Closed-Loop Active Model Diagnosis Using Bhattacharyya Coefficient Application to Automated Visual Inspection

Noom, J.; Nguyen, Hieu Thao; Soloviev, O.A.; Verhaegen, M.H.G.

DOI
10.1007/978-3-030-71187-0_60

Publication date
2021

Document Version
Accepted author manuscript

Published in
International Conference on Intelligent Systems Design and Applications

Closed-Loop Active Model Diagnosis using Bhattacharyya Coefficient: Application to Automated Visual Inspection*

Jacques Noom(✉), Nguyen Hieu Thao, Oleg Soloviev, and Michel Verhaegen

Delft Center for Systems and Control (DCSC),
Delft University of Technology, The Netherlands
j.noom@tudelft.nl

Abstract. This manuscript presents an improvement of state-of-the-art Closed-Loop Active Model Diagnosis (CLAMD). The proposed method utilizes weighted Bhattacharyya coefficients evaluated at the vertices of the polytopic constraint set to provide a good trade-off between computational efficiency and satisfactory input choice for separation of candidate models of a system. A simulation of a dynamical system shows the closed-loop performance not being susceptible to the combination of candidate models. Additionally, the broad applicability of CLAMD is shown by means of a demonstrative application in automated visual inspection. This application involves sequential determination of the optimal object inspection region for the next measurement. As compared to the conventional approach using one full image to recognize handwritten digits from the MNIST dataset, the novel CLAMD-approach needs significantly (up to 78%) less data to achieve similar accuracy.

Keywords: Active fault diagnosis · Model discrimination · Auxiliary signal design · Bhattacharyya coefficient · Machine vision

1 Introduction

Model diagnosis is a key element in automated systems and will play an essential role in the production lines of Industry 4.0. Applications range from fault detection and isolation in dynamical systems [4] to object recognition in images [11], and beyond. In general, passive approaches are adopted, meaning that models are discriminated using observations only. In contrast, active approaches apply an auxiliary input to distinguish models with higher reliability. This auxiliary input can be either precomputed, or determined during operation. Early studies define this as offline and online (experiment) design [24], whereas modern contributions shift to the conventions open-loop and closed-loop [5], respectively.

* This project has received funding from the ECSEL Joint Undertaking (JU) under grant agreement No 826589. The JU receives support from the European Union's Horizon 2020 research and innovation programme and Netherlands, Belgium, Germany, France, Italy, Austria, Hungary, Romania, Sweden and Israel.

In contrast to open-loop, closed-loop designs offer the possibility to tune the input online in order to (further) optimize the model diagnosis. Generally, this leads to less intrusive inputs for obtaining model estimates with similar or higher confidence levels. Humans perform such sequential action-perception process unconsciously in everyday life, yet this form of intelligence is nowadays not widely spread in automation.

Among the first contributions on Closed-Loop Active Model Diagnosis (CLAMD) are [7] for discriminating between two models and [3] for multiple models. Zhang and Zarrop [25] were the first to apply CLAMD in dynamical systems [5] by using the well-known sequential probability ratio test of Wald [20]. Further developments use applications mainly in the detection and isolation of faults, whereas they generally contribute to the underlying theory of discriminating models. Consequently, the term *active fault diagnosis* [5] has been adopted standardly, thereby neglecting the equally relevant problem of discrimination between models instead of just faults. The focus of this paper is on model diagnosis.

Deterministic and stochastic approaches to CLAMD are developed to account for bounded and probabilistic uncertainties, respectively. For bounded uncertainty, a deterministic approach enables guaranteed diagnosis of the correct model (e.g. [13, 15, 22]). For probabilistic uncertainty such guarantee does not exist. Instead, a stochastic (or: probabilistic) approach considers the probability of misclassification. Whereas deterministic approaches anticipate worst-case scenarios, probabilistic approaches "often achieve reasonable accuracy with much less aggressive input signals" [14]. It is more appropriate to formulate CLAMD as a stochastic problem, since explicit bounds on signals are often difficult to define in practice.

