# Dynamic Bayesian Network-Based Anomaly Detection for In-Process Visual Inspection of Laser Surface Heat Treatment

[1]Alberto Ogbechie, [1,2]Javier Díaz-Rozo, [1]Pedro Larrañaga, and [1]Concha Bielza

[1]Department of Artificial Intelligence, Technical University of Madrid, School of Computer Science. Madrid, Spain
[2]Ikergune A.I.E., San Antolín 3, 20870 Elgoibar, Spain
a.ogbechie@upm.es

**Abstract.** We present the application of a cyber-physical system for in-process quality control based on the visual inspection of a laser surface heat treatment process. To do this, we propose a classification framework that detects anomalies in recorded video sequences that have been pre-processed using a clustering-based method for feature subset selection. One peculiarity of the classification task is that there are no examples with errors, since major irregularities seldom occur in efficient industrial processes. Additionally, the parts to be processed are expensive so the sample size is small. The proposed framework uses anomaly detection, cross-validation and sampling techniques to deal with these issues. Regarding anomaly detection, dynamic Bayesian networks (DBNs) are used to represent the temporal characteristics of the normal process. Experiments are conducted with two different types of DBN structure learning algorithms, and classification performance is assessed on both anomaly-free examples and sequences with anomalies simulated by experts.

**Keywords:** Dynamic Bayesian network, anomaly detection, laser heat treatment, visual inspection, cyber-physical system.

## 1 Introduction

During the production of a steel product, there are specific needs regarding the modification by heat treatment of its surface mechanical properties in order to meet final application requirements. In this respect, laser heating is capable of operating on small and localized areas, ensuring the repeatability, reproducibility and robustness demanded by the manufacturing sector. Laser beams are high-density energy sources that induce fast heating-cooling cycles that are sensed by contactless pyrometers or thermal high-speed cameras. Consequently, in-process monitoring systems manage high volumes of data, increasing the computational power required to give timely feedback.
The objective of this paper is to continue the research reported in [1] and take advantage of the capability of cyber-physical systems (CPS) to handle large

amounts of information through embedded processing capabilities [2]. In this paper, we further investigate the visual inspection of the laser superficial heat treatment of steel cylinders monitored by a high-speed thermal camera in order to build a classification system. This system should analyze the video sequence from the laser heat treatment of a cylinder and decide whether it has been correctly processed before the next production step begins.
However, the task of classifying "normal" and "abnormal" products is complex because major irregularities seldom occur in efficient industrial processes. This is an obstacle to training automated systems with statistical learning models [3]. One-class classification is the anomaly detection technique used in machine learning for binary classification when information on only one of the classes is available. In this scenario, a model of normality is learned and used to assign anomaly scores to previously unseen examples [4]. The larger the anomaly score, the more "abnormal" the example is. Different models have been proposed in the literature for learning the normal behavior of temporal systems. Hidden Markov models (HMMs) are the most used probabilistic graphical models [5]. We propose to use dynamic Bayesian networks (DBNs), which are a generalization of HMMs [6], since we are concerned with providing an interpretable representation of uncertain knowledge [7] avoiding the lack of physical meaning of hidden variables. DBNs, have already proved to successfully describe the spatial-temporal relationships of monitored variables in different domains without the use of hidden variables, e.g., neuroscience for learning the temporal connections of brain regions [8], bioinformatics for inferring the interactions of DNA microarrays [9], or engineering for fault detection of autonomous spacecrafts [10].
Finally, data collection from some industrial applications is expensive. In these cases, sample sizes are small and generalization cannot be guaranteed. To overcome this, we implement sampling and cross-validate estimation techniques.
The paper is organized as follows. Section 2 describes the proposed classification system and explains its theoretical groundwork. The framework is applied to an industrial laser heat treatment process presented in Section 3 and discussed in Section 4. Finally, Section 5 summarizes the conclusions.

## 2 Methodology

This section explains the general aspects of the proposed classification framework for image sequences, and a schematic flowchart is illustrated in Fig. 1.

