Deep learning with filtering for defect characterization in pulsed thermography based non-destructive testing

Sethu Selvi Selvan[1], Sharath Delanthabettu[2], Menaka Murugesan[2], Venkatraman Balasubramaniam[2]
[1]Department of Electronics and Communication, Ramaiah Institute of Technology, Bangalore, India
[2]Safety, Quality and Resource Management Group, Indira Gandhi Center for Atomic Research, Kalpakkam, India

Article Info

ABSTRACT

Pulsed thermography is widely used for non-destructive testing of various materials. The temperature profile obtained after pulse heating is used to characterize the underlying defects in an object. In this paper, the automation of the process of defect visualization and depth quantification in pulsed thermography through various deep learning algorithms is reported. Stainless steel plate with artificial defects is considered for analysis. The raw temperature data is smoothed using moving average, Savitzky-Golay and quadratic regression filters to reduce noise. Thermal signal reconstruction, the conventional method to eliminate noise, is also used for generating filtered datasets. Defect visualization refers to identifying and locating the defects in an image sample and Mask region convolutional neural network (Mask R-CNN) is considered for not just detecting the defects but also locating them on the image. The located defects are utilized for depth estimation using the following networks-multi-layer perceptron (MLP), long short-term memory (LSTM) and gated recurrent units (GRU). The input to the networks is the temperature contrast characteristics which symbolizes the difference in temperature over defective and non-defective areas measured over 250 time points and output of the networks is the estimated depth. The study shows that LSTM based approach provides the least percentage error of 5.5% and is a very suitable approach for automation of defect characterization in pulsed thermography.

Corresponding Author:

Sethu Selvi Selvan
Department of Electronics and Communication, Ramaiah Institute of Technology
Bengaluru, Karnataka-560 054, India
Email: selvi@msrit.edu

1. INTRODUCTION

Non-destructive evaluation (NDE) approaches are employed for detection and characterization of underlying defects in various structural materials. Numerous NDE methods are widely used in industries such as nuclear power plants, and chemical industries. One of the advanced NDE approaches is pulsed thermography (PT), which provides a quick inspection of near surface defects in various materials, through a non-contact mechanism. In this approach, the material under inspection is heated up uniformly using a very short optical pulse. Any near surface defects, such as bonding faults and corrosion are observed by measuring the surface temperature response using a thermographic camera. The defects provide a different temperature signature compared to non-defective areas, which are easily picked up in thermal images. The noise in PT measurements affects the evaluation approaches and it is necessary to improve the signal to noise ratio (SNR) without losing crucial information required for defective visibility. This aids in accurate defect depth estimation.

Most of the reported methods involve complicated processes of noise suppression. The objective of this paper is to propose a defect characterization system with minimal human intervention and is based on deep learning architectures as they are expected to provide higher accuracy for defect characterization. One network is designed for locating the defects and another network for estimating the depth of the defect.

A stainless-steel plate with artificial defects is considered for analysis and raw temperature data is filtered. Mask regional convolutional neural network (Mask R-CNN) is used for defect detection while multi-layer perceptron (MLP), long short-term memory (LSTM), and gated recurrent unit (GRU) networks are used for depth estimation. Depth estimation accuracy is considered as a performance measure for choosing the best deep learning network architecture.

The paper is organized as follows: section 2 elaborates on the theoretical background and the experimental setup for PT. Section 3 presents a literature review on existing deep learning-based algorithms for defect characterization. Section 4 elaborates on the proposed algorithm for defect characterization with filtering to enhance the SNR. Section 5 depicts the performance comparison of the proposed algorithms through simulation experiments. Conclusions and future directions for the proposed work are provided in section 6.

2. THEORETICAL BACKGROUND AND EXPERIMENTAL SETUP
2.1. Pulsed thermography

Pulsed thermography (PT) is meant for quickly inspecting materials for near surface defects and bonding weaknesses. In this approach, the surface of the sample material is bombarded with a short and high energy light pulse. The light energy is absorbed by the surface and the temperature on the material rises instantaneously. The generated thermal waves travel through the material due to which the surface temperature decreases. The decay in temperature with respect to time is portrayed in Figure 1. If there is no defect in the material, the propagation of thermal waves happens without any disturbance as shown in plot (a) of Figure 1. If there is a defect, some of the thermal waves are reflected back to the surface as they get interfered with by the defect. This increases the temperature on the defective area, which results in slower decay of temperature as depicted in plot (b) of Figure 1. The acquired thermal images are stored as a 3D matrix where the pixel positions are indicated through the spatial x-coordinate and y-coordinate and time is represented by the z-coordinate.

2.2. Defect detection

In non-defective areas, the temperature profile is a continuous non-periodic signal that decays as the square root of time. Once the decrease in temperature stabilizes, temperature changes are negligible. The behavior of temperature vs time characteristics is similar for both defective and non-defective areas, as the applied heat has not reached the defect. As the thermal energy reaches the defective area, surface temperature of the defective area increases and is more than the temperature on the non-defective area until stabilization as shown in Figure 1.

2.3. Experimental setup

The setup for pulse thermography is as shown in Figure 2. A flashlight excites the inspected surface resulting in instantaneous rise in surface temperature. An infrared camera records the drop in surface temperature after the flash. A CEDIP Silver 420 infrared camera is employed to record the thermal images. The camera consists of an indium antimonide (InSb) focal plane array detector with a resolution of 320×250 pixels together with a Stirling based cooling system. The maximum achievable temperature resolution is 25 mK and the maximum achievable frame rate is 176 Hz. The camera collects and detects infrared radiations in the wavelength range of 3-5 μm. Two Xenon flash lamps with a power of about 1,600 W each and a flash duration of less than 2 ms are considered for the experiments. The experiment is carried out in reflection mode. The distance between the camera and the object is fixed at 35 cm and lamp to object distance is 30 cm [1].

2.4. Material

Two samples of AISI grade 316L stainless steel (SS) plate of dimension 150×100×3.54 mm drilled with square defects from the backside are considered for the experiment. Sample 1 has square defects of depths 0.4, 1.13, 1.78, 2.48, 3.17, and 3.36 mm, from front surface and sample 2 has square defects of depths 0.5, 1.15, 1.8, 2.55, 3.2, and 3.36 mm. The sizes of the defects considered are 10×10, 8×8, 6×6, 4×4, 2×2 mm which are machined using electrical discharge machining (EDM) process. The schematic diagram of two samples is as shown in Figure 3. All the dimensions are in mm.
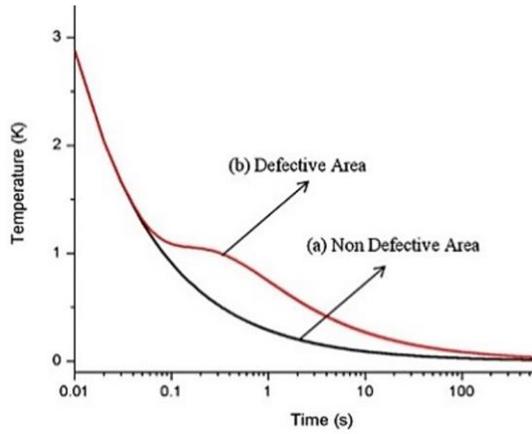
Figure 1. Temperature vs time characteristics for non-defective and defective areas [1]
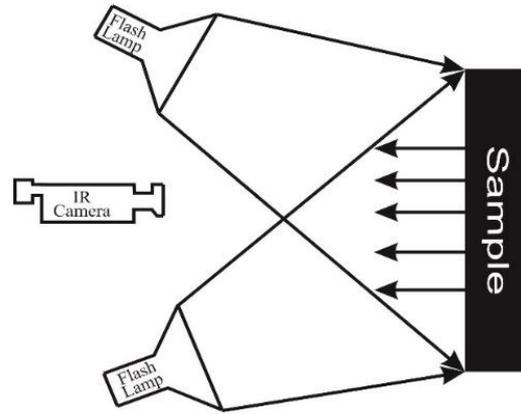

