IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 28, 2019, accepted July 2, 2019, date of publication July 10, 2019, date of current version July 26, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2927162

A Neural Network Model for Material Degradation Detection and Diagnosis Using Microscopic Images

WOOSUNG CHOI[1], HYUNSUK HUH[2], BAYU ADHI TAMA[2], GYUSANG PARK[1], AND SEUNGCHUL LEE[2,3]

[1] Power Generation Laboratory, KEPCO Research Institute, Daejeon 34056, South Korea
[2]Department of Mechanical Engineering, POSTECH, Pohang 37673, South Korea
[3]Institute for Convergence Research and Education in Advanced Technology, Yonsei University, Seoul 03722, South Korea

Corresponding author: Seungchul Lee (seunglee@postech.ac.kr)

This work was supported in part by the Research and Development project of Korea Electric Power Corporation (KEPCO) under Grant R17GA08, in part by the Technology Innovation Program (smart state monitoring system for maintenance of ship rotating machinery) of the Ministry of Trade, Industry and Energy (MOTIE) under Grant 10080729, and in part by the Korea Institute for Advancement of Technology (KIAT) grant funded by the MOTIE, Korea (The Competency Development Program for Industry Specialist), under Grant N0008691.

ABSTRACT Detection and diagnosis of material degradation are of a complex and challenging task since it is presently hand-operated by a human. Therefore, it leads to misinterpretation and avoids correct classification and diagnosis. In this paper, we develop a computer-assisted detection method of material failure by utilizing a deep learning approach. A deep convolutional neural network (CNN) model, combined with an image processing technique, e.g., adaptive histogram equalization, is trained to classify a real-world turbine tube degradation image data set, which is retrieved from a power generation company. The experimental result demonstrates the effectiveness of the proposed approach with predictive classification accuracy is up to 99.99% in comparison with a shallow machine learning algorithm, e.g., linear SVM. Furthermore, performance evaluation of a deep CNN with and without an above-mentioned image processing technique is exhibited and benchmarked. We successfully demonstrate a novel application in constructing a deep-structure neural network model for material degradation diagnosis, which is not available in the current literature.

INDEX TERMS Material degradation, deep learning, creep damage, convolutional neural network, histogram equalization, boiler tube, high temperature.

I. INTRODUCTION

In an engineering field, degradation is specified as an unexpected condition of particular properties of material which occurs continuously due to gradual exposure during the operating condition. The greatest root causes that significantly contribute to the degradation of engineering material such as exceedingly high temperature, deviation within the chemical composition, mechanical loading, and other environment-related factors [1]. Among these elements, a higher operating temperature from those specified leads to the major changes in the microstructure [2]. The degradation induced by an environmental effect is commonly restricted to the near-surface zone and involves certain processes such as corrosion and its various forms, oxidation, creep, and so forth [3].

Since the changes brought about by degradation include material loss or changes in physical properties, most industries have released a tremendous expense against material degradation. Moreover, it becomes the central problems in particular industries, i.e. marine environments, oil and gas production, and energy conversion and generation systems. The process of analyzing the degradation of concrete roughly composes the following actions: (i) site visit in order to obtain representative failure samples, (ii) samples test, i.e. microscopic examination, and (iii) analysis, interpretation, and diagnosis of failure [4]. However, failure diagnosis requires an intensive prior knowledge of an expert which occasionally leads to misinterpretation as some forms of failure might not be visibly apparent. For instance, surface corrosion might be immediately observed, yet another form of corrosion could be solely verified microscopy.

Moreover, major power plants operate at high temperature operating conditions and are exposed to high temperature

The associate editor coordinating the review of this manuscript and approving it for publication was Madhusanka Liyanage.

environments for a long time, so it is necessary to accurately evaluate and manage the state of high temperature components for a stable operation. Such power plant components, i.e. boiler tubes, piping, and turbine, are exposed to high temperature environments for a long period of time because of the purpose of generating and transporting steam, resulting in deterioration due to continuous material damage, particularly creep damage. Creep damage means permanent deformation at high temperature over time. If high stress occurs for a long period of time under high temperature environment, the strength of the part decreases, which shortens useful life of high temperature components. When creep damage occurs in boiler tubes, piping and turbine, the pores grow in the micro structure to create cracks and increases the size of the carbides.

