Product Quality Control in Assembly Machine under Data Restricted Settings

Fatemeh Kakavandi*, Roger de Reus†, Cláudio Gomes*,
Negar Heidari*, Alexandros Iosifidis*, Peter Gorm Larsen*
*Department of Electrical and Computer Engineering, Aarhus University, Aarhus, Denmark
†Device Manufacturing Development Department 2540, Novo Nordisk A/S, Hillerød, Denmark
Email: fateme.kakavandi@ece.au.dk

2022 IEEE 20th International Conference on Industrial Informatics (INDIN) | 978-1-7281-7568-3/22/$31.00 ©2022 IEEE | DOI: 10.1109/INDIN51773.2022.9976173

***Abstract*—Evaluating the product quality in an assembly machine is critical yet time-consuming since, in product assessment in batch manufacturing, a certain amount of products should be investigated in an invasive manner. However, continuous manufacturing ensures product quality assessment during assembly with high efficiency and traceability. This paper proposes a quality assessment method for an industrial use case. First, the data is prepared based on two indicators and expert knowledge. Then two data classification approaches (one-class classification and binary classification) are applied to evaluate the products' quality by analysing the related data. Finally, the most efficient model is selected to predict the product labels and deviate anomalies from normal products. For the studied use case and the limited number of products, the binary classifier guarantees to detect 100% of defective products. The proposed approach can provide the engineers and operators with understandable extracted process knowledge, and can therefore be adapted to a high-speed manufacturing line where large data volume and process complexity can be problematic.**

***Index Terms*—Medical Device Assembly, Anomaly Detection, Product Quality Assessment, One Class Support Vector Machine, Binary Classifier.**

I. Introduction

Continuous Manufacturing has become of great importance in pharmaceutical industry recently, to support the use of modern manufacturing technology and to simplify the manufacturing processes by, e.g. using an integrated process with fewer process steps and shorter processing times; supporting an enhanced development approach; enabling real-time product release; and providing flexible operation. The intention is that this will be to the benefit of both industry and patients [1].

The mechanical process steps in the assembly of medical devices are needed in order to secure high precision of the assembly steps during high-speed manufacturing. Quality control of the mechanical processes is usually carried out by using calibrated equipment and by following specific control steps measuring defined output characteristics, e.g. examining whether the force applied during a step in the production of a product is within a pre-specified range or the product should be declined. Using such rule-based approaches for quality control increases the number of control steps and the complexity of the entire assembly process, specifically when there is a limited data collection and, thereby, limited insight in actual process performance.

However, with the use of new sensors, introduction of new technology and use of Machine Learning (ML) algorithms, one can develop a quality assessment model for the assembly process of medical devices by identifying critical process parameters, process responses and quality characteristics. ML algorithms have been rarely used in quality control systems for medical device assembly, and to the best of our knowledge, there is only one example of ML-based quality control system for this purpose [2]. It is expected that adopting ML-based quality assessment models in assembly of medical devices will increase device quality, lower manufacturing costs, decrease the lead time from assembly to release of device and, in the end, improve availability of quality devices to the patients.

Anomaly detection models are beneficial for product quality assessment since they can detect any deviation from the normal status especially when there are not sufficient samples of faulty products. Different approaches to address anomaly detection in time series have been proposed [3], [4]. However, these complex and time consuming computational methods are not suitable for real-time quality control systems. Furthermore, the high complexity of the suggested algorithms often results in poor explainability of the model. Support Vector Machine (SVM) is one of the well-known ML methods that have shown good performance in different fields; e.g. a quality monitoring model for a plastic injection molding process based on SVM was designed in [5]. Variants of the SVM targeting one-class classification problems have also been shown to be effective in novelty detection problems [6], [7].

To overcome the current drawbacks of the conventional quality assessment tools in medical device assembly lines, we propose an anomaly detection method for an industrial use case. In this method, first, the data related to a specific step in the assembly process is identified based on two different recorded measurements and expert knowledge. This data preparation and feature extraction method can reduce the computational time in training an efficient ML method. Then the conventional data preprocessing techniques like noise filtering and data normalization are applied. Two ML-based approaches are assessed to find the best candidate for detecting anomalies from normal products for the existing data. This process can give the operators reliable feedback about the current status of the assembly line for a more informed troubleshooting process.

