Leveraging Unstructured Image Data for Product Quality Improvement

Oliver Nalbach[1], Maximilian Derouet[1], Anil Kumar Erappanakoppal Swamy[1], and Dirk Werth[1]

[1] AWS-Institut für digitale Produkte und Prozesse gGmbH, Saarbrücken, Deutschland
{<first_name>.<last_name>@aws-institut.de}

Abstract. Recently, traditional quality assurance methods, which often require human expertise, have been accompanied by more automated methods that use machine learning technology. These methods offer manufacturers to reduce error rates and, consequently, to increase margins as well. In particular, predictive quality assurance (PreQA) allows to minimize expenses by feeding back information from product returns and quality checks into the early product development. However, PreQA requires detailed information about previous quality problems which is not always readily available in a structured form. In this paper, we therefore discuss the potential of leveraging initially unstructured information in the form of images, taken either during quality checks or by customers when returning a product, to the end of product quality improvement. We furthermore show how this might be realized in practice using the case of fashion manufacturing as an example.

Keywords: quality assurance, image analysis, data science

1 Introduction

Manufacturers strive to avoid product defects as they have the potential to diminish margins and are detrimental to their public image. Traditionally, quality problems have been combatted by quality checks during production to avoid defective products slipping through to the customer and manual analyses of problem causes to avoid the initial bad decisions that might lead to quality problems further down the road [1]. Recently, these prevalent methods, which often require some form of human expertise, have been accompanied by more automated methods that use machine learning technology. In particular, the idea of predictive quality assurance (PreQA) [2] allows to minimize expenses due to product failures by feeding back information from product returns and quality checks into the early product development phase where then data-based predictions are made.

Predictive quality assurance requires as detailed information as possible about previous quality problems. However, quality checks are laborious and, for the case of product returns, while information is provided "for free" by the customer, it is often also very imprecise and coarse. A common situation is that of a customer taking

14th International Conference on Wirtschaftsinformatik,
February 24-27, 2019, Siegen, Germany

photographs of her defective product to point out the respective problems, but not using any kind of systematic annotation. This is the case for online purchases in particular as there is no sales staff to discuss issues with.

In such cases, information is available but not in a structured form so that it cannot be immediately used for machine learning purposes. To retrieve the knowledge hidden in photographs, structured information has to be mined from them. In this paper, we will discuss the potential of leveraging the initially unstructured information in the form of images, taken either during quality checks or by customers, to the end of product quality improvement, and how this might be realized in practice. The latter is demonstrated using the case of fashion manufacturing as an example.

After reviewing previous work in the next section, we summarize the predictive quality assurance approach [2] and discuss the role of unstructured data in that context (Sec. 3). In Sec. 4, we then outline an application scenario in fashion industry and show how a processing pipeline can use state-of-the-art image processing and vision methods to mine detailed information about product defects from photographs of different garments. First results presented in Sec. 5 underline the feasibility of the approach in practice. We conclude with an outlook towards future work on the topic (Sec. 6).

2 Previous Work

Existing literature overlapping with the topics of this paper comes from the domains of quality assurance, machine learning and computer vision.

2.1 Machine Learning for Quality Improvement

Applying machine learning to quality assurance is not a recent idea but has been around for a few decades. Early approaches however used to be rather passive in nature, focusing on the mere prediction of quality-related problems based on features and properties of a product and its production process. Examples include predicting whether a particular software component is error prone based on source code properties [4], or forecasting product quality in injection molding processes [5] and semiconductor manufacturing [6] given the manufacturing parameters. While such predictions can help to avoid producing low quality products they do not automatically lead to the production of higher quality products – for that, optimization approaches are necessary. E.g., given a model predicting the quality of an injection molding process in terms of properties of the resulting product, the process parameters may be tweaked automatically using a genetic algorithm [7]. Such pure optimization approaches have recently been accompanied by assistance systems which do not only allow to tweak existing products but which can already assist their users during product design and which allow the transfer of knowledge to completely new products instead of being tied to a specific process for a specific product. In particular, the predictive quality assurance architecture [2] provides a framework to implement such an assistance system. None of these existing approaches have however demonstrated how unstructured information about product quality problems can be mined and then used for an eventual quality improvement, so far.

