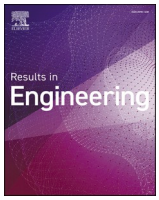
Results in Engineering 22 (2024) 102003

Contents lists available at ScienceDirect

Results in Engineering

journal homepage: www.sciencedirect.com/journal/results-in-engineering



Deep Learning and Image data-based surface cracks recognition of laser nitrided Titanium alloy

Muhammad Rizwan Awan [a,*], Chi-Wai Chan [a], Adrian Murphy [a], Dileep Kumar [b], Saurav Goel [c,d], Caroline McClory [e]

[a] School of Mechanical and Aerospace Engineering, Queen's University Belfast, UK
[b] Department of Electronic Engineering, Universitat Politechnica De Catalonia (UPC) Barcelona Tech, Spain
[c] School of Engineering, London South Bank University, London, UK
[d] Department of Mechanical Engineering, Shiv Nadar University, Gautam Budh Nagar, Greater Noida, India
[e] Northern Ireland Technology Centre (NITC), Queen's University Belfast, UK

ARTICLE INFO

Keywords:
Laser nitriding
Crack recognition
Deep learning
Transfer learning
Anomalies
Binary classification

ABSTRACT

Laser nitriding, a high-precision surface modification process, enhances the hardness, wear resistance and corrosion resistance of the materials. However, laser nitriding process is prone to appearance of cracks when the process is performed at high laser energy levels. Traditional techniques to detect the cracks are time consuming, costly and lack standardization. Thus, this research aims to put forth deep learning-based crack recognition for the laser nitriding of Ti–6Al–4V alloy. The process of laser nitriding has been performed by varying duty cycles, and other process parameters. The laser nitrided sample has then been processed through optical 3D surface measurements (Alicona Infinite Focus G5), creating high resolution images. The images were then pre-processed which included 2D conversion, patchification, image augmentation and subsequent removal of anomalies. After preprocessing, the investigation focused on employing robust binary classification method based on CNN models and its variants, including ResNet-50, VGG-19, VGG-16, GoogLeNet (Inception V3), and DenseNet-121, to recognize surface cracks. The performance of these models has been optimized by fine tuning different hyper parameters and it is found that CNN base model along with models having less trainable parameters like VGG-19, VGG-16 exhibit better performance with accuracy of more than 98% to recognize cracks. Through the achieved results, it is found that VGG-19 is the most preferable model for this crack recognition problem to effectively recognize the surface cracks on laser nitrided Ti–6Al–4V material, owing to its best accuracy and lesser parameters compared to complex models like ResNet-50 and Inception-V3.

1. Introduction

Surface-coated components and structures find extensive application in various industries, including medical equipment manufacturing, medical implants, optical device production, precision cutting tool fabrication, molding can production, and aesthetic gear manufacturing [1–5]. Detecting surface defects becomes crucial as they can provide valuable insights into the quality of the surface coating technique employed and the superiority of the raw materials used for both the coating and the base [6]. The development of reliable and efficient crack recognition techniques for coated surfaces have been a subject of extensive research in the past such as eddy current testing, ultrasonic testing, laser testing and microwave testing [7–10]. Such techniques are time consuming, costly and require human expertise to interpret the results which can lead to variability and errors.

To overcome the limitations of the noted traditional techniques, advancements in image processing and computer vision techniques have provided new opportunities for crack recognition on metal surfaces and coatings. By analysing digital images or optical profiles of surfaces through machine vision, it becomes possible to recognize the presence of cracks with high accuracy and efficiency [11–14]. In computer vision, convolutional neural networks (CNN) have emerged as a focal point of research for metal surface defect recognition [15–18]. CNNs are one of the foundational and most used models in computer vision domain and can automatically perform end-to-end learning by integrating feature extraction and classification into a single framework with minimum

* Corresponding author.
E-mail address: m.awan@qub.ac.uk (M.R. Awan).

https://doi.org/10.1016/j.rineng.2024.102003
Received 2 November 2023; Received in revised form 26 January 2024; Accepted 8 March 2024
Available online 18 March 2024

number of parameters compared to the complex DL models. A number of authors have directed their efforts to use CNN and its variants for surface defect recognition. For instance, Nguyen et al. [19] employed a CNN-based multi class classification algorithm to classify surface defects in cast products, including blow hole, chipping, crack and wash defects. The developed CNN model achieved accuracy of more than 95% for all of the studied defects. Yang et al. [20] applied transfer learning using AlexNet as the proposed model on a small data set to detect damage on the surface of wind turbine blades. The authors used a combination of different data set arrangements varying in ratios between training, validation and testing, and compared the model's results in different modes of transfer earning and without transfer learning. Transferred learnt models achieved accuracy of 92% in all data arrangements. Comparison of different CNN models is a proven technique to find the optimum model for defect classification. In this direction, Gao et al. [21] developed a semi-supervised model using CNN variants such as ResNet-50, DenseNet-201, VGG-19 and EfficientNet-b0 to detect steel surface defects by incorporating pseudo labels to utilize unlabelled samples. His results demonstrated that EfficientNet-b0 was the most accurate and the fastest model with 100% accuracy to detect steel surface defects. Cheng et al. [22] proposed a CNN based Microcrack-Net model to improve the feature detection of noisy images of microcracks on the terminal electrode of tantalum capacitors. The proposed model gave superior performance in comparison to the more advanced variants of CNN like Resnet-50, resNext50, VGG16, SegDecNet and Racki-Net to detect micro cracks.

Similar kind of efforts have been made to detect defects in surface coatings, for instance, Lu et al. [23] trained different variants of CNN using transfer learning to classify the accuracy of thermal barrier coatings using high resolution SEM images. Zhang et al. [24] used multi-scale faster R–CNN to detect cracks in thermal barrier coatings of turbine blades. Xiao et al. [25] employed the deep CNN using 2D SEM images to detect the cracks in air plasma sprayed thermal barrier coatings with 98% accuracy. Li et al. [26]detected cracks during a laser cladding process by using a combination of signal processing of acoustic emission signals and deep neural networks. Signal processing was used to extract different features from acoustic emission signals and CNN was used to identify the cracks based on the extracted feature vector. Zhao et al. [27] proposed a CNN based classification method for detecting coating surface defects. Experimental results show that the CNN models, particularly ResNet-50, achieve high precision and accuracy of more than 90% in classifying various coating defects. These studies show that Deep Learning (DL) is an efficient technique to recognize defects in surface coatings.

Laser nitriding is one of the most efficient surface modification techniques for the improvement of surface performance [28,29]. During this process, a metal surface is irradiated by sufficient laser energy in nitrogen containing atmosphere. Laser nitriding results in the formation of a titanium nitride (TiN) layer on the metal surface. This layer enhances the surface properties, increasing its hardness and resistance to both corrosion and wear [30]. However, a problem associated with laser nitriding process is the appearance of cracks on the surface which can initiate the process of mechanical and corrosion failures. These cracks often arise due to rapid thermal expansion and contraction during the laser nitriding process, creating stress within the Ti material [31]. The formation of cracks in nitrided layers of Ti materials is primarily attributed to the introduction of excessive residual stress during the laser nitriding process. As outlined by Holmberg et al. [32], there are two primary mechanisms through which residual stress leads to crack formation. The first involves the generation of excessively high lateral tensile residual stresses in the semi-brittle layer. Such stresses can result in the formation of tensile cracks, which typically develop perpendicular to the interface. The second mechanism pertains to the occurrence of critically high compressive stresses, leading to a regular pattern of shear cracks. It is important to note that the spatial distribution of these stresses within the material layers plays a pivotal role in determining the likelihood and nature of crack formation. Recent research conducted by Shirazi et al. [33] indicates the impact of tensile stress within the nitrided layer as a direct contributor to crack formation. When the tensile stress surpasses the material threshold for withstanding elongation, cracks begin to emerge. In contrast, the compressive stress present in the Ti substrate acts as a counterbalancing force to the tensile stress in the nitrided layer. While compressive stress typically inhibits crack formation, the differential in stress types and intensities between the nitrided layer and the Ti substrate can induce instability at their interface. It is this interplay of tensile stress in the nitrided layer and compressive stress in the Ti substrate that influences the formation and propagation of cracks.

Ti–6Al–4V is a titanium alloy widely employed across diverse industries, including aerospace, medical, and automotive, owing to its exceptional combination of strength, lightweight characteristics, and superior corrosion resistance [34]. The selection of Ti–6Al–4V for the laser nitriding process is grounded in its prevalence in applications where a balance of robust mechanical properties and reduced weight is paramount [35,36]. The urgency of addressing crack formation in Ti–6Al–4V surfaces, particularly laser nitriding, is underscored by the critical applications of this material in key industries. For instance, in aerospace engineering, the structural integrity of components made from Ti–6Al–4V is paramount for ensuring the safety and reliability of aircraft [37].

