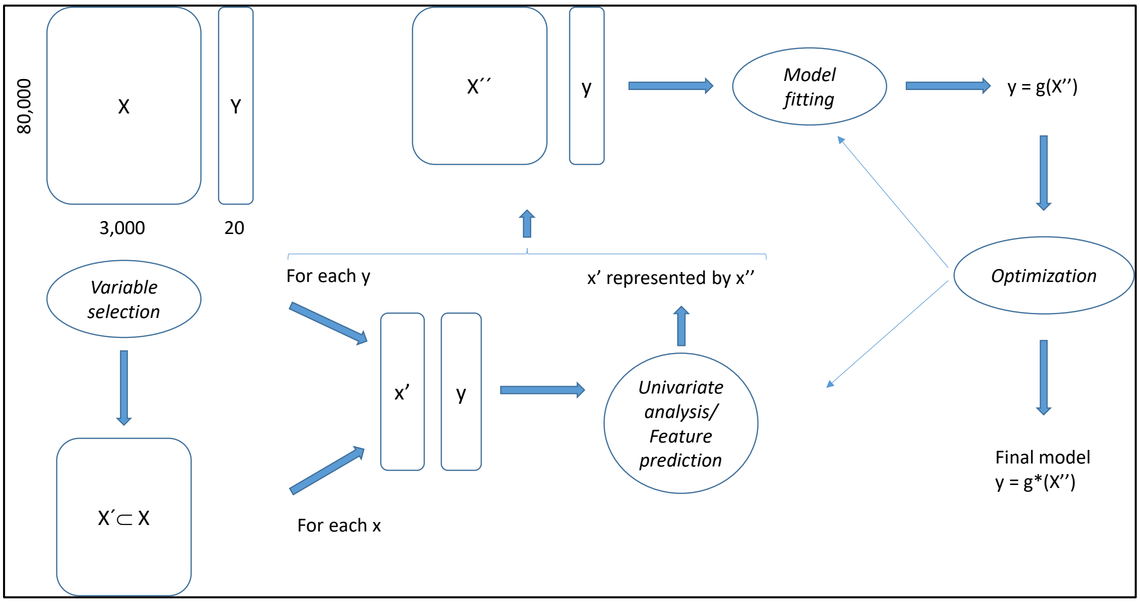


Figure 6: The principal steps of the modeling. The relationship between the explanatory variables X and the response variable of interest (y) was modelled applying external variable selection (reducing the dimension of X), univariate analysis (specifying for example a linear or quadratic relationship between a specific CSV (x) and y) and a modeling and optimization step where the model's hyper-parameters were tuned.

of quality variables (y) and the matrix X, i.e. to fit a function *g* such that y can be predicted as y = g(X). The modelling include three principal steps: external variable selection, feature prediction via univariate analyses and model fitting combined with optimization, see Figure 6.

As a first step of the modelling highly correlated CPVs were identified. In the case the correlated variables were derived from the same variable specification (see Figure 5) one representative were selected and included in downstream analyses. For the relative rare case when the CPVs were derived from "independent specifications" one of the variables were included in the model fitting, but all variables influenced the prediction.

In the second step we conducted univariate analyses in order to specify the relationships between the individual CPVs and the quality variables. Visualization tools were implemented and used to generate several diagnostic plots. The aim was to detect problems such as erroneous sensor data and outliers as well as long and short time trends, see Figure 7. In addition, these analyses were used for feature prediction. For example, suppose that the univariate analysis revels that the relationship between the explanatory variable x and the risk of dust defects is quadratic with the best quality observed at a value b, then it may be convenient to replace x with x´ = (x-b)^2 prior to the modeling, see Figure 7.

The univariate analyses resulted in several interesting findings including that sanding by machine at repair station 2 and spending long time within some buffers resulted in significantly higher risk of crater and dust defects at quality station 3. These results are discussed further in the following sections.

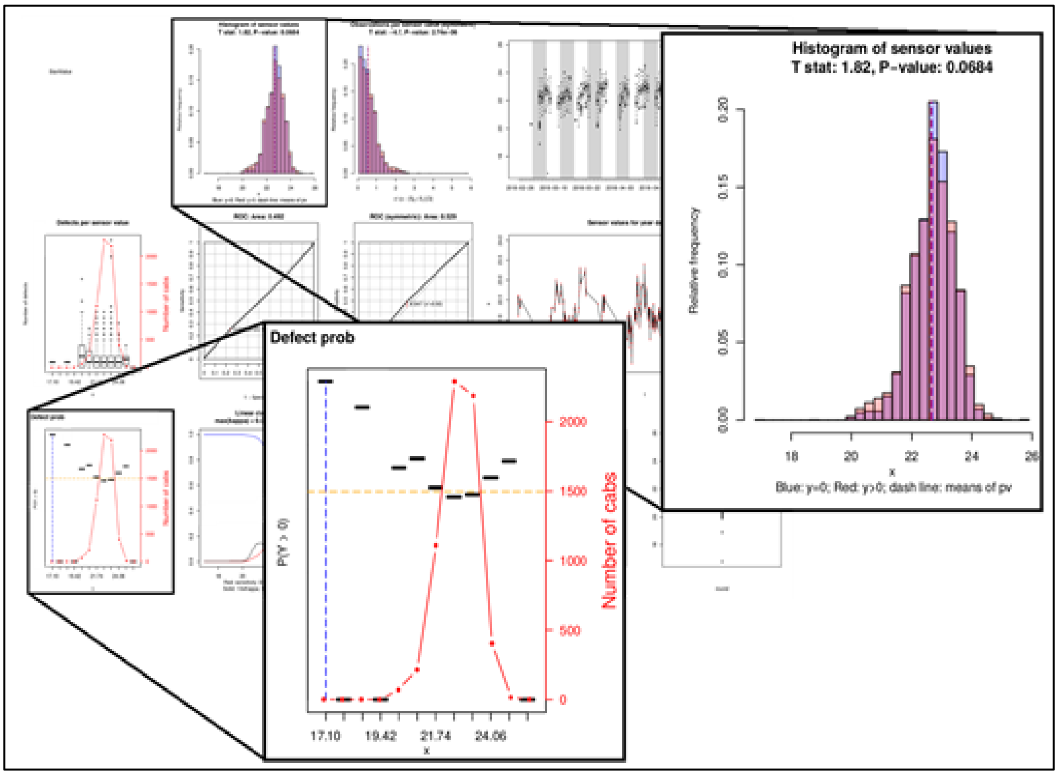


Figure 7: Visualizations for a CSV x and a quality variable y.