2.2 Image Processing and Computer Vision

Research in image processing and computer vision has produced methods which are relevant to this work in two ways. First, there is a history of methods with a similar objective that use image processing to detect defects on objects either during maintenance or quality checks in production [8]. Second, for our proposed processing pipeline (Sec. 4), we transfer algorithms originally designed for different tasks such as people detection and pose estimation to serve a new, additional purpose as sub-steps in the mining of defect information from images.

For identifying defects and failures from images, there are both supervised methods, i.e. methods requiring a labeled data set to be trained on [9-11], and unsupervised methods [12-14], purely based on detecting statistical irregularities. Some aim at generic anomaly detection [14] while others have fixed application domains ranging from wood boards [13] to steel rails [11] to textiles [10]. The vast majority of those methods assumes their input images to be of a very regular nature, only showing a relevant patch of an object's surface texture, not the whole corresponding objects. This requires images to be taken under pre-defined conditions either automatically using calibrated equipment at a production site, e.g., at a fixed distance, angle and under pre-defined lighting conditions, or manually by an instructed expert. Consequently those approaches cannot easily be transferred to the much more irregular photographs taken by customers using varying hardware and showing different views of a full object or even just showing drastically varying complex objects such as different garments.

We post-pone the discussion of related work originally developed in other contexts, which we propose for different processing steps, to Sec. 4 where we will introduce our defect mining methodology.

3 Background

In this section we first summarize the main ideas behind the predictive quality assurance architecture [2] before going on to discuss the role and benefits that image data may provide in this context.

3.1 Predictive Quality Assurance

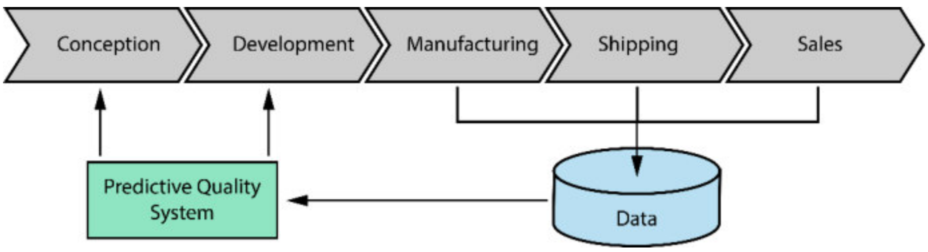

Figure 1. Predictive quality assurance: Data about products and the problems they exhibited previously is collected throughout the depicted process from conception to sales. The predictive quality system analyzes this data to gain insights which are fed back into the early process steps.

Predictive quality assurance (PreQA) uses historical data about manufactured goods and the defects they exhibited to eventually increase the quality of future products. At the core of the method are classification algorithms which are trained to predict the expected occurrence rates for different types of product defects, given a structured description of product features as input. In particular, predictions can also be made given only a subset of all possible features so that the method can already be applied in early stages of product development where not all existing features have been specified yet (Fig. 1). Using these predictions, issues such as bad product designs can be identified early-on. An integrated assistance component then allows to correct adverse decisions made about the product by suggesting alternative options with significantly lower expected probabilities of failure.

3.2 The Value of Image Data for Quality Assurance and Improvement

As PreQA is based on classification algorithms, the availability of a high quality dataset to train these algorithms on is crucial to the method's success. This not only concerns the level of detail found in the specifications of existing products, which should comprise a large set of expressive features, but furthermore especially the descriptions of previous cases of defect.

Using the case of fashion industry as an example, product defects can be of drastically varying nature: holes, stains, fragile seams, broken zippers, lost buttons, bleached colors or strong pilling all are reasons for a customer to return a piece of clothing. But they have very different causes. A bad material mix may affect pilling behavior but arguably has little influence on oil stains caused by a production machine. A weak thread can lead to loose buttons but will not cause problems with a zipper. Thus, if different types of failure are clustered together into a single class of defect, valuable potential for a detailed analysis and a subsequent reliable prediction and failure avoidance is lost irretrievably.