However, no significant effort has been made in the recent past to recognize the surface defects on laser intruded Ti–6Al–4V using DL. To bridge this gap, the present study aims to put forth a robust DL based crack recognition process for laser nitride surfaces, which has not been attempted before.

In this study, state of the art laboratory equipment has been used to perform the laser nitriding of Ti–6Al–4V, and for high resolution image collection. Convolutional neural network (CNN) and its different variants have then been employed using transfer learning for binary classification between the cracked and no crack images. Further, comparative analysis of CNN and its variants have enabled deep insight on the accuracy and speed of the approach.

Compared to existing research, the main scientific contribution of this work is as follows:

- Development of consistently high-quality nitrided surfaces, specially tailored for accurate crack detection
- Robust binary classification method to recognize cracks on the surface of Ti–6Al4V
- Development of better image acquisition and pre-processing techniques for complex laser nitrided images
- Comparative assessment of CNN and its variants to select the best model for surface crack recognition with best possible accuracy.

Furthermore, the paper is organized into sections, section 2 provides a brief overview of the methodology, data collection, and image-processing. Section 3 details the architecture description of the transfer learning models used for the crack recognition and performance evaluation metrics. Section 4 presents and discusses the achieved results. In section 5 and 6, the conclusions and recommendations for future improvement are provided.

2. Overview of the methodology

To perform the crack recognition using DL, data has been collected and processed in different stages as shown in Fig. 1.

In the first stage, state of the art equipment has been used to perform the laser nitride coating process on the material, which has been prepared through machining and grinding process. Then high-resolution image data, of about 500 times magnification, of the laser nitrided sample has been collected, across a range of nitriding energy levels. The image data is then pre-processed to convert the images into a form

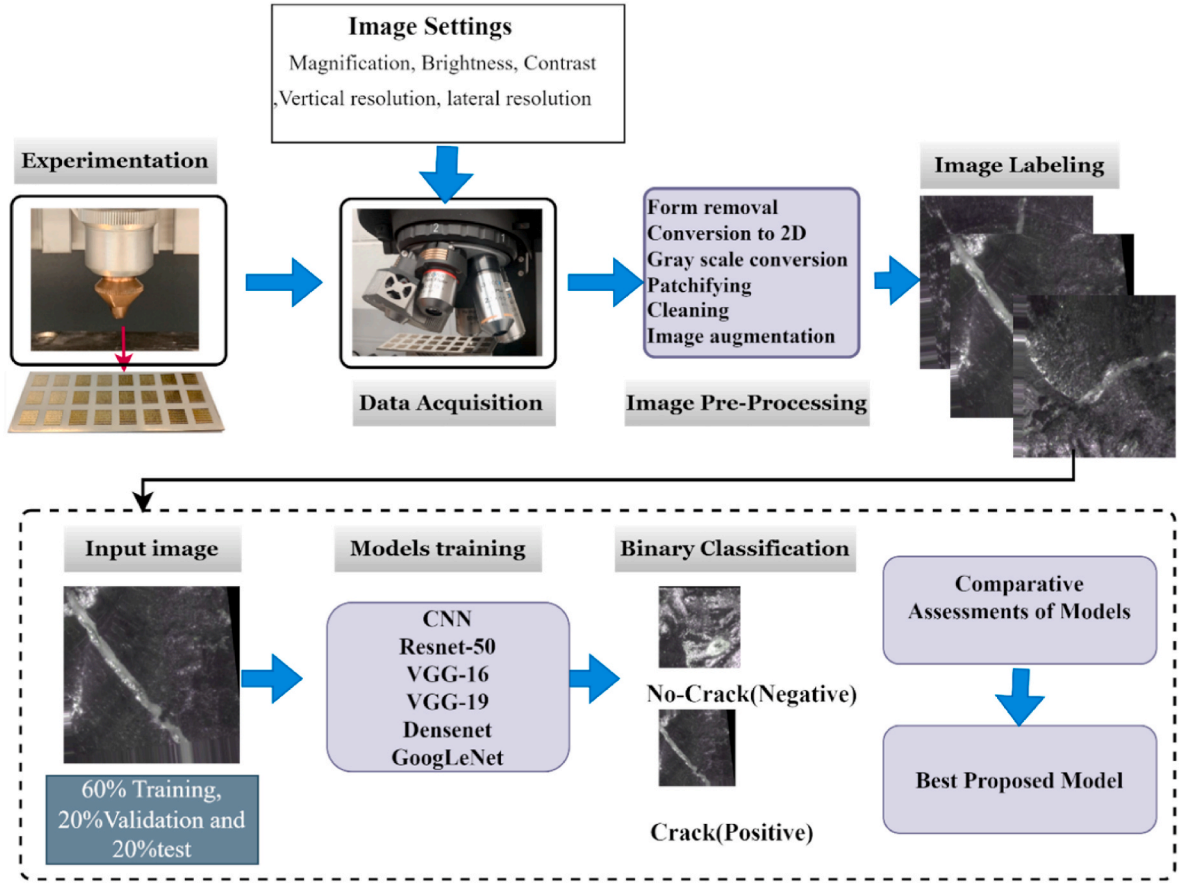


Fig. 1. Overview of the methodology.

readable by algorithms, and is cleaned to remove unwanted anomalies. CNN model and its variants including ResNet-50, DenseNet-121, and Google Net (Inception V3) have been employed as transfer learning approaches, to recognize the presence of cracks. These models are trained, validated and tested with a data division of 60%, 20% and 20%, respectively. Then the accuracy of these models has been calculated and analysed using: confusion matrix, metrics of precision, recall and F1 score.

2.1. Material preparation for experimentation

The material employed in this study is Ti–6Al–4V alloy was provided by American elements in plate form having dimensions of 250mmx250mmx2mm. To prepare the sample for the laser nitriding process, spark cutting was performed using Electric Discharging Machining (EDM) to cut the material in to smaller sizes of 125mmx30mmx2mm. To improve the surface quality of the samples, polishing and grinding was performed with silicon carbide (SiC) sand-papers ranging from 120 grit to 1000 grit. Subsequently, the polished samples were meticulously cleaned and degreased by immersing them in

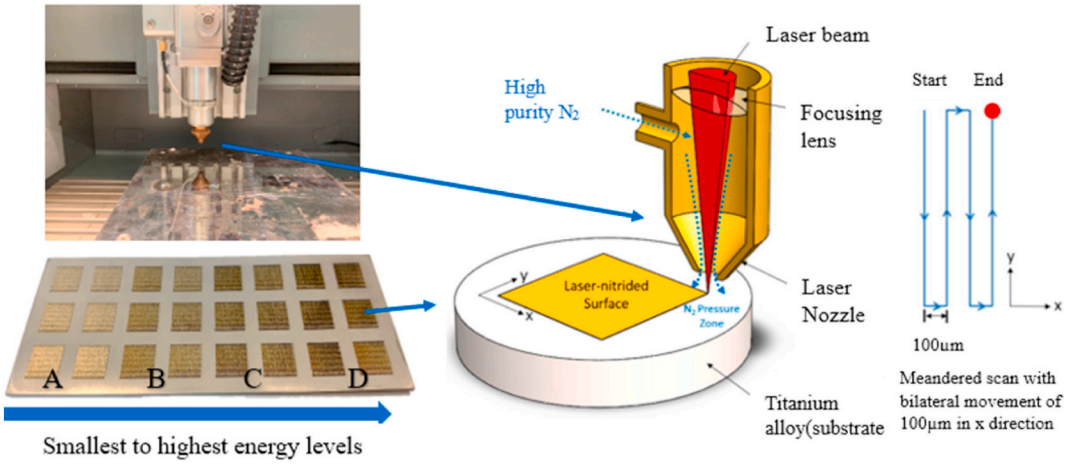


Fig. 2. Experimentation of laser nitriding [38].

an ultrasonic water bath containing acetone for 10 min. After the cleaning process, the samples were rinsed with deionized water and dried thoroughly using a stream of cool air.

2.2. Experimentation

The laser nitriding process was performed on the polished and cleaned samples using a SPI 200 W fiber laser machine in open air with local gas purging. The experimental setup of the laser nitriding process has been shown in Fig. 2.