A major challenge in CLAMD is to provide a satisfactory input choice, while suppressing computational effort. Although approaches for nonlinear systems are currently being explored (e.g. [17, 18]), approaches for linear systems deserve attention as improvements are still possible and linear system theory is still highly relevant in practice. Further, it should be remarked that linear system theory often serves as a basis for extension to nonlinear systems.

This paper further develops a recent linear stochastic approach in [12], in which the sum of pairwise Bhattacharyya distances is maximized with respect to the input in order to minimize the probability of misclassification. Since the system uncertainties are Gaussian and the input constraints are assumed to be polytopic, the sum of Bhattacharyya distances is convex. It is evaluated at all vertices of the constraint set to find the maximum. The algorithm is implemented using a receding horizon in order to achieve a closed loop. In contrast to the reinforcement learning based linear stochastic approach in [16], no learning stage is required to find a satisfactory input policy. Though computationally efficient, the approach in [12] adversely separates all candidate models simultaneously, disregarding current belief states. The performance hereby becomes susceptible to the combination of candidate models being used. The proposed approach overcomes this vulnerability, while suppressing the computational effort.

The novelty of our approach is to evaluate a *weighted* sum of Bhattacharyya *coefficients* instead, at identical vertices of the constraint set. This sum benefits from large affinity with probability of misclassification, and therefore it is expected to generate more appropriate inputs for discrimination of models. Since the latter method uses a number of evaluations equivalent to the method in [12], the increase in computational effort is only moderate.

The second aim of this paper is to extend the applicability of CLAMD to visual defect detection based on two-dimensional images. Whereas existing 2D-approaches to *passive* fault detection (e.g. [21]) convert a 1D-signal of a *dynamical* system to 2D-images, the application of our *active* approach inherently deals with 2D-images of *static* objects. The auxiliary input defines what regions of the object to inspect. The application closely relates to robotic view planning [23] and sensor management [6, 9], yet it is distinctive by formulating it as a general CLAMD-problem, regarding the views or sensors as inputs to the discrimination problem. Hereby this work facilitates a valuable connection between theory of CLAMD and applications including robotic view planning and sensor management.

The manuscript is structured as follows. First, the methods are presented in Sect. 2 with the state-of-the-art and the proposed approach. Then two simulations of a dynamical system in Sect. 3 and of automated visual inspection in Sect. 4 substantiate the use of the proposed approach. A conclusion is drawn in Sect. 5.

2 Methods

Below, a summary of the approach in [12] for sequentially determining the auxiliary input, follows. Afterwards, the novel approach is presented. The notation used here largely corresponds to the notation in [12], with a concatenated vector $u_{a:b} = [\, u_a^\top \quad u_{a+1}^\top \quad \dots \quad u_b^\top \,]^\top$ or using $\tilde{u}_a = u_{a:a+N-1}$ with prespecified horizon length N, $\|\cdot\|_p$ indicates the ℓ_p-norm, $\|\cdot\|_0$ the ℓ_0-"norm" counting the number of nonzero elements, $P(\cdot)$ the probability of an event, $p(x)$ the probability density function (PDF) of a random variable x, and $x|y$ a random variable x conditioned on y.

Regard the linear time-varying models of a system

$$M_i : \begin{cases} x_{k+1} = A_k^{[i]} x_k + B_k^{[i]} u_k + E_k^{[i]} w_k \\ y_k = C_k^{[i]} x_k + F_k^{[i]} v_k \end{cases} \tag{1}$$

with k the time step and $x \in \mathbb{R}^{n_x}$, $u \in \mathbb{R}^{n_u}$, $y \in \mathbb{R}^{n_y}$, $w \sim \mathcal{N}(0, Q \in \mathbb{R}^{n_w \times n_w})$ and $v \sim \mathcal{N}(0, R \in \mathbb{R}^{n_v \times n_v})$ the state, input, output, Gaussian process noise and Gaussian measurement noise, respectively. The system is operating under one of the models with discrete index $i = \{0, 1, \dots\}$. The goal is to minimise the probability of error given by the Bayes risk using the PDFs $p(\tilde{y}_{k+1|k}|M_i, \tilde{u}_{k|k})$ and the belief states (or: prior probability) $P(M_i|y_{0:k})$. The belief states are

