Towards real-time in-situ monitoring of hot-spot defects in L-PBF: a new classification-based method for fast video-imaging data analysis

Matteo Bugatti[1] · Bianca Maria Colosimo[1]

Abstract

The increasing interest towards additive manufacturing (AM) is pushing the industry to provide new solutions to improve process stability. Monitoring is a key tool for this purpose but the typical AM fast process dynamics and the high data flow required to accurately describe the process are pushing the limits of standard statistical process monitoring (SPM) techniques. The adoption of novel smart data extraction and analysis methods are fundamental to monitor the process with the required accuracy while keeping the computational effort to a reasonable level for real-time application. In this work, a new framework for the detection of defects in metal additive manufacturing processes via in-situ high-speed cameras is presented: a new data extraction method is developed to efficiently extract only the relevant information from the regions of interest identified in the high-speed imaging data stream and to reduce the dimensionality of the anomaly detection task performed by three competitor machine learning classification methods. The defect detection performance and computational speed of this approach is carefully evaluated through computer simulations and experimental studies, and directly compared with the performance and computational speed of other existing methods applied on the same reference dataset. The results show that the proposed method is capable of quickly detecting the occurrence of defects while keeping the high computational speed that would be required to implement this new process monitoring approach for real-time defect detection.

Keywords Image-based process monitoring · In-situ defect detection · Machine learning · Neural network · Laser Powder Bed Fusion (L-PBF)

Abbreviations

AER	Actual error rate
AM	Additive manufacturing
DL	Deep learning
HS	Hot-spot
IC	In-control
LHZ	Laser heated zone
L-PBF	Laser Powder Bed Fusion
KM	K-means
ML	Machine learning
NN	Neural network
OOC	Out-of-control
PCA	Principal component analysis
ROI	Region of interest
SPM	Statistical process monitoring
SVM	Support Vector Machine

✉ Matteo Bugatti
matteo.bugatti@polimi.it

Bianca Maria Colosimo
biancamaria.colosimo@polimi.it

[1] Dipartimento di Meccanica, Politecnico di Milano, via La Masa 1, 20156 Milan (MI), Italy

Introduction

In recent years, the increasing availability of low-cost machine vision systems and the advances in computational capabilities for image and video processing have pushed the adoption of these systems for change detection and process monitoring. High space and temporal resolution data streams from machine vision systems have found their application also in metal additive manufacturing (AM) process monitoring (Everton et al. 2016; Grasso and Colosimo 2017; Mani et al. 2017; Spears and Gold 2016; Tapia and Elwany 2014). The layer-by-layer manufacturing mechanism of AM allows to monitor the process at a very high detail level, i.e. during the production of each layer up to melt pool dynamics level (at thousands fps). Correlation between observable features

(proxies) and final defects must then be carefully evaluated in order to implement a robust defect detection strategy. AM process monitoring is expected to be one of the key features of the new generation of AM machines (Grasso and Colosimo 2017) to limit the process variability that burdens this technology since its birth. In the last few years, machine builders (e.g. Renishaw, Trumpf) and independent companies (Sigma Labs) have started to implement monitoring sensors on industrial machines and to develop robust monitoring strategies for defect detection.

When the level of detail required for the analysis is high, e.g. high spatial and temporal resolution of video images, fast data processing capabilities and efficient data handling algorithms are needed to make sense of the big data stream. This may not be fundamental when monitoring is applied to analyze the data off-line to provide relevant information for post-process inspection, but it becomes crucial when in-line monitoring is implemented, as the big data stream needs to be processed and analyzed in real-time to really unlock the true potential of this kinds of systems, i.e. onset defect detection for closed-loop process control or at least alarm raising to stop the build and reduce the waste.

For this reason, in this work particular attention was put on data handling and on assessing the real-time applicability of the developed process monitoring method: the new approach aims at detecting the onset of overheating, a.k.a. hot-spot, phenomena during the L-PBF process by analyzing the big data stream coming from the high-speed videos acquired during laser scanning.

The paper is organized as follows:

- in "State of the art" section, a short state-of-the-art review on statistical process monitoring (SPM) with image data is reported.
- in "Case study and experimental setup" section, the case study on which the proposed method was tested is presented together with the experimental setup used for data acquisition.
- in "Methodology" section, the developed methodology is described starting from the big data handling strategy to the final defect localization and detection methods based on machine learning.
- in "Discussion of results" section, a critical discussion of the defect detection performance obtained with the developed methodology and a comparison with other state-of-the-art approaches on the same case study are reported. In "Computational cost, sensitivity analysis and realtime applicability" section the real-time applicability of the developed algorithm is quantitatively assessed and a simple strategy for faster defect detection is proposed. In "Simulation study" section a simulation study is performed to further test the three competitor ML classification methods on artificially injected defects.

- in "Conclusion" section, the final conclusions are drawn and a possible research path with some ideas on how to improve the results is tracked for future work on this topic.

State of the art

The most straightforward and easy approach to apply statistical process monitoring (SPM) to image data is to extract synthetic information from the images using computer vision techniques and study the in-control variability of the extracted dataset with standard control charting techniques. This is extremely effective when the quantities that can be extracted from the acquired images are accurate and give a good representation of what needs to be monitored, e.g. dimensions or other product properties measurable via machine vision techniques (Horst and Negin 1992; Lyu and Chen 2009; Nembhard et al. 2003; Tong et al. 2005; Wang and Tsung 2005; Park and Shrivastava 2014).

Scanning statistics is another simple way to deal with anomaly detection and is based on dividing the images into regions and monitoring any set of low-dimensional features (Li et al. 2013; Megahed et al. 2012; He et al. 2016) to detect both the spatial location of a defect and the change-point in time within the image stream.

Other anomaly detection schemes applied to images are kernel and basis representation methods, which model the in-control space with a certain representation prior, e.g. sparsity, and detect anomalies by judging the distance of new observations from that prior depending on their kernel or basis representation. This method has found many applications in quality control (Carrera et al. 2016) but the computational time required to compute a basis representation of a 2D matrix, i.e. an image, is usually not compatible with the speed required for real-time analysis of high-speed videos and even less suitable when a 3D matrix representation needs to be computed to assess the temporal structure of the anomaly.

All these methods focus on process monitoring for image data, where the images used during the learning phase are a sequence of uncorrelated random replicates of an in-control pattern. Anomalies related to temporal dynamics rather than simple image analysis still represent a challenge for most of the current state of the art techniques and require *ad hoc* extensions to monitor the temporal evolution of the phenomenon that is being observed. In addition, when dealing with a complex, highly dimensional dataset, several authors Nti et al. (2021), Mahato et al. (2020) and Bai et al. (2019) have underlined the importance of pre-processing via dimensionality reduction to extract a synthetic representation of the dataset and to analyze it more efficiently with conventional machine learning techniques. However, when applied to temporally and spatially auto-correlated datasets (i.e. videos),

dimensionality reduction can lose valuable information about the temporal and spatial structure of the observed anomalies. This problem has been addressed by the researchers who proposed new approaches based on principal component analysis (PCA) on video imaging. In their works Celik (2009), Grasso et al. (2016) and Colosimo and Grasso (2018) they have applied the dimensionality reduction capabilities of PCA to the complex spatio-temporal dataset, often combining it with clustering algorithms for defect localization. One interesting extension of these PCA-based algorithms is the spatially weighted version proposed by Colosimo and Grasso (2018) to consider the spatial and temporal correlation structure in video images: this approach makes the PCA-based algorithm more suitable for the analysis of locally correlated datasets.

Finally, neural network (NN) and especially deep learning (DL) architectures are becoming more and more popular nowadays thanks to the impressive results achieved with their image analysis and classification capabilities (Redmon and Farhadi 2018), but there is still debate on what is the best way to feed temporal information to a DL architecture (Lipton et al. 2015) and on how to achieve the same good results with shallower, and thus faster, networks able to keep up with the typical acquisition rate of high-speed video imaging. Nevertheless, in the last few years the first examples of application of neural networks to AM process monitoring have been developed, demonstrating their automatic feature extraction capabilities and classification performance when applied to in-situ process data (Kwon et al. 2020; Li et al. 2020; Gonzalez-Val et al. 2020). However, most of them focus on classifying the static observations, i.e. single images, with respect to a known process parameter level rather than focusing directly on defect detection. One relevant exception is the work of Gonzalez-Val et al. (2020), who focused on the slower direct energy deposition (DED) process and developed a convolutional NN for defect detection in single track welds starting from on-axis MWIR images of the melt pool. This study represents the first promising result for in-process defect detection using NN. Unfortunately, most of the works involving DL or NN in general for process monitoring, do not report the speed of their network for evaluating new observations: this is a crucial point, especially for fast dynamics mPBF processes, and should be taken into consideration before developing a deep and computationally expensive NN.

