International Journal of Thermal Engineering (IJTE)
Volume 12, Issue 2, July– December 2024, pp. 1–20, Article ID: IJTE_12_02_001
Available online at https://iaeme.com/Home/issue/IJTE?Volume=12&Issue=2
ISSN: 2390-4299, Journal ID: 8053-4025
Impact Factor (2024): 10.15 (Based on Google Scholar Citation)

A REVIEW OF THERMAL IMAGING BASED INTERNAL CRACK DETECTION USING DEEP LEARNING (AI)

Prashant Jadhav, Sandeep Thorat, Sachin Pawar, B.K. Patle

ABSTRACT

Thermographic Non-destructive Testing (TNDT) has gained increasing importance in various industry fields. It can provide rapid, non-contact, and robust non-invasive detection of both surface and internal damage. Artificial Intelligence (AI) is an emerging technology that shows increasing potential in almost all fields and has recently attracted significant interest in TNDT. Thermal signals from TNDT have relatively low signal-noise-ratio (SNR), and most thermal images have the common weakness of edge blurring. The abovementioned obstacles lead to high requirements of field expertise and subjectivity in TNDT inspections. One of the purposes of developing AI is substituting human work more efficiently and objectively. The above mentioned weaknesses in TNDT can be overcome with help of AI technologies deep learning. This paper offers a review of state-of-art researches on AI deployment in TNDT, discussing the current challenges and a roadmap for application expansion. Deep Learning is the most commonly used AI technology since it has powerful feature extraction and pattern recognition capabilities for imaging processing and computer vision. Most existing research adopted Convolutional Neural Network (CNN) models utilizing only spatial information in thermal images to detect defects such as U-net, VGG, Yolo, etc. Except for defect detection, automated defect depth estimation is another focus in the deep learning method. Recurrent Neural Networks (RNNs) such as LSTM and GRUs are usually applied for extracting the temporal feature from thermal sequences, which is sensitive to defect depth. Furtherly, different deep model variations and integrated algorithms are also reviewed, which improves the performance of defect detectability. Method followed in way as preparing dataset, building the model, training the model and testing the model.

Keywords: Thermal Non-destructive Testing (TNDT); Internal Crack Detection, Subsurface Defects Detection; Deep Learning (DL); Convolutional Neural Network (CNN); Recurrent Neural Network (RNN); Image Processing

Cite this Article: Prashant Jadhav, Sandeep Thorat, Sachin Pawar, B.K. Patle, A Review of Thermal Imaging Based Internal Crack Detection Using Deep Learning (AI), International Journal of Thermal Engineering (IJTE),12(2), 2024, pp. 1–20.
https://iaeme.com/MasterAdmin/Journal_uploads/IJTE/VOLUME_12_ISSUE_2/IJTE_12_02_001.pdf

INTRODUCTION

As the main storage device for dangerous chemicals, pressure vessel tanks will directly affect the safety of transportation chemicals and hazardous gases. Due to the defects of the inner wall steel plate during the manufacturing process, the stress concentration and the long-term fatigue load, cracks will occur in the inner wall of the tank. If the crack is not found in time, it will become a huge safety hazard, so the crack of the steel plate of the tank car detection is especially important.

Non-destructive testing (NDT) techniques that can inspect specimens without causing any damages to the object being tested. Different types of NDT techniques have been developed during the past few decades to meet the broad requirements in various fields of modern industry. The most typical methods are visual and optical testing, electromagnetic testing, radiographic testing, ultrasonic testing, and thermography testing

The most popular method in thermography testing (also named Thermal Nondestructive Testing, TNDT) is Infrared Thermography which adopts infrared rays to conduct the inspection. Compared with other techniques such as X-rays, ultrasonic, and eddy current, infrared thermography can be applied to much broader materials due to its excellent performance for metallic and non-metallic materials, especially for widely used composite materials such as Carbon-Fiber-Reinforced Polymer(CFRP) materials.

However, TNDT process requires lot of expertise to read thermal images associated with the manual entry of various parameters to perform the inspection. Though the inspection procedure has been adapted to specific applications, the overall process flow of parameter selection, experimental setup, data acquisition, data post-processing and analysis with the final set of results has not changed over the years as they are directly dependent on expert knowledge.

The primary objective of this study is to present suitable deep-learning frameworks that can aid in the automatic detection of defects through pulsed thermography for thermal non-destructive evaluations. These frameworks aim to accurately and efficiently extract and separate different types of defects, including less visible cracks, internal defects, and delamination structures, even when the data are limited.

The key contributions of this review is as follows: 1. A comprehensive and systematic investigation and comparison of classical deep learning methods were conducted to analyze the accuracy and efficiency of defect detection using pulsed thermography. 2. An innovative instance segmentation method was introduced to predict the irregular shape of each defect instance in thermal images, enabling efficient defect segmentation and identification for each defect type across different specimens. 3. Experimental modelinananalysis

In recent years, with the rapid development of Artificial Intelligence (AI), some researchers deployed Deep learning with TNDT and achieved a highly automated inspection process powered by AI.

In this paper, technology of applying to TNDT are reviewed. It has been observed that most studies focused on utilising DL-based approaches to realise automated defects detection, segmentation, or classification, while some of them set foot in defect depth measurement. Different types of deep learning models have been developed for this application studies. Typical models include Convolutional Neural Network (CNN) 8 based models such as VGG network, Unet, SegNet, FCN8, Yolo-v3 and Faster R-CNN, and Recurrent Neural Network (RNN) 9 based models like Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRUs). CNN-based models are primarily used for defects detection, segmentation, and classification in thermal images since their excellent performance in dimensionality reduction and image segmentation.

Temporal information ofthermal image sequences (cooling curves) is alwaysextracted and processed by RNN-based models due to their strong ability in processing time-series information. Most studies only applied one of these deep learning models, whereas few combined two ormore methods. Section 2 to will present threegroups of DL methods: spatial-oriented methods,temporal-based methods, and integrated methodsprocessing both spatial and temporal information.Different deep learning models reviewed in thispaper and a brief classification and interrelationbased on the application scope are shown in Fig.2.

1. Crack Detection Method

When thermally excited from one side of the object, the steady state conditions of the object are destroyed. At this time, heat is diffused in the horizontal direction and the vertical direction. Due to heat transfer, when the heat diffused in the horizontal direction encounters a crack, heat is concentrated on one side of the crack, and the temperature difference between the two sides of the crack increases. The detection principle is shown in figure 1.

Thermal excitation Surface temperature field distribution Receiving energy Horizontal heat transfer Figure 1. Horizontal heat conduction diagram

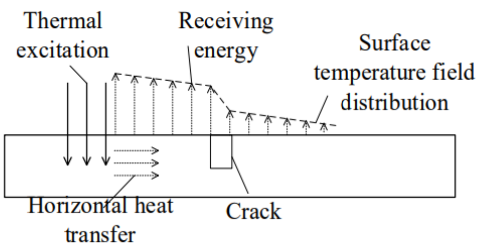


Fig.1 Horizontal heat conduction diagram

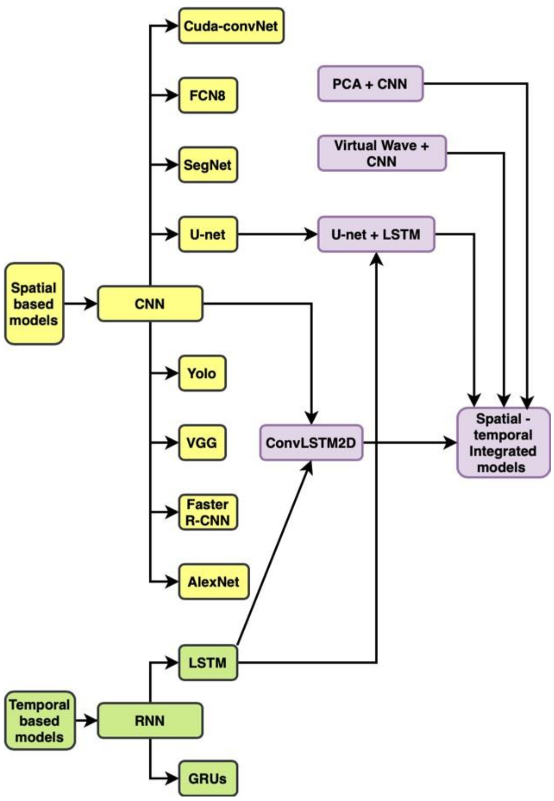


Fig. 2 Spatial-Temporal based classification of different deeplearning models

2. Spatial Deep Learning Methods

Most deep learning methods utilised to detect or classify thermal image defects are Convolutional Neural Network (CNN) based. CNN has a superior ability in feature extraction and pattern recognition which makes it perform well in thermal images segmentation which has been challenging using traditional methods due to thermal data's low Signal Noise Ratio (SNR)

Yolo detected most defects compared with the other combinations. A Yolo-v3 deep learning algorithm was brought forward by Q. Fang et al to realise automatic subsurface defects detection in four different materials, Plexiglass, Carbon Fiber Reinforced Polymer (CFRP), Glass Fiber-Reinforced Polymer (GFRP) and steel. Optical pulsed thermography was adopted in this study for obtaining thermal image sequences. . The training accuracy of the four different materials with this supervised learning is 0.9867 for Plexiglass, 0.992 for CFRP, 0.991 for GFRP and 0.9935 for steel, respectively. Another noteworthy point from the result is that the Probability of Detection (POD) with each bounding box reached an average of 0.99 for different materials with this Yolo-v3 model.