The laser machine was the integration of two components, Micro Laser systems BV (Driel, Gelderland, the Netherlands) and a fiber laser system having 1064 nm in wavelength which is manufactured by SPI Lasers UK Ltd. (South Hampton, Hampshire, UK). The laser machine offered working in both continuous wave (CW) and modulated modes, with modulation capabilities up to 100 kHz and modulated pulses less than 5 μs The specific laser processing parameters were determined based on a series of preliminary experiments. They were set as follows: laser power of 45 W, duty cycle (DC) ranging from 40% to 100%, laser scanning speed of 25 mm/s, stand-off distance (SD) of 1.5 mm (referring to the distance between the laser nozzle and the sample surface) and purging with pure N2 gas at a pressure of 6 bar, which was delivered coaxially with the laser beam. The calculated laser spot size at the SD of 1.5 mm was approximately 100 μm. In the laser nitriding experiments, only the DC parameter was varied, while the other parameters remained constant.

The duty cycle (DC) is defined as the ratio of the ON time to the total time (ON + OFF). In the context of these experiments, a DC of 40% indicates that the laser was in the ON state for 40% of the total time, while a DC of 100% corresponds to continuous wave (CW) mode where the laser remained continuously ON. The samples were labelled as DC40, DC60, DC80, and DC100, indicating the nitriding at DC values of 40%, 60%, 80%, and 100%, respectively. The nitrided areas on the samples measured 10 × 10 mm^2, encompassing approximately 100 laser tracks. To ensure repeatability, each DC condition was repeated six times.

2.3. Image data acquisition

As the probability of cracks are greatest at the highest energy level, i.e. DC 100; and can be logically expected to be smallest at the lowest energy level, i.e. DC40, these two process areas were considered for image data collection. Fig. 3 elaborates the employed image data acquisition process.

The image data has been collected using Alicona Infinite Focus G5, which is a 3D coordinate measurement optical instrument based on Focus-Variation optical technology. 2 × 10mm^2 areas were captured at DC 40 and DC 100 as shown in Fig. 3. To get high resolution images of the area, extra-large image field datasets were collected using the ×50 objective lens. This approach provided surface magnification up to level of 50×. Measurement areas of 10 mm^2 were divided into 11 separate image fields and 396 million surface data points were captured per image field. To get appropriate images of any cracks, several arrangements were tried for exposure and contrast adjustment to get the right combination. About 25 high resolution images were collected for each small square of these 11 areas in a single nitride square. A total of 600 high resolution images of 5117 × 3491 pixels were collected for DC40 and DC100 energy levels.

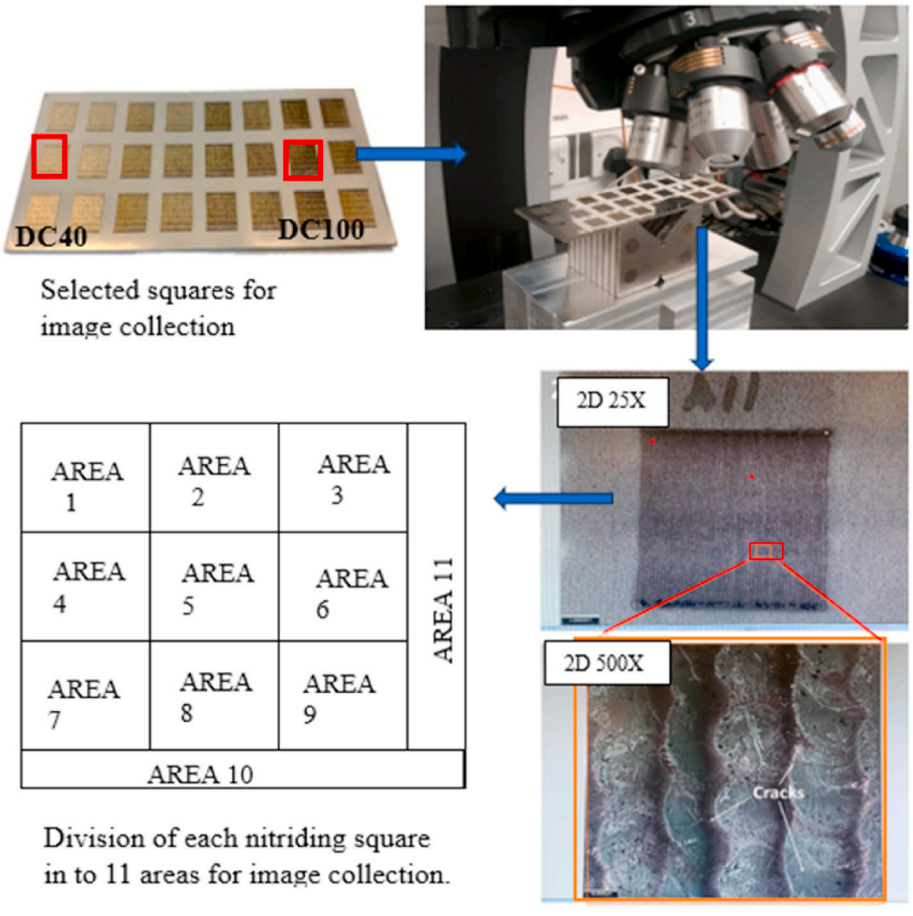


Fig. 3. Image data acquisition.

2.4. Image pre-processing

The images have been pre processed in three stages, as presented in Fig. 4. In the first stage, images have been pre processed using different images settings within Alicona. First of all, 3D profile of the laser nitrided surface has been captured. As the adopted models of CNNs have been inherently designed to work with 2D images, and working with 2D images reduces the complexity of neural network model and computational resources required, so using these 3D images or surface profiles are further converted in to 2D coloured images. In the next step, background noise of these images is reduced, which is termed as form removal. As shown in Fig. 4, step 2, cracks are not discernible in the coloured images, highlighting the height contrast of features in the nitrided surfaces, so images were converted into gray scale images which show the clear cracked lines.

The size of images is very large as shown in step 3, which makes it difficult for the model to process these images. So these images were further patchified (cropped) in to small number of images. This second step of the image pre processing has been completed using python library named as patchify. Cropping the images in to smaller sizes did not alter the crack features such as shape and orientation. So, image of size 5117 × 3491 was patchified into 25 number of smaller images, which increased the data set size to almost 3000 for cracked images. The image dataset was cleaned to remove images at the boundary line containing both the nitride surface and base metal. Images containing the surface analyser spots where it could not take measurement were also removed from the dataset.

3. Model architecture

3.1. CNN

Like other neural network models, CNNs utilize layers of interconnected neurons that enable them to learn hierarchical representations. These layers are connected via weights and biases. The input layer receives the image data as the initial layer, while the output layer provides predictions, such as classifying between cracks and no-cracks. Between these layers, there are hidden layers that transform the input's feature space to align with the desired output [39].

The standard CNN architecture adopted for this research has been shown in Fig. 5. It consists of three components, convolutional layer, pooling layer and classification layer. Crack features are extracted using the convolutional and pooling layers. The classification layer is usually connected to the last layer to recognize the presence of cracks. 16 kernels or filters of size (3 × 3) has been used in the first convolution layer and 32 kernels or filters of size (3 × 3) have been used in the 2nd convolution layer. For the convolution layer, ReLU is chosen as the activation function due to higher computation efficiency. After the convolution layer, the pooling layer is added to perform subsampling of the feature extracted by the convolutional layer and activated through ReLU. In this model, max pooling has been used to its ability of fast convergence and robustness. Max pooling is represented in Equation (2), where R is receptive field representing the local area within the input image, and $(u_{m,n})$ shows the value of node location (m,n) on the receptive field [21].