To overcome the spatial, temporal and computational limitations of the methods available in literature and to match the stringent requirements related to the fast image analysis application described in "Case study and experimental setup" section, a hybrid computer vision and machine learning (ML) algorithm is proposed in "Methodology" section which reframes the temporal dynamics anomaly detection problem into a computationally less expensive classification task. In fact, the approach proposed in this paper aims at developing a region of interest (ROI) based method to reduce the dimensionality, and thus the computational effort, of the temporal anomaly detection problem in a 3D dataset, i.e. video, by extracting the relevant spatio-temporal information only from ROIs identified in static frames with computer vision techniques. The dimensionality reduction obtained in the data extraction step significantly simplifies the anomaly detection task of the implemented ML methods and can then be combined with the ROI position information for precise defect localization.

Case study and experimental setup

Metal additive manufacturing processes allow to produce parts with shapes and features that are otherwise impossible to make with conventional technologies, but part complexity often hides manufacturing challenges that can potentially lead to process defects, with a consequent detrimental effect on the final characteristics of the part. One particular process defect, the hot-spot, can occur when the laser beam is repeatedly focused on a thermally insulated region, i.e. areas mostly surrounded by powder (e.g. overhanging walls, acute corners, etc.). This would result in a very localized overheating that leads to the following defects or inhomogeneities:

- high surface roughness: excessive heat input can lead to melting unwanted areas of powder and partially melted powder particles attach to the surface, thus increasing its roughness;
- change in microstructure of the material: normal melting zone are characterized by a high cooling rate that leads to finer grain formation, while overheating regions tend to develop a coarser microstructure due to the slower cooling transient;
- porosity formation: if the region is already hot, new laser scans may lead to material vaporization, hence unstable keyhole formation which is often correlated with porosity.

To detect any out-of-control (OOC) behavior like hot-spot phenomena, the process should be monitored with sensors which are able to evaluate the fast cooling dynamics that occur during the process. High-speed imaging is potentially suitable to characterize fast thermal phenomena related to the laser-material interaction. An example of an in-situ monitoring setup proposed in Grasso et al. (2016) is shown in Fig. 2. It consists of an Olympus I-speed 3 camera (CMOS sensor) placed outside the build chamber viewport. A sampling frequency of $f = 300$ fps was selected as a compromise between the capability of capturing the laser kinematics while keeping computational cost compatible with the in-
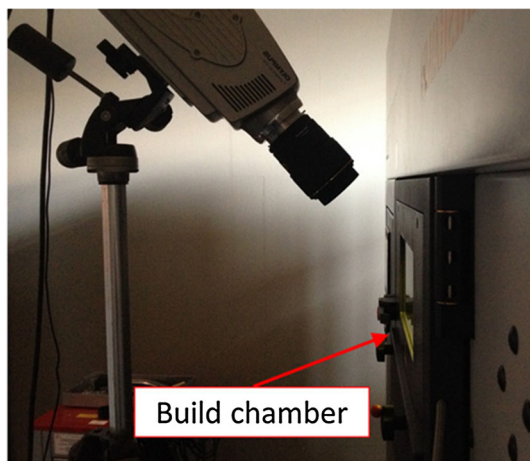
process image analysis. Using this experimental setup for monitoring, a complex shape of about 50 × 50 × 50 mm (Fig. 2a) produced via L-PBF of AISI 316L powder (average particle size of about 25–30 μm). Different hot-spot events occurred during the process in correspondence of acute corners belonging to overhanging areas shown in Fig. 2b. Figure 2c shows that these hot-spot events produced local geometrical deformations and increased roughness in the printed part. Further details are provided in Grasso et al. (2016).

Three high-speed videos (8-bit grey-scale images) were acquired during the L-PBF of different slice geometries in three consecutive layers. Figure 7 shows these triangular features belonging to the sliced CAD model of the case study geometry and indicates the acute corners where the geometrical defect was found. The three corresponding image streams were denoted as OOC scenarios 1, 2 and 3, respectively. After acquisition, since the regions of interest corresponding to the scanned areas represent only a portion of the whole image, a crop operation was performed to extract only the ROIs. The resulting image size for each of the three OOC scenarios was 121 × 71 pixel.

In the next sections, a new method for hot-spot in-situ detection and localization is described together with a critical discussion on its real-time applicability for future implementation of closed-loop process control strategies.

Methodology

The proposed approach reframes the problem of process monitoring for hot-spot localization and detection into a standard classification framework that exploits the brightness evolution of the pixels inside the bright regions detected in each frame to find areas with an anomalous cooling rate, i.e. the hot-spots.
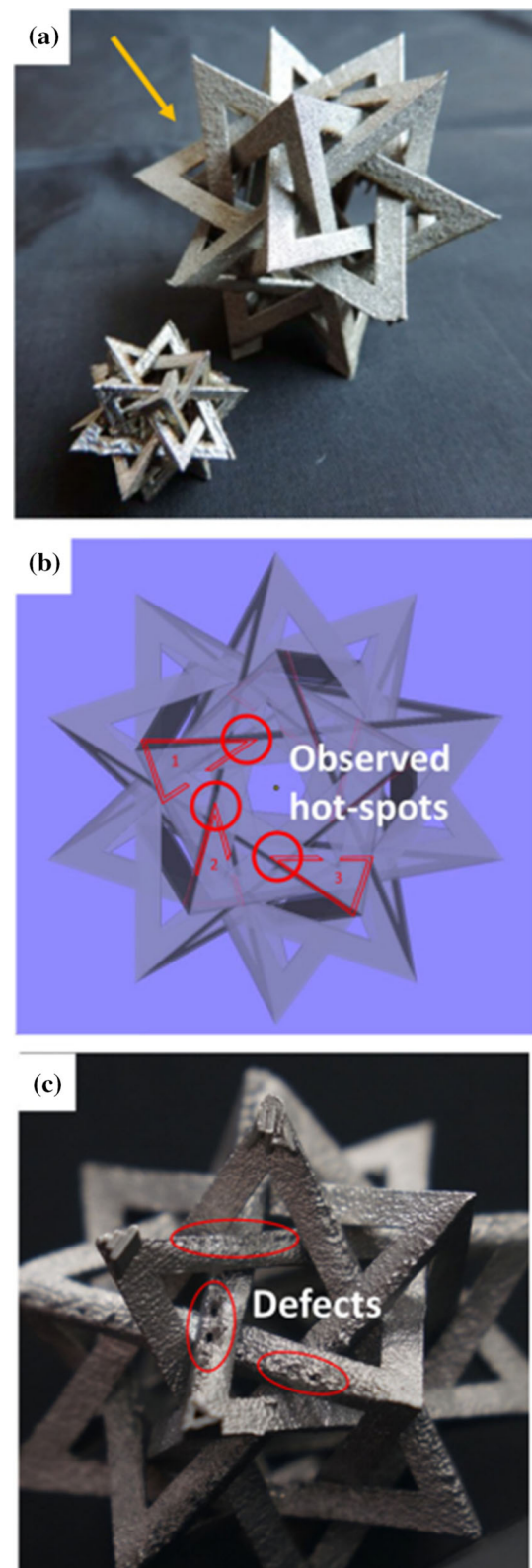

Fig. 2 a Complex shape part used to test the proposed approach; b examples of triangular portions of the sliced CAD model; c local defects corresponding the acute corners of those triangles. (Grasso et al. 2016)

Each classification technique implemented in this work works on the same dataset and the dataset extraction method from the raw data is the core idea of the presented framework. Normal (i.e. laser, spatters) and defect-related (i.e. hot-spots) bright regions often coexist in the same frame, they are very similarly shaped but their brightness evolution is different: the idea is to exploit the fast dynamics of normal bright regions to correctly separate them from the defective bright regions, which exhibit much slower dynamics. In order to do this, an efficient region-based data extraction algorithm was implemented to get synthetic information about the bright regions' dynamics. The main steps of the data extraction algorithm are reported in the following and displayed in Fig. 3:

1. *Thresholding* simple image thresholding is performed to identify bright regions (laser, spatters and hot-spots) in each frame. For this step, an arbitrary brightness level threshold was set to twice the background level (~ 200).
2. *Region isolation* the pixels inside each identified bright region are isolated.
3. *Data extraction* the mean brightness of the pixels in the isolated region is extracted from the L following frames. The resulting time series are the training/testing dataset of the classification techniques described in the following subsections "Unsupervised classification—k-means functional data clustering" to "Supervised classification—neural network". In this study, a $L = 10$ frames long mean brightness history was extracted for the analysis.