The paper is structured as follows. Background on ML methodologies is provided in Section II. The industrial use case is described in Section III. The adopted methodological steps are described in Section IV, followed by the method evaluation in Section Section V. Conclusions are drawn in Section VI.

II. Background

In this section, we introduce the concept of anomaly detection, and then we narrow down to the two ML models used in this paper: one class SVM and Binary SVM.

A. Anomaly Detection

Anomaly detection is similar to outlier detection in statistics and is used interchangeably with novelty detection in the data analysis field. As can be derived from the phrase, anomaly detection means observing unusual and unexpected instances.

Developments in ML methods and the ability to collect extensive data sets make it possible to improve anomaly detection methods. However, to be useful in applications, data collected at a point in time or over an extended period must be accurate and reliable. In this paper, we try to improve the accuracy and reliability of the detection model by choosing those measurements that are sensitive to deviation from normal situations.

Anomaly detection endeavours to identify data occurrences that are dissimilar from normal samples. So far, there have been many applications requiring identification of anomalies [8], [9]. Anomaly detection is also advantageous for production system monitoring in digitalization and continuous manufacturing. However, finding an anomaly with traditional methods in complex industrial devices is not easy; unknown types of anomalies, lack of traceability, system complexity, and equipment degradation are some of these challenges.

The SVM technique is introduced in [10] as a classification method. SVM can be applied to define a linear hyperplane depending on the data distribution when the classes forming a classification problem can be divided linearly. In more complex cases, where data forming the classes is distributed non-linearly, the raw data should be mapped into a new feature space in which classes can be discriminated linearly.

B. One Class Support Vector Machine

One Class Support Vector Machine (OCSVM) is one of the methods that have been proposed for solving one-class classification problems [11]. In this method, data samples with normal labels are mapped to another feature space, in order to find a hyperplane with the largest distance from the origin, while all mapped data are placed on the opposite side of the hyperplane. Assuming that $\mathbf{x}_i$ is the feature vector of a data sample and $\mathbf{w}$ is the weight vector, the defined hyperplane in OCSVM is formulated as $\mathbf{w}^T\boldsymbol{\phi}(\mathbf{x}_i) - \rho = 0$. This goal can be formulated as the following primal optimization problem:

$$\begin{aligned} &\min_{\mathbf{w},\boldsymbol{\xi},\rho} \frac{1}{2}\|\mathbf{w}\|^2 + \frac{1}{\nu N}\sum_{i=1}^{N}\xi_i - \rho \\ &\text{s.t. } \mathbf{w}^T\boldsymbol{\phi}(\mathbf{x}_i) - \rho + \xi_i \geq 0, \quad \xi_i \geq 0, \quad \forall i. \end{aligned} \tag{1}$$

In (1), $\nu \in (0, 1]$ is the regularization factor, which expresses a trade-off between model complexity and training error; it controls the number of training samples excluded by decision boundary. ξ_i's are the slack variables to soften the margin and $\boldsymbol{\phi}(\cdot)$ is the mapping function. Furthermore $\mathbf{w}$ and ρ are a weight vector and an offset, respectively, parameterizing a hyperplane in the feature space. The dual form of the 1 is as follows:

$$\begin{aligned} &\max_{\boldsymbol{\alpha}} -\frac{1}{2}\sum_{i,j=1}^{N}\alpha_i\alpha_j K(\mathbf{x}_i, \mathbf{x}_j), \\ &\text{s.t. } \sum_{i=1}^{N}\alpha_i = 1, \quad 0 \leq \alpha_i \leq \frac{1}{\nu N}, \quad \forall i, \end{aligned} \tag{2}$$

where $K(\mathbf{x}_i, \mathbf{x}_j) = \boldsymbol{\phi}(\mathbf{x}_i)^T.\boldsymbol{\phi}(\mathbf{x}_j)$, the kernel function, is the inner product of mapped data, and α_i is the dual variable. One of the nonlinear kernels that is commonly used is Radial Basis Function (RBF) $K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma\|\mathbf{x}_i - \mathbf{x}_j\|^2)$, where γ is a scalar that defines how much influence a single training sample has.