D. Sanding by machine at repair station 2 increases the risk of craters at quality station 3

At repair station 2 (R2) 25 % of the cabs were sanded in order to remove dust or crater defects identified at quality station 2. The defected cabs undertook different repair actions: sanding by machine (one or several times), sanding by hand, sanding by both machine and hand or no repair action. The investigation revealed that cabs that were sanded by machine at R2 had an increased risk (base risk % + 3.7 %) of having crater defects detected at quality station 3 (QC3) while the risk of dust defects at QC3 were unchanged. Sanding only by hand did not affect the risk of having crater or dust defects at QC3, see Figure 8. Clearly the analysis do not address how the risk of dust defects at QC3 would have been affected if sanding by machine at R2 would have been omitted. However, the results motivated an intervention study were 20 % of the cabs detected with dust at QC2 were not treated by sanding. Preliminary results suggests that omitting sanding do not increase the risk of dust defects at QC3. Hence, this suggest that the quality control at QC2 and the repair made at R2 can be omitted, leading to a substantial cost reduction and a decreased risk of crater defects. The sanding investigation is presented in a technical report [6].

E. Spending long time within certain buffers increases the risk of dust defects at quality station 3

A buffer in the production line is a place where several cabs can be stored for a longer time. In addition to the major buffers in Figure 4 there are several additional smaller buffers in the production line. If a cab is produced during *uninterrupted production* (i.e. production with no planned and no longer unplanned breaks) it will not spend long time in any buffer while a cab produced with an overnight/weekend interruption is likely to spend more than 5 hours in one of the buffers.

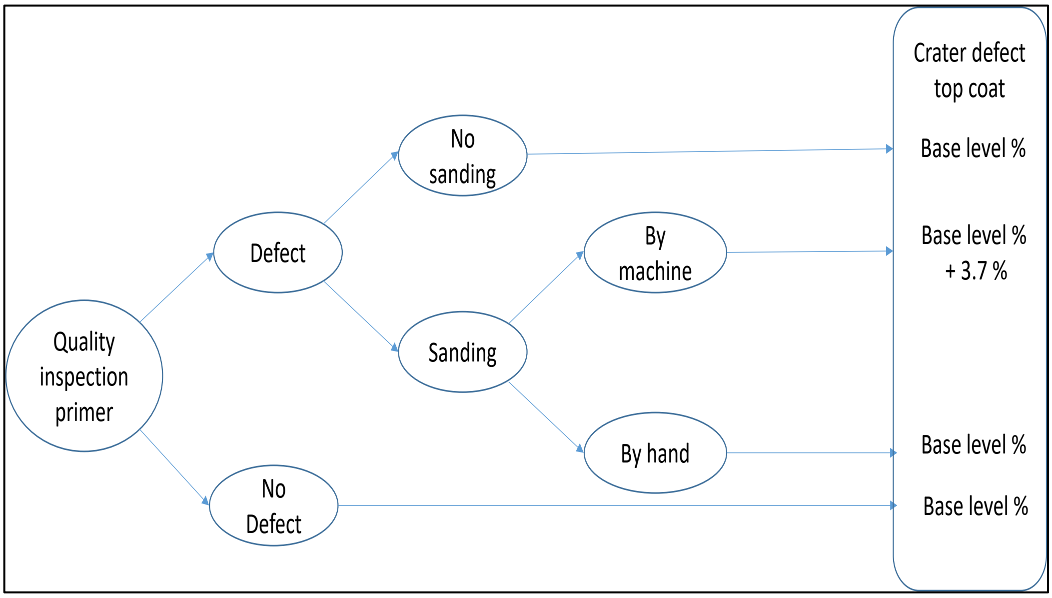


Figure 8: Summary of the sanding study. Sanding by machine at the repair station following the primer process step increases the risk of having crater defects after the top coat process.

Eight buffers were considered and the cabs were divided in nine disjoint categories: spend less than five hours in all buffers, more than 5 hours in buffer 1, …, more than 5 hours in buffer 8. For each category the risk of dust and crater defects at the following quality station (either QC2 or QC3) were observed. The base level risk of dust at QC2 was increased by 6 % units if the cab overnighted in buffer B2, see Figure 4. For cabs spending time in B4 the risk for dust and crater defects at QC3 were increased with 7 and 4.3 % units respectively. The buffer investigation is presented in a technical report [7].

F. Modell fitting, optimization and evaluation

The aim is to build models that allow us to link changes in quality to abnormalities in the production. The major idea is that the risk of a certain type of defect (e.g. craters) is a function of how the cab was produced, which in turn can be described by the observed CPVs. The data for our modelling efforts have been the CPV-matrix X and a binary quality variable y, where the values are one if the cab is defect and zero otherwise. Modell fitting was performed in two principal ways. In the first approach y was modelled as a function of X with the objective to maximize the accuracy or some related measure. The drawback with this direct approach is that the modelling doesn’t take the structure of the underling risk into account. The second approach (the indirect approach) estimated the underlying risk (here the estimated risk is denoted z) based on the observed defects (y), for example by applying a sliding window as shown in Figure 9, and then model z as a function of X with the objective to maximize some distance D. In the simplest case D was the mean square error (MSE) for z and the predicted value of z, i.e.

