

Research Article

Deep Learning Technology for Weld Defects Classification Based on Transfer Learning and Activation Features

Chiraz Ajmi ^{1,2,3} Juan Zapata,² Sabra Elferchichi,^{3,4} Abderrahmen Zaafouri,¹ and Kaouther Laabidi⁴

¹University of Tunis, National Superior Engineering of Tunis, Street Taha Hussein, Tunis, Tunisia

²Universidad Politécnica de Cartagena, Departamento Tecnologías de la Información y las Comunicaciones, Campus la Muralla, Edif. Antigones 30202, Cartagena, Murcia, Spain

³University Tunis El Manar, National Engineering School of Tunis, Analysis Design and Control of Systems Laboratory (LR11ES20), Tunis 1002, Tunisia

⁴University of Jeddah, CEN Department, Jeddah, Saudi Arabia

Correspondence should be addressed to Chiraz Ajmi; chiraz.ajmi@hotmail.fr

Received 13 March 2020; Revised 4 July 2020; Accepted 24 July 2020; Published 14 August 2020

Academic Editor: Marco Cannas

Copyright © 2020 Chiraz Ajmi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Weld defects detection using X-ray images is an effective method of nondestructive testing. Conventionally, this work is based on qualified human experts, although it requires their personal intervention for the extraction and classification of heterogeneity. Many approaches have been done using machine learning (ML) and image processing tools to solve those tasks. Although the detection and classification have been enhanced with regard to the problems of low contrast and poor quality, their result is still unsatisfying. Unlike the previous research based on ML, this paper proposes a novel classification method based on deep learning network. In this work, an original approach based on the use of the pretrained network AlexNet architecture aims at the classification of the shortcomings of welds and the increase of the correct recognition in our dataset. Transfer learning is used as methodology with the pretrained AlexNet model. For deep learning applications, a large amount of X-ray images is required, but there are few datasets of pipeline welding defects. For this, we have enhanced our dataset focusing on two types of defects and augmented using data augmentation (random image transformations over data such as translation and reflection). Finally, a fine-tuning technique is applied to classify the welding images and is compared to the deep convolutional activation features (DCFA) and several pretrained DCNN models, namely, VGG-16, VGG-19, ResNet50, ResNet101, and GoogLeNet. The main objective of this work is to explore the capacity of AlexNet and different pretrained architecture with transfer learning for the classification of X-ray images. The accuracy achieved with our model is thoroughly presented. The experimental results obtained on the weld dataset with our proposed model are validated using GDxray database. The results obtained also in the validation test set are compared to the others offered by DCNN models, which show a best performance in less time. This can be seen as evidence of the strength of our proposed classification model.

1. Introduction

During the construction of water pipes, internal and external welding works will be carried out for fixing the metal parts. Because of the imperfection of junctions, different types of welding defects such as cracks or porosities can be observed by the human expert, which could cause a limited lifetime of pipelines. Therefore, a quality control is required in order to ensure a good quality of weld. The verification of the

pipelines should be done without destruction of the component. Traditionally, this type of control is performed through ultrasonic techniques. Currently, online nondestructive testing (NDT) is being tested using industrial vision techniques. It is a testing and analysis technique used by industries to evaluate that the exigency characteristics of a material, structure, or system are fulfilled without damaging the original part. Within the framework of this subject, welds can be tested with NDT techniques such as radiography

using radiation that passes through a test tube to detect faults. X-rays are used for thin materials, while gamma rays are used for thicker items. The results can be scanned using film radiography, computer-assisted radiography, computed tomography.

Currently, these films are digitized to be processed on a digital computer. Since the digitized images are characterized with poor quality and low contrast, weld defect inspection could become a challenging task. Unluckily, these qualities effect the interpretation and the classification. Therefore, digital computer vision and machine learning are invented to help the expert in judging the results. This classification is a process of categorization, where objects are recognized and understood. Certainly, the classification of images will be in the coming years because it is a well-known field of computer vision.

In the same frame of reference, computer-aided diagnosis method is a very important research approach that replaces the human expert in the classification of the weld defects for a neutral, nonsubjective, and less expensive way. Generally, this kind of proposal is based on three steps: A first preprocessing step, followed by a second step based on segmentation and features extraction, and a last step where the classification is obtained. Classically, the feature extraction step is usually done by human experts, and therefore it is a time-consuming process and is not accurate. In fact, this problem of accuracy is due, firstly, to the variety of welding defect classes [1], as is shown in Figure 1 and, secondly, to the generated features that are inadequate to achieve a good quality of recognition in order to detect the defect effectively. Following the conventional scheme of the machine learning (ML) application and engaging in features extraction and ML algorithms are well investigated for many researches. However, the simplicity of implementation of some artificial neural network (ANN) algorithms gets paid with the classification time and the bad accuracy. But now deep learning in particular is used to leverage a huge dataset and to make an automatic feature extraction step that will be effective for the reconnaissance of the flaws.

In the literature, there is research on deep learning models for classification of welding, especially convolutional neural network (CNN) models [2]. We are led to work with this new trend, which is the industrial vision. Nevertheless, the use of specific deep learning architecture for welding control to match the command of classification with particular data remains a challenge and can be examined from different angles. To overcome such restriction of previous works and to invest in our dataset composed with the major two defects types lack of penetration and porosities, a pretrained network is applied based on the known technology transfer learning [3]. Our goal is then to preprocess a suitable network model to fit with our case and allow classification of new dataset with respectable accuracy. So, we applied one of the first popular pretrained networks "AlexNet." On one hand, in this work, we enhanced the quality of our few original images and enlarged them with cropping and rescaling by keeping the same aspect ratio to be sure not to destroy the images. This is for generating new images also with different numbers of defects in the same

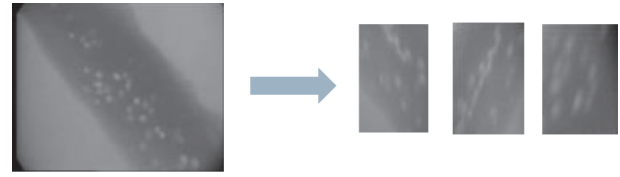


FIGURE 1: From left to right: original and constructed welding radiographic images of porosities defects; original and constructed welding radiographic images of lack of penetration defects.

image and finally applying data augmentation (Section 4.3) to our data. On the other hand, there are no available pretrained networks previously trained on such data, so we trained our radiographic images of pipeline welding with the model and modified hyperparameters to fit with our case. This task is getting better also with transfer learning technology, considering available resources and short time. In this paper, an AlexNet classification network for welding images is proposed using five convolutional layers (Section 3.2).

In the following section, some researches are detailed on welding detection and classification starting by the traditional methods and getting to the new models based on deep learning; next, in the first section of the proposed method in Section 3, the welding detection problems particularly in the dataset are detailed and the factors affecting the convolutional neural network's performance are examined and learned in detail. In Section 3.2, the structure of the network is described; after that, in Section 4, the detailed parameters are described and DCFA method is detailed. In Section 5, the application result is cited, and the comparison term with the Deep Activation Features Network and various CNN models is deduced. In the end, Section 9 shows conclusions and future work. So, the application of deep network which was previously trained with extracted activated features on a huge, fixed amount of classification tasks of the dataset is evaluated and compared with transfer learning. The efficacy of depending on different levels of the network is defined on a fixed feature and gives new results which have been exceeded on various significant challenge views. The experimental results on the images show that the introduced methods have efficient accuracy of classification and transfer learning improves current computer-aided diagnosis methods while providing higher accuracy.

2. Related Work

Detection of industrial X-ray weld images defects is an important research field in nondestructive testing (NDT) [4]. Generally, this kind of proposal is based on three steps: a first preprocessing step, followed by a second step based on segmentation and features extraction, and a last step where the classification is obtained. Therefore, numerous works were done with this purpose. Tong et al. [5] applied a weld defect segmentation method based on morphological and thresholding aspect. A system of defects detection [6] was provided for classifying the digitized welding images on the basis of traditional processes: segmentation, feature

extraction, and classification. These previous approaches are standard features extraction procedures. In [7], an artificial neural network (ANN) is implemented for classification based on the geographical features of the welding heterogeneity dataset. Another work [8] compared an ANN for linear and nonlinear classification. In addition, Kumar and al. [9] described a defect classification method based on texture features extraction to train a neural network, where an accuracy over 86% is obtained.

The weld defects inspection based on approaches of this kind is still semiautomatic and can be influenced by several factors because of the need for expertise, especially for the segmentation and features extraction steps. But now deep learning is used with an automatic feature extraction step for the reconnaissance of the flaws. In addition, deep models have been verified to be more precise for many types of objects detection and classification [10]. Convolutional neural network (CNN) (as can be seen in Section 3.2) is a known deep learning algorithm. Firstly, it enables the object detection process by reducing design effort related to feature extraction. Secondly, it achieves remarkable performance in difficult tasks of visual recognition and equivalent or better performance and accuracy compared to a human being [11], such as classification, detection, and tracking of objects. In [10], three CNN architectures are applied and compared for two different datasets: CIFAR-10 and MNIST dataset. The most acceptable result by comparing the three architectures is the one that was formed with a CNN in CIFAR-10 dataset. The accuracy score reached was over 38% but these networks had limitations related to the image quality, scene complexity, and computational cost. Furthermore, the authors in [12] applied a model of features adaptation to recognize the casting defects. The accuracy score was low for the detection. In [13], a new deep network model is suggested to improve the patch structure by distinguishing the input of the convolutional layers in CNN. The test error results of the combination of data augmentation and dropout technique give the highest score. The application of multilayers and its combinations has resulted in some limitations, especially in comparing layers and reducing their performance.