Another research of applying deep learning techniques to identify subsurface defects in composite materials was conducted by H. T. Banga *et al.* . Faster R-CNN was adopted in this research due to its fast processing speed compared with Fast R-CNN, R-CNN, and typical CNN structures. The majority of Faster R-CNN is similar to Fast R-CNN but with one significant difference in identifying proposed regions. Faster R-CNN utilises a separate network to propose candidate regions from original feature maps while Fast R-CNN resorts to selective search algorithm. The model was trained with a public dataset and was validated by two types of manufactured composite specimens (composed of polypropylene and glass fibre with different proportions) with various artificial defects. The proposed system was capable of identifying defects as well as measuring their size and shape, even for defects with irregular shapes.

3. Temporal-Based Deep Learning Methods

The temperature evolution information of the tested specimen can be extracted from time-series information of thermal image sequences. The essential information implied in temperature evolution curves is the thermal contrast (temperature contrast) between defective and sound areas. Thermal contrast can indicate whether a pixel of the specimen belongs to a defective area or sound area. Furthermore, the defect depth has a direct relationship to the time when the cooling curve has a deviation from the one of sound area.

LSTM is an updated method based on RNN. David Müller, Udo Netzelmann and Bernd Valeske proposed a LSTM-RNN model to estimate defects depth in thermal NDT. In this study, raw thermal data and data pre-processed by Thermal Signal Reconstruction (TSR) were chosen as two comparative input sources to test their performance in the LSTM-RNN model. The result showed that TSR data outperformed raw data in depth measurement in this LSTM model. The proposed LSTM-RNN model was also compared with RNN, CNN, and LSTM, respectively, and the result suggested its better ability in learning features.

4. Methods Based on The Combination of Time Spatial Domain

Since both spatial and temporal data of thermal sequences can provide helpful information in defect identification, shown in Fig. 2, some studies also proposed integrated deep learning models to learn both special and temporal information. A neural network was adopted to classify specific defects such as oil, water and air in a homogeneous material in the study, FBHs were produced in a stainless-steel sample to simulate impurities like air, oil and water

A multilayer NN with a feed-forward pattern was applied in this study, the network transformed the 3D information of thermal imagesequences (2D for spatial information and 1D for temporal information) into 2D data by vectorisation.Then the vectorised data was feed into the proposedneural network for defect classification. Raw cooling data of the sample and TSR coefficients were chosen as the two types of input. Two differentNN models were constructed and trained, respectively. The research result showed that both NNs displayed good performance in certain types ofdefects detection with a recall rate of more than 97%. The NN using TSR coefficients outperformed the counterpart model using raw data in defect classification with an average recall rate of 96% and89%

In another study conducted by Q. Luo *et al.*, Pulsed Thermography was chosen as the inspectionmethod due to its fast detection. Five image segmentation techniques, including UNet, VGG- UNet, SegNet, VGG-SegNet and FCN8 wereapplied to detect subsurface defects in the spatial domain in this study. The comparison of their detection results and performance are elaborated. Considering all aspects together, the VGG-UNet model showed the best performance in both high- power and low-power OPT platforms, with an average POD value of 0.667. In terms of time domain, a 3-layer Long-Short-Term-Memory (LSTM) model was adopted to make predictions of transient properties. This model worked well in heat-collecting CFPR materials and coating materials, whereas it performed poorly with heat-dissipating CFPR materials and R shaped samples. Another technique utilised in this study is cross-network learning, which means integrating two different deep learning models. The VGG-U-net model proposed in this research is a good illustration of thisstrategy. U-net cannot learn adequate features from a small dataset in relatively complex cases. But withthe help of some pre-trained weights implanted in the encoder stage which generated from a VGG-16 module, this drawback can be tackled. Apart from VGG-UNet, a PCA-VGG-UNet conjoint model wasalso proposed as a cross-network learning method. In this model, Principal Components Algorithm (PCA) was utilised to extract principal components from thermal sequences . The extracted temporal-spatial features (3-dimensional information) were then converted to a 2-dimensional image which canbe processed by spatial-oriented models such asVGG-UNet. This spatial-temporal conjoint model can be used to detect relatively deeper defects sincethe information obtained from the time domain can help the spatial model achieve a higher detection rate. An integrated deep model ConvLSTM2D whichcombined the CNN method and LSTM to realise defect reconstruction, was proposed in the study.The model utilised CNN's strength in processing spatial information for defect detection and LSTM'sspeciality in predicting defect depth by analysing time-series information of thermal sequences.

5. Preparation of Training dataset

Unlike the normal photographic images, which are always composed of RGB colours including three channels: red, green and blue, thermal images only have one channel. To be more specific, a thermal image can be regarded as a kind of intensityimage representing temperature of each pixel. Another difference between photographic imagesand thermal images is the tonal range. Most RGB images are 8bit, which means there are 256 (2^8) tonalvalues for each colour channel, while thermal images are usually 16-bit which exponentially increases the amount of possible tonal values to 65536 (2^{16}); this also indicates a higher sensitivity to minor changes in temperature. The most meaningfuldifference between thermal images in NDT and photographic images appears when employed in deep learning. Thermal images always come in the form of image sequences which can reflect the temperature propagation and decay along time. Temperature changes over time contain crucial information for subsurface defects identification anddefects depth measurement. Therefore, the form of multi-input and single-output for deep learning model is used to detect subsurface damages in TNDT.

One challenge in applying deep learning models to TNDT is obtaining enough training data from real specimens with different defects. AI models need to be fully trained to achieve good evaluation performance on the testing data, which usually requires a large dataset containing sufficient variants. Data collected from experiments on specimens with artificial defects or real defects are intrinsically limited due to high cost and time consumption. There are three main approaches to tackle this challenge in current studies: 1. Real data from tested samples and data augmentation. 2. Data from simulation. 3. Data from both real samples and simulation or public database. In the first scenario, real thermal data collected from inspection of specimens with subsurface defects are used for training model and data augmentation such as scaling, rotation and flipping are adopted to expand the dataset. The drawback of this method is the lack of variaton of defects in real specimens. The second method uses data generated from simulation, such as FEM. In this method, no real data and real samples are involved. The main disadvantage of this method that the simulated data cannot be identical to real thermal data, which means the model performance could be compromised in real cases. The third approach is also called "transfer learning". The model is pre-trained using simulated data and then fed real data from thermal inspection for the final

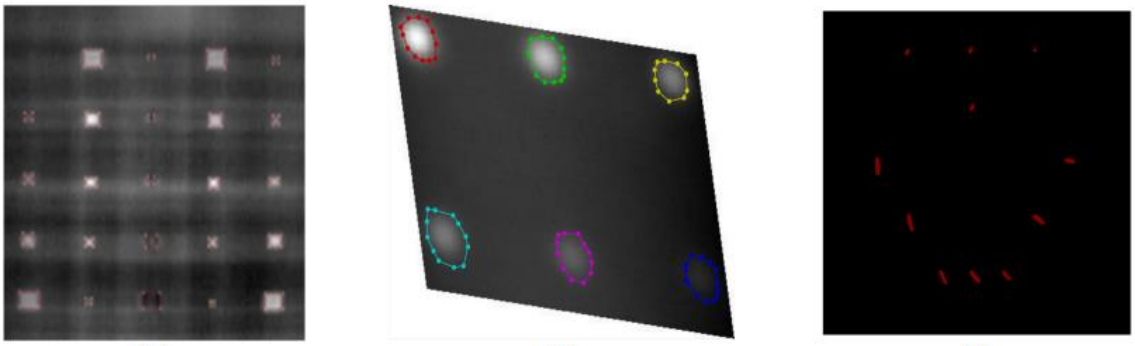
This method may achieve satisfactory performance for specific cases. However, there is still a concern that the deep learning model trained with the dataset containing artificial data and low correlation data (generic data) may not be robust and reliable enough. To maximize the probability of detection, we independently sampled 4000 thermal images in total from the pulsed thermography experiment in three types of materials (plexiglas, carbon fiber-reinforced polymer (CFRP), and steel) to build a training and testing database from pulsed thermography data. As the images used for training should be the same size, the database was split into 512 × 640 pixels.

