

A Prestudy of Machine Learning in Industrial Quality Control Pipelines

Jože Ravničan¹, Anže Marinko², Gjorgji Noveski², Stefan Kalabakov², Marko Jovanović³, Samo Gazvoda⁴ and Matjaž Gams¹

E-mail: joze.ravnican@unior.com, anze.marinko@ijs.si, gjorgji.noveski@ijs.si, stefan.kalabakov@ijs.si, marko.jovanovic@smm.si, samo.gazvoda@gorenje.com, matjaz.gams@ijs.si

¹UNIOR Kovaška industrija d.d., Zreče, Slovenia

²Department of Intelligent Systems, Jožef Stefan Institute, Ljubljana, Slovenia

³SMM proizvodni sistemi d.o.o., Maribor, Slovenija

⁴Cooking Appliances Division Gorenje Group, Velenje, Slovenia

Keywords: machine learning, manufacturing, quality control, home appliances, car industry

Received: January 24, 2022

Today's fast paced industrial production requires automation at multiple steps during its process. Involving humans during the quality control inspection provides high degree of confidence that the end products are with the best quality. Workers involved in the control process may have an impact on production capacity by lowering the throughput, depending on the complexity of the control process at the time the control is carried out, during the process which is a time-critical operation, or after the process is completed. Companies are striving to fully automate their quality control stages of production and it comes naturally to focus on using various machine learning methods to help build a quality control pipeline which will offer high throughput and high degree of quality. In this paper we give an overview of applying several machine learning approaches in order to achieve an autonomous quality control pipeline. The applications for these approaches were used to help improve the quality control pipeline of two of the biggest manufacturing companies in Slovenia. One of the most challenging part of the study was that the tests had to be performed only on a small number of defective products, as is in reality. The motivation was to test several methods to find the most promising one for later actual application.

Povzetek: Z nekaj prototipi je bila narejena analiza možnosti uporabe umetne inteligence v nekaj velikih slovenskih podjetjih.

1 Introduction

In the world of machinery, products are expected to be available in high quantities as well with high quality. To make this possible, many companies around the world introduce an efficient production process which is streamlined as much as possible. To achieve this, many go to such lengths by having a zero defects policy. This policy suggests that product defects, while unavoidable, should be kept to the lowest percentage possible. The reason for this is to increase profits by lowering the cost of failure. Defects can occur at any stage during the production pipeline. It is important to catch these product defects as early as possible so they don't have a chance to propagate to later stages causing even more damage. Some approaches, such as [4], reduce defect propagation by shifting quality control at earlier stages of the production process.

Quality control is a process that is present throughout all aspects of production, ensuring standards and quality criteria are met before products leave the assembly line. In order for this process to be as efficient as possible, it needs to be integrated into the manufacturing process. Depending on the products manufactured, quality control might need a separate physical area devoted to it in order to perform mul-

tle analyses, although this segregation of tasks in a production environment is often cumbersome and error prone. Traditionally, the quality control process requires specialized labor who are skilled in correctly identifying mistakes or inconsistencies in the developed product. The downside of using such a specialized labor force is that humans are prone to error [9]. Additionally, repetitive tasks are not suited for one person to perform them for a long period of time with consistent quality. One solution is to use computer aid in order to achieve a higher quality of the overall processes.

Implementing a machine learning quality control pipeline is a challenging task. Difficulties arise from the specifics of the industry, requiring expert knowledge to be used to develop a system which is capable of satisfying the quality needs while also being fast. In [20] a comprehensive overview regarding the scope of those challenges is given and one could see the plethora of categories that machine learning has been applied to.

Looking at machine learning as an enabler to industrial processes, [15] provide with an overview of the effort put into machine learning in manufacturing. Most noticeably are the uses of neural networks and support vector machines (SVMs) since they are generically applicable and

provide competitive results with a low investment time. Regarding industrial maintenance, [18] provides an overview on current and future trends which are shaping companies towards Industry 4.0. Prevalent in these efforts is the inter-connectivity of different sensors and systems, namely internet-of-things (IoT).

In this paper we provide a technical overview of various applications of machine learning into a manufacturing environment. The solutions presented here are provided for industrial products each having their own specific purpose. Due to the different types of products, we will test which methods are more suited than others. We also show how making use of transfer learning in the image recognition domain can provide encouraging results on domains that the model was not trained on.

The products that were used for our experiments included **metallic forks**, **metallic surfaces**, **oven fans** and **broaches**. These are included in the following sections and the data is structured as follows. Section 2 presents similar papers which focus of quality control in manufacturing environments. In Section 3 we present our solutions for quality control on metallic forks from car parts. Regarding home appliances, applicable solutions are presented in Section 4 and Section 5 where we examine oven parts. Section 6 examines broaches in metalworking to determine the amount of wear. Discussion is provided in section 7. Lastly a conclusion is presented in Section 8.