The gap between the carbides increases and the hardness decreases. At present, micro-structures on the surface of the damaged part are replicated by the film at the power plant site, brought to the laboratory, observed using a microscope, and used by the human-eyes. These evaluation methods depend on individual subjective judgment and experience, and it takes time from field measurement to laboratory analysis and evaluation. Also, in practice, it is not possible to distinguish between classes utilized in the evaluation criteria. Therefore, quantitative, systematic and objective evaluation methods are necessarily important.

With the rapid evolution of information discovery approaches, artificial intelligence and machine learning remain to demonstrate a prominent role in an overabundance of applications ranging from technological and non-technological domain [5]. They have brought a tremendous impact in the purview of engineering by providing a smart way of manufacturing process, for instance [6]. Smart manufacturing is defined as a new manufacturing paradigm in which machines are fully connected via wireless networks, monitored by sensors, and controlled by next-generation artificial intelligent techniques in order to efficiently enhance the whole manufacturing process [6]–[8]. Furthermore, data-driven strategies, i.e. big data analytics are more possible to be implemented thereby enabling companies to increase their competitive advantages [9], [10].

There is presently a lack of literature discussing deep neural network for material degradation detection and diagnosis. Nevertheless, there have been constant studies of employing deep neural network for gearbox fault identification and diagnosis [11], [12]. The papers combine wavelet analysis and deep convolutional network to solve gearbox fault classification problem. More recently, deep learning in fault diagnosis of rotating machinery has been outlined in [13], in which the paper also reviews other machine learning techniques, i.e. k-NN, support vector machine, and naive Bayes. A research presented in [14] proposes stacked autoencoder to extract features from mechanical vibration signals, then uses the softmax regression to characterize the machine health conditions. The stacked autoencoder is also harnessed by [15] to detect fault in rotating machinery. The detection performance of the stacked autoencoder is evaluated over multiple data sets, i.e. four roller bearing and a planetary gearbox data set.

Furthermore, it is worth mentioned that other researchers have applied deep convolutional neural network in a plethora of engineering applications. Work in [16] have developed convolutional neural network model for condition monitoring of rotating machinery. With similar fashion, [17] and [18] exploit convolutional neural network for bearing fault diagnosis, followed by [19] and [20] who take into account convolutional neural network for defect detection and motor fault detection, respectively. Finally, a real time vibration-based structural damage detection is recommended by [21]. The merit of the proposed method includes an automatic extraction of damage features from the input which does not require any extra processing.

In this paper, we utilize deep convolutional neural network for identifying and classifying the material degradation. We take into account a preliminary processing technique of image data set by incorporating an image-processing methods, e.g. histogram equalization. Histogram equalization is a technique, where image intensity/contrast is adjusted in such a way that the intensities could be better distributed on the histogram. It enables a lower contrast to obtain higher contrast, hence the distribution of pixels is not too far from uniform [22]. We prove that an above-mentioned preprocessing technique, combined with deep convolutional neural network yield higher detection performance than a deep convolutional neural network model with no preprocessing technique.

The remaining part of the paper is decomposed into following parts. Section II briefly describes an overview material degradation and its current diagnosis method. Section III presents a brief overview of histogram equalization and deep convolutional neural network. The experimental result is discussed in Section IV, and finally we draw the outline of the paper in Section V.

II. AN OVERVIEW OF MATERIAL DEGRADATION AND ITS ASSESSMENT METHOD

Some major damages in a power plant such as creep, material degradation, fatigue, creep-fatigue, thermal shock, erosion, and various types of corrosion are the principle of deterioration mechanisms in hot-section equipment of a power plant [23]. Creep damage is one of the most critical high temperature deterioration mechanisms. Depending on the characteristic of the material, classification of creep damage in a power plant components, i.e. turbine and boiler has been carried out qualitatively [24]. They characterize the damages based on the distribution of creep voids and microcracks conducted by an extensive observation on steam pipes in Germany power plants, then classify the cavity evolution in steels into four stages, i.e. isolated cavities, oriented cavities, microcracks, and macrocracks. Finally, they devise recommendations in accordance with the four stages of degradation. For class A damage, no action would be necessary. For class B damage, re-inspection within the designated period

is necessary. For class C, repair or replacement would be required. For class D, an immediate repair is mandatory [25].