Unfortunately, when returning an item bought from a store due to a defect, customers are usually required to specify their reasons for doing so only in a very coarse way. This is because customers, on one hand, cannot be expected to analyze the specific nature of product failures as they are no domain experts and, on the other hand, because it would take additional time and lead to increased frustration at the customer's end. While the provided coarse defect information can give first insights into problems connected to specific products the imprecision makes it difficult to connect particular defects to individual product features using machine learning. This is reflected in the fact that the prediction confidence of PreQA is usually lower for unspecific defect labels while being higher for more specific ones [2].

Attaching unstructured data in the form of images to product return cases offers a way to mitigate this problem: A picture is worth a thousand words. Using photographs, we can generate detailed descriptions of situations in the fraction of a second. Taking a picture of her defective item takes the customer less effort than filling out a detailed return form by selecting appropriate defect descriptions from dropdown lists but still encodes virtually all information about the occurred problem into a compact visual description. The same applies to quality checks of final products performed in companies. While in this case the staff is specialized and experienced in assessing defects, just taking a photograph could significantly decrease the time required for checks which would allow a denser sampling of the full production yield given the same time budget and increase the quality of the items sold by itself.

3.3 Manual Visual Inspections in Fashion Companies

An automated analysis of images showing defective products should ideally produce output which is identical to the one expected from a human expert assessing the case manually. We will thus briefly describe the way visual inspections proceed for the exemplary case we are dealing with in this paper, the fashion industry.

While there are some industrial norms for quality checks in the form of ISO or DIN standards in clothing industry they mainly concern apparel which is used in safety-critical contexts. When it comes to quality checks of casual clothing, each company defines its own inspection protocols to follow. Nevertheless, there are certain best practices performed in most companies.

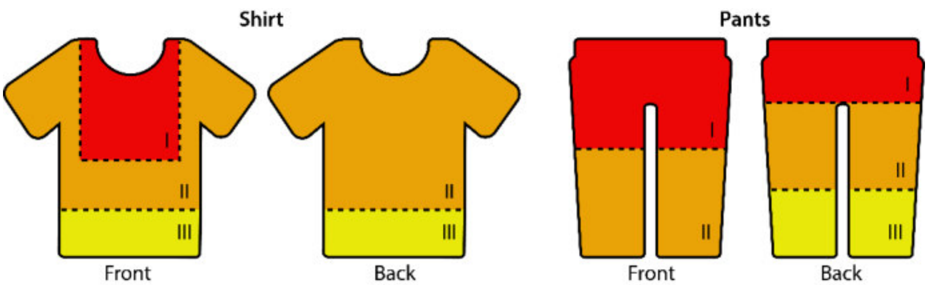

Figure 2. Typical defect zones used in visual garment inspections, according to [19]

First, a list of possible defects is defined. These potential quality flaws are then considered one by one during visual inspections. Second, garments are commonly subdivided into different zones called e.g., zone A to C or I to III, as depicted in Fig. 2.

These zones often correspond to different levels of visibility, and consequently also severity, of defects. Usually, the highest severity is assigned to parts visible when presenting the product in a store (Fig. 2, zones I), the second highest to areas more visible when wearing it (Fig. 2, zones II) and so on. Third, each company defines combinations of defect types and defect locations, in terms of zones, which lead to the rejection of a piece of clothing or its re-use in outlet stores, respectively.