C. Binary Classifier with Support Vector Classifier

Support Vector Network is one of binary classification methods first proposed in [12] and later the parameter ν which is proved to be an upper bound on the fraction of training errors and a lower bound of the fraction of support vectors was introduced in [13]. The ν-Support Vector Classifier (ν-SVC) is the binary classifier that is used for separating the normal and abnormal samples in this paper. The objective of this model is to separate the two classes of data with a hyperplane and unlike the OCSVM model, the binary classifier is a supervised method. Similar to the previous section we assume $\mathbf{x}_i$ as the input vector and $y_i \in \{-1, +1\}$ as the data label where the $+1$ shows the normal data and the -1 shows the anomalous samples. The primal optimization problem is as following:

$$\begin{aligned} &\min_{\mathbf{w},\boldsymbol{\xi},b,\rho} \frac{1}{2}\|\mathbf{w}\|^2 - \nu\rho + \frac{1}{N}\sum_{i=1}^{N}\xi_i \\ &\text{s.t.} \quad y_i(\mathbf{w}^T\boldsymbol{\phi}(\mathbf{x}_i) + b) \geq \rho - \xi_i, \quad \xi_i \geq 0, \quad \forall i. \end{aligned} \tag{3}$$

As mentioned in the previous subsection, ξ_i's are the slack variables and $\boldsymbol{\phi}(\cdot)$ is the mapping function while $\mathbf{w}$, ρ and b are weight vector and offsets parameterizing a hyperplane in the feature space. The dual problem is as the following:

$$\begin{aligned} &\min_{\boldsymbol{\alpha}} \frac{1}{2}\sum_{i,j=1}^{N}\alpha_i(y_i y_j K(\mathbf{x}_i, \mathbf{x}_j))\alpha_j, \\ &\text{s.t.} \quad 0 \leq \alpha_i \leq \frac{1}{N}, \quad \sum_i \alpha_i \geq \nu, \quad \sum_i y_i\alpha_i = 0, \quad \forall i. \end{aligned} \tag{4}$$

D. Leave-One-Out Cross-Validation

Leave-One-Out cross-validation is a popular method employed in many studies to evaluate the performance of a classification model, especially when the number of samples either in the data set or in one specific class is small [14]. This validation method aims to obtain a reliable accuracy estimate of the classification model. This method leaves one sample out from the training set; e.g. assume n samples in the data set, n-1 are used for training the model, and one is used for validation. Cross-validation is an advantageous technique for assessing the effectiveness of the model, particularly in need of decreasing overfitting. It is also of use in determining the model's hyperparameters, in the sense that which parameters will result in the lowest test error.

In this paper, an adapted version of the Leave-One-Out cross-validation method, called One-Anomaly-out cross-validation, is applied for hyperparameter selection in OCSVM and ν-SVC. This method leaves one abnormal sample out in each iteration while the rest of the abnormal data set participates in the validation process. For tuning ν in OCSVM, the training set consists of only normal samples; however, the validation set includes normal and abnormal samples. In ν-SVC hyperparameter selection, specific ratios of normal and abnormal data sets are used for both training and validation processes. In both OCSVM and ν-SVC hyperparameter tuning, one abnormal sample and some normal samples are combined to test the model in each iteration.

III. Industrial Use Case

The use case generating the data that the two classifiers are tested on features a snap process from a real-world pharmaceutical manufacturing pilot line. In the related station two parts of the medical device called sub-assemblies are assembled by a linear motor with vertical displacement.