$$MSE = \frac{\sum_{i=1}^{n} \left(z_i - \hat{z}_i \right)^2}{n} .$$

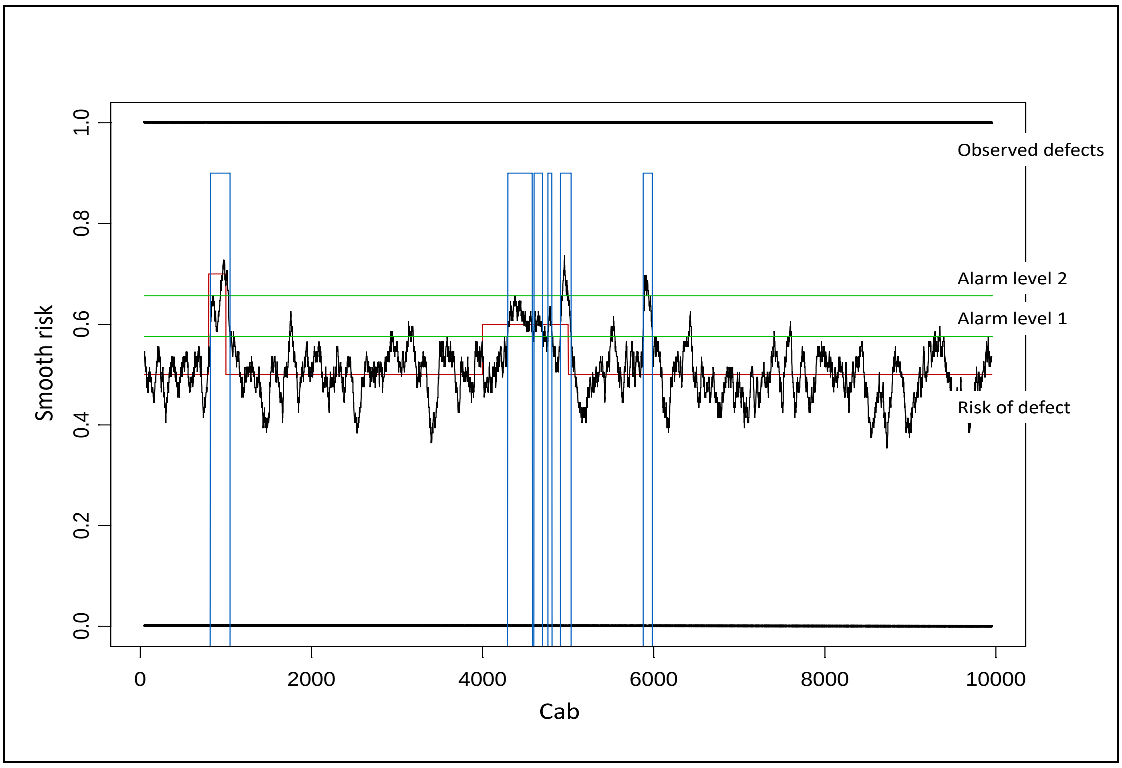


Figure 9: Simulated data where the red line shows the underling risk of defect with an increased risk for cabs 700:900 and 4000:5000. The observed data were either 0 (no defect) or 1 (defect) and the black line shows the smooth defect risk for a window size = 99. The alarm levels were obtained via Monto Carlo Simulations and shows the 95th (alarm level 1) 99th (alarm level 2) percentiles of the smooth risk when the underling risk for defects was 0.5. The blue line shows the alarms, where an alarm was triggered if the smooth risk acceded alarm level 1 for more than 30 cabs in a row.

A drawback is that this distance is sensitive to outliers in the X-data. An alternative is to replace the predicted z vector with a smooth z-vector, e.g. by applying a sliding window to the predicted z-vector or to a predicted binary y-variable constructed from the predicted z-variable.

For the first approach, where the response variable is binary, we identified a number of suitable machine learning algorithms including: penalized logistic regression [9], penalized linear discriminant analysis [9], random forest [10], support vector machines [11], Adaboost [12], Rusboost [13], artificial neural network [14] and deep artificial neural network [15]. In a first attempt we applied penalized logistic regression and used accuracy to evaluate the performance, see Master thesis [19]. As expected the accuracy was low, but despite this the work shoved that the CSVs can be used to model quality and that the models can be used for root cause analysis.

In the indirect approach the response variable is a continuous random variable for which a number of machine learning algorithms are available, including: regression analysis [16], LASSO regression [17], Beta regression [18], support vector regression [9] and random forest regression [9]. This work is ongoing and is carried out by industrial PhD-student funded mutually by Umeå University and AB Volvo. Preliminary results suggest that this is a promising approach that are better than the direct approach to capture quality trends in the, see Figure 10.

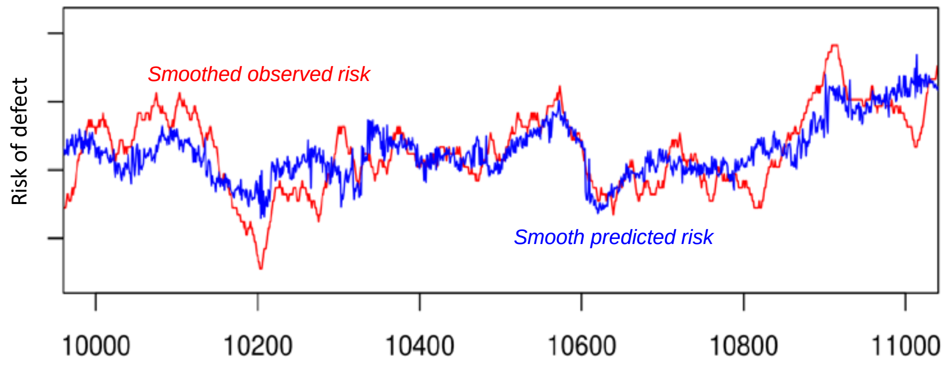


Figure 10: Smooth risk (red line) and predicted defect risk (blue) for 1000 cabs observed 2018.

Independent of approach, the data were split into a training data set (80 % of the observations) and a test data set in order to enable unbiased estimation of the models' performances. Cross validation was applied to tune the models' hyper parameters and to avoid overfitting.

G. Root cause analysis

The derived models link quality outcomes to the production process and can be the basis for process adjustments. This can be achieved either by using the predicted quality to decide when to adjust the process, i.e. an alarm system, or by using the contributions of the variables utilized by the model to infer what adjustments are appropriate, i.e. a root cause analysis.

We have implemented prototypes for both these approaches. The components of our alarm and root cause analysis system (AR-system) include: data on CSVs and quality outcome, predictive models, alarm functions, visualization tools and quantification of risk factors.

Our pilot system is built on historical data, but in order to implement the system data need to be accessed in real time, for example by using database queries or push protocols. An important feature is to determine when to alarm, which can be derived using either the observed quality data or the predicted quality. In the former case the alarm function can be determined via Monte Carlo simulations assuming a base line defect probability, see Figure 9. For linear models (e.g. penalized regression and LDA models) with standardized explanatory variables (x) the coefficients describe (β) describe the relative importance of the CSVs and the products βx, where x is observed just prior the alarm, describe to what degree the CSVs contributed to the alarm. For non-linear models an alternative is to use model-independent feature importance metrics such as mean decrease in impurity or Shapley Additive Explanations, see Master thesis [20]. The pilot AR-system shows the predicted quality, highlights CSVs and sub processes that contribute to the alarm, see Figure 11. A description of the AR-system is presented in a technical report [8].