$$y_{max} = \max_{m,n \epsilon R} \left(u_{m,n}\right) \quad (2)$$

The classification layer consists of Global average pooling layer,

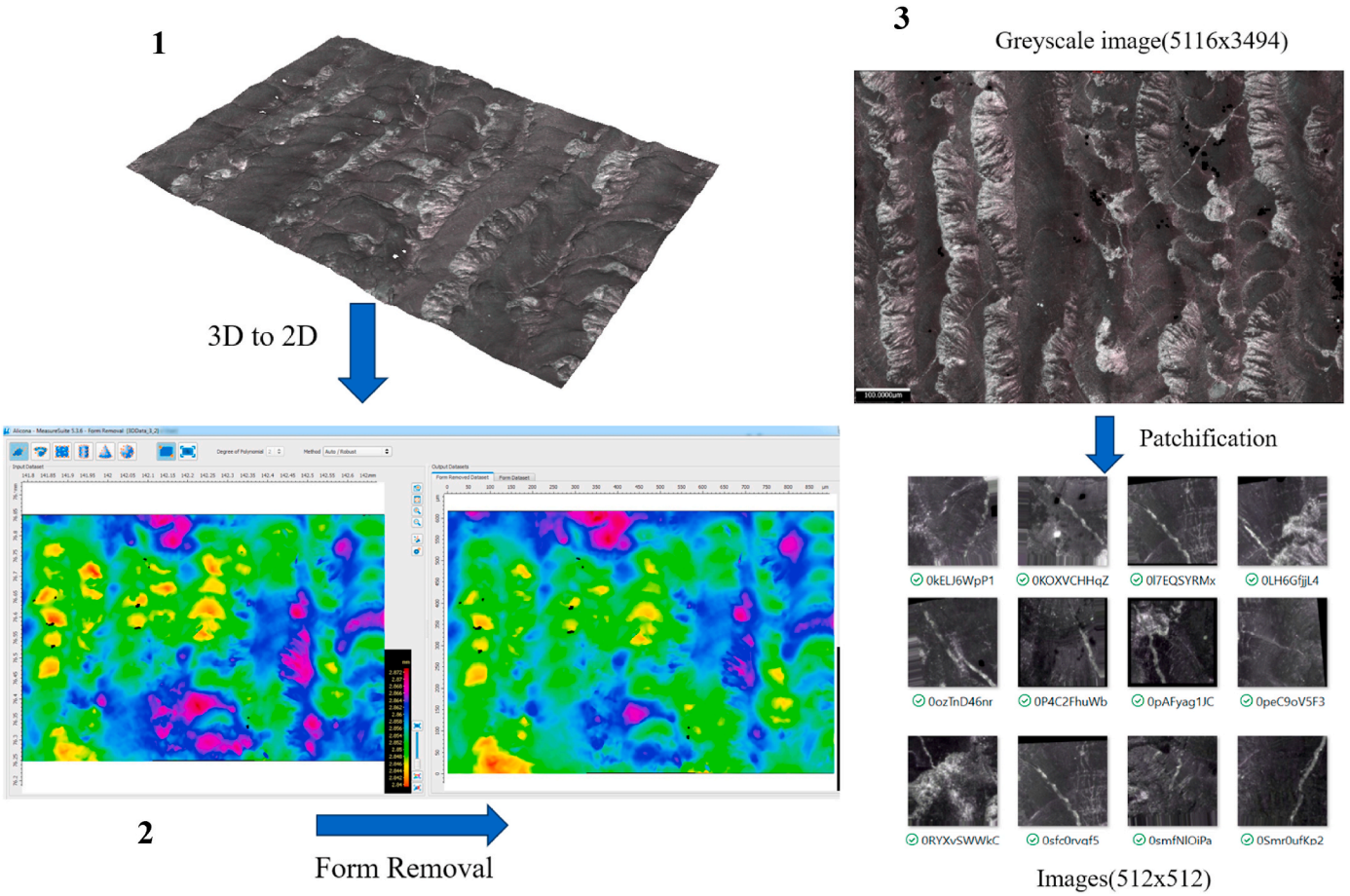


Fig. 4. Image pre-processing.

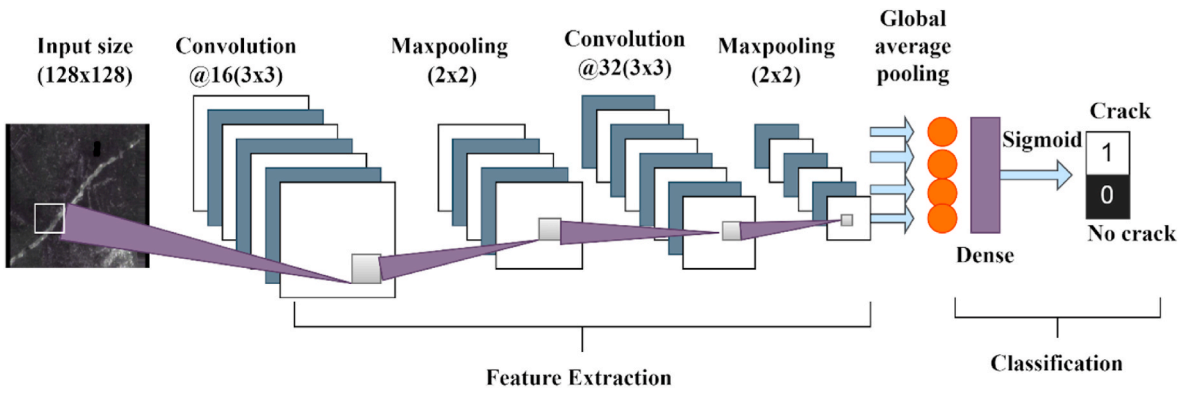


Fig. 5. CNN architecture.

dense layer and sigmoid layer. For this study, global average pooling layer (GAP) has been preferred over the fully connected layer in similar kind of studies [28,30]. After, Global average pooling dense layer has been used. The purpose of the used dense layer is to map the learned feature to the desired output. The dense layer functions as the final stage of feature extraction and mapping. After the dense layer, sigmoid activation function is used for binary classification, which shows the probability score for crack and no crack class.

3.2. Transfer learning models

In this research work, five pre trained CNN models (VGG19, VGG16, Densenet-121, Inception V3 and Resnet-50) have been used through transfer learning. The general archietcture of the model is shown in Fig. 6. All of these networks have been pretrained on ImageNet data set [40].

VGG-19 (Visual Geometry Group network), a widely known CNN architecture, was developed by Simonyan and Zisserman for image recognition tasks [41]. Due to an optimum number of layers in the network, VGG-19 has demonstrated proficiency in handling complex functions. The first five blocks of the architecture have been focused, in which each block has four CONV layer of 3×3 filters and a max pooling layer of 2×2 filters. At the end of the blocks, a global average pooling layer of 512 units has been used. For binary classification, the last two dense layers have 64 and 1 units respectively.

VGG-16. The architecture of the VGG-16 only differs from VGG-19 in terms of number of convolutional layers in each block. The first two blocks have two, and last three blocks have four CONV layer of 3×3 filters. The convolutional layer is followed by max pooling layer of 2×2 filters at the end of each block. For classification two dense layers of 64 units and 1 unit have been used.

GoogLeNet In the same year, Szegedy et al. [42] introduced the twenty-two layers deep CNN GoogLeNet, it was the best performing network at the image net large scale visual recognition challenge (ILSVRC) in 2014.

ResNet-50 was developed by He et al. [43]. Its unique architecture of skipping connections enhanced its accuracy, with reasonable convergence rates. At the (ILSVRC) 2015, ResNet emerged as the best network during the Common Objects in Context (COCO) challenge. The adopted model in this project uses 7×7 initial convolution layers to extract fundamental features, which is followed by a batch normalization and ReLU activation. After that a max pooling layer is applied to reduce the dimensions. A key part of ResNet-50 are the residual blocks which contain multiple convolutional layers with batch normalization and ReLU activation. After the residual blocks, global average pooling is applied to average out the feature maps. After that a single-unit fully connected layer is used for binary decision making.

DenseNet-121 Huang et al. [44] developed the densely connected layers deep CNN DenseNet, which stands out as it has fewer parameters, along with feature propagation and feature reuse. In this model, the initial convolutional layers are followed by dense blocks consisting of multiple densely connected layers. Within the dense blocks, the transition blocks control feature dimensions. Following the dense blocks, global average pooling is applied to extract the feature map, and then a fully connected layer with one unit is followed by a sigmoid activation function. Among the different DenseNet architectures (DenseNet-201, DenseNet-160, DenseNet-121), this study used DenseNet-121.

These five pre trained CNN models (VGG-19, VGG-16, Densenet-121, GoogleNet and Resnet-50) have been used as the feature extractors in this study. The base layers are frozen and custom classification layers are added on top of the base layers as shown in Fig. 6. The input images of the surface are provided to the classification layers to recognize the

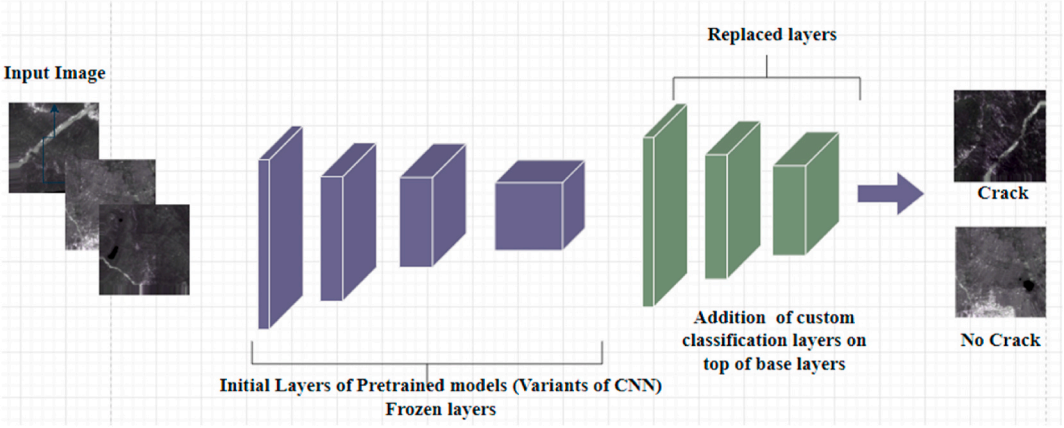


Fig. 6. Transfer learning architecture.