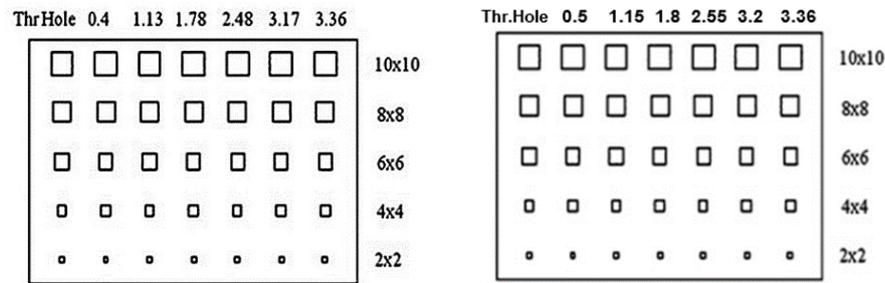Figure 2. Schematic diagram of PT experimental setup

Figure 3. Schematic diagram of sample 1 and sample 2

The emission capacity of SS is small as it has a lower emissivity of 0.7. To increase the emissivity capacity of the material, a uniform and thin coat of black paint is applied to the surface. As thermal diffusivity of 316L SS metallic material is comparatively higher, thermal images are acquired with a frame rate of 125 Hz for 2 s. Raw data of standard 316L SS material is obtained and pre-processed through background subtraction to remove the non-uniformity from the thermal images.

3. DEEP LEARNING ALGORITHMS FOR DEFECT CHARACTERIZATION-EXISTING METHODOLOGIES

The rapid development of deep learning architectures for defect characterization enables PT to be more intelligent and highly automated, thus considerably increasing its usage in number of applications. The work in [1] elaborates on the potential of pulsed thermography for defect depth estimation and establishes the efficiency of three different PT approaches namely temperature contrast method, contrast derivative method and logarithmic second derivative method for defect depth estimation and the performance is compared with finite difference modelling. The contrast derivative and logarithmic second derivative methods provide better accuracy in determining the depth and logarithmic second derivative displays least error for depth prediction.

A review on the principle, cameras, thermal imaging based machine vision and deep learning architectures is presented in [2]. Artificial intelligence (AI) is considered in combination with infrared thermography to detect and localize various defects on laminates [3]. Segmentation is performed on mid-wave and long-wave infrared measurements obtained during PT experiments through a deep neural network for each wavelength. The F1-score for mid-wave images-based model is 92.74%, while for long-wave images is 87.39%.

An artificial neural network (NN) is employed to detect depth of the defects in composite samples, coupled with PT [4]. A multi-physics finite element method (FEM) based simulation model of the inspection process is considered for generating a training dataset and the proposed NN validated the accuracy and performance experimentally. The accuracy of the NN obtained for synthetic data is more than 97% and 90% for experimental data. A neural network based approach for defect characterization through thermal images is

presented in [5]. Better performance is achieved by using a smoother activation function for the output units and this algorithm proved extremely satisfactory, when compared to conventional procedures for thermal image analysis. A novel statistical processing algorithm is proposed [6] and performance comparison is provided for Perceptron and Kohonen based architectures. An interference-based technique is explored which is a two-step approach comprising of detection with phase and characterization with amplitude. Applications based on thermographic non-destructive evaluation (TNDE) are limited due to the complex non-linearity effect in defect depth estimation. A three dimensional thermal model for carbon fiber reinforced plastic (CFRP) is proposed in [7] with back propagation neural networks (NN) as defect detector and depth estimator. A new algorithm based on a multilayer NN post-processor is proposed [8] to predict defect depths in real-time with an accuracy of 97.5% for synthetic data with defect depth up to 1 mm and a worst-case accuracy of 90% for a 0.5 mm depth. A neural network based infrared thermography is proposed in [9] to classify defects, such as air, oil, and water, which degrades performance of materials. Raw data and thermographic signal reconstruction coefficients are employed to train, and test the two multilayers, feed forward NN models, which proved to be more precise and better test repeatability.

In [10] Mask region based convolutional neural network (Mask R-CNN) is considered for image segmentation. The thermal images are obtained through experiments and synthetic images through FEM in COMSOL software. The proposed network yields better performance when the dataset is combined with synthetic images for training. 25 flat bottom holes (FBH) in circle or square shapes are inserted at various depths on a CFRP composite and plexiglass samples of size 30×30 cm. In this work, the confidence score is defined as 0.75, which is the acceptance criterion based on the percentage of the ground truth boundaries overlapping with the predicted defective area boundaries.

Paper [11] evaluates different models based on deep learning such as Fast R-CNN, Faster R-CNN, RetinaNet, and YOLOv3 for two standard datasets: a small object dataset, a filtered dataset from PASCAL VOC 2007. The inference from this work is that two-stage methods like Faster R-CNN perform better and YOLO is best if training time is not important, to provide a trade-off between accuracy and speed. A novel unsupervised auto-encoder (AE) based deep learning approach is proposed in [12] for increasing the SNR and improve the visualization of the defects. The SNR and defect detection performance is compared with higher order statistics (HOS), principal component thermography (PCT) and partial least square regression (PLSR) thermography and proves that AE approach provides better SNR at the cost of defect detectability.

The authors of [13] consider gated recurrent units (GRUs) in a depth quantifying algorithm for composite material samples via PT. The dataset consists of CFRP specimens with flat bottom holes simulated through FEM modeling with precisely controlled depth and geometrics of the defects. The complicated and non-linear issues involved in estimating defect depths in composite materials via infrared thermography with a GRU learning model are also addressed. The accuracy obtained for depth estimation is 90% before data normalization and increased to 95% after normalization.

Saeed et al. [14] provide a thorough analysis of employing neural networks as a post-processor for quantifying defect depths from thermograms of CFRP. The hyper parameters of the network and thermogram specific factors such as frame rate, and number of considered points per defect location, are explored to prove the robustness of the network which is more sensitive to network architecture, noise, training hyperparameters and the type, amount and diversity of training data. Deep feed forward (DFF) network with rectified linear unit (ReLU) activation function for the hidden layer estimates the depth of defects in CFRP structures through a regression model. The network consists of 5 hidden layers and the succeeding layer has two-thirds the number of nodes of the preceding layer.

