

Novelty Detection in Images by Sparse Representations

Giacomo Boracchi, Diego Carrera

Dipartimento di Elettronica, Informazione e Bioingegneria
Politecnico di Milano, Milano, Italy
Email: {giacomo.boracchi, diego.carrera}@polimi.it

Brendt Wohlberg

Theoretical Division, Los Alamos National Laboratory
Los Alamos, NM, USA
Email: brendt@lanl.gov

Abstract—We address the problem of automatically detecting anomalies in images, i.e., patterns that do not conform to those appearing in a reference training set. This is a very important feature for enabling an intelligent system to autonomously check the validity of acquired data, thus performing a preliminary, automatic, diagnosis.

We approach this problem in a patch-wise manner, by learning a model to represent patches belonging to a training set of normal images. Here, we consider a model based on sparse representations, and we show that jointly monitoring the sparsity and the reconstruction error of such representation substantially improves the detection performance with respect to other approaches leveraging sparse models. As an illustrative application, we consider the detection of anomalies in scanning electron microscope (SEM) images, which is essential for supervising the production of nanofibrous materials.

I. INTRODUCTION

We consider monitoring systems acquiring and processing images, such as those employed in biomedical or industrial-control applications. Often, in these application scenarios, images acquired under *normal* conditions are characterized by specific structures or patterns, as such, regions that do not conform to these are considered *anomalies*; see Figure 1 for an illustrative example. Anomalies might indicate a change in the operating conditions, an unexpected evolution of the monitored process, or a fault in the sensing apparatus. For these reasons, it is important that an intelligent system automatically detects these anomalies, and properly locates the anomalous regions within the image.

Typically, the problem of identifying patterns that are different (or novel) with respect to those in a training set of normal data is referred to *novelty detection* [1], [2], [3], and is formulated as a one-class classification problem [4]. It is often not unrealistic to assume that a training set of normal data is available, since in several application scenarios these are easy to collect, while it is usually difficult to gather enough examples of anomalies, and sometimes impossible to represent all the possible anomalies that might occur. *Anomaly detection* instead [5], refers to the general problem of detecting unexpected patterns both in supervised scenarios (when normal, or normal and anomalous samples are provided for training) and in unsupervised scenarios (when no labels of the training data are provided). In this paper, we shall refer to the patterns being detected as anomalies, despite the novelty detection context, since this is a more appropriate description in the type of application we consider.

Most often, novelty-detection techniques learn a predictive or approximating model from a training set of normal data and then, during operation, assess the goodness of fit of the learned model to each test data to determine whether it is normal or anomalous. Typically, the reconstruction error, i.e., the discrepancy between the model output and the test data, is monitored. This is a very general approach and any approximation model can be, in principle, used for this purpose. However, most of the solutions presented in the literature resort to neural networks [1].

We here tackle the novelty-detection problem by learning a model providing sparse representations of normal data. Sparse representations have become lately very popular in several application domains [6], however, very little work has been done for novelty detection using sparse representations. In this work, we show that when adopting sparse models for novelty detection, the sparsity of the representation has to be explicitly considered. In particular, we suggest to jointly monitor the sparsity and the reconstruction error, and this substantially outperforms other solutions exploiting standard criteria to assess how close the test data are to the sparse learned model.

We consider the novelty-detection problem in images like those of Figure 1(a), having normal regions well characterized by specific *local structures* (patterns), and therefore we operate in a patch-wise manner. Patches, namely small image regions having a predefined shape, are thus the core objects of our analysis, and each patch will be independently analyzed to determine whether it is normal or anomalous. By doing so, we use the learned model to detect anomalies in the local image structure, rather than outliers in the pixel values or changes affecting the whole image.

As a reference application scenario, we consider the production of nanofibrous materials by an electrospinning process [7], [8], which is monitored by a scanning electron microscope (SEM). In normal conditions, these SEM images depict a situation similar to in Figure 1(a), which clearly displays the peculiar structure of these nanofibres. Anomalous regions, like those highlighted in Figure 1(b), typically indicate problems in the nanofibres production process: an intelligent monitoring system is expected to autonomously detect and possibly measure these regions. In this scenario, automatic novelty detection [1] becomes of paramount importance since human inspection is infeasible because of the large number of images acquired and the small size of the anomalies to be detected. Our experiments indicate that sparse representations, and the