Assessment of the material degradation relies on the inspection, or an estimation of the accumulation of degradation based on a model for degradation accumulation, or both. According to [23], assessment procedures of creep damage can follow the general practice summarized in the American Petroleum Institute (API) Recommended Practice 579, ''Fitness-for-Service'', which it was firstly published in 2000. These series of steps are as follows: (i) identification of deterioration mechanism, (ii) identification of the applicability of the assessment procedures applicable to the particular deterioration mechanism, (iii) identification of the data requirements, (iv) evaluation of the assessment procedures and techniques, (v) evaluation of the remaining life of material, (vi) remediation if necessary, (vii) monitoring, and (viii) documentation, allowing suitable records for further evaluation.

Hitherto, there have been increased approaches of predicting creep damage and remaining material life assessment, e.g. continuum damage mechanics [26]–[29], in terms of practical use. The growth in cracks in equipment operating at high temperature can be appraised using predictive methods, for instance, by [30]. However, rather than employing a conventional technique using complex human-eye analysis, in this study, we aim at simplifying the method of creep damage evaluation by taking into account an automatic classification using deep learning model while ensuring accuracy and reliability.

III. PREPROCESSING AND CLASSIFICATION TECHNIQUE

A combination of histogram equalization and deep convolutional neural network is proposed for material degradation. In this section, a brief theory of histogram equalization is firstly discussed, followed by the illustration of image classification using deep convolutional neural network.

A. HISTOGRAM EQUALIZATION

Histogram equalization is an effective method in enhancing the appearance of a poor image. It transforms the histogram of the resultant image as flat as possible (the overall shape of histogram remains the same). The fundamental theory for histogram equalization involves probability theory, where the intensity levels in an image might be viewed as a random variable of the gray levels [22], [31]. The resultant image is obtained by mapping each pixel in the input image with intensity ratio r_k into a corresponding pixel with level s_k in the output image using Equation 1 [22].

$$s_k = T(r_k) = (L-1)\sum_{j=0}^{k} p_r(r_j) = \frac{(L-1)}{MN}\sum_{j=0}^{k} n_j \quad k = 0, 1, 2, \ldots, L-1 \tag{1}$$

where MN is the total number of pixels in the image, n_k is the number of pixels that have intensity r_k, and L is the number of possible intensity levels in the image. The mapping $T(r_k)$ in Equation 1 is known as *histogram equalization*.

Generally speaking, the histogram equalization process for digital images can be broken down into four steps: (i) find the running sum of histogram values, (ii) normalize the values from step (i) by dividing by the total number of pixels, (iii) multiply the values from step (ii) by the maximum gray-level value and round, and (iv) map the gray-level values to the results from step (iii) using a one-to-one correspondence [31]. Histogram equalization of an image will not commonly offer a histogram that is faultlessly flat, yet it will make the distribution of gray-level values as evenly as possible.

The preprocessing so far has been applied uniformity to the whole part of the image. However, it is not always useful for a typical image which is a mixture of bright and dark areas. Adaptive Histogram Equalization (AHE) [32] is an image processing technique that divides an image into tiles of small size and partially applies histogram equalization. Using the AHE, it has an excellent effect in enhancing the partial contrast of the image and detecting the contour, rather than using the original histogram equalization. However, there is a drawback that it is sensitive to the noise of the image. Thus, image-processing algorithms called contrast limited adaptive histogram equalization (CLAHE) are used to prevent amplification of noise. CLAHE sets a limit of contrast, and if the value of a particular pixel exceeds that value, it redistributes the value over the other area. Figure 1 illustrates the difference between original histogram equalization and CLAHE.

B. DEEP CONVOLUTIONAL NEURAL NETWORK

Deep learning is made up of a number of processing layers to learn representations of data with numerous levels of abstraction. Unlike conventional machine learning techniques which require a corresponding feature representation from which a classification algorithm could detect or classify patterns, deep learning enables a machine to be inserted with raw data and automatically find out the representations required for detection and classification [33]. In addition, a deep learning technique transforms the representation from one level into a higher abstract level, hence very complex functions are able to be learned. The differences between deep learning and conventional machine learning models can be summarized in Table 1.