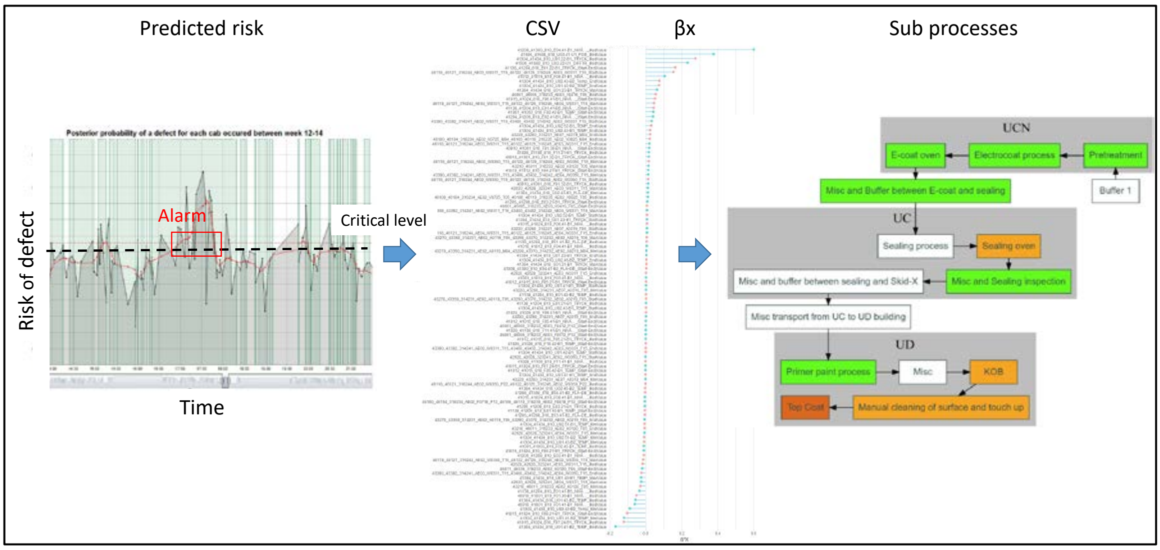


Figure 11: Selected output from the alarm and root cause analysis system. Predicted smooth risk of defect with one indicated alarm (red box). CSVs explaining the increased risk at the alarm by using the products βx. Sub processes contributing to the increased risk, where orange (some contribution) and red (high contribution) indicates that the process is likely to have contributed to the alarm.

6.3 Pilot systems in a production environment

A pilot system has been built within the Volvo's production facility in Umeå. The system contains two modules: one for automatic defect detection and classification, and one module for alarm and root cause analysis. The system is located within the paint shop and placed physically in the process after the top coat oven. The target for the system is the upper part of the luggage lid for the left side of the cab.

The installed equipment includes the following hardware:

- One 55" monitor.
- Two Fujinon HF16SA1 cameras.
- Fixture for 55" monitor and cameras.
- Cables for power, camera, monitors.
- Two monitors from Dell (for visualization).
- One Table (for the visualization monitors).
- One Dell computer, Intel Xeon 5122 CPU @ 3,6 GHz, 160Gb RAM.
- Windows 10, set up as a standard Volvo production environment computer.
- The software were built in MatLab R2018R and Python v3.7.

A patent screening was made within the project. The conclusion was that there was a patent opportunity, especially within WP2. However, it was also concluded that the value for Volvo to continue a patent process was too low since infringements would be difficult to detect. Below follows a detailed description of the two modules.

1. Automatic defect detection and classification

This module consists of three parts: Camera system, prediction model and visualization. The system is trained on dust and crater defects and can detect and classify these defects by nearly 100 %. The system also states if the defect must be repaired or not.

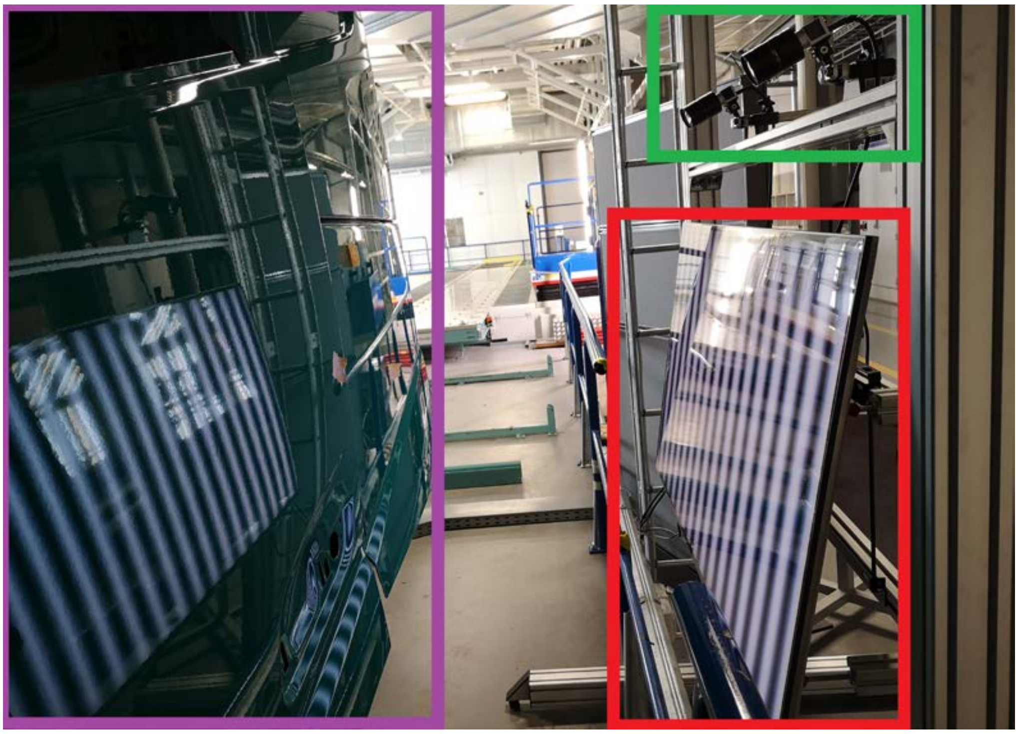


Figure 12: Monochrome camera and Sinusoidal pattern display in a production environment

A. Camera system

The camera system is based on the principle PMD (Phase Measuring Deflectometry) and projects 16 images with a sinusoidal pattern in the paint on the cab. The projection comprises a surface of about 0.2 m^2 on the cab's left boot lid. Two cameras captures the 16 reflections. The camera management system was built in MatLab. The system communicates with the production control system and is thus integrated with the process equipment. The communication is made with a standard ODBC client. Once the cab arrives at the station, the system fetches the cab id from the production environment and triggers 16 sinusoidal patterns on the 55” monitor.

B. Image handling and prediction

A unique pattern is called a channel. The images from the two cameras are put together into one big image for each channel. The images are then stored locally on the computer in a bmp-format. The software creates smaller image patches based on the 16 images. Two prediction models can be used as a base in the system: Support Vector Machines (SVM) + HOG feature vector or Convolutional Neural Networks (CNN).

For SVM the patch pixel size is chosen as 50x50 pixels and for CNN, the chosen pixel size is 128x128 pixels. Based on the pixel information, the system predicts and classifies defects per patch.

The classification result is stored in a text file, containing the following information: Cab id, phase/frequency info, patch position number, probability, type of defect, repair info.