5.1. Calibration of the Data

Calibration of the Data The marking process was conducted with the two labelling software based on the model type: Colabeler toolkit (YOLO–V3; Faster–RCNN); LabelMe 2.5 toolkit (Mask– RCNN; Center–Mask; U-net; Res–U-net). Each representative image file from the four types of samples was extracted from the sfmov.format sequence files or matrix raw files. These samples created multiple shapes of defects in the database, such as squares and rectangles.In the Colabeler toolkit, only one label (square-shape label) was used for all of the different kinds of marks. The bounding boxes were prepared by hand for each of the images, then exported to a .xml file by Colabeler. Each bounded defect was used as training for the algorithm. The process has to be repeated for all images used for training. In the Labelme toolkit, a different labeling curve from the procedure will be provided regardless of the shape of the defects for segmentation, a labeling curve on each object in the images is then exported to a json.file by Labelme to transform into a large scale object segmentation database (COCO). The elaborate labeling procedure has been explicitly depicted in Figure 4a–c, providing a comprehensive representation of the precise steps involved in the processing of the data.

5.2. Preprocessing and Data Augmentation

In the case of the overfitting issue during the training, data augmentation plays a significant role. We encourage this model to learn the invariant and transformations by using rotation and flipping for the raw images. Since the defects in these materials remain in permanent positions and shapes, they lead to a requirement of capturing images in diverse conditions. As known, the defect is not clear because of the shaping process and/or the specifications of materials that lead to captured images on cluttered background.



(a) (b) (c)

Figure 4. Processing of labelling. (a) bounding box labelling; (b) circle labelling; (c) irregular labelling. images before entering them into a deep-learning network, which is important. Partial images for the training are undertaken in a preprocessing stage. We adapted the preprocessed sequence images from feature extraction methods, including Principal Component Thermography (PCT), which extracts meaningful features by dimension reduction and reflects the intuitions of the data. For example, when the data arise from the high dimensional form (sparse and unstable estimation), the PCT can give more redundancy to our classier to enable them to make a better decision

6. Methodologies

Defect Detection Methods by Deep Learning Algorithms As shown in Figure 5 below, three main deep-learning feature-extraction methods and their implementation steps were introduced: A. Objective localization algorithms: Method 1. Single-stage real-time algorithm-You Only Look Once (YOLO-V3), and Method 2. Two-stage real-time algorithm—Faster Region-based Convolutional Neural Networks—Faster–RCNN; B. Semantic segmentations: Method 3. U-net, and Method 4. Res–U-net; C. Instance segmentation: Method 5. Mask–RCNN and Method 6. Center– Mask; D. Regular thermal segmentation: Method 7. The absolute thermal contrast with global threshold.

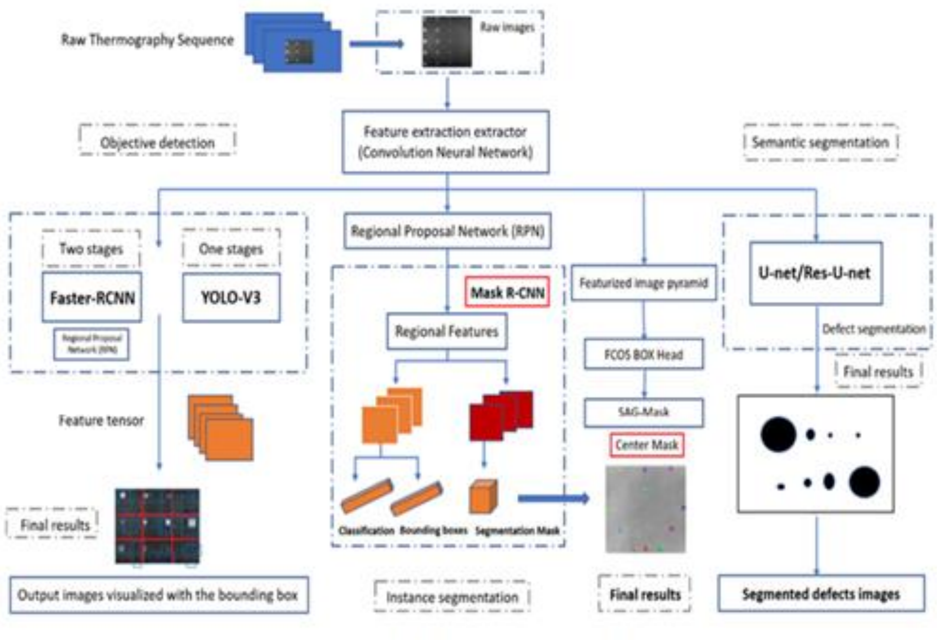


Figure 5. Three types of deep-learning methods (objective detection; instance segmentation; semantic segmentation).

6.1. Objective Localization Algorithms Method 1:

Real-time defect localization (YOLO-V3) YOLO-v3 is a proposed supervised deep-learning algorithm that has excellent detection capability both on the large or small objects due to its concatenation involving the merging of the features from the earlier layer with the features from the deeper layer, especially during the infrared nondestructive evaluation with an automatic defect detection task (subsurface defects case).

Processing images with YOLO v3 is quite fast and simple, allowing defects to be detected and localized directly. To perform the feature extraction, residual networks and successive 3 × 3 and 1 × 1 convolutional layers are localized in YOLO-v3 in Figure 6. The skip-connections mechanism was achieved by residual networks through multiple residual units [9,10], which was proposed to improve the performance of object detection, and also solve the gradient vanishing issue. In this research, the YOLO-v3-based deep architecture neural network is proposed to perform the detection of defects (of various sizes). This algorithm includes the implementation of three steps. First, the pictures are resized as the input size. Then, an entire convolutional network is run on these pictures. Lastly, we threshold the detection results based on the model confidence scores. In Figure 7, an example is shown of an original image (a) and a detected image (b) from the YOLO-V3 network. The CNN was able to distinguish the components, which have a similar thermal pattern with defects during the processing of thermal diffusion, which indicated that the supervised learning method (YOLO-V3) is less influenced by the boundary information in the components

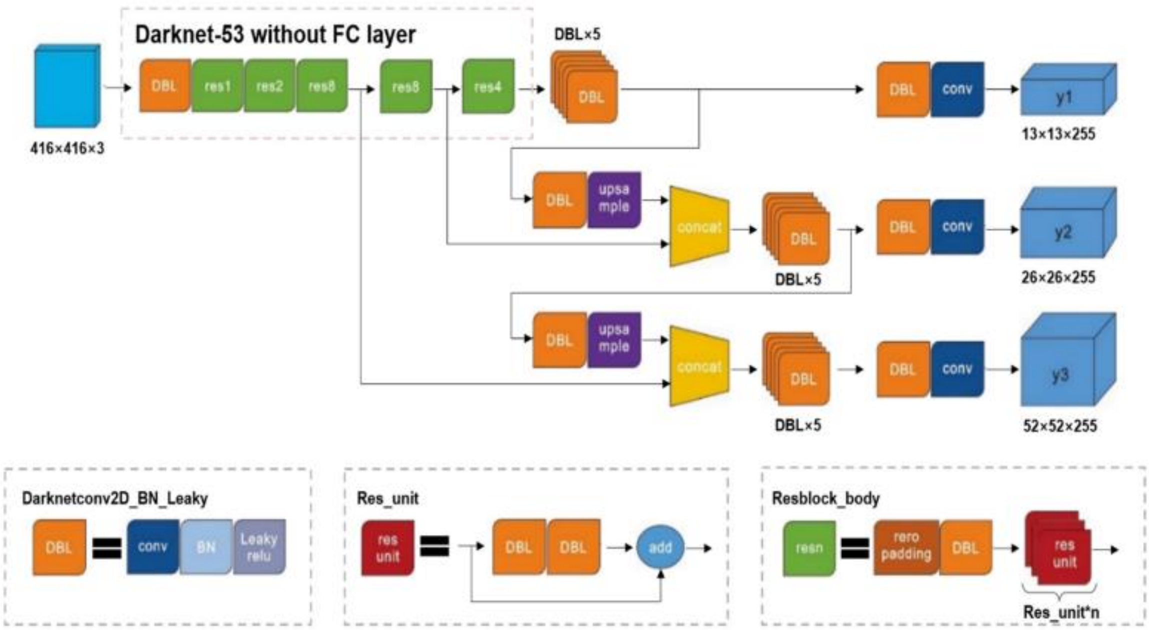
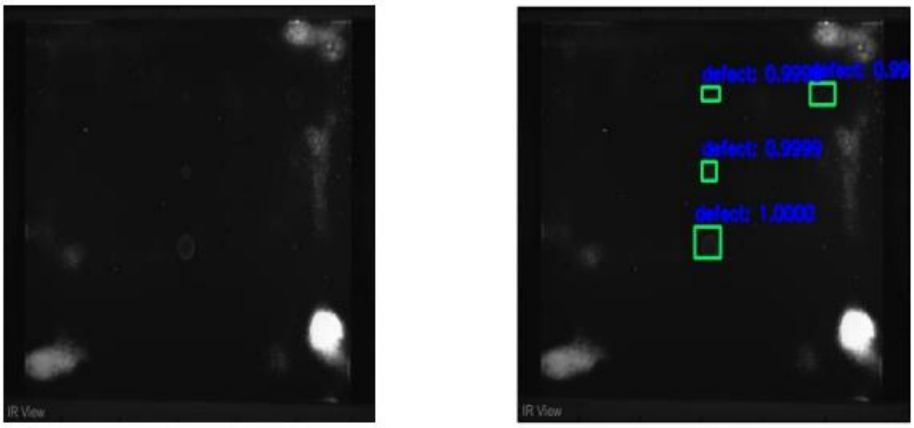


