





## Article

# Enhanced Defect Detection in Carbon Fiber Reinforced Polymer Composites via Generative Kernel Principal Component Thermography

Kaixin Liu <sup>1</sup>, Zhengyang Ma <sup>1</sup>, Yi Liu <sup>1,\*</sup>, Jianguo Yang <sup>1</sup> and Yuan Yao <sup>2,\*</sup>

<sup>1</sup> Institute of Process Equipment and Control Engineering, Zhejiang University of Technology, Hangzhou 310023, China; kxliu@zjut.edu.cn (K.L.); zymazjut@163.com (Z.M.); yangjg@zjut.edu.cn (J.Y.)

<sup>2</sup> Department of Chemical Engineering, National Tsing Hua University, Hsinchu 30013, Taiwan

\* Correspondence: yliuzju@zjut.edu.cn (Y.L.); yyao@mx.nthu.edu.tw (Y.Y.); Tel.: +886-3-5713690 (Y.Y.)

**Abstract:** Increasing machine learning methods are being applied to infrared non-destructive assessment for internal defects assessment of composite materials. However, most of them extract only linear features, which is not in accord with the nonlinear characteristics of infrared data. Moreover, limited infrared images tend to restrict the data analysis capabilities of machine learning methods. In this work, a novel generative kernel principal component thermography (GKPCT) method is proposed for defect detection of carbon fiber reinforced polymer (CFRP) composites. Specifically, the spectral normalization generative adversarial network is proposed to augment the thermograms for model construction. Sequentially, the KPCT method is used by feature mapping of all thermogram data using kernel principal component analysis, which allows for differentiation of defects and background in the dimensionality-reduced data. Additionally, a defect-background separation metric is designed to help the performance evaluation of data analysis methods. Experimental results on CFRP demonstrate the feasibility and advantages of the proposed GKPCT method.

**Keywords:** infrared non-destructive assessment; polymer composite; deep learning; generative adversarial network; thermographic data analysis; kernel principal component analysis



**Citation:** Liu, K.; Ma, Z.; Liu, Y.; Yang, J.; Yao, Y. Enhanced Defect Detection in Carbon Fiber Reinforced Polymer Composites via Generative Kernel Principal Component Thermography. *Polymers* **2021**, *13*, 825. <https://doi.org/10.3390/polym13050825>

Academic Editor: Javier González-Benito

Received: 31 January 2021

Accepted: 2 March 2021

Published: 8 March 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Infrared thermography (IRT), an attractive non-destructive assessment technique, is applied widely in the quality assessment of composite structures due to its fast response, wide-range of measurement, and intuitive results [1–4]. However, especially in the task of defect assessment, IRT is difficult to achieve the desired inspection results. The non-uniform background and noise in thermal images often hinder the accuracy and efficiency of IRT detection due to the influence of experimental setup and environment [5,6]. Therefore, it has become common to adopt thermographic data analysis methods to improve the IRT performance.

In recent years, machine learning methods have provided alternative solutions and achieved promising results for industrial data/image analysis tasks [7–9]. For IRT defect identification, several thermographic data analysis methods have been developed to enhance the visibility of defects in thermograms. These methods include principal component thermography (PCT) [10] and its variants (such as sparse PCT [11], candid covariance-free incremental PCT [12], generative PCT [13]), thermographic signal reconstruction [14], penalized least squares [15], pulsed phase thermography [16], independent component thermography [17], manifold learning methods [18–20], autoencoder methods [21,22], other deep learning methods [23–25], etc. Most of the above-mentioned methods can be considered as a kind of metric learning [26]. For example, PCT is an unsupervised linear metric learning method with simplicity and scalability, which can be used to improve the defect detection performance of IRT.

However, the thermographic data recorded by IRT are often nonlinear due to uneven heating and thermal anisotropy of composite materials [21]. Direct observation of this property is that thermal images contain heavily inhomogeneous background and noise, which masks defect information. Therefore, nonlinear models may be more suitable than linear candidates in the task of defect identification. Nevertheless, there are only a few nonlinear thermographic data analysis methods developed [18–20]. Another challenge of thermographic data analysis is that the thermograms are typically limited because of the shallow detection depth of the thermal imager, which indeed restricts the defect detection capabilities of thermal imaging methods. While this difficulty is solved by using a data expansion strategy in the generative PCT (GPCT) approach [13], there is still room for improvement in the performance of defect identification because the nonlinearity of thermal image data is not considered in it.

In this work, a generative kernel principal component thermography (GKPCT) method for defect detection of carbon fiber reinforced polymers is proposed to enhance the visibility of defects. The main contributions are summarized as the following three aspects. First, a kernel principal component thermography (KPCT) method is developed to extract nonlinear features from thermographic data. Second, a spectral normalized generative adversarial network (SNGAN) [27] is integrated with KPCT to build a GKPCT modeling and analysis framework. A data augmentation strategy is proposed to generate a number of informative thermographic images to improve the feature mapping performance of the KPCT model, thus improving the defect detection. Third, a local pixel standard deviation (LPSD) evaluation index is proposed for comparing different thermal image analysis methods for defect detection. Two case studies demonstrate the feasibility and advantages of the GKPCT approach.

The rest parts are so organized. In Section 2, the framework of GKPCT with its algorithmic details is proposed. Section 3 introduces the experimental research of IRT defect detection on two carbon fiber reinforced polymer (CFRP) specimens. In Section 4, experimental results are reported. Finally, the conclusions are summarized in Section 5.

## 2. Methodology

### 2.1. KPCT–Nonlinear Thermographic Data Analysis Approach

Since thermographic data tend to be nonlinear, a linear thermography approach based on dimensionality reduction can lose the high-dimensional structure of the data. Therefore, a nonlinear model may be applicable and worth trying in the analysis of thermographic data. In this work, the KPCT method is proposed for learning nonlinear features in thermograms, so that linearly indistinguishable defects are separated from the background. Its principle can be summarized as mapping the original thermographic data into a high-dimensional feature space  $H$  by a nonlinear transformation  $\Phi$ , followed by principal component analysis (PCA) in  $H$  [28]. As a result, KPCT solves the problem of nonlinearity and feature extraction of thermal image data, and the defect morphology is identified. As a point of attraction for this work, it is worth illustrating that nonlinear mappings are learned through kernel methods, which is different from the developed methods for manifold and deep learning thermographic data analysis [18–25]. The following details can help to understand the technical process of the KPCT method.

The KPCT approach begins with IRT experiments. After the experiment, an infrared dataset is obtained, which can be expressed as a three-dimensional (3-D) matrix with dimensions of  $n_t \times n_x \times n_y$ . It means that  $n_t$  frame thermal images with a size of  $n_x \times n_y$  are selected as the original data for modeling. To meet the needs of KPCT data analysis later, this 3-D data is converted into a two-dimensional (2-D) matrix  $X$  with  $n_t \times (n_x \times n_y)$  dimensions. Subsequently, a two-step data preprocessing procedure is executed on the  $X$ . In detail, first, the rows of the matrix are subjected to a max-min normalization to unify the scale to 0–1, after which the columns are centered to remove the inhomogeneous background of the data.

After the data is prepared, KPCT selects pixel temperature signals (column vectors  $\mathbf{x}_i$  of matrix  $\mathbf{X}$ ) as samples and projects them to high-dimensional feature space  $H$  through nonlinear mapping  $\Phi$ . The covariance matrix in the  $H$  is:

$$\mathbf{S}^H = \frac{1}{m} \sum_{i=1}^m \Phi(\mathbf{x}_i) \Phi(\mathbf{x}_i)^T, \quad (1)$$

where  $\Phi(\mathbf{x}_i)$  is a sample in feature space  $H$ ,  $m$  is the number of samples and  $m = n_x \times n_y$ .

Because  $\Phi(\mathbf{x}_i)$  is unknown, the eigenvalues and eigenvectors of  $\mathbf{S}^H$  cannot be obtained directly through PCA. Herein, defining an  $m \times m$  kernel matrix  $\mathbf{K} = \Phi(\mathbf{x}_i)^T \Phi(\mathbf{x}_i)$ , and the eigenvalue decomposition of  $\mathbf{K}$  can be expressed as:

$$\mathbf{K}\mathbf{w} = \lambda\mathbf{w}, \quad (2)$$

where  $\lambda$  is the eigenvalue matrix of  $\mathbf{K}$ ,  $\mathbf{w}$  is the corresponding eigenvector matrix.

Equation (2) is multiplied by  $\frac{1}{m} \sum_{i=1}^m \Phi(\mathbf{x}_i)$  on the left to perform identity transformation to obtain:

$$\frac{1}{m} \sum_{i=1}^m \Phi(\mathbf{x}_i) [\Phi(\mathbf{x}_i)^T \Phi(\mathbf{x}_i) \mathbf{w}] = \frac{1}{m} \sum_{i=1}^m \Phi(\mathbf{x}_i) \lambda \mathbf{w}. \quad (3)$$

The combined Equations (1) and (3) are equivalent to the following equation:

$$\mathbf{S}^H [\Phi(\mathbf{x}) \mathbf{w}] = \lambda [\Phi(\mathbf{x}) \mathbf{w}]. \quad (4)$$

Observing the above equation shows that the loading matrix of  $\mathbf{S}^H$  is exactly  $\Phi(\mathbf{x}) \mathbf{w}$ . Although  $\Phi(\mathbf{x})$  is unknown, the loading matrix is not the ultimate goal. The component of the sample  $\Phi(\mathbf{x})$  in the projection direction  $\Phi(\mathbf{x}) \mathbf{w}$  is  $\Phi(\mathbf{u})^T \Phi(\mathbf{u}) \mathbf{w}$ . Consequently, the score matrix  $\mathbf{T}$  of the KPCT of can be calculated as:

$$\mathbf{T} = \mathbf{K}\mathbf{w}. \quad (5)$$

At this point, the key to solving the matrix  $\mathbf{T}$  is to calculate the  $\mathbf{K}$  matrix. After that, the  $\mathbf{w}$  matrix is calculated by the formula (2). In this work, the  $\mathbf{K}$  matrix is calculated by the Gaussian kernel function, and its mathematical expression is as follows:

$$\mathbf{K}_{ij} = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{\tau}\right), \quad (6)$$

where  $\mathbf{x}_i, \mathbf{x}_j$  for the  $i$ -th and  $j$ -th samples of the original data, respectively, and  $\tau$  is the width of the Gaussian kernel.