2 Related work

Many applications of machine learning proved successful in a wide set of industries, ranging from automotive to pharmaceutical. Each industry specializes on its problem domain to achieve a great end product or a service.

Looking at the field of pharmaceuticals, high-throughput screening (HTS) is a method to quickly test a large number of chemicals and/or compounds. Using robotics, data processing, sensitive detectors and more, this allows a researcher to conduct millions of chemical, genetic and pharmacological tests. Because of the necessity to conduct a large amount of tests, machine learning can be used here to speed up the process. In [10], a large problem of the process represents data mining and data storage of the images acquired during testing. There exists a trade-off between quality and high throughput. More time is needed to get results with higher quality though that lowers the overall throughput. Often, image analysis is a bottleneck in this process, meaning good and fast machine learning methods for visual tasks are paramount.

An interesting implementation of HTS is presented in [2], where a whole quality control pipeline is created using open source software. Their goal is eliminating images which deviate from the standard quality baseline. Their tool of choice can implement multiple image processing modules all of which are configurable. The modules implement machine learning models which achieve good results.

In manufacturing environments, where products need to conform to high standards, [5] use a knowledge-based intelligent supervisory system to find events of rare quality. These rare quality events can be considered as finding defects in around one million opportunities. By using several methods, such as logistic regression and feature elimination, they have achieved great results, showcasing what kind of exceptional quality machine learning can bring.

In the automotive industry, timely detection of car body defects is important since a bad end product can cost tens of thousands of dollars. To ensure good car body dimensions, [12] use XGBoost [3], random forest, and other methods to detect dimensional defects. From their production line we can see that they carry out quality control in two stages while the stages that are in-between are human operated.

Integrating machine learning solutions into industrial processes needs careful selection of best performing models and parameters. The deployed models have to provide a prediction in less time than a unit needs to be manufactured. Because of this, not all production processes can utilize machine learning easily. In [14], an integrated quality inspection solution is presented which uses the power of machine learning and edge computing. Edge computing was used to provide faster model training and evaluation in order to find the one that gave the best results. They work on improving the quality inspection in electronics manufacturing which uses surface mount technology (SMT) assembly. Printed circuit boards (PCBs) are assembled and go through a solder paste inspection (SPI) phase, which is the phase that the predictive models are applied. They used several models such as Naïve Bayes, Decision Tree, Logistic Regression, SVM and Gradient Boosted Tree. Since in their use case the manufacturing plant was outputting a device every single second, the model needs to be trained and predict in less time than that. All of the models were capable of achieving this except Gradient Boosted Tree where the training time took 2 second. It was shown that this approach could improve the product throughput by lessening the need for units to go to additional quality checks.

Related to food, [1], explore the usefulness of using machine vision in food packaging. Specifically, they analyze the seals of thermoforming packaging using three convolutional neural networks (CNNs): ResNet [6], VGG [16], DenseNet [8]. These tasks are still done by human operators which presents a possibility to greatly automate and improve the production line. Line workers were observed to see which packages were accepted and which were rejected. With that knowledge, a system was built to classify packages with improper seals that could cause food to spoil. Five different datasets were created depending on the type of product that was produced using an inline image acquisition system. Their solution detects 99.93% of defects in production. It is worth mentioning that this was tested in both laboratory conditions and also in a real scenario.

Efforts are made into automating quality control processes in order to be closer to Industry 4.0. In [11], they mention several manufacturing paradigms with the main

focus being on Holonic Shop Floor Control Systems. To aid in bridging the flow of information between a management level and a manufacturing process level, several CNNs were trained in order to find flaws in cast products, namely impellers of submersible pumps. In order to obtain more samples for training, the authors used several image augmentation techniques such as: different axis rotations, zooming in or out and shear mapping. The best performing CNN had a total of 7 layers and achieved precision of 99.82%.

In order to get a better representation of the efforts which are put into building quality control system, a summary of related work, their description and results is shown in Table 1.

3 Metallic forks

In this section, we describe our quality control solutions which were tested in different manufacturing environments. Our goal is to utilize machine learning accompanied by preprocessing techniques in order to classify if a given industrial product contains a defect or not. The solutions created must be easy to implement and have little interference during the manufacturing of products. To detect defects on metal forks we used three approaches: vibration analysis, 2D and 3D machine vision analysis.