### 2.1 Image Sequence Preprocessing and Feature Subset Selection

Each frame of the recorded videos can be seen as a feature vector whose values (pixel color information) across the different frames form a multivariate time series. The bits per pixel, which define the range of possible discrete values of a pixel, depend on the properties of the camera. However, this range will always be in excess of the number of categories that discrete statistical models can deal

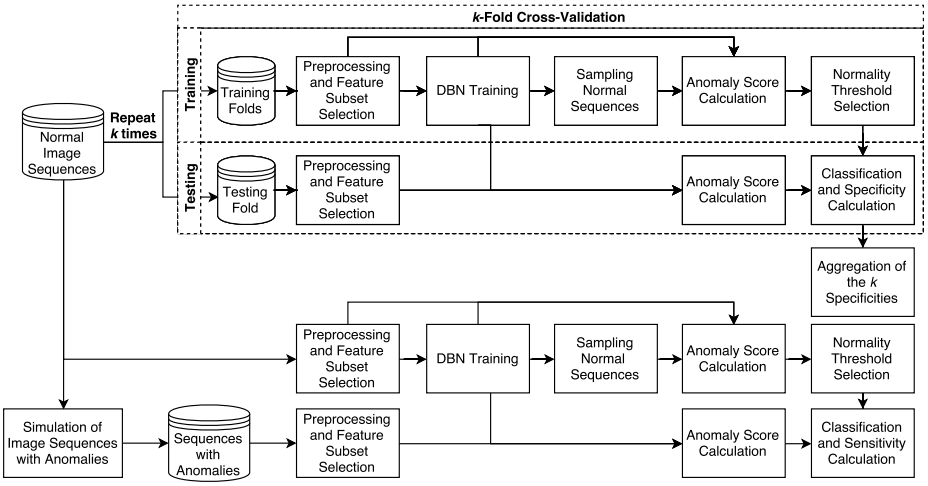


Fig. 1: Schematic flowchart of the classification system.

with. Therefore, they have to be integrated into a number of partitions. The chosen technique was the equal-width interval binning discretization method.
Modeling the time series of every pixel in a video sequence is a high-dimensional problem. Considering that we need an in-process response from the classification system implemented in the CPS, it will be too time consuming and computationally prohibitive to model this problem. Therefore, the number of features has been reduced based on spatial correlations among pixels, where highly correlated pixels are grouped into clusters, i.e., regions with similar properties. These regions are static, and they are the same for all video sequences. We employed hierarchical clustering with an agglomerative strategy as the clustering algorithm to perform this segmentation. This algorithm uses the relationships among data to cluster the features [11]. Apart from the pixel color information in each frame, the algorithm has access to another relationship: each pixel's neighbors within the frame space summarized in a connectivity matrix. The role of this matrix is to avoid the agglomeration of unconnected pixels. In order to select the number of clusters, experts defined a qualitative threshold criterion in terms of the maximum number of clusters in the frame that do not include any artifacts. Artifacts are construed as being pixel clusters that have no physical explanation in regard to the analyzed process, i.e., very small regions or unconnected areas.
Finally, information is extracted from the pixel values of each cluster for each frame. These cluster features will feed the machine learning algorithm.

### 2.2 Dynamic Bayesian Networks and Structure Learning

Bayesian networks (BNs) are probabilistic graphical models that compactly describe the joint probability distribution of a vector of random variables $\mathbf{X} = (X_1, ..., X_n)$ by explicitly representing their probabilistic conditional dependen-

cies. Any variable $X_i$ in a BN is independent of its non-descendants given the value of its parents $\mathbf{Pa}(X_i)$, i.e., nodes at the origin of the arcs pointing to $X_i$.

DBNs, which are temporal extensions of BNs, expand the above concepts to stochastic processes. In this case, the variables of interest are observed sequentially at discrete times. Thus, a joint probability distribution of the possible values of such variables can be built. This distribution could be extremely complex; hence, two hypotheses are normally assumed. The first one is the first-order Markov assumption, where the future is independent of the past given the present. The second assumption is stationarity, i.e., the transition probabilities are independent of time. Under these assumptions, the joint probability distribution of a time series ending at time $T$ can be written as

$$P(\mathbf{X}[0], ..., \mathbf{X}[T]) = P(\mathbf{X}[0]) \prod_{t=1}^{T} P(\mathbf{X}[t] \,|\, \mathbf{X}[t-1]) = \prod_{t=0}^{T} \prod_{i=1}^{n} P(X_i[t] \,|\, \mathbf{Pa}(X_i[t])). \tag{1}$$

This type of DBNs are called two-time-slice Bayesian networks (2-TBNs) since their transition model can be specified by a BN unrolled in two consecutive time slices. Much of the research into static BNs can be applied when learning the structure of discrete 2-TBNs from complete data. The key algorithm proposed by Friedman et al. [12] is based on hill-climbing search (*DHC*). Recently, Trabelsi et al. [13] developed a scalable algorithm called dynamic max-min hill climbing (*DMMHC*), that adapts *MMHC* [14] to consider the temporal dimension.