After the  $\mathbf{T} = \mathbf{R}^{m \times d}$  matrix is obtained, an image reconstruction procedure is applied to visualize the defects. Specifically, the column vector of the matrix  $\mathbf{T}$  is reconstructed as a 2-D matrix of size  $n_x \times n_y$ , and  $d$  principal components (PCs) thermal images are obtained accordingly. However, checking each PC image is both time-consuming and laborious. As a variant of the PCT method, as with PCT, a few PCs retain most of the information about the original thermographic data [8]. Therefore, only the first few columns of  $\mathbf{T}$  need to be investigated to reconstruct the PC images. These PC images contain more pronounced geometric properties of defects than the original thermograms, which shows the defect information (including position, shape, and size) more clearly than the original thermal images.

## 2.2. GKPCT-KPCT Approach Based on Image Augmentation

Recently, the GAN family has become increasingly popular and has demonstrated attractive performance in data augmentation tasks [29,30]. SNGAN, an improved GAN method using weight normalization (also known as spectrum normalization) technique

was proposed to address the problem of unstable GAN training [27]. Like GAN, SNGAN has basically the same structure: a generator  $G$  for learning the model distribution and a discriminator  $D$  for distinguishing the model distribution from the target. Different from GAN, in SNGAN’s discriminator network, the weight matrix of each layer is spectrally normalized. The advantage of this is that SNGAN is stable and easy to implement during training [27]. The details of the algorithm can be described as follows:

The weight matrix  $\mathbf{W}^l$  is spectrally normalized for each layer of  $D$  so that it satisfies the Lipschitz constraint  $\sigma(\mathbf{W}^l) = 1$ :

$$\mathbf{W}_{SN} = \frac{\mathbf{W}^l}{\sigma(\mathbf{W}^l)} \tag{7}$$

where  $\mathbf{W}^l$  represents the weight matrix of layer  $l$  ( $l = 1, \dots, L$ ),  $\sigma(\mathbf{W}^l)$  denotes the spectral norm of  $\mathbf{W}^l$ ,  $\mathbf{W}_{SN}$  is the weight matrix after spectral normalization.

Here, SNGAN is introduced as an image augmentation strategy into the IRT data analysis model to expand the number of thermal images, and a SNGAN-based GKPCT is proposed to improve the defect detection performance of KPCT. The motivation for this is that thermal image analysis models often perform poorly due to insufficient training data. As shown in Figure 1, the entire GKPCT framework consists of two parts: the SNGAN thermal image generation structure on the left and the KPCT thermal image analysis structure on the right. With the initial infrared dataset, SNGAN architecture is implemented to generate thermal image sequences. The convolutional layers of the two networks both use the rectified linear unit (ReLU) activation function to prevent the gradient from over-expanding. The biggest difference between them is that each layer weight matrix of the  $D$  network is spectrally normalized. In addition, all convolution kernel dimensions are set to  $3 \times 3$ , the optimizer selected is Adam, and the best learning rate is set to 0.0002.

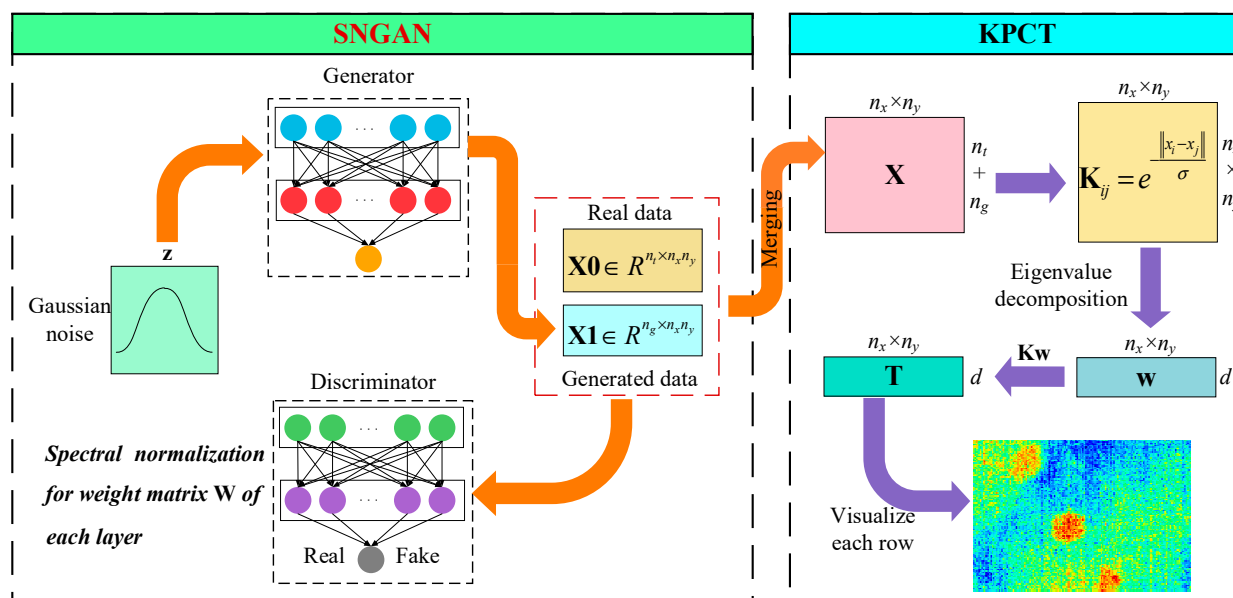


Figure 1. The GKPCT modeling and analysis framework.

After the above SNGAN model was trained,  $n_g$ -frame fake images with a similar distribution to the original thermal image data were generated and expanded to the training data. A merged thermographic data with a size of  $(n_t + n_g) \times (n_x \times n_y)$  is used for subsequent KPCT analysis. Consequently, GKPCT eliminates most of the inhomogeneous background and makes the geometric characteristics of the defect areas visible in the PC images, thus improving the efficiency and accuracy of defect assessment by IRT techniques.

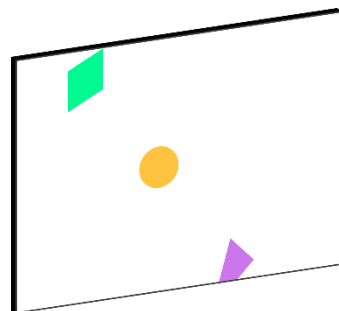


### 3. Experiments

#### 3.1. Manufacturing and Data Acquisition of CFRP Specimen 1

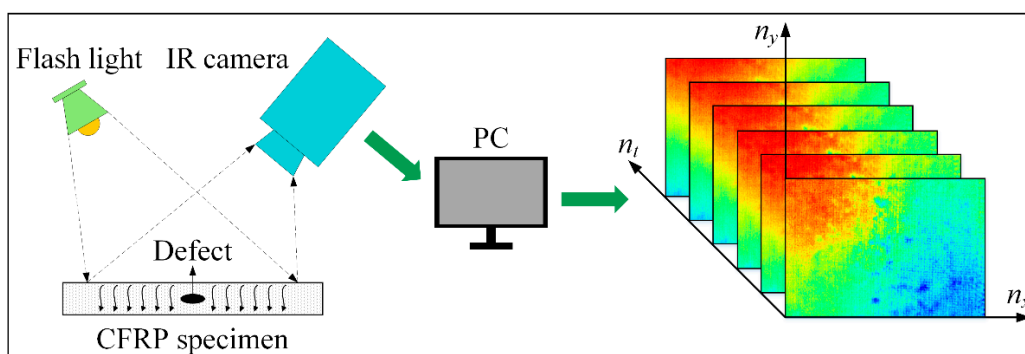
To test the effectiveness of the proposed GKPCT, IRT defect detection and data analysis were performed on the CFRP specimen. Considering that the defect properties of the natural CFRP are unknown, it is difficult for GKPCT to judge the performance of its defect assessment. In this experiment, an artificial CFRP specimen containing multiple defects was manufactured. It is made from 20 carbon fiber sheets and epoxy resin using the resin transfer modeling [31] process, with an overall thickness of about 1 cm and dimensions of about 18 cm × 18 cm. In more detail, the CFRP was fabricated by injecting thermosetting resin into fiber preformed rods placed in a closed mold. Prior to resin injection, three Teflon strips (of different shapes but all equal to about 3 cm<sup>2</sup> in area) were inserted into the carbon fiber sheet to form the preset defects.

As shown in Figure 2, the defect pattern inside the artificially fabricated CFRP specimen 1 can be observed. From the top view, the deepest diamond-shaped defect is in the upper left corner where the Teflon strip is covered by three layers of carbon fiber sheets (about 0.15 cm deep). The shallowest Teflon strip is covered by a layer of carbon fiber sheet (about 0.05 cm deep), which is a trapezoidal defect located in the lower right corner. The one remaining defect is circular (about 0.10 cm deep) in the middle position made of a Teflon strip covered by two layers of carbon fiber sheets. It needs to be noted that these defects are invisible after the epoxy is injected. The purpose of this work is to develop a thermography data analysis method that can facilitate IRT defect detection performance, leading to more accurate defect estimation.