updated using Bayes' theorem [12]. It is the task to minimize the Bayes risk by choosing the optimal input from the constraint set $\tilde{\mathcal{U}}$:

$$\tilde{u}^*_{k|k} = \arg\min_{\tilde{u}_{k|k}\in\tilde{\mathcal{U}}} P(\text{error}|y_{0:k}, \tilde{u}_{k|k}). \tag{2}$$

From [2] we have that

$$P(\text{error}|y_{0:k}, \tilde{u}_{k|k}) \leq \sum_i \sum_{j>i} \sqrt{P(M_i)P(M_j)}\mathfrak{B}_{ij}(\tilde{u}_{k|k}) \tag{3}$$

with

$$\mathfrak{B}_{ij}(\tilde{u}_{k|k}) = \int \sqrt{p(\tilde{y}_{k+1|k}|M_i, \tilde{u}_{k|k})p(\tilde{y}_{k+1|k}|M_j, \tilde{u}_{k|k})}\ \mathrm{d}\tilde{y}_{k+1|k} \tag{4}$$

the Bhattacharyya coefficient between models i and j. For Gaussian process and measurement noise this coefficient equals $\mathfrak{B}_{ij}(\tilde{u}_{k|k}) = \exp\left(-d_{ij}(\tilde{u}_{k|k})\right)$ with $d_{ij}(\tilde{u}_{k|k})$ the convex quadratic Bhattacharyya distance, constructed from the state-space and covariance matrices as illustrated in [12].

Now it can be observed that $\mathfrak{B}_{ij}(\tilde{u}_{k|k})$ is minimized by maximizing the convex quadratic function $d_{ij}(\tilde{u}_{k|k})$. For applying this concept to multiple models, a further relaxation is made by taking the sum of pairwise Bhattacharyya distances $\sum_i \sum_{j>i} d_{ij}(\tilde{u}_{k|k})$ as objective to maximize. Solutions can be found at the border of the constraint set $\tilde{\mathcal{U}}$. Since this "set $\tilde{\mathcal{U}}$ is assumed to be a closed and bounded convex polytope" [12], the latter sum only needs to be evaluated at a finite number of points q to find its global optimum. The resulting optimization reads as

$$\tilde{u}^*_{k|k} = \arg\max_{\tilde{u}_{k|k}} \sum_i \sum_{j>i} d_{ij}(\tilde{u}_{k|k}) \qquad \text{s.t. } \tilde{u}_{k|k} \in \{\tilde{u}^{[1]}, \tilde{u}^{[2]}, \ldots, \tilde{u}^{[q]}\}. \tag{5}$$

As closed-loop implementation this maximization is implemented in a receding-horizon: only the first element of $\tilde{u}^*_{k|k}$ is applied to the system and the input sequence is updated after next measurement.

2.1 Proposed Approach

Instead of evaluating (5) at all vertices of the constraint set, the proposed approach evaluates the right-hand side of (3) at the same vertices. The advantage is that the belief states are now considered in the right proportions regarding (3), instead of completely neglecting those. The result of the proposed minimization

$$\tilde{u}^*_{k|k} = \arg\min_{\tilde{u}_{k|k}} \sum_i \sum_{j>i} \sqrt{P(M_i)P(M_j)}\mathfrak{B}_{ij}(\tilde{u}_{k|k}) \quad \text{s.t. } \tilde{u}_{k|k} \in \{\tilde{u}^{[1]}, \tilde{u}^{[2]}, \ldots, \tilde{u}^{[q]}\} \tag{6}$$

is not necessarily identical to the result of maximization (5) but evidently it is equal or closer to the global optimum of (2). Furthermore, with Gaussian uncertainties $\mathfrak{B}_{ij}(\tilde{u}_{k|k})$ just equals the exponent of $-d_{ij}(\tilde{u}_{k|k})$. Together with

the fact that (6) is evaluated at an equivalent number of vertices as (5), the increase in computational effort is only moderate.