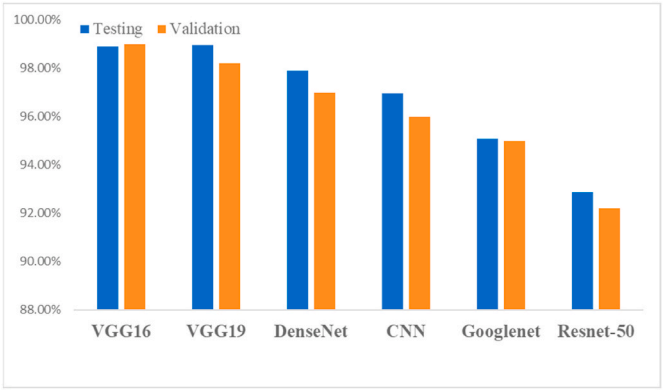


Fig. 7. Comparison of testing and validation accuracy.

presence of cracks. Models size, depth and parameters are given in Table 1.

3.3. Experimental setup and assessment criteria

After image preprocessing, the dataset was meticulously curated, leveraging domain expert knowledge, and cracked and non-cracked images were systematically organized into separate folders, labelled as 0 for non-cracked and 1 for cracked images, for precise binary classification during the model training process (see Table 2). These images were further augmented to increase the data set. In image augmentation, a combination of rotation, flipping, scaling, zooming and shearing was applied to enhance the image data set. As the probability of no cracks in the two image sets (DC40, and DC100) are low, the number of no cracked images was lower than the number of cracked images and thus the data set was unbalanced. The total number of images after augmentation were 6944 and 12,761 for cracked and no cracked conditions respectively. The images were further divided in to 60% training, 20% for validation and 20% for testing.

The process of binary classification of the coated surface images was performed by running the models on an Alienware 13th Gen Intel@Core-i9 Windows11 computer. A GPU of NVIDIA GeForce RTX 4080 with 12 GB DDR6 Ram was used to accelerate the model training and achieve faster calculations. To access the performance of the binary classification of the defects, a confusion matrix is used, which provides the comparison of the actual and the predicted classes of the input images [45]. To evaluate the accuracy of the model, different terms are used in the confusion matrix. True positive (TP), the number of instances correctly predicted as positives (Belong to the positive class). True Negative (TN) the number of instances that are correctly predicted as negative (belonging to the negative class). Here Cracked images have been defined as the positive class and no cracked images have been defined as the negative class. The false positives (FP) indicated here that number of actual non cracked images turned out to be cracked images. The false negatives (FN) indicates the number of cracked images incorrectly predicted as non-cracked images. The confusion matrix shows the detailed classification reports through performance metrics of: accuracy, precision, Recall and F measure.

The term Accuracy represents the overall performance of the model in terms of the number of correct predictions out of the total number of test cases, Equation (3).

$$\text{Accuracy} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{TN} + \text{FP} + \text{FN})} \tag{3}$$

Precision shows the correctly predicted positive cases from all the positive predictions. For instance, it will show how many of the images where correctly predicted to be cracked images out of the total number of true positive and false positive images, Equation (4).

$$\text{Precision} = \frac{\text{TP}}{(\text{TP} + \text{FP})} \tag{4}$$

Recall calculates how many of the actual positive (cracked images) where predicted correctly from the true positive and the true negative cases, Equation (5).

$$\text{Recall} = \frac{\text{TP}}{(\text{TP} + \text{TN})} \tag{5}$$

F1 score is the harmonic mean of precision and recall, it provides the mutual view of precision and recall.

Table 1
Models specifications.

Network	Depth	Size	Total Parameters	Trainable Parameters
VGG-19	19	574 MB	20,057,281	32,897
VGG-16	16	533 MB	14,747,585	32,897
Densent-121	121	33 MB	7,103,169	65,665
GoogLeNet	22	40 MB	22, 065,185	262,401
ResNet-50	50	167 MB	23, 718, 913	131,201

Table 2
Dataset generation.

Descriptor	Crack	No-Crack
Total Number of images	6944	12,761
Images for training (60% of total images)	4166	7656
Images for validation (20% of total images)	1328	1740
Images for testing (20% of total images)	1328	1740
Image size	128 × 128 × 3	128 × 128 × 3

4. Results and discussion

The formation of cracks during laser nitriding can be attributed to various factors including: Heat accumulation, thermal cycling effects, melt pool dynamics, residual stresses, microstructural evolution, depth of the nitrided layer, dendritic growth and rapid solidification, and micro segregation in dendritic regions.

Firstly, in continuous wave (CW) mode, lasers perform at a consistently high power (DC100), leading to an immediate temperature increase and subsequent heat accumulation. This rapid temperature change can introduce thermal stresses in materials, potentially causing cracks. Conversely, in modulated mode, lasers work intermittently, allowing materials periodic cooling periods, which can reduce these stresses. Secondly, using modulated mode, especially with duty cycles (DC) ranging from 40% to 60%, introduces thermal cycling. This periodic heating and cooling can foster a more consistent grain growth and mitigate residual stresses in the nitrided layer, reducing the likelihood of crack formation. The rapid cooling characteristic of CW mode can result in pronounced residual stresses. When the residual stress surpasses the yield strength of the material, it can lead to cracking. Yet, in modulated mode, the periodic heating can facilitate stress relaxation, decreasing the chances of crack formation [46,47].

The presence of cracks or no crack on the laser nitrided surface has been investigated using binary classification. The performance of the 5 models (Table 1) to recognize the cracks on the surface of laser nitrided Ti–6Al–4V have been analysed using testing and validation accuracy, confusion matrix and classification report. The results show the comparative outlook for binary classification of surface cracks on laser nitride surfaces using deep learning model of CNN, and its variants of Densenet-121, GooLeNet, Resnet50, VGG16 and VGG19, which have been applied through transfer learning. These models have been optimized by fine tuning their hyper parameters to achieve the best possible performance. These models provide the robust binary classification to recognize the surface cracks of laser nitride coating on Ti–6Al–4V. The purpose to include the baseline model of CNN in comparison to its variants is to gauge the trade-off between complexity and performance

for this data set. It also helped to benchmark the performance of CNN in comparison to its variants for this dataset.

Fig. 7 shows the comparison of the validation and testing accuracy of the applied models. Overall, all of the applied models show a testing and validation accuracy of more than 92%, which indicates that all of the models achieve good performance to classify cracks in laser nitrided surface coatings. The best validation and testing accuracy of 99% is achieved by the VGG16 and VGG19 models through transfer learning. Following closely, the DenseNet-121 demonstrates very good performance with an accuracy of 98%. The base model, CNN, demonstrated good performance with validation and testing accuracy of 96%. However, the transfer learnt models, GoogLeNet (Inception V3) and Resnet-50, showed lower accuracies of 95% and 92% respectively.

Looking at the number of total parameters and trainable parameters in Table 1, and the performance of the models in Fig. 7, it is important to note that models with less training parameters like VGG-19, VGG-16, and DenseNet-121 have better validation and testing accuracy. On the other hand, models like GoogLeNet and ResNet-50 with more trainable parameters have lower validation and testing accuracy. In comparison the base model of CNN with significantly fewer training parameters performed much better than the ResNet-50 and GoogLeNet (InceptionV3). For this data set, the network architecture of VGG16, and VGG19 were found to be more suitable and provided the testing and validation accuracy of 99% to recognize the cracks on the surface of laser nitrided Ti–6Al–4V.

Fig. 8 and Table 3 show the classification report of the models along with accuracy and testing time in milliseconds. In terms of testing time and accuracy, VGG-19 along with base model CNN provided the most optimum results, with accuracies of 99 and 97% and testing times of 12300 ms and 11145 ms respectively. VGG-16 and DenseNet-121 also exhibited very good accuracies of 99 and 98% but are slow in comparison to the CNN and VGG-19, with testing times of 13696 ms and 14250 milli seconds respectively. The DenseNet-121 model achieved accuracy of 98% with slightly more training time than the VGG models. ResNet-50 displayed lowest testing accuracy of 92%, but has a fast testing time. GoogleNet (Inception-V3) also has the low accuracy of 95%with very slow testing time of 13828 ms.