3.1 Vibrations

Looking at utilizing vibrations, we measured the signal of displacement of the rack which holds and shakes the product and measured the signal of displacement for the product. To measure the displacement we used a laser distance meter with high accuracy. In the end we observed the relationship between the two displacement signals. Before going deeper with this technique, we did a basic validity test in which we glued a small piece of insulating tape on the fork. Results showed that the displacement signals had very high deviation from the baseline which gives hope that real defects will also be noticeable on the displacement signals. Then we tested the real products by comparing ones with and without an imperfection. Each output signal was 10 seconds long which was obtained from using a total of 24 products. Four different vibration settings were used which were:

1. Amplitude: 0,389 Vpp; frequency: 50Hz
2. Amplitude: 0,389 Vpp; frequency: 60Hz
3. Amplitude: 0,2026 Vpp; frequency: 60Hz
4. Amplitude: 0,2026 Vpp; frequency 50Hz

We split each 10 second measurement into 10 pieces of 1 second long data points. On each of these pieces, 22 features were expertly selected based on the characteristics

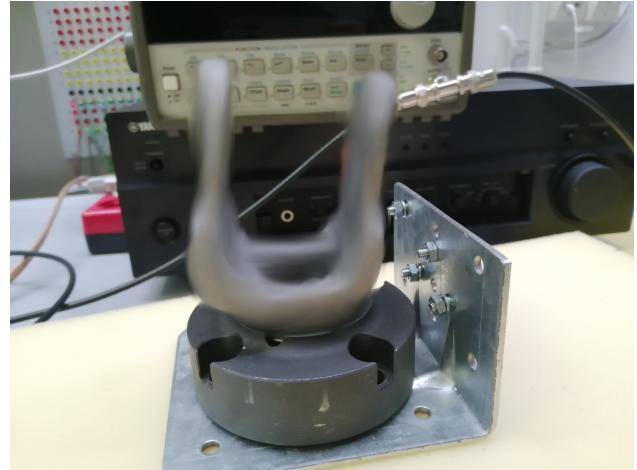


Figure 1: Vibration analysis.

from the time and frequency spectrum. On the dataset 5-fold cross validation was done. For the vibration analysis a convolutional neural network was also used, having 2 hidden layers and an input of 256 neurons. The input was the whole 10 second measurement of the signal. The results obtained from the CNN were: Accuracy 48%, Recall 42%, Precision 91%. There was no point going further with this approach, so other methods had to be tested.

3.2 2D machine vision analysis

The major problem with machine learning in our case was lack of faulty products. For the machine vision approach we had 9 defective products in total from which 46 images were captured. Each of the images was manually annotated with a region of interest which encompasses the area with an imperfection. This was done so we could have an annotated dataset we can use to train our deep neural network. This dataset was further divided into a training and testing dataset.

In order to perform image analysis we used a pre-trained convolutional neural network, Faster RCNN [13]. This network consists of several convolutional layers and at the end several fully connected layers. The convolutional layers are good at extracting features from the image such as shapes and edges, whereas the fully connected layers learn to classify these features for the task at hand. The whole network consists of two parts, a segmentation part and a classification part, both of which were used to detect defects on the surfaces. The first part, namely segmentation, was used in order to extract specific regions of interest in the picture which in our case was regions with defects. These regions are then fed into the second part of the neural network whose task is to classify if a certain region contains an imperfection or not. The pre-trained weights were used from the Faster RCNN network and only the fully connected layers were trained on our training set, effectively freezing the convolutional layers so we can utilize the features learned in previous object detection tasks.

Paper	Description	Results
[5]	Using feature elimination and classification threshold search to find best models for classifying good or bad welds	Logistic regression model - maximum probability of correct decision (MPCD): 1
[12]	Testing several machine learning models to predict dimensional defects in cars	Both XGBoost and Random Forest - ROC AUC: 0.97
[14]	Training machine learning models to detect defective units during PCB manufacturing	Gradient boosted tree - precision: 93.1%, recall: 89.9%, accuracy: 92.6%
[1]	Proper seal detection in thermoforming food packaging	ResNet18 - accuracy: 95%
[11]	Convolutional neural networks in detecting faults in cast metal products	Custom CNN - accuracy: 99.82%

Table 1: Overview of related work.

During training we used 2900 epochs on the segmentation network and 200 epochs on the classification network. In Figure 2 we can see a schematic view of the metallic fork that was used in our experiments.

The test set used the segmentation network from the Faster RCNN which contained 21 images with an imperfection, 7 without, are presented in Table 2, whereas the results on the classification network on the testing dataset which had 18 images with an imperfection, 1 without are presented in the Table 3.

TP	FP	TN	FN	Accuracy	Recall	Precision
17	4	3	4	71.4%	81%	81%

Table 2: Results from the testing dataset on the segmentation network.

TP	FP	TN	FN	Accuracy	Recall	Precision
1	0	1	17	10%	5%	100%

Table 3: Results from the testing dataset on the classification network.