In next section, a proof-of-concept simulation of a dynamical system is presented advocating the use of (6) over (5). Subsequently, the use of (6) is demonstrated in a static imaging system with non-Gaussian uncertainties.

3 Simulation of Dynamical System

As proof-of-concept, the simulation of the dynamical system consists of five candidate models as illustrated in Fig. 1(a). Consider two groups of models. Models 0 and 1 are similar, except for a small change in damping. Models 2, 3 and 4 are again very similar, but now except for a different DC-gain. The constraints on the auxiliary input (which is the only input in this case), are:

$$|u_k| \leq u_{\max} = 2.5, \qquad |u_{k+1} - u_k| \leq \Delta u_{\max} = 2$$

The horizon length for calculating optimal input sequences is $N = 4$, and the prior probabilities at time step $k = 0$ are $P(M_0) = 0.6$ and $P(M_{1:4}) = 0.1$. The system is operating under model M_3. The initial conditions are $x_0 = \begin{bmatrix} 0 & 0 \end{bmatrix}^\top$ and $\Sigma_0 = 0.5 I_2$. Noise covariances are $R = Q = 0.2$.

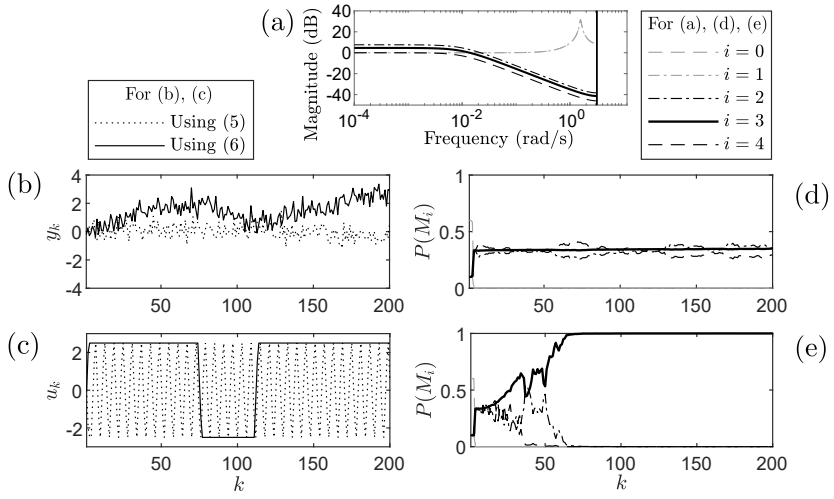

Fig. 1. Simulation of dynamical system with (a) magnitude plots of candidate models. Models M_0 and M_1 have microscopic difference in damping, and models $M_{2,3,4}$ differ in DC-gain. The system operates under model M_3. In (b) is the system output against measurement number, using auxiliary inputs plotted in (c); and in (d) and (e) the belief states using (5) and (6), respectively. Maximization (5) finds an input for separating all models simultaneously, whereas minimization (6) takes into account the belief states, leading to proper separation with significant increase in decision certainty.

The results are plotted in Fig. 1(b-e). The original approach computes the optimal auxiliary input for separating all 5 models simultaneously. It fails to diagnose a (correct) model within 200 time steps. On the contrary, the adapted approach computes the inputs while regarding the current belief states. Once the probabilities for models M_0 and M_1 approximate zero, it starts generating a (nearly) constant input for checking the DC-gain, after which it correctly diagnoses model M_3 in 64 time steps with 98% certainty. The computational time per measurement using an Intel i7-9750H CPU is 0.87 ms and 1.32 ms for the original and proposed approach, respectively. In view of the increase in decision confidence, the increase in computational effort is moderate.