After the synthetic data extraction step, different machine learning (ML) algorithms have been implemented to test their capability at distinguishing between normal and abnormal brightness decay history. The advantages of this technique with respect to previous state of the art literature on this topic are the:

- *Synthetic representation of complex dataset and dimensionality reduction* it is not necessary to store the raw data, i.e. the images, to perform defect detection; with this method the high flow of data is filtered and the amount of information that needs to be processed to efficiently perform the defect detection task is significantly reduced.
- *Region-based instead of pixel-based hot-spot detection* the use of regions of pixels instead of single pixels simplifies the defect localization task, as it will correspond with the centroid of the identified bright region. In addition, each region brightness evolution collects information about a group of connected pixels, increasing the overall robustness of the algorithm with respect to considering the single pixel brightness history for classification.
- *Easy hot-spot localization* by handling pixels as a region, the hot-spot localization is a simple consequence of its

detection because the position of each bright region is stored together with the synthetic functional output.
- *Scalability* the proposed approach can also be easily extended to handle higher frame rates and higher resolution images as the increased computational demand only scales linearly with the number of frames. Further details about the computational speed of the presented algorithm and comparison with other already published hot-spot detection strategies is reported in "Discussion of results" section.

In the next subsections, the machine learning classification methods that leverage the functional output of the data extraction algorithm will be described and discussed, outlining their pros and cons, while the classification performances are reported in "Discussion of results" section.

Unsupervised classification—k-means functional data clustering

K-means (KM) clustering for functional data is the functional version of the popular k-means classification algorithm used in multivariate statistics. Let $X = \{x_1, x_2, \ldots, x_n\}$ be a given functional dataset of size n to be analyzed, where x_i belongs to $\mathbb{R}^m$, and $V = \{v_1, v_2, \ldots, v_c\}$ be the functional set of cluster centers, where c is the number of clusters and v_i belongs to $\mathbb{R}^m$. KM iteratively computes cluster centroids in order to minimize the sum with respect to the specified performance measure (e.g. mean square error). KM algorithm aims at minimizing an objective function known as the squared error function given as follows:

$$J_{KM}(X; V) = \sum_{i=1}^{c} \sum_{j=1}^{n_i} D_{ij}^2 \tag{1}$$

where D_{ij}^2 is the squared chosen distance measure which can be any p-norm:

$$D_{ij} = \|x_{ij} - v_i\|^p \tag{2}$$

with $1 \leq i \leq c$, $1 \leq j \leq n_i$ and where n_i represents the number of data points in ith cluster. For c clusters, KM is based on an iterative algorithm minimizing the sum of distances from each observation to its cluster centroid. The observations are moved between clusters until the sum cannot be decreased any more. For the implementation of this technique, the presence of 2 clusters which can be distinguished according to the different brightness decay history, i.e. normal and defective (hot-spot) bright region, is assumed during the training phase. The results of the training phase are 2 functional data centroids associated to normal, fast-decaying bright regions and defective, slow-decaying hot-spot regions (see Fig. 4).
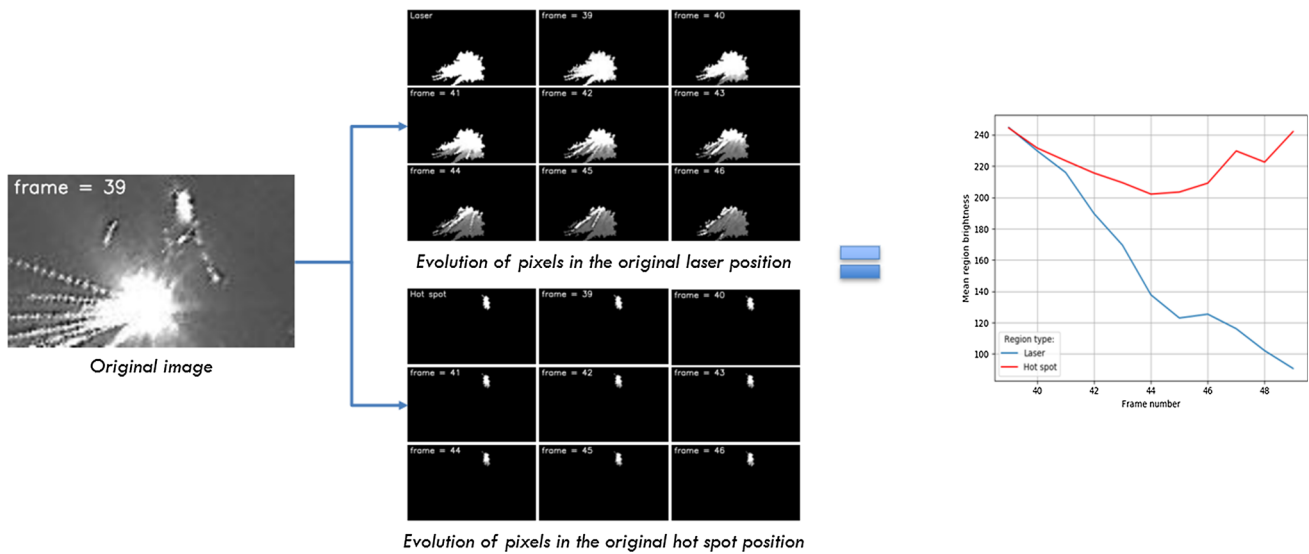
Fig. 3 Graphical representation of the algorithm and of its functional output
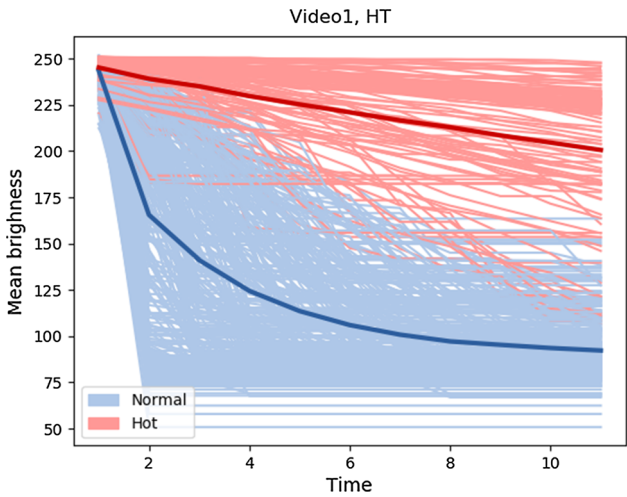

Fig. 4 Training phase: Normal and Hot functional centroids (solid colors) and assigned classes (Color figure online)

To improve the classification performance, an additional pre-processing step is performed on the brightness decay history to prevent the mean brightness to increase back again after the initial decay:

$$b_{adj}(t+1) = min(b(t), b(t+1)) \qquad (3)$$

where $b(t)$ is the mean region brightness at frame t. This rule imposes a never increasing trend in the mean region brightness history, which filters out some of the possible false indication deriving from the potential overlapping of successive bright regions (e.g. due to laser rescanning) with the currently analyzed region in one of the successive frames.

Supervised classification—Support Vector Machine

The second classification method implemented for this purpose is a 2-class Support Vector Machine (SVM). The principle of SVM for classification consists in the definition of the optimal separating hyperplane between the 2 classes (see Fig. 5). One of the main advantages of SVM is that no assumption on data probability distribution within classes is required, thus making it applicable to highly dimensional but small datasets (with few observations). Even though some efforts are being made to apply SVM to functional data, most common codes and routines apply SVM to simple multivariate datasets. For this reason, the disadvantage of applying SVM is the need to perform an additional feature extraction step to gather a finite set of variables from the functional data themselves. In particular, for each functional dataset, the following features have been extracted:

- Mean gradient

$$\overline{\Delta_1 b} = \frac{1}{L-1} \sum_{t=1}^{L-1} \Delta_1 b(t) = \frac{1}{L-1} \sum_{t=1}^{L-1} (b(t+1) - b(t))$$

$$(4)$$

- Maximum mean brightness drop between consecutive frames

$$\Delta_{1,max} b = \max_{1 \le t < L} \Delta_1 b(t) \qquad (5)$$

The upside of this approach consists in the possibility of adding other discrete information about the bright region to the multivariate dataset: in this case, the shape and size of the
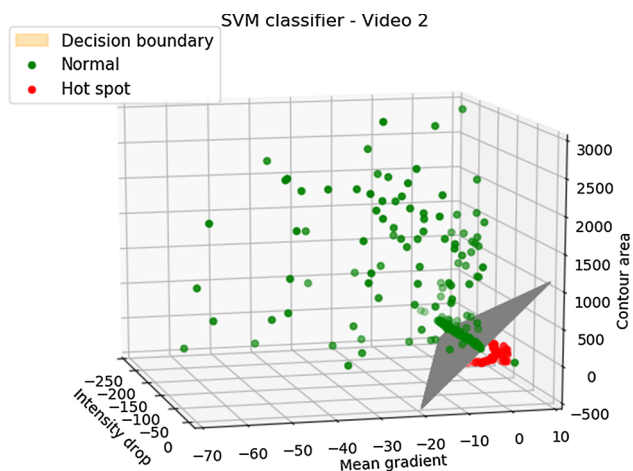
Fig. 5 Visualization of the optimal separating hyperplane between the two classes (linear kernel)

regions have been included into the dataset to enrich it with additional information.