A. Pilot Line

The pilot line consists of a test bench for medical device assembly by Stevanato Group/SVM. As shown in Fig. 1(a), the assembly line contains a base and top frame (1 and 2). The transport system (3) is a modular XTS linear motion platform by Beckhoff Automation. The linear motors (5 and 6, Linmot PR02-52) provide both vertical and rotary motion and are equipped with force and torque transducers for continuous process monitoring with sampling frequencies up to 20 kHz. A height sensor (4) is used to verify the assembly process. The actual image of the machine is depicted in Fig. 1(b). The mover, or pallet, is mounted on the transport system. It contains the sub-assemblies and component to be assembled and can move to an arbitrary position along its axis of movement.

B. Product

To complete the medical device assembly two modules (or sub-assemblies) and one component must be assembled (see Fig. 2(a)) . The modules and components are kept together by a snap connection.

C. The Snap Process

This process consists of two sub-assemblies snap together with linear displacement along the arrow direction shown in Fig. 2(a). However, prior horizontal alignment and rotational orientation are needed for the two sub-assemblies to be in the required position with respect to each other. Next, to mount the two modules and stabilise the assembly a snap-fit method as shown in Fig. 2(b) consisting of two main snaps and two minor snaps is applied. To study the behaviour of the snap process, we look into different sensors signals, as shown in Fig. 3. The force profile and displacement are correlated with the snap process, which means for mounting two sub-assemblies, the gripper moves down, resulting in negative values in displacement measurement. At the same time, the applied force increases for interlocking the two parts. The torque measurement does not carry valuable knowledge since the interlocking process lacks rotational movement. The velocity measurement is also the derivative of the vertical displacement.

IV. Methodology

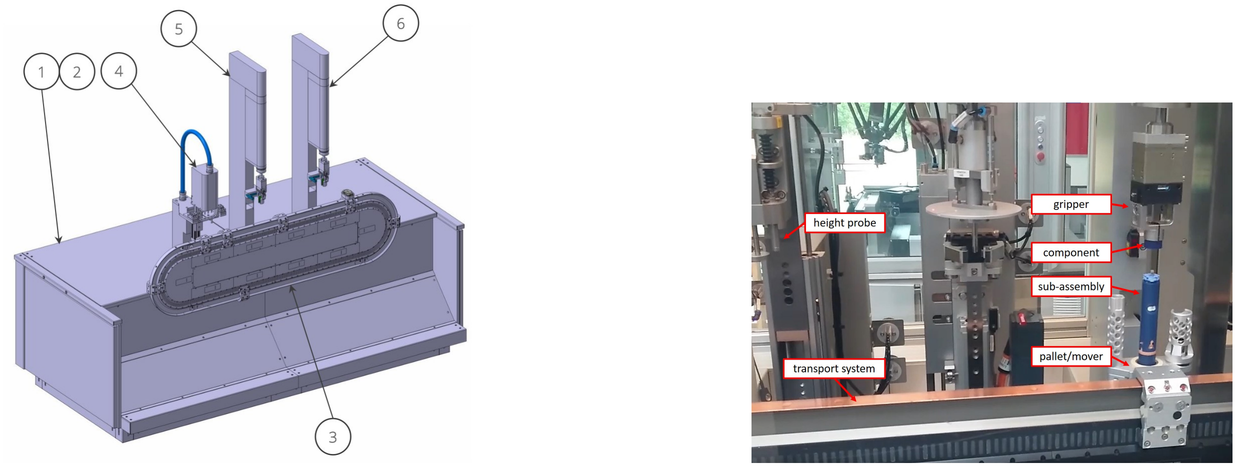
We propose the following methodology. First, data generation is carried out, where normal and abnormal products are assembled, and related sensors are recorded. Then, sensors related to the snap process and suitable for anomaly detection are selected. Since the data is collected continuously, it forces us to use a data preparation process to find the exact data segment related to the snap process. Finally, the anomaly detection model is trained on the prepared data, and the output of this model is the process quality label. The data preparation method, considers supervised dimension reduction, aims for faster implementation and testing of the anomaly detection model for future data and more accurate performance.