4 Application to Automated Visual Inspection

In this section, two examples illustrate how CLAMD can be applied in automated visual inspection. Both examples rely on choosing the most appropriate region of the object to inspect, based on previous measurements. A *single-image case* diagnoses the sample using one single realization of a class. This can be interpreted as single-image object recognition, in which the images are sequentially inspected at small image subregions which are believed to carry most useful information for classification. It will reduce the required amount of data for classification, hereby making integrated visual inspection less dependent on the limited data transfer rate between camera and computer.

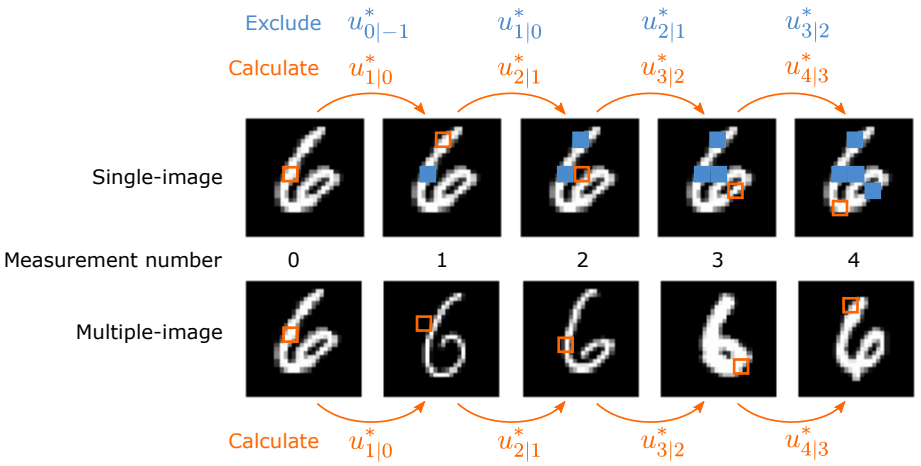

Fig. 2. Illustration of CLAMD in automated visual inspection using a single image (top) and multiple images from the same class (bottom), taken from the MNIST handwritten digit dataset [10]. After each measurement, the optimal next 4×4-pixel inspection region $u^*_{k+1|k}$ is calculated. In the single-image case, previously inspected area is excluded, whereas the multiple-image case can have equivalent inspection regions for different measurement numbers.

In contrast, the *multiple-image case* uses more realizations of the same class for diagnosis. This is useful e.g if an industrial machine produces consecutive products with similar faults, and inspection of an entire object is costly. As more specific example, one could think of inspection of wafers in semiconductor industry, where complete inspection is very time consuming due to mechanical constraints on the visual inspection machine [1]. By inspecting only the small parts of the wafers which are believed to carry most useful information for classification, the inspection time per product is decreased while ensuring classification with high reliability.

Both cases are tested on the MNIST handwritten digit dataset [10], as illustrated in Fig. 2. The 28×28-pixel images are split into 49 small 4×4-pixel regions. Each time step, the CLAMD-algorithms decide which image region to inspect. In the single-image case the previously inspected regions are excluded in future optimizations to prevent double usage of information.

4.1 Formulation of CLAMD-problem

The inspection problem is shaped to a CLAMD-problem as follows. The surrogate measurement model for this problem is

$$y_{k+1} = f(u_k) \tag{7}$$

with u_k binary inputs indicating what regions to inspect, and $f(\cdot)$ a nonlinear function formed by a neural network. The corresponding minimization becomes

$$\begin{aligned} \tilde{u}^*_{k|k} = \ & \arg\min_{\tilde{u}_{k|k}} \sum_i \sum_{j>i} \sqrt{P(M_i)P(M_j)}\mathfrak{B}_{ij}(\tilde{u}_{k|k}) \\ & \text{s.t. } \left\|\tilde{u}_{k|k}\right\|_0 = 1, \qquad \left\|\tilde{u}_{k|k}\right\|_1 = 1. \end{aligned} \tag{8}$$