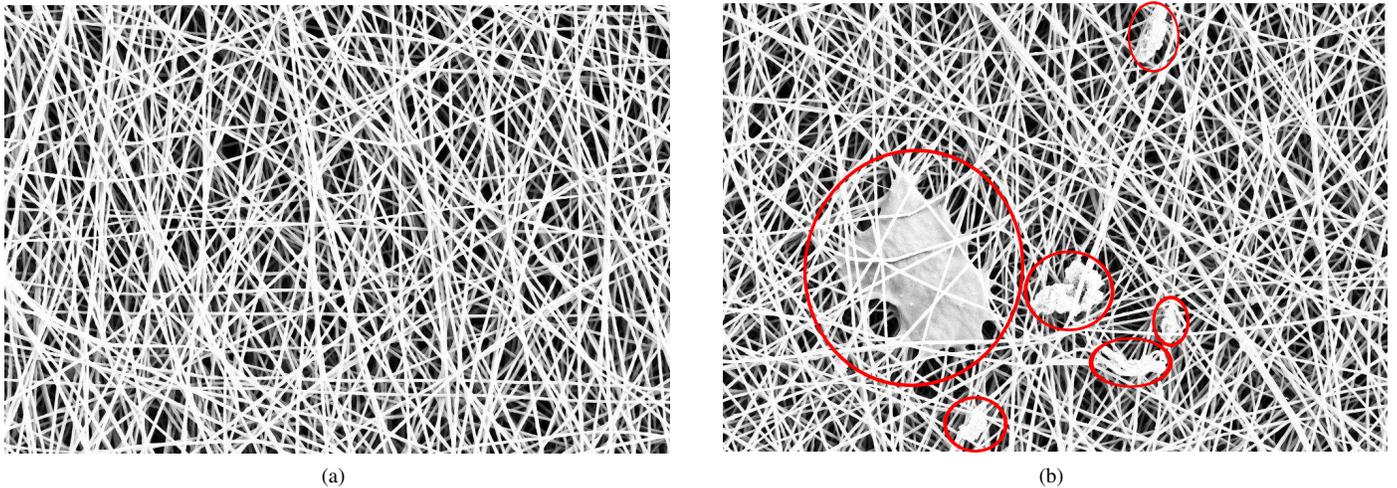


Fig. 1. Examples of SEM images depicting a nanofibrous material produced by an electrospinning process: Figure (a) does not contain anomalies, and is characterized by specific structures and patterns also at local-level. Figure (b) highlights anomalies that are clearly visible among the thin fibres.

proposed solution in particular, are very effective in detecting anomalies in textures and in the SEM images like those in Figure 1.

II. RELATED WORKS

An overview of novelty-detection methods on image data can be found in [1]. Here, we only point out few works that are particularly related to ours. In particular, [9] addresses novelty detection in mammograms to identify image regions indicating abnormal tissues. Five features are extracted from the images, and Parzen windows are used to estimate their joint probability density function. Then, anomalies are detected by drawing decision boundaries over the estimated density. In [10], the distribution of pixels in normal regions is modeled as a multivariate Gaussian Markov random field, and the detection of anomalous regions is performed by means of the Neyman-Pearson hypothesis test.

Only recently have sparse models been exploited for novelty detection. In [11], a specific sparse-coding procedure was designed to detect anomalies as data that do not admit a sparse representation with respect to a given dictionary: in practice, the recovery of sparse representations and the identification of anomalous components are simultaneously performed while processing the input data. In [12], a sparse model was exploited in a sequential monitoring application to detect structural changes in a stream of signals. All the above solutions identify anomalies by assessing the reconstruction error of the sparse model. We instead jointly monitor both the reconstruction error and the sparsity of the representation that is obtained when solving an unconstrained optimization problem, and our experiments demonstrate the effectiveness of this solution. Sparse representations have been also used for detecting unusual events in video sequences [13], [14], by monitoring the values of the functional minimized during the sparse coding stage.

It is worth mentioning signal-detection problems [15], [16], [17], [18], which are also performed by means of sparse representations and that have been mainly addressed in the

compressive sensing literature. In contrast to the novelty detection problem considered here, these works assume that in normal conditions only noise is observed, and are meant to detect the presence of any structured signal by analyzing few compressive measurements of the observations.

III. PROBLEM FORMULATION

Let us denote by $s : \mathcal{X} \rightarrow \mathbb{R}^+$ a grayscale image, where $\mathcal{X} \subset \mathbb{Z}^2$ is the regular pixel grid corresponding to the image domain¹. We denote by

$$\mathbf{s}_c = \{s(c+u), u \in \mathcal{U}\}, \forall c \in \mathcal{X} \quad (1)$$

the patch centered in a specific pixel c , having the support defined by \mathcal{U} , which is a neighborhood of the origin. While in principle patches \mathbf{s}_c can be defined over arbitrary shapes, in practice, \mathcal{U} is typically a square neighborhood of $\sqrt{m} \times \sqrt{m}$ pixels, where m is the cardinality of \mathcal{U} . Note that \mathbf{s}_c will usually be considered as a column vector representation in \mathbb{R}^m . We assume that patches in anomaly-free images are drawn from a stationary, stochastic process \mathcal{P}_N , and we refer to these as normal patches. In contrast, we assume that anomalous patches are generated by a different process \mathcal{P}_A , which yields unusual structures that do not conform to those generated by \mathcal{P}_N . We further assume that a training set of l normal patches is provided as a matrix $\mathbf{T} \in \mathbb{R}^{m \times l}$. While it is possible to learn a model approximating normal patches from \mathbf{T} , the same does not hold for anomalous patches, since no (or not enough) training samples are provided: thus, \mathcal{P}_A is completely unknown.

Novelty detection is performed at the patch level: each patch \mathbf{s}_c is tested to determine whether it *does or does not conform* with the model learned to approximate \mathcal{P}_N . This allows the automatic identification of regions of Figure 1(b) where the local image structure differs from the images in the training set. Thus, we are not simply interested in determining whether an image s contains anomalies or not, which in

¹We assume grayscale images here, but the method can be easily extended to colour images by defining patches over multiple colour bands.

practice would mean classifying Figure 1(a) as normal and Figure 1(b) as anomalous.

IV. NOVELTY DETECTION BY MEANS OF SPARSE REPRESENTATIONS

In this section we introduce the model we use for approximating patches generated by \mathcal{P}_N , and we show how to effectively detect patches that are not consistent with the learned model.