A. Data Generation

The assembly machine assembles products with normal status to generate the data set labeled normal. Some experiments have been conducted to introduce abnormal behaviour in the assembly process to evaluate the anomaly detection model. Furthermore, we produced products labeled "abnormal" with different process characteristics. The anomalous samples are performed in two different types, *Type1* and *Type2*. With fault Type1, we change the setting of the assembly machine, so the gripper offset differs from the calibrated working point. The calibrated value is $180°$; however, the gripper position gets different values as $\{170°, 175°, 185°, 190°\}$, as shown in Table I. While in fault Type2, we change the structure of the sub-assemblies (see Section III-B). We remove some deformation structures in the products, as shown in Table I, with varying number of structures to see various abnormal behaviour.

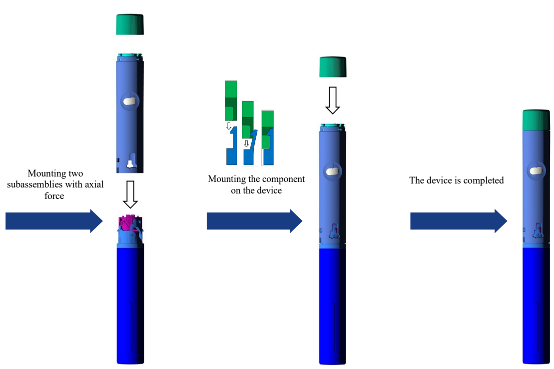
B. Sensor Selection and Data Preparation

For the various recorded measurements of the snap process, the force profile and displacement best show the process response. As described in Section III-C, the vertical displacement with axial force is applied for snapping the sub-assemblies. At the same time, there is no rotational movement for interlocking the two sub-assemblies. Therefore, the torque sensor does not reveal extra information about the process. The displacement and velocity measurements are recorded directly



(a) (b)

Fig. 1. (a) 3D schematic of the assembly equipment consists of (1) base frame, (2) top frame (not shown in figure) (3) transport system, (4) sensor system, (5) and (6) linear motors. (b) Shows the linear motors perform the assembly steps, and the sensor system verifies the assembly [15].



(a) (b) (c)

Fig. 2. The product consists of two sub-assemblies and a component, the vertical displacement with applied force snap two sub-assemblies (a), the snap-fit method is used for interlocking the two modules (b), the product is finished after the snap process (c)

from the linear motor and are used to indicate the snap process period. Therefore, we focus on the specific part of the force profile that shows the snapping action using displacement and velocity measurements, as shown in Fig. 3. The green line depicts when the gripper is holding the sub-assembly and waiting for the vertical down movement. The red line specifies when the velocity is zero, no further action is applied, and the sub-assemblies are mounted completely. Consequently, we focus only on the specific time where the gripper holds the top sub-assembly and mount it on the other one. Therefore, the redundant part of the force profile where the machine is not in contact with the device is ignored. By selecting the correct part of the force measurement, we do not lose the critical part of the data, and we supervise dimension reduction, which will be beneficial for saving time for training the anomaly detection model.

TABLE I. Different types of faults that have been applied

Fault Types	Value
Type1	175°
	170°
	185°
	190°
Type2	4
	2
	2
	3
	all

C. Anomaly Detection

To evaluate the snap quality, we use two efficient ML models, described in Section II. Since we have a handful of anomalous samples and a larger set of normal samples, we try two different approaches to address the issue of the small, imbalanced data set. The first approach is to find the boundary around the normal samples by OCSVM, where the model is only trained on normal data and afterwards is tested both with normal and abnormal sets. The strength of this method is to focus mainly on the normal data set and try to verify the distribution related to normal samples. Since we have an imbalanced data set with a larger amount of normal samples, it is beneficial to find the normal sample distribution.

The second approach is to train a binary classifier that benefits from both normal and abnormal samples in training time. Finally, we try to see which method performs best and consider this model as the winning candidate for the anomaly detection.