Since the MNIST dataset does not include any dynamics in its samples, the Bhattacharyya coefficient is only dependent on the prior knowledge of the (non-Gaussian) noise parameters and can be calculated offline using (4). Note that the proposed approach in Sect. 2.1 guarantees global optimality of (8). On the contrary, the method in [12] is only suitable under Gaussian noise and relies on a varying Bhattacharyya coefficient. Therefore, the proposed closed-loop approach will be compared to an open-loop approach instead. This open-loop approach calculates the optimal input sequence offline based on Bhattacharyya coefficients. In the single-image case, it will result in an equivalent sequence for each experiment and in the multiple-image case, it results in constantly inspecting the same region.

For each 4×4-pixel region, a neural network is constructed using the 60 000 training samples of the MNIST handwritten digits dataset, which are normalized and contain noise with variance 10^{-4}. All 49 neural networks have 3 layers with 128, 10 and 1 neurons, respectively. The first two layers have as activation function a rectified linear unit (ReLU). The mean squared error of the 1-dimensional predictions with the training labels is used as loss function. For optimizing the

variables in the neural network, the Adam optimizer [8] is utilized. As reference, for full image classification a similar neural network design is taken using 2 layers with 128 and 10 neurons, respectively, with the first layer having a ReLU activation (TensorFlow 2 quickstart for beginners, in [19]).

Using the same data as for training the neural network, the PDFs are constructed using Gaussian kernel density estimation from the Python-command `scipy.stats.gaussian_kde()`. In some regions e.g. in the corners, the 4×4-pixel regions convey very little information, as those regions are mostly black for all classes. This results in very sharp peaks in the histograms, frustrating determination of accurate PDFs. Therefore, image regions with histograms of which more than 40% of the data is within a neural network output of 0.02, are disregarded for CLAMD. This results in 15 remaining valid inspection regions.

A final decision was made if the belief state satisfies $P(M_i) > 0.98$ for any class i. In the single-image case there is a finite number of regions to inspect. Therefore, if $P(M_i) \leq 0.98$ for all classes i after inspecting all valid regions in the single-image case, the full image is considered for classification.

4.2 Results

The results of the validation test using 10 000 realizations, of both single-image and multiple-image case are given in Table 1 and Fig. 3. In the single-image case, Table 1 shows a moderate decrease in data usage from 784 to 489 and 470 data points (or: pixels) using open- and closed-loop implementation, respectively. However, the accuracy also decreases. A possible reason for this is inaccuracy in construction of PDFs, which can result in erroneous belief states. In the multiple-image case, the decrease in number of data points is considerable with 171 data points for the closed-loop implementation. Furthermore, the accuracy increased to 98.9%. This means that a higher accuracy was achieved in classifying objects using 78.2% less data, compared to inspecting one full image.

The mean decision time for the full 28×28-pixel image is notably less than the open- and closed-loop implementations. They are, however, not directly comparable with each other. Moreover, full image inspection only regards the classification step, whereas the open- and closed-loop designs consider the integrated vision system including data acquisition and transfer, which can be done in a pipeline in parallel with the computations.

Fig. 3 shows that the digits 0 and 1 are in general recognized in fewer measurements than the other digits. This is probably due to their clear spatial dissimilarity with the digits 2 to 9, especially in the middle of the images, where the first measurements are taken. In agreement with Table 1, the closed-loop approach is remarkably favorable in the multiple-image case. The open-loop approach separates all classes simultaneously by taking the single best region for this. Instead, the closed-loop approach only separates the classes of which the belief states are high, which is considerably more efficient for most digits.