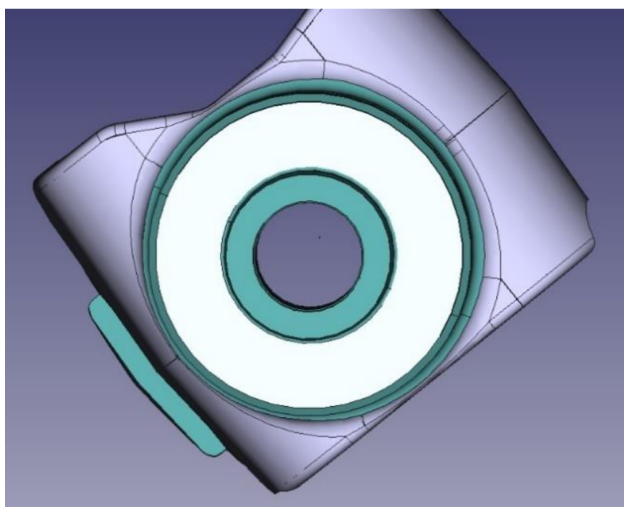


Figure 2: Schematic view of the metallic fork.

The small number of learning examples was insufficient for a hard real-life problem, as expected.

3.3 3D machine vision analysis

The last technique for detecting imperfections on metallic surfaces that was tested was through using a 3D hardware system. This is the most expensive option from all the others but is the one that was expected to deliver the most promising results. We scanned a total of 8 items, 4 of which contained a scratch on their surface and 4 without. The hardware system in question is 3D camera "Ranger 3" manufactured by SICK with laser light source of 660 nm (red laser). It is a line scanner which means it scans objects not at once, but line by line. During scanning, either the object or the scanner might be moved in order to get a complete scan. In our case, the scanner was mounted on a fixed point and only the object was moved along one axis. Figure 3 showcases the setup used to scan objects, in this example a saw tooth blade. The point clouds obtained from the system are 3D representations of the scanned object. Each point in the point cloud has a specific color value associated with it, which represent the reflectance value of that point. The scratches in the point cloud exhibited lower number of 3D points at their location, and additionally a change in reflectance values. The observed scratches varied by length and width, some of them even forming curved lines. Concerning width, a single scratch might also have varying width along its whole length. In reality, the width of the scratches varies between 0.15mm to 2mm.

In order to have a bigger dataset to work with, new samples were generated by adding artificial imperfections on samples or adding some noise. Random noise was introduced by moving each point 0.001 mm in any direction, ending up with a new point cloud. In order to generate a point cloud with a scratch, the following steps were taken:

1. Two points on the surface are chosen which will represent the beginning and the end of a scratch.
2. A random point on that line will represent an area where the scratch is the widest.
3. All points that are contained within a circle which its center is defined by the random chosen point are affected points.
4. Randomly sample affected points. The probability of a point being sampled and removed is higher when

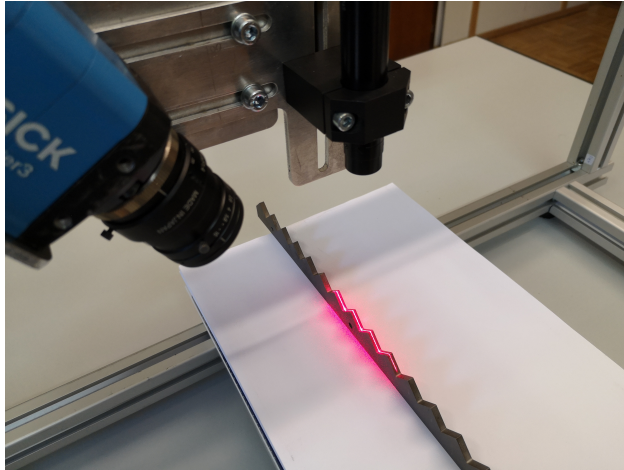


Figure 3: 3D vision system setup.

it is closer to the circle center and center line of the scratch. The radius of the circle is random and varies between 0.15mm to 2mm.

5. The points which were not removed in the previous step are sampled again using the same logic but this time instead of being removed their reflectance is lowered.

Lastly we reduce the problem dimensionality by removing the Z-axis. In order not to lose the Z-axis information it is encoded as brightness of the pixels. Next we select an area of interest on which we will do our analysis. After selecting an area of interest, a median 3x3 filter is applied to remove parabolic lines from the scan since those are a by-product of the whole scanning process. After blurring, a binary threshold is applied using a constant $T=60$. Deciding whether the region of interest includes a scratch, the image is compared to an error-free surface using a metric called "structural similarity index". The values returned by this metric are between 1 and 0.

Using this metric as the only feature, a random forest classifier is trained and the results are shown in Table 4. The classifier achieves a perfect score.

True\predicted	No error	With error
No error	113	0
With error	1	106

Table 4: Confusion matrix of the random forest classifier.

In summary we tested three approaches to solve quality control of metallic car objects: by vibrations, by using a 2D hardware system and by a 3D hardware system. The first two approaches soon turned out unsuccessful. Only the last, hardware costing around 10.000€ enabled good results with machine learning.