A. Sparse Approximation for Normal Patches

We adopt the classical model of sparse representations [6] and we learn a dictionary of patches \widehat{D} to yield an accurate and sparse representation for any patch s_c generated by \mathcal{P}_N , i.e.,

$$s_c \approx \widehat{D}x_c = \sum_{i=1}^n \widehat{d}_i x_i. \quad (2)$$

In (2), the *dictionary* $\widehat{D} = [\widehat{d}_1 | \dots | \widehat{d}_N]$ is a matrix of $\mathbb{R}^{m \times n}$ whose columns, $\widehat{d}_i, i \in \{1, \dots, n\}$, are referred to as *atoms*. The coefficient vector $x_c \in \mathbb{R}^n$ is assumed to be *sparse*, i.e. $\|x_c\|_0 = L \ll n$, where the ℓ^0 “norm” of x_c is the number of non-zero components in x_c . In (2), a patch s_c is approximated by $\widehat{D}x_c$ which occupies a low-dimensional subspace of \mathbb{R}^m generated by the patches corresponding to nonzero coefficients in x_c . Therefore, the sparse approximation of a patch can be seen as the projection onto the best union of low-dimensional subspaces spanned by few dictionary atoms.

We consider two different formulations of *sparse coding*, namely the estimation of a sparse representation for a specific patch s_c with respect to a given dictionary \widehat{D} :

- The constrained problem

$$x_{c,0} = \arg \min_{x \in \mathbb{R}^n} \|\widehat{D}x - s_c\|_2 \quad \text{such that} \quad \|x\|_0 \leq L, \quad (3)$$

which minimizes the reconstruction error of the best L -sparse representation of s_c . Here $L > 0$ represents the desired sparsity of the solution and the subscript 0 in $x_{c,0}$ indicates that the sparse coding was obtained by solving an optimization problem constraining ℓ^0 “norm” of the coefficient vector. Finding an exact solution of (3) is a computationally intractable problem; the standard approach being greedy algorithms such as Orthogonal Matching Pursuit (OMP) [19].

- The unconstrained problem

$$\widehat{x}_{c,1} = \arg \min_{x \in \mathbb{R}^n} J_\lambda(x, \widehat{D}, s_c), \quad (4)$$

where the $J_\lambda(\cdot)$ is a convex loss function defined as

$$J_\lambda(x, \widehat{D}, s_c) = \frac{1}{2} \|\widehat{D}x - s_c\|_2^2 + \lambda \|x\|_1, \quad (5)$$

and $\lambda > 0$ is a regularization parameter that balances the reconstruction error $\|\widehat{D}x - s_c\|_2^2$, and the sparsity $\|x\|_1$ of the solution measured by the ℓ^1 norm. There are a number of methods for solving this Basis Pursuit DeNoising (BPDN) [20] problem, including Alternating Direction Method of Multipliers (ADMM) [21].

The corresponding *dictionary learning* problems can be formalized as a joint optimization over both the dictionary and coefficients of a sparse representation of the training matrix T . These problems are non-convex, and are usually approached via alternating optimization with respect to the sparse representation and the dictionary. One of the best known dictionary learning algorithm is K-SVD [22], which usually employs OMP in its sparse coding stage.

According to the standard novelty detection scheme, we would like \widehat{D} to provide accurate and sparse representations exclusively for normal patches.

B. Detecting Anomalous Patches

The dictionary \widehat{D} can be used as a tool to measure the extent to which a test patch is close to those generated by \mathcal{P}_N , and thus the extent to which it is normal or anomalous. In particular, to quantitatively assess how close a patch is to the normal ones, we use *anomaly indicators* that measure the degree to which its approximation by \widehat{D} is accurate and sparse. When approximating anomalous patches by linear combinations of a few atoms of \widehat{D} , we expect a substantial deviation in the sparsity or reconstruction error – and thus also in the anomaly indicators – since atoms of \widehat{D} were learned to represent normal structures.

To this purpose, a viable anomaly indicator for the patch s_c would be the reconstruction error given by the solution of the constrained sparse coding (4),

$$e(s_c) = \|\widehat{D}x_{c,0} - s_c\|_2, \quad (6)$$

where $x_{c,0}$ comes from (3). The reconstruction error (6) was used also in [12] to detect structural changes in a sequential change-detection problem. It is important to remark that $e(s_c)$ constitutes a meaningful anomaly indicator, since the coefficient vector $x_{c,0}$ is L -sparse. This indicator is the same that was used in [12] for sequential monitoring applications.

In contrast, in the BPDN formulation (4), (5) the sparse coding of a patch s_c does not constrain the sparsity of $x_{c,1}$, the functional minimization allowing a trade-off between reconstruction error $\|\widehat{D}x_{c,1} - s_c\|_2$ and sparsity $\|x_{c,1}\|_1$. Thus, the reconstruction error alone is not a valid anomaly indicator. The most natural choice would be to use, for each patch s_c , the value of the functional J_λ (5),

$$f(s_c) = J_\lambda(x_{c,1}, \widehat{D}, s_c) = \frac{1}{2} \|\widehat{D}x_{c,1} - s_c\|_2^2 + \lambda \|x_{c,1}\|_1, \quad (7)$$

where $x_{c,1}$ is given by (4). The value of the functional (7) combines both the sparsity of $x_{c,1}$ and the reconstruction error of $\widehat{D}x_{c,1}$, and is the most straightforward option to assess the success of the unconstrained sparse coding of s_c with respect to \widehat{D} .

Here we adopt instead a bivariate anomaly indicator, thus jointly accounting for both the reconstruction error and the sparsity of the approximation given by \widehat{D} . In particular, given a patch s_c , we compute the sparse coding $x_{c,1}$ solving the BPDN (4) problem, and we define the vector

$$g(s_c) = [\|\widehat{D}x_{c,1} - s_c\|_2, \|x_{c,1}\|_1], \quad (8)$$

as the bivariate anomaly indicator.