Figure 6: The architecture of residual units in Yolo-v3.



(a) (b)

Figure 7. An example of used Method A: (a) the original thermal image; (b) the detected image.

Method 2

Faster–RCNN is a real-time detector that achieved satisfying accuracy with several previous object localization applications in NDT [31]. In 2018, the Faster–RCNN was used for crack detection in an eddy current thermography diagnosis system. The neural network based on a deep architecture was proposed to deal with the problem of accurate crack detection and localization via the preprocessing unsupervised method (Principal Component Analysis). The deep architecture of Faster–RCNN is composed of several modules (Figure 8): 1. A fully convolutional network, which included five blocks of basic convolutional layers and a Relu layer with a pooling layer to extract feature from the input images. 2. A region proposal network (RPN) connected with the fully convolutional network to obtain the region of interest (RPI). 3. A Fast–RCNN detector using the feature region extracted in the (1)–(2) to achieve bounding box regression and SoftMax classification. The Faster R–CNN trained from multi-properties, rather than the regular unsupervised method, was limited with respect to certain properties that the defect information contained. An example image detected from Faster–RCNN, as well as a corresponding original thermal image, is shown in Figure 9. Figure 8. Faster–RCNN defect detection for infrared data. Figure 7. An example of used Method A: (a) the original thermal image; (b) the detected image.

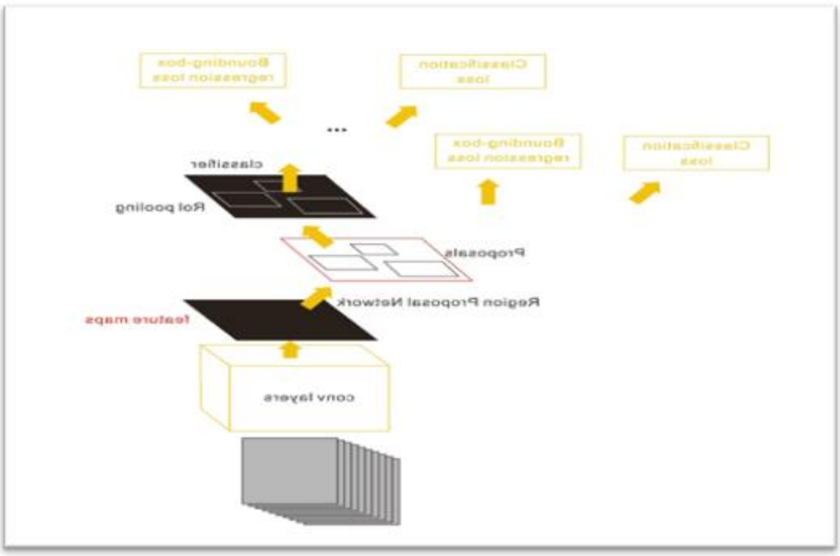


Figure 8 Faster–RCNN defect detection for infrared data.

Figure 9. An example of Method B: (a) The original thermal image; (b) The Faster–RCNN-detected image.

Semantic Defect Segmentation Method

Method3-defect-segmentation method with U-net network the U-net is an excellent auto-encoder format model to handle the training data with dimensionality reduction and data augmentation. It is worth evaluating the performance of semantic segmentation by U-net after extracting objective features from the temporal infrared sequence. In the previous article, the U-net was employed for the segmentation of wildland and forest fires as a deep-fire convolutional network obtaining very good performance. The convolutional architecture of U-net is inspired from the auto-encoder network architecture, as indicated in Figure 10. Contracting path maps from the original image to a low dimension vector by extracting meaningful feature representations, and the expansive path reconstructs the output of the desired feature maps. The contracting path is composed of a group of convolutional blocks: convolutional layers; rectified linear unit (ReLU); and max pooling (dimension reduction). The expansive path included groups of reconstruction blocks to upsample the feature: up-conv (half-reduce the feature channels), concatenation with a feature map from cropping in the contracting.

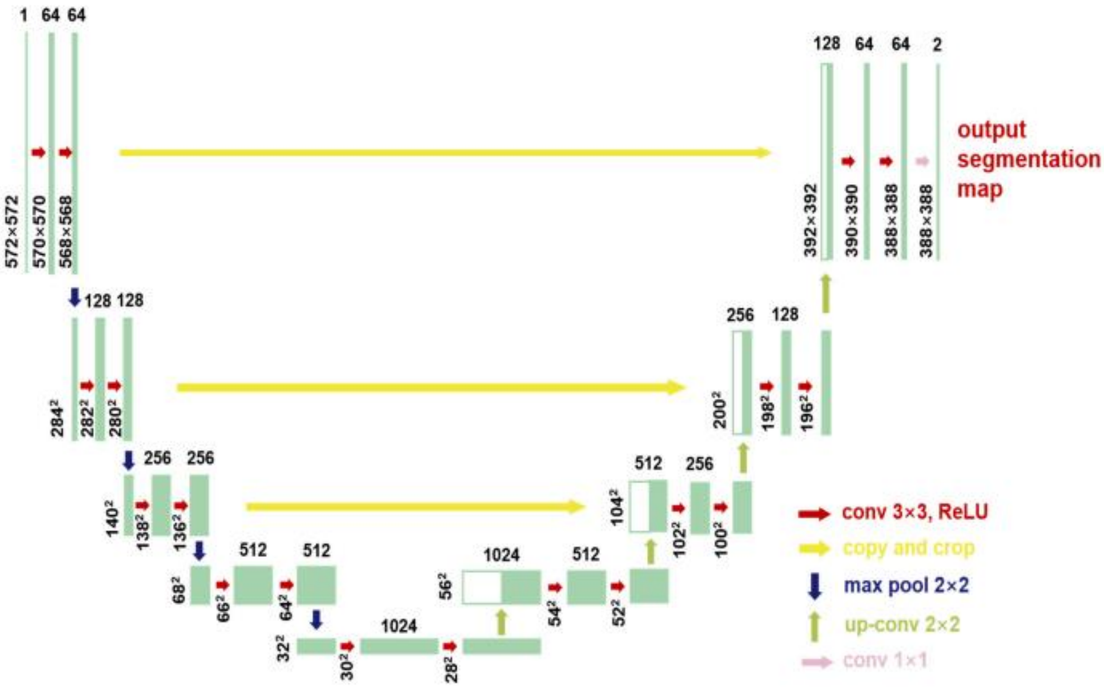


Figure 10. U-net model structure.