4 A Processing Pipeline for Mining Textile Defects from Images

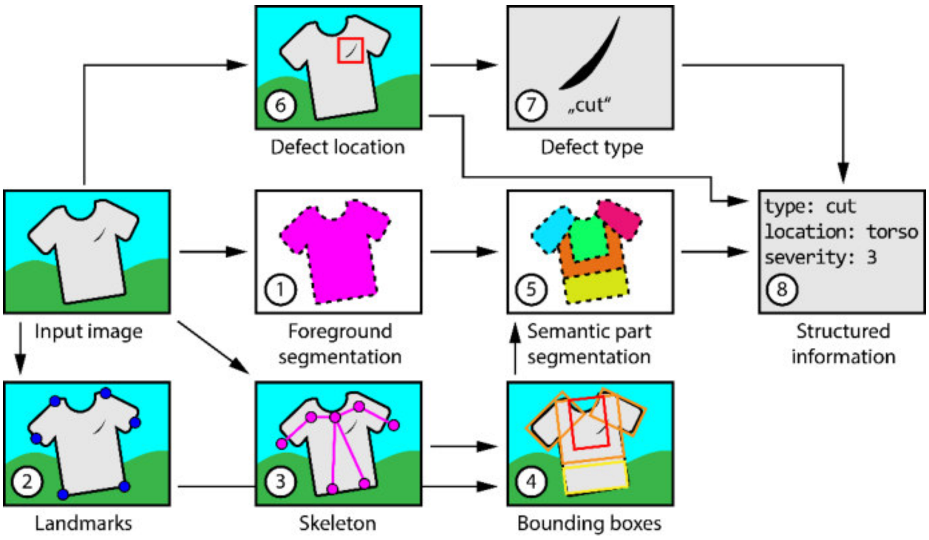

Figure 3. Image processing pipeline to mine details about product defects from images for the case of fashion industry. Steps are labelled by the order in which they are described.

We will now outline how computer vision techniques can be used to compose a processing pipeline to mine textile defects from photographs of pieces of clothing. An overview is shown in Fig. 3. It is inspired by the best practice approaches used in companies (Sec. 3.3) on one hand and the requirements set by machine learning algorithms on the other hand. The input to the pipeline is a photograph of a defective garment. No additional annotation is required. The output of the pipeline consists of a defect description regarding three properties: the type of the defect, its location with respect to the garment and a severity rating depending on the type and location which can be used to decide how to proceed further with the specific item. The pipeline consists of seven processing steps. For each of these steps, we will describe which output they compute, why they are necessary and which existing methods can be used to implement them.

4.1 Foreground Segmentation

Customers and customer care staff may take photographs of defective products in front of varying backgrounds. The first processing step therefore consists of segmenting the actual object of interest from that background (Fig. 3, step 1). In technical terms, each pixel of the original image is labeled as either belonging to the piece of clothing or to the background. This step is necessary for computing a more detailed segmentation of the object into different zones later on (Fig. 3, step 5) and also provides the precise areas of the image where defects can be expected to be localized later on (Fig. 3, step 6). There are several approaches which deal with this problem in different contexts, for example from the domains of image retrieval [15] or semantic scene description [16]. Some of them work best on photographs in which the clothing is worn by a person as they are based on detecting human features in a pre-processing step [15] while others are solely based on segmentation, such as the method by Borras et al. [16], and could therefore be applied in a more general setting.

4.2 Landmark Detection

In the second step, different semantic landmarks are detected (Fig. 3, step 2). Those are important image locations corresponding, e.g., to the end points of sleeves or of the collar. The output of this computation step is a list of detected landmarks – which may also correspond to only a subset of all known types of landmarks, depending on the type of garment and the specific view – and of their corresponding positions with respect to the image. The landmarks allow for a subsequent computation of 2D bounding geometry that then serves as guidance for the full semantic segmentation of the object (Fig. 3, step 5). Fashion landmark detection constitutes a very new area of research and has been introduced only recently by Liu et al. [17]. Their method employs deep neural networks and is robust regarding whether an image contains a person or not and whether the full item is depicted or only part of it (cf. Fig 4, second column).

4.3 Skeleton Fitting

In addition to the landmarks computed in step 2 which correspond to different extremal points of the pieces of clothing, further important key points can be derived using so-called pose estimation algorithms which try to fit an approximate skeleton to person depicted in an image (Fig. 3, step 3). These skeletons consist of a number of detected joints, connected by lines corresponding to different limbs (Fig. 4, third column). While those methods are originally used to determine the poses of humans in photographs or videos they often work surprisingly well for images without people in them (cf. Fig. 4, row 1) as the underlying neural networks have learned to use clothing as an important feature. Specifically, we employ OpenPose [20] to gather the additional pose information when possible. This framework outputs a number of human joint locations with respect to the input image. When particular joints cannot be detected reliably, only the subset which is detected confidently is output.

