DETECTION OF THINNING OF HOMOGENEOUS MATERIAL USING ACTIVE THERMOGRAPHY AND CLASSIFICATION TREES

Sebastian Dudzik, Grzegorz Dudek

Czestochowa University of Technology, Faculty of Electrical Engineering, Al. Armii Krajowej 17, 42-200 Częstochowa, Poland (✉ *sebdud@el.pcz.czest.pl*, +48 34 325 02 25, *dudek@el.pcz.czest.pl*)

Abstract

Active thermography is an efficient tool for defect detection and characterization as it does not change the properties of tested materials. The detection and characterization process involves heating a sample and then analysing the thermal response. In this paper, a long heating pulse was used on samples with a low thermal diffusivity and artificially created holes of various depths. As a result of the experiments, heating and cooling curves were obtained. These curves, which describe local characteristics of the material, are recognized using a classification tree and divided into categories depending on the material thickness (hole depths). Two advantages of the proposed use of classification trees are: an in-built mechanism for feature selection and a strong reduction in the dimensions of the pattern. Based on the experimental study, it can be concluded that classification trees are a useful tool for the thinning detection of homogeneous material.

Keywords: active thermography, classification tree, defect detection and characterization, material thinning detection.

1. Introduction

Material defects (*i.e.* discontinuities in material structures) can be detected using *non-destructive testing methods* (NDT). These methods allow the quality of the construction elements of machines and devices or their components to be controlled without changing the properties of materials from which they were made. NDTs have found widespread application in the chemical and petrochemical industries, aerospace industry, construction industry and power industry, among others.

To ensure high production quality, various NDT methods can be employed, including eddy-current, ultrasonic or radiographic. Passive and active infrared thermography [1–4] can also be used for defect detection.

During experiments with active thermography the following steps can be distinguished [1, 5–7]:

– the surface of the investigated material sample is subjected to a specific thermal excitation (lock-in, stepped heating and others),

– the transient temperature field occurring on the surface of the investigated material sample is recorded either during the heating or cooling process.

In the first step, various heating sources can be applied, for example, the stepped heating method (SH) which involves heating with a long thermal pulse, or lock-in thermography which uses periodical excitation. In the second step, the temperature field occurring on the investigated material surface is recorded over time in the form of sequences of thermograms, most frequently using an infrared camera [1, 2]. After the experimental phase of non-destructive testing, analysis is carried out. In the analysis, the sequences recorded in the experiments play the main role, with the ultimate goal of analysis being material defect detection. Depending on the experimental method employed, different data structures can be obtained and different kinds of analysis can be applied (*e.g.*: machine learning methods [5, 6, 8], heat transfer modelling methods or digital image processing methods). In our work, we use the stepped heating method in the experiments. In this case, data with many features is obtained as the result of the experiment. This data takes the form of vectors representing temporal temperature waveforms (the features are temperature values recorded at specified time steps). Due to a large number of features and in order to obtain an effective classification, further processing usually uses dimensional reduction methods (*e.g.* *Principal Component Analysis* – PCA). In this work, classification trees were used for thinning detection of homogeneous material based on analysis of a sequence of thermograms. An advantage of this classification method is the in-built mechanism for feature selection.

In this paper, the time-resolved infrared radiometry method (so-called stepped heating method) was applied in experiments [1, 5]. In this method, a two-layer model describes the geometry of the investigated sample, wherein "layer one" denotes a material without defects and "layer two" represents a subsurface defect at a specified depth [1]. The real and simplified geometry models are shown in Fig. 1. In this figure, "1" denotes the defect-free area, "2" denotes the defect area and z denotes the thickness of the first material layer (*i.e.* defect depth). For the simplified two-layer geometry model shown in Fig. 1, the rise in surface temperature caused by heating in the form of the step function, can be expressed by the following equation [1, 9]:

$$\Delta T\left(\tau\right) = C_c \sqrt{\tau} \left\{ 1 + \sum_{n=1}^{\infty} 2\left(-\Gamma\right)^n \left[\exp\left(-\frac{n^2 z}{a_0 \tau}\right) - \frac{nz\sqrt{\pi}}{\sqrt{a_0 \tau}} \mathrm{erfc}\left(\frac{nz}{\sqrt{a_0 \tau}}\right) \right] \right\}, K, \qquad (1)$$
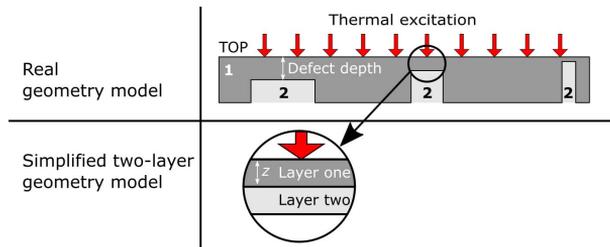

Fig. 1. Model of the investigated sample with two material layers.

In formula (1), the C_c describes the energy absorption coefficient dependent on the heating source characteristic, a_0 is thermal diffusivity of the defect-free area and the Γ coefficient is called

the mismatch factor. This can be defined as:

$$\Gamma = \frac{e_1 - e_0}{e_1 + e_0},\qquad(2)$$

where: e_0 – thermal effusivity of the "layer one" (*i.e.* material without defect), e_1 – thermal effusivity of the "layer two" (*i.e.* material with a defect).

During the experiment, the investigated sample is heated using constant excitation of low power. At the same time, the temperature rise is recorded using an infrared camera. As a result of recording, a sequence of thermograms of the surface of the investigated material sample is obtained. This sequence is the record of the transient temperature field caused by the applied thermal excitation. Another task undertaken during the thermal NDT is defect characterization. This relies on quantitative estimation of defect parameters such as depth, size and thermal resistance. Unfortunately, equation (1) cannot be used for defect characterization because it takes into account the one-dimensional heat conduction model which is valid only for very simplified assumptions [9]. Therefore, new methods of characterization are still being sought, such as Bayesian classification or artificial neural networks [6, 8, 10].

In this work, during the experiments, sequences of the surface thermograms of the tested material were obtained. On the basis of these sequences, *heating-cooling curves* (H-CC) expressing temperatures in Kelvins in each sample points (pixels in thermograms) for subsequent time steps were determined. Finally, a classification tree was used to analyse the H-CC curves. As a result of the analysis, areas of changes in material thickness were recognized on the surface of the sample under investigation. The input patterns for the tree are H-CCs.