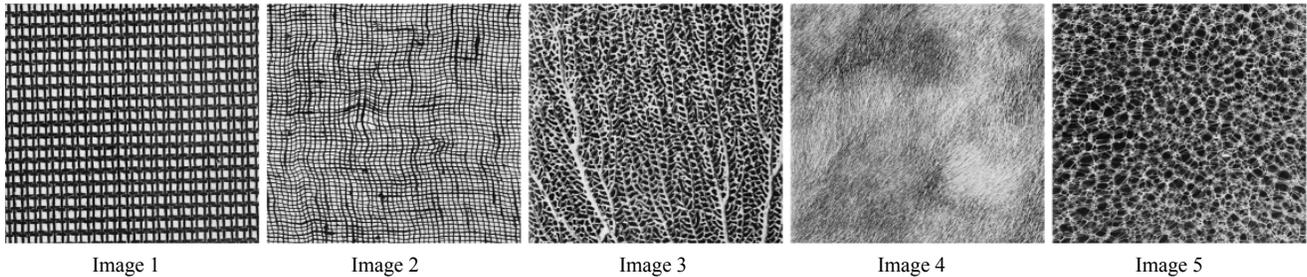


Fig. 2. The five texture images selected from the Brodatz dataset [23], since their structure can be well captured by 15×15 patches.

Anomaly indicators can be considered as maps from \mathbb{R}^m (the space of patches) into a low-dimensional feature space, where features from normal patches follow an unknown distribution \mathcal{F} . Patches yielding unusual values of these indicators can be considered anomalies. One of the most straightforward solution to detect anomalies consists of building a suitable confidence region and considering anomalous all the patches yielding anomaly indicators falling outside this region. In practice, a valid option for the scalar anomaly indicator (6), is

$$\mathcal{I}_\alpha^e = [q_{\frac{\alpha}{2}}, q_{1-\frac{\alpha}{2}}], \quad (9)$$

where q_α denotes the α -quantile of the empirical distribution of the reconstruction error (6). A patch \mathbf{s}_c is considered anomalous if

$$e(\mathbf{s}_c) \notin \mathcal{I}_\alpha^e. \quad (10)$$

Analogously, it is possible to define the region \mathcal{I}_α^f related to the indicator in (7), which leads to the following test:

$$f(\mathbf{s}_c) \notin \mathcal{I}_\alpha^f. \quad (11)$$

When the bivariate indicator (8) is used, we can build a two-dimensional region [24]

$$\mathcal{R}_\gamma = \left\{ \phi \in \mathbb{R}^2 : \sqrt{(\phi - \mu)^T \Sigma^{-1} (\phi - \mu)} \leq \gamma \right\}, \quad (12)$$

where μ and Σ are the expectation and the covariance matrix of \mathcal{F} , respectively, and γ is a suitably chosen threshold. Then, a patch \mathbf{s}_c is considered anomalous when it does not belong to \mathcal{R}_γ , i.e.,

$$\sqrt{(\mathbf{g}(\mathbf{s}_c) - \mu)^T \Sigma^{-1} (\mathbf{g}(\mathbf{s}_c) - \mu)} > \gamma. \quad (13)$$

While confidence regions \mathcal{I}_α^e and \mathcal{I}_α^f defined by the quantiles $q_{\frac{\alpha}{2}}$ and $q_{1-\frac{\alpha}{2}}$ have straightforward interpretation, the justification of (12) comes from the multivariate Chebyshev's inequality, which in this case ensures that, for a normal patch \mathbf{s}_c , holds

$$\Pr(\{\mathbf{g}(\mathbf{s}_c) \notin \mathcal{R}_\gamma\}) \leq \frac{2}{\gamma^2}, \quad (14)$$

where $\Pr(\{\mathbf{g}(\mathbf{s}_c) \notin \mathcal{R}_\gamma\})$ denotes the probability for a normal patch \mathbf{s}_c to lie outside the confidence region (false positive detection).

The parameters of the confidence regions can be estimated from the anomaly indicators computed from the training matrix \mathbf{T} or, when available, on additional anomaly-free images that can be used for validation purpose. In particular, the quantiles

$q_{\frac{\alpha}{2}}$ and $q_{1-\frac{\alpha}{2}}$ can be computed by analyzing the empirical distribution of the anomaly indicators over training or validation data, while μ and Σ are given by their corresponding sample estimators. In contrast, values of α and γ have to be empirically chosen according to the desired responsiveness of the anomaly detectors. In particular, the value of γ has to be tested to guarantee that the detector (13) yields an acceptable percentage of false positives. In fact, (14) provides an upper bound on the probability of having a false positive, which might substantially overestimate the expected percentage of false positives during operation. Of course, when the distribution of \mathcal{F} is known, confidence regions yielding tighter bounds of false positives can be used.

C. Preprocessing and Postprocessing

Before performing the dictionary learning and sparse coding, we subtract from each patch its average intensity value. This is a very common normalization procedure when using sparse representations for image-processing tasks. By doing so, we make our novelty-detection algorithm more sensitive to changes in the patch structure rather than in the overall patch intensity.

The final decision whether a patch is normal or anomalous ought be taken after some post-processing to spatially aggregate the decisions (i.e., normal/anomalous) at neighboring patches. Spatial aggregations is meant to filter out detections that very likely refer to normal regions, thus reducing the false positive rates of the novelty detector. In practical applications, these post-processing operations heavily determine the overall detection performance, since these can be designed exploiting specific information about the problem at hand (e.g., the minimum size of the anomalous regions, their shape or average intensity). Nevertheless, we do not consider here the post-processing since we are interested in investigating the effectiveness of sparse representations for novelty detection, rather than designing a specific novelty-detection solution.