4.4 Fitting of Bounding Boxes

In the fourth processing step, the shape of the garment is approximated coarsely by a collection of two-dimensional boxes (Fig. 3, step 4). Fitting such so-called geometric proxies greatly simplifies the definition of rules to use by the semantic segmentation step following next. For example, in the case of an upper-body garment such as a shirt or jacket we first compute a box covering the torso region by fitting it to contain the four landmarks corresponding to the collar and the bottom of the torso. Boxes corresponding to the different defect zones within the torso region (Fig. 2) can then be derived from it: The zone III region typically corresponds to the lower 20% of the torso region while zone I covers 50% of the torso height and 70% of its width as it corresponds to the portion of the garment visible when folded and presented in a shop. To determine the height of the sleeves' bounding boxes we use the sleeve-ends landmarks in combination with the detected shoulder joints from step 3, if applicable, and fall back to using the collar landmarks from step 2 otherwise. The width of the sleeve boxes is derived from the horizontal distance between the collar landmarks. Results for this step are shown in the fourth column of Fig. 4.

4.5 Semantic Segmentation

Given the coarse shape approximation from step 4, as well as a mask segmenting the item of clothing from the background (step 1), the derivation of a complete segmentation (Fig. 3, step 5) can be defined using a set of processing rules: Each pixel is assigned to the semantic part corresponding to the closest box. In case a pixel is contained in multiple boxes, precedence is given to boxes corresponding to higher severity and main boxes are preferred over boxes corresponding to sleeves or legs. The resulting pixel-wise labelling is used to derive the location of the defect can then be used to rate its severity. While there are previous methods to solve this segmentation problem sometimes also referred to as clothes parsing [18] many of them strictly require the garments to be worn by a person [15, 18] and none take the typical zone assignment used in textile quality checks into account.

4.6 Anomaly Detection

Before we can classify textile defects into different categories, it is useful to first determine the image region where the defect is located (Fig. 3, step 6). The classification following in step 7 can then be applied to a cropped image showing only that region which significantly simplifies the problem as the cropped images to be classified are more regular. Assuming that most items of clothing are largely homogeneous regarding the fabric they are made of, we can employ unsupervised learning methods [12-14] which detect statistical irregularities and output a bounding box around the anomaly which can be used to determine the region to crop. By limiting the detection to the garment, using the segmentation from step 1, we can avoid irregularities being detected outside of the actual product.

4.7 Defect Classification

In the final image processing step (Fig. 3, step 7) we have to classify defects on an otherwise homogeneous image showing a textile patch. This problem has been previously tackled to automate visual quality checks of (raw) fabrics which constitute the feedstock for apparel manufacturing. There even is a standard benchmark, the TILDA dataset [21], which contains several hundred images of fabric patches with different defects and has been commonly used to assess different classification methods [22-24]. Those conventional methods are based on a two-step approach where first statistical features are computed in order to reduce the dimension of the input data (a patch of just 512 by 512 pixels with 3 color channels corresponds to a vector of 786,432 components) and then a classification only based on these features is performed.

However, recently, deep artificial neural network structures based on convolutional layers as the main computational units, called (deep) Convolutional Neural Networks (CNNs), have become very popular for various image processing and computer vision tasks [25]. In particular for classification tasks such as the annual ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) [27] they have outperformed more traditional methods by a wide margin. We therefore compare their performance for the task of textile defect classification to that of existing methods [22-24] in Sec. 5.2.

4.8 Data Fusion

After we have gathered information about the different possible defect zones with respect to the piece of clothing (Step 5), the location of the defect (Step 6) and its type (Step 7), we can output corresponding structured information (Fig. 3, Step 8) which can then be used to train more precise classifiers for PreQA.

5 Results and Discussion

In this section we present initial results for several steps of our processing pipeline, specifically, the computation of bounding boxes and the final classification step.