### 2.3 Anomaly Score, Normality Threshold and Classification

After learning the normality model with DBNs, anomalies are detected as sequences of consecutive frames that are too far removed from normal frames. In anomaly detection, this distance is called anomaly score. Within this framework we define the anomaly score as the log-likelihood of a sequence with respect to the trained normal model.

The anomaly scores of the normal training sequences form a distribution function for which we can define a percentile $\alpha$ for a two-tailed test. $2\alpha$ can be interpreted as the probability of normal sequences that we are willing to mistakenly classify in order to be better able to detect the sequences with anomalies. The two values that limit the rejection regions are the normality thresholds. If they are to be reliably selected, we need a large training set, which we do not have. To overcome this problem, we use the probability distribution defined by the normality model to sample new normal sequences in order to augment the training sample size. Then, the anomaly scores of both the new and the training sequences are calculated and they are used to select the normality thresholds according to $\alpha$.

Now, a new unseen sequence can be classified by calculating its anomaly score. It is abnormal if it falls in the rejection region; otherwise, it is normal.

We define two figures of merit for assessing system performance[1]:

$$sensitivity = \frac{TP}{TP + FN} \qquad specificity = \frac{TN}{FP + TN} \ , \tag{2}$$

where $P$, $N$, $T$ and $F$ account for positive, negative, true and false, respectively. Since the sample size is small, a $k$-fold cross-validation method is defined in order to estimate the specificity of the system. The aim is to guarantee generalization. Normal sequences $\mathcal{D}$ are divided into $k$ groups $\mathcal{D} = D_1 \cup ... \cup D_k$. The normality model $M_i$ is learned with the training folds $\mathcal{D} \setminus \{D_i\}$. Then, the testing fold $D_i$ is used to estimate the specificity $\hat{p}_i$ of the classifier. The mean of the $k$ specificities is $\hat{p}_M$, $M$ being the normality model learned from $\mathcal{D}$.
Regarding the sensitivity calculation, it is required to simulate anomalies in the normal image sequences because this is a one-class classification scenario.

## 3 Experiments

The reason behind the development of this classification system is to perform in-process quality control through the visual inspection of sequences recorded from the radiation produced by the laser superficial heat treatment process of steel cylinders. As proposed by other visual inspection methods, the analysis of these emissions will help to detect defects during laser processes [3], [15].

### 3.1 Data Description

The experimental data were collected by recording the correct superficial laser heat treatment of 32 steel cylinders. The camera used was a NIT TACHYON 1024 $\mu$CORE at a rate of 1,000 frames per second and a FPA resolution of 32×32 (1,024 pixels). A total of 21,500 frames were obtained for each cylinder. Every pixel raw data item was a 10-bit value (ranging from 0 to 1,023) proportional to the temperature reading. The in-process response time for this application is three seconds. Fig. 2(a) shows an example frame of the laser process.
Experts simulated three different defects in the 32 normal videos in order to assess the response of the classification system to anomalies: (i) the spot is programmed within the normal heat treatment process to avoid an obstacle in the cylinders (Fig. 2(b)) and hence videos from obstacle-free processes are abnormal; (ii) the laser scanner control adjusts the energy that the beam deposits, where a power supply unit failure could lead to not enough heat being produced for the process; (iii) the camera is working in hazardous conditions due to heat, sparks and smoke that gradually wear out the sensors, producing noise.

[1] Normal examples are labeled as "negative", while examples with anomalies are labeled as "positive".

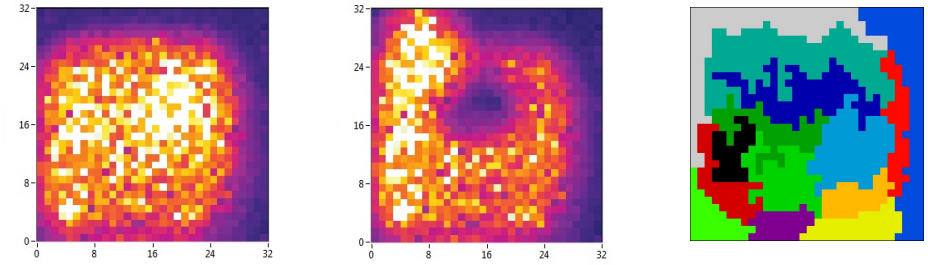


Fig. 2: (a) and (b) are example images of the laser heat treatment process taken by the NIT camera. Specifically, (b) pictures the moment when the laser spot avoids an obstacle. (c) shows the different clusters into which the pixels are divided.