The remaining sections of the paper are organized as follows. In Section 2, experimental investigations using active thermography and stepped heating method are described. In Section 3, a classification tree for detection and characterization of material thinning on the basis of thermograms is presented. Analysis of the thermograms and evaluation of the performance of the classification tree are shown in Section 4. Finally, Section 5 is a summary of our conclusions.

2. Experimental study

The stepped heating method involves heating the surface of the investigated material sample and registering the temperature field occurring on that surface. In this work, experimental studies were carried out to detect and characterize changes in material thickness using the stepped heating method. In these studies, the experimental setup presented in Fig. 2 was used.

Fig. 2. Experimental setup for stepped heating method.

As can be seen from Fig. 2, an IRS336-NDT infrared camera (long-wave 8 μm–14 μm, spatial resolution 336 × 256 pixels and an FPA detector with NETD < 50 mK) are used. The camera records the thermal image sequences. Two 250 W lamps (2) were used as heat sources. The lamps were supplied with a power amplifier (5) controlled by a heat excitation controller (3) and a digital acquisition card (4). The infrared camera was connected to a PC (6) using a GigEVision digital interface with GenICam. The PC had data acquisition and processing software. The software (IrNDT from Automation Technology) cooperated with the infrared camera and the heat excitation controller used to control the lamp power.

In the experimental study we used a sample made of Plexiglas (7). Nine blind, flat-bottomed holes were drilled in the sample. Each of them was drilled to a different depth ranging from 1 mm to 3.4 mm and with diameters ranging from 2.0 mm to 18 mm. The sample material used in the tests had a thickness of 10 mm. The heated sample surface was coated with a paint layer having a coefficient of band emissivity $\varepsilon \sim 0.98$. Figure 3 presents the shape and dimensions of the tested sample and the location of the holes.
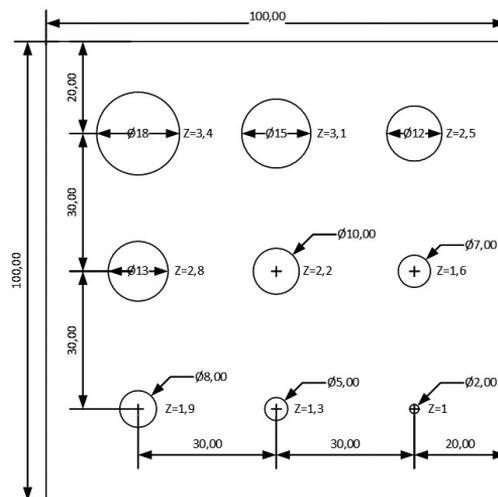

Fig. 3. The geometry of the investigated sample.

Experiments were conducted using the following procedure:
1. The heating phase. In this phase, the surface of the sample is heated using two lamps. The heating is conducted for 65 s. During the heating, a sequence of thermograms is recorded by means of an infrared camera (the frame rate of recording is equal to 15 Hz).
2. The cooling phase. After the end of the heating phase, a sequence of thermal images is recorded for a further 60 s. In this phase, there is no heat excitation.

As a result of experimental investigations, a sequence of 3175 thermograms was obtained. Each thermogram covers a field of view of 189 × 190 pixels. Selected thermograms recorded in the experimental study can be seen in Figs 4–6. In Fig. 4, we can see the early heating phase and in Figs 5 and 6, the end of the heating phase and the late cooling phase respectively.

Analysing the results of the experiment presented in Figs 4–6, it is possible to observe heterogeneity of surface heating on the tested sample, in particular during the heating phase. There are known methods of removing heterogeneities, for example those presented in [13]. This work, however, did deliberately not use any of these methods as we wanted to assess the
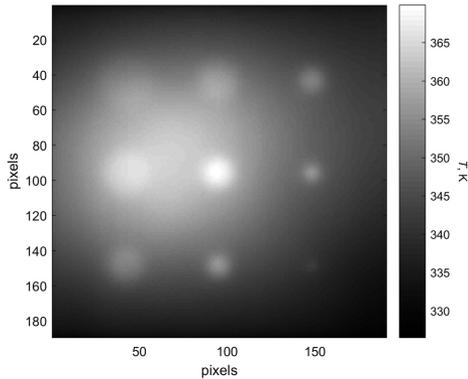
Fig. 4. Thermogram of the surface of the investigated
sample recorded at $\tau = 20$ s.
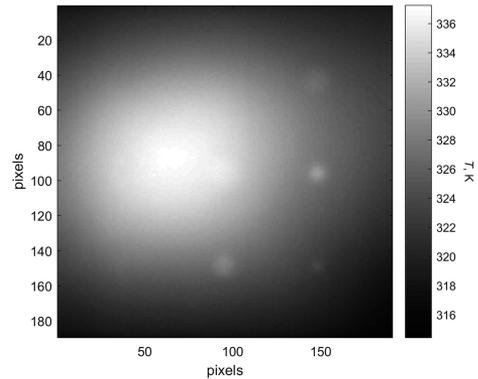
Fig. 5. Thermogram of the surface of the investigated
sample recorded at $\tau = 65$ s.
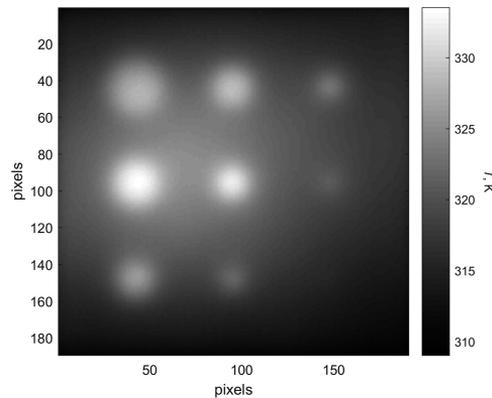
Fig. 6. Thermogram of the surface of the investigated sample recorded at $\tau = 125$ s.

effectiveness of detection and characterization of the thinned areas by means of classification trees for the worst case. In addition, on the thermograms in Figs. 4–6, you can observe a transition zone (*i.e.* a blur) between the thinned regions and the background area of full thickness. This is strictly dependent on the time step of recording. It is obvious that in the sample under investigation, the heat also flows in a direction parallel to its surface. Therefore, one can notice in the thermograms a "blurring" of the image of the temperature field in the transition zone. This is caused by a change in the temperature distribution on the investigated surface associated with the heat flow mentioned above. For this reason, a transitional zone appears between the thinned regions and the background area. The presence of such zones can cause problems at the data processing stage, especially when pixels from the transition zone are used as a training data source in machine learning algorithms. Due to the fact that changes in thickness of a homogeneous material may be caused by material faults (*e.g.*: cracks, dents and others), they will be referred to as "defects" later in this article.