Thermal signal reconstruction (TSR) is known to be computationally complicated despite being considered as the most accurate method. The method allows materials to be inspected for both near surface and subsurface defects quickly. This is a very fast algorithm and gives nearly accurate results for defect depth estimation. Object detection based algorithms like Mask R-CNN [10] and Faster R-CNN [11] are considered for defect detection. However, Mask R-CNN is chosen for defect visualization as its output includes the location of the detected defects, which is necessary for automation. The depth of the defect is calculated using neural networks like perceptron [14] and GRU [14] with different inputs and manually calculated using contrast derivative or log second derivative methods [15].

Two deep learning approaches are considered to recover temperature profiles from PT images [16]. A deep neural network (DNN) is trained with surface temperature measurements as input of the first approach and in the other approach, surface temperature measurements are converted to virtual waves. For 1,00,000 simulated temperature measurement images both approaches outperformed the baseline algorithms in terms of reconstruction accuracy. The virtual waves are features which yield better reconstructions with the same number of training samples and provides more compact network architectures.

A depth quantifying method is proposed [17] using GRUs for composite materials through PT. The model quantifies the depth of defects automatically and is evaluated for accuracy and performance of synthetic CFRP data from FEM. In [18], a spatial and temporal hybrid deep learning architecture is proposed which significantly minimizes uneven illumination and enhances the detection rate. Visual geometry group-UNet (VGG-UNet) provides a better contrast between defective and non-defective areas. Recent advancements in the field of AI can support post-processing of thermographic data, efficiently, especially for non-linear or complex thermography scanning routines. An autonomous post-processor capable of automatically detecting defects from given thermograms via a convolutional neural network (CNN) is proposed in [19] with a deep feed forward neural networks (DFF-NN) algorithm to estimate the defect depth through transfer learning. Paper [20] presents an overview of object detection models based on R-CNN, YOLO and SSD employed for defect detection. This contributes to an increase in productivity and detection of industrial cables in several hours.

4. DEEP LEARNING ALGORITHMS WITH FILTERING FOR DEFECT CHARACTERIZATION

The proposed deep learning based defect characterization algorithm with various filtering approaches is depicted in Figure 4. The algorithm employs filtered datasets together with the original noisy dataset and compares the performance for defect visualization and depth estimation. Moving average (MA), savitzky-golay, quadratic regression and thermal signal reconstruction filters are considered for filtering the dataset. The deep learning architectures are based on R-CNN, MLP, LSTM, and GRU.
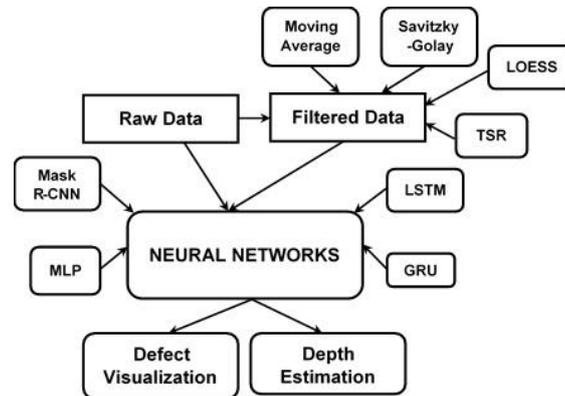

Figure 4. Block diagram of proposed methodology

4.1. Datasets

Two separate datasets are considered for training and testing the neural networks. Dataset 1 consists of 250 raw thermal images from sample 1, while dataset 2 consists of 250 images from sample 2. Dataset 1 is used for training while dataset 2 is used for testing. Eight filtered datasets are generated and considered for defect visualization and depth estimation. This 3-dimensional dataset is directly used for defect depth estimation.

4.2. Deep learning models for defect detection

Pulsed thermography (PT) is an approach that utilizes external energy heating. This approach is a non-linear process due to non-uniform heating, heat diffusion, cooling process, environmental noise, and limited parameters of the thermal imaging acquisition system. Neural networks (NN) are suitable for learning and modeling complex non-linear functions, with an added advantage of generalizability. R-CNN based networks are considered for defect visualization. As the temperature profile (time vs temperature) is considered for depth estimation, recursive neural networks based architecture meant for modelling time series data is considered. The distinct cooling process associated with each defect in pulsed thermography makes this a non-linear process.

Mask R-CNN [11] which is based on R-CNN is considered for defect characterization. This algorithm consists of two-stages of Faster R-CNN architecture, in which input images are scanned in the first stage, and a region proposal network (RPN) generates the region of proposals for the next stage. In the first branch of the second stage, classification scores of the category for a region of proposal are predicted to distinguish whether or not it is a background. In the second branch, a binary mask is added to each region of

proposal as a mini-semantic segmentation network to classify each pixel in the region proposal as whether or not it is an object. A certain number of bounding boxes is predicted by the model for the defects, and each defect area within the bounding box is segmented at the same time. The architecture is as shown in Figure 5.
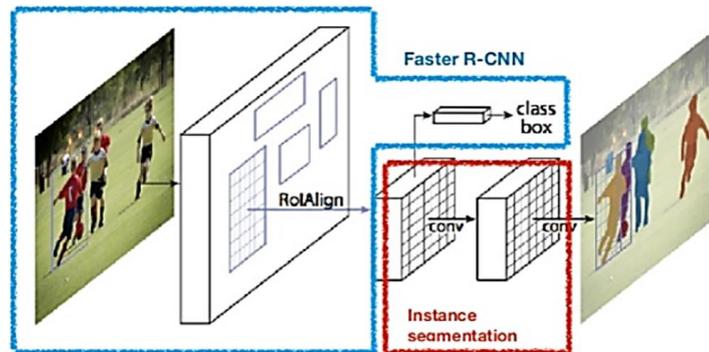

Figure 5. Mask R-CNN Architecture [15]

The backbone of the network is Resnet101 with a learning momentum of 0.9 and learning rate of 0.001. All layers are trained for 10 epochs. The dataset consists of 60 images each of size 226×320 pixels, 50 images are used for training and 10 for testing. As the size of the available dataset is very small, the model weights obtained for COCO image dataset is considered for better accuracy. An auto encoder is also considered which provides a well-trained model to obtain thermal images with better representation of features from 250 input images. The encoder part of the auto encoder is considered for dimensionality reduction, and for creating one image from 250 images in the dataset which improves the visibility of the defects.