CNN model is used for detection and classification of welds. In [14], an automatic defect classification process is designed to classify Tungsten Inert Gas (TIG) images generated from a spectrum camera. The work obtained a performance of 93.4% in accuracy. Further research is underway for the detection of defect pipelines welding. For example, in [15], a deep network model for classification is proposed, although this work does not allow the classification of different types of weld defects. Leo et al. [16] trained a VGG16 pretrained network for handed cut regions of weld defect rather than the entire image. A similar topic is proposed in [17]. Wang et al. proposed a pretrained network based on RetinaNet deep learning procedure to detect various types in a radiographic dataset of welding defects. Another fundamental computer vision section is named object detection. The main difference between this part and classification is the design of a bounding box around the interest object to situate it within the image. According to each case of detection, there are a specific number of objects

in each image. So, these objects can be situated in various spatial locations in the image. Thus, the selection of different regions of interest is applied to classify the presence of an object within these regions based on CNN network. This approach can lead to the blow-up, especially if there are many regions. Therefore, object detection methods such as R-CNN, Fast R-CNN, Faster R-CNN, and YOLO [47–51] are examined to fix these imperfections. Various developed object detection methods have been applied based on region convolutional neural network (R-CNN) architecture [47], including Fast R-CNN that mutually optimizes bounding box regression chores [48], Faster R-CNN that adds the regions proposal network for regions search and selection [49], and YOLO which performs object detection through a fixed-grid regression [50]. All of them present varying improvement degrees in detection performance compared to the main R-CNN and make precise, real-time object detection more feasible. The author in [48] introduced a multitasking loss in the bounding box regression and classification and proposed a new CNN architecture called Fast R-CNN. So, generally the image is processed with convolutional layers to produce feature maps. Thus, a fixed length feature vector is extracted from each region proposal with a region of interest (RoI) group layer. Every characteristic vector is then introduced into a sequence of FC layers before finally branching out into two sibling outputs. An exit layer is responsible for producing softmax probabilities for categories and the other output layer encodes the refined positions of the bounding box with four real numbers. To solve the shortage of previous R-CNN model, Ren et al. presented an additional Region Proposal Network (RPN) [49], which acts in a nearly cost-free way by sharing full-image convolutional features with detection network. The RPN is carried out with a totally convolutional network, which has the ability to predict the limits and scores of objects in each simultaneous position. Redmon et al. [50] proposed a new framework called YOLO, which uses the entire map of higher entities to predict both the confidences for several categories and the bounding boxes. Another research topic based on Single Shot MultiBox Detector (SSD) [51], which was inspired by the anchors, adopted RPN [49] and multiscale representation. Instead of the fixed grids adopted in YOLO, the SSD takes advantage of a set of default anchor boxes with different aspect ratios and scales to discretize the output space of the bounding boxes. To manage objects of different sizes, the network combines the predictions of various feature cards with different resolutions. These approaches share many similarities, but the latter is designed to prioritize the speed of evaluation and precision. A comparison of the different object detection networks is provided in [52]. Indeed, several promising instructions are considered as guidance for future work of object detection on welding defects based on neural network-based learning systems.

A research on CNN's basic principles related to tests series has made it possible to deduce the utility of this technique (transfer learning), to recognize the steps taken to adapt to a network, and to progress its integration in the welding fields. Our goal is the classification of defects in

X-ray images; in a previous work [18], we followed pre-processing, extraction of ROI, and segmentation steps. Now, we can get away from these last steps and go directly to the classification with pretrained CNN to classify the defects.

3. Proposed Method for Weld Defect Classification

3.1. Problem Statement. The acquisition system that generates the X-ray images is our original data set which is illustrated Figure 1. It is characterized by poor quality, uneven illumination, and low contrast. In addition, our dataset is small but with images with large-size dimensions between 640×480 and 720×576 pixels and each image has single weld defect or multiple weld defects with different dimension and quantity. Seeing that, the detection of weld defects is a complex, arduous mission. The classification task often used feature extraction methods that have been proven to be effective for different object recognition tasks. Indeed, deep learning reduces this phase by automating the learning and extracting the features through the network's specific architecture. Major problems limiting the use of deep learning methods are the availability of computing power and training data (see Section 4). To train an end-to-end convolution network on such hardware of ordinary consumer laptop and with the dataset's size would be enormously time-consuming. In this work, we had access to a high graphics processor applied for search goal. Moreover, convolution networks need a wide quantity of medium-sized training data. Since the collection and recording of a sufficiently large dataset require hard work, all the works in this topic focused on ready dataset. This is a problem because we have few available datasets of radiographic pipeline welding images which were previously pretrained and our data were characterised with very bad quality and different type not like the public dataset from GDXray dataset [19]. For the same reasons, we tried to generate more data from the original data we have of welding images presented in Figure 1. To be sure about not destroying the image, we cropped and rescaled it while keeping the same aspect ratio (Section 5.1). Furthermore, overfitting is a major issue caused by the small size of dataset; we will therefore apply other efficient methods to prevent this problem by adding training labelling examples with data augmentation and dropout algorithm explained in Section 4.3. The general steps of our proposed method are detailed, illustrating the adjustment process in the following figure (Figure 2).

3.2. Overview of CNN Structure and Detailed Network Architecture. Linear and nonlinear processes have been implicated with convolutional neural model that is a set of overlapping layers. They are acquired in common [2]. The CNN head structure blocks are constituted by convolutional layer, cluster layer, and Rectified Linear Units (ReLU) layer linked to a fully connected layer and a loss layer bottom. There are prominent results for many applications as the visual tasks [20]; also all fields of weld defects detection [15–18, 30] are well investigated. AlexNet [21] is a developed

CNN [20] for ImageNet 2012 and the challenge of large-scale visual recognition (ILSVRC-2012). Its architecture is as follows: max pooling layer is fulfilled in the fifth and the two convolutional layers. For each convolution, a nonlinear ReLU layer is stacked and the normalization is piled for both the first and the second ones. In addition, softmax layer and the “cross entropy” loss function are stacked after two completely fully connected layers. More than 1.2 million images in 1000 classes are trained with this network. In this part, AlexNet architecture is presented. According to the following Figure 3, the convolutional layers are presented with blue color; the max pooling layers are structured with the green one and the white layers represent the normalization. The fully connected layer is introduced as the final rectangle in the upper right corner of the summary structure. The output of this mentioned network is one-dimensional vector as a probability function with the number of elements to be classified which corresponds to two classes in this case. It indicates the extent to which the input images are trained and classified for each type of class.

AlexNet is one of pretrained CNN-based deep learning systems with a specific architecture. Five convolutional layers constitute the network with core sizes of 11×11 , 5×5 , 3×3 , 3×3 , and 3×3 pixels: Conv1, Conv2, Conv3, Conv4, and Conv5, respectively. Taking into account the specificity of weld faults such as the dimensions of images in the dataset, the resolution related to the first convolutional layer is 227×227 . It is stated to have 96 cores with a stride of 4 pixels and size of 11×11 . A number of 256 kernels with a stride of 1 pixel and size of size 5×5 are stacked in the second convolutional layer and filtered from the pooling output of the first convolutional layer. The output of the previous layer is connected to the remainder of convolutional layers with a stride of 1 pixel for each convolutional layer with 384, 384, and 256 kernels of size 3×3 and without pooling grouping. The following layer is piled to 4096 neurons for each of the fully connected layers and a max pooling layer [22] (Section 4.2). After all, the last fully connected layer's output is powered by softmax, which generates two class labels as shown in Figure 3. In this architecture, a max pooling layer is piled with 32 pixels' size and the stride of 2 pixels only after the two beginning and the fifth convolutional layers. The application of the activation function ‘ReLU nonlinearity’ for each fully connected layer improves the speed of convergence compared to sigmoid and Tanh activation functions. The full network requirement and the principal parameters of the CNN design are presented in Table 1 and detailed in the third section.