**Figure 2.** Illustration of defective regions in a CFRP specimen 1.

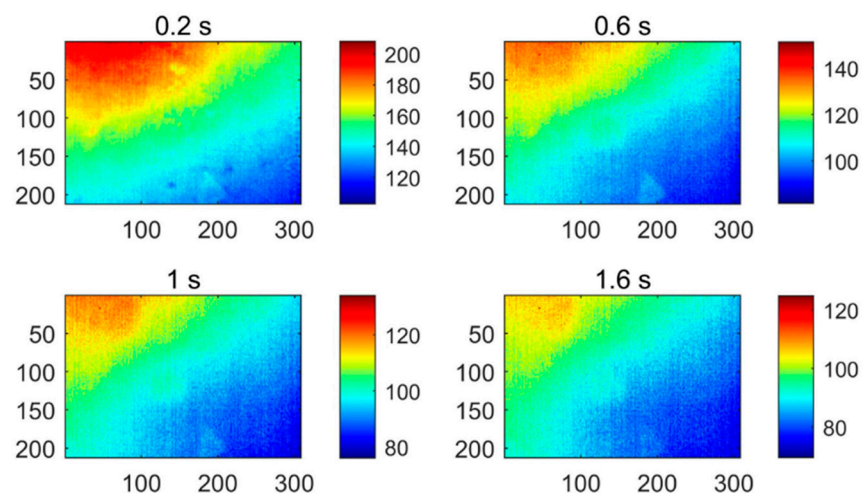
Pulsed thermography (PT), one of the fastest and simplest IRT techniques, was used in this study [16,17]. The experimental setup of the PT system in our laboratory is shown in Figure 3, which consists of a flashlight that can deliver 3200 W of energy for 3 ms duration, an infrared camera (TAS-G100EXD, NEC, Minato-ku, Tokyo, Japan.) with a COMS sensor and an uncooled focal plane array (UFPA) detector, and a computer for image storage and analysis. Specifically, the process of PT detecting CFRP internal defects is:



**Figure 3.** The PT inspection equipment and acquired thermogram sequence.

The CFRP specimen 1 is heated by a thermal pulse excited by a working flash, which passes through the specimen after the thermal front contacts the surface. Typically, surface temperatures in healthy areas will decrease evenly over time. However, subsurface discontinuities caused by defects can cause the surface to produce an abnormal temperature response. This temperature pattern can be recorded during the specimen 1 cooling phase by the infrared camera. The camera is sensitive to infrared light in the temperature range of  $-40$  to  $1500$  °C with a spectral range of  $8\text{--}14$   $\mu\text{m}$ . Its sampling frequency and spatial resolution of the camera are 30 frames per second and  $320 \times 240$  pixels, respectively. At the end of the experiment,  $308 \times 212$  pixels were selected as the region of interest (ROI) for each thermogram. As a result, 90 frames of thermal images were captured within 3 s for subsequent analysis.

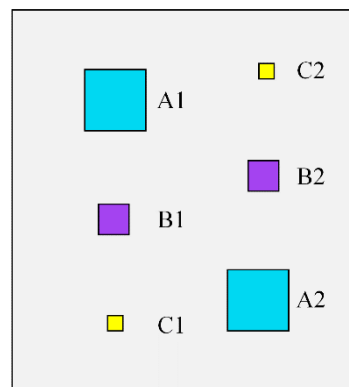
Figure 4 shows four of the thermograms obtained by implementing the PT technique for testing CFRP specimen 1, which were sampled at different instances. It can be found that the thermograms contain a large amount of uneven background and noise, which masks the defect information. This result can be attributed to the limited detection capability of the PT system itself. PT is easily affected by various factors, such as reflections from the environment, changes in emissivity, uneven heating, and changes in the surface geometry of the detection target. Although the effects of similar surface conditions can be controlled, energy inhomogeneity is an inevitable problem in the experimental environment. It is for this reason that more and more thermal image analysis methods are developed to promote the performance of PT.



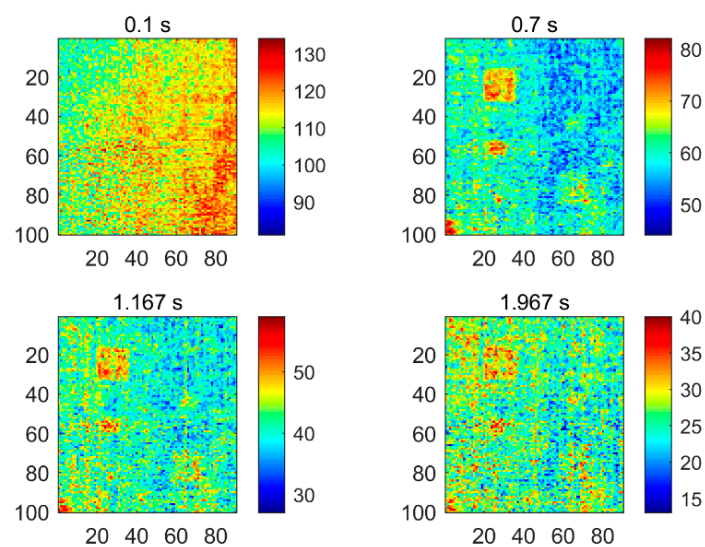
**Figure 4.** The PT thermograms of specimen 1 recorded at different sampling instances.

### 3.2. Manufacturing and Data Acquisition of CFRP Specimen 2

The second CFRP specimen was manufactured in a similar way to the first one, which contains six defects. Figure 5 shows the defect distribution in the region of interest of this specimen. The six defects are square and range in size from  $1.6\text{ cm} \times 1.6\text{ cm}$  to  $0.4\text{ cm} \times 0.4\text{ cm}$ . The three defects on the left, i.e., defects A1, B1, and C1, are located under one layer of carbon fiber sheet, while the defects on the right, i.e., defects A2, B2, and C2, are located under two layers of sheets. After the specimen was fabricated, the PT technique was used for nondestructive testing. A total of 59 thermal images with a size of  $90 \times 100$  were obtained in the experiment for the subsequent analysis. Figure 6 shows some of the original thermal images recorded at different moments, from which it can be observed that the defects are difficult to be identified due to the existence of inhomogeneous background and noise.



**Figure 5.** Illustration of defective regions in a CFRP specimen 2.



**Figure 6.** The PT thermograms of specimen 2 recorded at different sampling instances.

#### 4. Results and Discussion

To evaluate the experimental results more objectively, a commonly-used signal-to-noise ratio (SNR) index is introduced [10,17,32–34]. Its mathematical expression is as follows:

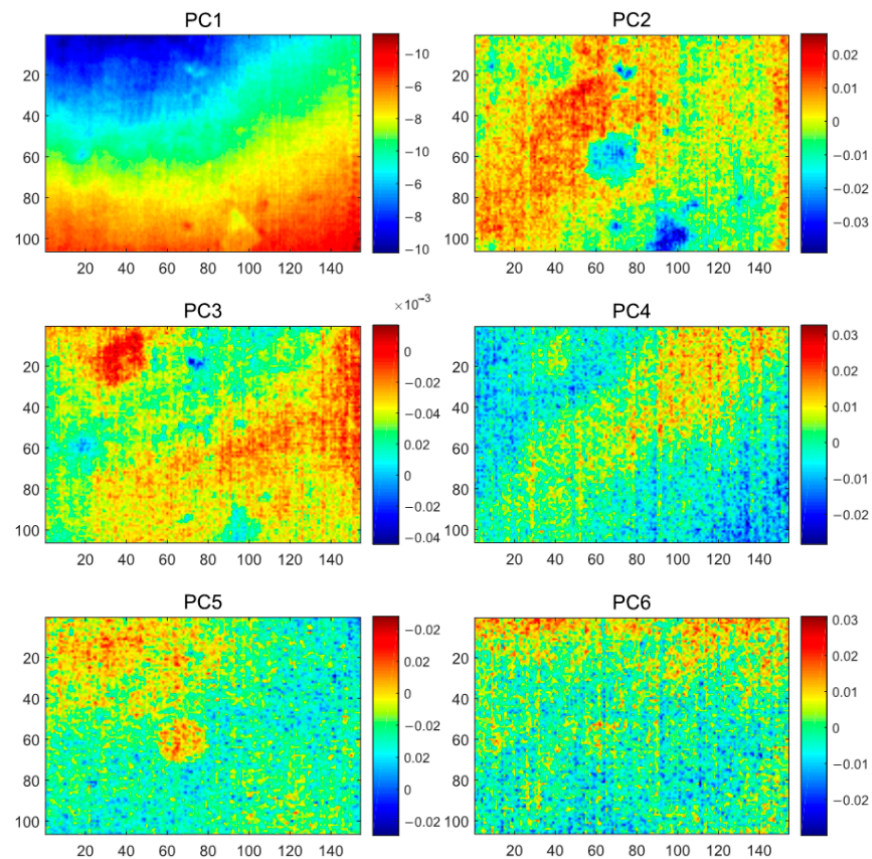
$$SNR = \frac{M_{def} - M_{in}}{\sigma_{in}}, \quad (8)$$

where  $M_{def}$  is the mean of pixel values in the defective area,  $M_{in}$  is the mean of pixel values in the non-defective area, and  $\sigma_{in}$  is the standard deviation of pixel values in the non-defective area. The higher the SNR value, the greater the ability of the method to identify defects. Therefore, the SNR index allows comparing the performance of different thermographic data analysis methods.

##### 4.1. Thermal Image Data Analysis of Specimen 1

In this section, PCT, GPCT are used to compare with the proposed KPCT and GKPCT methods. PCT, which uses PCA techniques to process and analyze thermal images [12]. As a representative data compression method, PCT has the advantages of high noise reduction capabilities, no parameter limitations, and low computational cost in the field of thermography. And GPCT is a deep learning thermographic data analysis method developed on the basis of PCT to address the shortage of training data. The results of PCT and GPCT on CFRP specimens with three defects are shown in Figures 5 and 6, respectively.