4.3. Defect depth estimation

In defective areas of raw data, temperature-time characteristics are generated, and thermal contrast curves are considered for defect depth estimation. The temperature curve of a defective area is subtracted from that of a non-defective area as in (1), where ΔT is the thermal contrast, T_{nd} and T_d are the temperatures at time t for non-defective and defective areas respectively.

$$\Delta T = T_d - T_{nd} \tag{1}$$

The mean temperature of a particular region is considered to obtain T_{nd} and T_d. The output is the estimated defect depth through regression. Twelve defects are chosen from dataset 1 for training and four defects from dataset 2 are chosen for testing. The input for defect depth estimation is the output of Mask R-CNN network with defect detection. This is used for locating the defects in the 3-dimensional dataset of size 226×320×250. The locations are considered across all the 250 images in the dataset. The quality of defect depth estimation [1] is measured in terms of the first derivative of the temperature contrast and peak time. The peak time t_s is the time at which contrast derivative peaks and the defect depth x is calculated as in (2) where α is the thermal diffusivity of the material.

$$x = \sqrt{\frac{\alpha t_s \pi^2}{3.64}} \tag{2}$$

The following are the networks considered for defect depth estimation.
a. Multilayer perceptron (MLP): The network consists of an input layer, hidden layers, and an output layer. The input layer has 250 neurons corresponding to the number of data points of thermal contrast curves. The hidden layers have two-thirds the number of neurons of the preceding layer. The output layer predicts the estimated depth. ReLU activation function and adaptive moment (ADAM) optimizer is employed, which adapts learning rates based on the average magnitudes of the gradients, resulting in fast convergence. The network is trained for 30 epochs, with a training time of 25 s and testing time of 3 ms.
b. Long short-term memory (LSTM) networks: These are recurrent neural networks (RNN) suitable for sequence prediction problems by learning order dependence. These networks consist of memory blocks

that are connected through layers, instead of neurons. The gates of a block predict the state of the block and output for an input sequence. Each gate employs an activation function for deciding whether it is triggered or not, which results in change of state and addition of information flowing through the block. The various gates within a unit are: Forget gate decides conditionally which information to retain and which to throw away from the block. Input gate decides conditionally which input values to retain for updating the memory state and output gate decides conditionally the output based on input and memory of the block. As LSTMs are sensitive to the scale of input data, the input is normalized to [0, 1] and also reshaped to a 3-dimensional array. Three hidden layers are considered, and the network is trained for 10 epochs, with a training time of 52 s and testing time of 1 s. The architecture is as shown in Figure 6.

c. Gated recurrent unit (GRU): This is the younger sibling of LSTM, and also a type of RNN. GRU exhibits superior speed with similar accuracy and effectiveness of LSTM. The GRU architecture is as shown in Figure 7. Similar to RNN, sequential input data is fed to the GRU cell at each time step together with the memory, or hidden state. GRU has one hidden state transferred between time steps and the hidden state is then re-fed into the cell together with the next input data in the sequence. The GRU cell contains two gates: reset gate which decides on which of the previously hidden states to be combined with current input for estimating a new hidden state. Update gate decides on which of the previous hidden state is to be retained and what portion of the new hidden state is to be added to the final hidden state. The gradient of the error obtained during training is employed to update the weights by the required magnitude in the required direction. The training time is 100 s and testing time is 6 s and about 240-270 neurons are selected for training. 15-20 epochs are considered based on the optimization needs with ReLU activation function. The architecture is as shown in Figure 7.
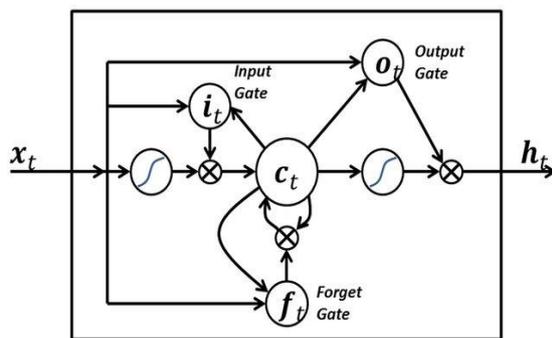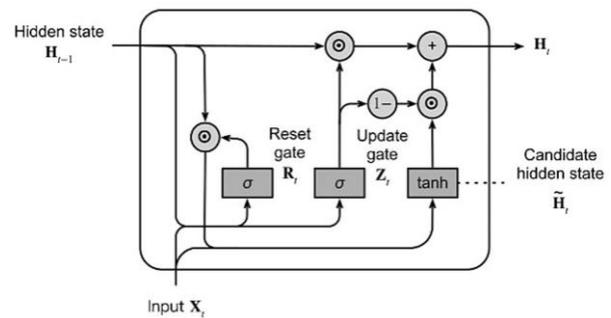

Figure 6. LSTM architecture [21]

Figure 7. GRU Architecture [22]

4.4. Filtering approaches

Filtering algorithms are considered to reduce noise and increase SNR across both temperature and time. Simple filters are preferable to reduce computational complexity. The 3D data acquired through PT consists of a tray of 250 images which is converted into a one-dimensional signal. The following one-dimensional filters are considered to reduce the spatial noise.

a. Moving average (MA) filter: This simple low pass finite impulse response (FIR) filter smooths data samples [23]. As a result, short-term fluctuations are smoothened, and long-term trends or cycles are highlighted.

b. Savitzky-Golay filter: This filter smooths noisy data samples occupying a large frequency span. As a result, less of the high-frequencies are filtered out compared to averaging FIR filters [24]. However, these filters are not suitable in conditions where noise levels are not particularly high.

c. Quadratic regression filter: Quadratic regression determines a parabola that best fits the data samples. Locally estimated scatter plot smoothing (LOESS) is a non-parametric approach which employs a quadratic function [25]. These fits multiple regressions in a local neighborhood and polynomial equation in (3) is used to calculate the weights, where y_k is the dependent variable and x_k is the independent variable.

$$y_k = a + bx_k + cx_k^2 \tag{3}$$

d. Thermal signal reconstruction (TSR): This approach is suitable for data reduction and noise filtering to improve the SNR. This is based on Fourier's 1-dimensional heat-transfer equation on a semi-infinite

surface subject to thermal excitation expressed as a Dirac delta function. TSR [1] models the temperature evolution of a sample through a nth degree polynomial and fitting the polynomial to a logarithmic sequence as in (4) where T is the increment in temperature as a function of time, $a_0, a_1, \dots, a_n$ are the polynomial coefficients and $T_{surf}(t)$ is the surface temperature at time t.

$$ln[T_{surf}(t)] = a_0 + a_1 \ln(t) + a_2[\ln(t)]^2 + \cdots + a_n[\ln(t)]^n \qquad (4)$$

5. PERFORMANCE COMPARISON

5.1. Dataset preparation

Masks are obtained for each image as in Figure 8 for all the different classes and the images are selected randomly from the validation dataset. The defects are labelled using visual geometry group (VGG) image annotator (VIA) using polygon shapes. The original image is displayed in Figure 8(a) and its corresponding mask is shown in Figure 8(b) with only one class defect.
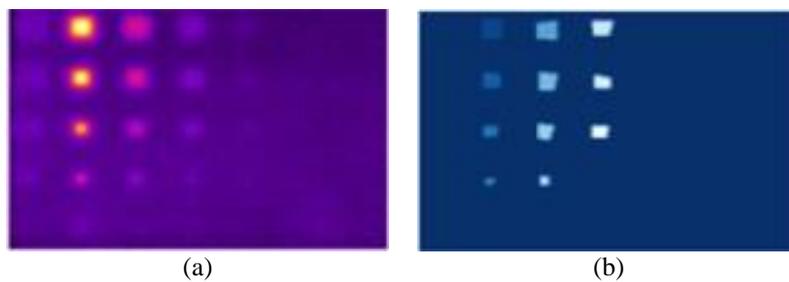