During the cooling period of the thermal data, a temperature change curve over time is obtained on the given image sequence. Therefore, each single thermal frame is fed into this model at the pixel level, and the thermal image can gradually capture the physical properties of temperature variation by U-net. The input values of U-net are thermal temporal evaluation vectors from each pixel. The output label is set either as 1 or 0 corresponding to the defect or non-defect region. During the validation stage, an obtained thermal sequence is selected as the input data after de-background and normalization. The output is a segmented image reconstructed from the predicted value as shown in Figure 11. Figure 11a is the corresponding original thermal image. (a) (b) Figure 11. An example of used Method C: (a) the original thermal image; (b) the being segmented image. Method 4: Res–U-net for defect semantic segmentation It is worth investigating comparatively to evaluate thermal sequence databases based on these different defect segmentation methods. As indicated in Figure 1, Res–U-net is an adapted novel encoder/decoder structure evolved from U-net in combination with several structures: residual connections ; atrous convolutions; pyramid scene Figure 10. U-net model structure. In the final layer, the feature vectors are classified into the target number of the class by 1 × 1 convolution. Moreover, this architecture relies heavily on data augmentation for its performance, which is explained below. The data augmentation strategy from the U-net architecture also brings a significant benefit for the performance for the training. Due to the characteristics of the spatial-thermal temperature sequence, the infrared thermal profile for the defect and non-defect pixels can be distinguished based on the labeling to the force implementation of the supervised learning method (U-net segmentation). During the cooling period of the thermal data, a temperature change curve over time is obtained on the given image sequence. Therefore, each single thermal frame is fed into this model at the pixel level, and the thermal image can gradually capture the physical properties of temperature variation by U-net. The input values of U-net are thermal temporal evaluation vectors from each pixel. The output label is set either as 1 or 0 corresponding to the defect or non-defect region. During the validation stage, an obtained thermal sequence is selected as the input data after de-background and normalization. The output is a segmented image reconstructed from the predicted value as shown in Figure 11b. Figure 11a is the corresponding original thermal image.

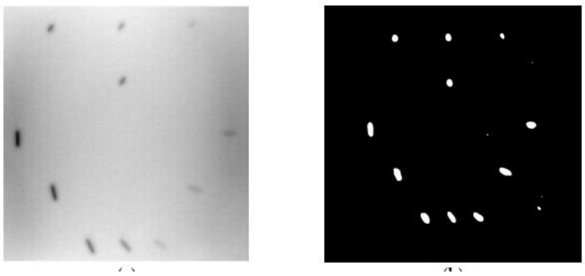


Fig 11 example of used Method C: (a) the original thermal image; (b) the being segmented image.

Method 4: Res–U-net for defect semantic segmentation It is worth investigating comparatively to evaluate thermal sequence databases based on these different defect segmentation methods. As indicated in Figure 12, Res–U-net is an adapted novel encoder/decoder structure evolved from U-net in combination with several structures: residual connections; atrous convolutions; pyramid scene parsing pooling [36]. Res–U-net can infer sequentially the boundary of the objects, the distance transforms of the segmentation mask, the segmentation mask, and a colored reconstruction of the input. Since residual blocks in Res–U-net can remove vanishing and exploding gradients to a great extent to improve the implementation efficacy of the learning mode and to achieve the pixel level of the segmenting of defects and classification, Res–U-net was compared with other state-of-the-art DL algorithms. The Res–U-net original was performed on the monotemporal aerial images for the task of semantic segmentation.: This reliable framework can perform semantic segmentation, resulting in high-resolution images.

To avoid the overfitting, the Res–U-net relied on the data augmentation strategy as well. Each image was rotated to the angle, zoom in/out, flip, and so on to enlarge the datasets of Res–U-net. In Figure 13, a segmented sample from Res–U-net (b) and the corresponding raw images (a) are shown.

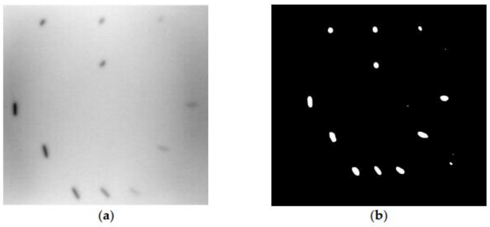


Fig 12 Example of used Method C: (a) the original thermal image; (b) the being segmented image

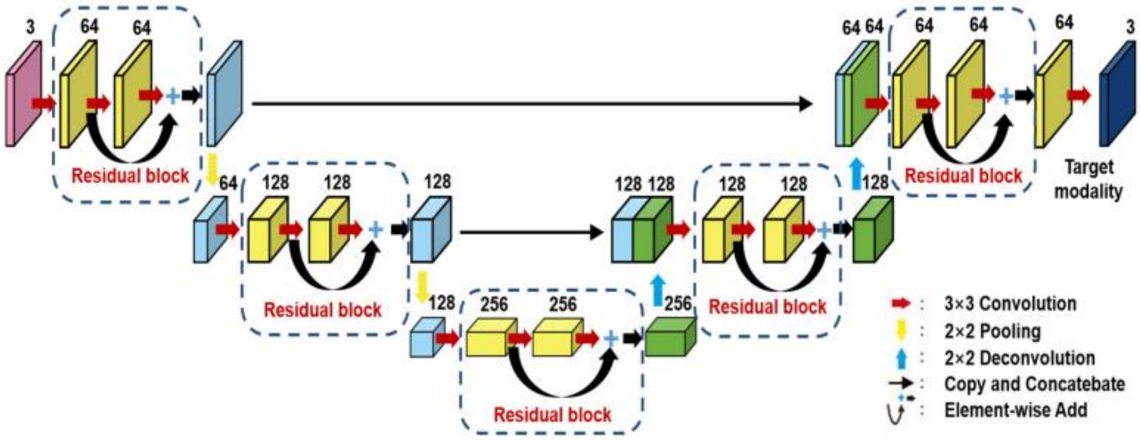


Figure 13. Res–U-net model structue

Figure 13. An example of Method 4: (a) the original thermal image; (b) the segmented image.

Method 5: MASK–RCNN for defect segmentation The Mask–RCNN detection procedure can be considered as either an object detection function or object segmentation function. Compared with the semantic segmentation, the instance segmentation associates each pixel of an image with an instance label. It can forecast a whole segmentation mask for each of those objects and predict which pixels in the input image correspond to each object instance. It also reduces the restriction to the position of defects rather than predicting a group of bounding boxes for the defects. Mask– RCNN is a classical instance segmentation method extended intuitively from Faster–RCN, which is an end-to-end trainable model to achieve pixel-to-pixel alignment segmentation between inputs and outputs of a convolutional backbone architecture.

ROI Align preserves spatial orientation of features with no loss of data for extraction over the entire image of the network. This approach efficiently detects objects in an image while simultaneously generating a high-quality segmentation mask for each instance. Each thermal image was fed into the backbone convolutional network from Mask–RCNN, once some learned region proposal was obtained from the backbone network.

These features projected learned region proposals onto convolutional feature maps. Mask–RCNN uses ROI aligning [39] to warp our feature from the convolutional feature map into the right shape then outputs it into two different branches. As shown in Figure 14, there are two different branches providing an output of predicted results. The top branch (blue line box) is a classification score of categories of region proposals and a bounding box for regression of coordinates in the output. In addition, at the bottom (red line box), a segmentation mask is predicted by the model for each of those region proposals to classify for each pixel in that input region proposal whether it is an object. Figure 15 provides an example of an original image from pulsed thermography (a) and a segmented image from Mask–RCNN (b).

Method 6: Central–Mask for defect segmentation Since the Mask–RCNN relies on the pre-defined anchors, its influence slowed down for the speed and accuracy in detection. Central–Mask is a simple yet efficient real-time anchor-free instance segmentation. Based on the structure, Central–Mask could be regarded as a novel spatial attention-guided mask (SAG–Mask) branch, adding a free anchor onestage object detector (FCOS) [40]. A segmentation mask head is located on each detected box with the spatial attention map that helps to aim attention at informative pixels and suppress noise. Figure 16 shows the overview architecture of Center–Mask. A feature pyramid extractor combines with the FCOS box head to predict classification scores and Sensorbounding box regression. A spatial attention-guided mask (SAG–MASK) predicts the segmentation map for the defects based on a spatial attention module [41] from each bounding box, which focuses on meaningful pixels and eliminates the noised influence. Central–Mask achieves a faster speed and surprising accuracy better than other state-ofthe-art instance segmentation approaches (Mask–RCNN). In this work, we adapted the Central–Mask network for feature extraction and defect segmentation. The main goal is to precisely detect and analyze defect information from the thermal images. The core strategy from this network is to extract the meaningful thermal pattern from the sequence for each specific defect. Figure 17 shows a raw thermal image (a) and a corresponding segmented thermal image (b) from Center–Mask. Each defect is precisely localized and segmented by the Mask. Sensors 2023, 23, x FOR PEER REVIEW 16 of 33 Mask–RCNN is a classical instance segmentation method extended intuitively from Faster–RCN, which is an end-to-end trainable model to achieve pixel-to-pixel alignment segmentation between inputs and outputs of a convolutional backbone architecture. ROI Align preserves spatial orientation of features with no loss of data for extraction over the entire image of the network. This approach efficiently detects objects in an image while simultaneously generating a high-quality segmentation mask for each instance. Each thermal image was fed into the backbone convolutional network from Mask– RCNN, once some learned region proposal was obtained from the backbone network. These features projected learned region proposals onto convolutional feature maps. Mask–RCNN uses ROI aligning [39] to warp our feature from the convolutional feature map into the right shape then outputs it into two different branches. As shown in Figure 14, there are two different branches providing an output of predicted results. The top branch (blue line box) is a classification score of categories of region proposals and a bounding box for regression of coordinates in the output. In addition, at the bottom (red line box), a segmentation mask is predicted by the model for each of those region proposals to classify for each pixel in that input region proposal whether it is an object. Figure 15 provides an example of an original image from pulsed thermography (a) and a segmented image from Mask–RCNN (b). Figure 14. Mask–R