V. EXPERIMENTS AND DISCUSSION

A. Considered Novelty Detection Techniques

In our experiments we consider the following novelty-detection techniques based on sparse models:

- **Reconstruction:** We use the OMP [19] algorithm to compute e in (6), and we use (10) as a criteria to detect anomalies. This indicator is the same used in [12] for a sequential change-detection algorithm.

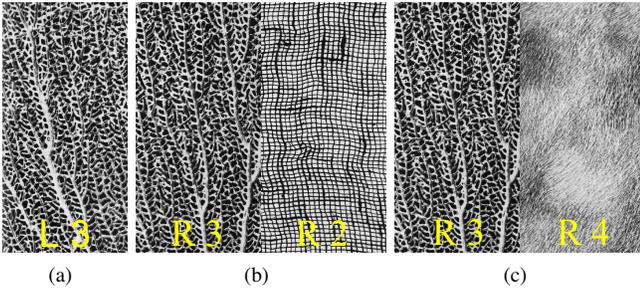


Fig. 3. Examples of test images used for the experiments in Section V-B. The letter (L or R) reported at the bottom of each half-image indicates whether this is the left or right-hand side of the original image, respectively, while the numbers refer to the index of the images in Figure 2 where the two halves have been taken from. As an illustrative example: the left-hand size of image 3 shown in (a) is exclusively used to learn a dictionary \hat{D} modeling this texture as normal, and the test images (b) and (c) are used to assess the novelty-detection performance when the anomalous patches are taken from textures of images 2 and 4, respectively. Note that in these test images, the false positive rate is correctly assessed using the right-hand size of image 3, and not patches used for training.

The responsiveness to anomalies is determined by the value of α , which defines the quantiles $q_{\frac{\alpha}{2}}$ and $q_{1-\frac{\alpha}{2}}$ in (9), thus the confidence region \mathcal{I}_{α}^e .

- **Functional:** We solve the unconstrained BPDN problem (4) by ADMM [21], then anomalies are detected by monitoring f in (7) by the confidence interval in (11). The responsiveness to anomalies is determined by the value of α .
- **Bivariate:** We monitor the indicator \mathbf{g} in (8) and we detect anomalies by means of (13). The responsiveness to anomalies is determined by the value of γ .
- **Coding:** We adopt the method presented in [11], where anomalies are detected while performing the sparse coding of the patches. This method is based on the assumption that anomalous patches cannot be well approximated by a linear combination of few atoms of \hat{D} . More specifically, the following approximation is considered:

$$\mathbf{s}_c \approx \hat{D}\mathbf{x}_c + \mathbf{a}_c, \quad (15)$$

where \mathbf{x}_c is the sparse coefficient vector of \mathbf{s}_c with respect to \hat{D} and \mathbf{a}_c gathers the components of \mathbf{s}_c that cannot be sparsely approximated. When the patch \mathbf{s}_c is normal, it can be well approximated by $\hat{D}\mathbf{x}_c$, then \mathbf{a}_c becomes negligible. In contrast, when \mathbf{s}_c is anomalous, \mathbf{a}_c collects almost all of the energy of the patch. Therefore, anomalies are detected by monitoring $\|\mathbf{a}_c\|_2$. More in details, the proposed detection ruled is

$$\|\mathbf{a}_c\|_2 > \tau, \quad (16)$$

where τ is the parameter determining the responsiveness to anomalies.

The above techniques have been tested in two application scenarios (Sections V-B and V-C) using patches of size 15×15 , where dictionaries \hat{D} were learned using the K-SVD algorithm [22]. Learned dictionaries are 4-times overcomplete, with $m = 225$ and $n = 900$; this is the configuration that provided

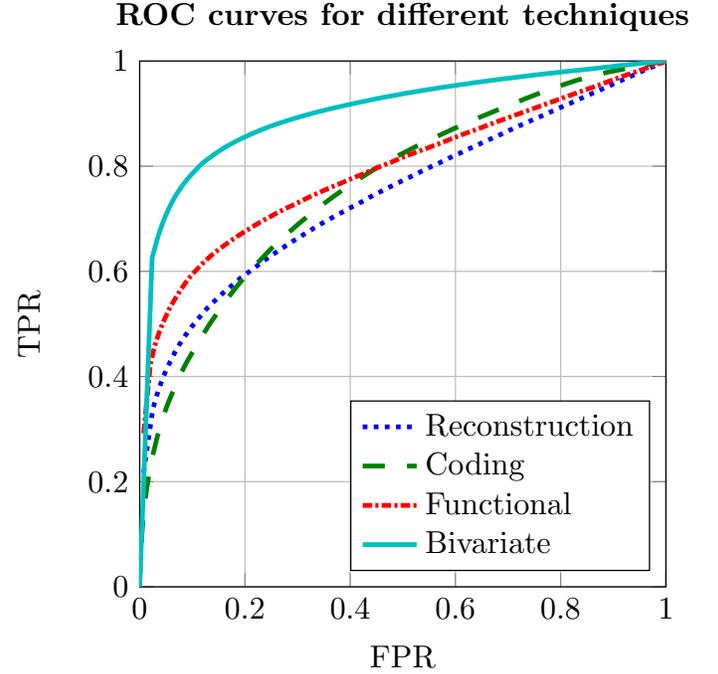


Fig. 4. ROC curves for the considered novelty-detection techniques in Section V-A, obtained by varying the corresponding parameters: α for *Reconstruction* and *Functional*, γ for *Bivariate* and τ for *Coding*. The FPR and TPR were averaged over all the test images of Section V-B. The ROC curve corresponding to the Bivariate is the closest to the upper-left corner, indicating that this novelty-detection solution outperforms the others.

the most satisfactory detection performance. By tuning the dictionary learning on few training and validation images, we set the degree of sparsity for the constrained sparse coding (3) as $L = 4$ and the penalization term in the BPDN problem (4) as $\lambda = 0.3$.