5.1 Clothes Segmentation

Using the publicly available implementations of OpenPose [20] as well as of the fashion landmark detection of Liu et al. [17] we implemented steps 2-4 of our pipeline. Fig. 4 demonstrates the output of these steps for the case of four upper-body garments. Results are shown for both images with (Fig. 4, rows 2+3) and without humans (rows 1+4) and for the cases of both successful (rows 1+2) and unsuccessful (rows 3+4) pose detection.

Figure 4. Pipeline input (first column) and outputs of processing steps 2 to 4 (second to fourth column) of our pipeline (Fig. 3). For rows three and four no full skeleton estimation was possible and the fallback approach (Sec 4.4) was used. The sample images have been taken from the fashion landmark detection benchmark dataset [17].

Table 1. Metrics achieved on the test set. As we deal with a multiclass classification problem the values have been averaged across the (evenly balanced) classes.

Metric	*Value*
Accuracy	90.1 %
Precision	87.8 %
Recall	87.6 %

5.2 Textile Defect Classification using Deep Learning

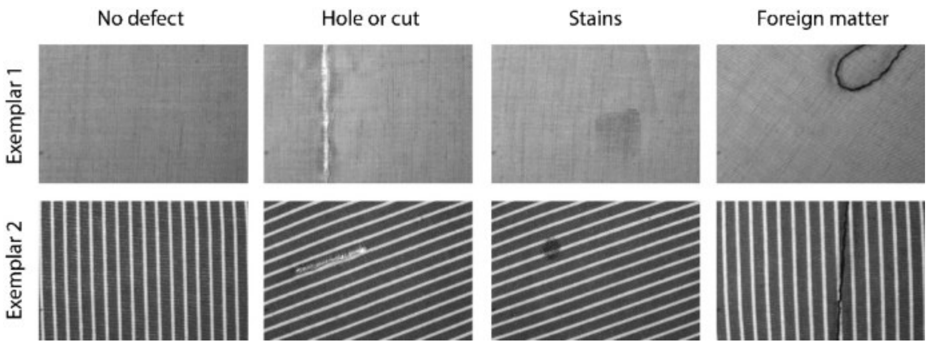

Figure 5. Sample images taken from the TILDA dataset [21] for four of the eight available classes (columns) and two of the included representatives (rows).

To evaluate the use of deep learning for textile defect classification, compared several popular network architectures from the ILSVRC contest [27], namely AlexNet [31], SqueezeNet [30] and GoogleNet [28]. We randomly split the TILDA dataset (cf. Fig. 5 for samples) into 90% of training and 10% of test images. The original balanced ratio of the classes (each class contributes 12.5%) was maintained. As TILDA only comprises about 3200 images, which is several orders of magnitude less than for most common deep learning datasets, we augmented our training data by adding mirrored and rotated copies of the original images. While TILDA contains high resolution images in a 4:3 format, all compared network architectures expect quadratic input images of lower resolution. We therefore padded images with white and then down-scaled them to the appropriate input resolution for each network. We first trained the candidate networks using the Adam solver [29] until their classification accuracy converged. As GoogleNet reached the highest initial accuracy, we selected it for further finetuning and increased the original drop-out rate to counteract potential overfitting to the small dataset, achieving an accuracy of 90% on the test set. A detailed evaluation using different common metrics is given in Table 1. This is on-par with values reported for competing methods [22] for which however no details, about whether a separate test set was used and which size it had, are given.

5.3 Discussion

The defect mining pipeline of Sec. 4 comprises eight steps, all of which constitute challenging problems by themselves. In the following, we discuss the current limitations of the most important steps and how these might be overcome in the future.

Fashion Landmark Detection While the method by Liu et al. applied by us for the problem of fashion landmark detection already achieves detection rates of about 70% when accepting only small misplacements of landmarks [17] this particular detection problem is a fairly new area of research. Improved, more robust methods can be expected to appear in the near future. This is also crucial for our defect mining approach as we use fashion landmarks as main guidance when constructing bounding geometry for different clothing zones.