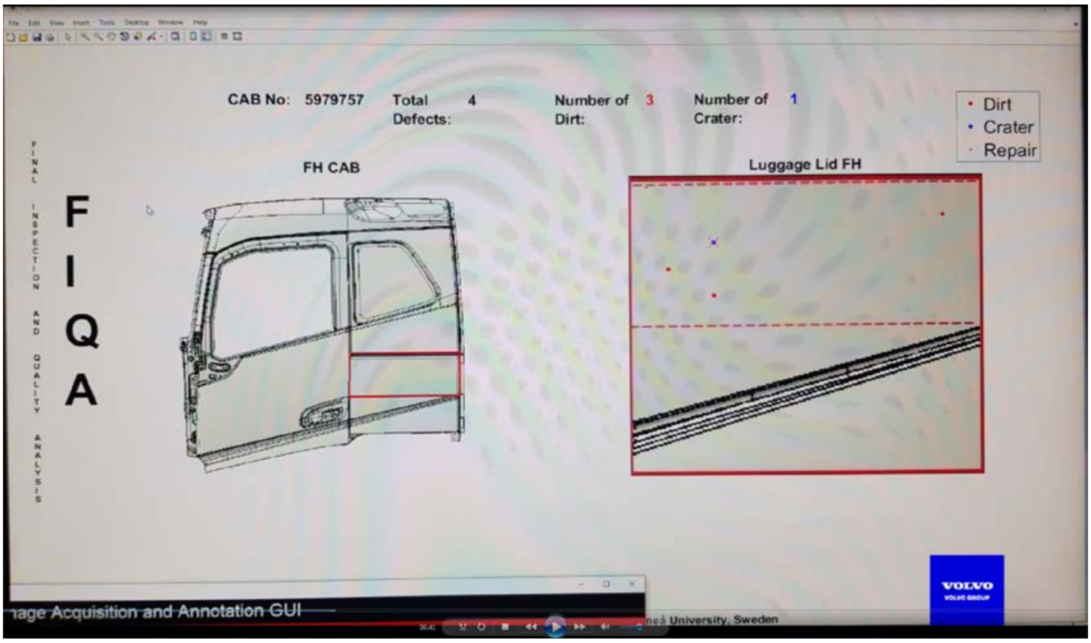


Figure 13: Visualization of defects

C. The visualization module

MatLab functions have been built to visualize the result of the prediction, see Figure 13. The figure shows the cab's surface with the identified defects. It displays different cab images depending on cab model. The system also states which defects to be repaired. The system delivers the prediction result, presented on the visualization screen, in about 3 seconds for the specified surface.

Based on the results we conclude that it will be possible to scale up build a fully automated system that scans the complete cab. An estimation gives that 3-4 robots containing one screen each and two cameras, should be sufficient to cover a cab. In order to scale the system it must be trained and optimized also for curved surfaces and edges. The system must also be able to handle more defect types. There is room for improvement by tuning the shutter speed for the camera, using an optimum number of images, optimizing the patch creation technique. Furthermore, the pilot system uses standard hardware which would help to minimize the implementation cost for a complete system.

The conclusion is that an implementation in practice will be carried out in two steps:

i) Assisting system to operators

An assisting system, covering the complete body and most defects, will show the operator the defects are and how they should be repaired, using a screen or a Virtual Reality device. The system will decrease operational cost as the judgement of defects are standardized, and repair cost and adjustments are minimized.

ii) Fully automated system

This system is completely automated, without any human interaction.
In despite if the system is fully automated or set up as an assisting system, the classified defect data produced by the system is a requirement in order to produce a complete and accurate alarm- and root cause analysis system.

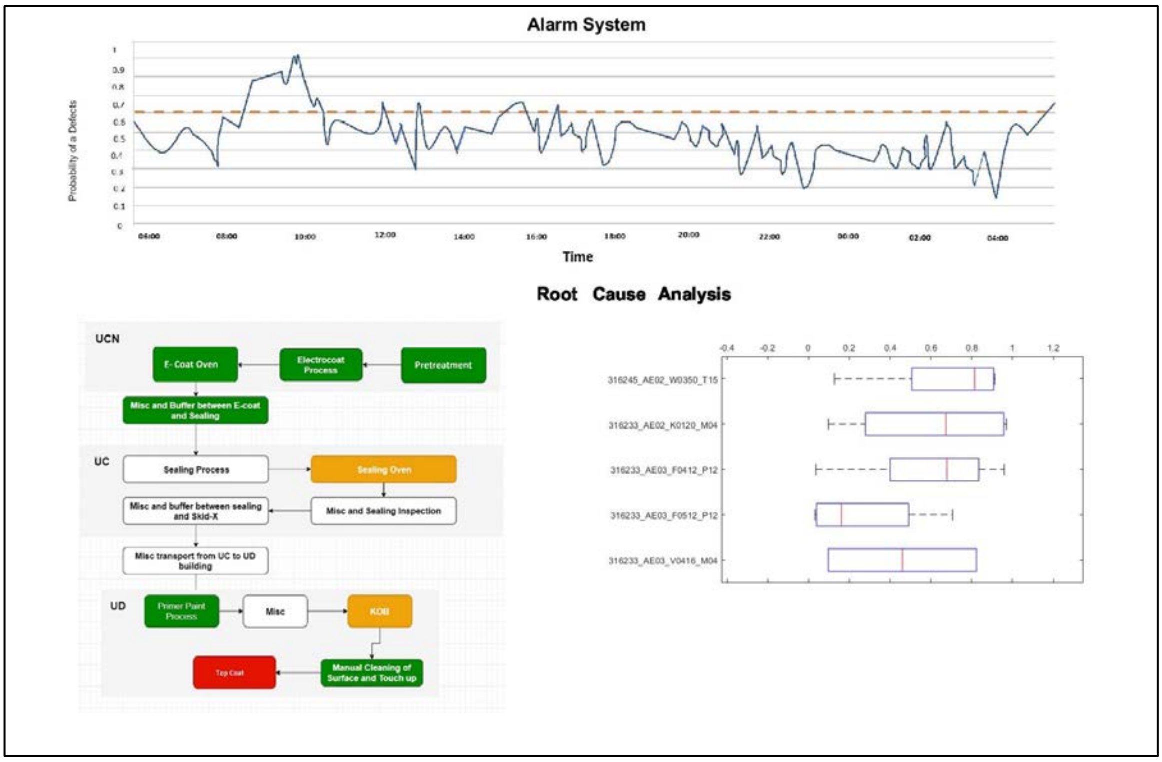


Figure 14: Alarm and root cause analysis visualization. Top part: the predicted risk defects over time. The bottom left: how the sub-processes contributes to the risk. The bottom right: the process variables that contributes most to the risk.

2. Alarm systems and root cause analysis

This module operates on the same computer as the one stated in module 1 and consists of three parts: Data retrieval, prediction calculation and visualization. This pilot system is more of a conceptual system.