(a) (b)

Figure 8. Mask extracted for a sample dataset image (a) original image and (b) mask

5.2. Mask R-CNN for defect detection

The defect detection capability of Mask R-CNN is depicted in Figure 9. A sample original image is depicted in Figure 9(a) and the proposed algorithm provides a very high confidence score of 0.99 for both the training and test datasets as shown in Figure 9(b). The coordinates obtained as output of Mask R-CNN are scaled down after removing the padding. These coordinate values are verified for the original image of size 226×320 by manually drawing rectangular boxes and checking if the defects are identified correctly. From Figure 9(c), it is observed that the defects are successfully identified, and these coordinates are the input to depth estimation network.
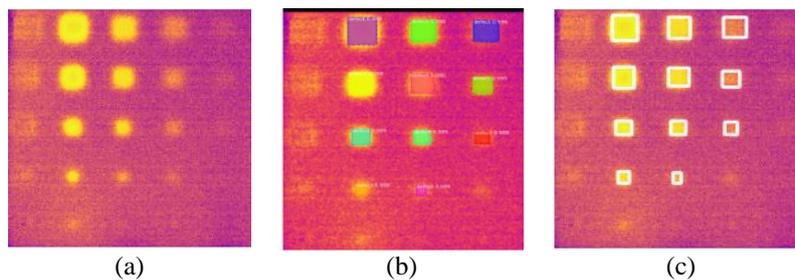

(a) (b) (c)

Figure 9. Defect detection based on mask R-CNN (a) original image, (b) output image,
and (c) defect location

5.3. Filtered dataset

The data is filtered through various filters to smooth the temperature vs time characteristics at each pixel location. The performance of various filters is compared in terms of quality and SNR. The output provided by Savitzky-Golay filter is presented. Figure 10 compares the filtered output with the noisy temperature profile measured at a particular pixel location across 250 images as specified in Figure 10(a). The blue line in Figure 10(b) represents the noisy temperature profile, while the red line represents the

filtered output of Savitzky-Golay filter. Figure 10(c) presents the filtered temperature contrast characteristics and proves that this filter yields the best filtered output suitable for detecting localization. Figure 11 presents the reconstructed images at different time instants. The filtered and reconstructed images at time instants of 0.1, 0.5, 1, and 1.5 s are depicted in Figures 11(a) to 11(d) respectively. It is observed that with increasing time deeper defects are visualized clearly.
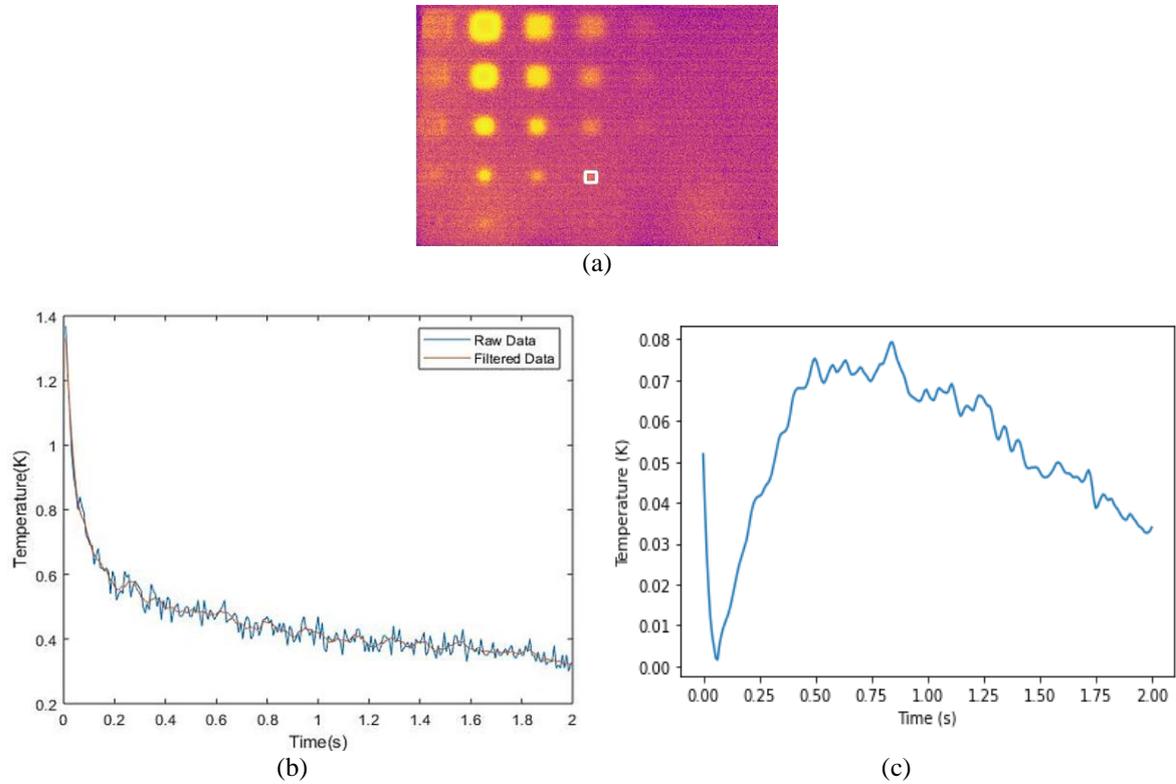

(a)

(b)

(c)

Figure 10. Dataset image filtered with Savitzky-Golay filter (a) defect location, (b) original noisy signal and filtered output, and (c) filtered temperature contrast
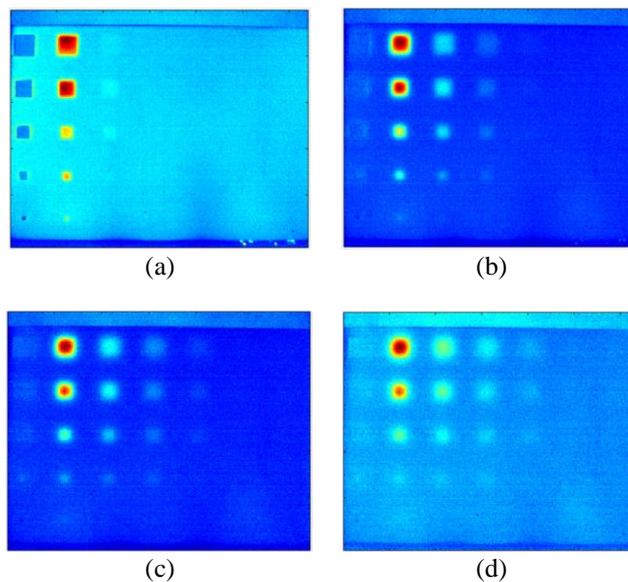