The algorithm shown in Fig. 4 illustrates the overall view of the methodology. First, the new product is assembled by

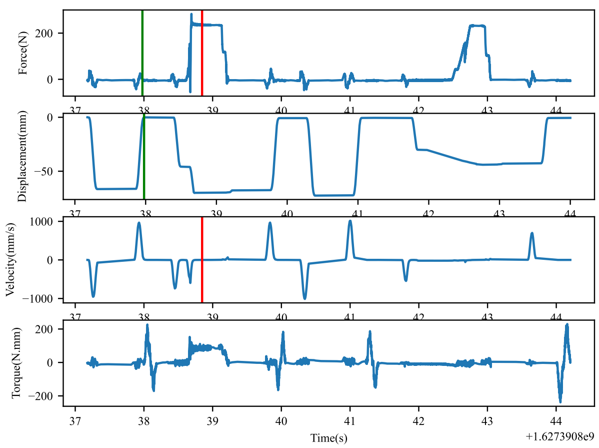


Fig. 3. Different sensors signals for the snap process. The specific force profile is selected by displacement and velocity measurements, where the second local maximum point of displacement and the fifth local minimum points in velocity are the indicators.

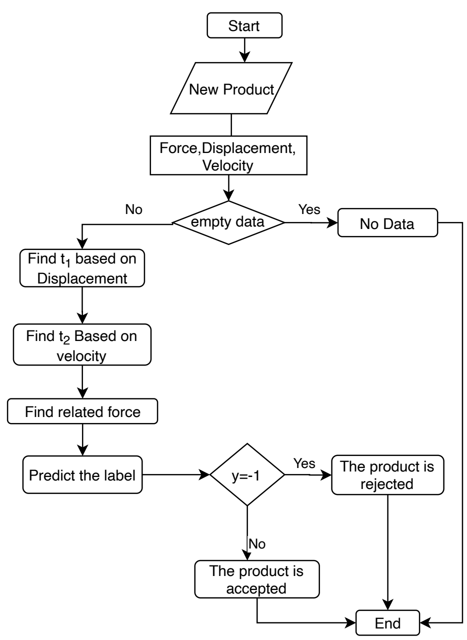


Fig. 4. The data preparation method for anomaly detection

the pilot line. Then the three critical measurements related to the snap process are extracted from the recorded data set. Afterwards, we find the starting point of the snap process t_1 from the displacement measurement. The end point of the snap process t_2 is derived from the velocity measurement. Based on the boundary $[\mathrm{t}_1, \mathrm{t}_2]$ found in the previous step, we select the correct segment of force measurement. Finally, the sample is tested by the anomaly detection model (the best performing model among OCSVM or the binary classifier) to predict the label of the product.

V. Evaluation

In this section, we evaluate the proposed algorithm for detecting anomalous products in the industrial use case described in Section III. We produced 102 normal and nine abnormal products for this purpose. After generating the data, we apply the data preparation process as the core preprocessing step, as mentioned in Section IV. To find the best hyperparameter ν for OCSVM and ν-SVC, we employ the One-Anomaly-Out validation method, see Section II. To get the most reliable accuracy, we employ the Micro-F1 score metric to evaluate the performance of the model due to the imbalanced data set. Since we only have two positive and negative classes which represent normal and abnormal samples respectively, the Micro-F1 score is calculated as $\frac{\mathrm{TP+TN}}{\mathrm{TP+FP+TN+FN}}$, where TP is the number of positive samples predicted as positive samples, etc.

A. Hyper Parameter Selection

The One-Anomaly-Out Validation method leaves one of the anomalous samples out for the testing process in each iteration. Therefore, these iterations are repeated nine times since there are nine abnormal samples. In the hyperparameter selection process, the normal and abnormal data sets are split into train, validation, and test categories. In OCSVM hyperparameter selection, the training set only includes normal samples; however, the validation and test sets consist of abnormal samples. Moreover, in ν-SVC, the abnormal samples exist in all three phases of train, validation and test. In the following, the differences between the two processes are described.

1) One Class Support Vector Machine: In each iteration of the One-Anomaly-Out validation method, 60% of the normal data train the one-class classifier. Then 20% of the normal data is combined with eight abnormal samples to validate the performance of the model with the candidate ν. Therefore, one abnormal sample is reserved with the last 20% of the normal data for testing in each iteration to report the model performance with candidate ν. This iteration is repeated nine times, and each time, one of the abnormal samples is included in the test set. Finally, the performance of the model for each specific ν is calculated as the average of all iterations' performance, as shown in Algorithm 1. Values in $[0.001, 0.01]$ have been investigated with a step length of 0.001 to find the best value for ν. The aim in One-Anomaly-Out hyperparameter selection is to choose the ν that has the most significant Micro-F1 score value. According to Table II all $\{0.003, 0.004, 0.006, 0.007\}$ have the same and the largest value of the Micro-F1 score. Therefore, we set the ν parameter to 0.004.