The other main disadvantage about this ML classification method is about the supervised nature of the algorithm itself. To find the optimal separating hyperplane between 2 classes, one of the inputs of SVM must be the ground truth, i.e. the true class of each bright region. To feed this unknown set of information to the SVM algorithm without introducing too much human bias for discerning between normal and abnormal bright regions, all bright regions whose centroid lied in the acute corner area that was found defective after final inspection have been labelled as hot-spots.

For the implementation of each supervised classification method, the dataset composed by 3 videos was split into training and testing, based on their acquisition order. The first video, *OOC Scenario 1* from layer 164, was used for the training phase. The second and third videos, *OOC Scenario 2* and *OOC Scenario 3* from layer 165 and 166, were used for testing. Further splitting of the training set into training and testing subsets was done to tune all the SVM hyperparameters based on their resulting classification performance (kernel, regularization parameter (cost) and tolerance).

Supervised classification—neural network

The last classification method implemented in this work is a fully connected neural network (NN). NN is a very flexible machine learning technique that can be employed for either regression or classification: in our case, the NNs have been implemented as a supervised classification algorithm that finds the optimal non-linear combination of input variables to distinguish between 2 or more classes. Just like SVM, also NN do not require any assumption on data probability distribution within classes. The main advantage of NN over SVM is the feature extraction capability that we have tested by

giving as input the functional data extracted from the video. Therefore, instead of training the classifier on extracted features, like for SVM, this input allows to fully leverage the feature extraction capabilities of NNs, but on the other hand we have no control over the feature extraction performed by the NN in the hidden layer(s). To have a fair comparison with SVM, also NN takes as input additional information about the shape and size of the identified bright regions. Just like any other supervised classification algorithm, also NNs need to be fed with true class labels during the training phase. The same approach described in "Supervised classification— Support Vector Machine" section was employed. Different architectures, i.e. number of hidden layers and size, have been tested and have been treated as hyperparameters of the ML technique to achieve the best classification performance in terms of accuracy. Figure 6 shows the final NN architecture and the activation functions used in the different layers. For training all NN classifiers, the Adam optimization algorithm was used on the binary cross entropy loss function.

Discussion of results

The results achieved with the methods described in "Methodology" section are discussed in "Classification results and comparison study" section together with a comparison with other state of the art approaches reported in literature on the same study case. "Computational cost, sensitivity analysis and realtime applicability" section presents an analysis on the real-time applicability of the presented methods with a focus on computational cost.
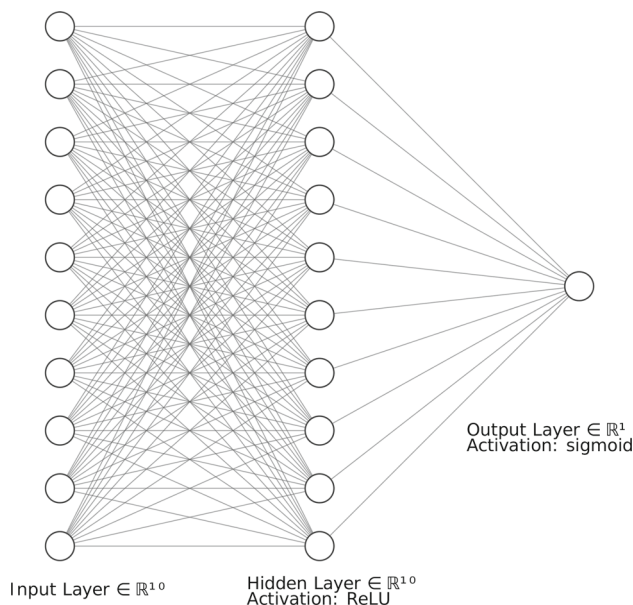

Fig. 6 Final NN architecture

Classification results and comparison study

Since the description of hot-spot is not uniquely defined, the classification performance of each method will be assessed by correlating the classification with the real defect position: if the centroid position of the region classified as defective lies in the area where the geometrical defect was found, i.e. the overhanging acute corner, the classification is considered correct. The following performance indicators are used for comparison:

- Time of first signal: frame of the first observation classified as hot-spot.
- Number of misclassifications: number of hot-spots identified outside the defective region.

To count the number of misclassifications, the standard confusion matrix is considered (Table 1), in which the predicted labels are compared with position-based labels whose value depend on the position of the centroid with respect to the defective corner (see Fig. 7). Since heat accumulation phenomena require a sufficient amount of heat to develop, hot-spots do not appear immediately at the beginning of the layer. To apply the position-based rule to define the ground truth while avoiding the inclusion of mislabelled observations, the first n_i frames of each video i were neglected. The number of frames to discard was selected according to the time, measured in frames, required to complete the first laser scanning phase corresponding to the slice contouring. Only labelled observations, i.e. observations extracted after the end of the contouring phase, were employed for training and testing the classifiers (Table 2).

The proposed methods were compared against three alternative approaches developed by Colosimo and Grasso in their previous works on the same hot-spot detection problem (Grasso et al. 2016; Colosimo and Grasso 2018). The first competitor represents the most intuitive and simple method to detect out-of-control states in video images. Since the resulting average intensity of pixels belonging to hot-spot regions is higher, the monitoring is performed on the mean pixel intensity map of the image stream. For each pixel, the mean intensity over the J observed frames is computed, $\bar{U} = \left\{ \bar{u}_{m,n} = (1/J) \sum_j u_{m,n,j} \right\}$, and a simple clustering-based alarm rule is applied to the final average intensity map $\bar{U}$.

The second and the third competitors are the PCA-based approaches, namely the T-mode PCA and the ST-PCA. The framework of both techniques is briefly reported in the following:

1. the image stream is rearranged into a multivariate dataset;
2. spatio-temporal weighting of the multivariate dataset (only for ST-PCA) is performed;

3. principal component analysis (PCA) is performed on the resulting multivariate dataset (simple or spatially weighted)
4. the first m PCs that capture at least 80% of the overall variability are kept;
5. Hotelling's $T^2(m, n)$ is computed and re-mapped onto the original frame shape;
6. cluster-based alarm rule is applied to the final $T^2(m, n)$ statistic map.

As clearly stated in the step sequence reported above, the ST-PCA is a spatially weighted version of the T-mode PCA which has been developed to characterize the temporal auto-correlation of pixel intensities over sequential frames while including the spatial information related to the pixel location within the image.

These techniques have been originally tested on the complete image stream, i.e. the full video, but two different approaches have also been developed to make these techniques more applicable for real-time:

- *recursive updating* the multivariate dataset increases its size as new frames are acquired. The side effect of this approach is that it may create problems with the computational time as the size of the multivariate dataset gets bigger and more difficult to handle.
- *moving windows* the multivariate dataset size is kept fixed and while new frames are added, the oldest ones get deleted.

For more details on these two techniques please refer to the original publications (Grasso et al. 2016; Colosimo and Grasso 2018). To keep a reasonable computational time, only the moving window updating scheme will be used for this comparison, with a default moving window length of $L = 50$ frames. To allow for a direct comparison between the proposed techniques and the competitors, all methods were applied only on the part of the *OOC Scenario* videos which were used for training and testing, i.e. all frames after laser contouring.

The results of the implemented methods are shown in Table 3 and are directly compared with the previously described competitor approaches. Please note that SVM and NN results for *OOC Scenario 1* are not reported because that dataset was used for training of these two supervised classifiers, as discussed in "Methodology" section.