H-CCs recorded at two exemplary points on the surface of the sample representing the area of the class "defects" and the area of the class "without defects" are presented in Fig. 7. It can be seen that the curve registered in the area of the "defect" class is above the curve registered in the

class area "non-defect". At "defect" points, where the material layer is thinner, there is smaller local heat capacity which creates differences in H-CCs presented in Fig. 7. In this work, this phenomenon (different locations of H-CCs for the "defects" and "non-defect" points) was used to detect the "defect" areas. The next section proposes a classification tree that classifies H-CCs recorded in the experiments and, in this way, identifies "defect" areas.
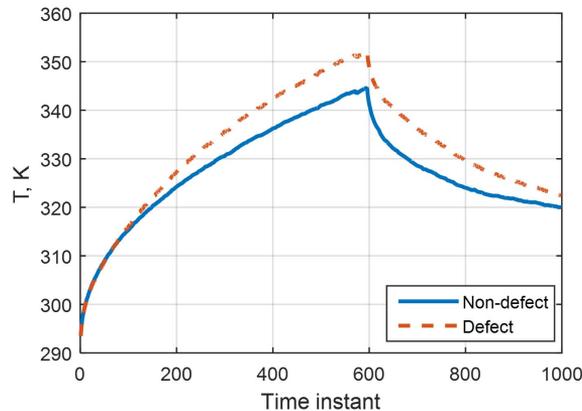

Fig. 7. Heating-cooling curves representing "defect" and "non-defect" classes [18].

3. Classification tree for characterization of material thickness

Decision trees (DT) are one of the most popular and effective tools of machine learning applied to both regression and classification problems [14]. DTs can extract useful information from large datasets and display it in easy-to-interpret visualizations. They have been widely used in several disciplines [15] because they are easy to use, free of ambiguity, and robust even in the presence of missing values. Both discrete and continuous variables can be used either as target variables or independent variables.

The most common strategy for learning DTs from data is a process of top-down induction of DTs using a greedy algorithm [16]. In the tree construction process, a series of decisions (splitting rules) are made to segment the predictor space into homogeneous subgroups. This process is called recursive partitioning and can be summarized in a tree with decision nodes and leaf nodes. The tree can be expressed alternatively by IF-THEN decision rules.

The advantages of DTs can be summarized as follows:
– use a white box model which is simple to understand,
– offer direct interpretability (decision rules),
– able to handle both numerical and categorical data as input and output variables,
– require little data preparation,
– fast training; the cost of using the tree is logarithmic in the number of training points,
– able to handle multi-output problems.

The disadvantages of DTs include: unstable learning (small variations in the data might result in a completely different tree), suboptimal final solution due to locally optimal decisions made at each node, and biased trees if some classes dominate.

In this work, we use the CART (Classification and Regression Tree) model of DT proposed by Breiman [17]. In the case of detection of changes in material thickness, the tree learns to

solve a classification problem where the inputs are n-dimensional vectors $x \in R^n$ which represent H-CCs, and the output represents a "defect" class: 0 for "non-defect", 1 for "defect 1", 2 for "defect 2", ..., 9 for "defect 9". Classes 1–9 correspond to the areas of nine "defects" (see. Fig. 3). The components of x express temperatures at consecutive time steps, *i.e.* each of N pixels has a corresponding x-vector containing the H-CC recorded for this pixel. The training set $\Omega = \{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\}$, on which the tree is constructed, contains N instances representing N pixels. An instance is a pair (x, y), where $x = [T_1 \, T_2 \ldots T_n]$ is a H-CC recorded for the pixel, and $\{y \in 0, 1, \ldots, 9\}$.

We assess the CART model using the accuracy defined as follows:

$$Acc = \frac{|\{x \in \Theta \mid h(x) = y\}|}{|\Theta|} \cdot 100, \tag{3}$$

where Θ is a set of instances (training or testing set) and $h(x)$ is the model output (predicted class). More precisely, we can assess the model using accuracy for each class:

$$Acc = \frac{|\{x \in \Theta \mid h(x) = k\}|}{|\Theta(k)|} \cdot 100, \tag{4}$$

where $\Theta(k)$ is a subset of instances from class k (in our case $k = 0, 1, \ldots, 9$).

Fig. 11 in the experimental part of the work shows an example of CART. As you can see from this figure, CART consists of two types of nodes: internal or decision nodes (marked with triangles) and terminal nodes – leaves (marked with points). The internal nodes perform tests on inputs, *i.e.* the components of x. The terminal nodes include class labels. CART is called a binary tree because internal nodes have at most two children.

The CART is built with a recursive tree-growing algorithm, from a root node and involves partitioning the data into subsets that contain instances with similar y-values. The sets of instances reaching internal nodes are split into two subsets. One of them contains the positive instances which meet the test performed in the node. The second one contains the negative instances which do not meet this test. For continuous variables, the test can be as follows:

$$\chi(x) = \begin{cases} 1, & \text{if} \quad x_i < s_i \\ 0, & \text{if} \quad x_i \geq s_i \end{cases}, \tag{5}$$

where x_i is a splitting variable (temperature at time step i in our case) and s_i is a split-point.

The splitting variable as well as the split-point are selected for each internal node individually using a greedy algorithm known as recursive binary splitting which considers all variables one by one (1874 consecutive temperatures forming the H-CC in our case) and all possible split-points for each of the variables. The maximized function is a reduction in a node impurity defined as follows:

$$\Delta I = I - p_1 I_1 - (1 - p_1) I_2, \tag{6}$$

where I, I_1 and I_2 are the impurities of the parent node and two children nodes, respectively, and p_1 is the proportion of instances which flow to the first children after test χ is performed in the parent node.