(a)

(b)

(c)

(d)

Figure 11. Reconstructed images at increasing time instants (a) 0.1 s, (b) 0.5 s, (c) 1 s, and (d) 1.5 s

The measured SNR is depicted in Figure 12 for 10×10 mm defective area of depth 1.13 mm with respect to time. The SNR increases for all the filters (blue line) when compared to that of the original raw data (red line) and has similar performance for the other 5 defects also. The SNR is compared for each defect in Table 1. All the filters show a significant improvement in SNR when compared to original raw data.
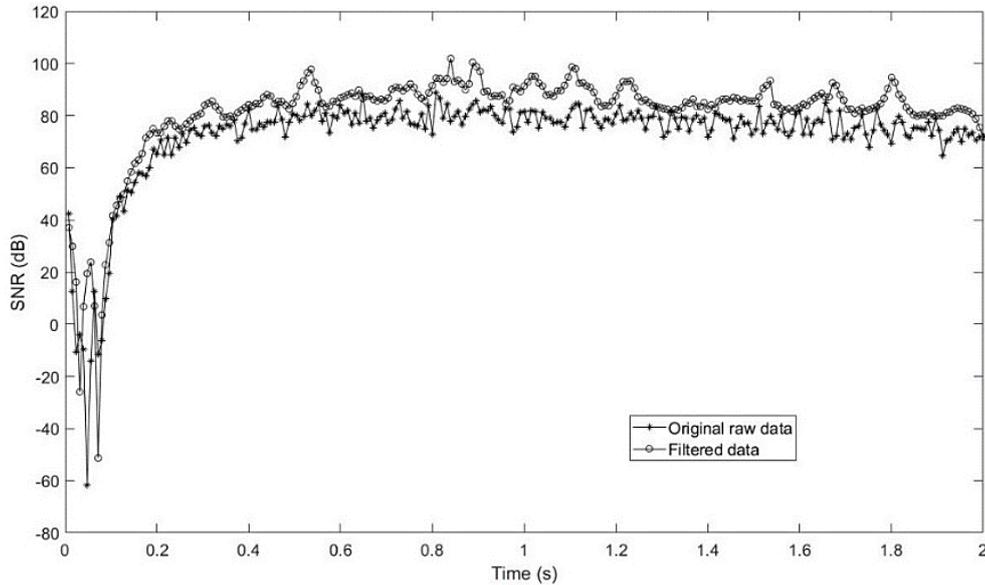

Figure 12. SNR comparison for Savitzky-Golay filter

Table 1. SNR for different defect sizes and depths

Defect Size (mm)	Defect Depth (mm)	Raw data (dB)	Savitzky-Golay filter (dB)
10×10	0.40	117.5	**128.0**
	1.13	88.8	**101.9**
	1.78	71.8	**85.5**
6×6	0.40	102.3	**110.2**
	1.13	83.4	**95.2**
	1.78	65.4	**79.5**

The quality of defect visibility is depicted in Figure 13 for a particular input raw data sample in Figure 13(a), where gray square indicates 'through hole', green square indicates 'clearly detected' defect, while blue square is 'barely detected' and red square is 'undetected' defect. The first 3 squares in the 5th column from the left are barely detected in the original raw image and they are visible in the Savitzky-Golay filtered output depicted in Figure 13(b). The contrast derivative plot for 10×10 mm defective area with a depth of 1.13 mm is provided in Figure 14. The first derivative of the temperature contrast between defective and non-defective areas is considered and the time at which the peak is obtained is noted as the peak time (t_s), which is 0.136 s in the example shown in Figure 14.
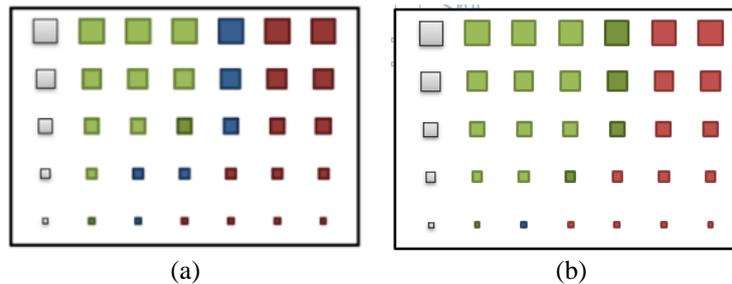

Figure 13. Sample defect visibility (a) raw data and (b) Savitzky-Golay filtered data

5.4. Depth estimation

The defect locations are identified in the original raw dataset and temperature curves are obtained for all 12 defects of 3 different depths 0.4, 1.13 and 1.78 mm. Temperature contrast curve is the input and the estimated depth is the output of the networks. These curves indicate temperature in y-axis with respect to time in x-axis. For LSTM and GRU the inputs are normalized. The proposed MLP consists of 4 hidden layers while LSTM and GRU have 5 hidden layers. The performance is compared based on percentage error which indicates the error between predicted and expected depth and root mean square error (RMSE).
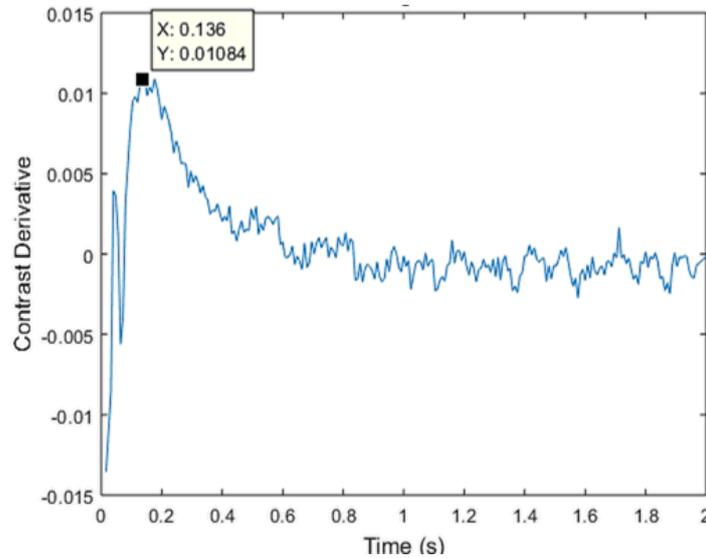

Figure 14. Peak time of temperature contrast derivative

5.5. Quality comparison

Table 2 provides the comparison of percentage error obtained through MLP, LSTM and GRU deep learning architectures for raw dataset. Two different datasets with dimensions 226×320×250 are used for training and testing. 11 contrast curves from sample 1 are used for training and 4 contrast curves from sample 2 are used for testing. It is observed that LSTM network has the least error percentage of 0.7% and a highest error percentage of 3.8% only, while the MLP network has the highest error percentage of 11.8% and GRU has 7.2%.