Deep convolutional neural network (CNN) is a type of deep learning models oftentimes used in image processing and classification. It comprises hierarchically arranged trainable stages that, besides can learn spatial hierarchies of patterns, it also needs fewer training samples to learn representations that have generalization power. These two properties make a CNN efficient when processing images [34], [35]. A CNN is designed to process data that come in the form of multiple arrays, called *feature maps*, in case of image processing is a 2D array. A typical CNN architecture (Figure 2) is composed of a series of stages containing

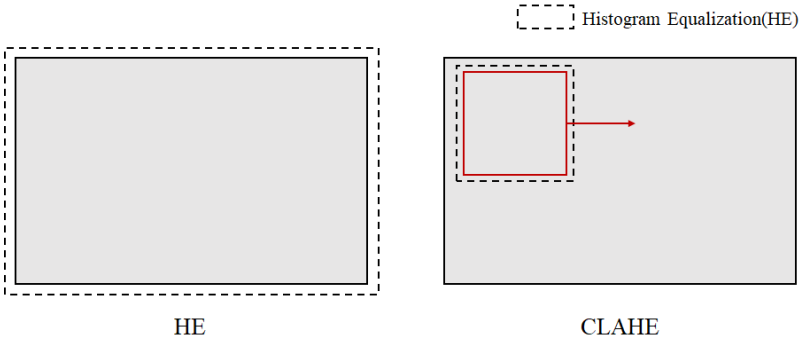

FIGURE 1. The difference between histogram equalization (left) and adaptive histogram equalization (right).

TABLE 1. The differences between deep learning and traditional machine learning models.

	Feature learning	Model building	Model training
Conventional machine learning	Require domain-expert to extract the features	Extracted features are used to build the model	Each training module is trained sequentially
Deep learning	Features are obtained by transforming the data into abstract representation	Hierarchical model structure with non-linear combination of many layers	Training parameters are learned simultaneously

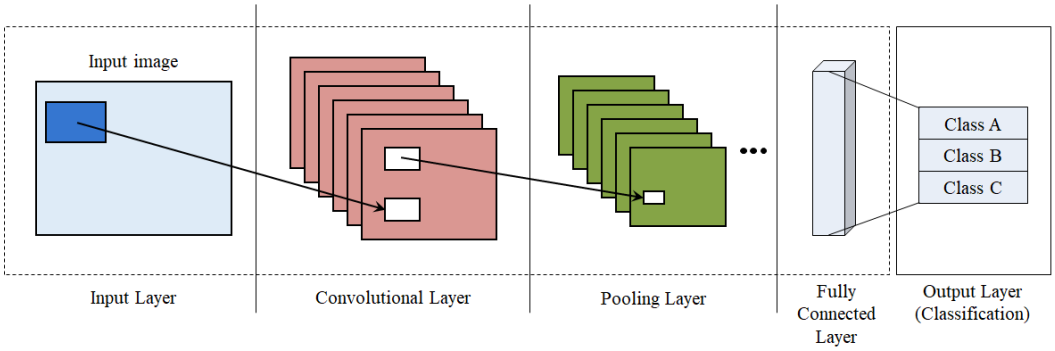

FIGURE 2. A typical architecture of deep convolutional neural network.

two types of layers, i.e. convolutional layers and pooling layers.

Let suppose the input to a convolutional layer is an image of size $a \times b$, and $c \times d$ becomes the size of k feature map of the convolutional layer, which is less than the size of the input image, then the output of the convolutional layer is a set of k feature maps of size $(a-c+1) \times (b-d+1)$ by stepping over one pixel. In order to acquire pixels in convolutional map, we assign a weight k_{ij} to each pixel in the input image, compute the weighted sums, and extract corresponding features in the image. A bias is appended to the weighted sums then send them to a non-linear activation function, i.e. *tanh*() or *sigm*(). The output of activation function y_r^t of a particular feature map r in convolutional layer t is defined as:

$$y_r^t = \xi\left[b_r^t + \sum_{i \in M_r^t} y_i^{t-1} \otimes k_{ir}^t\right] \tag{2}$$

where ξ is non-linear activation function, whilst b_r^t denotes bias for the t^{th} layer. M_r^t and $\otimes$ represent the selected feature maps i in the $(t-1)^{th}$ layer and convolutional operator, respectively.

The operation continues within pooling layer which secures the pixels processed in previous layer are carried forward. The output of the pooling layer leads to a dimensional reduction, with respect to the selected stride. The activation output f_h^r after reduction-sampling the feature map r into a feature map h in a layer t is specified as:

$$f_h^t = \zeta\left(g_r^t, N^t\right) \tag{3}$$

where ζ is the reduction-sampling function, i.e. mean or maximum function that reduces the samples by a factor of N^t and g_r^t is the convoluted feature map to be reduced. Finally, the last layer of the network is characteristically a classification layer. The output of classification layer o can be as follow.

$$o = \psi\left(b_o + Wz\right) \tag{4}$$

TABLE 2. Fault classes representing five different deterioration conditions.