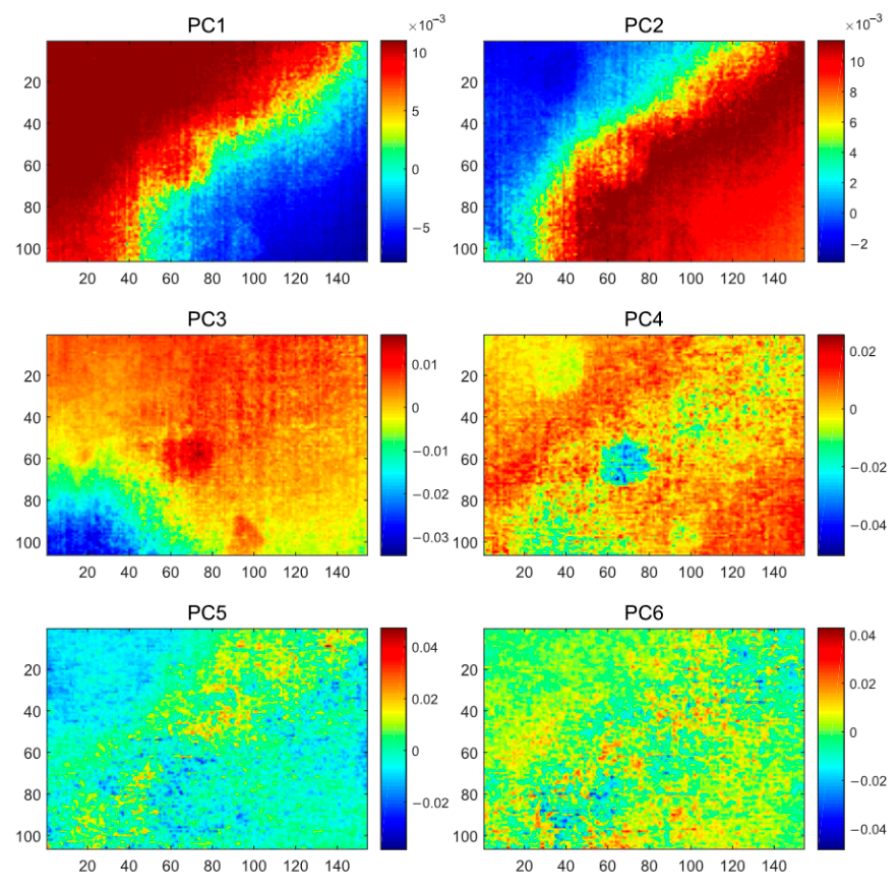
The PCT results, as shown in Figure 7, the first six PC plots show three types of information as inhomogeneous background, noise, and defects. PC1 shows a heavily inhomogeneous background, while PC2 and PC3 highlight the geometric properties of the defects at different depths. Compared to the original PT images, the PCT method better reveals the number and rough location of the defects. However, as a linear analysis method, PCT neither completely separates the background and defects in the thermographic data, nor does it incorporate all the defects that are centralized in a single PC for easy manual inspection.



**Figure 7.** The PCT analysis results of the CFRP specimen 1.

As with PCT, the first six PCs of GPCT were visualized as in Figure 8. Even though GPCT overcame the challenge of insufficient training data for the PCT method, the expected improvement in defect detection was not achieved. PCs 1–4 contained severe inhomogeneous backgrounds, which added to the difficulty of defect identification. Compared with PCT, the defect detection ability of GPCT becomes worse, which may be attributed to the poor nonlinear processing ability of the GPCT method that extracts the inhomogeneous background as the main information after the data augmentation, and the defect information is not focused on. However, although the defect information is masked by the inhomogeneous background, the boundaries of defects in the GPCT results (PC3 and PC4) are still clearer than those in the PCT results (PC2) under careful observation.

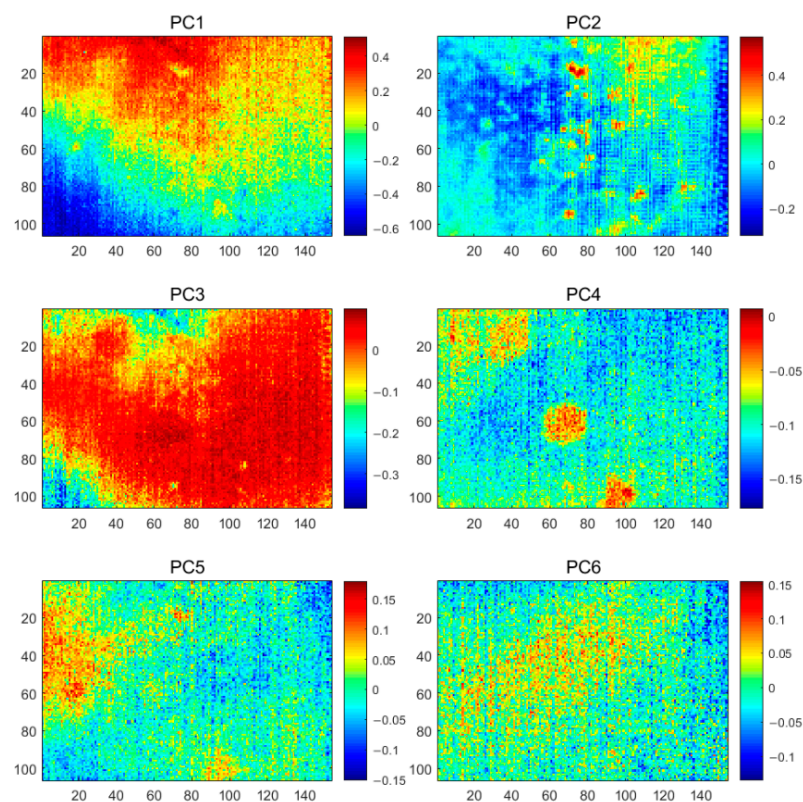




**Figure 8.** The GPCT analysis results of the CFRP specimen 1.

KPCT, application of the kernel PCA algorithm to thermography. The value of the Gaussian kernel width was chosen to be 1.4 in the execution of the algorithm. By mapping the nonlinearly indistinguishable thermographic data into a high-dimensional space, it can be linearly separable. In doing so, the performance of KPCT is usually improved over PCT (besides the identification of diamond-shaped defects), which has a stronger ability to extract defect information. Figure 9 shows the first six PC images obtained by KPCT. It can be observed from the figure that the PC4 image shows all the prefabricated defects, while PC1 and PC3 focus on the severe uneven background. Compared to PCT, the defects were concentrated in a single PC4 thermal image in the results of KPCT. However, the KPCT still detects defects with poor visibility and the profile of the defects is not clear. As shown in the PC3 image, the faintly visible defect is covered by a severely uneven background. One of the main reasons for the unsatisfactory performance of KPCT is that the proportion of information of uneven background and defect features is unbalanced in the original thermal image. The nonlinear mapping capability of KPCT is used more to extract the inhomogeneous background. Comparing Figures 7 and 9, the first six PCs of KPCT contribute more inhomogeneous background information (KPCT has two PC images, i.e., PC1 and PC3, whereas PCT has only one PC1 image).





**Figure 9.** The KPCA analysis results of the CFRP specimen 1.

In this work, the GKPCT method is applied to analyze the thermograms to enhance the visibility of defects. The algorithm is calculated and implemented based on the TensorFlow framework on a computer with 16 G memory, Intel<sup>®</sup> Core<sup>™</sup> i7 CPU, and Windows 7 system. The value of the Gaussian kernel width  $\tau$  was chosen to be 5.6 in the algorithm. As the core of the GKPCT approach, there is a need to determine how many thermal images generated by SNGAN are beneficial for overall model performance. To this end, this work investigates the defect detection results of GKPCT when generating 20, 40, 60, 80, 100, 120, 140, 160 thermal images. It was found that the analysis results of the GKPCT method had the largest *SNR* values when 80 thermal images were generated. Some thermal images generated by SNGAN are shown in Figure 10. As with the original thermal image, the generated thermal image contains noise and severe background inhomogeneity, which makes it difficult to visually detect defective features in the image.

To further study the characteristics of SNGAN-generated data, t-distributed stochastic neighborhood embedding (t-SNE) is used to visualize the distribution of original and generated data [35]. t-SNE results for 2D visualization are shown in Figure 11, where the order of original thermograms in the time series is marked with numbers. The SNGAN-generated images in the figure are distributed around the middle of the original PT image sequence. A reasonable explanation is that the first few frames of the PT test thermograms contain a lot of noise and inhomogeneous background, while the last few frames have very little temperature decay. Compared to the two phases, the thermograms in the middle period contain more defect detection information and become the main information for SNGAN training and learning. Eventually, SNGAN generates thermal images with the same distribution of training data as the middle part.

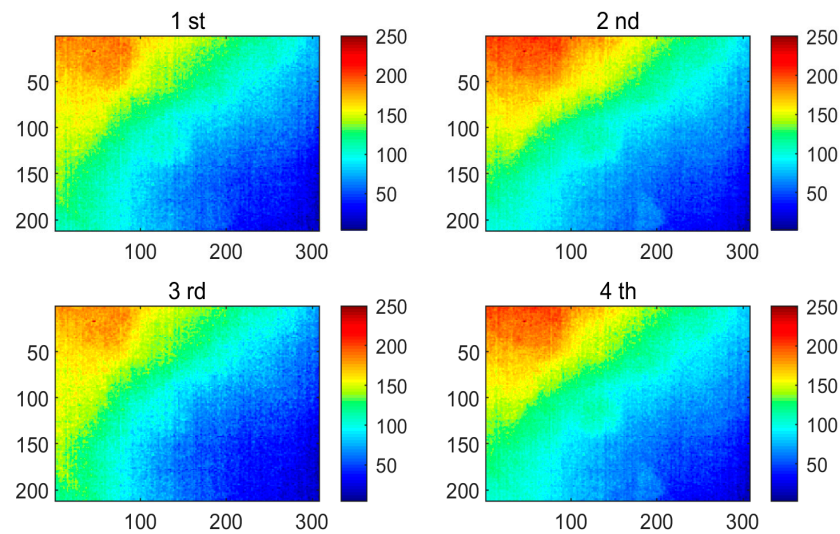


Figure 10. Several simulated SNGAN thermograms of the CFRP specimen 1.