Table 2. Comparison of percentage error

Algorithms	Expected Depth (mm)	Predicted Depth (mm)	Error
MLP	0.50	0.38	11.8%
LSTM		0.54	3.8%
GRU		0.55	3.8%
MLP	1.15	1.22	7.0%
LSTM		1.14	0.7%
GRU		1.21	4.6%
MLP	1.80	1.84	4.1%
LSTM		1.82	1.6%
GRU		1.88	6.5%

From Table 3, it is noted that LSTM network provided the least RMSE for Savitzky-Golay filtered dataset, followed by quadratic regression based filtered dataset. The largest RMSE value is for the moving mean filtered dataset. But in terms of error percentage, the least error is 3.9% for Savitzky-Golay filtered dataset followed by quadratic regression and TSR dataset with an error percentage of 5.5% for both. The efficiency of employing filtered datasets is proved by comparing the estimated depth with that of depth obtained through manual calculations as depicted in Table 4.

Manual calculation of defect depths for different filtered datasets are obtained through (2). The depth of the defect is estimated using the time at which the temperature contrast peak is obtained. For raw

dataset the manual calculation was not done, as the noise in raw signal is enhanced by the derivative operator making it impossible to locate the peak time from the curve. The highest percentage error values obtained for each dataset for different depths are tabulated in Table 4.

From Table 4, it is clear that there is a significant reduction of error in deep learning approaches over manual calculations. Overall, the error measured in deep learning networks is always lesser than manual calculation. MLP network performed second best overall in terms of accuracy. LSTM network showed the highest reduction in error compared to manual calculation. Thus, deep learning approaches have significantly improved depth estimation accuracy. From the table it is also observed that filtering always does not help in reducing the error in depth estimation. For example, in the case of LSTM, the moving average filter slightly increased the error in depth estimation when compared to raw data. The reduction in error is not very significant for filtered data when compared to raw data. For example, in the case of LSTM, the error reduced from 5.9% for raw data to 5.5% for quadratic regression filtered data.

Table 3. Comparison of error and RMSE for LSTM network

Dataset	Percentage error			
	MLP	LSTM	GRU	Manual calculation
Raw data	11.8%	5.9%	7.2%	-
Moving average	7.1%	8.6%	9.0%	15%
Savitzky-Golay	7.6%	3.9%	13%	23%
Quadratic Regression	8.2%	5.5%	6.1%	23%
TSR	5.8%	5.5%	5.7%	25%

Table 4. Comparison of percentage error with manual calculations for different filters

Dataset	Expected Depth (mm)	Predicted Depth (mm)	Error	RMSE
Raw data	0.50	0.53	3.8%	0.0364
	1.15	1.14	0.7%	
	1.79	1.81	1.6%	
Moving Average Filter	0.50	0.51	1.4%	0.0655
	1.15	1.21	6.6%	
	1.79	1.87	7.1%	
Savitzky-Golay Filter	0.50	0.50	0.09%	0.0264
	1.15	1.18	3.5%	
	1.79	1.83	3.9%	
Quadratic Regression	0.50	0.52	2.2%	0.0383
	1.15	1.18	3.3%	
	1.79	1.76	3.4%	
TSR	0.50	0.55	5.5%	0.0413
	1.15	1.18	3.4%	
	1.79	1.84	4.1%	

6. CONCLUSION

In this paper deep learning algorithms have been successfully adopted for localizing the defects and estimating their depth. The raw data was filtered using various filtering approaches. Both raw data and filtered data were used for depth quantification using various deep learning network architectures. The accuracy of depth estimation for filtered datasets using deep learning networks is improved in comparison with manually calculated depth measurements. LSTM network gives the least error of 5.9% for the raw dataset. For the filtered datasets, considering the RMSE values and highest error percentages, it is concluded that the Savitzky-Golay filtered dataset gives the best performance with least percentage error for LSTM network.

Comparing the different networks across different filtered datasets, MLP performs the best for the moving average filter with least error of 7.1% and LSTM performs the best for Savitzky-Golay filter with least error of 3.9%. LSTM also performs the best for quadratic regression filter and thermal signal reconstruction (TSR) with error of 5.5% for both which is the least among the networks. Hence, on average it is concluded that LSTM is the best performing neural network across different filtered datasets and raw dataset as well. The depth estimation accuracy does not improve significantly for the filtered data when compared to raw data. The proposed methodology is for depth estimation as part of defect characterization and can be extended to estimate defect size. General adversarial network (GAN) can be considered to augment the dataset and increase its size, which would aid in obtaining better performance.

REFERENCES

[1]   D. Sharath, M. Menaka, and B. Venkatraman, "Defect characterization using pulsed thermography," *Journal of Nondestructive Evaluation*, vol. 32, no. 2, pp. 134–141, Jun. 2013, doi: 10.1007/s10921-012-0166-4.

[2]   Y. He *et al.*, "Infrared machine vision and infrared thermography with deep learning: A review," *Infrared Physics and Technology*, vol. 116, Aug. 2021, doi: 10.1016/j.infrared.2021.103754.

[3]   Z. Wei, H. Fernandes, H.-G. Herrmann, J. R. Tarpani, and A. Osman, "A deep learning method for the impact damage segmentation of curve-shaped CFRP specimens inspected by infrared thermography," *Sensors*, vol. 21, no. 2, Jan. 2021, doi: 10.3390/s21020395.

[4]   N. Saeed, Y. Abdulrahman, S. Amer, and M. A. Omar, "Experimentally validated defect depth estimation using artificial neural network in pulsed thermography," *Infrared Physics and Technology*, vol. 98, pp. 192–200, May 2019, doi: 10.1016/j.infrared.2019.03.014.

[5]   G. Manduchi, S. Marinetti, P. Bison, and E. Grinzato, "Application of neural network computing to thermal non-destructive evaluation," *Neural Computing and Applications*, vol. 6, no. 3, pp. 148–157, Sep. 1997, doi: 10.1007/BF01413826.

[6]   S. Vallerand and X. Maldague, "Defect characterization in pulsed thermography: a statistical method compared with Kohonen and Perceptron neural networks," *NDT and E International*, vol. 33, no. 5, pp. 307–315, Jul. 2000, doi: 10.1016/S0963-8695(99)00056-0.

[7]   A. Darabi and X. Maldague, "Neural network based defect detection and depth estimation in TNDE," *NDT and E International*, vol. 35, no. 3, pp. 165–175, Apr. 2002, doi: 10.1016/S0963-8695(01)00041-X.

[8]   N. Saeed, M. A. Omar, and Y. Abdulrahman, "A neural network approach for quantifying defects depth, for nondestructive testing thermograms," *Infrared Physics and Technology*, vol. 94, pp. 55–64, Nov. 2018, doi: 10.1016/j.infrared.2018.08.022.

[9]   Y. Duan *et al.*, "Automated defect classification in infrared thermography based on a neural network," *NDT and E International*, vol. 107, Oct. 2019, doi: 10.1016/j.ndteint.2019.102147.

[10]   Q. Fang, C. Ibarra-Castanedo, and X. Maldague, "Automatic defects segmentation and identification by deep learning algorithm with pulsed thermography: synthetic and experimental data," *Big Data and Cognitive Computing*, vol. 5, no. 1, Feb. 2021, doi: 10.3390/bdcc5010009.