One of the most popular impurity measures is a Gini diversity index defined as follows:

$$I = 1 - \sum_{k=0}^{K} p_k^2, \tag{7}$$

where p_k is the proportion of class k instances in the node and K is the number of classes.

S. Dudzik, G. Dudek: DETECTION OF THINNING OF HOMOGENEOUS MATERIAL USING ACTIVE THERMOGRAPHY . . .

The Gini index takes on a small value if all of the p_k's are close to zero or one. This means that a node contains predominantly instances from a single class. Such a node has a high degree of purity.

The stopping criteria used when constructing the tree are:
- a node is pure, *i.e.* there are only instances of one class in the node,
- no reduction in impurity (6) is observed,
- a node includes less than L instances.

When, during the tree construction process, a node meets one of these conditions it becomes a leaf. The minimal leaf size is *L*. This parameter controls the generalization property of the model. If the value of *L* is small, the tree is deeper, having many leaves and tending to overfit. Larger values of *L* lead to a shallower tree which, while robust and easy to interpret, might not capture the important structure of the data. Optimal tree size should be adaptively chosen from the data.

4. Simulation study

In [18], we analysed the simple case of a thermogram for a sample fragment of 39×39 pixels around the middle hole $\phi 10$ (see: Fig. 3). In the experiment, a stepped heating method using a halogen lamp with a power of approx. 1.75 kW was applied. The heating phase lasted 120 s and the cooling phase was 80 s. Thermograms were recorded at 5 Hz frame rate. Finally, a sequence of 1000 thermograms was obtained. A CART model was built based on this data. The training set was composed of $39 \times 39 = 1521$ instances. Input vectors represent H-CCs for each pixel, so their length was 1000. Some of them had the label "non-defect" ($y = 0$), and others "defect" ($y = 1$).

The goal of this study was to show how the CART model deals with thermograms. For a 39×39 pixels sample fragment we get 1521 H-CCs which form two overlapping bands, see Fig. 8. The upper band is composed of curves representing pixels belong to the thinning region and the lower band is composed of curves representing pixels in unchanged material thickness area. The curves lying in the overlapping band are difficult to recognize. Figure 8 shows that the overlapping region is not very wide. This means that the classes "defect" and "non-defect" are well separated and that temperature has a large discriminative power. When we analyse the overlapping band over time, we can see that its width is variable. At the beginning of the heating
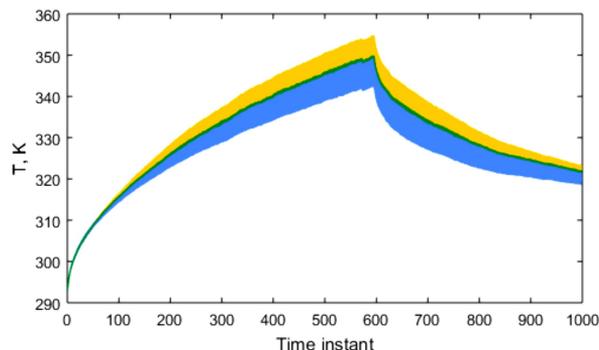

Fig. 8. The heating-cooling curves representing the "defect" (upper band)
and "non-defect" (lower band) classes for all pixels [18].

process most of the pixels belong to the overlapping band, see Fig. 9. When the heating process is more advanced, the number of pixels in the overlapping band decreases rapidly and stays at a low level. The pixels lying in the overlapping region depending on the time step are shown in Fig. 10. As you can see from this figure, at the initial stage of the heating most pixels from the upper part of the thermogram and from the "defect" area belong to the overlapping band. This is because they have a similar temperature. After $\tau = 20$ s the overlapping region decreases. Only 74 pixels from the transition zone between the "defect" and "non-defect" regions lie in the overlapping band. Further heating reduces the number of such pixels to only 18 after $\tau = 120$ s. Thus, the discriminative power of temperature increases with heating time.
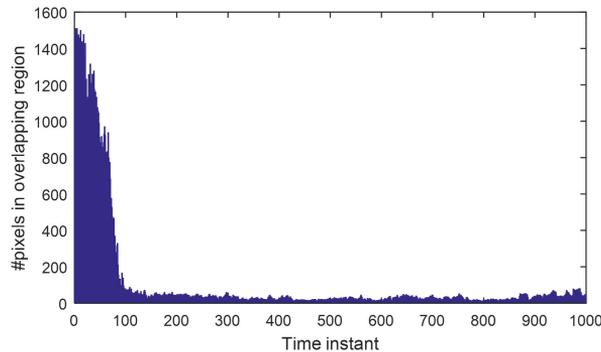

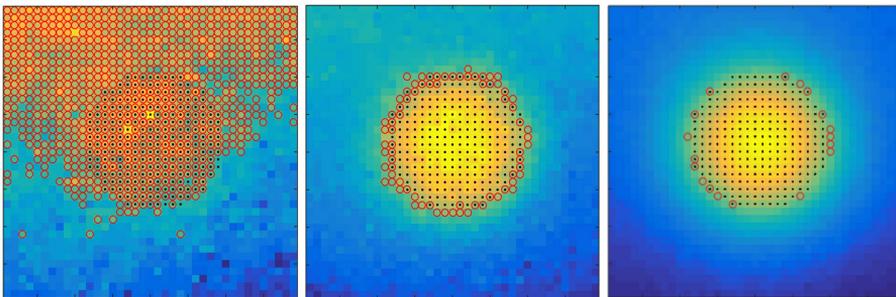Fig. 9. Number of pixels for which the heating-cooling curves lie in the overlapping band [18].

Fig. 10. Thermograms recorded at time step 50 (left), 100 (middle), and 600 (right). "Defect" areas are marked with black dots. Pixels lying in overlapping region are marked with rings [18].