3.3. Pretraining and Fine-Tuning Learning. Transfer learning computer vision technology [3] is a common way because it permits us to configure precise models quickly. A considerable labelled data amount with great power of processing is advised for novel model of learning. So, we avert starting from scratch by taking benefit of preceding training. With transfer learning, when the process is with various issues, we proceed with performed forms to solve a different problem. It is usually expressed using a big reference dataset of a

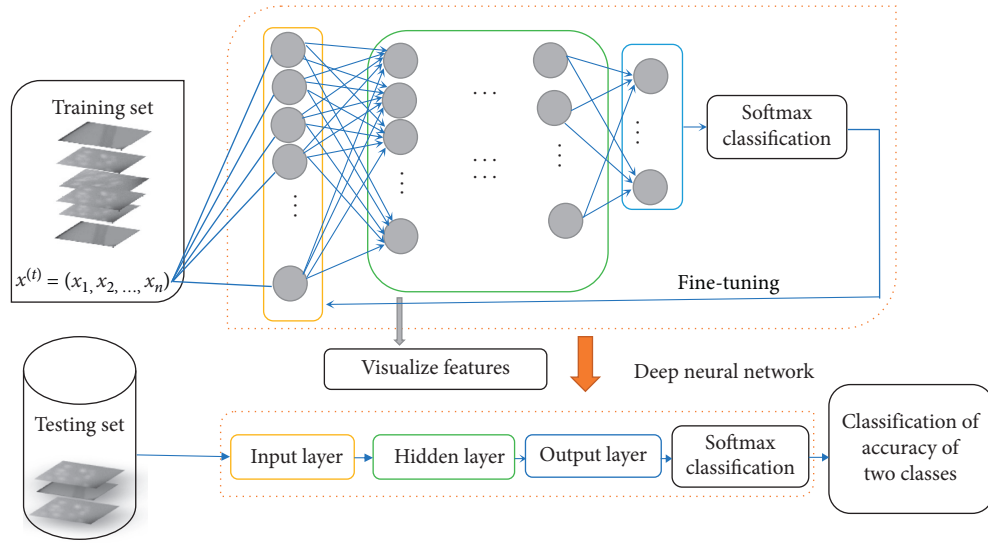


FIGURE 2: The proposed method's flow chart.

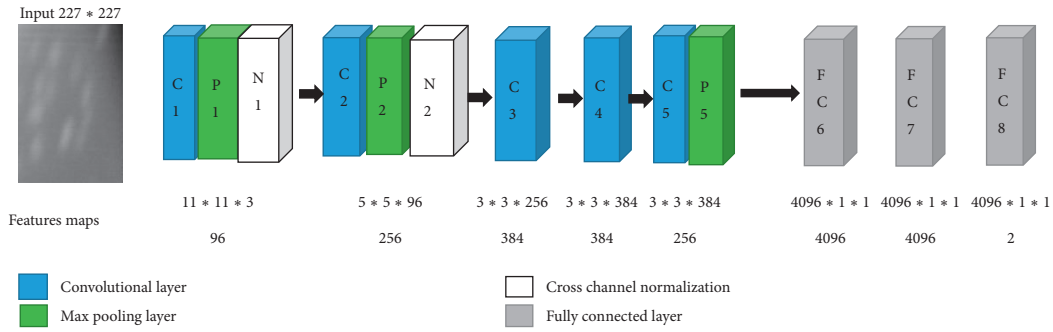


FIGURE 3: The detailed pretrained network scheme.

TABLE 1: The principal parameters of the CNN design.

1	Data	Image input	$227 \times 227 \times 3$ images with zero center normalization
2	Conv1	Convolution	96 $11 \times 11 \times 3$ convolutions with stride [4, 4] and padding [0, 0, 0, 0]
3	Relu1	ReLU	ReLU
4	Norm1	Cross channel normalization	Cross channel normalization with 5 channels per element
5	Pool1	Max pooling	3×3 max pooling with stride [2, 2] and padding [0, 0, 0, 0]
6	Conv2	Convolution	256 $5 \times 5 \times 48$ 4 convolutions with stride [1, 1] and padding [2, 2, 2, 2]
7	Relu2	ReLU	ReLU
8	Norm2	Cross channel normalization	Cross channel normalization with 5 channels per element
9	Pool2	Max pooling	3×3 max pooling with stride [2, 2] and padding [0, 0, 0, 0]
10	Conv3	Convolution	348 $3 \times 3 \times 256$ convolutions with stride [1, 1] and padding [1, 1, 1, 1]
11	Relu3	ReLU	ReLU
12	Con4	Convolution	348 $3 \times 3 \times 192$ convolutions with stride [1, 1] and padding [1, 1, 1, 1]
13	Relu4	ReLU	ReLU
14	Con5	Convolution	256 $3 \times 3 \times 192$ convolutions with stride [1, 1] and padding [1, 1, 1, 1]
15	Relu5	ReLU	ReLU
16	Pool5	Max pooling	3×3 max pooling with stride [2, 2] and padding [0, 0, 0, 0]
17	FC6	Fully connected	4096 fully connected layers
18	Relu6	ReLU	ReLU
19	Drop6	Dropout	50% dropout
20	FC7	Fully connected	4096 fully connected layers
21	Relu7	ReLU	ReLU
22	Drop7	Dropout	50% dropout
23	FC8	Fully connected	2 fully connected layers
24	Prob	Softmax	Softmax
25	Output	Classification output	Cross entropy

pretrained model to resolve the task like the studied case. Many studies have shown the feasibility and efficiency of dealing with new tasks using the preformed model with various types of datasets [23]. They reported how the functionality of this layer can be transferred from one application to another. So, the amelioration of the generalization is accomplished by initialing transferred features of any layer of network after fine-tuning to new application. In our case, the fine-tuning [24] of the corresponding welding dataset is done with the weights of AlexNet. All layers are initialized only the last layer which correspond to the number of labels categories of the weld defects data. The loss function is calculated by label's class computed for our new learning task. A small learning rate of 10-3 is applied to improve the weight of convolutional layers. As with the fully connected layers, the weightings are randomly initialized and the learning rate for these layers is the same. In addition, for updating the network weights of our dataset, we have used stochastic gradient descent (SGD) algorithm with a mini batch size of 16 examples and momentum of 0.9. The network is trained with approximately 80 iterations and 10 epochs, which took 1 min on GPU GeForce RTX 2080.

4. Parameters Setting and DCFA Description

4.1. Normalization. The application of normalization in our CNN helps to get some sort of inhibition scheme. The cross channel local response normalization layer performs channel normalization and in general comes after ReLU activation layer. This process replaces each element with a controlled value and this is accomplished by selecting the elements of certain neighboring channels. So, the training network estimates with following equation a normalized value x' for each component:

$$x' = \frac{x}{(K + (\alpha * \text{ss}/\text{window ChameSize}))^{\beta}} \quad (1)$$

Note that ss is the elements squares sum in the normalization window and K , α , and β are the hyper-parameters in the normalization. In our case, after the first and the second blocks of convolution, two cross channel normalizations with five channels per element are implemented.

4.2. Overlap Pooling. For CNN networks, similarly in map of the kernel, the neuron neighboring groups outputs are re-capped in the grouping layers [22]. In the network, the reduction of the parameters number is done gradually by decreasing the representation spatial size. On each feature map, pooling layer performs separately. To be more accurate, the assembly layer can be considered as a connection of pooling units spaced by s pixels; each one summarizes a neighborhood of size $k \times k$ which is centered on the pooling unit position. In addition, the set of $s = k$ is employed for traditional local pooling and if $s < ik$ we obtain the overlapping pooling. For this work, we put $k = 3$ and $s = 2$. The greatest shared method used in pooling is max pooling [3]. This method takes the defined grid maximum value and

intervenes in the downsampling of the width and height of the image without changing the depth. After the application of ReLU, the nonnegative numbers are presented only from the resulting matrix. Indeed, overfitting is avoided and the dimensions are reduced by max pooling. Anyways, in CNN, the similarity of the number of outputs from the fully connected neurons layer and the final convolution layer is the result of the adjustment of the max pooling layer. In our case, the applied pooling layer is between neighbouring windows with 3×3 sizes and stride of 2.

4.3. Dropout and Data Augmentation. The objective is to learn many parameters without knowingly growing the metrics and extensive overfitting. There are two major ways to prevent that and they are detailed as follows. Dropout layers showed a particular implementation to contract with the task of CNN regularization. According to many studies, dropout technique prevents from overfitting even if it increases twofold the needed iterations number to converge. Moreover, the "dropout" neurons do not contribute to direct transmission and in back-propagation. Specific nodes are removed at every learning's step [25], with a probability p or with a probability $(1 - p)$. As is mentioned in Table 1, for this work, the dropout is applied with a ratio probability = 0.5 in the first two fully connected layers.

For deep learning training, the required dataset must be huge. The simplest and most common form to widen the dataset artificially is by using label-preserving transformations. There are many different methods to data size increasing using data augmentation [26]. According to our network architecture, there are three kinds of data augmentation which permit a very small calculation for the creation of the new images from our original images: translations, horizontal reflections, and the intensities of the RGB channels modification. In this work, these transformations are employed to get more training examples with wide coverage. In addition, our dataset is constructed with radiographic images in gray level; thus, we used the color processing of the gray to RGB which is explained more in the next section of experimental results.

4.4. Activation Features Network. Deep Convolutional network with Activation Function (DCFA) is used to check the efficiency of the structure applied before; we have tried the modern method available for general comparison on the same set of data. We tried to use a pretrained CNN as an experienced copywriter. In fact, the convolutional neural network (CNN) is a great method for automated learning from the field of deep learning. Indeed, they are trained using large collections of diverse images to pick up the rich features' representations of each image. These entity representations often exceed hand-created features such as HOG, LBP, or SURF [27, 28]. One of the simple ways to leverage CNN's power without investing time and effort in training is to use CNN, which has already been tested as a feature extractor. In this comparative example, our images of the weld dataset are classified using a multilayer linear SVM that guides the CNN features extracted from the images. This

methodology to classify the image category surveys the usual training off-the-regular using features extracted from images. An explicated schema of the whole procedure is shown in Figure 4 as follows.