B. Experiments on Texture Images

Texture images are taken from the Brodatz dataset [23], and suitably combined to prepare test images for assessing novelty-detection performance. We considered the five 640×640 textures² displayed in Figure 2 and we learn a dictionary $\hat{D}_i, i \in \{1, \dots, 5\}$ from all the patches in the left half of each image (thus on a 320×640 image). The left half of each of these images is not considered for testing, while the right half is instead used to prepare test images. Test images are vertical juxtaposition of two different texture and are processed using the dictionary \hat{D}_i that refer to one of the two textures. By doing so, half of the test image have to be considered normal, while the other half anomalous. Figure 3 illustrates how these images have been prepared. Overall, we tested all the possible combinations of texture images in Figure (3), obtaining twenty test images.

The following figures of merit are used to assess the performance of the novelty detectors in Section V-A:

- **FPR**, the false positive rate, i.e. the percentage of normal patches labeled as anomalous.

²These textures were selected because it seems to be possible to capture their structure within patches of 15×15 pixels.

- **TPR**, the true positive rate, i.e. the percentage of anomalies correctly detected.

These figures of merit heavily depend on the values of γ , α and τ . To determine which of these novelty-detection techniques was the most effective one, we tested a wide range of values for each of these parameters and we plotted the FPR and TPR averaged over the 20 test images. The results are reported in the receiver operating characteristic (ROC) curves of Figure 4. This plot shows a clear gap between the performance achievable by the bivariate indicator and the others, since its curve is much closer to the upper-left corner.

C. Experiments on the NanoTwice Images

The NanoTwice dataset consists of SEM images obtained during the quality control of an electrospinning process for producing nanofibres. Electrospinning is an electro-hydrodynamic process taking place between a spinning head with a capillary opening and a static plate. The spinning head is connected with a reservoir filled with a polymer solution under pressure. A high voltage is applied to the spinning head, whereas the plate is usually grounded, the difference in the voltage between the spinning head and the plate resulting in an electrically driven polymer solution jet. The solvent rapidly evaporates from the jet during the run and, under optimal conditions, a continuous nano-sized filament is deposited on the plate.

SEM images of these nanofibres show a very peculiar, non-woven, structure like those reported in Figure 1. In order to control and optimize the nanofibres production process, it is important to detect anomalies in such nanofibres, which are typically classified as

- *Beads*: limited pieces of the fibre whose diameter is significantly larger than the rest of the fibre.
- *Films*: thin, flat layer of polymer lying among the nanofibers.
- *Holes*: large dark areas which are not covered by nanofibers.

We learned the dictionary \hat{D} from more than 166000 patches extracted from 10 anomaly-free regions of size 400×400 that were manually cropped from some images. We report the detection performance over the three images shown in the top row of Figure 6, where each technique was tuned to yield approximately the same number of detections. Since these anomalies are rather easy to locate, we resort to visual inspection for performance assessment. In Figure 6, pixels marked in red correspond to coordinates c where the patch s_c was considered anomalous.

Results reported in Figure 6 show that the proposed Bivariate indicator is better at locating anomalies as it concentrates more than others the detections over beads, films and holes in these images. Similar results can be achieved by monitoring the Functional and the Reconstruction indicators, even though some small beads in the second image and part of the films can not be detected. Conversely, the Coding indicator is not able to detect films and holes, since, given their simple structure, patches in these areas can be accurately reconstructed and cannot be detected by (16).

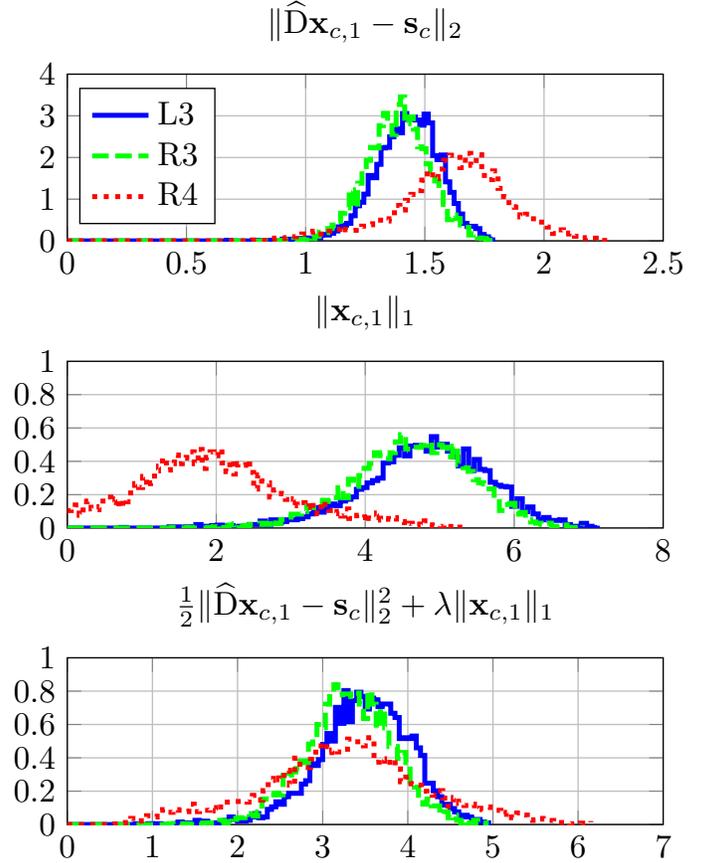


Fig. 5. Empirical distributions of the anomaly indicators (7) and (8) on the Images in Figure 3 (a) and Figure 3 (c). In this specific case the Functional indicator is not able to correctly identify anomalies, which can be instead successfully detected by jointly monitoring the Bivariate anomaly indicator (8). Each histogram has been normalized to yield area equal to 1.