A construction algorithm of the classification tree selects the time step in each internal node and performs a test on the temperature values at this time step. The selected time steps have the strongest discriminative power as splitting variables to ensure that the training set is partitioned into classes with the smallest error. How complex the tree is (how many internal nodes there are) depends on the distribution of data within the component space. CART constructs a decision boundary partitioning the space into different class regions in a specific way. The boundary is approximated using hyperplanes which are perpendicular to the axis of the coordinate system. A test in form of (5) creates a hyperplane $x_i = s_i$, which is perpendicular to axis x_i. The decision boundary, which is a hyperplane perpendicular to one of the axes, can be formed using only one internal node. However, in practice, decision boundaries are usually nonlinear and not perpendicular to the axis. In such cases, many internal nodes are needed to model them.

97

In Fig. 11, the CART constructed for our simple example is shown. As can be seen from this figure, the tree has only three internal nodes which means that only three H-CC points were selected as splitting variables: T_{776}, *i.e.* the temperature after $\tau = 155.2$ s, T_{73} ($\tau = 14.6$ s), and T_{138} ($\tau = 27.6$ s). The classification accuracy estimated using 10-fold cross-validation was 99.47%.
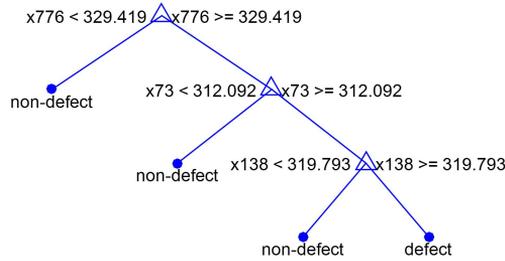

Fig. 11. Optimal CART.

Each H-CC when reduced to only three values $(T_{776}, T_{73}, T_{138})$ can be expressed as a point in a 3-dimensional space. Figure 12 shows these points in 2D for better clarity. The vertical line in this figure, which expresses the split-point for T_{776}, divides the data points almost perfectly into two classes. The two instances from the "non-defect" class, which are wrongly classified in the root node, are separated from the "defect" class points by a horizontal line (split-point T_{73}). Another wrongly classified point is separated in the third node by the third split for T_{138} (made in the third dimension, not shown here).
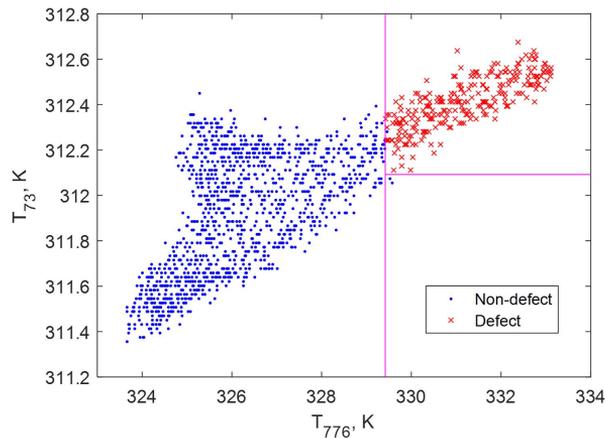

Fig. 12. Training data divided into classes by CART.

The sample described in Section 2 has nine holes, each representing a different thinning region (which we call "defect" in this study). Our task is to recognize the "defect" type (class) corresponding to hole diameter and depth. Let us number the classes of "defects": 1 for the hole in the upper left corner ($\phi\,18$), 2 for hole $\phi\,15$, ..., and 9 for the hole in the bottom right corner ($\phi\,2$). Class 0 is assigned to the "non-defect" pixels. The total number of pixels is $189 \times 190 = 35910$. For each pixel we get H-CC of the length of 1874 time steps. In this case the training set Ω is

composed of 35910 instances and the input vectors x have 1874 components. Each instance is labelled with a class number from 0 to 9. Thus, the problem is much more complex than the problem described above. Instead of two, now the tree recognizes ten classes. Moreover, the classes are not represented equally. The "non-defect" class is the biggest one with 32393 pixels. The most numerous "defect" class, corresponding to hole ϕ 18, has 973 pixels, and the smallest, class 9, has only 21 pixels. In such imbalanced cases, the average classification accuracy is misleading because it can only reflect the underlying class distribution. In our case, the classifier assigning class 0 to all pixels has over 90% accuracy.

To properly evaluate the classifier performance, we use a confusion matrix summarizing the number of correct and incorrect predictions with count values broken down by each class. This gives an insight not only into the average error values being made by the classifier but, more importantly, the types of errors that are being made. The confusion matrix for CART, which was obtained using 10-fold cross-validation, is shown in Fig. 13. As we can see from this figure, the average classification accuracy is 98.72%. As expected, the prediction accuracy for class 0 is the highest: 99.48%. The lowest accuracy is for class 9: 76.19%. 16 instances from this class are recognized correctly and 5 incorrectly as belonging to class 0. Despite the imbalanced distribution of the training data set, the classification results achieved by CART should be considered satisfactory. CART does not ignore any class, even the smallest one containing only 0.0585% instances. Real and predicted class distributions are shown in Fig. 14.

Fig. 13. Confusion matrix for CART.

The ultimate performance of CART is dependent on the L parameter (the threshold number of instances in the node) used in the stopping criterion. To find the optimal L value, CART was trained for L in a range from 1 to 100 and evaluated using 10-fold cross-validation. This
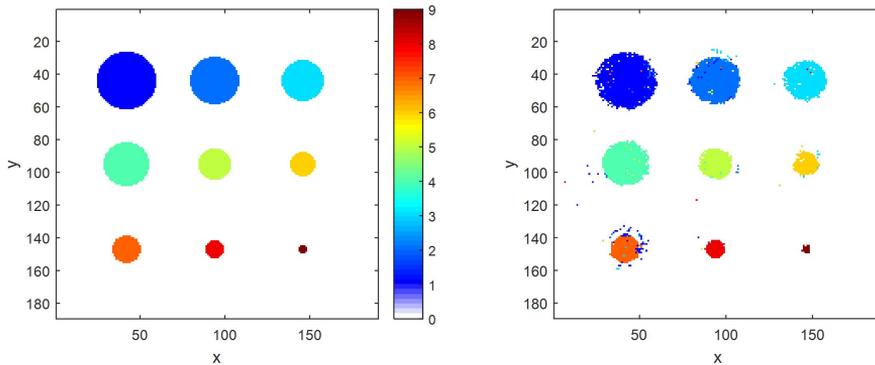
Fig. 14. Real and predicted class distributions.

procedure outputs $L = 3$, the value which achieved the highest score. Such a small value of L may be surprising due to the tendency of model overfitting, but it should be viewed in the context of the size of classes. To ensure correct classification of the smallest classes, the value of this parameter cannot be large. The average tree size was 260 nodes (130 internal nodes and 130 leaves). In the next experiment, we test the CART model when using only:

– the heating fragment of the H-CCs, *i.e.* the fragment from T_1 up to T_{975},
– the cooling fragment of the H-CCs, *i.e.* the fragment from T_{976} up to T_{1874}.