Class	Description	Remarks
A	Normal	No creep damage, e.g. cavitation detected. Some evidence of thermal degradation may be seen.
B	Isolated	Isolated cavities are observed. It is not possible to deduce the direction of maximum principal stress from the damage seen.
C	Oriented	Cavities are observed, usually with multiple cavities on the same boundary. A clear alignment of damages boundaries can be seen, indicating the axis of maximum principal stress.
D	Microcracked	Cavities are observed on boundaries normal to maximum principal stress. Some boundaries have separated due to the inter-linkage of cavities on them to form microcracks. Typically less than 2 mm. Might be detected by a conventional nondestructive evaluation method.
E	Macrocracked	In addition to cavities and microcracks being observed, macrocracks have joined together and widened to form macrocracks many grain boundaries long. Commonly more than 2 mm and should be detected by a conventional nondestructive evaluation method.

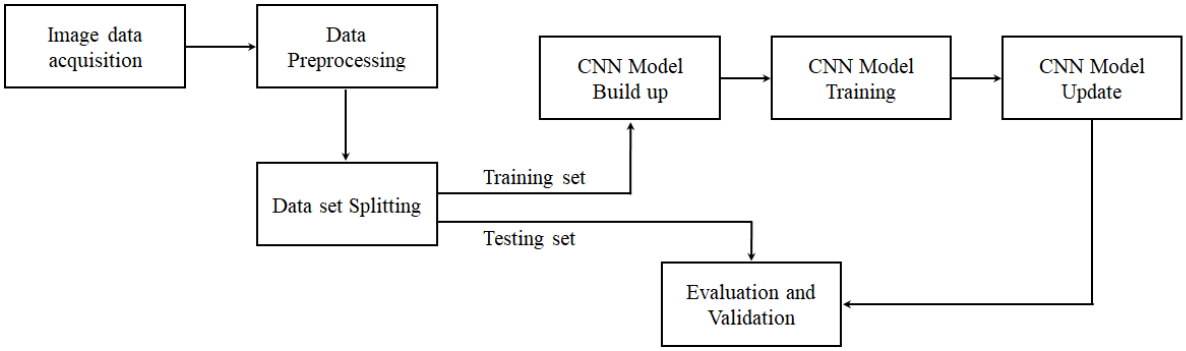

FIGURE 3. Flowchart of the material degradation classification and diagnosis.

where b_o is the bias of the output layer, W is weight matrix between the second-to-last layer and the output layer, and z is the concatenated feature maps of the second-to-last layer.

CNN relies on several learning parameters, i.e. b_r^t, k_{ir}^t, b_o, and W during the training process. A stochastic gradient descent method, i.e. back propagation is utilized at the training stage which aims at minimizing the error between the actual output and the desired output. The error function of the back propagation method is calculated as:

$$E = \frac{1}{2}\sum_{n=1}^{N}\sum_{k=1}^{c}(t_k^n - y_k^n)^2 \tag{5}$$

where N depicts the number of training images, c is the number of output neurons, t_k^n is the k^{th} element of the target output for the n^{th} pattern and y_k^n is the actual output of the network for the n^{th} pattern.

IV. IMPLEMENTATION AND RESULT

This section describes the data used in the experiment, steps of the experiment, architecture of deep convolutional neural network, as well as experimental results and discussions.

A. EXPERIMENTAL DATA

Real-world material degradation images that represent the grade of boiler tubes in power plants are used in the experiment. Data set is collected from a power generation company in Korea. It is made up of a number of actual microscopic ($< 100\mu m$) image samples containing five class label attributes. The images are labeled with five different fault assessments performed by some experts in this field with respect to a work of [24]. Since the industry is engaging with power generation, thus the material degradation detection and diagnosis are very critical in order to maintain the continuity and reliability in supplying power to the end users. The five classes of the faults employed for classification and analysis are outlined in Table 2.