5. Experimental Results

5.1. Dataset Processing and Training Method. The images that create the dataset were obtained from an X-ray image acquisition system. X-ray films can be digitized by various systems such as flat panel detectors or a digital image capture device. Our model is based on X-ray source radiations of total power Y.Smart 160E 0.4/1.5; the X-ray object which is a spiral welded steel pipe of different thickness and diameter and the image intensifier (II) detects the x-rays beams generated by the X-ray source, filters them after penetrating the exposed object up to 20 mm, amplifies them, and converts them into visible light using a phosphor screen. A digital video recorder is used to capture the images from the image intensifier analog camera (with 720p AHD) with 720 horizontal resolution and provides an image with 1280×720 pixels and converts it into a digital format. In this work, this digital device was used to digitize the signal generated from the Spa Maghreb Tubes 1. This resolution was adopted for the possibility of detecting and measuring faults of hundredths of a millimeter, which, in practical terms of radiographic inspection, is much greater than the usual cases. The quality of the video generation is controlled by the environment conditions which can be related to the object's thickness, the exposure environment, the motion of the pipe, and the ADC (analog-to-digital converter) conversion circuits that affect the images' quality and the use of image enhancement filters is needed. According to the statistics on the welding defects dataset, the defect images variable resolution is ordered between 640×480 and 720×576 pixels. So, we implemented an algorithm with MATLAB for cropping and rescaling the original image to new images; after that, we applied preprocessing methods of noise removal with Wiener and Gaussian filter and contrast enhancement with stretching as is mentioned in a previous work [18]. An example of generated images is presented in Figure 5. After that, the generated images are resized by maintaining the same aspect ratio to a stable determination or a specific size of 250×200 . Deep learning deals with specific input size and mostly not big size. To match the necessities of a continuous input dimension of the classification system, we describe welding defects to the extreme level and decrease the complexity at the same time. Although, from a strictly aesthetic point of view, the aspect ratio of the image when resizing an image should be maintained, most neural networks and convolutional neural networks applied to the task of image classification assume a fixed size input which means that the dimensions of all images that must be passed through the network must be the same. Common choices for width and height image sizes used as input in convolutional neural networks include 32×32 , 64×64 , 224×224 , 227×227 , 256×256 , and 229×229 . So, mostly the aspect ratio here is 1 because the width and the height are equal. Before proceeding with

the training, the images had to be resized to the size of the model, which is 227×227 , without keeping the image aspect of the width to the height. Finally, the neural network model is trained by the database that includes only two folders of 695 X-ray welding images. Each folder represents a class of defect. Each class has a number of images that differ from the rest of the folders. Therefore, the lack of penetration group has 242 images and the porosity group has 453 images, which are very known defect types. The images are split for training set and testing set with the same size. So, we have selected 80%–20% for training and testing, respectively. So, in general, we have 139 testing images and 556 training images. This neural network is implemented with MATLAB R2020a GeForce RTX 2080.

5.2. Data Augmentation Implementation and Layers Results.

In deep learning investigations, enormous amount of data is needed to avoid many grave problems of unnecessary learning. Below diverse uses, changing the image geometrically is applied to raise this amount of data. In our case, three types of data augmentation are used: image translations, horizontal reflections, and altering the intensities of the RGB channels. The first technique of data generation is to create image translations and horizontal reflections. To do this, we extract 227×227 random patches (and their horizontal reflections) from 256×256 images and form our network on these extracted patches. The subsequent training examples are highly reliant due to the size rise of the training, which is defined by the factor 4096. If we will not apply this process for dataset, the network will suffer significant overfitting. For the testing time, prediction is made by taking five patches of 227×227 as well as their horizontal reflections (and thus ten patches in total). The second one is changing our images from gray level channel to RGB channels for training and testing images to ensure good performance specifically for our type of images because the pretrained network was trained on colors images of ImageNet. A training image can generate translated ones as seen in Figure 6, an example of lack of penetration defect image, and the augmented images which are most similar to each other. Anyways, if the original dataset is used, the difference between the learning and validation errors indicates the occurrence of an overfitting problem. The learned form is not able to capture the large variation within the small learning dataset. Nevertheless, increasing the dataset size can ameliorate the performance of training. As mentioned earlier, translating the dataset has also been possible to solve the problem of partial replacement by reducing the gap between learning and validation errors. Indeed, the performance of the pretrained AlexNet model after being fine-tuned for 10 epochs, with mini batch size of 16, when training using the augmented dataset outdid the training without data augmentation. So, we conclude that the rescale and translation can be of benefit in the process of training and in the accuracy results. In addition, the performance of the model has been significantly improved.

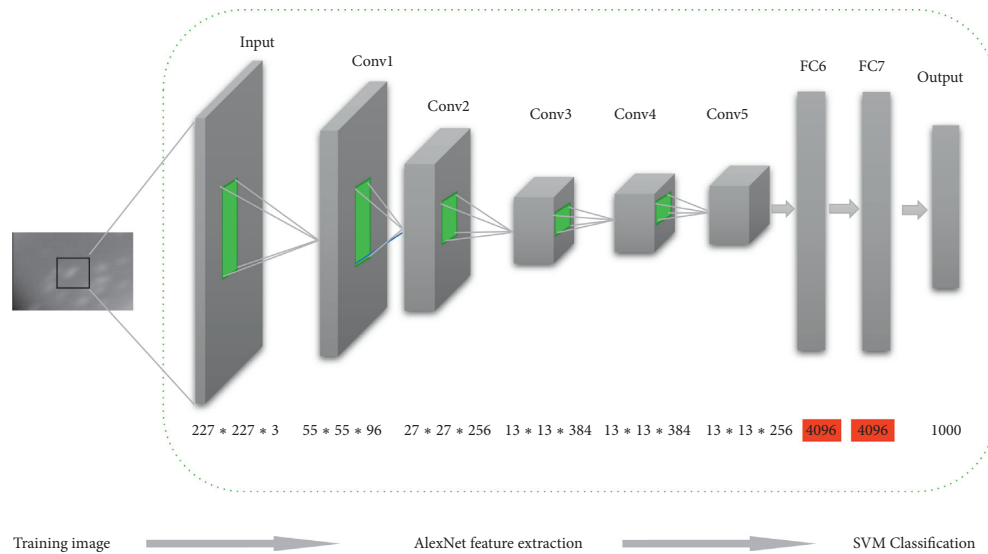


FIGURE 4: The flow chart of the DCFA.

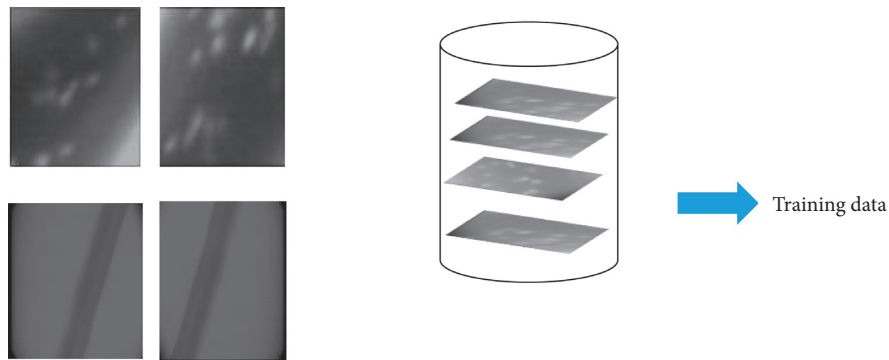


FIGURE 5: Generated images after the crops and resizing.

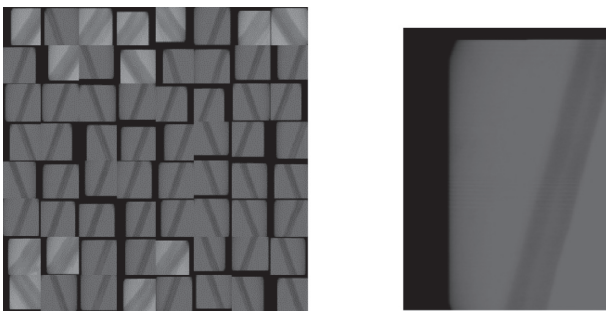


FIGURE 6: Data augmented for lack of penetration defect.