The difference in performance increase using CLAMD between the single-image and multiple-image case is so large for two reasons. First, the accessible information is more valuable using multiple images of the digits (i.e. similar

Table 1. The results for single- and multiple-image cases, computed using Intel i7-9750H CPU. In the single-image case, the closed-loop approach results in a moderate decrease in number of data points with respect to the open-loop approach. Whereas the accuracy is lower compared to single full image classification, the open- and closed-loop approaches reduce data consumption with from 784 to 489 and 470 data points, respectively. In the multiple-image case, the closed-loop approach uses only 171 data points, while obtaining a higher accuracy than single-image classification.

	Average #measurements	Average #data points	Mean decision time	Accuracy
Single-image case				
Full image	1	784	0.758 ms	97.8%
Open-loop	12.4	489	43.2 ms	95.2%
Closed-loop	11.8	470	50.9 ms	95.4%
Multiple-image case				
Open-loop	37.4	598	136 ms	96.3%
Closed-loop	10.7	171	56.8 ms	98.9%

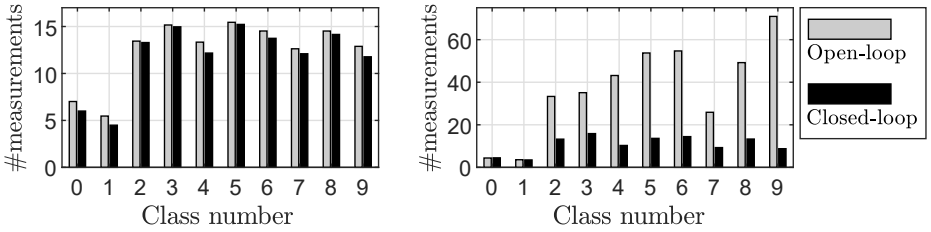

Fig. 3. Average number of measurements per class number, for the single-image (left) and the multiple-image case (right), using open-loop and proposed closed-loop approach, respectively. Whereas the advantage of closed-loop diagnosis in the single-image approach is only moderate, it becomes more prevailing for the multiple-image case. In specific, for the digits with high spatial similarities (i.e. 2 to 9) the number of measurements decreases significantly using the closed-loop method.

regions can be inspected in multiple images). Second, inspection of only one region as being done in the open-loop multiple-image case, is clearly inadequate for recognizing most digits. Altogether, this experiment advocates the use of CLAMD in a multiple-image case for classification using little data.

5 Conclusions

This paper improves the state-of-the-art in CLAMD and shows its broad applicability. A simulation of a dynamical system shows the existence of certain cases with which the proposed CLAMD-approach can deal, while a recently developed approach fails in discriminating the models. Moreover, it is expected that the proposed approach either equals, or outperforms the existing approach in all cases. The CLAMD-approach is applied to automated visual inspection, in which most informative object regions are sequentially inspected based on most recent measurements. In recognizing the MNIST handwritten digits, a reduction in data usage of 78% is achieved, while maintaining accuracy with respect to inspecting one full image using a similar neural network design. This makes the CLAMD-approach particularly suitable for efficiently handling costly and/or time-consuming measurements in inspection problems.

Future perspectives count on development of efficient algorithms for handling more types of constraints, such as bounding the input energy. Possible applications include, but are not limited to medical imaging systems in which exposure to radiation should be minimized while information gain should be maximized. CLAMD enforces artificial agents to thoroughly interpret reality by acting on it, just like humans do. A fine interpretation enhances reliability, efficiency and intelligence. With the goal of Industry 4.0 to increase both throughput and fidelity, implementation of CLAMD seems indispensable for efficiently picking the right information for accurate product examination.