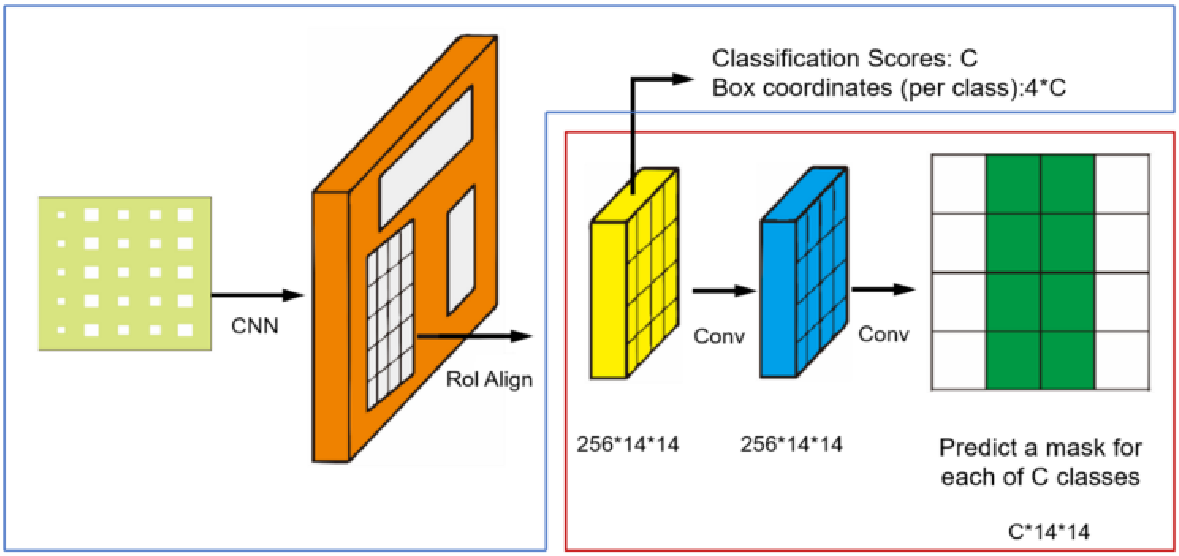


Figure 14. Mask–RCNN processing architecture.

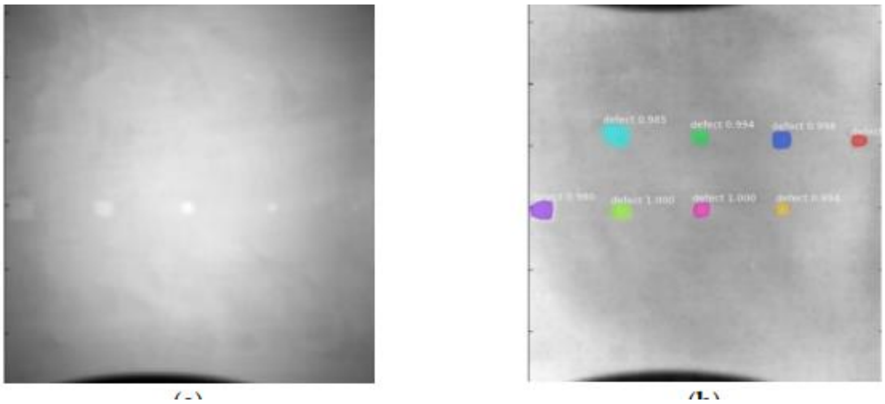


Figure 15. An example of used Method 5: (a) the original thermal image; (b) the detected image.

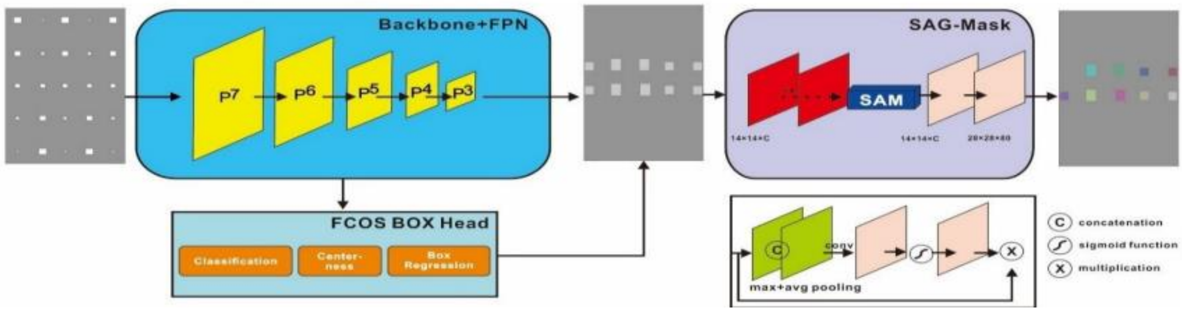


Figure 16. The structure of Center–Mask

Figure 17. An example of used Method: 6(a) the original thermal image; (b) the detected image.

Method 7: Absolute thermal contrast (ATC) with global threshold (GT) In combination with a global threshold method (GT), the ATC was adapted for the procedure of segmenting defects areas. The vital concept of this method was to compare the grey level of the pixel in the image coordinated (x, y) with the average grey level of a sound region of the sample, and it is often adapted in infrared image processing. Equation (3) Sensors 2023, 23, 4444 17 of 33 describes how this method works: where Tatc is the grey level in the ATC image in the coordinate (x, y) of the ATC image. Td (x, y) is the average grey level of the group pixels in the defect region and Ts(x, y) is the average temperature of a nearly sound region. Tatc = Td(x, y) − Ts(x, y) (3) Figure 18 provides an example of the segmentation with this method: (a) The raw image from pulsed thermography; and (b) The corresponding segmented image in Method 7. This method made it possible to reduce the effect from non-uniform heating and remove some thermal pattern noises. Sensors 2023, 23, x FOR PEER REVIEW 18 of 33 Figure 18 provides an example of the segmentation with this method: (a) The raw image from pulsed thermography; and (b) The corresponding segmented image in Method 7. This method made it possible to reduce the effect from non-uniform heating and remove some thermal pattern noises

(a) (b)

Figure 18. An instance of Method 7 applied on the thermal image (a) the original thermal image; (b) the detected image.

7. Evaluation Metrics

F-score and the probability of detection [45] are introduced to analyze the capability of detection of each detection deep-learning model, which is interpreted by Equations (4)–(7). The precision means the ratio from the cases contain the defects over the cases that are recognized by the system that contains the defects, which represent how accurate the system is in identifying the defects. The recall means the system correctly recognized the defects over the cases that actually contained the defects. The precision and recall values heavily depend on the confidences scores that the system is setting. The F-score is a method to estimate the detection and segmentation capability from these algorithms. β is a value to represent the weight between the precision and recall value. In this work, the recall is a metric that is more influential in evaluating the performance. Therefore, β is equal to 2. The POD reveals the accuracy of the method to detect the defects, which are always calculated at a specific confidence score value. Although the POD keeps the same mathematical format as the recall in the equation, POD represents a further explanation in quantifying research with NDT inspectors. In this work, we set the threshold for CTS at 75% for POD metric. Precision = TP TP + FP (4) Recall = TP TP + FN (5) POD = TP TP + FN (CTS = 75%) (6) F score = β 2 + 1 Precision × Recall (β 2 × Precision) + Recall (7) where TP is true positive, and FN is the false negative representing the number of the defects that have not been detected. Meanwhile, FP is the false positive defect representing the defects that are wrongly detected as defects when they are in fact not defects. Moreover, the confidence threshold score (CTS) was defined as a standard for measuring the accuracy of detecting corresponding objects in each dataset. CTS is a simple measurement standard that can be used for any task that yields a prediction range (bounding boxes, segmented maps) in the output regarding the ground truth.