2) Binary Support Vector Classifier: Similar to OCSVM, for hyperparameter tuning in binary classifier ν-SVC, we use One-Anomaly-Out validation method. In this case, in each iteration, one anomalous sample is out for the testing, combined with 20% of the normal samples, while six abnormal samples are joined with 60% of the normal data set to shape the training set. Furthermore, two remnants of the abnormal data set are merged with 20% of the normal data set for validating the model. The performance for each specific model

Algorithm 1 One-Anomaly-Out cross-validation algorithm for OCSVM hyperparameter tuning.

Require: $\mathcal{D}_{normal} = \{x_i^n\}_{i=1}^N$, $\mathcal{D}_{abnormal} = \{x_j^m\}_{j=1}^M$, $x_i^n, x_j^m \in \mathbb{R}^l$, $\nu = \{\nu_r\} = \{0.001 \times r\}_{r=1}^{10}$
Require: $\mathcal{M}$ OCSVM model
Ensure: Optimum ν^* for $\mathcal{M}$ with highest Micro-F1 score
Split $\mathcal{D}_{normal}$ to $\mathcal{D}_{Train}, \mathcal{D}_{Val}, \mathcal{D}_{Test}$
$MicF1ValAll \leftarrow [\,]$
$MicF1TestAll \leftarrow [\,]$
for $r \leftarrow 1$ to 10 **do**
  $MicF1ValSum \leftarrow 0$
  $MicF1TestSum \leftarrow 0$
  Train $\mathcal{M}$ on D_{Train} with ν_r
  for $j \leftarrow 1$ to M **do**
    Compute validation Micro-F1 score $MicF1Val$ for $\mathcal{M}$ with $D_{abnormal} \setminus \{x_j^m\} \bigcup D_{Val}$
    $MicF1ValSum \leftarrow MicF1ValSum + MicF1Val$
    Compute $MicF1Test$ for $\mathcal{M}$ with $D_{Test} \bigcup \{x_j^m\}$
    $MicF1TestSum \leftarrow MicF1TestSum + MicF1Test$
  end for
  $MicF1ValAvg \leftarrow MicF1ValSum/M$
  $MicF1TestAvg \leftarrow MicF1TestSum/M$
  $MicF1ValAvg$ append to $MicF1ValAll$
  $MicF1TestAvg$ append to $MicF1TestAll$
end for
$MicF1MaxIdx \leftarrow \arg\max(MicF1ValAll)$
$\nu^* \leftarrow \nu[MicF1MaxIdx]$

with the candidate ν is calculated by averaging all nine iterations' results.

According to the results of One-Anomaly-Out validation process, all the values of ν in the $[0.001, 0.09]$ with steps of 0.01 have the same performance, which shows the binary classifier can ideally detect the anomalous samples.

B. Results

The results of both one-class and binary classifier are reported for the best value of ν in Table III. Since we need a high certainty in model prediction for our use case (to avoid waste of time and resources), the binary classifier seems to be more reliable. However, we need a large enough data set with sufficient amount of abnormal samples in both train and test data splits. On the other hand, data generation demands time and resources. Increasing the model generalizability by having limited amount of data is considered as future research direction.