All presented classifiers are able to detect the hot-spots and outperform both the *Average intensity* and the PCA-based methods in terms of detection speed, but their accuracy is generally lower. To this regard, all methods would benefit the implementation of additional alarm rules to add robustness to the monitoring method at the cost of some identification speed: for example, a location-based alarm could be trained

Table 1 Confusion matrix, position-based versus classification-based prediction

| | | Position-based prediction | |
		Normal	Hot-spot
Classif. prediction	Normal	True negatives	False negatives
	Hot-spot	False positives	True positives
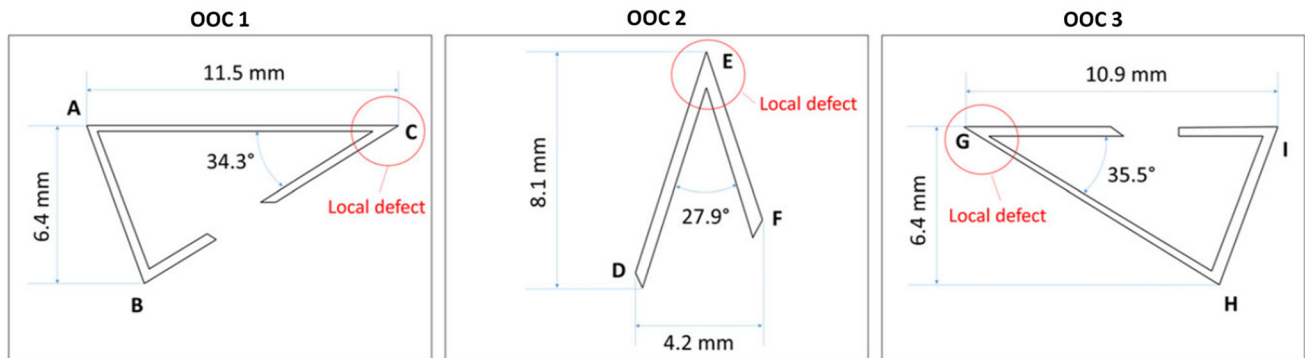
Fig. 7 Dimensions of the triangular portions of the part slices monitoring in this study (complex-shape part); the area of the defective corners used for the position-based labelling are highlighted in red (Color figure online)

Table 2 Dataset description

	OOC Scenario 1	OOC Scenario 2	OOC Scenario 3
Format	8-bit grayscale video	8-bit grayscale video	8-bit grayscale video
Framerate [fps]	300	300	300
Dimensions	$121 \times 71p$		
Nr. of frames	329	219	336
Total nr. of observations	414	324	367
Contouring time [frames]	60	35	56
Nr. of observations after contouring	307	239	278

to raise an alarm only after n anomalous bright regions are detected in the same area.

The confusion matrices reported next to each iteration of the implemented methods reveal that, for each OOC Scenario, less than 7% of the classifications are wrong. One main source of misclassification can be identified, which is related to the overlapping between the characteristics of normal and hot-spot regions. In fact, the hypothesis of distinguishing between normal and defect-related bright regions based on the brightness dynamics discussed in "Methodology" section might be incorrect in some cases because the potential overlap between the features of normal regions and hot-spots can lead to misclassifications. Visualization of classification results in the original videos allows to see that the main condition that leads to misclassification problems is related to laser rescanning (see Fig. 8). This happens when the analyzed region is near the border, or in general any position close to two or more shortly repeating laser tracks: this condition results in a slowly decaying brightness history or in a sequence of brightness decay and increase, which can

only be partially filtered out using the approach described in "Unsupervised classification—k-means functional data clustering" section. When laser rescanning occurs, the fundamental hypothesis on which the whole data-extraction algorithm is based, i.e. fast-dynamics normal versus slow-dynamics defective regions, is not valid anymore and normal laser heated zones (LHZ) can be confused with hot-spots. Figure 8 highlights a region classified as hot-spot outside of the defective area (orange contour): the root cause of the misclassification is the continuous rescanning of the laser in the subsequent frames, which keeps the mean brightness level high as if it were defect-related.

One possible way to increase the capability of the presented approach would be to combine the classifiers prediction into one. Since all classifiers are trained on slightly different datasets, their combination can result in a more robust implementation of the classification-based hot-spot detection method. Table 4 shows the results obtained by using together more than one classifier at a time: a hot-spot alarm is raised only when all the considered classifiers agree on

Table 3 Hot-spot detection results

	Approach	Time of first signal (frame index)	Signaled defect location (X,Y) (pixels)	Classif. prediction		Position-based prediction	
						Normal	Hot-spot
OOC Scenario 1 (frames 60 to 329)	KM	74 (+10)	(109, 18)	Classif. prediction	Normal	205	10
					Hot-spot	8	84
	SVM	OOC1 used for training					
	NN	OOC1 used for training					
	Average intensity	No detection	–				
	T-mode PCA	198	(104.25, 52.71)				
	ST-PCA	75	(108.7, 52.8)				
OOC Scenario 2 (frames 35 to 219)	KM	38 (+10)	(73, 12)	Classif. prediction	Normal	132	4
					Hot-spot	12	91
	SVM	38 (+10)	(73, 12)	Classif. prediction	Normal	141	5
					Hot-spot	3	90
	NN	38 (+10)	(73,12)	Classif. prediction	Normal	132	4
					Hot-spot	12	91
	Average intensity	No detection	–				
	T-mode PCA	No detection	–				
	ST-PCA	92	(74.81, 55.98)				
OOC Scenario 3 (frames 56 to 336)	KM	93 (+10)	(30, 16)	Classif. prediction	Normal	167	9
					Hot-spot	11	91
	SVM	93 (+10)	(30,16)	Classif. prediction	Normal	178	9
					Hot-spot	0	91
	NN	78 (+10)	(19,11)	Classif. prediction	Normal	166	6
					Hot-spot	12	94
	Average intensity	173	(33.88, 55.73)				
	T-mode PCA	168	(31.19, 56.25)				
	ST-PCA	153	(29.21, 55.75)				

For the new algorithms' *Time of signal*, 10 frames should be added because they all use a 10 frames brightness history to perform the classification
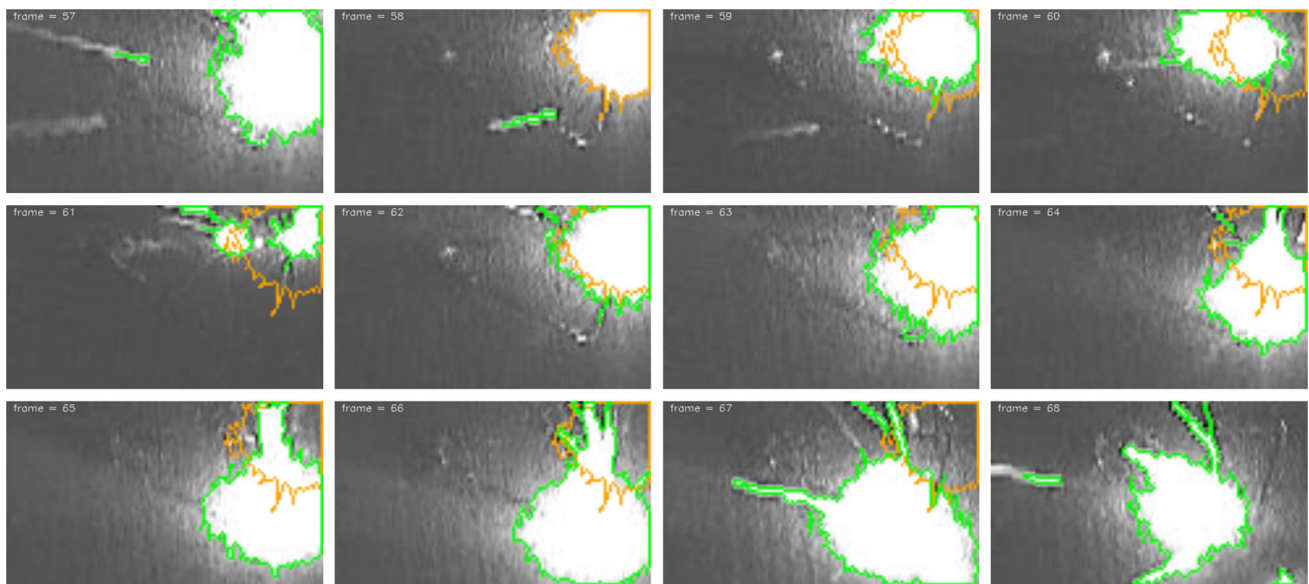
Fig. 8 Extract of OOC Scenario 3 with classified bright regions (frames 57 to 68). Example of false positive (orange contour) due to laser rescanning signaled at frame 58, other bright normal regions are displayed with a green contour (Color figure online)

the defect detection. In most cases, the use of more than one classifier at a time allows to filter out almost all false positives, while keeping a faster detection speed with respect to all competitor approaches. In "Conclusion" section a few other ideas on how to limit this problem are discussed.

Computational cost, sensitivity analysis and real-time applicability

Real-time applicability is a crucial point for all AM-related monitoring methods as the fast process dynamics require both a very high data acquisition rate and algorithms that are able to keep up with the big data stream to provide useful information during the AM process itself and to support, in the future, the implementation of process control strategies. For this reason, the computational cost of the proposed method was accurately tested on a laptop equipped with an Intel i7-8550U CPU and 16 GB of RAM. The results shown in this section are intended as wall time required for the analysis and region classification of one video frame. All the frames from the videos of *OOC Scenarios 2* and *3* were analyzed and the mean wall time required for each of the steps outlined in Fig. 9 was extracted and reported in the attached table.