### 3.2 Classification System Experimental Setup

The agglomerative clustering identified 14 artifact-free clusters within the image. The resulting clustered image is shown in Figure 2(c). The clusters adjacent to the edges of the image were discarded because they are considered to be background, leaving nine clusters to be used by the model. Four feature types were extracted from the pixel values of each cluster for every frame: (1) median was chosen because pixel values had a skewed distribution, (2) standard deviation is used to quantify the amount of instability in the region, (3) maximum and (4) minimum values are useful to know the cluster temperature range. The values of these features were discretized into 10 intervals of 102.4 units.
The normality models were learned with both *DHC* and *DMMHC* algorithms so as to assess their performance. They were implemented by extending the bnlearn R package [16]. Constraints were placed on their arcs in order to obtain coherent relationships among cluster variables: arcs between features of the same type are the only ones allowed if the features belong to different clusters, whereas any arc is permitted if it connects features from the same cluster.
The parameter $k$ of the cross-validation method was set to 8. The number of normal sequences sampled from the DBNs was 72. The response of the system was studied for different values of $\alpha$, namely 0%, 1%, 2%, 3%, 4% and 5%.
The first defect (obstacle) was simulated by removing the 4,200 frames that take the laser to avoid the obstacle. The second defect (offset) was simulated by decreasing the pixel values with an offset of 4% of the pixel range (41 units). The third defect (noise) was simulated by adding Gaussian noise to pixel values with mean 0 and a standard deviation of 2.5% of the pixel range (26 units).

### 3.3 Results

Table 1 reports the specificity and sensitivities of the classification system for the different combinations of normality models, and $\alpha$ values.

Table 1: Results for specificity (normal) and sensitivities (obstacle, offset, noise) of the system when learning the normality model with *DHC* or *DMMHC* for different $\alpha$.

| **Normality model** | **Test** | **Specificity and sensitivities for different $\alpha$ values** | | | | | |
|---|---|---|---|---|---|---|---|
| | | *0%* | *1%* | *2%* | *3%* | *4%* | *5%* |
| *DHC* | *Normal* | 100% | 100% | 96.9% | 93.7% | 90.6% | 87.5% |
| | *Obstacle* | 78.1% | 78.1% | 78.1% | 81.2% | 81.2% | 81.2% |
| | *Offset* | 90.6% | 90.6% | 93.7% | 100% | 100% | 100% |
| | *Noise* | 100% | 100% | 100% | 100% | 100% | 100% |
| *DMMHC* | *Normal* | 90.6% | 90.6% | 90.6% | 81.2% | 75.0% | 65.6% |
| | *Obstacle* | 90.6% | 90.6% | 90.6% | 93.7% | 93.7% | 93.7% |
| | *Offset* | 81.2% | 81.2% | 100% | 100% | 100% | 100% |
| | *Noise* | 25.0% | 25.0% | 43.7% | 100% | 100% | 100% |

The normality model learned with *DHC* outperformed *DMMHC* when classifying normal sequences. *DHC* correctly classified 100% of the normal sequences with $\alpha < 2\%$ and maintained a high specificity value (87.5%) for the highest $\alpha$. *DMMHC* was just over 90% with low $\alpha$ and fell to 65.6% with the highest $\alpha$. *DHC* also outperformed *DMMHC* when detecting anomalies produced by offset and noise. Regarding offset, *DHC* had a sensitivity of 90.6% with $\alpha < 2\%$, while *DMMHC* had 81.2%. Both reached 100% with $\alpha > 2\%$. *DHC* detected 100% of sequences with noise, while the best result for *DMMHC* was 43.7% for $\alpha < 3\%$. *DMMHC* only outperformed *DHC* when detecting the absence of obstacles in data with a sensitivity of 90.6% to 93.7%, while *DHC* scored 78.1% to 81.2%. Finally, the proposed methodology met the in-process classification requirement of three seconds with both *DHC* and *DMMHC*.

## 4 Discussion

It is vital in industrial applications to detect sequences with errors (high sensitivity) without triggering false alarms (high specificity). Depending on the specific application, either option could be more important than the other. In this particular laser application, the aim is to reach a balanced trade-off between both measures. The best option then is to use the *DHC* algorithm to learn the normality model of the classification system with $\alpha = 3\%$. This ensures a specificity above 90% with sensitivities greater than 80% for the different types of anomalies.