The CNN base model along VGG19 have turned out to be more accurate and fast, while inception V3 and have lower accuracy is slow in testing. The highest accuracy and faster speed of the models is attributed to the less training parameters used in the architecture of VGG19, VGG16 and CNN. The high accuracy of these models along with the lowest training times also indicate that the data set is more suitable for simpler networks with less trainable parameters.

The confusion matrix along with the classification report results as shown in Fig. 9 and Table 3, indicate that the unbalanced data set did not influence the models to become biased. As, the present models are designed for recognition of cracks, so precision is the most decisive parameter. Among the trained models VGG16, VGG19, Densenet-121 and CNN gave the same precision values to recognize the surface images with cracks and no cracks. However, a marginal difference in accuracy is seen in the case of GoogLeNet (Inception-V3) and Resnet-50, as shown in Table 3.

Except VGG19 and VGG16 models, all of the other models have marginal difference in their Recall Values for cracks and no cracks images.

Considering the testing and validation accuracy, and classification report, it appears that this data set was efficient to recognize the cracks on laser nitrided surface of Ti–6Al4V. Comparing the base model of CNN with its variants, it can deduced that simpler models with less trainable parameters were more accurate to recognize cracks in less time. Based on this, best model for this kind of data set is VGG19 to recognize and classify cracks.

Figs. 10 and 11 show that training, and validation loss and accuracy of the proposed model along with the number of epochs. The VGG 19 model was so accurate for this data set that it achieved the 98% accuracy at the 10th epoch and reached to the minimum loss of 0.05, at 99% accuracy at the 60th epoch.

This study is limited to recognizing the presence of surface cracks through image processing and the application of deep learning-based models. To comprehensively understand the evolution of cracks, including crack initiation, propagation, and growth, additional data collection is necessary. This involves new experimentation and conducting different material tests, such as XRD (X-ray diffraction) and SEM (Scanning Electron Microscopy).

5. Conclusion

The research demonstrated the DL based binary classification method to recognize the surface cracks on laser nitrided samples of Ti–6Al-4. A state-of-the-art experimental data set and data acquisition system has been deployed to process high resolution surface images of laser nitrided samples of Ti–6Al4V. The images were pre-processed and a base model of a CNN along with its five variants have been applied

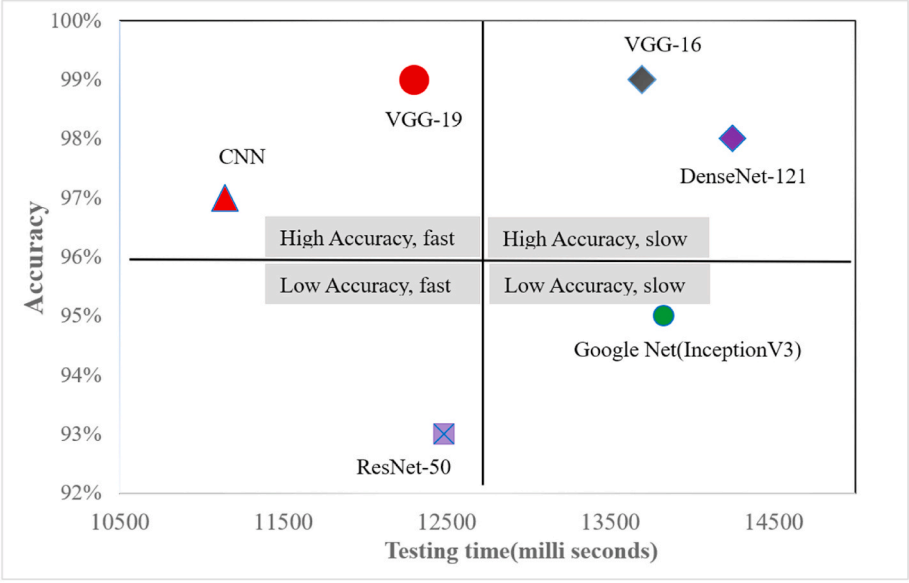


Fig. 8. Performance evaluation of the models based on their Accuracy and testing time.

Table 3
Classification report and testing time.

Network		Precision	Recall	F1 Score	Support	Accuracy	Testing Time
CNN	Crack	0.97	0.98	0.97	1328	0.97	11145 ms
	No-Crack	0.97	0.96	0.96	1740		
VGG19	Crack	0.99	0.99	0.99	1328	0.99	12300 ms
	No-Crack	0.99	0.99	0.99	1740		
Resnet-50	Crack	0.91	0.93	0.92	1328	0.93	12482 ms
	No-Crack	0.93	0.94	0.94	1740		
VGG16	Crack	0.99	0.99	0.99	1328	0.99	13696 ms
	No-Crack	0.99	0.99	0.99	1740		
Googlenet	Crack	0.96	0.93	0.94	1328	0.95	13828 ms
	No-Crack	0.95	0.97	0.96	1740		
Densenet	Crack	0.98	0.99	0.98	1328	0.98	14250 ms
	No-Crack	0.98	0.97	0.98	1740		

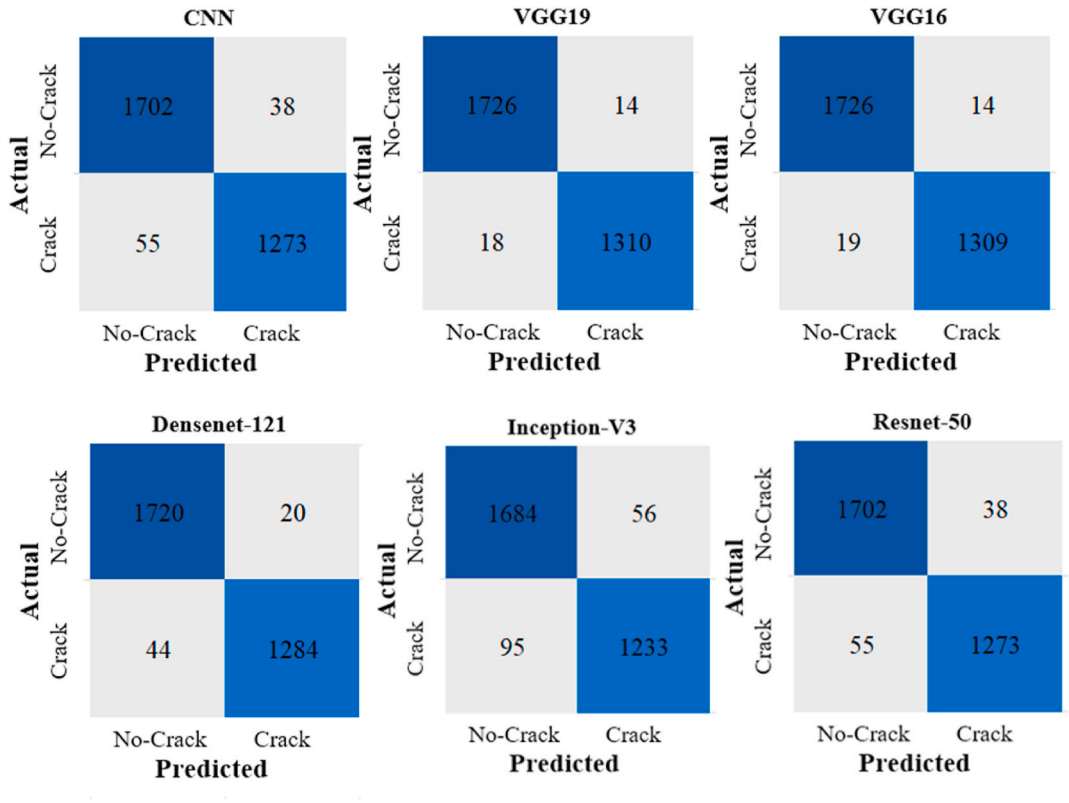


Fig. 9. Confusion matrixes.

through transfer learning to recognize the surface cracks. Model testing and validation accuracy, confusion matrix and classification report has been used as the assessment criteria. All of the models exhibited very good accuracy of more than 90% to recognize surface cracks. The CNN variants used with the transfer learning approach, with the smallest number of training parameters, were found to be fast and more accurate to make the predictions with an accuracy of 98%. Moreover, VGG-19 provides best crack recognition accuracy of 99%. For this kind of data set, the future work is to localize the cracks within the image. For this purpose, object detection using combination of Yolo variants and faster RCNN could be applied. The other method to detect cracks could be the image segmentation using the different variants of UNET models.