TABLE II. The One-Anomaly-Out validation results for OCSVM

ν	**Micro-F1 Score**
0.001	0,4643
0.002	0,8571
0.003	0,8929
0.004	0,8929
0.005	0,8571
0.006	0,8929
0.007	0,8929
0.008	0.8571
0.009	0.8571
0.01	0.8571

TABLE III. The Performance of OCSVM and ν-SVC

Model	**Micro-F1 Score**
OCSVM	0.8571
ν-SVC	1

VI. Conclusion

In this work, an ML-based anomaly detection approach was developed and evaluated for quality assessment in an industrial medical device assembly line. The method consists of different steps, including data generation, data preparation, and ML model training. After producing some normal and abnormal products and recording different sensors, in the data preparation step, the sensor selection and feature extraction occur in a supervised manner. Later on, an efficient ML model (binary classifier) is selected to predict the quality of the products. The results show that this anomaly detection approach can identify defective products with 100% accuracy. However, this method is in the early development stage and applied only for the snap process with limited data. In the future, the model will be trained on a large volume of data, reflecting the various pilot line processes. Additionally, the anomaly detection model will be integrated into the existing dashboard for the pilot line to make the model more accessible to operators.

Acknowledgment

We would like to thank the following people from Novo Nordisk A/S; Thomas Algot Søllested, Jeppe Wind, Tinna Dofradóttir, and Sebastian Dengler for providing the physical system, conducting data collection, and insights to understand the assembly process. Additionally, we extend our gratitude to Innovation Foundation Denmark for funding the MADE FAST project, which this work is a part of.

References

[1] U.S. Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER), "Quality considerations for continuous manufacturing guidance for industry," https://www.fda.gov/regulatory-information/search-fda-guidance-documents/quality-considerations-continuous-manufacturing, (accessed May 2022).

[2] S. Dengler, S. Lahriri, E. Trunzer, and B. Vogel-Heuser, "Applied machine learning for a zero defect tolerance system in the automated assembly of pharmaceutical devices," *Decision Support Systems*, vol. 146, p. 113540, 2021.

[3] Z. Niu, K. Yu, and X. Wu, "LSTM-based VAE-GAN for time-series anomaly detection," *Sensors*, vol. 20, no. 13, p. 3738, 2020.

[4] M. Munir, S. A. Siddiqui, A. Dengel, and S. Ahmed, "DeepAnT: A deep learning approach for unsupervised anomaly detection in time series," *IEEE Access*, vol. 7, pp. 1991–2005, 2018.

[5] B. Ribeiro, "Support vector machines for quality monitoring in a plastic injection molding process," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 35, no. 3, pp. 401–410, 2005.

[6] J. Raitoharju and A. Iosifidis, "Generalized reference kernel for one-class classification," *arXiv preprint arXiv:2205.00534*, 2022.

[7] V. Mygdalis, A. Iosifidis, A. Tefas, and I. Pitas, "Graph embedded one-class classifiers for media data classification," *Pattern Recognition*, vol. 60, pp. 585–595, 2016.

[8] J. An and S. Cho, "Variational autoencoder based anomaly detection using reconstruction probability," *Special Lecture on IE*, vol. 2, no. 1, pp. 1–18, 2015.

[9] U. Yokkampon, S. Chumkamon, A. Mowshowitz, R. Fujisawa, and E. Hayashi, "Improved variational autoencoder anomaly detection in time series data," in *International Conference on Systems, Man, and Cybernetics*. IEEE, 2020, pp. 82–87.
[10] A. J. Smola, "Learning with kernels," Ph.D. dissertation, Technical University of Berlin, Germany, 1998.
[11] B. Schölkopf, J. C. Platt, J. Shawe-Taylor, A. J. Smola, and R. C. Williamson, "Estimating the support of a high-dimensional distribution," *Neural Computation*, vol. 13, no. 7, pp. 1443–1471, 2001.
[12] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
[13] B. Schölkopf, A. J. Smola, R. C. Williamson, and P. L. Bartlett, "New support vector algorithms," *Neural Computation*, vol. 12, no. 5, pp. 1207–1245, 2000.
[14] T.-T. Wong, "Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation," *Pattern Recognition*, vol. 48, no. 9, pp. 2839–2846, 2015.
[15] F. Kakavandi, C. Gomes, R. D. Reus, J. Badstue, J. Langdal, P. G. Larsen, and A. Iosifidis, "Towards developing a digital twin for a manufacturing pilot line: An industrial case study," in *Digital Twin Driven Intelligent Systems*. Springer Nature, Unpublished.