## 5 Conclusion

We have reported an in-process classification system learned from a small number of anomaly-free examples for detecting anomalies in large video sequences of the laser superficial heat treatment process of steel cylinders.

We are working on implementing this classification system into a CPS for automated visual inspection in order to provide timely feedback about the quality of the process and minimize product failures and waste. To be precise, wrongly processed cylinders will be immediately marked and removed from the production line for later manual inspection.
Additionally, experts are studying the learned DBN structures of normal laser processes in order to gain insight into the thermodynamic and spatial behavior occurring in the region where the laser spot is moving. This should improve the adjustment of different process parameters, e.g., the movement pattern of the spot and its frequency, or the energy that the beam should deposit depending on the position of the spot.

### Acknowledgments

This research has received funding from the Spanish Center for the Development of Industrial Technology (CDTI) as part of project TIC-20150093 and partial funding from the Spanish Ministry of Economics and Competitiveness as part of project TIN2013-41592-P and from Madrid Regional Government as part of project S2013/ICE-2845-CASI-CAM-CM.

## References

1. Díaz, J., Bielza, C., Ocaña, J.L., Larrañaga, P.: Development of a cyber-physical system based on selective Gaussian naïve Bayes model for a self-predict laser surface heat treatment process control. In: Machine Learning for Cyber Physical Systems: Selected papers from the International Conference ML4CPS 2015. pp. 1–8. Springer (2016)
2. Baheti, R., Gill, H.: Cyber-physical systems. In: The Impact of Control Technology, 12, pp. 161–166. IEEE Control Systems Society (2011)
3. Jäger, M., Knoll, C., Hamprecht, F.A.: Weakly supervised learning of a classifier for unusual event detection. IEEE Trans. Image Process. 17(9), pp. 1700–1708 (2008)
4. Chandola, V., Banerjee, A., Kumar, V.: Anomaly detection: A survey. ACM Comput. Surv. 41(3), 15, pp. 1–58 (2009)
5. Barber, D., Cemgil, A.T.: Graphical models for time-series. IEEE Signal Process. Mag. 27(6), 18-28 (2010)
6. Murphy, K.P.: Dynamic Bayesian Networks: Representation, Inference and Learning. Doctoral dissertation, University of California, Berkeley (2002)
7. Bielza, C., Larrañaga, P.: Discrete Bayesian network classifiers: A survey. ACM Comput. Surv. 47(1), 5, pp.1–43 (2014)
8. Rajapakse, J.C., Zhou, J.: Learning effective brain connectivity with dynamic Bayesian networks. NeuroImage. 37(3), pp. 749–760 (2007)
9. Husmeier, D.: Sensitivity and specificity of inferring genetic regulatory interactions from microarray experiments with dynamic Bayesian networks. Bioinformatics. 19(17), pp. 2271–2282 (2003)
10. Codetta-Raiteri, D., Portinale, L.: Dynamic Bayesian networks for fault detection, identification, and recovery in autonomous spacecraft. IEEE Trans. Syst., Man and Cybern. 45(1), pp. 13–24 (2015)

11. Xu, D., Tian, Y.: A comprehensive survey of clustering algorithms. Ann. Data Sci. 2, pp. 165–193 (2015)
12. Friedman, N., Murphy, K., Russell, S.: Learning the structure of dynamic probabilistic networks. In 14th Conference on Uncertainty in Artificial Intelligence. pp. 139–147 (1998)
13. Trabelsi, G., Leray, P., Ayed, M. B., Alimi, A.M.: Dynamic MMHC: A local search algorithm for dynamic Bayesian network structure learning. In: Advances in Intelligent Data Analysis XII. pp. 392-403. Springer (2013)
14. Tsamardinos, I., Brown, L.E., Aliferis, C.F.: The max-min hill-climbing Bayesian network structure learning algorithm. Mach. Learn. 65(1), 31–78 (2006)
15. Alippi, C., Braione, P., Piuri, V., Scotti, F.: A methodological approach to multisensor classification for innovative laser material processing units. In 18th IEEE Instrumentation and Measurement Technology Conference. 3, pp. 1762–1767 (2001)
16. Scutari, M.: Learning Bayesian networks with the bnlearn R package. J. Stat. Softw. 35(3), pp. 1–22 (2010)