D. Discussion

Our experiments clearly show the potential of sparse representations for detecting anomalies in the image structure, and the ROC curves in Figure 4 indicate that the Bivariate detector (8) can definitively outperform the others. This suggests that anomalous patches sometimes yield reconstruction errors and sparsity values that are not anomalous by themselves even though, when jointly considered, they fall outside the confidence region (12). Remarkably, the results emerging from the NanoTwice dataset are consistent with those on synthetic images.

It is however surprising that the functional in (7), which is actually the figure of merit minimized by the sparse coding (4) – thus the most natural indicator of the goodness of fit for a sparse model – is far less effective than the Bivariate anomaly indicator (8). Probably, one of the motivations is illustrated in Figure 5, where the distribution of indicators over the test image of Figure 3 is reported. In this specific situation, the sparse coding (4) achieves very sparse representations for the anomalous patches (central plot) yielding comparable reconstruction errors to the normal ones (top plot). This behavior is motivated by the fact that anomalous patches depict structures (see image 4 in Figure 2) that are simpler than the

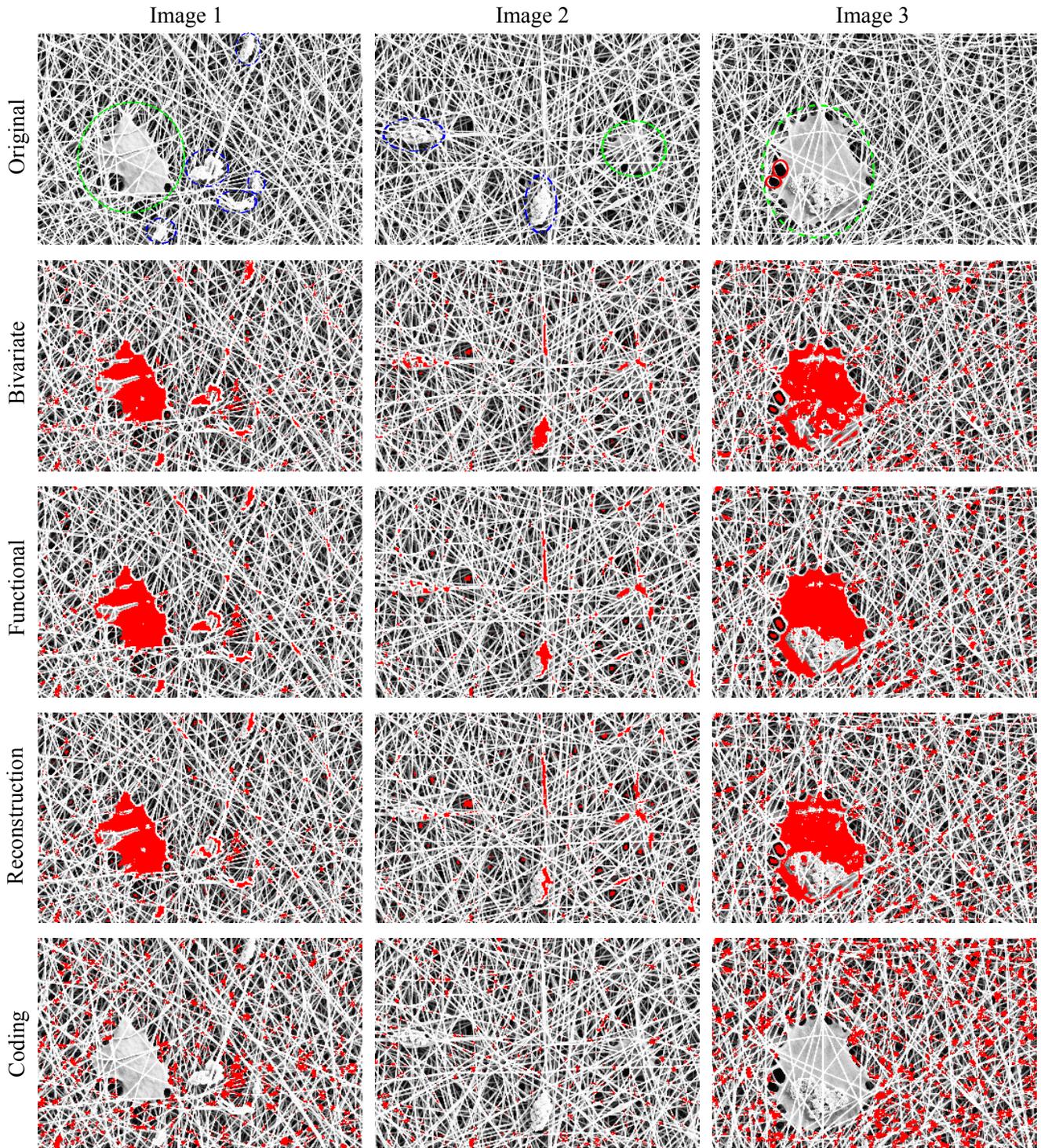


Fig. 6. Anomaly-detection performance of the considered solutions on the SEM images from the NanoTwice dataset. The top row depicts the original images, highlighting anomalies: films (dashed green lines), beads (dotted blue lines) and holes (solid red lines). Images from the second to fifth row show the output of the considered novelty-detection techniques: each red pixel represents the center of a patch considered anomalous. The novelty detector based on Bivariate indicator is able to correctly identify almost all anomalies, with a small number of false positives. Similar results hold for the Functional and the Reconstruction indicators. The Coding indicator is instead not able to detect films and holes, because patches belonging to these regions are very simple and admit an accurate reconstruction. These results show that it is essential to take into account also the sparsity of the representation when performing novelty detection. The parameters α , γ and τ , have been set to provide approximately the same number of patches detected as anomalous in each image.

normal patches (see image 3 in Figure 2). Then, the functional somehow balances the contribution of the two terms (bottom plot), making impossible the separation of anomalies from normal patches. In contrast, the bivariate indicator is able to detect these situations since they fall outside the confidence region (12). Similar arguments holds for the Reconstruction (6) and the Coding indicators, that are far more sensitive to variations in the reconstruction error.