In the first case, the accuracy was 97.47%, and in the second case, the accuracy was 98.14%. Figure 15 shows accuracies observed for each class compared to the case when the whole H-CCs are used as CART inputs. Figure 16 shows predicted class distributions. In both cases the results are only slightly worse than for whole H-CCs.
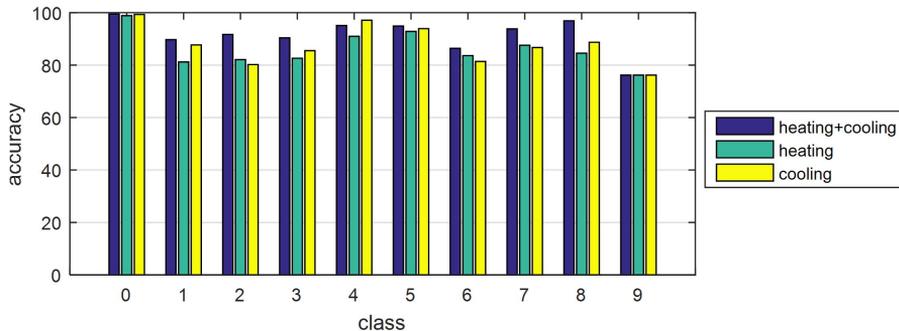

Fig. 15. Accuracies for each class when using the whole H-CCs or their heating and cooling parts.

From a practical point of view, a question arises of how reducing the heating time will affect the classification accuracy. To investigate this, we shortened the length of the patterns x to m first components which correspond to the first H-CC parts, from T_1 to T_m. When m is smaller than 975, *i.e.* the time step of finishing the heating, the x-patterns contain only fragments of the heating parts. When $m > 975$ they contain the whole heating parts and also fragments of the cooling parts. Classification accuracies depending on m in Fig. 17 are shown. As you can see from this figure, accuracies increase with m for all classes with the exception of class 9 where the deterioration in accuracy is observed for x-patterns containing the cooling fragments.
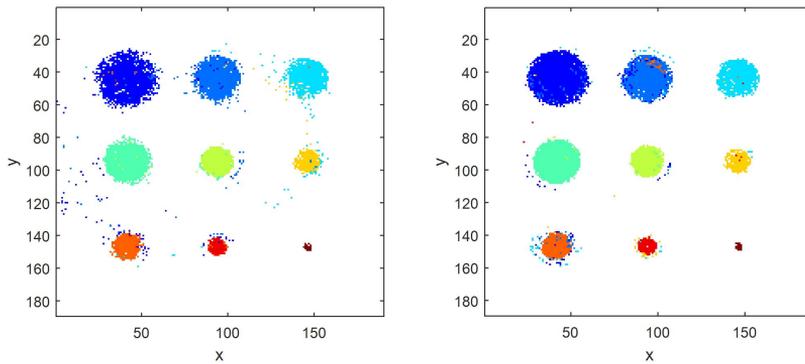
Fig. 16. Predicted class distributions for heating (left) and cooling (right) parts of H-CCs.
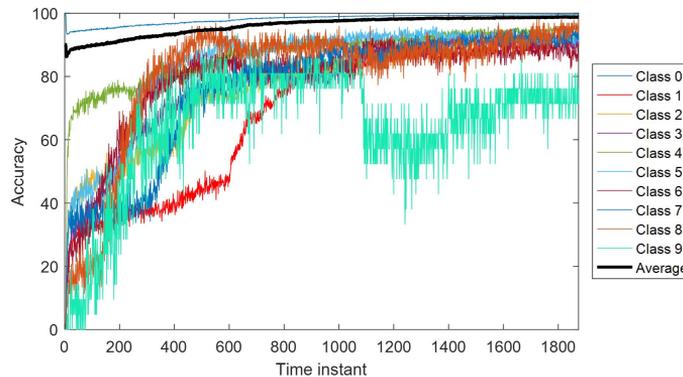

Fig. 17. Classification accuracy depending on real and predicted class distributions for heating (left) and cooling (right) parts of H-CCs.

The H-CCs were recorded at a frequency of 15 Hz. An interesting question is whether the results will be as accurate when we reduce this frequency. A lower frequency means less data resulting in a shorter learning time. At 1 Hz, H-CCs have only 125 time points instead of 1874. A confusion matrix for this case is shown in Fig. 18 and the predicted class distribution in Fig. 19. The average tree size was 302 nodes divided halfway into internal nodes and leaves. As can be seen from the confusion matrix, the classification accuracy has not deteriorated when compared to the 15 Hz case. Average accuracy was 98.81%.

Figure 20 shows accuracy for lower frequencies *i.e.* from 1 Hz to 0.0167 Hz. This corresponds to the length of H-CC from 125 to just 2 points. As you can see in this figure, the accuracy stays at a level of over 98.80% up to 0.1 Hz. In this case, H-CC is composed of only 12 points.

The H-CC can be expressed using PCA. In this case a set of strongly correlated HC-C points are transformed into a set of linearly uncorrelated principal components (PCs). Usually, a small set of the first PCs explains the data variance and can be used instead of the original data. The set of H-CC represented by the first two PCs is shown in Fig. 21. As you can see in this figure, the classes overlap which makes it is difficult to recognize them perfectly. The accuracy of the CART model when k first PCs are used as the model inputs instead of H-CC is depicted in Fig. 22. Note that only five PCs give an accuracy of over 99%.