The classification-based methodology presented here allows for an unparalleled fast analysis of hot-spot phenomena, with a speed which is approximately two times the acquisition speed of the proposed case study (300 fps). The fastest implementation of the PCA-based methods described in "Classification results and comparison study" section requires ≈ 8 seconds to analyze each batch of 50 frames. Figure 9 reports the computational cost of the algorithm and

splits it into two macro tasks: *Data extraction* and *Classification*, with the 3 alternative classification methods previously discussed. Once trained, all classifiers are extremely fast at classifying the observations obtained in the data extraction step and they only require less than 3% of the total time needed for the analysis of each frame. The most computationally expensive task of the whole algorithm is the functional data extraction step, which accounts for more than 94% of the total algorithm time. This step involves image manipulation, masking and mean pixel value extraction. Additional code optimization efforts should be put on this step to cut its execution time.

For what concerns the real-time applicability, one should note that the algorithm is based on the extracted mean brightness time series, therefore it should be considered that classification of bright regions detected at frame n will only occur after the $(n + 10)$th frame is acquired and the algorithm for data extraction and classification is run. In this specific case, considering the 300 fps video frame rate and the algorithm wall time (≈ 2 ms), the hot-spot is detected ≈ 35 ms after its first appearance into the frame. This value may not seem too high but, if we consider a standard 1000 mm/s laser speed, it means that the classification is made after the laser has completed a 35 mm track, or 7 5 mm parallel scan lines, if a standard raster pattern is considered. This may not be enough as, during that time, the laser may have already rescanned a potential hot-spot region. For this reason, further time reduction strategies should be explored.

Overall, the algorithm time accounts for very little of the total time required for detection, therefore a significant improvement can only be achieved by decreasing the number of

Table 4 Classification performance of combined classifiers: hot-spot detection in non-defective region dropped significantly

	Approach	Time of first signal (frame index)	Signaled defect location (X,Y) (pixels)	Classif. prediction		Position-based prediction	
						Normal	Hot-spot
OOC Scenario 2 (frames 35 to 219)	KM+SVM	38 (+10)	(73, 12)	Classif. prediction	Normal	141	5
					Hot-spot	3	90
	KM+NN	38 (+10)	(73, 12)	Classif. prediction	Normal	137	5
					Hot-spot	7	90
	SVM+NN	38 (+10)	(73, 12)	Classif. prediction	Normal	141	5
					Hot-spot	3	90
	KM+SVM+ NN	38 (+10)	(73, 12)	Classif. prediction	Normal	141	5
					Hot-spot	3	90
OOC Scenario 3 (frames 56 to 336)	KM+SVM	93 (+10)	(30, 16)	Classif. prediction	Normal	178	10
					Hot-spot	0	90
	KM+NN	93 (+10)	(30, 16)	Classif. prediction	Normal	174	9
					Hot-spot	4	91
	SVM+NN	93 (+10)	(30,16)	Classif. prediction	Normal	178	9
					Hot-spot	0	91
	KM+SVM+ NN	93 (+10)	(30, 16)	Classif. prediction	Normal	178	10
					Hot-spot	0	90

Step	Duration (ms)	TIMELINE (ms)		
Data extraction				
Thresholding	0.003			
Contour detection	0.045			
Contour-based mask creation	0.010			
Functional data extraction	1.600			
Features extraction (SVM only)	0.010			
Classification			**Total (ms)**	**Speed (fps)**
KM	0.018		1.676	596.7
SVM	0.005		1.673	597.7
NN	0.050		1.708	585.5

0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0 1.1 1.2 1.3 1.4 1.5 1.6

Fig. 9 Breakdown structure of the algorithm steps with their wall time expressed in milliseconds

frames required for classification (this would also incidentally decrease the computational cost related to functional data extraction). A sensitivity analysis of the three classifiers against the time series length L is shown in Fig. 10, where the qualitative Actual Error Rate (AER) as a function of L is displayed. All tests have been performed on *OOC Scenarios 2* and *3*, and the total AER is computed from a weighted average of the AER_k from the k considered OOC scenarios:

$$AER = \sum \frac{1}{obs_k} AER_k$$
$$= \sum \frac{1}{obs_k} \frac{\sum_i \sum_j C_{k,ij} - \sum_i C_{k,ii}}{\sum_i \sum_j C_{k,ij}} \qquad (6)$$

where C_k is the confusion matrix and obs_k is the number of observations, i.e. regions, analyzed in the kth OOC scenario. Considering the qualitative classification performance metric and the random initialization of NN parameters, we can conclude that supervised methods perform roughly the same, and longer time series only seem to offer a marginal improvement over very short time series for data extraction: this means that the total time required for single frame classification can be further reduced to 10–15 ms. As expected, both supervised classification methods perform better than KM, as the lack of input ground truth results in a longer time series required to get to a reasonable level of accuracy. The steady decrease in the AER may hint that KM would benefit even longer region mean brightness time series.

Another potentially viable option is to increase the frame rate up to the algorithm speed limit ($\approx$ 600 fps) because the shorter exposure time would allow for more precise localization of LHZs and hot-spots, canceling out some of the blurriness and the light trails due to spatters observed in some frames of the analyzed videos, but this possibility has not been explored in this work.
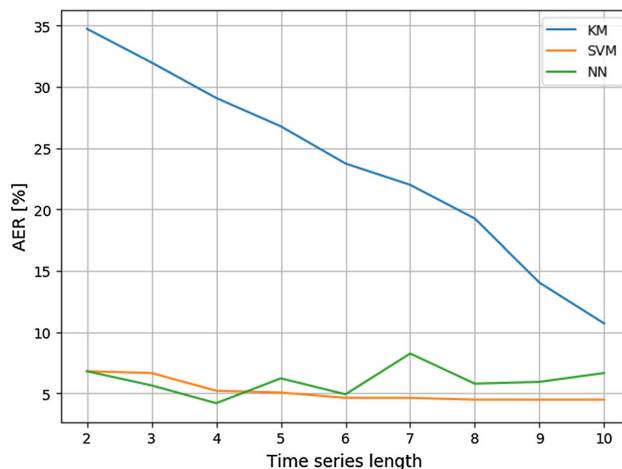

Fig. 10 Sensitivity analysis of the proposed classification methods

Simulation study

To further investigate the capabilities of the presented method, a simulation study was performed by randomly injecting user defined hot-spots in a new high-speed video of the process, right after the laser scan. The video employed for this study records at 150 fps the scanning of a cylindrical shape of diameter 16 mm and it was acquired using the same setup shown in Fig. 2. Since no anomalous behavior was observed and the final part quality was considered acceptable, this video was deemed to be representative of an in-control (IC) process condition, therefore it could be modified to verify the performance of the methods.

To simulate the hot-spot phenomena, a sigmoid function was used to modify the pixel brightness history of a number k of adjacent pixels around the centroid of the LHZ detected at frame i.

$$u_{m,n,j}(\tau) = \frac{255}{1 + exp(0.2(j - 0.95\tau))} \tag{7}$$

where u is the grayscale level, m, n are the pixel position indices, $j = i + 1, \ldots, i + \tau$ is the frame index and τ is the simulated hot-spot duration. This function allows to accurately capture the real cooling behavior of defective pixels, as shown in the comparison reported in Fig. 11. For this study, three levels of hot-spot size were analyzed by varying the number of adjacent pixels k with a sigmoid cooling history, i.e. $k = 9$ (small), $k = 20$ (medium) and $k = 45$ (large) (see Fig. 12). Each hot-spot size was studied in separate simulation runs, in which hot-spots were randomly injected in 100 different locations and with a variable duration $\tau \in [1, 50]$.

The results obtained with the three trained methods are shown in Fig. 13. The pictures show the number of correctly detected hot-spots within 10 frames after their injection for each hot-spot size and duration combinations. KM method, which relies only on the brightness history for classification, shows a very similar trend for each of the 3 hot-spot sizes and fails to detect artificial hot-spots with a duration $\tau < 8$. The reason for this is twofold: (i) the profiles of the short hot-spots obtained by applying the sigmoid function are closer to the normal cooling behavior and (ii) the peak brightness of short hot-spots quickly falls below the threshold employed

for bright regions detection, therefore, even if the hot-spot is present in the analyzed frame, the underlying thresholding method on which all three ML methods rely for dataset extraction fails at detecting a bright region to analyze (see Fig. 14)

SVM classifier results show a similar trend to KM but it performs slightly better at detecting shorter duration hot-spots. This confirms the lower sensitivity of the SVM-based classifier to time series length already discussed in "Computational cost, sensitivity analysis and realtime applicability" section, which means that only the first few entries of the mean brightness time series are enough to extract relevant information for classification. Also NN-based classifier confirms its ability of detecting very short hot-spots but, unlike its KM and SVM counterparts, it was found to be more inclined to raise false alarms, i.e. hot-spots detected far from the hot-spot centroid. Table 5 offers a more detailed overview of the performances of the three methods, giving insights about the detection speed, accuracy and false positives. Any HS detected in the unmodified image stream of the IC video was counted as false positive, while for the simulation study, any HS detected with an Euclidean distance from the injected HS higher than 5 pixels was considered a false positive. In fact, before starting with the simulation study, the IC image stream was analyzed with the three different methods to check for