B. EXPERIMENT SETTING

The flowchart presented in Figure 3 depicts the steps performed in the experiment. The experiment is begun with image data collection. The output of this step is raw data set that requires preprocessing methods prior to be used as input of classifier. Images collected through data acquisition are not suitable for determining degradation ratings due to non-uniform reflection of illumination, non-uniformity of the surface, and unclear patterns. We used an image processing technique, called histogram equalization in order to transform the raw image data set in such a way that the intensity of pixels are distributed uniformly. The acquired image is partially cut through the histogram equalization algorithm to sharpen the contrast, and the color of an image is converted into a gray-scale one, so that it is robust against illumination

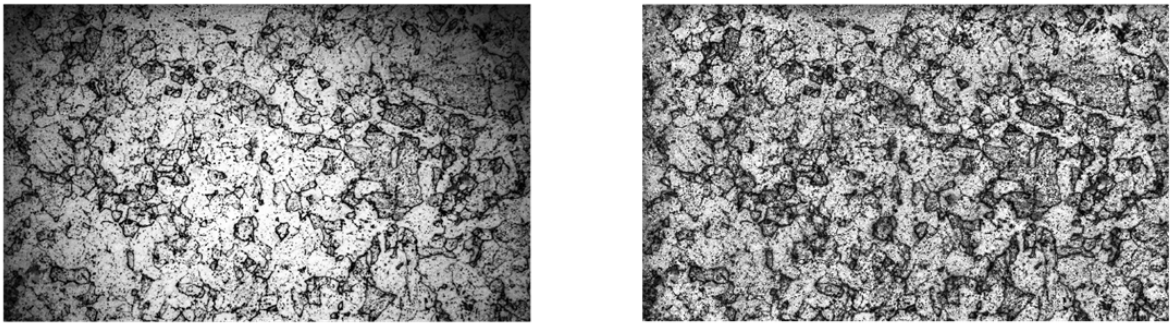

FIGURE 4. The result of applying histogram equalization (left) and CLAHE (right).

and reflection. In addition, the image file is converted to a *pickling* file so that we can load into Python environment.

More precisely, since degraded images used to build our CNN model are not always obtained in the same lighting condition and the number of samples is not too large, histogram equalization (HE) algorithm is applied to generalize the lighting condition of acquired data for more stable learning and reducing an aleatory uncertainty deriving from in-homogeneous characteristics of specimen materials. However, the lighting of the microscopic image tends to be concentrated in the center of the subject, surface of boiler tube, thus we adopt the CLAHE algorithm to solve such different local illumination conditions. The result of applying CLAHE for degradation image is shown in Figure 4.

All degraded images are labelled with A, B, C, D, and E classes. There are in total 90 sample images, where 10 images for class A and 20 images for class B to class E, respectively. In the experiment, 80% (72 images) of them are used as a training set and 20% (18 images) are used as testing set. However, 72 images are not enough to build the deep learning model as a small training data set might lead to *overfitting*. Therefore, we randomly cut, flip and rotate the degraded images in order to raise the number of training samples using only images in original training set. This method is also known as *data augmentation* (see Figure 6). The microstructures analyzed in the sample image have the same degradation characteristics regardless of their location. Thus, data augmentation for each pixel can produce $400 \times 440 \times 8$ images. The pattern of degraded images are evenly distributed throughout the image, so it is possible to learn stably by cropped images. A cropped image (400×400) used for training after applying an image processing technique is shown in Figure 5.

In the training phase, we build an effective CNN architecture for image classification. The structure of the CNN model is described in detail in Section IV-C. We train the model, then save the model for further validation/testing analysis. In the learning process, the *softmax* cross entropy is used to store low loss parameters. In the CNN model update process, the stored parameters are retrieved through learning, and

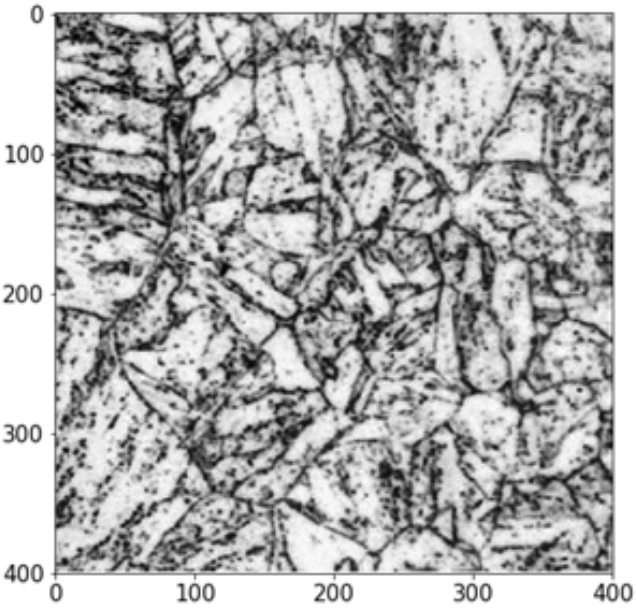

FIGURE 5. A cropped image used as input of CNN.