Within the paint shop numerous sensors are installed to measure different kinds of conditions in different areas, for example air pressure, air humidity, air flows and temperatures. These sensors are connected to a surveillance control system. An OPC communication protocol was set up and data from 139 different sensors are continuously stored in a database. These values consist of an array with sensor id and the measured value. The system communicates with the production data base with a standard ODBC communication interface made in MatLab. It's continuously monitoring the database and fetches new sensor values online. The sensor id is also linked to a production process position. The system is thus integrated to the paint shop production equipment.

A prediction file, trained on historical data, gives coefficients in order to calculate the posterior risk for cab defects online. The risk is predicted in real time and the system provides the operator with information on how the sub-processes contributes to this risk, see Figure 14.

6.4 Reference list

The list below is a basis for the results presented in this chapter.

[1] M.S.L. Khan and S. ur Réhman. (2018). Computer vision approach towards final inspection quality analysis. *FIQA Research Report 1*.

[2] N. Pya Arnqvist, B. Ngendangenzwa, L. Nilsson, E. Lindahl, and J. Yu. (2018). Defect detection and classification: Statistical learning approach – Part I. *FIQA Research Report 2*

[3] N. Pya Arnqvist, B. Ngendangenzwa, L. Nilsson, E. Lindahl, and J. Yu. (2019). Defect detection and classification: Statistical learning approach – Part II. FIQA Research Report 3.

[4] N. Pya Arnqvist, B. Ngendangenzwa, L. Nilsson, E. Lindahl, and J. Yu. (2019). Efficient surface finish defect detection using reduced rank spline smoothers and probabilistic classifiers. Under review for publication in *Econometrics and Statistics*.

[5] N. Fries, X. Liu, S. Bertilsson, U.Andersson, K. Sundberg, E. Lindahl and P. Rydén. (2019) Compiling a dataset for statistical learning in a manufacturing environment. *FIQA Research Report 4.*

[6] X. Liu, N. Fries, E. Lindahl and P. Rydén. (2019) Study on how sanding of the primer affects the risk of crater defects in the topcoat. *FIQA Research Report 5.*

[7] N. Fries, X. Liu, E. Lindahl and P. Rydén. (2019) The effect of buffer times on the paint job quality. *FIQA Research Report 6.*

[8] N. Fries, J. Harbs, A. Kumar Dey, E. Lindahl and P. Rydén. (2019) A framework for deploying an automated alarm and root cause analysis system. *FIQA Research Report 7.*

[9] T. Hastie, R. Tibshirani and J. Friedman. (2009). *Elements of Statistical Learning: Data Mining, Inference and Prediction.* 2nd edition, NY: Springer.

[10] T.K. Ho. (1995). Random decision forests. *Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC*, 14-16.

[11] C. Cortes and V.N. Vapnik. (1995). Support vector networks. *Machine learning*. 20 (3): 273-297.

[12] Y. Freund and R.E. Schapire. (1997). A decision-theoretic generalization of on-lin learning and application to boosting. *Journal of Computer and System Sciences*. 55: 119-139.

[13] C. Seiffert, T.M. Khoshgoftaar and J.V. Hulse (2010). RUSBoost: A hybrid approach to alleviating class imbalance. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 40(1): 185-197.

[14] M. Minsky and S. Papert. (1969). *Percptrons: An Introduction to Computational Geometry*. MIT Press.

[15] D. Ciresan, U. Meier and J. Schmidhuber. (2012). Multi-column deep neural networks for image classification. arXiv preprint arXiv:1202.2745.

[16] D.G. Kleinbaum, L.L. Kupper, K.E. Muller and A. Nizam. (1988). *Applied Regression Analysis and Other Multivariable Methods* (Vol. 601). Belmont, CA: Duxbury Press.

[17] R. Tibshirani. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):.267-288.

[18] S. Ferrari and F. Cribari-Neto. (2004). Beta regression for modelling rates and proportions. *Journal of Applied Statistics*, 31(7):799-815.

[19] J. Svensson and J. Harbs. (2018). Automated alarm and root-cause analysis based on real time high-dimensional process data. Master thesis at the Department of Mathematics and Mathematical Statistics, Umeå University.

[20] E. Conradsson and V. Johansson. (2019). A model-independent methodology for a root cause analysis system. Master thesis at the Department of Mathematics and Mathematical Statistics, Umeå University.

[21] Y. Mi. (2017) Specular Surface Inspection based on Phase Measuring Deflectometry. Master thesis at the Department of Applied Physics and Electronics, Umeå University.
[22] N. Dalal and B. Triggs. (2005). Histograms of oriented gradients for human detection. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05),* volume 1: 886–893.
[23] F. Suard, A. Rakotomamonjy, A. Bensrhair, and A. Broggi. (2006). Pedestrian detection using infrared images and histograms of oriented gradients. *Intelligent Vehicles Symposium*, *2006 IEEE*, 206–212.
[24] T. Mäenpää and M. Pietikäinen. (2005). Texture analysis with local binary patterns. *Handbook of pattern recognition and computer vision*, pages 197–216.
[25] A. Bruce and H.Y. Gao. (1996). *Applied Wavelet Analysis with S-plus*. Springer-Verlag NY, Inc.
[26] S.N.Wood. (2017). *Generalized Additive Models: An Introduction with R*. CRC Press.
[27] B.K. Kwon, J.S. Won, and D.J. Kang. (2015). Fast defect detection for various types of surfaces using random forest with vov features. *International Journal of Precision Engineering and Manufacturing*, 16(5):965–970.
[28] N. Pya. (2017). scam: Shape Constrained Additive Models. R package version 1.2-2.
[29] I. Goodfellow, Y. Bengio, and A. Courville. (2016). *Deep Learning*. MIT Press.
[30] B. Ranneby and J. Yu. (2011). Nonparametric and probabilistic classification using NN-balls with environmental and remote sensing applications. In *Advances in Directional and Linear Statistics*, pages 201–216. Springer.

7 Dissemination and publication

The project has been widely communicated within different fora's within Sweden. Besides the automotive industry, actors like RISE, Umeå municipality, Region Västerbotten and companies working within the field of forestry, agriculture and unions are informed.

The project is also communicated globally at different international scientific conferences and globally within the Volvo group and Volvo Cars global organizations.