101

Confusion Matrix

	1	2	3	4	5	6	7	8	9	10	
1	32229 / 89.7%	56 / 0.2%	36 / 0.1%	38 / 0.1%	10 / 0.0%	11 / 0.0%	17 / 0.0%	5 / 0.0%	4 / 0.0%	2 / 0.0%	99.4% / 0.6%
2	46 / 0.1%	889 / 2.5%	22 / 0.1%	0 / 0.0%	0 / 0.0%	0 / 0.0%	0 / 0.0%	5 / 0.0%	0 / 0.0%	0 / 0.0%	92.4% / 7.6%
3	27 / 0.1%	23 / 0.1%	605 / 1.7%	2 / 0.0%	1 / 0.0%	0 / 0.0%	1 / 0.0%	0 / 0.0%	0 / 0.0%	0 / 0.0%	91.8% / 8.2%
4	28 / 0.1%	1 / 0.0%	0 / 0.0%	448 / 1.2%	0 / 0.0%	0 / 0.0%	0 / 0.0%	0 / 0.0%	1 / 0.0%	0 / 0.0%	93.7% / 6.3%
5	20 / 0.1%	0 / 0.0%	0 / 0.0%	0 / 0.0%	555 / 1.5%	7 / 0.0%	0 / 0.0%	0 / 0.0%	0 / 0.0%	0 / 0.0%	95.4% / 4.6%
6	10 / 0.0%	0 / 0.0%	0 / 0.0%	0 / 0.0%	11 / 0.0%	275 / 0.8%	0 / 0.0%	0 / 0.0%	0 / 0.0%	0 / 0.0%	92.9% / 7.1%
7	17 / 0.0%	0 / 0.0%	0 / 0.0%	0 / 0.0%	0 / 0.0%	0 / 0.0%	158 / 0.4%	0 / 0.0%	1 / 0.0%	0 / 0.0%	89.8% / 10.2%
8	9 / 0.0%	4 / 0.0%	2 / 0.0%	0 / 0.0%	0 / 0.0%	0 / 0.0%	0 / 0.0%	215 / 0.6%	0 / 0.0%	0 / 0.0%	93.5% / 6.5%
9	5 / 0.0%	0 / 0.0%	0 / 0.0%	1 / 0.0%	0 / 0.0%	0 / 0.0%	1 / 0.0%	0 / 0.0%	91 / 0.3%	0 / 0.0%	92.9% / 7.1%
10	2 / 0.0%	0 / 0.0%	0 / 0.0%	0 / 0.0%	0 / 0.0%	0 / 0.0%	0 / 0.0%	0 / 0.0%	0 / 0.0%	19 / 0.1%	90.5% / 9.5%
	99.5% / 0.5%	91.4% / 8.6%	91.0% / 9.0%	91.6% / 8.4%	96.2% / 3.8%	93.9% / 6.1%	89.3% / 10.7%	95.6% / 4.4%	93.8% / 6.2%	90.5% / 9.5%	98.8% / 1.2%

Output Class (vertical), Target Class (horizontal)

Fig. 18. Confusion matrix for the sampling rate of 1 Hz.
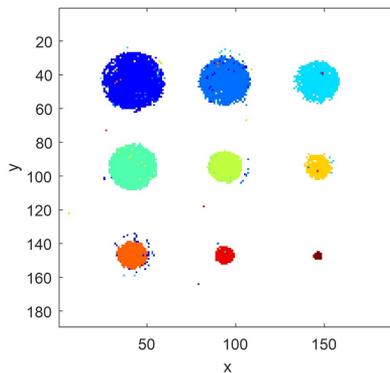

Fig. 19. Predicted class distribution for sampling rate of 1 Hz.
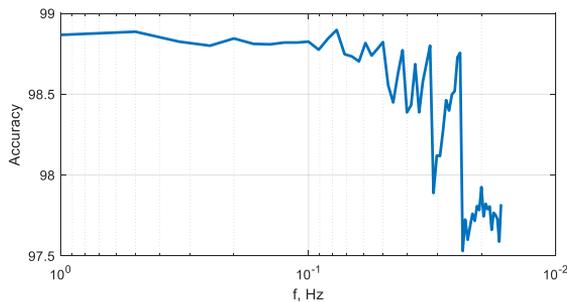

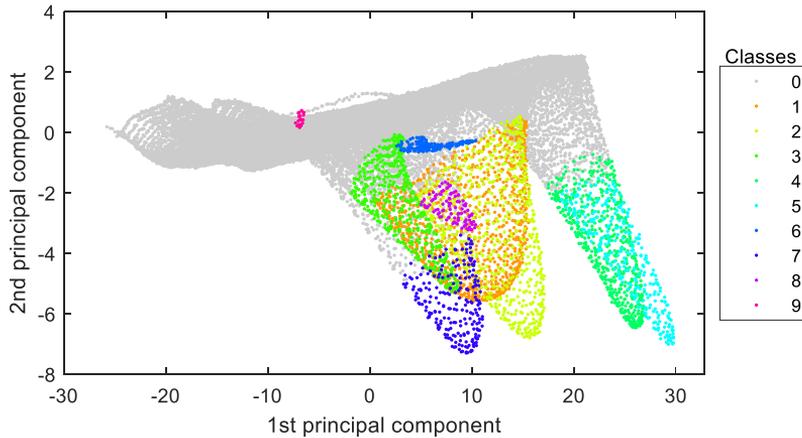Fig. 20. Classification accuracy depending on the sampling frequency.

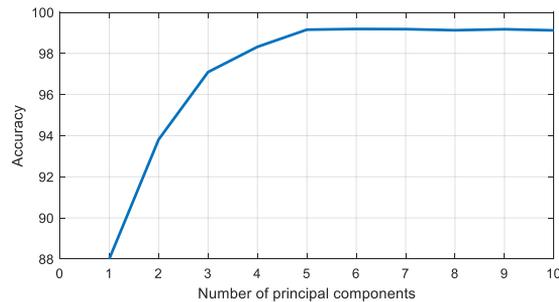Fig. 21. Class distributions in the principal component space.

Fig. 22. Classification accuracy depending on the number of principal components.

It is worth noting that PCA is a loss data compression method. Thus, we lose information on the H-CC curves. But it turns out that this information is redundant and is not necessary for correct classification. The information remaining in principal components is sufficient to obtain an accuracy of over 99%.