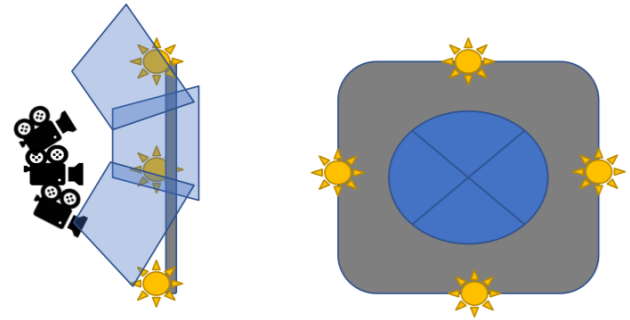


Figure 4: Camera and lighting arrangement.

4 Oven surfaces

This section is focused on detecting imperfections on oven faceplate's metallic surface and the interior surface of the oven. During the final stages of the production of a commercial oven, a key step is visual inspection on its surfaces. For the decision of image acquisition we opted to use a simple RGB camera that will have almost no interference on the production line. The amount of oven faceplates we obtained was again very small, in total 5 faceplates from which we took our images. Out of the 5 faceplates, 3 had scratches on the metallic surface and 2 did not have any scratches.

Before using machine learning models, we also tested some feature extraction methods, such as Speeded Up Robust Features (SURF) and Scale Invariant Feature Transform (SIFT). These were used in order to compare images with and without defects from 13 different key areas on the oven. These key areas are places where the most errors occur, such as panel buttons, door alignment, etc. These methods proved to be effective, achieving an almost perfect F1 score, or at least above 0.9 on all the 13 key areas.

4.1 Machine vision

Alongside detecting scratches on the front metallic surface of the oven, pictures were taken from inside the oven at varying angles and light conditions. In Figure 4 we can see the triple camera arrangement with 4 different light sources. Multiple light sources were used and these caused a certain amount of noise and glare in the final picture, remedying it by performing post-processing of the images. This resulted in improved stability in the classification algorithms. To increase the number of faults on the metallic surface, adhesive tape was used to mimic imperfections alongside with using a sharp metal object to create new scratches on the interior oven surface.

In order to be able to use deep learning models, segmentation and augmentation was used to increase the number of images. Before segmentation, we manually annotated the images that contained an imperfection using a binary mask, 1 indicating a pixel which is part of an imperfection, 0 otherwise. We used various methods to increase the number of images we could get from the faceplates, from

Architecture	Simple-CNN	VGG16	InceptionV3	ResNet101V2
(Macro) F1-score	9.8%	13.0%	58.75%	60.75%

Table 5: Average model F1 score after leave-one-out cross-validation.

covering individual scratches with adhesive tape, to coloring with a felt-tip pen and finally using image augmentation. To segment an image we used a "sliding window" approach where a window slides over and extracts smaller parts of the image. The segments which overlap with the binary mask from before are labeled as images that contain an imperfection. Figure 5 shows an example of how an oven faceplate is segmented, avoiding the display area.

After segmentation, class-invariant transformations are performed in order to increase the dataset size. Traditional image augmentation techniques such as flipping and rotating perform well on the results of a classifier as can also be seen in [19]. With all these methods combined, we managed to obtain more than 20,000 images. We have chosen 4 convolutional neural network models for this task, one simple CNN and the others which are VGG16 [16], InceptionV3 [17], ResNet101V2 [7]. In order to train all of these with transfer learning, we disregarded the last fully connected layers, substituting them with our own layers with randomly assigned weights. During training the weights of all except the final fully connected layers were updated. By making use of transfer learning we shorten the time needed to train a neural network and simultaneously use knowledge which was previously obtained by the pre-trained network on a wide domain of images. The performance of the models was evaluated using leave-one-out cross-validation (LOOCV), with 4 folds. For an evaluation metric F1 macro average score was used. The reason for using the F1 metric with averaging is because of our imbalanced dataset where the majority of samples belong to the "No imperfection" class.

The results of all the models are shown in Table 5. As we can see our simple CNN and VGG16 weren't able to capture the difference between images with or without imperfections. On the other hand, InceptionV3 and ResNet101V2 showed better and similar results to each other.

5 Oven fans

Also regarding commercial ovens, the next quality control solution focuses on oven fans. In order to inspect if a fan which is located inside an oven is in working condition, videos were recorded with a frame rate of 30 frames per second and from those videos images were obtained totalling 7200 images. Out of all those frames 4000 were of a working fan and 3200 were with one which is not working. The visual data was obtained through a closed oven door, since opening the door in a manufacturing environment would take too much time.

Preprocessing of the images consisted of three steps:

1. Object detection

2. Glare reduction

3. Thresholding

5.1 Extracting image features

In object detection the Hough Gradient Method [21] was used to extract certain shapes from images. We are only interested in the area where the fan is located which happens to be a circular opening. Before running the method to detect circles, a median filter is applied to blur the image, since using the original image results in many false positive detection of circles. The output of this method was a single circle although the radius of each circle from all the images varied. Because of this the mean value of the circles' radius was calculated and used.