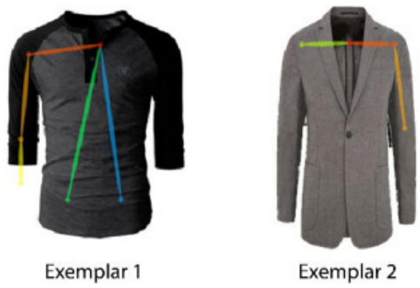

Figure 6. Two cases in which skeleton fitting failed for clothing-only images. The images shown are part of the fashion landmark detection benchmark dataset [17].

Skeleton Fitting We use skeleton fitting with OpenPose [20] as an additional input to help improve the subdivision of garments. While skeletons can sometimes also be derived for images where an item of clothing is not worn by a human (Fig. 4, first row), this cannot be expected to work in all cases. Fig. 6 shows two examples where skeleton fitting fails partially. Currently, we completely ignore the pose information in these cases. In future work, it might be useful to still extract available information for those parts of the skeleton that could be fit with significant confidence. For instance, for the left exemplar in Fig. 6, all joints but the ones corresponding to the left arm have been placed correctly and could have contributed information for fitting bounding boxes.

Bounding Boxes At the moment, we use simple boxes as the geometric proxies for different zones on pieces of clothing. This is sufficient in many cases but there are cases where a box provides a poor fit to the actual shape of the object, e.g., in cases where the sleeves of a shirt are not shown straight but bent. In such situations the fitting could be improved either by using multiple boxes, one for the upper and one for the lower part of a sleeve, or by replacing the boxes by more complex shapes such as two-dimensional generalized cylinders [26].

False Positive Anomalies We propose to use anomaly detection to spot locations of defects with respect to pieces of clothing. This assumes that their material is mostly homogeneous. While advanced detection methods also work for textured image regions, modern fashion sometimes uses defects such as holes or abrasion as design elements. Even for humans it is not always possible to decide with certainty whether a particular defect has been created voluntarily or not, unless additional exemplars of the same product are available to compare to. Consequently, it is debatable whether artificial intelligence can be expected to perform this kind of distinction more reliably.

6 Conclusion and Future Work

In this paper, we have discussed the importance of detailed training data when applying machine learning for quality improvement of manufactured goods by means of predictive quality assurance and demonstrated how image data can become a significant part in it. The latter was achieved by presenting the first pipeline to mine detailed structured defect information from images of defective products. While this pipeline is currently still of prototypical nature, preliminary results for several of the computational steps show its potential. Future work will see a full implementation of the described approach as well as a transfer to additional application scenarios in different industries, such as furniture manufacturing, where analogous methods to realize the individual processing steps, e.g., segmentation or anomaly detection, already exist within different contexts in computer vision research.

References

1. Morlock, S.: Qualitätsgesicherte Produktentwicklung in der Bekleidungsindustrie, https://www.hohenstein.de/de/inline/pressrelease_45248.xhtml?excludeId=45248 (Accessed: 30.08.2018)
2. Nalbach, O., Linn, C., Derouet, M., Werth, D.: Predictive Quality: Towards a New Understanding of Quality Assurance Using Machine Learning Tools. In: 21st International Conference on Business Information Systems. Springer, Cham. (2018)
3. Ng, H.F.: Automatic thresholding for defect detection. Pattern recognition letters, vol. 27(14) (2006)
4. Lanubile, F., Visaggio, G.: Evaluating predictive quality models derived from software measures: lessons learned. Journal of Systems and Software, vol. 38(3) (1997)
5. Ali, I.G., Chen, Y.-T.: Design quality and robustness with neural networks. IEEE Transactions on Neural Networks, vol. 10(6) (1999)
6. Chen, W.-C., Lee, A.H.I., Deng, W.-J., Liu, K.-Y.: The implementation of neural network for semiconductor PECVD process. Expert Systems with Applications, vol. 32(4) (2007)
7. Shen, C., Wang, L., Li, Q.: Optimization of injection molding process parameters using combination of artificial neural network and genetic algorithm method. Journal of Materials Processing Technology, vol. 183(2-3) (2007)
8. Xie, X.: A review of recent advances in surface defect detection using texture analysis techniques. ELCVIA Electronic Letters on Computer Vision and Image Analysis, vol. 7(3) (2008)
9. Wiltschi, K., Pinz, A., Lindeberg, T.: An automatic assessment scheme for steel quality inspection. Machine Vision and Applications, vol. 12(3) (2000)
10. Kumar, A. Neural network based detection of local textile defects. Pattern Recognition, vol. 36(7) (2003)
11. Mandriota, C., Nitti, M., Ancona, N., Stella, E., Distante, A.: Filter-based feature selection for rail defect detection. Machine Vision and Applications, vol. 15(4) (2004)
12. Escofet, J., Navarro, R. F., Garcia-Verela, M. S. M., Pladellorens, J. M.: Detection of local defects in textile webs using Gabor filters. Optical Engineering, vol. 37(8) (1998)
13. Kim, C. W., Koivo, A. J.: Hierarchical classification of surface defects on dusty wood boards. In: Pattern Recognition Letters, vol. 15(7), pp. 713–721. (1994)