9. DETECTION RESULTS

This model provided the shape and location of each defect detection results based on the labeled images with ground truth. In Table 3f, the noise of the input image is the main factor affecting the segmentation results. As indicated in the U-net result, the segmented image is not clear. The segmentation boundary is still blurry. A preprocessed image from principal component analysis (PCA) Sensors 2023, 23, 4444 20 of 33 was added in the validation database to verify whether the segmentation effect will be better after denoising in the Res–U-net model training. From the results, it seems the performance improved to some extent, and the test result of Resnet–U-net gave a better performance than the original U-net. Table 3. Results with semantic segmentation and object localization algorithms. Res–U-Net U-Net Faster–RCNN Yolo-v3 (a) Sensors 2023, 23, x FOR PEER REVIEW 22 of 35 seems the performance improved to some extent, and the test result of Resnet–U-net gave a better performance than the origin

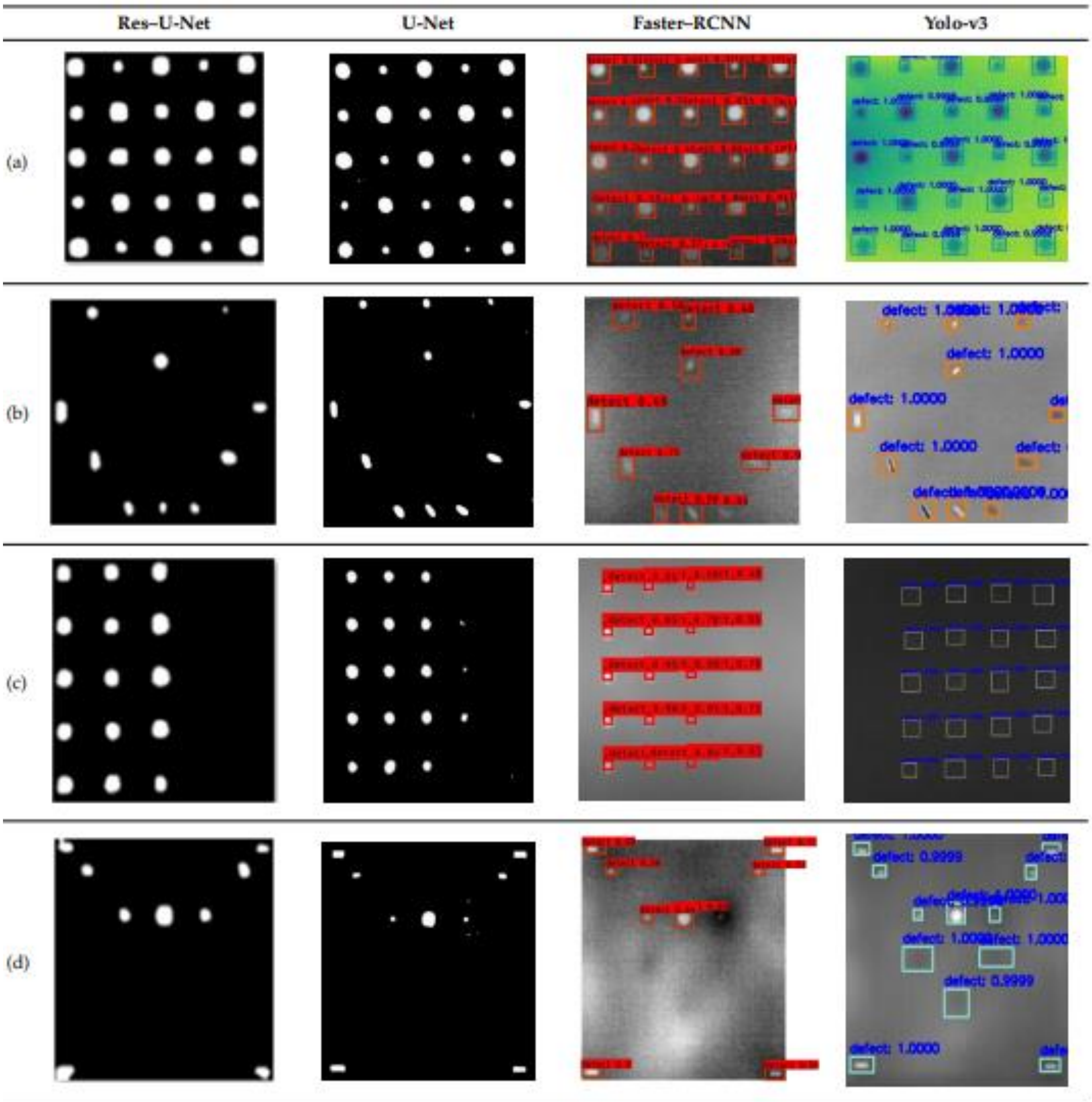


Table 3. Results with semantic segmentation and object localization algorithms.

Average frame per second for each deep-learning model. 7. Results Analysis The deep segmentation models gave attractive results for the Plexiglass/CFRP/Steel materials defects identification evaluation. This project focused on building and fine-tuning the training parameters for those defects. To improve the accuracy of the detection model, the way the dataset is built has a significant impact. According to the results obtained, the following analyses and points of this experiment were concluded below: 1. To implement a robust detection model, the databases must include enough samples. One way to effectively improve is to increase the size of the dataset by including multiscale images. A database composed of images on different scales (larger or smaller), enables the training to be sensitive to those new dimensions. This would increase the robustness of the deep segmentation algorithms facing larger defects, as well as improve the results on blurry pictures. To help reduce false alarms in the algorithm results and be more convenient for the user, implementing different types of labels is necessary. In the case of this project, each section was labeled with a defect in the spatial segmentation training (Mask–RCNN; U-net; Res–U-net). The proposal is to add different classifications. For example, including the name of the shape of the defect: circle, triangle, or some false positive cases (lighting spots, scratches) would be beneficial. This would allow the algorithm to not detect these shapes as a defect, and, thus, reduce the number of false alarms. 2. Another critical point in this experiment to be considered is the marking process. In comparison to other objective detection methods, Mask–RCNN/Center–Mask especially involves a pixel-based marking approach that could mark the defects accurately, as opposed to marking a considerable area around each defect. It can rapidly and easily annotate the object without the bounding boxes restrictions in most cases. In comparison with an instance-segmentation method, U-net and Res–U-net are the auto-encoder format DL models that can be trained based on each pixel level to semantically segment defect pixels from sound pixels. However due to the burden of tackling massive temporal data of thermal frames, U-net and Res–U-net have less time efficiency and high time complexity on the thermal data in comparison to the instance-segmentation model. Therefore, building and creating more diverse and Sensors 2023, 23, 4444 30 of 33 representative training samples is the key point in the future work in this research. There are several ways in which the size of the dataset can be effectively increased. Through data augmentation involving rotation, horizon flipping, and vertical shifts, the deep neural network model could learn the transformations further. By having different scales of larger or smaller training images, the learning procedure will be more sensitive to those new dimensions. This would also enhance the robustness of the algorithm to train for the detection of large defects and improve the results of grayscale images. 3. In addition, the specific training gave results for specific defects in the academic samples. In this work, training only involved using square, circle, and rectangle defects of plexiglass, CFRP, and steel samples. The detection results indicate that similar defects could be detected on other types of training samples. However, the results also show that if the learning model is tested on other defects that the model did not learn on, it would not be an accurate system to rely on. Hence, to use the deep-learning algorithm for training, we should clearly define the type of sample we are working on and enlarge the robustness of the system to learn this type of sample during the neural network training procedure. In addition, due to the time limitation, we simply labeled all the visible defects of each sample in this experiment. However, if we want to extract the feature map completely for each defect area, the positioning of less visible defects in infrared data will be a significant but challenging issue in further research. 4. A specific limitation of the objective localization algorithms is the influence of the labeling process. Although fast and efficient to use, the bounding boxes also led to some restrictions in most cases. As can be seen, when the circle is present in bounding box, this involves a defect that is totally bounded by the box.