5.3. Layers' Fine-Tuning Experimental Results. For classification means, transfer learning methods use new deep grids to leverage information from the preliminary test delivered by a pretrained network to apply it for new patterns in new data. Usually, the training of data from scratch is slower than using transfer learning and fine-tuning method. Therefore, using these types of networks allows us to pick up new works without configuring a new network and with a great graphics processor. We evaluated the network by training on the welding images. For test, 20% of the two categories were

chosen as test and validation dataset means: 48 for lack of penetration and 91 for porosities. The images were employed with fixed resolution of 227×227 as the input of the network that would convolve and pool the activation repeatedly and then forward the results into the fully connected layers and classify the data stream into 2 categories. To prevent decreasing error caused by the low amount of data, the initial learning rate (LR) base is adjusted as 0.001. Furthermore, the softmax output layer FC8 is characterized by 2 categories and the hidden layers FC6 and FC7 are piled with 4096 neurons. Inaccuracy of predictions in classification is presented by the loss function (LF) which measures the optimal strategy. The system is performing well according to the smallest value of LF. As can be seen in Table 2, after 80 iterations, the loss curve tends to zero, while the classification accuracy curve tends to 1, meeting the requirements of the optimization objectives. The validation classification accurately reaches as high as 0.65 when the iteration is 1, while it increases to 1 when the iteration is 80. The effect of fine-tuning layers of the network, according to Table 2, is shown in the form of AlexNet6-8 where the network will be fine tuned from layer 6 to layer 8 while the previous layers are kept constant with no update. In addition, we stated that the

TABLE 2: Training results for the first epoch.

Epoch	Iteration	Time elapsed, hh:mm:ss	Validation accuracy (%)	Validation loss	Base learning rate
1	1	00:00:01	65.47	1.0146	0.0010
1	3	00:00:15	66.19	0.6706	0.0010
1	6	00:00:16	87.77	0.3102	0.0010
1	9	00:00:16	91.37	0.3085	0.0010
1	12	00:00:17	92.09	0.2529	0.0010

generation of dataset can give a high-quality image by enhancing the quality and eliminating the portion of periodic noise in it [18], which is one of the main supporting successes of the proposed algorithm. Last, it is required to evaluate trained network. For each of the images in testing set, it must be presented to the network and ask it to predict what it thinks about the label of the image. The predictions of the model for an image in the testing set must be evaluated. Finally, these model predictions are compared to the ground-truth labels from testing set. The ground-truth labels represent what the image category actually is. In our work, two categories are applied for experiments and the results in the testing dataset are compared with the ground-truth labels. Their mean classification accuracy was taken as the final results in Figure 7. From there, the predictions of our classifier were correct according to the Confusion Matrix reports used to quantify the performance of network as a whole.

5.4. Final Results and Discussion. We can conclude from Figure 8 that the learning and validation curves are significantly improved and the network tends to converge from the second epoch and the cross entropy loss tends to zero with the increase of the epochs. The accuracy curves of train set and validation set are shown in the same figure. Each point of the precision curve means the correct prediction rate for the set of train or validation images. The accuracy curve adopts the same smooth processing as loss curve. We can see that the accuracy of train set tends to 100% after 2 epochs as well as the validation set accuracy. The test accuracy of 139 weld defect testing images is 100%. Additionally, we can note that the results of testing are mentioned in Table 3 of the Confusion Matrix for the validation data. The performance measurement is with four different combinations of predicted and target classes which are the true positive, false positive, false negative, and the true negative. In this format, the number and percentage of the correct classifications performed by the trained network are indicated in the diagonal. For example, 48 of defects are correctly classified as lack of penetration (LP). This corresponds to 34.5% of all 139 testing welding images. In the same way, 91 defects are correctly classified as porosities. This matches 65.5% of all testing welding images. So, all the classes of defects are correctly classified and the overall accuracy is that 100% of predictions are correct and 0% are incorrect.

As seen in Figure 9, a negative impact on system performance results by fine-tuning the fourth convolutional layer because the information gets very rudimentary about data, such as edges and points of layers, and we learn specific

structures of the dataset. It is noted that fine-tuning the earlier layers end to end with the fully connected layers is less good than only fine-tuning of the fully connected layers. However, according to Table 4 of the Confusion Matrix, for the validation data, the fine-tuning of the network has been decreased dramatically to less than 23% compared to our model's accuracy result. The decrease in performance during training of our network is very low compared to that of the deep convolutional features activation system. This can be seen as evidence of the strength of the AlexNet model for classification task. Given that decrease, 77% validation accuracy for the DCFA system has been recovered. Obviously, 38 of defects are correctly classified as lack of penetration (LP). This corresponds to 27.3% of all 139 testing welding images. In the same way, 69 defects are correctly classified as porosities. This matches 49.6% of all testing welding images. The results show that an extensive and deep neural network is able to deliver unprecedented results on a very challenging dataset using supervised learning. It should be noted that the performance of our network is deteriorating if only one convolutional layer is removed. Depth also is really important to achieve our results. We did not use any prior unsupervised training [29] because of simplicity, although we expected it to be useful, especially if we have sufficient computing power to considerably increase the size of the network without growing the amount of labelled data. The time used for processing each patch is 0.01 s, which is also faster than DCFA method mentioned before. This has proved that the deep transfer learning is suitable for detecting weld defects in X-ray images. In the end, we would like to use very large and deep convolutional networks on video streams whose timeline provides very useful visible information or huge dataset of images to provide good results.

6. Performance Comparison of Transfer Learning-Based Pretrained DCNN Models with the Proposed Model for Our Dataset

6.1. Methodology of Transfer Learning in DNNs Models. The idea of transfer learning is related more to its efficiency in implementing DCNN [32, 33] trained on big datasets and to "transferring" their situational capabilities in training to newer image categorization than DCNN formation from scratch [45]. With correct fit, pretrained DCNN succeeded in performance trained DCNN from scratch. Actually, this mechanism using any pretrained model will be described shortly in Figure 10. The well-known pretrained DCNN

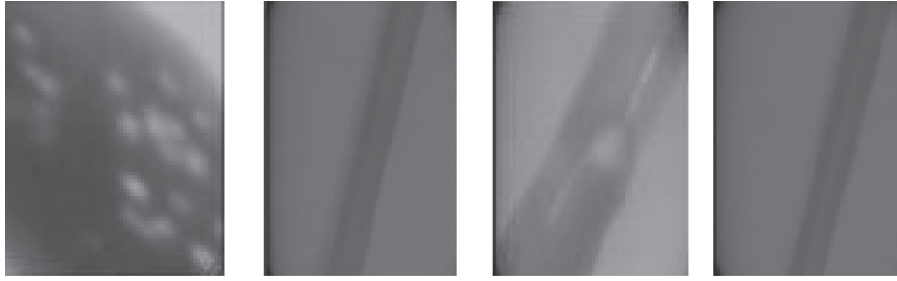


FIGURE 7: Predicted images (from left to right): well-classified porosity defects image, well-classified lack of penetration defect image, well-classified lack of penetration defect image, and well-classified lack of penetration defect image.

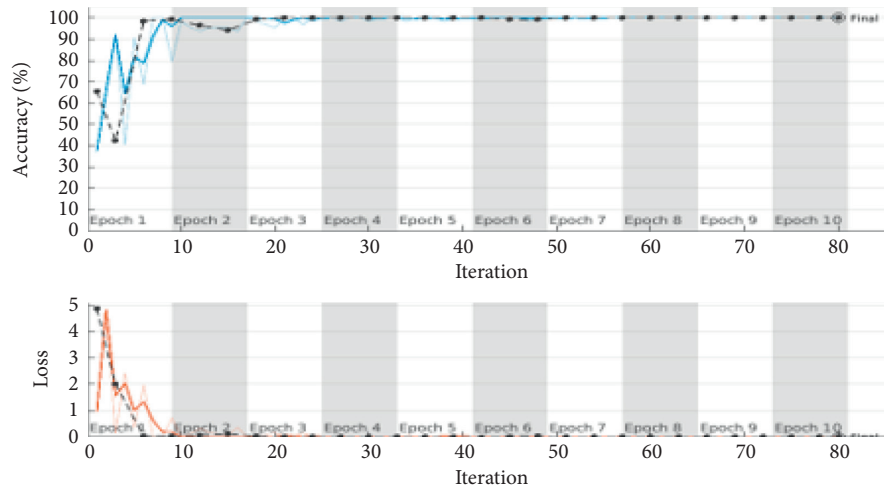


FIGURE 8: Training process curves: training accuracy and loss curves are represented by a straight line; validation accuracy and loss curves are represented by a dashed line.

TABLE 3: Confusion Matrix for validation data.

Output Class	LP	48	0	100%
		34.5%	0.0%	0.0%
	Porosities	0	91	100%
		0.0%	65.5%	0.0%
		100%	100%	100%
		0.0%	0.0%	0.0%
	LP			
	Porosities			
	Target class			

models like GoogLeNet [34], ResNet50 [35], ResNet101 [35], VGG-16, and VGG-19 [31] have been used for efficient classification. The organization shown in Figure 10 is composed of layers from the pretrained model and few new layers. In this work, for all the models, only the last three layers have been replaced to suit the new image categories. The modification of each pretrained network is detailed as follows:

- (i) For VGG-16 and VGG-19, only the last three layers of the pretrained network with a set of layers are modified (fully connected layer, softmax layer, and classification output layer) to classify images into respective classes.