References

1. Aryan, P., Sampath, S., Sohn, H.: An overview of non-destructive testing methods for integrated circuit packaging inspection. Sensors **18**(7), 1981 (2018)
2. Blackmore, L., Williams, B.: Finite horizon control design for optimal discrimination between several models. In: Proceedings of the 45th IEEE Conference on Decision and Control, pp. 1147–1152 (2006)
3. Box, G.E., Hill, W.J.: Discrimination among mechanistic models. Technometrics **9**(1), 57–71 (1967)
4. Gao, Z., Cecati, C., Ding, S.X.: A survey of fault diagnosis and fault-tolerant techniques—part I: Fault diagnosis with model-based and signal-based approaches. IEEE Transactions on Industrial Electronics **62**(6), 3757–3767 (2015)
5. Heirung, T.A.N., Mesbah, A.: Input design for active fault diagnosis. Annual Reviews in Control **47**, 35–50 (2019)
6. Hero, A.O., Cochran, D.: Sensor management: Past, present, and future. IEEE Sensors Journal **11**(12), 3064–3075 (2011)

7. Hunter, W.G., Reiner, A.M.: Designs for discriminating between two rival models. Technometrics **7**(3), 307–323 (1965)
8. Kingma, D.P., Ba, J.L.: Adam: A method for stochastic optimization. In: 3rd International Conference for Learning Representations (2015)
9. Lauri, M.: Sequential decision making under uncertainty for sensor management in mobile robotics. Ph.D. thesis, Tampere University of Technology (2016)
10. LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. Proceedings of the IEEE **86**(11), 2278–2324 (1998)
11. Lowe, D.G.: Object recognition from local scale-invariant features. Proceedings of the IEEE International Conference on Computer Vision **2**, 1150–1157 (1999)
12. Paulson, J.A., Heirung, T.A.N., Braatz, R.D., Mesbah, A.: Closed-loop active fault diagnosis for stochastic linear systems. In: Proceedings of the American Control Conference, pp. 735–741 (2018)
13. Raimondo, D.M., Braatz, R.D., Scott, J.K.: Active fault diagnosis using moving horizon input design. In: 2013 European Control Conference, pp. 3131–3136 (2013)
14. Scott, J.K., Marseglia, G.R., Magni, L., Braatz, R.D., Raimondo, D.M.: A hybrid stochastic-deterministic input design method for active fault diagnosis. In: Proceedings of the IEEE Conference on Decision and Control, pp. 5656–5661 (2013)
15. Shen, Q., Yong, S.Z.: Active model discrimination using partition-based output feedback input design. In: 2020 European Control Conference, pp. 712–717 (2020)
16. Škach, J., Punčochář, I., Lewis, F.L.: Optimal active fault diagnosis by temporal-difference learning. In: 2016 IEEE 55th Conference on Decision and Control, pp. 2146–2151 (2016)
17. Škach, J., Straka, O., Punčochář, I.: Efficient active fault diagnosis using adaptive particle filter. In: 2017 IEEE 56th Annual Conference on Decision and Control, pp. 5732–5738 (2017)
18. Straka, O., Punčochář, I.: Decentralized and distributed active fault diagnosis: Multiple model estimation algorithms. International Journal of Applied Mathematics and Computer Science **30**(2), 239–249 (2020)
19. TensorFlow: Large-scale machine learning on heterogeneous systems (2015), https://www.tensorflow.org/
20. Wald, A.: Sequential analysis. John Wiley & Sons, Inc., New York (1947)
21. Wen, L., Li, X., Gao, L., Zhang, Y.: A new convolutional neural network-based data-driven fault diagnosis method. IEEE Transactions on Industrial Electronics **65**(7), 5990–5998 (2018)
22. Yang, S., Xu, F., Wang, X., Liang, B.: A novel online active fault diagnosis method based on invariant sets. IEEE Control Systems Letters **5**(2), 457–462 (2021)
23. Zeng, R., Wen, Y., Zhao, W., Liu, Y.J.: View planning in robot active vision: A survey of systems, algorithms, and applications. Computational Visual Media **6**(3), 225–245 (2020)
24. Zhang, X.J.: Auxiliary signal design in fault detection and diagnosis. Lecture Notes in Control and Information Sciences **134**, Springer-Verlag, Berlin (1989)
25. Zhang, X.J., Zarrop, M.B.: Auxiliary signals for improving on-line fault detection. In: 1988 International Conference on Control, pp. 414–419 (1988)