14. Tsa, D. M., Wu, S. K.: Automated surface inspection using Gabor filters. In: The International Journal of Advanced Manufacturing Technology, vol. 16(7) (2000)
15. Yamaguchi, K., Kiapour, M. H., Ortiz, L. E., Berg, T. L.: Parsing clothing in fashion photographs. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3570–3577. IEEE (2012)
16. Borras, A., Tous, F., Lladós, J., Vanrell, M.: High-level clothes description based on colour-texture and structural features. In: Iberian Conference on Pattern Recognition and Image Analysis, pp. 108–116. Springer (2003)
17. Liu, Z., Yan, S., Luo, P., Wang, X., Tang, X.: Fashion landmark detection in the wild. In: European Conference on Computer Vision, pp. 229–245. Springer (2016)
18. Yamaguchi, K., Kiapour, M. H., Ortiz, L. E., Berg, T. L.: Parsing clothing in fashion photographs. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3570–3577. IEEE (2012)
19. Fashion2Apparel, http://fashion2apparel.blogspot.com/2017/01/defects-zone-apparel-inspection.html (Accessed: 29.08.2018)
20. Cao, Z., Simon, T., Wei, S. E., Sheikh, Y.: Realtime multi-person 2d pose estimation using part affinity fields. arXiv preprint (2016)
21. Technische Universität Hamburg-Harburg. Ein Referenzdatensatz zur Evaluierung von Sichtprüfungsverfahren für Textiloberflächen. Technical Report (1996)
22. Murino, V., Bicego, M., Rossi, I. A.: Statistical classification of raw textile defects. In: Proceedings of the 17th International Conference on Pattern Recognition, vol. 4, pp. 311–314. IEEE (2004)
23. Schael, M.: Texture defect detection using invariant textural features. In: Joint Pattern Recognition Symposium, pp. 17–24. Springer, Berlin (2001)
24. Sezer, O. G., Erçil, A., Ertuzun, A.: Using perceptual relation of regularity and anisotropy in the texture with independent component model for defect detection. In: Pattern Recognition, 40(1), pp. 121–133 (2007)
25. Krizhevsky, A., Sutskever, I., Hinton, G. E.: Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems, pp. 1097–1105 (2012).
26. Fang, X. S.: The extruded generalized cylinder: A deformable model for object recovery. In: 1994 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 174-181. IEEE (1994)
27. Russakovsky, O., Deng J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg A. C., Fei-Fei, L.: ImageNet Large Scale Visual Recognition Challenge. IJCV (2015)
28. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Rabinovich, A.: Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1–9 (2015)
29. Kingma, Diederik P., and Jimmy Ba.: Adam: A method for stochastic optimization. arXiv preprint (2014)
30. Iandola, F. N., Han, S., Moskewicz, M. W., Ashraf, K., Dally, W. J., Keutzer, K. : Squeezenet: Alexnet-level accuracy with 50x fewer parameters and< 0.5 mb model size. arXiv preprint (2016)
31. Krizhevsky, A., Sutskever, I., Hinton, G. E.: Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems. 1097-1105. (2012)