- (ii) For GoogLeNet, also the network's last three layers are changed. These later layers loss3-classifier, prob, and output are altered by a fully connected layer, softmax layer, and a classification output layer. Afterwards, the last transferred layer remaining on the network (pool5_drop 7×7 _s1) is linked by the new layers.
- (iii) For ResNet50, the last three layers fc1000, fc1000_softmax, and ClassificationLayer_fc1000 of the network are substituted by fully connected layer, a softmax layer, and a classification output layer. Then, the last remaining transferred layer on the network (avg_pool) is connected to the novel layers.
- (iv) For ResNet101, predictions of the network's last three layers which are fc1000, prob, and ClassificationLayer_ are altered with a fully connected layer, a softmax layer, and a classification output layer, respectively. Finally, the new layers and the last transferred layer remaining in the network (pool5) are connected.
- (v) Train the network.
- (vi) Test the new model on the testing dataset.



FIGURE 9: Results of application of DCFA (from left to right): porosities defects image classified as follows: lack of penetration defect image, well-classified lack of penetration defect image, well-classified lack of penetration defect image, and well-classified porosities defects image.

TABLE 4: Confusion Matrix of DCFA for validation data.

		38	22	63.3%
	LP	27.3%	15.8%	36.7%
		10	69	87.3%
	Porosities	7.2%	49.6%	12.7%
Output Class		79.2%	75.8%	77.0%
		20.8%	24.2%	23.0%
		LP	Porosities	
		Target class		

6.2. *Comparison of the Proposed Model with the DCNNs Procedures: Results and Discussion.* The transferred DCNN models are investigated in this work to compare the results and to prove the efficiency of our proposed model. According to the details mentioned in 6.1, we designed each model in MATLAB R2020a and we changed the layers of each one to fit with our needs. The transferred models were trained using stochastic gradient descent with momentum (SGDM). The mini batch size was taken as 16, the maximum number of epochs was 10, and the learning rate was 0.001 and in general the other parameters were taken to be the same as the proposed structure and the details of each network are replaced as described in subsection 6.1. The ranking performance of each CNN is summarized in Table 5 which illustrates performance comparison for all pretrained models using all chosen performance metrics. The tabular results show that the pretrained AlexNet model reaches the best results followed by VGG-16 and VGG-19. The performance measures presented in Table 5 indicate that the DCNNs models VGG-16, VGG-19, and GoogLeNet proved to be stellar by attaining 95%, 97.8%, and 99.3% accuracy. Moreover, ResNet50 and ResNet101 have accomplished 100% but with more computation time compared to our proposed model. The training time is considered as a comparative metric for the efficiency of each model. In addition, AlexNet model reaches the highest accuracy level in the lowest time opposed to other transferred DCNN models. For the challenging X-ray dataset, the models with transfer learning yielded specific significant results thanks to their implementation using a strong hardware with GPU capability (NVIDIA RTX 2080). So, the processing time is mostly not high. The training progress curves and the Confusion Matrix of the best model which is AlexNet are presented in Table 3 and Figure 8. From the training

progress curve, it is apparent that the AlexNet model attains the highest level of accuracy for the dataset in only 10 epochs. Also, it is able to make correctly classified samples from datasets with a limited false positive rate and a superior true fraction value. The value of performance metrics shows the advantage of transfer learning in minimizing overfitting and increasing the convergence speed.

7. Comparative Performance Using the GDXray Images

The GDXray (Grima X-ray) database covers five categories of radiographic images: welds, castings, settings and natural objects, and luggage. As is known, until now there have been no public databases of digital radiographic images for X-ray testing. We are interested in the welding data which are grouped into 3 series including 88 images. So, these X-ray images were generated from the BAM Federal Institute for Materials Research and Testing, Berlin, Germany, as shown in Figure 11.

Experiments with these data were done in different works [36–39]. Many efforts have been made in the field of weld failure detection, especially using the GDXray database as it is the public and available benchmark. For classification purpose usually, the data is preprocessed as it contains few images with big size, different quality, and various defect types in each image, which is not the same case as that in our images. In addition, as previously mentioned, researches done applied a method to extract the image to different patches or regions to augment the data and after that classify it. This data was used usually for detection aim with any machine learning algorithm as the focused data and it is cropped, resized, and ameliorated and then trained and finally tested. In our case, we will apply this benchmark for validation of our proposed classification model, which means that the data was not trained before with the specificity of the pretrained network. Thereby, at first, the preprocessing is needed to group the data into two defect types: porosities and lack of penetration; secondly, cropping from the original images, defects positions according to the labels class to reduce the identification burden, and finally, resizing the cropped regions to small version according to the input layer of the AlexNet network. Thus, the generated images are resized, keeping the same aspect ratio in a stable determination or a

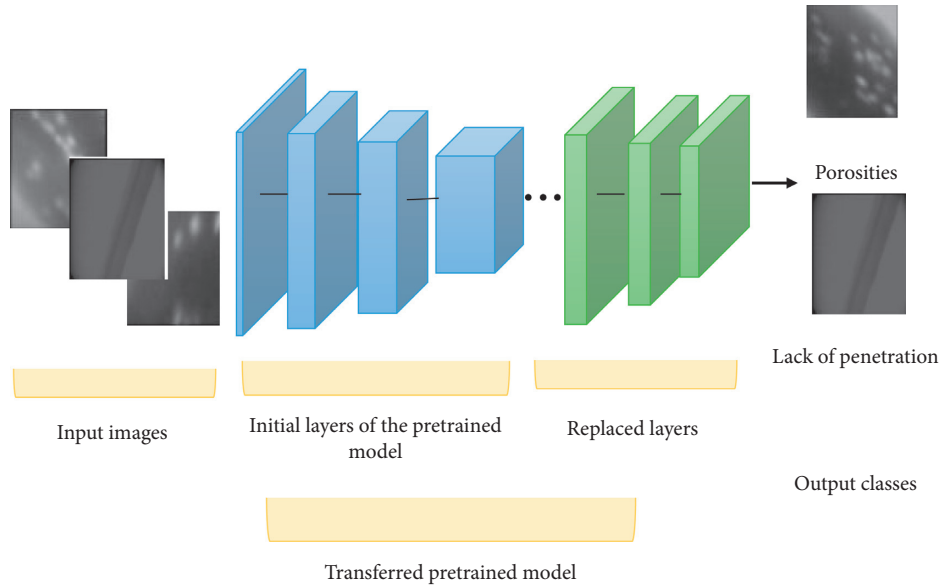


FIGURE 10: Mechanism of transfer learning using pretrained models.

TABLE 5: Performance metrics of DCNN models with processing time.

Pretrained model	Accuracy	Error	Sensitivity	Specificity	Precision	False positive rate	Time
AlexNet	100	0	100	100	100	0	47 s
VGG-16	95	0.05	100	85	92	0.145	3 min 34 s
VGG-19	97.8	0.0216	100	93.75	96.8	0.0625	3 min 55 s
GoogLeNet	99.3	0.0072	98.9	100	100	0	1 min 38 s
ResNet50	100	0	100	100	100	0	3 min 12 s
ResNet101	100	0	100	100	100	0	6 min 52 s

specific size of the network, 227×227 . According to the statistics on this database, the defect images are few with variable resolution ordered between 11668×3812 pixels or less with multiple various dimensions. So, the process takes place in two stages: first cropping each image to patches in the region of defect so that we get 108 images and after that resizing it to the network model. Normally, for the testing time, it is needed to apply data augmentation for changing our images from gray level channel to RGB channels as the pretrained network was trained on colors images of ImageNet to fit with the input data of the model. These images are not large enough to test a deep convolutional neural network. We adopted our proposed network model to expand the dataset with translation and reflections so that we avoid the overfitting problem. In this comparative aspect, two categories are applied for experiments and the results in the testing dataset are compared with the ground-truth labels. Their mean classification accuracy was taken as the final results in Figure 12. From there, the predictions of our classifier are shown in the Confusion Matrix reports used to quantify the network validation performance as a whole. The classification performance of the proposed transferred model is evaluated on GDXray, and the results of testing are mentioned in Table 6 of the validation database Confusion Matrix. Performance measurement consists of four different combinations of predicted and

target classes that are true positive, false positive, false negative, and true negative. In this format, the correct classifications number and percentage reached by the network are presented diagonally. For example, 43 were fully and correctly classified as lack of penetration (LP). This corresponds to 39.8% of the 108 weld test images. Also, 49 faults are correctly classified as porosities, which corresponds to 45.4% of all the weld test images. Thus, the overall accuracy is that 85.2% of the predictions are correct. Moreover, the training on the GDXray images is presented in Table 7, which reached the overall accuracy of 87%. We can conclude that the results are accurate and performant enough according to the images quality and amount and the volatility inside the patches. The advantage of applying each of these models was quantified through ablation tests. Based on the transferred DCNN models, we compare the results and prove the efficiency of our proposed model on GDXray images. The ranking performance of each CNN is summarized in Table 8 which illustrates performance comparison for all pretrained models using all chosen performance metrics on GDXray images. The tabular results show that the pretrained AlexNet model reaches the best results. The performance measures presented in Table 8 indicate that the DCNNs models VGG-16, VGG-19, and GoogLeNet proved to be stellar by attaining 84.3%, 80.6%, and 74.1% accuracy.