Funding

All authors gratefully acknowledge the support provided by UKRI via Grants No.: EP/T024607/1 & EP/S036180/1.

CRediT authorship contribution statement

Muhammad Rizwan Awan: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing – original draft. **Chi-Wai Chan:** Data curation, Methodology, Resources. **Adrian Murphy:** Funding acquisition, Project administration, Resources, Supervision, Writing – review & editing. **Dileep Kumar:** Data curation, Formal analysis, Investigation, Visualization, Writing – review & editing. **Saurav Goel:** Funding acquisition, Project administration. **Caroline McClory:** Data curation, Resources.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

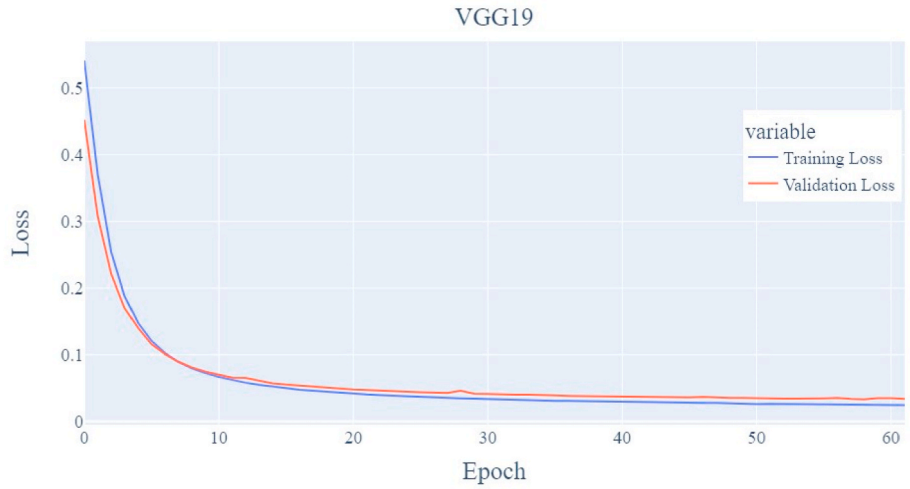


Fig. 10. Loss curve with number of epochs.

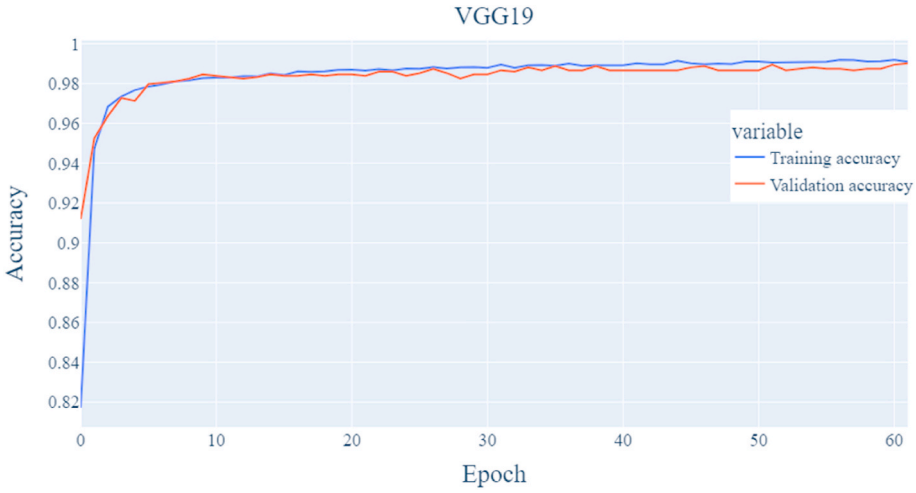


Fig. 11. Accuracy curve with number of epochs.

Data availability

Data will be made available on request.

References

[1] Y. Zhang, W. Li, C. Zhang, H. Liao, Y. Zhang, S. Deng, A spherical surface coating thickness model for a robotized thermal spray system, Robot. Comput. Integrated Manuf. 59 (Oct. 2019) 297–304, https://doi.org/10.1016/J.RCIM.2019.05.003.

[2] G. Liu, et al., Development of Bioimplants with 2D, 3D, and 4D additive manufacturing materials, Engineering 6 (11) (Nov. 2020) 1232–1243, https://doi.org/10.1016/J.ENG.2020.04.015.

[3] G. Hu, K. Guan, L. Lu, J. Zhang, N. Lu, Y. Guan, Engineered functional surfaces by laser microprocessing for biomedical applications, Engineering 4 (6) (Dec. 2018) 822–830, https://doi.org/10.1016/J.ENG.2018.09.009.

[4] Y. Wang, et al., Challenges and solutions for the additive manufacturing of biodegradable magnesium implants, Engineering 6 (11) (Nov. 2020) 1267–1275, https://doi.org/10.1016/J.ENG.2020.02.015.

[5] H. Khatun, M. Rahman, S. Mahmud, M.O. Ali, M. Akter, Current advancements of hybrid coating on Mg alloys for medical applications, Results in Engineering 18 (Jun. 2023) 101162, https://doi.org/10.1016/J.RINENG.2023.101162.

[6] Y. Aslam, N. Santhi, N. Ramasamy, K. Ramar, Localization and segmentation of metal cracks using deep learning, J. Ambient Intell. Hum. Comput. 12 (3) (Mar. 2021) 4205–4213, https://doi.org/10.1007/S12652-020-01803-8/TABLES/4.

[7] J. García-Martín, J. Gómez-Gil, E. V.-S.- Sensors, and undefined 2011, 'Non-destructive techniques based on eddy current testing', *mdpi.comJ García-Martín*, J Gómez-Gil, E Vázquez-SánchezSensors, 2011•mdpi.com 11 (2011) 2525–2565, https://doi.org/10.3390/s110302525.

[8] S. Dixon, S. Burrows, B. Dutton, Y.F. Ultrasonics, undefined, Detection of Cracks in Metal Sheets Using Pulsed Laser Generated Ultrasound and EMAT Detection, Elsevier, 2011 [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0041624X10000818. (Accessed 26 June 2023).

[9] C. Yeh, R. Z.-I. T. on I. and, and undefined 1994, 'A novel microwave method for detection of long surface cracks in metals', ieeexplore.*ieee.org*, Accessed: June. 26, 2023. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/328896/.

[10] N. Wang, R. Zoughi, Moment Method Solution for Modeling the Interaction of Open Ended Coaxial Probes and Surface Cracks in Metals, 2002 [Online]. Available: https://scholarsmine.mst.edu/ele_comeng_facwork/588/. (Accessed 26 June 2023).

[11] Y. Gao, L. Gao, X. Li, A hierarchical training-convolutional neural network with feature alignment for steel surface defect recognition, Robot Comput Integr Manuf 81 (Jun. 2023) 102507, https://doi.org/10.1016/J.RCIM.2022.102507.

[12] L. Lu, J. Hou, S. Yuan, X. Yao, Y. Li, J. Zhu, Deep learning-assisted real-time defect detection and closed-loop adjustment for additive manufacturing of continuous fiber-reinforced polymer composites, Robot. Comput. Integrated Manuf. 79 (Feb. 2023) 102431, https://doi.org/10.1016/J.RCIM.2022.102431.

[13] H. Golnabi, A. Asadpour, Design and application of industrial machine vision systems, Robot. Comput. Integrated Manuf. 23 (6) (Dec. 2007) 630–637, https://doi.org/10.1016/J.RCIM.2007.02.005.

[14] X. Qi, G. Chen, Y. Li, X. Cheng, C. Li, Applying neural-network-based machine learning to additive manufacturing: current applications, challenges, and future perspectives, Engineering 5 (4) (Aug. 2019) 721–729, https://doi.org/10.1016/J.ENG.2019.04.012.

[15] R. Tian, M. Jia, DCC-CenterNet: a rapid detection method for steel surface defects, Measurement 187 (Jan. 2022) 110211, https://doi.org/10.1016/J.MEASUREMENT.2021.110211.