For glare reduction, each image was decomposed into a color, saturation and brightness component (HSV). After this, particularly bright areas of the image were identified and their pixel values inpainted in respect to their surrounding pixel. We followed few rules while reducing glare. Because reflections coming from the light source are white, any pixel that contains glare will not have saturation, since white has no color or saturation. Then we filtered out the areas that have low saturation. Following that, the non-saturated areas were reduced by an erosion morphological operator and the brightness values of the saturated pixels were set to 0. Lastly, the final glare mask was obtained by choosing pixels that have high brightness (threshold 130). The pixels from the original image that overlap with the glare mask were then interpolated with an inpainting operation. The final image without glare is almost identical to the original image.

The last step is thresholding. In an image frame where a fan is not working, the fan stands out more, and in a frame where a fan is working the fan is slightly blurred. Using this information we decided to use binary thresholding and count the number of white pixels in an image. The idea is that if the thresholded image represents a non working fan, then the number of white pixels would be greater than in a thresholded image representing a working fan. For the binary thresholding we used a constant of $T=90$. Next, to calculate the white pixel count threshold by which we decide the class of the image, that is working/non-working we calculate the number of pixels in all binary-thresholded images of a non-working fans and take the 5th percentile of that. In the end, there were some abrupt changes in classification during consecutive frames and to remedy that the class of each frame was taken as the majority of the last 20 frames.

The confusion matrix for the results is shown in Table 6.

	Non-working	Working
Non-working	3117	82
Working	280	3720

Table 6: Confusion matrix from the results.

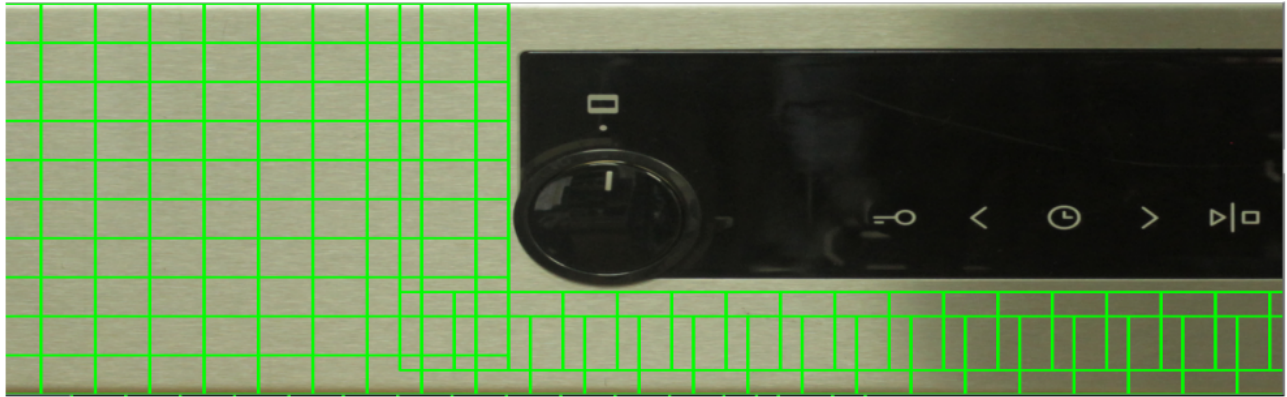


Figure 5: Example of image segmentation.



Figure 6: Oven fan and corresponding glare mask.

5.2 Audio analysis

A technique that is worth mentioning is the use of signal analysis in order to detect the workings of the oven fans. Using an ordinary microphone from a telephone, samples were gathered which then were analysed. The mobile phone was placed on top of the oven housing in order to record the audio signal. It was observed that the operation of the working oven fan was present in the frequency spectrum at the peak of 100Hz. These results were obtained in a controlled laboratory setting. Recording samples were also obtained in a production environment where more noise is present. This added noise introduced some disturbances in the final signal spectrogram which made the detection of a working oven fan more difficult, nevertheless, correct detection of the oven fan was still possible. The ovens are equipped with two fans, a cooling fan and a heating one. During the period when they are operating, they both exhibit a signal of 100Hz on the spectrogram. This makes their separation and identification very difficult since the signals overlap. With this technique, at any given time when the fans are working, it can be concluded that at least one is operational, but not both because of the previously mentioned problem.

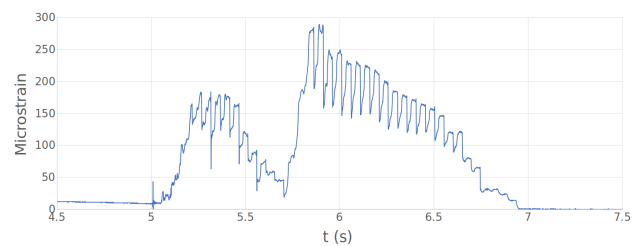


Figure 7: Microstrain parameter over time.