7.1 Dissemination of knowledge and results

How should the project result be used and disseminated?	Mark with x	Comment
Increase knowledge in the field	x	FIQA will be used in Volvo Group University's training in AI Awareness.
Passed on to other advanced technological development projects	x	A PhD student will continue working within the WP2 area. The student is co-funded by Umeå University and AB Volvo.
Will be passed on to product development projects		
Introduced in the market		
Used in investigations / regulations / permit cases / political decisions		

The following is a list of the main fora's where FIQA has been presented:

1) J. Yu. Research in Mathematical Statistics at UmU. *Presentations as a part of UmU delegation for Valencia University (Valencia), Polytechnic University of Catalonia (Barcelona) Universitat Jaume I (Castellón), 17-23 February 2017.*
2) P. Rydén and J. Yu. FIQA. *Inauguration of MIT Place, 2 May 2017, Umeå, Sweden*
3) E. Lindahl. FIQA. *Presentation for the Robust Polishing Vinnova project, November 2017, Umeå, Sweden.*
4) E. Lindahl, J. Yu., and P. Rydén, S. ur Réhman, M.S.L. Khan. Invigning av Volvo GTO Pilot Plant. *Presentation för regionala aktörer i Västerbotten, Dec 2017, Umeå, Sweden*
5) S. ur Réhman, Final Inspection and Quality Analysis. *AI4X – Collecting Ideas and Identifying Challenges for Future AI Research in Sweden, February 12, 2018, Stockholm*, Sweden.
6) P. Rydén. FIQA - Automatiserad kvalitetsstyrning med hjälp av storskaliga produktionsdata och avancerad statistisk modellering. *Presentation for Northvolt delegation at Umeå University, 3 April, 2018, Umeå Sweden.*
7) J. Yu. FIQA. *SASUF2018, 14-18 May 2018, Pretoria, South Africa.*
8) K. Sundberg, J. Yu., and P. Rydén. FIQA. *Klusterkonferensen 2018, 23-24 May 2018, Katrineholm, Sweden.*
9) M. Silver. FIQA, *Volvo Cars Manufacturing Engineering: Advanced Engineering Technical Meeting, June 2018 and Jan 2019, Gothenburg, Sweden*
10) E. Lindahl, FIQA, *Volvo GTO global technical management, Oct 2018, Umeå, Sweden.*

11) E. Lindahl, L Marklund, B. Ngendangenzwa, FIQA, *Presentation for IF Metall Volvo, IF Metall Chairman Marie Nilsson and CEO for Volvo Group Martin Lundstedt, Okt 2018, Umeå, Sweden*
12) M. Silver. FIQA, *Volvo Cars Manufacturing Engineering: Paint Core Group Technical Meeting, June 2017 and Oct 2018, Gothenburg, Sweden*
13) N. Pya Arnqvist, B. Ngendangenzwa, L. Nilsson, E. Lindahl and J. Yu. Automated surface finish defect detection using statistical learning approach. *The 6th Swedish Workshop on Data Science - SweDS18, 20-21 November 2018, Umeå, Sweden.*
14) N. Fries, X.J. Liu, and P. Rydén. Statistical learning in a manufacturing environment. *The 6th Swedish Workshop on Data Science - SweDS18, 20-21 November 2018, Umeå, Sweden.*
15) K. Sundberg. FIQA. *Teknikfrukost Uminova Innovation, Dec 2018, Umeå, Sweden*
16) M. Silver. FIQA, *Volvo Cars Manufacturing Engineering: Global Technical and Project Control meeting, Oct 2018, Jan 2019, Gothenburg, Sweden*
17) E. Lindahl. Erfarenheter från att jobba med AI och big data ur ett industriperspektiv *Teknikfrukost Uminova Innovation, Feb 2019 Umeå, Sweden*
18) J. Yu. Statistical modelling for spatiotemporal data. *The Hawassa Math&Stat Conference 2019, 11-15 February 2019, Hawassa, Ethiopia.*
19) N. Pya Arnqvist, B. Ngendangenzwa, L. Nilsson, E. Lindahl and J. Yu. A statistical learning approach for defect detection and classification on specular cab body surfaces. *The 46th Winter Conference in Statistics – Machine Learning, 10-14 March 2019, Hemavan, Sweden.*
20) N. Pya Arnqvist, B. Ngendangenzwa, L. Nilsson, E. Lindahl and J. Yu. Efficient surface finish defect detection using reduced rank spline smoothers. *CRoNoS & MDA 2019, 14-16 April 2019, Limassol, Cyprus*.
21) E. Lindahl, C, Declomesnil, FIQA, *Volvo GTO Top management team for Europe and Brazil Manufacturing, Maj 2019, Marstrand, Sweden.*
22) E. Lindahl, J. Yu., and P. Rydén. FIQA. *Klusterkonferensen 2019, 22-23 May 2019, Katrineholm, Sweden.*

Besides above presentations, the FIQA project is used as a part of the Volvo Group University training for AI awareness and also presented at numerous occasions within Volvo GTO sites in France, Russia, United States and Sweden.

7.2 Publications

Within the project, eight (9) technical reports have been produced during the work, where of 3 scientific reports within WP1 and 4 scientific reports within WP2. The last report is a technical documentation report for the pilot system built in the Volvo GTO facility in Umeå.

23) M.S.L. Khan and S. ur Réhman. (2018). Computer vision approach towards final inspection quality analysis. *FIQA Research Report 1.*
24) N. Pya Arnqvist, B. Ngendangenzwa, L. Nilsson, E. Lindahl, and J. Yu. (2018). Defect detection and classification: Statistical learning approach – Part I. *FIQA Research Report 2*
25) N. Pya Arnqvist, B. Ngendangenzwa, L. Nilsson, E. Lindahl, and J. Yu. (2019). Defect detection and classification: Statistical learning approach – Part II. FIQA Research Report 3.
26) N. Pya Arnqvist, B. Ngendangenzwa, L. Nilsson, E. Lindahl, and J. Yu. (2019). Efficient surface finish defect detection using reduced rank spline smoothers and probabilistic classifiers. Under review for publication in *Econometrics and Statistics.*

27) N. Fries, X. Liu, S. Bertilsson, U. Andersson, K. Sundberg, E. Lindahl, and P. Rydén. (2019). Compiling a dataset for statistical learning in a manufacturing environment. *FIQA Research Report 4.*
28) X. Liu, N. Fries, E. Lindahl, and P. Rydén. (2019). Study on how sanding of the primer affects the risk of crater defects in the topcoat. *FIQA Research Report 5.*
29) N. Fries, X. Liu, E. Lindahl, and P. Rydén. (2019). The effect of buffer times on the paint job quality. *FIQA Research Report 6.*
30) N. Fries, J. Harbs, A.K. Dey, E. Lindahl, and P. Rydén. (2019). A framework for deploying an automated alarm and root cause analysis system. *FIQA Research Report 7.*
31) A.K. Dey, (2019), Pilot system documentation.