[16] T. Wang, Y. Chen, M. Qiao, H. S-T I J of, undefined, A fast and robust convolutional neural network-based defect detection model in product quality control, Springer

94 (9–12) (2018) 3465–3471, https://doi.org/10.1007/s00170-017-0882-0. Feb. 2018.
[17] M. Marei, S. El Zaatari, W. L.-R. and Computer-Integrated, and undefined 2021, 'Transfer learning enabled convolutional neural networks for estimating health state of cutting tools', Elsevier, Accessed: June. 27, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0736584521000302.
[18] Z. Pan, S.L.H. Lau, X. Yang, N. Guo, X. Wang, Automatic pavement crack segmentation using a generative adversarial network (GAN)-based convolutional neural network, Results Eng. 19 (Sep. 2023) 101267, https://doi.org/10.1016/J.RINENG.2023.101267.
[19] T.P. Nguyen, S. Choi, S.J. Park, S.H. Park, J. Yoon, Inspecting method for defective casting products with convolutional neural network (CNN), Int. J. Precision Eng. Manufactur. - Green Technol. 8 (2) (Mar. 2021) 583–594, https://doi.org/10.1007/S40684-020-00197-4.
[20] X. Yang, Y. Zhang, W. Lv, D. Wang, Image recognition of wind turbine blade damage based on a deep learning model with transfer learning and an ensemble learning classifier, Renew. Energy 163 (Jan. 2021) 386–397, https://doi.org/10.1016/J.RENENE.2020.08.125.
[21] Y. Gao, L. Gao, X. Li, X. Yan, A Semi-supervised convolutional neural network-based method for steel surface defect recognition, Robot Comput Integr Manuf 61 (Feb. 2020) 101825, https://doi.org/10.1016/J.RCIM.2019.101825.
[22] M. Cheng, C. Xu, J. Wang, W. Zhang, Y. Zhou, J. Zhang, 'MicroCrack-Net: A deep neural network with outline profile-guided feature augmentation and attention-based multiscale fusion for MicroCrack detection of tantalum capacitors', IEEE Trans Aerosp Electron Syst 58 (6) (Dec. 2022) 5141–5152, https://doi.org/10.1109/TAES.2022.3181117.
[23] Y. Lu, et al., Deep learning-based models for porosity measurement in thermal barrier coating images, Int. J. Multimed. Data Eng. Manag. 11 (3) (Oct. 2020) 20–35, https://doi.org/10.4018/IJMDEM.2020070102.
[24] C. Zhang, et al., Evaluation of internal cracks in turbine blade thermal barrier coating using enhanced multi-scale faster R-CNN model, Appl. Sci. 12 (13) (Jun. 2022) 6446, https://doi.org/10.3390/APP12136446, 2022, Vol. 12, Page 6446.
[25] X. Shan, et al., Automatic recognition of microstructures of air-plasma-sprayed thermal barrier coatings using a deep convolutional neural network, Coatings 13 (1) (Dec. 2022) 29, https://doi.org/10.3390/COATINGS13010029, 2023, Vol. 13, Page 29.
[26] K. Li, T. Li, M. Ma, D. Wang, W. Deng, H. Lu, Laser cladding state recognition and crack defect diagnosis by acoustic emission signal and neural network, Opt. Laser Technol. 142 (Oct. 2021) 107161, https://doi.org/10.1016/J.OPTLASTEC.2021.107161.
[27] H. Zhao, Y. Lv, J. Sha, R. Peng, Z. Chen, G. Wang, Research on detection method of coating defects based on machine vision, in: 2021 IEEE International Conference on Artificial Intelligence and Computer Applications, ICAICA 2021, Jun. 2021, pp. 519–524, https://doi.org/10.1109/ICAICA52286.2021.9498238.
[28] C.W. Chan, X. Chang, M.A. Bozorgzadeh, G.C. Smith, S. Lee, A single parameter approach to enhance the microstructural and mechanical properties of beta Ti-Nb alloy via open-air fiber laser nitriding, Surf. Coat. Technol. 383 (Feb. 2020) 125269, https://doi.org/10.1016/J.SURFCOAT.2019.125269.
[29] Y. Yu, S. Bai, S. Wang, A. Hu, Ultra-short pulsed laser manufacturing and surface processing of microdevices, Engineering 4 (6) (Dec. 2018) 779–786, https://doi.org/10.1016/J.ENG.2018.10.004.
[30] P. Schaaf, Laser nitriding of metals, Prog. Mater. Sci. 47 (1) (Jan. 2002) 1–161, https://doi.org/10.1016/S0079-6425(00)00003-7.
[31] A.M. Kamat, S.M. Copley, A.E. Segall, J.A. Todd, Laser-sustained plasma (lsp) nitriding of titanium: a review, Coatings 9 (5) (Apr. 2019) 283, https://doi.org/10.3390/COATINGS9050283, 2019, Vol. 9, *Page 283*.
[32] K. Holmberg, et al., Residual stresses in TiN, DLC and MoS2 coated surfaces with regard to their tribological fracture behaviour, Wear 267 (12) (Dec. 2009) 2142–2156, https://doi.org/10.1016/J.WEAR.2009.01.004.
[33] H.A. Shirazi, C.W. Chan, S. Lee, Elastic-plastic properties of titanium and its alloys modified by fibre laser surface nitriding for orthopaedic implant applications, J. Mech. Behav. Biomed. Mater. 124 (Dec. 2021) 104802, https://doi.org/10.1016/J.JMBBM.2021.104802.
[34] R. Bammidi, D. Sreeramulu, H. Madivada, P.K. Rejeti, M. Venkatesh, Towards an understanding of Ti-6Al-4V machining and machinability, Mater Today Proc (2023), https://doi.org/10.1016/J.MATPR.2023.09.094.
[35] A.N. Aufa, M.Z. Hassan, Z. Ismail, N. Harun, J. Ren, M.F. Sadali, Surface enhancement of Ti–6Al–4V fabricated by selective laser melting on bone-like apatite formation, J. Mater. Res. Technol. 19 (Jul. 2022) 4018–4030, https://doi.org/10.1016/J.JMRT.2022.06.135.
[36] G. Miranda, et al., Design of Ti6Al4V-HA composites produced by hot pressing for biomedical applications, Mater. Des. 108 (Oct. 2016) 488–493, https://doi.org/10.1016/J.MATDES.2016.07.023.
[37] R. Jones, R.K.S. Raman, A.P. Iliopoulos, J.G. Michopoulos, N. Phan, D. Peng, Additively manufactured Ti-6Al-4V replacement parts for military aircraft, Int. J. Fatig. 124 (Jul. 2019) 227–235, https://doi.org/10.1016/J.IJFATIGUE.2019.02.041.
[38] C.W. Chan, S. Lee, G.C. Smith, C. Donaghy, Fibre laser nitriding of titanium and its alloy in open atmosphere for orthopaedic implant applications: investigations on surface quality, microstructure and tribological properties, Surf. Coat. Technol. 309 (Jan. 2017) 628–640, https://doi.org/10.1016/J.SURFCOAT.2016.12.036.
[39] T. Kattenborn, J. Leitloff, F. Schiefer, S. Hinz, Review on convolutional neural networks (CNN) in vegetation remote sensing, ISPRS J. Photogrammetry Remote Sens. 173 (Mar. 2021) 24–49, https://doi.org/10.1016/J.ISPRSJPRS.2020.12.010.
[40] J. Deng, W. Dong, R. Socher, L.-J. Li, Kai Li, Li Fei-Fei, ImageNet: A Large-Scale Hierarchical Image Database, 2010, pp. 248–255, https://doi.org/10.1109/CVPR.2009.5206848.
[41] K. Simonyan, A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition, in: 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings, 2014 [Online]. Available: https://arxiv.org/abs/1409.1556v6. (Accessed 3 October 2023).
[42] C. Szegedy, et al., Going Deeper with Convolutions, 2015, pp. 1–9.
[43] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.
[44] G. Huang, Z. Liu, L. Van Der Maaten, K.Q. Weinberger, Densely connected convolutional networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 4700–4708.
[45] S.A. Singh, A.S. Kumar, K.A. Desai, Comparative assessment of common pre-trained CNNs for vision-based surface defect detection of machined components, Expert Syst. Appl. 218 (May 2023) 119623, https://doi.org/10.1016/J.ESWA.2023.119623.
[46] H.A. Shirazi, C.W. Chan, S. Lee, Elastic-plastic properties of titanium and its alloys modified by fibre laser surface nitriding for orthopaedic implant applications, J. Mech. Behav. Biomed. Mater. 124 (Dec. 2021) 104802, https://doi.org/10.1016/J.JMBBM.2021.104802.
[47] C.W. Chan, X. Chang, M.A. Bozorgzadeh, G.C. Smith, S. Lee, A single parameter approach to enhance the microstructural and mechanical properties of beta Ti-Nb alloy via open-air fiber laser nitriding, Surf. Coat. Technol. 383 (Feb. 2020) 125269, https://doi.org/10.1016/J.SURFCOAT.2019.125269.