FIGURE 11: Some images of group welds series of GDxRay.

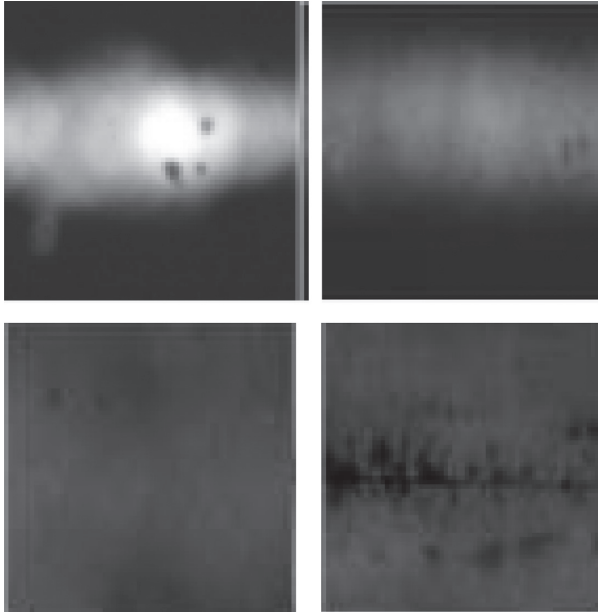


FIGURE 12: Results of validation with GDxRay (from left to right and from top to bottom): well-classified porosities defects image, porosities defects image classified as lack of penetration, well-classified porosities defects image, and well-classified lack of penetration defect image.

8. Comparison of the Transfer Learning-Based AlexNet Model with Studies on Welds Classification

Table 5 shows that the transferred DCNN models AlexNet, ResNet50, ResNet101, VGG-16, and VGG-19 have achieved performant accuracy. Indeed, a comparative analysis in Table 9 is given with the existing works on welds database. An overview of studies using supervised machine learning algorithms for the classification step is presented in Table 9. According to this table, various works are based on ANN (artificial neural networks) [40–42] and fuzzy logic systems [43]. Moreover, support vector machines have been utilized [44]. In [43], the credibility level is considered for failure proposals detected by a fuzzy logic system. The aim of this proposal is to decrease the rate of false alarms (or false positives) instead of increasing the detection accuracy subject to a low false positive rate. The authors in [40] covered the classification related to welding fault that integrates the ANN classifier with three different sets of functionalities. This manuscript is based on preprocessing, segmentation, feature extraction (both texture and geometric), and finally classification in terms of individual and

TABLE 6: Confusion Matrix of validation database GDxRay.

Output Class	LP	43	13	6.8%
		39.8%	12.0%	23.2%
	Porosities	3	49	94.2%
		2.8%	45.4%	5.8%
		<93.5%	79.0%	85.2%
		6.5%	21.0%	14.8%
	LP			
	Target class		Porosities	

TABLE 7: Confusion Matrix of training database GDxRay.

Output Class	LP	40	8	83.3%
		37.0%	7.4%	16.7%
	Porosities	6	54	90.0%
		5.6%	50.0%	10.0%
		87.0%	87.1%	87.0%
		13.0%	12.9%	13.0%
	LP			
	Target class		Porosities	

TABLE 8: Comparison of the performance metrics of pretrained DCNN models.

	Accuracy (%)	TP (%)	TN (%)	FN (%)	FP (%)
Our method	85.2	39.8	54.4	2.8	12.0
VGG-16	84.3	37	47.2	5.6	10.2
VGG-19	80.6	25.9	54.6	16.7	2.2
GoogLeNet	74.1	21.3	52.8	21.3	4.6
ResNet50	68.5	29.6	38.9	13.0	18.5
ResNet101	74	25.0	49.1	17.6	8.3
DCFA	71.3	22.2	49.1	20.4	8.3

combined features. It is necessary to improve the precision of the classification using textures and geometric characteristics based on the ANN classifier. The classification accuracy based on the combination of the features' characteristics led to an acceptable accuracy between 84% and 87%. It should also be noted that [46] investigates the fake defects that are inserted into the radiographs as a training set with positive results. General shape descriptors, which are used to characterize weld failures, have been studied, and an optimized set of shape descriptors has been identified to efficiently classify weld failures using a multiple comparison procedure. However, the rate of successful categorization varies from 46% to 97%. It is important to note that a successful categorization depends largely on the variability of the entry. Most of the studies in Table 9, as well as the machine learning study, have focused on how to correctly detect and classify as many defects as possible based on features extractions and classifiers. The

TABLE 9: Comparison of the transfer learning-based pretrained DCNN models with works.

Study	Method	Probability of classification (%)
(23) (Kaftandjian et al., [43])	Region-based + reference + morphological geometrical fuzzy logic	80% at 0% false alarms
(29) (Valavanis et al., [44])	Graph-based (region growing + connectivity similarities) Geometrical, texture SVM, ANN, k-NN	46%–85%
(28) (Zapata, vilar, Ruiz, [42])	Region edge growing; geometrical + PCA multiple-layer ANN, adaptive network-based fuzzy inference system (ANFIS)	79%–83%
(26) (Kumar, et al. [40])	Edge, region growing, watershed texture, geometrical ANN	87%
(19) (Lim, Ratnam [46])	Thresholding, geometrical, multiple-layer ANN	97%
(27) (Zahran et al., [41])	Morphological + wavelet + region growing; Mel-frequency cepstral coefficients (MFCCs) and power density spectra ANN	75%

comparative analysis has been restricted to the measures reported in the literature, that is, classification accuracy, number of features, and computation time. The measures reported show that transferred DCNN models reached higher level of accuracy compared to the methods devised in previous cited works. The main pretrained DCNN models' benefit in comparison with the studies is that they do not require a feature extraction mechanism or an intermediate feature selection phase. Furthermore, AlexNet model with transfer learning achieves a recognition rate of 100% with 25 learning layers. In addition, the transferred AlexNet model rendered 100% recognition, with short computation time.

9. Conclusions

This paper proposed the transfer learning method based on AlexNet for weld defects classification. The previously learned network architecture has been modified to fit the new classification problem by resizing the output layer and adding two dropout layers to avoid the overfitting problem. The application of the fully convolutional structure needs the same size input for training as the similar processing, even for the testing images. Few datasets of two low quality classes that contain 695 X-ray welding images are applied to study learning performance. In order to ameliorate the network's performance, dataset augmentation is used through translations, horizontal reflections, and modification to RGB channels. It is noted that the implementation result has achieved 100% of classification accuracy. Comparing the results of our network with the DCFA network and different pretrained DCNN models with transfer learning, we note that our AlexNet network achieves the highest performance. As expected, it has been verified that it is recommended to fine-tune the fully connected layers of the AlexNet network instead of doing it for the whole network, specifically for small datasets. The benefits of applying pretrained DCNN models with transfer learning for classification are varied: at first, the classification mechanism is fully automated; besides it eliminates the conventional step of feature extraction and selection and then no inter- and intraobserver biases are there and the predictions done by the pretrained DCNN models are reproducible with a high accuracy level. For classification and defect detection applications, the system

required the generation of a new welding dataset that represented different types of welds with good quality. Now, the defect detection step is really needed to be investigated to recognize the defect and its dimensions. Several algorithms can be implemented for object detection, which belong to the RCNN family. This can be the focus of the upcoming article.

Data Availability

The data used to support the findings of this study are currently under embargo, while the research findings are commercialized. Requests for data will be considered by the corresponding author.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work has been partially funded by the Spanish Government through Project RTI2018-097088-B-C33 (MINECO/FEDER, UE).

References

- [1] N. Boaretto and T. M. Centeno, "Automated detection of welding defects in pipelines from radiographic images DWDI," *NDT & E International*, vol. 86, pp. 7–13, 2017.
- [2] B. T. Bastian, J. N. S. K. Ranjith, and C. V. Jiji, "Visual inspection and characterization of external corrosion in pipelines using deep neural network," *NDT & E International*, vol. 107, Article ID 102134, 2019.
- [3] H.-C. Shin, H. R. Roth, M. Gao et al., "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," 2016, <http://arxiv.org/abs/1602.03409>.
- [4] Z. Fan, L. Fang, W. Shien et al., "Research of lap gap in ber laser lap welding of galvanized steel," *Chinese Journal of Lasers*, vol. 41, no. 10, Article ID 1003011, 2014.
- [5] T. Tong, Y. Cai, and D. Sun, "Defects detection of weld image based on mathematical morphology and thresholding segmentation," in *Proceedings of the 2012 8th International*