Fig. 11 Example of pixel-intensity times series in the presence of a real hot-spot in OOC scenario 1 (left panel) and example of pixel-intensity time series with simulated hot spot of duration $\tau = 75$ (right panel). (Colosimo and Grasso 2018)
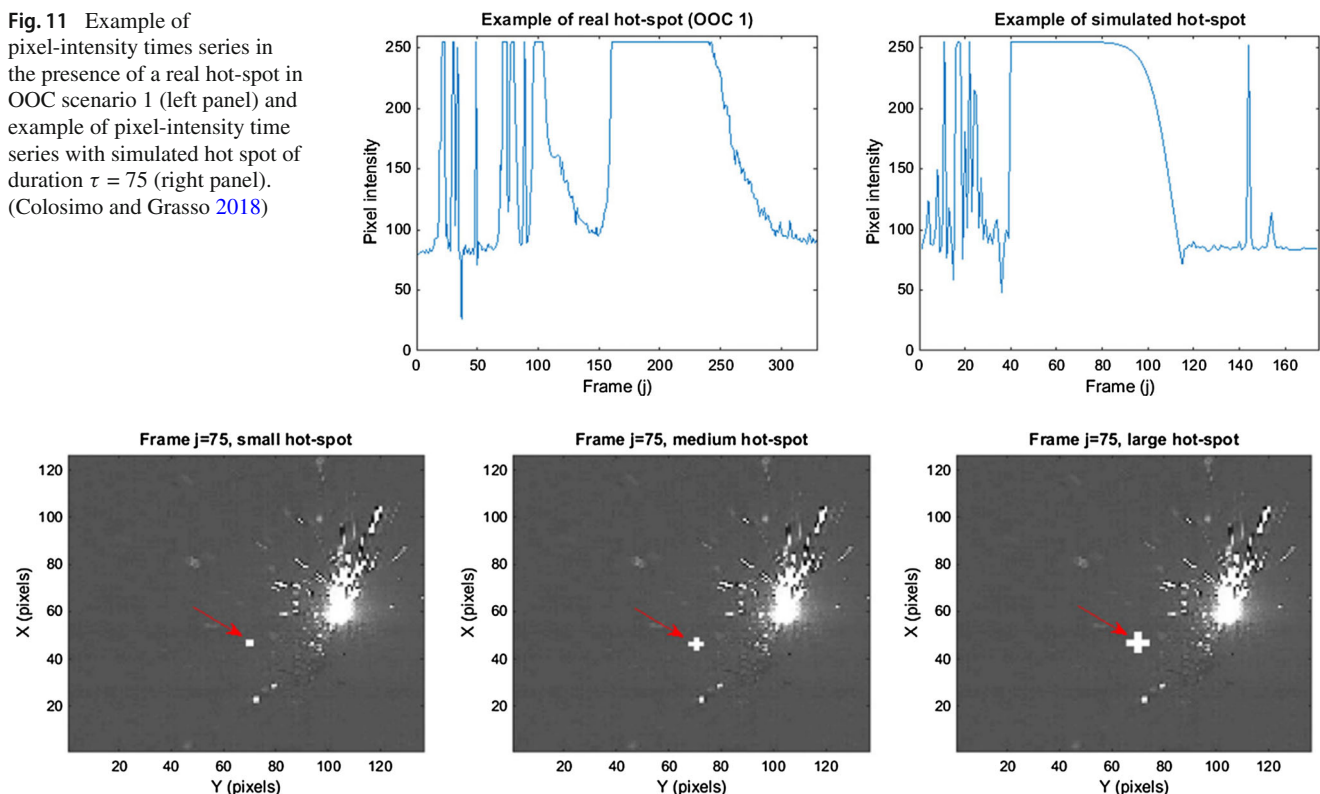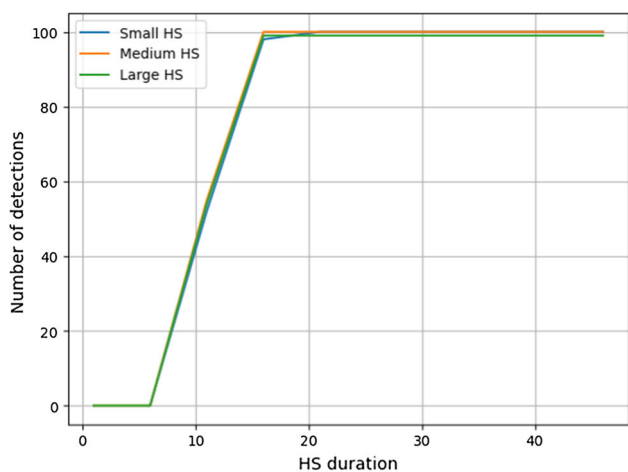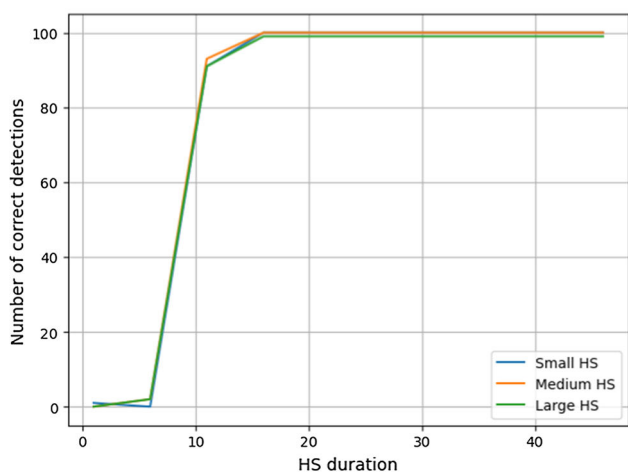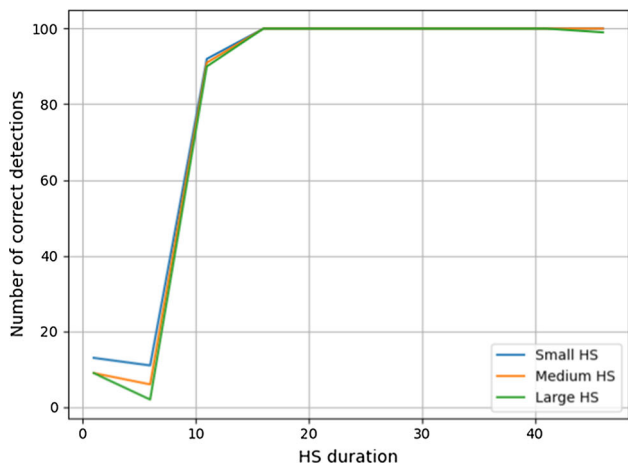

Fig. 12 Examples of one single frame where hot-spot events of different sizes were injected (arrows indicate the spatial location of the hot-spot) (Colosimo and Grasso 2018)

(a) KM

(b) SVM

(c) NN

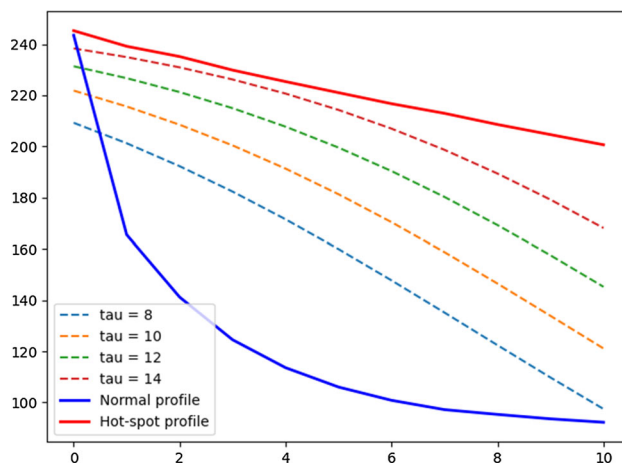**Fig. 13** Comparison of simulation results obtained with the three pre-trained classifiers

Fig. 14 Comparison between short duration hot-spots profiles obtained from Eq. 6 for different levels of τ (dashed lines) and trained K-Means computed profiles for normal and hot-spot (solid lines)

false alarms: from this test, only KM and SVM-based classifiers resulted in 0 detected HS, while NN classifier raised a significant number of false alarms. All methods seem to perform very similarly in terms of HS detection time and accuracy but the SVM and NN-based classifiers are characterized by a higher false positive rate in the simulation study compared to KM. However, it is worth pointing out that all the false positives detected in the simulation study are concentrated in the short duration HS simulations ($\tau < 8$). In fact, when there is a lack of brightness decay information, the SVM and NN algorithms tend to rely on size and shape information for HS classification, thus leading to a higher sensitivity but also to less accurate detections with respect to their size-agnostic, KM-based counterpart.