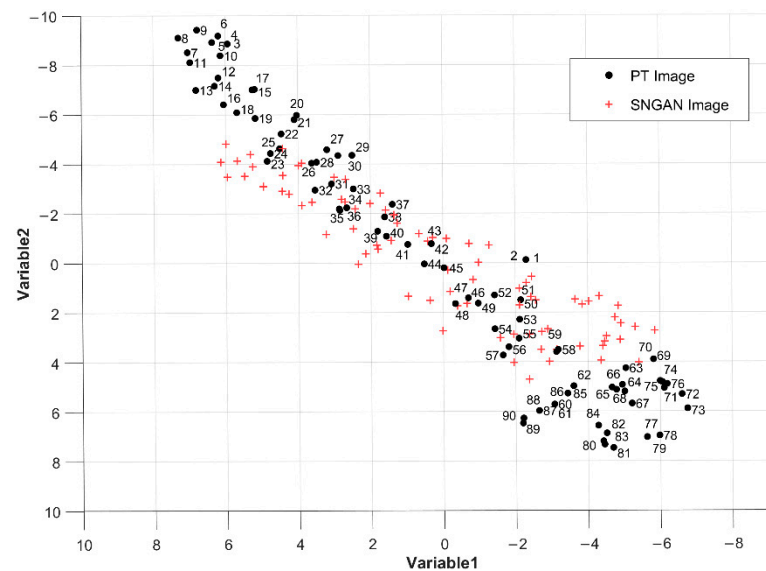
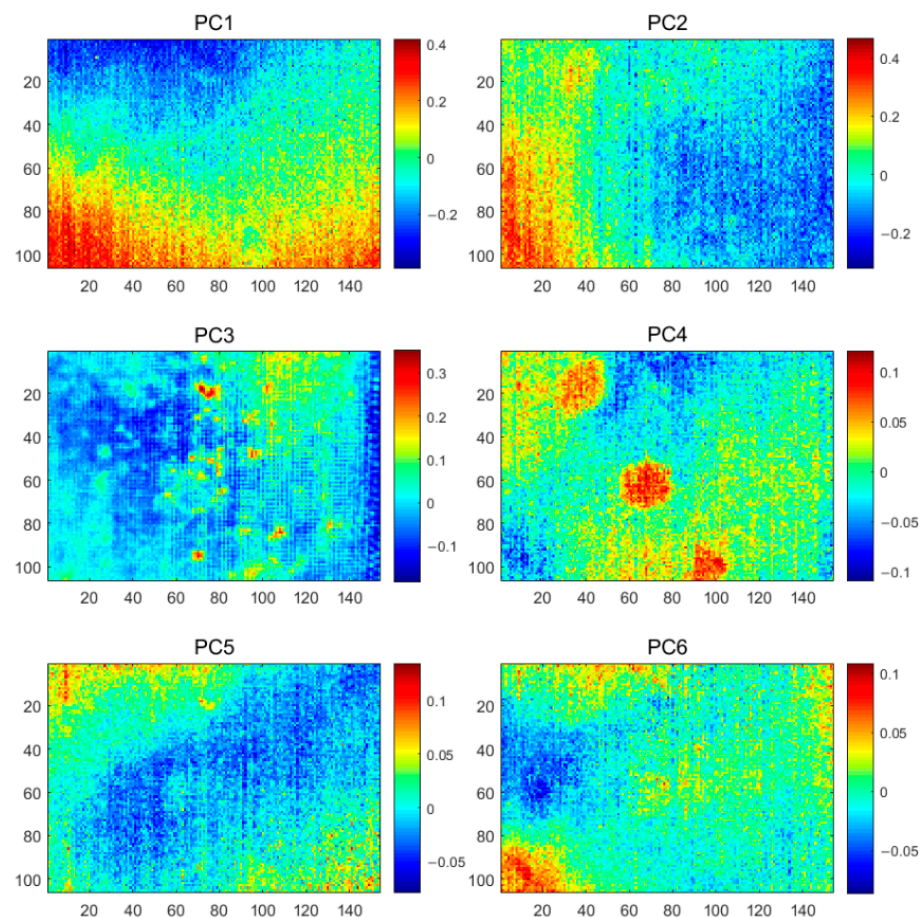


Figure 11. The t-SNE visualization of PT thermograms and SNGAN thermograms.

The results of the proposed GKPCT method are shown in Figure 12, where the PC4 image extracts the morphology of all defects with clear defect boundaries. The PC1 and PC2 images have extracted uneven background. Compared to GPCT and KPCT, the inhomogeneous background information is reduced in the GKPCT results, which is attributed to the fact that SNGAN expands the dataset to increase the training data of the model and reduces the degree of information imbalance. Additionally, although this paper is to validate the feasibility of GKPCT by man-made defects, some defects not intended for man-made purposes can also be detected (cavity defects showed by PC3 in Figure 12). This also indirectly indicates the ability of GKPCT to be generalized in practical applications. These results show that GKPCT has a good ability to separate information and identify defects, which is an effective thermographic data analysis method that can improve IRT's ability to evaluate internal defects in CFRP specimen.



**Figure 12.** The GKPCT analysis results of the CFRP specimen 1.

Table 1 compares the *SNR* values of the original thermal images and several methods, and the largest *SNR* values obtained by each method are listed in the table. The *SNR* values and defect visualization map performance of each method are basically the same: 1. PCT detects the presence of defects compared to the original thermal image, and performs well especially in the detection of diamond-shaped defects (which is superior to other methods); 2. GPCT, a data-augmented PCT method, detects a large improvement in the contour and clarity of defects in the resulting thermal image compared to PCT (in the detection of circular defects good), but the *SNR* value is calculated to be small due to the severe inhomogeneous background in the thermal image masking the defects; 3. KPCT, a method proposed for the characteristics of the nonlinearity of thermal image data, the defect information is effectively separated from the inhomogeneous background, noise and other information. Additionally, the inspection results are concentrated in one thermal image, which saves inspection time and effort. However, subject to the factor of insufficient training data, the visibility of defects in the detection results of the KPCT method is still low, which in scenarios requiring higher detection accuracy, the accurate detection of defects can only depend on the knowledge and skills of the inspector and can easily lead to ambiguity in the defect evaluation results; 4. GKPCT, a data-augmented and nonlinear PCT method, has strong defect-recognition compared with several comparison methods capability (sub-high *SNR* for diamond-shaped defects, similar to the detection results of PCT methods). The method further improves the performance of the KPCT model in the task of defect detection and identification with the data augmentation strategy, which overcomes the challenge of insufficient training data and also solves the problem of difficult information separation due to the nonlinearity of thermal image data, and enhances the visibility of defects in the detection results.

**Table 1.** SNR values of different methods for CFRP specimen 1.

|                 | SNR            |                 |                    |             |
|-----------------|----------------|-----------------|--------------------|-------------|
|                 | Diamond Defect | Circular Defect | Trapezoidal Defect | All Defects |
| Original images | 1.50           | 0.48            | 1.14               | 0.72        |
| PCT [11]        | 1.97           | 2.04            | 3.35               | 1.67        |
| GPCT [13]       | 1.74           | 2.56            | 2.25               | 1.35        |
| KPCT            | 1.85           | 2.22            | 2.33               | 2.00        |
| GKPCT           | 1.91           | 2.61            | 3.64               | 2.25        |

#### 4.2. Performance Analysis of GKPCT Method on Specimen 1

The GKPCT method solves the problem of insufficient model training data and performs well in the evaluation of nonlinear thermographic data defects. Generally, the increase in training data enables the model to better learn and fit data features. However, it is important to ensure that the additional training data contains valid information that can help improve model performance. For the task of IRT defect detection, the accurate identification of defects lies in comparing the local temperature information in the thermogram. In this work, a local pixel standard deviation (LPSD) index is introduced to measure the local smoothness of the thermal image.