VI. CONCLUSIONS

In our previous work [12], we demonstrated the potential of sparse representations for detecting structural changes in streams of signals. In particular, in [12] we achieved promising change-detection performance by monitoring the reconstruction error of a sparse representation computed when constraining the sparsity of the solution. The present paper extends this work to novelty detection in images, confirming the potential of models providing sparse representations for detection purposes. Furthermore, we show the advantages of jointly monitoring the reconstruction error and the sparsity of the solution to the unconstrained BPDN problem.

Future works will address the application of these results to the sequential monitoring scenario, and the investigation of novel learning algorithms to provide dictionaries that are specifically meant to perform change/novelty detection rather than data reconstruction. We will also investigate how the detection performance and the number of required training patches vary when the patch size increases.

ACKNOWLEDGMENTS

The authors would like to thank the Institute of Applied Mathematics and Information Technology (IMATI), Milan, Italy, and the Institute of Science and Technology for Ceramics (ISTEC), Faenza, Italy, for providing them the images of NanoTwice dataset, and Dr. Adler for the codes and the support to run the method in [11].

REFERENCES

- [1] M. A. Pimentel, D. A. Clifton, L. Clifton, and L. Tarassenko, "A review of novelty detection," *Signal Processing*, vol. 99, pp. 215 – 249, June 2014.
- [2] M. Markou and S. Singh, "Novelty detection: a review – part 1: statistical approaches," *Signal Processing*, vol. 83, no. 12, pp. 2481–2497, 2003.
- [3] —, "Novelty detection: a review – part 2: neural network based approaches," *Signal Processing*, vol. 83, no. 12, pp. 2499 – 2521, 2003. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0165168403002032>
- [4] B. Schölkopf, R. C. Williamson, A. J. Smola, J. Shawe-Taylor, and J. C. Platt, "Support vector method for novelty detection," in *Advances in Neural Information Processing Systems 12 (NIPS)*, November 2000, pp. 582–588.
- [5] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Computing Surveys*, vol. 41, no. 3, pp. 15:1–15:58, July 2009.
- [6] A. M. Bruckstein, D. L. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," *SIAM Review*, vol. 51, no. 1, pp. 34–81, February 2009.
- [7] W. E. Teo and S. Ramakrishna, "A review on electrospinning design and nanofibre assemblies," *Nanotechnology*, vol. 17, no. 14, p. R89, July 2006.
- [8] L. M. Bellan and H. G. Craighead, "Applications of controlled electrospinning systems," *Polymers for Advanced Technologies*, vol. 22, no. 3, pp. 304–309, March 2011.
- [9] L. Tarassenko, P. Hayton, N. Cerneaz, and M. Brady, "Novelty detection for the identification of masses in mammograms," in *Proceedings of IEEE International Conference on Artificial Neural Networks*, June 1995, pp. 442–447.
- [10] G. G. Hazel, "Multivariate gaussian mrf for multispectral scene segmentation and anomaly detection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 3, pp. 1199–1211, May 2000.
- [11] A. Adler, M. Elad, Y. Hel-Or, and E. Rivlin, "Sparse coding with anomaly detection," in *Proceedings of IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, September 2013, pp. 1–6.
- [12] C. Alippi, G. Boracchi, and B. Wohlberg, "Change detection in streams of signals with sparse representations," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2014, pp. 5252 – 5256.
- [13] B. Zhao, L. Fei-Fei, and E. P. Xing, "Online detection of unusual events in videos via dynamic sparse coding," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2011, pp. 3313–3320.
- [14] Y. Cong, J. Yuan, and J. Liu, "Sparse reconstruction cost for abnormal event detection," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2011, pp. 3449–3456.
- [15] J. Haupt and R. Nowak, "Compressive sampling for signal detection," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 3, April 2007, pp. 1509–1512.
- [16] Z. Wang, G. R. Arce, and B. M. Sadler, "Subspace compressive detection for sparse signals," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2008, pp. 3873–3876.
- [17] M. A. Davenport, P. T. Boufounos, M. B. Wakin, and R. G. Baraniuk, "Signal processing with compressive measurements," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 445–460, April 2010.
- [18] M. F. Duarte, M. A. Davenport, M. B. Wakin, and R. G. Baraniuk, "Sparse signal detection from incoherent projections," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 3, May 2006.
- [19] Y. C. Pati, R. Rezaifar, and P. S. Krishnaprasad, "Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition," in *Proceedings of Asilomar Conference on Signals, Systems, and Computers*, November 1993, pp. 40–44.
- [20] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33–61, January 1998.
- [21] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1–122, January 2010.
- [22] M. Aharon, M. Elad, and A. M. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, November 2006.
- [23] P. Brodatz, *Textures: A Photographic Album for Artists and Designers*. Peter Smith Publisher, Incorporated, 1981.
- [24] R. Johnson and D. Wichern, *Applied multivariate statistical analysis*. Prentice Hall, 2002.