However, this shows that although the entire defect is contained, the bounding box also extracted the non-defect area, which possibly introduces multiple errors and less accuracy in the results. The proposal is to make a pixel-based labeling to achieve integrity in the image segmentation, which would only label the defects and not a considerable area around each defect. This proposition can be further clarified by segmentation methods. The results presented here lead to a more reliable defects characterization with pulsed thermography (PT). 5. A good defect characterization is essential to not replace parts that could yet be used and to not leave critically damaged components without the needed repair. Therefore, these results are important, especially, e.g., in the designing of autonomous diagnosis NDT systems, which can make decisions regarding the integrity of the inspected part by themselves. In this work, three different types of automatic detection, being intelligent techniques, to combine with infrared thermography could improve the detection with industrial applications based on each group of results in the previous section. The critically damaged components could be easier identified and maintained the component that could be used by those algorithms with a high AP rate (81.06%). However, the instance segmentation (e.g., Center–Mask) provided the highest detection rate associated with vivid segmentation results among three different algorithms to provides the better solution of detection capability compared with the conventional thermal inspection method in industries. Therefore, it could be able to apply and contribute to current industrialized infrared inspection and controlling system. 6. Future work includes: (a) Tests that can be performed with the instance segmentation method and other NDT techniques based on images like stereography and holography; (b) The best technique, method instance-segmentation method (Center–Mask), which can still be improved by tuning achieves excellent performance, other network architectures must be tested and compared in the future to specify the best intelligent tool for defect measurement with infrared images. Sensors 2023, 23, 4444 31 of 33 8. Conclusions In this work, six spatial deep-learning models, involving instance segmentation (Mask– RCNN; Center–Mask), autoencoder format semantic segmentation (U-net; Res–U-net), and the object localization model (YOLO-V3; Faster–RCNN) are applied for defect detection in infrared thermography. The evaluated results and analysis from different geometric specimens of plexiglass, CFRP, and steel specimen with different aspect ratios (size/depth) are indicated in Section 6. Each POD curve is related to the defect sizes that assess the quality of the results to land smoothly in the case of catastrophic failure results. These spatial deep-learning models are separately and comparatively discussed in brief. Future work will focus on the detection of more complicated structured materials through the modification and combination of different spatial and transient deep-learning models.

REFERENCE

[1] Roy, R., Shaw, A., Erkoyuncu, J. A. & Redding, L. Through- life engineering services. Measurement and Control 46, 172–175 (2013).

[2] Hao, Q., Xue, Y., Shen, W., Jones, B. & Zhu, J. A decision support system for integrating corrective maintenance, preventive maintenance, and condition-based maintenance. in Construction Research Congress 2010: Innovation for Reshaping Construction Practice 470–479 (2010).

[3] Hull, J. B. & John, V. Non-destructive testing. (Macmillan International Higher Education, 2015).

[4] Vavilov, V. Thermal non destructive testing: short history andstate-of-art. in QIRT vol. 92 1928–1992 (1992).

[5] Balageas, D. et al. Thermal (IR) and other NDT techniques forimproved material inspection. Journal of nondestructive evaluation 35, 18 (2016).

[6] Schroeder, J. A., Ahmed, T., Chaudhry, B. & Shepard, S. Non- destructive testing of structural composites and adhesively bonded composite joints: pulsed thermography. Composites Part A: Applied Science and Manufacturing 33, 1511–1517 (2002).

[7] LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. nature 521, 436–444 (2015).

[8] Albawi, S., Mohammed, T. A. & Al-Zawi, S. Understanding of a convolutional neural network. in 2017 International Conference on Engineering and Technology (ICET) 1–6 (2017).

[9] Connor, J. T., Martin, R. D. & Atlas, L. E. Recurrent neural networks and robust time series prediction. IEEE transactions on neural networks 5, 240–254 (1994).

[10] Akula, A. & Sardana, H. K. Deep CNN-based Feature Extractor for Target Recognition in Thermal Images. in TENCON 2019-2019 IEEE Region 10 Conference (TENCON) 2370–2375 (2019).Yousefi, B. et al. Application of deep learning in infrared non-destructive testing. QIRT 2018 Proceedings (2018).

[11] Saeed, N., King, N., Said, Z. & Omar, M. A. Automatic defects detection in CFRP thermograms, using convolutional neural networks and transfer learning. Infrared Physics & Technology 102, 103048 (2019).

[12] Cuda-convnet: High-performance C++/CUDA implementation of convolutional neural networks. https://code.google.com/archive/p/cuda-convnet/ (2021).

[13] Krizhevsky, A., Sutskever, I. & Hinton, G. E. ImageNet Classification with Deep Convolutional Neural Networks. http://code.google.com/p/cuda-convnet/.

[14] Redmon, J., Divvala, S., Girshick, R. & Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. (2015).

[15] Fang, Q., Nguyen, B. D., Castanedo, C. I., Duan, Y. & Maldague II, X. Automatic defect detection in infrared thermography by deep learning algorithm. in Thermosense: Thermal Infrared Applications XLII vol. 11409 114090T(2020).

[16] Bang, H. T., Park, S. & Jeon, H. Defect identification in composite materials via thermography and deep learning techniques. Composite Structures 246, (2020).

[17] Ren, S., He, K., Girshick, R. & Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. arXiv preprint arXiv:1506.01497 (2015).

[18] Müller, D., Netzelmann, U. & Valeske, B. Defect shape detection and defect reconstruction in active thermography by means of two-dimensional convolutional neural network as well as spatiotemporal convolutional LSTM network. Quantitative InfraRed Thermography Journal 1–19 (2020).

[19] Fang, Q. & Maldague, X. A method of defect depth estimation for simulated infrared thermography data with deep learning. Applied Sciences 10, 6819 (2020).

[20] Duan, Y. et al. Automated defect classification in infrared thermography based on a neural network. NDT & E International 107, 102147 (2019).

[21] Luo, Q., Gao, B., Woo, W. L. & Yang, Y. Temporal and spatial deep learning network for infrared thermal defect detection. NDT & E International 108, 102164 (2019).

[22] Burgholzer, P., Thor, M., Gruber, J. & Mayr, G. Three- dimensional thermographic imaging using a virtual wave concept. Journal of Applied Physics 121, 105102 (2017).

[23] Kovács, P. et al. Deep learning approaches for thermographic imaging. Journal of Applied Physics 128, 155103 (2020).

[24] C. Ibarra-Castanedo, J. R. Tarpani, and X. Maldague, "Nondestructive testing with thermography," Eur. J. Phys., vol. 34, pp. S91–S109, Oct. 2013.

[25] B. Gao, L. Bai, W. L. Woo, G. Y. Tian, and Y. Cheng, "Automatic defect identification of eddy current pulsed thermography using single channel blind source separation," IEEE Trans. Instrum. Meas., vol. 63, no. 4, pp. 913-922, Apr. 2013.à

[26] Shepard S M. Advances in pulsed thermography[C]//Thermosense XXIII. International Society for Optics and Photonics, 2001, 4360: 511-515.

[27] N. Rajic, “Principal component thermography for flaw contrast enhancement and flaw depth characterisation in composite structures,” Compos. Struct., vol. 58(4), pp. 521–528, Dec. 2002.

[28] Maldague X, Galmiche F, Ziadi A. Advances in pulsed phase thermography[J]. Infrared physics & technology, 2002, 43(3-5): 175-181.

[29] Balageas D L, Roche J M, Leroy F H, et al. The thermographic signal reconstruction method: a powerful tool for the enhancement of transient thermographic images[J]. Biocybernetics and biomedical engineering, 2015, 35(1): 1-9.

[30] Deng L, Yu D. Deep learning: methods and applications[J]. Foundations and trends in signal processing, 2014, 7(3–4): 197-387.

[31] Fang Q, Nguyen B D, Castanedo C I, et al. Automatic defect detection in infrared thermography by deep learning algorithm[C]//Thermosense: Thermal Infrared Applications XLII. International Society for Optics and Photonics, 2020, 11409: 114090T.

[32] Fang Q, Ibarra-Castanedo C, Maldague X. Automatic defects segmentation and identification by deep learning algorithm with pulsed thermography: Synthetic and experimental data[J]. Big Data and Cognitive Computing, 2021, 5(1): 9.

[33] Hu J, Xu W, Gao B, et al. Pattern deep region learning for crack detection in thermography diagnosis system[J]. Metals, 2018, 8(8): 612.

[34] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, et al., “Generative adversarial nets,” In Advances in Neural Information Processing Systems, 2014, pp. 2672–2680.

Citation: Prashant Jadhav, Sandeep Thorat, Sachin Pawar, B.K. Patle, A Review of Thermal Imaging Based Internal Crack Detection Using Deep Learning (AI), International Journal of Thermal Engineering (IJTE),12(2), 2024, pp. 1–20.

Abstract Link: https://iaeme.com/Home/article_id/IJTE_12_02_001

Article Link:
https://iaeme.com/MasterAdmin/Journal_uploads/IJTE/VOLUME_12_ISSUE_2/IJTE_12_02_001.pdf

✉ **editor@iaeme.com**