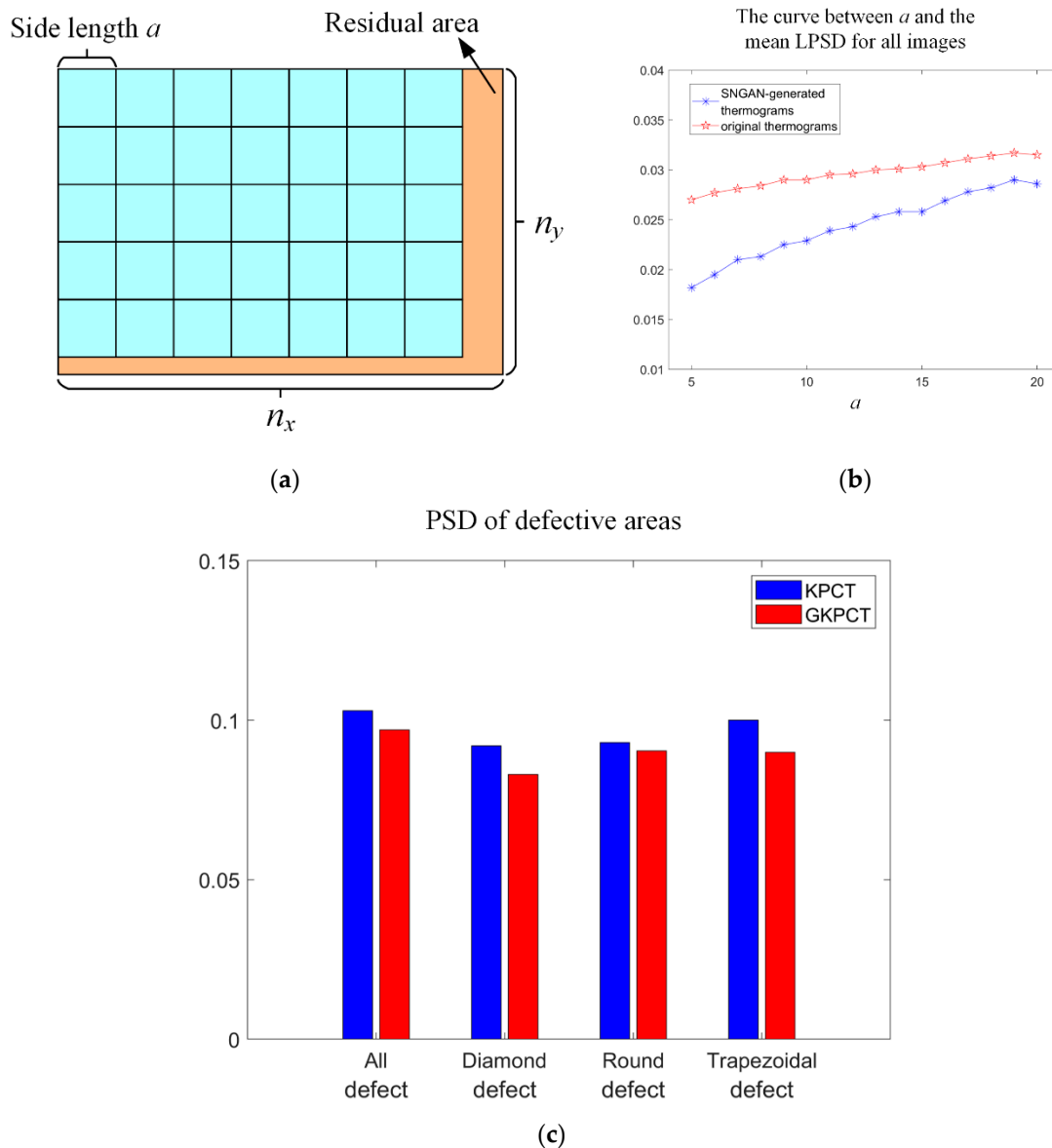
The LPSD design process is exemplified in Figure 13a with the following details. First, the pixel values of the thermal images are Min-Max normalized to (0, 1). Each image is then divided into several square regions with side length  $a$ , and the pixel standard deviation (PSD) values in each region is calculated. Note that the value of  $a$  should not be set too large, otherwise LPSD loses its usefulness as a measure of local quality. Finally, the average value of all PSD is taken as the LPSD value of the image, which is a measure of the smoothness of the image. Thus, the LPSD is calculated as:

$$\text{LPSD} = \frac{\sum_{i=1}^n \sigma_{\text{local}_i}}{n}, \quad (9)$$

where  $n$  is the number of squares blocks with side length  $a$  obtained after partitioning and  $n = \text{floor}(n_y/a)$ , here, floor is a downward integer function that returns the largest integer not larger than the result,  $n_y$  is the length of the thermogram.  $\sigma_{\text{local}_i}$  is the standard deviation of the pixel values contained within the  $i$ -th square. A small LPSD value means that the image is smoother in the local area, or it can be understood that the pixel values of the local area are not discrete, and the concentrated pixel values are reflected in the thermogram as almost the same temperature color. The local difference helps to identify the contour of the defect.

As shown in Figure 13b, the LPSD values of the thermal images (i.e., 90 PT thermal images and 80 SNGAN thermal images) used for the GKPCT modeling were calculated and plotted in the figure, respectively. It can be seen that the LPSD values of the SNGAN-generated thermal images are much smaller when  $a$  is set from 5 to 30, which means that the SNGAN images have a locally higher smoothness compared to the original images. This may be due to the low noise of the generated thermal image caused by the convolution operation included in the SNGAN structure. To further investigate the promotion of SNGAN image augmentation on GKPCT, the PSD values of defective regions (local locations) in the PC4 images of KPCT and GKPCT were calculated separately. The comparison between the two is shown in Figure 13c, where the pixel standard deviation value of GKPCT is lower for both individual defects and the entire defect region. This indicates that the pixels in the defect area are numerically closer and the color of the defect area becomes more visually consistent. As a result, GKPCT has a better defect identification ability than KPCT, and the size and boundaries of defects are clearer in its visualization results.





**Figure 13.** (a) Schematic diagram of area division; (b) the relationship between LPSD and  $a$ ; (c) the PSD of the defective areas in KPCT and GKPCT.

#### 4.3. Thermal Image Data Analysis of Specimen 2

Similar to the analysis procedure of specimen 1, first, the GKPCT method was performed to analyze the PT results of CFRP specimen 2. In detail, SNGAN was used to generate 60 fake thermograms which were combined with the original thermograms. In doing this, the dataset was enlarged, which was then analyzed by implementing KPCT. The results are shown in Figure 14. The processing results of PCT, GPCT, and KPCT are depicted in the same figure for comparison. It is observed that the results of GKPCT have the best visibility of the defects, especially the three defects in the right side which locate deeper than those in the left. The quantitative evaluation results based on the  $SNR$  index are listed in Table 2. GKPCT has the largest  $SNR$  values for defects A2, B2, and C2. In the meantime, its overall performance is also the best. However, concerning defects A1, B1, and C1, the  $SNR$  values associated with GKPCT are not the largest, although these defects can be easily identified from the feature images provided by GKPCT as shown in Figure 14. The reason why the  $SNR$  values are against the intuition is attributed to the fact that  $SNR$  compares the defect area with the global background, whereas actually the



human judgment of the defects only compares the defective regions with the neighboring area. This is a major disadvantage of the SNR index.

The above discussions can be verified by equally dividing the image into left and right parts and calculating the SNR values for each method based on the half image, as illustrated in Figure 15. In other words, the SNR values associated with defects A1, B1, and C1 are calculated based on the left half image alone, while the SNR values of A2, B2, and C2 are calculated based on the right half image. The results are shown in Table 3, from which it is obvious that GKPCT outperforms other methods.

Table 2. SNR values of different methods for CFRP specimen 2.

|                 | SNR         |             |             |             |             |             |             |
|-----------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
|                 | Defect A1   | Defect B1   | Defect C1   | Defect A2   | Defect B2   | Defect C2   | All Defects |
| Original images | 0.83        | 0.93        | 1.09        | 0.84        | 0.90        | 1.06        | 0.17        |
| PCT [11]        | 1.67        | 1.52        | 1.33        | 1.25        | 1.27        | 1.14        | 0.67        |
| GPCT [13]       | 2.01        | 1.82        | <b>1.57</b> | 1.32        | 1.45        | 1.22        | 1.16        |
| KPCT            | <b>2.17</b> | <b>1.86</b> | 1.46        | 1.40        | 1.53        | 1.19        | 1.22        |
| GKPCT           | 1.94        | 1.78        | 1.54        | <b>1.43</b> | <b>1.56</b> | <b>1.21</b> | <b>1.27</b> |

The value in bold indicates the maximum SNR value corresponding to the defect.