Thirteen (13) students have worked in FIQA and completed eight (8) master thesis and one (1) bachelor thesis within Umeå University. These publications are listed below:

32) O. Melander and P, Öhlund (2016). Modelling of surface quality defects related to production process data. Master thesis at the Department of Mathematics and Mathematical Statistics, Umeå University.
33) Y, Mi. (2017) Specular Surface Inspection based on Phase Measuring Deflectometry. Master thesis at the Department of Applied Physics and Electronics, Umeå University.
34) M, Augustian (2017). Neural Network based fault detection on painted surface. Master thesis at the Department of Applied Physics and Electronics, Umeå University.
35) Ngendangenzwa (2017). Specular surface inspection based on phase measuring deflectometry. Master thesis at the Department of Mathematics and Mathematical Statistics, Umeå University.
36) J. Svensson and J. Harbs. (2018). Automated alarm and root-cause analysis based on real time high-dimensional process data. Master thesis at the Department of Mathematics and Mathematical Statistics, Umeå University.
37) J. Rönnlund and J. Sjölund. (2019). A Deep Learning Approach to Detection and Classification of Small Defects on Painted Surfaces. Master thesis at the Department of Mathematics and Mathematical Statistics, Umeå University.
38) E. Conradsson and V. Johansson. (2019). A model-independent methodology for a root cause analysis system. Master thesis at the Department of Mathematics and Mathematical Statistics, Umeå University.
39) W. Xiao. (2019). Data augmentation and using deep learning for specular surface defect estimation. Master thesis at the Department of Applied Physics and Electronics, Umeå University.
40) V. Kröger. (2019). Orange peel, Primer correlation with top coat. Bachelor thesis at Department of Mathematics and Mathematical Statistics, Umeå University.

8. Conclusions and further research

Continuous digitalization and automation of manufacturing processes are necessary in order for industries to stay competitive. The necessary steps involve: generating large amount of high quality manufacturing data, making data easily accessible (e.g. using efficient data bases), modelling data (e.g. using statistical learning and machine learning) and build systems that utilize the models to automatically adapt, improve and optimize the manufacturing processes (e.g. using artificial intelligence). The aim of the FIQA project was to digitalize and automate the quality control and root cause analysis processes at the paint processes at Volvo GTO factory in Umeå. This involved:

- Developing a system for automated detection and classification of defects generating high quality defect data, which included developing a camera system (data generation) and predictive models (modeling).

- Developing a system that connect manufacturing data with defect data (modeling) enabling an automated quality control and root cause analysis (improvement of the quality processes).

- Set up a pilot system in production that incorporated the above systems in daily production.

The FIQA project has delivered the desired pilot system, meeting most of the specifications, and has in addition generated general knowledge, specific improvements and a platform for future research and development as described below:

- Generating data is easy getting access to quality data is hard and demand time and a lot of work. It is clear that Volvo needs to develop novel data solutions that are flexible (e.g. allowing new sensors) and can be used to access large and complex data in real time.

- Raw data are not necessarily of high quality. Data are generally “dirty” and need to be pre-processed, which includes replacing outliers and identifying broken sensors. A lot of time need to spend on understanding, identifying, and constructing relevant variables. Furthermore, the quality outcome determined by the manual inspections at the quality stations have a low resolution with a high variation and are not optimal for modelling.

- Statistical learning and machine learning can be used to predict defects and model the relationship between process variables and quality outcome. For example, the agreement between manually detected defects and defects predicted via statistical learning nearly 100%.

- The collaboration between academia (Umeå University) and industry (Volvo GTO and Volvo cars) has for most times worked smoothly. Some major success factors are: commitment from all parties, support by managements (both at Volvo and the University) and a healthy budget together with a support from Vinnova FFI.

- Working with data is rewarding. As discussed in WP2 the working process resulted in that we identified that sanding generates craters and that spending time in some buffers increases the risk of dust defects. The former finding has resulted in changed practice saving money as well as increasing quality. The novelty of these findings were only partially new, but the systematic work resulted in evidence that enabled actions.

- Although the systems for automated detection and root cause analysis can be further developed, the result shows that the approaches works conceptually. Hence, raw data can be converted, through modeling, to information and effective tools. Importantly, there are no principal obstacles why these approaches wouldn't be possible to apply on a large scale and implemented production. However there are some challenges:

  - The developed camera system needs to be replaced by a large scale commercial system. The FIQA project can be used to determine the specification of this system.

  - All data from the manufacturing process needs to be available and pre-processed in real time.

  - Additional repair data, cleaning data and robot data are data sources that should be included in future.

  - The two approaches should be combined

Our root cause analysis system has so far mainly used manually detected defect data and not the automated defect data generated by WP1. The two main reason for this are that we need automated data from more cabs to fit the root cause analysis models and that the automated system only generates defect data from a relatively small part of the cab.
In a longer perspective there are some very interesting problems that can be addressed. The current system detects problems and identify sub-processes responsible for the decreased quality. We believe that it also would be possible to build a similar system that automatically tune parameters (e.g. various set points), suggest repair actions of machines so that the capacity is maximized.

The purpose today with a manual inspection and notation is to facilitate so that car bodies or cabs can be adjusted. In FIQA, a comparative study has been made of the response data that is currently collected when cabs are manually inspected compared to a dataset that has been more carefully collected to train the classification algorithms. One conclusion is that an automatic detection system is required as a base, and that all defects are identified, not just the defects to be repaired to make a satisfactory prediction. To do a full-scale implementation of alarm systems and root cause systems, the following steps are needed:

- System for detecting and classifying defects that covers the complete vehicle body (either as assisting of a full scale system).

- System logic for managing large amounts of data (Big data) from different sources online.

- Build a flexible system where it's possibility to add more sensor data overtime that might affect quality.

The trend in the market is that many actors work with machine learning and big data. During the FIQA project, the commercial solutions have also been further developed in part 1. Today we can see examples of suppliers who also applied machine learning techniques to detect and classify defects. At present, these systems are relatively unusual in the automotive industry, but the interest is high from many manufacturers to implement solutions. These solutions seems also isolated to one unique process (as different processes has different surface properties). These systems only states where defects occur, has limited classification results, and does not give any information why the defects occurs.

The experience from FIQA gives us understanding that working with big data is a complex task. We believe companies must change their mind sets in order to stay competitive then it comes to use big data. Big data affects all organizations, and collaboration in between these organizations is essential in order to succeed. This is challenging, but it also gives a lot of business opportunities when the academy meets the industry.

9. Contributing parties and contact persons

Volvo Group Truck Operations (GTO)
Eric Lindahl
eric.lindahl@volvo.com
+46 728628939

Umeå University
Jun Yu
jun.yu@umu.se
+46 702490822

Patrik Rydén
patrik.ryden@umu.se
+46 704823555

Volvo Car Corporation
Mattias Silver
mattias.silver@volvocars.com
+46 729774161