6 Broaches

During metalworking it is necessary to form a metal into a certain shape. Broaching comes into play because it cuts and shapes metal using so called broaches. Broaches come in different shapes and sizes with multiple blades on them, all with varying size. In the broaching process, internal or external broaching can be used and in our approach we experiment with internal broaching. These broaches need to be replaced after prolonged use and our goal is to identify how many operational cycles can be done using the broach before replacement is needed. If the blades of the broach are too worn but not damaged, then they can also be sharpened instead of replacing the whole broach, which is cheaper. To find out when a broach reaches its maximum operational period before irreversible damage, we measure the microstrain parameter by the cut time, which tells us the strain change in parts per million.

6.1 Signal features

The normal amount of operational cycles for a single broach measures around 1500. Using a broach beyond this number has a very high probability that the cuts done by it will not be accurate or that a tooth from it might be damaged and in the end, broken. External factors such as heat and material lubrication affect the rate at which broaches are getting worn out. In defect recognition the most important factors are the number and shape of peaks in the measured stretch signal. In Figure 7 we can see a signal

of the amount of stretch over time which was taken during a cut of the broach and we can see how it changes over time when different sized blades pass through the object. Automatic peak recognition was done by measuring when the signal rises from its standard deviation. Doing this we were able to isolate a time window in which the signal was generated when the broach was used for cutting.

Various features were extracted from the signal such as minimum and maximum signal value, frequencies, patterns etc. The features were filtered using a statistical significance test and after using this test alongside with the Benjamini–Yekutieli procedure, it turned out that there are only three most important features which are: area, number of peaks and maximum value of the signal. These features were used to train several machine learning models such as:

1. Linear regression
2. Gradient boosting
3. AdaBoost
4. K-nearest neighbour

The results show that for broaches that are worn, the area of the signal is larger since more force is needed to make a cut resulting in more stretch. This is analogous to a dull knife requiring more force to make a cut. What is also noticed is that the number of peaks is inversely proportional to the number of cuts done by the broach. Since the blades are getting worn out, they don't scrape enough metal which is why we see a lower amount of peaks. For the regression metrics we used Mean Absolute Error (MAE). Our model achieved a MAE of 27.58 when predicting the current cut cycle. That corresponds to 98.18% accuracy of the model that predicts in which cycle the needle needs to be replaced. In Table 7 we can see the MAE results on all the tested regressors.

Regressor	MAE
Linear regression	101.25
Gradient boost	27.58
AdaBoost	165.44
KNN	74.16

Table 7: Regressors and their corresponding MAE metrics.

7 Discussion

Manufacturing is a large field which encompasses many different production areas, from automotive to pharmaceutical, to also food and packaging. Our proposed quality control models are from a range of applications and each is tailored to the specific use case. Several papers from the Related Work section are comparable to our solutions. For example, we use CNNs in detecting imperfections in oven faceplates, which similar application is used in [11]. Although the accuracy of their proposed network is higher, we

showcase that there is underlying knowledge which can be used in pre-trained networks that brings forth faster training. In the future, custom networks can be built to facilitate a more specific problem learning. Moving on to oven fans, we see that good results are achieved by using simpler techniques such as extracting image features, glare reduction, thresholding and so on. While the work in Table 1 analyzes only static products, our scenario posed a different problem because of the movement of the oven fans. Taking a look at random forest classifiers, comparable results are achieved in our 3.3 section in regards to the efforts of [12]. The results from our random forest classifier can be seen in Table 4.

8 Conclusion

In this paper we presented tests of many different methods in improving the quality control pipeline in manufacturing, using typically only a couple of imperfect products. Our contribution includes novel methods used in conjunction with machine learning such as processing 3D point clouds and vibration analysis of industrial products. An addition to that is the application of various machine learning models on data obtained from an industrial environment. We presented augmentation techniques to increase our dataset size since obtaining defective industrial products in large numbers is difficult. The results from each applied method were presented in the previous sections and they are encouraging. With more data it is possible to train better classifiers which will have better performance in classifying defective samples from the production line. This research tests different kinds of industrial products for some of which only a sparse number of samples could be obtained. It is worth mentioning that some of the niche industrial products have such specific defects that so far using machine learning methods proved to be unsatisfactory. The current classifiers have difficulties learning meaningful information from those samples, nevertheless results on other industrial products give hope that new models with better quality can be learned if sufficient data is acquired.

The purpose of this study was to test which hardware and which methods should be applied in the second stage of a project, that is, actual implementation. The study showed that several approaches demand special care for our specific real-life quality control, namely:

1. Some methods like DNNs demand lot of learning examples. In some cases, it was possible to artificially generate these learning examples and the machine learning approach was successful, otherwise the approach failed.
2. Depending on the difficulty of the quality control, cheaper or more expensive hardware was needed. For several visual problems, quality cameras were sufficient, but for demanding visual problems the only hardware that enabled good results was the 10.000 euro 3D hardware system.

3. Machine learning proved successful when sufficient number of learning examples were given, either real or artificially generated. However, each approach needed special machine learning methods. For example, DNs sometimes performed best and sometimes worst. In summary: having quality input data enables creation of proper machine learning models.