- Conference on Wireless Communications, Networking and Mobile Computing*, pp. 1–4, Barcelona, Spain, October 2012.
- [6] D. Mery and M. A. Berti, “Automatic detection of welding defects using texture features,” *Insight - Non-Destructive Testing and Condition Monitoring*, vol. 45, no. 10, pp. 676–681, 2003.
 - [7] J. Hassan, A. M. Awan, and A. Jalil, “Welding defect detection and classification using geometric features,” in *Proceedings of the 2012 10th International Conference on Frontiers of Information Technology*, IEEE, Washington, DC, USA, pp. 139–144, 2012.
 - [8] R. R. da Silva, L. P. Calôba, M. H. S. Siqueira, and J. M. A. Rebello, “Pattern recognition of weld defects detected by radiographic test,” *NDT & E International*, vol. 37, no. 6, pp. 461–470, 2004.
 - [9] J. Kumar, R. Anand, and S. Srivastava, “Multi-class welding flaws classification using texture feature for radiographic images,” in *Proceedings of the 2014 International Conference on Advances in Electrical Engineering (ICAEE)*, pp. 1–4, Vellore, India, 2014.
 - [10] U. Boryczka and M. Ba lchanowski, “Differential evolution in a recommendation system based on collaborative filtering,” in *Proceedings of the International Conference on Computational Collective Intelligence*, pp. 113–122, Springer, Halkidiki, Greece, 2016.
 - [11] “Image recognition,” 2018, <http://www.tensorflow.org/tutorials/imagerecognition>.
 - [12] Y. Yongwei, D. Liuqing, Z. Cuilan, and Z. Jianheng, “Automatic localization method of small casting defect based on deep learning feature,” *Chinese Journal of Scientific Instrument*, vol. 6, p. 21, 2016.
 - [13] M. Lin, Q. Chen, and S. Yan, “Network in network,” 2013, <http://arxiv.org/abs/1312.4400>.
 - [14] D. Bacioiu, G. Melton, M. Papaalias, and R. Shaw, “Automated defect classification of ss304 TIG welding process using visible spectrum camera and machine learning,” *NDT & E International*, vol. 107, Article ID 102139, 2019.
 - [15] W. Hou, Y. Wei, J. Guo, Y. Jin, and C. Zhu, “Automatic detection of welding defects using deep neural network,” *Journal of Physics: Conference Series*, vol. 933, Article ID 012006, 2018.
 - [16] B. Liu, X. Zhang, Z. Gao, and L. Chen, “Weld defect images classification with vgg16-based neural network,” in *Proceedings of the International Forum on Digital TV and Wireless Multimedia Communications*, pp. 215–223, Shanghai, China, November 2017.
 - [17] Y. Wang, F. Shi, and X. Tong, “A welding defect identification approach in X-ray images based on deep convolutional neural networks,” in *Proceedings of the International Conference on Intelligent Computing*, pp. 53–64, Springer, Tehri, India, 2019.
 - [18] C. Ajmi, S. El Ferchichi, and K. Laabidi, “New procedure for weld defect detection based-gabor filter,” in *Proceedings of the 2018 International Conference on Advanced Systems and Electric Technologies (IC ASET)*, pp. 11–16, Hammamet, Tunisia, March 2018.
 - [19] D. Mery, V. Rifio, U. Zscherpel et al., “GDxray: the database of X-ray images for nondestructive testing,” *Journal of Nondestructive Evaluation*, vol. 34, no. 4, p. 42, 2015.
 - [20] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” *Advances in Neural Information Processing Systems*, pp. 1097–1105, 2012.
 - [21] X. Han, Y. Zhong, L. Cao, and L. Zhang, “Pre-trained AlexNet architecture with pyramid pooling and supervision for high spatial resolution remote sensing image scene classification,” *Remote Sensing*, vol. 9, no. 8, p. 848, 2017.
 - [22] D. Scherer, A. Muller, and S. Behnke, *Evaluation of Pooling Operations in Convolutional Architectures for Object Recognition, International Conference on Artificial Neural Networks*, Springer, Berlin, Germany, 2010.
 - [23] E. T. Moore, W. P. Ford, E. J. Hague, and J. Turk, “An application of CNNs to time sequenced one dimensional data in radiation detection,” in *Algorithms, Technologies, and Applications for Multispectral and Hyperspectral Imagery XXV*, International Society for Optics and Photonics, Bellingham, WA, USA, 2019.
 - [24] S. Pan and Q. Yang, “A survey on transfer learning,” *IEEE Transaction on Knowledge Discovery and Data Engineering*, vol. 22, no. 10, 2010.
 - [25] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: a simple way to prevent neural networks from overfitting,” *The Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958, 2014.
 - [26] A. Miko larczyk and M. Grochowski, “Data augmentation for improving deep learning in image classification problem,” in *Proceedings of the 2018 International Interdisciplinary PhD Workshop (IIPhDW)*, pp. 117–122, Swinoujscie, Poland, May 2018.
 - [27] D. Mery and C. Arteta, “Automatic defect recognition in X-ray testing using computer vision,” in *Proceedings of the 2017 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 1026–1035, Santa Rosa, CA, USA, March 2017.
 - [28] Y. Mu, S. Yan, Y. Liu, T. Huang, and B. Zhou, “Discriminative local binary patterns for human detection in personal album,” in *Proceedings of the 2008 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8, Washington, DC, USA, January 2008.
 - [29] N. C. Chung, B. Mirza, H. Choi et al., “Unsupervised classification of multi-omics data during cardiac remodeling using deep learning,” *Methods*, vol. 166, pp. 66–73, 2019.
 - [30] M. Ben Gharsallah and E. Ben Braiek, “Weld inspection based on radiography image segmentation with level set active contour guided off-center saliency map,” *Advances in Materials Science and Engineering*, 2015.
 - [31] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” 2014, <http://arxiv.org/abs/1409.1556>.
 - [32] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the inception architecture for computer vision,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2818–2826, 2016.
 - [33] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, “Inception-v4, inception-ResNet and the impact of residual connections on learning,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, San Francisco, CA, USA, 2017.
 - [34] C. Szegedy, W. Liu, Y. Jia et al., “Going deeper with convolutions,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–9, Boston, MA, USA, 2015.
 - [35] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778, Las Vegas, NV, USA, 2016.
 - [36] M. Carrasco and D. Mery, “Segmentation of welding defects using a robust algorithm,” *Materials Evaluation*, vol. 62, no. 11, pp. 1142–1147, 2004.
 - [37] D. Mery, “Automated detection of welding defects without segmentation,” *Materials Evaluation*, vol. 69, no. 6, pp. 657–663, 2011.

- [38] D. Mery and M. A. Berti, "Automatic detection of welding defects using texture features," *Insight-Non-Destructive Testing and Condition Monitoring*, vol. 45, no. 10, pp. 676–681, 2003.
- [39] D. Mery, "Automated detection of welding defects without segmentation," *Materials Evaluation*, vol. 69, no. 6, pp. 657–663, 2011.
- [40] J. Kumar, R. S. Anand, and S. P. Srivastava, "Flaws classification using ANN for radiographic weld images," in *Proceedings of the 2014 International Conference on Signal Processing and Integrated Networks (SPIN)*, IEEE, Florence, Italy, pp. 145–150, May 2014.
- [41] O. Zahran, H. Kasban, M. El-Kordy, and F. E. A. El-Samie, "Automatic weld defect identification from radiographic images," *NDT & E International*, vol. 57, pp. 26–35, 2013.
- [42] R. Vilar, J. Zapata, and R. Ruiz, "Classification of welding defects in radiographic images using an adaptive-network-based fuzzy system," in *Proceedings of the International Work-Conference on the Interplay between Natural and Artificial Computation*, pp. 205–214, Springer, La Palma, Spain, 2011.
- [43] V. Kaftandjian, O. Dupuis, D. Babot, and Y. M. Zhu, "Uncertainty modelling using Dempster–Shafer theory for improving detection of weld defects," *Pattern Recognition Letters*, vol. 24, no. 1–3, pp. 547–564, 2003.
- [44] I. Valavanis and D. Kosmopoulos, "Multiclass defect detection and classification in weld radiographic images using geometric and texture features," *Expert Systems with Applications*, vol. 37, no. 12, pp. 7606–7614, 2010.
- [45] Y. Bar, I. Diamant, L. Wolf, S. Lieberman, E. Konen, and H. Greenspan, "Chest pathology detection using deep learning with non-medical training," in *Proceedings of the 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, IEEE, Washington, DC, USA, pp. 294–297, April 2015.
- [46] T. Y. Lim, M. M. Ratnam, and M. A. Khalid, "Automatic classification of weld defects using simulated data and an MLP neural network," *Insight-Non-Destructive Testing and Condition Monitoring*, vol. 49, no. 3, pp. 154–159, 2007.
- [47] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 580–587, Columbus, OH, USA, June 2014.
- [48] R. Girshick, "Fast R-CNN," in *Proceedings of the 2015 International Conference on Computer Vision*, Araucano Park, Las Condes, Chile, December 2015.
- [49] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: towards realtime object detection with region proposal networks," in *Proceedings of the Neural Information Processing Systems*, pp. 91–99, Montreal, Canada, December 2015.
- [50] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: unified, real-time object detection," in *Proceedings of the 2016 Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, June 2016.
- [51] W. Liu, D. Anguelov, D. Erhan et al., "Ssd: Single shot multibox detector," in *Proceedings of the European Conference on Computer Vision*, Amsterdam, The Netherlands, October 2016.
- [52] J. Huang, V. Rathod, C. Sun et al., "Speed/accuracy trade-offs for modern convolutional object detectors," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, July 2017.