Conclusion

The new data extraction strategy based on direct image analysis developed in this paper allowed to implement a monitoring strategy based on machine learning to observe local overheating phenomena in laser powder bed fusion processes. The high computational efficiency of this approach enables fast processing of the data coming from a high-speed imaging system and it opens new possibilities for the application of real-time monitoring to process phenomena with very fast dynamics.

Three machine learning classification techniques have been trained to detect hot-spot defects in high-speed videos starting from the synthetic functional dataset extracted from the regions of interest that appear during the video. The methods were then tested on a real case study in laser powder bed fusion of a complex geometry and compared with other state of the art competitor approaches.

Table 5 Simulation results

	KM	SVM	NN
No. of HS detected in the IC video	0	0	17
Avg. detection time [frames]	1.70	1.73	1.74
Avg. distance first closest [px]	0.03	0.03	0.12
Avg. distance first [px]	0.51	0.40	1.31
No. of HS detected far from real HS position	0	10	139

A detailed computational cost analysis revealed that the presented algorithm, with all the classifier combinations, can process data 2 times faster than the video acquisition rate, and up to 80 times faster than the state of the art PCA-based methods used for comparison. In addition, the developed methodology can be used for a frame-by-frame analysis, compared to the other methods that are based on the analysis of frames batches to perform their defect detection. The scalability of this method is also improved with respect to the competitor methods as it can be extended to analyze higher resolution images with no significant increase in computational effort.

A sensitivity analysis of the classifiers' performance against the extracted time series length hints that the elapsed time between data acquisition and defect detection can be further reduced by training and running the classifiers on an even more synthetic dataset extracted from the videos.

An overall accuracy between 93 and 97% was observed for all the 3 machine learning classification methods but, despite being faster in defect detection compared to all competitor approaches, some obvious false alarms are raised, resulting in a sub-optimal performance of the algorithm. Future research will be aimed at further improving the classification performance, by tackling the misclassification with the addition of other rules, which can be added on top of classification models at the cost of some detection speed, or novel data fusion techniques to combine multiple sources of information and improve accuracy. For example, the following ideas may be implemented:

- *Classifiers combination* as discussed in "Classification results and comparison study" section, combining the prediction from multiple classifiers trained on different datasets allows to add robustness to the whole framework and to reduce the number of hot-spots detected outside of the defective region.
- *Location-based alarm* more than *n* bright regions identified in the same area and classified as hot-spot shall be present before an alarm is raised. This would reduce the number of false alarms at the cost of some detection speed, depending on *n*.
- *LHZs filter* as most classification errors correspond to misclassified LHZs, the fusion of high-speed video data

and real-time laser position would lead to a massive simplification of the analysis because it would allow to automatically discard from the analysis any bright spot found in correspondence with the laser heated zone (LHZ) and to focus only on the classification of the remaining bright regions, i.e. spatters and hot-spots.

In addition to the implementation of some of the approaches discussed above, further work may address the extendibility of the method to other types of defect and other image acquisition setups (e.g. thermal cameras).

In conclusion, this approach has proved its efficacy in hot-spot detection task and its proven real-time applicability will enable its application to in-situ process monitoring and, in the future, process control to mitigate the frequency of hot-spot phenomena in complex geometries.

References

Bai, Y., Sun, Z., Zeng, B., Long, J., Li, L., de Oliveira, J. V., et al. (2019). A comparison of dimension reduction techniques for support vector machine modeling of multi-parameter manufacturing quality prediction. *Journal of Intelligent Manufacturing*, *30*(5), 2245–2256.

Carrera, D., Manganini, F., Boracchi, G., & Lanzarone, E. (2016). Defect detection in SEM images of nanofibrous materials. *IEEE Transactions on Industrial Informatics*, *13*(2), 551–561.

Celik, T. (2009). Unsupervised change detection in satellite images using principal component analysis and *k*-means clustering. *IEEE Geoscience and Remote Sensing Letters*, *6*(4), 772–776.

Colosimo, B. M., & Grasso, M. (2018). Spatially weighted PCA for monitoring video image data with application to additive manufacturing. *Journal of Quality Technology*, *50*(4), 391–417.

Everton, S. K., Hirsch, M., Stavroulakis, P. I., Leach, R. K., & Clare, A. T. (2016). Review of in-situ process monitoring and in-situ metrology for metal additive manufacturing. *Materials and Design*, *95*, 431–445.

Gonzalez-Val, C., Pallas, A., Panadeiro, V., & Rodriguez, A. (2020). A convolutional approach to quality monitoring for laser manufacturing. *Journal of Intelligent Manufacturing*, *31*(3), 789–795.

Grasso, M., & Colosimo, B. M. (2017). Process defects and in situ monitoring methods in metal powder bed fusion: A review. *Measurement Science and Technology*, *28*(4), 044005.

Grasso, M., Laguzza, V., Semeraro, Q., & Colosimo, B. M. (2016). In-process monitoring of selective laser melting: Spatial detection of defects via image data analysis. *Journal of Manufacturing Science and Engineering*, *139*(5), 051001.

He, Z., Zuo, L., Zhang, M., & Megahed, F. M. (2016). An image-based multivariate generalized likelihood ratio control chart for detecting and diagnosing multiple faults in manufactured products. *International Journal of Production Research*, *54*(6), 1771–1784.

Horst, R. L., & Negin, M. (1992). Vision system for high-resolution dimensional measurements and on-line SPC: Web process application. *IEEE Transactions on Industry Applications*, *28*(4), 993–997.

Kwon, O., Kim, H. G., Ham, M. J., Kim, W., Kim, G.-H., Cho, J.-H., et al. (2020). A deep neural network for classification of melt-pool images in metal additive manufacturing. *Journal of Intelligent Manufacturing*, *31*(2), 375–386.

Li, X., Jia, X., Yang, Q., & Lee, J. (2020). Quality analysis in metal additive manufacturing with deep learning. *Journal of Intelligent Manufacturing*, *31*(8), 2003–2017.

Li, W., Mahadevan, V., & Vasconcelos, N. (2013). Anomaly detection and localization in crowded scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *36*(1), 18–32.

He, Z., Zuo, L., Zhang, M., & Megahed, F. M. (2016). An image-based multivariate generalized likelihood ratio control chart for detecting and diagnosing multiple faults in manufactured products. *International Journal of Production Research, 54*(6), 1771–1784.

Lyu, J. J., & Chen, M. N. (2009). Automated visual inspection expert system for multivariate statistical process control chart. *Expert Systems with Applications*, *36*(3), 5113–5118.

Mahato, V., Obeidi, M. A., Brabazon, D., & Cunningham, P. (2020). Detecting voids in 3d printing using melt pool time series data. *Journal of Intelligent Manufacturing,* 1–8.

Mani, M., Feng, S., Lane, B., Donmez, A., Moylan, S., & Fesperman, R. (2017). *Measurement science needs for real-time control of additive manufacturing powder-bed fusion processes*. In Additive manufacturing handbook: Product development for the defense industry.

Megahed, F. M., Wells, L. J., Camelio, J. A., & Woodall, W. H. (2012). A spatiotemporal method for the monitoring of image data. *Quality and Reliability Engineering International*, *28*(8), 967–980.

Nembhard, H. B., Ferrier, N. J., Osswald, T. A., & Sanz-Uribe, J. R. (2003). An integrated model for statistical and vision monitoring in manufacturing transitions. *Quality and Reliability Engineering International*, *19*(6), 461–476.

Nti, I. K., Adekoya, A. F., Weyori, B. A., & Nyarko-Boateng, O. (2021). Applications of artificial intelligence in engineering and manufacturing: A systematic review. *Journal of Intelligent Manufacturing,* 1–21.

Park, C., & Shrivastava, A. K. (2014). Multimode geometric-profile monitoring with correlated image data and its application to nanoparticle self-assembly processes. *Journal of Quality Technology*, *46*(3), 216–233.

Redmon, J., & Farhadi, A. (2018). YOLO v.3. Tech report.

Spears, T. G., & Gold, S. A. (2016). In-process sensing in selective laser melting (SLM) additive manufacturing. *Integrating Materials and Manufacturing Innovation*, *5*(1), 16–40.

Tapia, G., & Elwany, A. (2014). A review on process monitoring and control in metal-based additive manufacturing. *Journal of Manufacturing Science and Engineering, 136*(6).

Tong, L.-I., Wang, C.-H., & Huang, C.-L. (2005). Monitoring defects in IC fabrication using a hotelling t/sup 2/control chart. *IEEE Transactions on Semiconductor Manufacturing*, *18*(1), 140–147.

Wang, K., & Tsung, F. (2005). Using profile monitoring techniques for a data-rich environment with huge sample size. *Quality and Reliability Engineering International*.