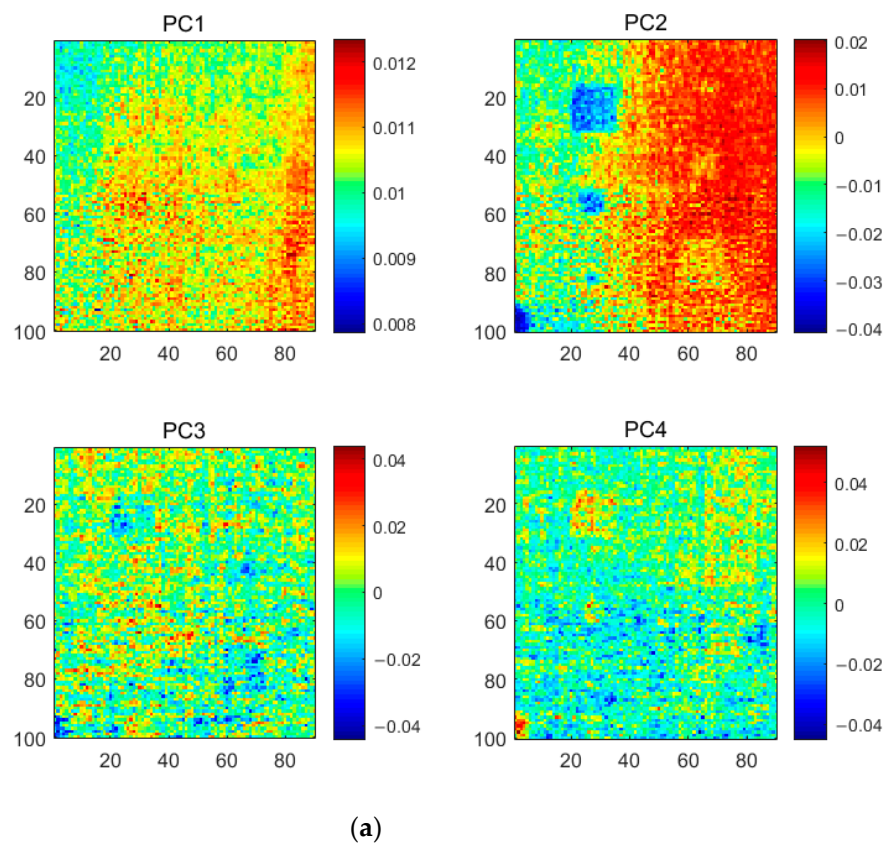
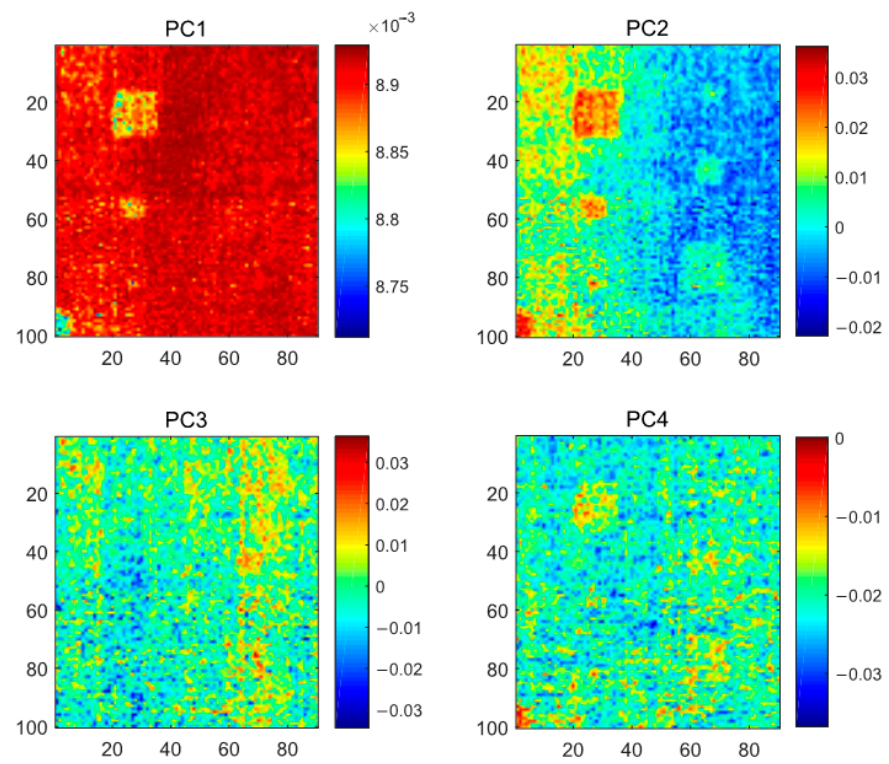
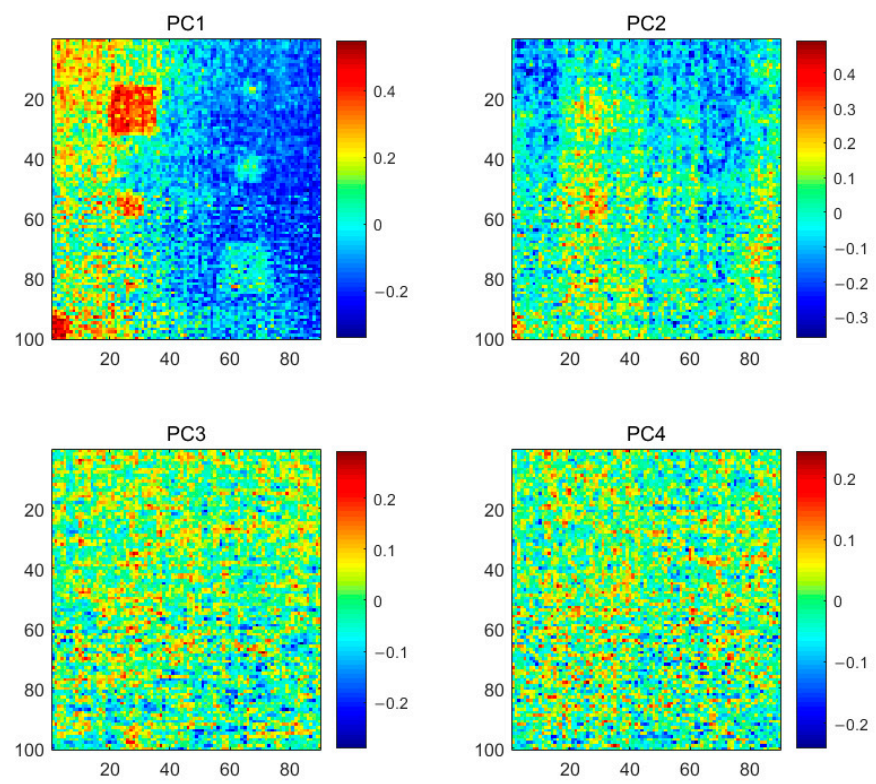


Figure 14. Cont.

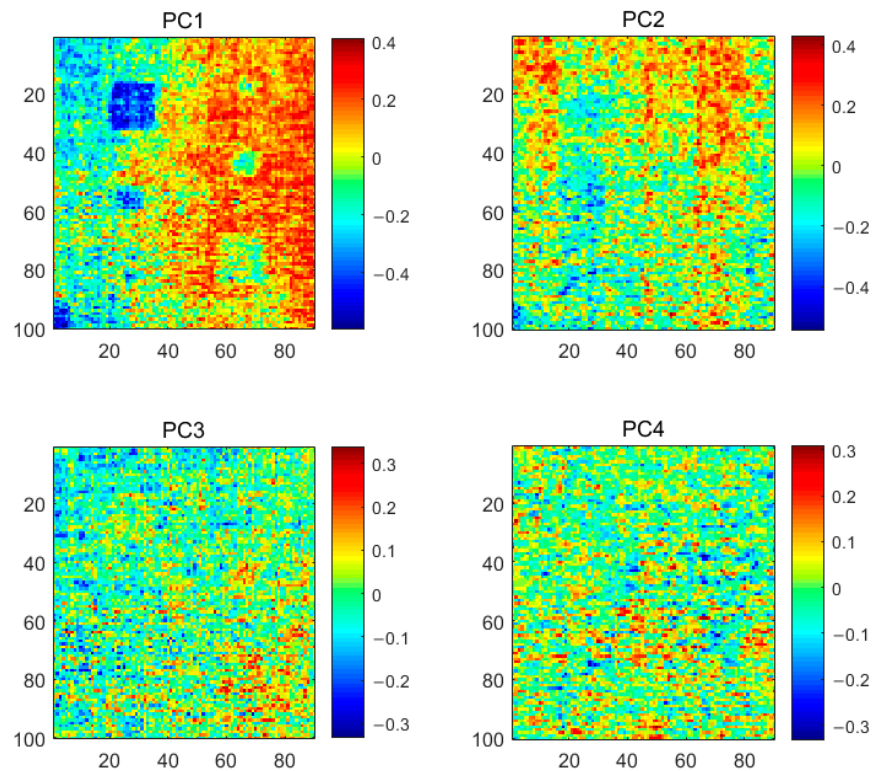


(b)



(c)

Figure 14. Cont.



(d)

Figure 14. Analysis results of CFRP Specimen 2 based on (a) PCT, (b) GPCT, (c) KPCT (with a Gaussian kernel whose width is set to  $1/\tau = 0.2$ ), and (d) GKPCT ( $1/\tau = 0.2$ ).

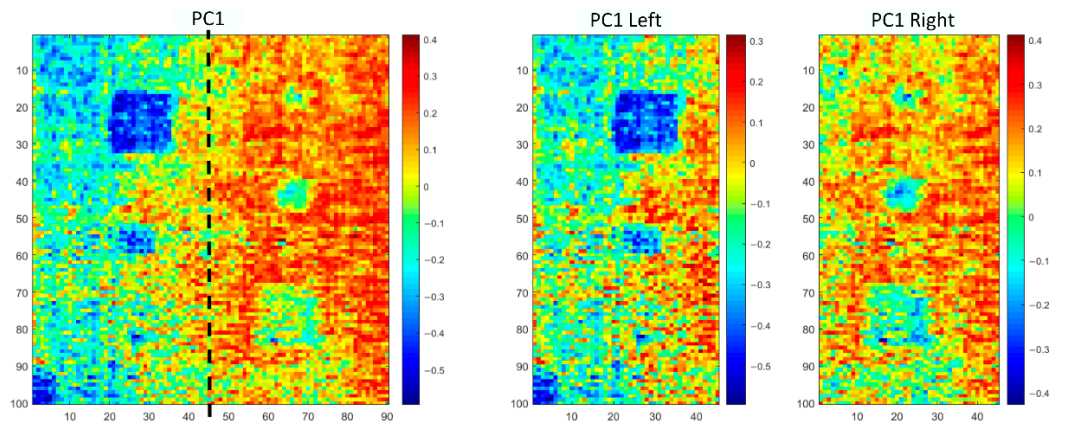


Figure 15. Schematic diagram of image segmentation with the improved SNR calculation.

Table 3. SNR values based on half images.

|           | SNR         |             |             |             |             |             |             |
|-----------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
|           | Defect A1   | Defect B1   | Defect C1   | Defect A2   | Defect B2   | Defect C2   | All Defects |
| PCT [11]  | 1.83        | 1.74        | 1.49        | 1.18        | 1.02        | 0.96        | 0.67        |
| GPCT [13] | 2.08        | 1.92        | 1.68        | 1.24        | 1.29        | 1.07        | 1.16        |
| KPCT      | 2.25        | 1.99        | 1.57        | 1.35        | 1.47        | 1.13        | 1.22        |
| GKPCT     | <b>2.54</b> | <b>2.18</b> | <b>1.71</b> | <b>1.46</b> | <b>1.60</b> | <b>1.28</b> | <b>1.27</b> |

The value in bold indicates the maximum SNR value corresponding to the defect.

In contrast to the SNR index, the LPSD index proposed in this work focuses more on the contrast between the defect region and its surrounding neighbors, which is in line with the human intuition. The LPSD values of different methods are listed in Table 4, where the results show the outer performance of the proposed GKPCT method.

**Table 4.** LPSD values of different methods for CFRP specimen 2.

|          | LPSD          |               |               |               |               |               |               |
|----------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
|          | Defect A1     | Defect B1     | Defect C1     | Defect A2     | Defect B2     | Defect C2     | All Defects   |
| PCT [11] | 0.0112        | 0.0131        | 0.0111        | 0.0122        | 0.0123        | 0.0123        | 0.0114        |
| KPCT     | 0.1390        | 0.1716        | 0.1422        | 0.1011        | 0.0949        | 0.0891        | 0.1772        |
| GKPCT    | <b>0.1561</b> | <b>0.2081</b> | <b>0.1731</b> | <b>0.1279</b> | <b>0.1324</b> | <b>0.1035</b> | <b>0.1973</b> |

The value in bold indicates the maximum LPSD value corresponding to the defect.