[11]   N.-D. Nguyen, T. Do, T. D. Ngo, and D.-D. Le, "An evaluation of deep learning methods for small object detection," *Journal of Electrical and Computer Engineering*, vol. 2020, pp. 1–18, Apr. 2020, doi: 10.1155/2020/3189691.

[12]   S. S. Selvan *et al.*, "A deep learning approach based defect visualization in pulsed thermography," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 11, no. 3, pp. 949–960, Sep. 2022, doi: 10.11591/ijai.v11.i3.pp949-960.

[13]   Q. Fang and X. Maldague, "A method of defect depth estimation for simulated infrared thermography data with deep learning," *Applied Sciences*, vol. 10, no. 19, Sep. 2020, doi: 10.3390/app10196819.

[14]   N. Saeed, H. Al Zarkani, and M. A. Omar, "Sensitivity and robustness of neural networks for defect-depth estimation in CFRP composites," *Journal of Nondestructive Evaluation*, vol. 38, no. 3, Sep. 2019, doi: 10.1007/s10921-019-0607-4.

[15]   K. He, G. Gkioxari, P. Dollar, and R. Girshick, "Mask R-CNN," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 2, pp. 386–397, Feb. 2020, doi: 10.1109/TPAMI.2018.2844175.

[16]   P. Kovács, B. Lehner, G. Thummerer, G. Mayr, P. Burgholzer, and M. Huemer, "Deep learning approaches for thermographic imaging," *Journal of Applied Physics*, vol. 128, no. 15, Oct. 2020, doi: 10.1063/5.0020404.

[17]   Q. Fang, F. A. Mamoudan, and and Xavier. Maldague, "Defect depth estimation in infrared thermography with deep learning," *Preprints*, Aug. 2020, doi: 10.20944/preprints202008.0585.v1.

[18]   Q. Luo, B. Gao, W. L. Woo, and Y. Yang, "Temporal and spatial deep learning network for infrared thermal defect detection," *NDT and E International*, vol. 108, Dec. 2019, doi: 10.1016/j.ndteint.2019.102164.

[19]   N. Saeed, N. King, Z. Said, and M. A. Omar, "Automatic defects detection in CFRP thermograms, using convolutional neural networks and transfer learning," *Infrared Physics and Technology*, vol. 102, Nov. 2019, doi: 10.1016/j.infrared.2019.103048.

[20]   A.-A. Tulbure, A.-A. Tulbure, and E.-H. Dulf, "A review on modern defect detection models using DCNNs-Deep convolutional neural networks," *Journal of Advanced Research*, vol. 35, pp. 33–48, Jan. 2022, doi: 10.1016/j.jare.2021.03.015.

[21]   M. Akram and C. El, "Sequence to sequence weather forecasting with long short-term memory recurrent neural networks," *International Journal of Computer Applications*, vol. 143, no. 11, pp. 7–11, Jun. 2016, doi: 10.5120/ijca2016910497.

[22]   A. Zhang, Z. C. Lipton, M. Li, and A. J. Smola, "Dive into deep learning," *arXiv preprint arXiv:2106.11342*, Jun. 2021.

[23]   I. Svetunkov and F. Petropoulos, "Old dog, new tricks: a modelling view of simple moving averages," *International Journal of Production Research*, vol. 56, no. 18, pp. 6034–6047, Sep. 2018, doi: 10.1080/00207543.2017.1380326.

[24]   W. Dai, I. Selesnick, J.-R. Rizzo, J. Rucker, and T. Hudson, "A nonlinear generalization of the Savitzky-Golay filter and the quantitative analysis of saccades," *Journal of Vision*, vol. 17, no. 9, Aug. 2017, doi: 10.1167/17.9.10.

[25]   W. Sharabati and B. Xi, "A neighborhood regression approach for removing multiple types of noises," *EURASIP Journal on Image and Video Processing*, vol. 2018, no. 1, Dec. 2018, doi: 10.1186/s13640-018-0259-9.

BIOGRAPHIES OF AUTHORS

Sethu Selvi Selvan 🆔 Professor Department of ECE, Ramaiah Institute of Technology obtained her Ph.D. from Indian Institute of Science in 2001 under Prof. Anamitra Makur in the area of image compression. She completed her B. E from Thiagarajar College of Engineering, Madurai in 1992 and M. E from Anna University in 1994. She joined the Faculty of Department of Electronics and Communication at M. S. Ramaiah Institute of Technology, Bangalore, in 2002 as assistant professor. She has numerous publications to her name in the field of machine learning, pattern recognition and signal and image processing. Her fields of interests are digital image processing, machine/deep learning, video processing, character recognition and biometrics. She has authored a chapter titled "Image Algebra and Image Fusion" in the book "Data Fusion Mathematics: Theory and Practice", CRC Press, 2017 and has been listed as a noteworthy technical contributor by Marquis Who's Who (World), 2009. She can be contacted at email: selvi@msrit.edu.

Sharath Delanthabettu 🆔 📇 sc ↻ is working as a Scientific Officer at Indira Gandhi Centre for Atomic Research, Kalpakkam. He has obtained his Ph.D. from Homi Bhabha National Institute, IGCAR Kalpakkam Campus in 2015. He is working in the areas of InfraredImaging and its applications in non-destructive evaluation and health care, and image processing. He has published 20 articles in international and national journals. He can be contacted at email: sharathd@igcar.gov.in.

Menaka Murugesan 🆔 📇 sc ↻ is a postgraduate in physics and has over 19 years of experience in the field of NDE for materials characterization. She has specialized in the areas of thermal imaging, image processing and digital radiography. She is presently heading the radiation application and metrology section at Radiological Safety Division of IGCAR, Kalpakkam. Her field of interests are material characterization using thermal NDE and thermal imaging as diagnostic tool in healthcare. She is an American Society for NDT certified Level-III in Infrared Thermal Testing. She can be contacted at email: menaka@igcar.gov.in.

Venkatraman Balasubramaniam 🆔 📇 sc ↻ is distinguished scientist and director of IGCAR, Kalpakkam and CMD of BHAVINI, Kalpakkam. He post-graduated in physics from St. Joseph College (Autonomous), Tiruchirappalli and obtained his Ph.D. from Madras University. With a research career spanning 37 years, he has combined the physics of non-destructive evaluation (NDE) with engineering and technology and consistently provided excellent R and D support and robust NDE-based solutions to technologically challenging problems in nuclear and other strategic and core industries. His significant milestone activities for the nuclear industry include procedures for X-ray and neutron radiography of highly irradiated fuel pins, comprehensive NDE for evaluation of tube-to-tube sheet welds of PFBR steam generator and radiometric testing of shielding structures. He has been primarily responsible for establishing the conventional and digital X-ray, neutron radiography and thermal imaging facilities at IGCAR. He is a recipient of various prestigious awards and fellowships. He can be contacted at email: bvenkat@igcar.gov.in.