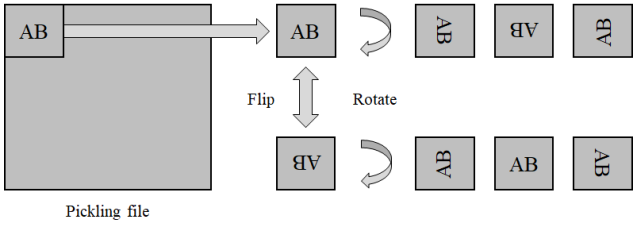

FIGURE 6. Illustration of data augmentation.

parameters of the learned model are obtained and rewritten when necessary.

C. NETWORK ARCHITECTURE

The network architecture used to evaluate the grade of the T91 steel boiler tubes is summarized in Table 3. The first layer following the input layer is two convolutional layers, with 16 feature maps of 3×3 kernel size each. This is followed by

TABLE 3. The proposed CNN structure in our experiment.

Layer	Type	Output shape	Kernel size	#Feature maps	#Trainable parameters	Stride
0	Input	400×400	-	-	0	-
1	2D Convolution	400×400	3×3	16	144	1
2	2D Convolution	400×400	3×3	16	2304	1
3	2D Maxpooling	200×200	2×2	-	0	2
4	2D Convolution	200×200	3×3	32	4608	1
5	2D Convolution	200×200	3×3	32	9216	1
6	2D Maxpooling	100×100	2×2	-	0	2
7	Fully-connected	1	-	-	-	-
				Total	16,272	

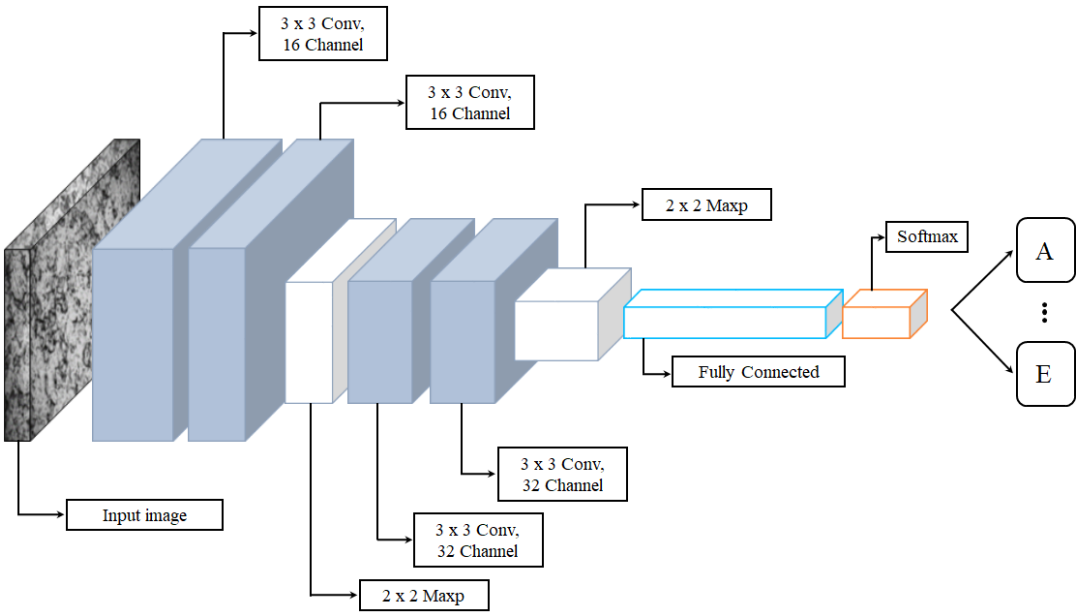

FIGURE 7. The proposed deep learning architecture.

a max-pooling layer of kernel size 2×2. The next layers are two convolutional layers which share the same kernel size of 3×3 and 32 feature maps, followed by a 2×2 max-pooling layer. The output layer is made up of 5 neurons representing five different degradation classes (see Figure 7). We employ *ReLu* as the activation function.

All layers are fully-connected. A gradient descent method is utilized with learning and decay rate is 0.001 and 0.965, respectively. The batch size is set to 50, whilst the network is trained for 30,000 epochs. To optimize our object functions, we used *AdamOptimizer* implemented in *TensorFlow*. The experiment takes several minutes on a GPU-based computer. Randomly chosen 50 cropped images are used for every iterations. The change in error rate for training and testing is visualized in Figure 8. As shown in the figure, the more iteration we perform, the more chance we have of getting small error rate.