## 5. Conclusions

In this work, a novel nonlinear thermographic data analysis method, namely GKPCT, is proposed to improve the detection accuracy of subsurface defects in CFRP. The SNGAN-generated images are mixed with real thermal images for better feature extraction based on nonlinear dimensionality reduction, which results in enhanced defect detection performance. The experimental results of two CFRP experiments show that the combination of image augmentation and nonlinear analysis of thermal images helps to extract features and improve the visibility of defects in the detection results. In the future, the efficacy of the method on other types of composites will be further tested.

**Author Contributions:** Data curation: K.L. and Y.Y.; funding acquisition: Y.L. and Y.Y.; investigation: K.L., Z.M., Y.L., and J.Y.; methodology: K.L., Y.L., and Y.Y.; project administration: Y.L.; writing—original draft: K.L. and Y.L.; writing—review and editing: Y.L. and Y.Y. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Natural Science Foundation of China (grant Nos. 62022073 and 61873241) for their financial support. Yao was supported in part by Minister of Science and Technology, ROC, under project number MOST 108-2221-E-007-068-MY3.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data presented in this study are available on request from the corresponding author.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Ibarra-Castanedo, C.; Tarpani, J.R.; Maldague, X.P.V. Nondestructive testing with thermography. *Eur. J. Phys.* **2013**, *34*, S91–S109. [[CrossRef](#)]
- Amer, S.; Al Zarkani, H.; Sfarra, S.; Omar, M. Infrared thermography approach for pipelines and cylindrical based geometries. *Polym.* **2020**, *12*, 1616. [[CrossRef](#)]
- Sun, S.; Han, Z.; Fu, H.; Jin, H.; Dhupia, J.S.; Wang, Y. Defect characteristics and online detection techniques during manufacturing of FRPs using automated fiber placement: A review. *Polymer* **2020**, *12*, 1337. [[CrossRef](#)]
- Titman, D. Applications of thermography in non-destructive testing of structures. *NDT E Int.* **2001**, *34*, 149–154. [[CrossRef](#)]
- Ibarra-Castanedo, C.; González, D.; Klein, M.; Pilla, M.; Vallerand, S.; Maldague, X. Infrared image processing and data analysis. *Infrared Phys. Technol.* **2004**, *46*, 75–83. [[CrossRef](#)]
- Yan, Z.; Chen, C.-Y.; Luo, L.; Yao, Y. Stable principal component pursuit-based thermographic data analysis for defect detection in polymer composites. *J. Process. Control.* **2017**, *49*, 36–44. [[CrossRef](#)]
- LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)] [[PubMed](#)]
- Xuan, Q.; Chen, Z.; Liu, Y.; Huang, H.; Bao, G.; Zhang, D. Multiview generative adversarial network and its application in pearl classification. *IEEE Trans. Ind. Electron.* **2019**, *66*, 8244–8252. [[CrossRef](#)]



9. Deng, H.; Yang, K.; Liu, Y.; Zhang, S.; Yao, Y. Actively exploring informative data for smart modeling of industrial multiphase flow processes. *IEEE Trans. Ind. Inform.* **2021**, *1*. [[CrossRef](#)]
10. Rajic, N. Principal component thermography for flaw contrast enhancement and flaw depth characterisation in composite structures. *Compos. Struct.* **2002**, *58*, 521–528. [[CrossRef](#)]
11. Wu, J.-Y.; Sfarra, S.; Yao, Y. Sparse principal component thermography for subsurface defect detection in composite products. *IEEE Trans. Ind. Inform.* **2018**, *14*, 5594–5600. [[CrossRef](#)]
12. Yousefi, B.; Sfarra, S.; Castanedo, C.I.; Maldague, X.P. Comparative analysis on thermal non-destructive testing imagery applying candid covariance-free incremental principal component thermography (CCIPCT). *Infrared Phys. Technol.* **2017**, *85*, 163–169. [[CrossRef](#)]
13. Liu, K.; Li, Y.; Yang, J.; Liu, Y.; Yao, Y. Generative principal component thermography for enhanced defect detection and analysis. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 8261–8269. [[CrossRef](#)]
14. Balageas, D.L.; Roche, J.-M.; Leroy, F.-H.; Liu, W.-M.; Gorbach, A.M. The thermographic signal reconstruction method: A powerful tool for the enhancement of transient thermographic images. *Biocybern. Biomed. Eng.* **2015**, *35*, 1–9. [[CrossRef](#)]
15. Zheng, K.; Chang, Y.-S.; Yao, Y. Defect detection in CFRP structures using pulsed thermographic data enhanced by penalized least squares methods. *Compos. Part B Eng.* **2015**, *79*, 351–358. [[CrossRef](#)]
16. Maldague, X.; Galmiche, F.; Ziadi, A. Advances in pulsed phase thermography. *Infrared Phys. Technol.* **2002**, *43*, 175–181. [[CrossRef](#)]
17. Liu, Y.; Wu, J.Y.; Liu, K.; Wen, H.L.; Yao, Y.; Sfarra, S.; Zhao, C. Independent component thermography for non-destructive assessment of defects in polymer composites. *Meas. Sci. Technol.* **2019**, *30*, 044006. [[CrossRef](#)]
18. Zhang, X.; He, Y.; Chady, T.; Tian, G.Y.; Gao, J.; Wang, H.; Chen, S. CFRP impact damage inspection based on manifold learning using ultrasonic induced thermography. *IEEE Trans. Ind. Inform.* **2018**, *15*, 2648–2659. [[CrossRef](#)]
19. Liu, Y.; Liu, K.; Yang, J.; Yao, Y. Spatial-neighborhood manifold learning for nondestructive testing of defects in polymer composites. *IEEE Trans. Ind. Inform.* **2020**, *16*, 4639–4649. [[CrossRef](#)]
20. Liu, Y.; Liu, K.; Gao, Z.; Yao, Y.; Sfarra, S.; Zhang, H.; Maldague, X.P. Non-destructive defect evaluation of polymer composites via thermographic data analysis: A manifold learning method. *Infrared Phys. Technol.* **2019**, *97*, 300–308. [[CrossRef](#)]
21. Wu, H.; Zheng, K.; Sfarra, S.; Liu, Y.; Yao, Y. Multiview learning for subsurface defect detection in composite products: A challenge on thermographic data analysis. *IEEE Trans. Ind. Inform.* **2020**, *16*, 5996–6003. [[CrossRef](#)]
22. Xu, C.; Xie, J.; Wu, C.; Gao, L.; Chen, G.; Song, G. Enhancing the visibility of delamination during pulsed thermography of carbon fiber-reinforced plates using a stacked autoencoder. *Sensors* **2018**, *18*, 2809. [[CrossRef](#)]
23. Ramesh, V. A review on application of deep learning in thermography. *Int. J. Eng. Manage. Res.* **2017**, *7*, 489–493.
24. Manduchi, G.; Marinetti, S.; Bison, P.; Grinzato, E. Application of neural network computing to thermal non-destructive evaluation. *Neural Comput. Appl.* **1997**, *6*, 148–157. [[CrossRef](#)]
25. Bang, H.-T.; Park, S.; Jeon, H. Defect identification in composite materials via thermography and deep learning techniques. *Compos. Struct.* **2020**, *246*, 112405. [[CrossRef](#)]
26. Kulis, B. Metric Learning: A Survey. *Found. Trends Mach. Learn.* **2013**, *5*, 287–364. [[CrossRef](#)]
27. Miyato, T.; Kataoka, T.; Koyama, M.; Yoshida, Y. Spectral normalization for generative adversarial networks. *arXiv* **2018**, arXiv:1802.05957.
28. Kim, K.I.; Jung, K.; Kim, H.J. Face recognition using kernel principal component analysis. *IEEE Signal Process. Lett.* **2002**, *9*, 40–42. [[CrossRef](#)]
29. Creswell, A.; White, T.; Dumoulin, V.; Arulkumaran, K.; Sengupta, B.; Bharath, A.A. Generative adversarial networks: An overview. *IEEE Signal Process. Mag.* **2018**, *35*, 53–65. [[CrossRef](#)]
30. Guei, A.-C.; Akhloufi, M. Deep learning enhancement of infrared face images using generative adversarial networks. *Appl. Opt.* **2018**, *57*, D98–D107. [[CrossRef](#)]
31. Wei, B.-J.; Chang, Y.-S.; Yao, Y.; Fang, J. Online estimation and monitoring of local permeability in resin transfer molding. *Polym. Compos.* **2016**, *37*, 1249–1258. [[CrossRef](#)]
32. Ibarra-Castanedo, C.; Piau, J.-M.; Guilbert, S.; Avdelidis, N.P.; Genest, M.; Bendada, A.; Maldague, X.P.V. Comparative study of active thermography techniques for the nondestructive evaluation of honeycomb structures. *Res. Nondestruct. Eval.* **2009**, *20*, 1–31. [[CrossRef](#)]
33. Meola, C.; Carlomagno, G.M. Recent advances in the use of infrared thermography. *Meas. Sci. Technol.* **2004**, *15*, R27–R58. [[CrossRef](#)]
34. Shrestha, R.; Kim, W. Non-destructive testing and evaluation of materials using active thermography and enhancement of signal to noise ratio through data fusion. *Infrared Phys. Technol.* **2018**, *94*, 78–84. [[CrossRef](#)]
35. Maaten, L.V.D.; Hinton, G. Visualizing data using t-sne. *J. Mach. Learn. Res.* **2008**, *9*, 2579–2605.