References

- [1] Núria Banús et al. “Deep learning for the quality control of thermoforming food packages”. In: *Scientific Reports* 11.1 (2021), pp. 1–15. DOI: 10.1038/s41598-021-01254-x.
- [2] Mark-Anthony Bray and Anne E Carpenter. “Quality control for high-throughput imaging experiments using machine learning in cellprofiler”. In: *High Content Screening*. Springer, 2018, pp. 89–112. DOI: 10.1007/978-1-4939-7357-6_7.
- [3] Tianqi Chen and Carlos Guestrin. “Xgboost: A scalable tree boosting system”. In: *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*. 2016, pp. 785–794. DOI: 10.1145/2939672.2939785.
- [4] D Coupek et al. “Proactive quality control system for defect reduction in the production of electric drives”. In: *2013 3rd International Electric Drives Production Conference (EDPC)*. IEEE, 2013, pp. 1–6. DOI: 10.1109/edpc.2013.6689762.
- [5] Carlos A Escobar and Ruben Morales-Menendez. “Machine learning techniques for quality control in high conformance manufacturing environment”. In: *Advances in Mechanical Engineering* 10.2 (2018), p. 1687814018755519. DOI: 10.1177/1687814018755519.
- [6] Kaiming He et al. “Deep residual learning for image recognition”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 770–778. DOI: 10.48550/arXiv.1512.03385.
- [7] Kaiming He et al. “Identity mappings in deep residual networks”. In: *European conference on computer vision*. Springer, 2016, pp. 630–645. DOI: 10.1007/978-3-319-46493-0_38.
- [8] Gao Huang et al. “Densely connected convolutional networks”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 4700–4708. DOI: 10.48550/arXiv.1608.06993.
- [9] Anil Mital, M Govindaraju, and B Subramani. “A comparison between manual and hybrid methods in parts inspection”. In: *Integrated Manufacturing Systems* (1998). DOI: 10.1108/09576069810238709.
- [10] Antje Niederlein et al. “Image analysis in high content screening”. In: *Combinatorial chemistry & high throughput screening* 12.9 (2009), pp. 899–907. DOI: 10.2174/138620709789383213.
- [11] Przemysław Oborski and Przemysław Wysocki. “Intelligent Visual Quality Control System Based on Convolutional Neural Networks for Holonic Shop Floor Control of Industry 4.0 Manufacturing Systems”. In: *Advances in Science and Technology. Research Journal* 16.2 (2022), pp. 89–98. DOI: 10.12913/22998624/145503.
- [12] Ricardo Silva Peres et al. “Multistage quality control using machine learning in the automotive industry”. In: *IEEE Access* 7 (2019), pp. 79908–79916. DOI: 10.1109/access.2019.2923405.
- [13] Shaoqing Ren et al. “Faster r-cnn: Towards real-time object detection with region proposal networks”. In: *Advances in neural information processing systems* 28 (2015), pp. 91–99. DOI: 10.1109/tpami.2016.2577031.
- [14] Jacqueline Schmitt et al. “Predictive model-based quality inspection using Machine Learning and Edge Cloud Computing”. In: *Advanced engineering informatics* 45 (2020), p. 101101. DOI: 10.1016/j.aei.2020.101101.
- [15] Michael Sharp, Ronay Ak, and Thomas Hedberg Jr. “A survey of the advancing use and development of machine learning in smart manufacturing”. In: *Journal of manufacturing systems* 48 (2018), pp. 170–179. DOI: 10.1016/j.jmsy.2018.02.004.
- [16] Karen Simonyan and Andrew Zisserman. “Very deep convolutional networks for large-scale image recognition”. In: *arXiv preprint arXiv:1409.1556* (2014). DOI: 10.48550/arXiv.1409.1556.
- [17] Christian Szegedy et al. “Rethinking the inception architecture for computer vision”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 2818–2826. DOI: 10.1109/cvpr.2016.308.
- [18] Chris J Turner et al. “Intelligent decision support for maintenance: an overview and future trends”. In: *International Journal of Computer Integrated Manufacturing* 32.10 (2019), pp. 936–959. DOI: 10.1080/0951192x.2019.1667033.
- [19] Jason Wang, Luis Perez, et al. “The effectiveness of data augmentation in image classification using deep learning”. In: *Convolutional Neural Networks Vis. Recognit* 11 (2017), pp. 1–8. DOI: 10.48550/arXiv.1712.04621.
- [20] Jing Yang et al. “Using deep learning to detect defects in manufacturing: a comprehensive survey and current challenges”. In: *Materials* 13.24 (2020), p. 5755. DOI: 10.3390/ma13245755.

- [21] HK Yuen et al. “Comparative study of Hough transform methods for circle finding”. In: *Image and vision computing* 8.1 (1990), pp. 71–77. DOI: 10 . 1016/0262-8856(90)90059-e.