5. Conclusions

In this paper, we have proposed using decision trees for analysing thermogram features in the context of detection of thinning regions in homogeneous material. The simulation studies have demonstrated that the heating-cooling curves carry valuable information on the thinning regions. Based on the set of heating-cooling curves recorded for each pixel of the sample in the active thermography process, the tree classifier recognizes nine classes representing different material thicknesses and sizes of the thinning regions. The classification problem is imbalanced *i.e.* the number of instances in the "non-defect" class significantly outnumbers the number of instances in each "defect" class. This usually leads to a bias towards the majority class. The proposed approach copes surprisingly well with imbalanced data. The classifier did not overlook any class. All "defect" classes representing changes in material thickness were recognized by the classifier with high accuracy. The average accuracy was higher than 98%.

In experimental studies, a large heterogeneity of heating of the tested samples was observed. Despite this, the accuracy of the proposed method was found to be high. Using principal components to express heating-cooling curves allows for further compression of data describing the thermograms without a deterioration in the accuracy of the CART model.

The recursive tree-growing process is equipped with an in-built feature selection mechanism. This allows the tree to select only those H-CC points which have the greatest discriminative power and are relevant for recognition of each class of material thickness. This mechanism is one of the great advantages of the tree model. Other advantages, which can be very useful in the detection of changes in material thickness using active thermography, are very fast classification and interpretability, *i.e.* expressing the tree through logical rules which are easy for people to read and understand.

References

[1] Maldague, X. P. (2001). *Theory and practice of infrared technology for non-destructive testing*, John Wiley & Sons.

[2] Minkina, W. & Dudzik, S. (2009). Infrared thermography, John Wiley & Sons, Ltd, Chichester.

[3] Ciampa, F., Mahmoodi, P., Pinto, F. & Meo, M. (2018). Review: Recent Advances in Active Infrared Thermography for Non-Destructive Testing of Aerospace Components. *Sensors*, *18*(2), 609. https://doi.org/10.3390/s18020609

[4] Montanini, R. (2010). Quantitative determination of subsurface defects in a reference specimen made of Plexiglas by means of lock-in and pulse phase infrared thermography. *Infrared Physics & Technology*, *53*(5), 363–371. https://doi.org/10.1016/j.infrared.2010.07.002

[5] Marani, R., Palumbo, D., Galietti, U., Stella, E., & D'Orazio, T. (2016). Automatic detection of subsurface defects in composite materials using thermography and unsupervised machine learning. *2016 IEEE 8th International Conference on Intelligent Systems (IS)*, Bulgaria, 516–521. https://doi.org/10.1109/IS.2016.7737471

[6] Dudzik, S. (2012). Application of the naive Bayes classifier to defect characterization using active thermography. *Journal of Nondestructive Evaluation*, *31*, 383–392. https://doi.org/10.1007/s10921-012-0149-5

[7] Popow, V., & Gurka, M. (2019). Possibilities and limitations of passive and active thermography methods for investigation of composite materials using NDT simulations. *Smart Structures and NDE for Energy Systems and Industry 4.0 (Vol. 10973, p. 109730K)*. https://doi.org/10.1117/12.2518226

[8] Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer-Verlag New York.

[9] Carslaw, H. S. & Jaeger, J. C. (1959). *Conduction of heat in solids* (2nd ed.). Oxford University Press.

[10] Saeed, N., Omar, M. A. & Abdulrahman, Y. (2018). A neural network approach for quantifying defects depth, for nondestructive testing thermograms. *Infrared Physics & Technology*, *94*, 55–64. https://doi.org/10.1016/j.infrared.2018.08.022

[11] Duan, L., Yao, M., Wang, J., Bai, T., & Zhang, L. (2016). Segmented infrared image analysis for rotating machinery fault diagnosis. *Infrared Physics & Technology*, *77*, 267–276. https://doi.org/10.1016/j.infrared.2016.06.011

[12] Lu, P., Gao, B., Feng, Q., Yang, Y., Woo, W. L., & Tian, G. Y. (2017). Ensemble variational Bayes tensor factorization for super resolution of CFRP debond detection. *Infrared Physics & Technology*, *85*, 335–346. https://doi.org/10.1016/j.infrared.2017.07.012

[13] Dudzik, S. (2010). Approximation of thermal background applied to defect detection using the methods of active thermography. *Metrology and Measurement Systems*, 17(4), 621–636. https://doi.org/10.2478/v10178-010-0051-3

[14] Rokach, L., & Maimon, O. Z. (2008). *Data mining with decision trees: theory and applications*. World Scientific Pub Co Inc. https://doi.org/10.1142/9789812771728_0001

[15] Hastie, T. J, Tibshirani, R. J, & Friedman, J. H. (2009). *The Elements of Statistical Learning: Data Mining Inference and Prediction* (Second Edition). Springer. https://doi.org/10.1007/978-0-387-84858-7

[16] Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, *1*, 81–106. https://doi.org/10.1007/BF00116251

[17] Breiman, L., Friedman, J. H., Olshen, R. A. and Stone, C. J. (1984). *Classification and regression trees*, Wadsworth, Inc.

[18] Dudek, G. & Dudzik, S. (2017). Classification tree for material defect detection using active thermography. *Proceedings of 38th International Conference on Information Systems Architecture and Technology ISAT*, Poland, 118–127. https://doi.org/10.1007/978-3-319-67220-5_11

Sebastian Dudzik. Associate Professor, works at Czestochowa University of Technology (CUT), Faculty of Electrical Engineering. His research interest include the applications of active infrared thermography, artificial neural networks and neuro-fuzzy models of heat exchange and non-destructive testing as well as the application of machine learning algorithms in data mining systems. He is the author of two books concerning infrared thermography and applications of artificial neural networks to detection and characterization of material defects and dozens scientific papers.

Grzegorz Dudek received his PhD in electrical engineering from Czestochowa University of Technology, Poland, in 2003 and habilitation in computer science from Lodz University of Technology, Poland, in 2013. Currently, he is an associate professor at the Department of Electrical Engineering, Czestochowa University of Technology. He is the author of two books concerning machine learning methods for load forecasting and evolutionary algorithms for unit commitment and over 100 scientific papers. His research interests include machine learning, artificial intelligence, and their application to practical classification, regression, forecasting and optimization problems.