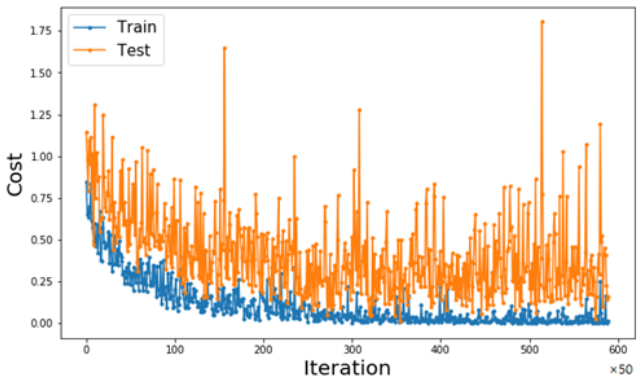

FIGURE 8. Change in error rate for training and testing set.

D. RESULT AND DISCUSSION

In order to verify the CNN model that we have built in the training, supplied test set is used for validation analysis. We take into account the performance accuracy as a metric. The accuracy rate is calculated as the ratio of the number correctly classified validation samples to the total number of validation samples. Our CNN model predicts the classes of the cropped images with an accuracy of 99.80%. However, because the actual data is a full-size image and not a cropped

one, in order to estimate the class of the full-size image, the following algorithm is taken into account.

$$argmax(pred(image_{i,j})) \tag{6}$$

where $image_{i,j}$ is cropped image from full-size images, and *pred* is a function to get predicted class from each cropped images, and *argmax* is a function to get the maximum prediction of given values. Hence, the above expression notes that the predicted class of a full-size degradation image is determined by the *argmax* function of the cropped image classes. The prediction result can be represented as a probability value with respect to the degradation class defined beforehand using *argmax* function. In this case, full-size degradation images are predicted by our model with 99.90% accuracy. Figure 9 shows a prediction result of a full-size image using cropped images. Different colors in the image belong to different degradation classes, which the smaller number of image pieces represent false prediction for a particular class.

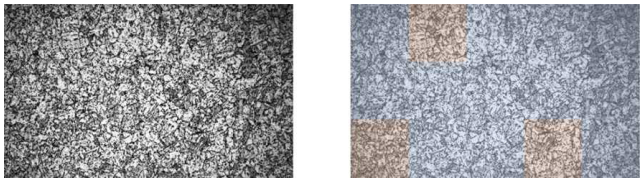

FIGURE 9. An original image (left) and the cropped images are used to predict a full size image (right).

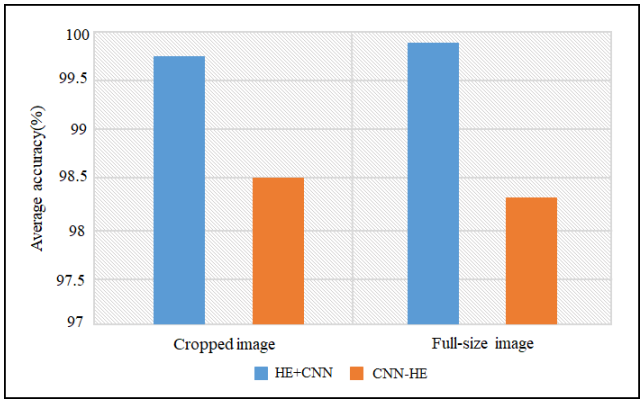

FIGURE 10. Model comparison of CNN model with and without histogram equalization.

The average accuracy of the proposed CNN architecture with and without histogram equalization are also benchmarked in order to indicate the effectiveness of histogram equalization in improving a deep learning model to classify degradation images. As shown in Figure 10, the proposed histogram equalization combined with convolutional neural network (HE+CNN) significantly be able to outperform the convolutional neural network with no histogram equalization is applied (CNN-HE). In the absence of histogram equalization, the CNN model yields an average classification accuracy of 98% and 97.78% in predicting cropped and full-size image, respectively. The result signifies the importance of such preprocessing technique, e.g. histogram equalization, so as an improved prediction accuracy of a CNN model could be obtained considerably. Furthermore, it is aligned with previous works presented in the current literature such as [36] and [37].

We are also interested to compare the impact of different kernel sizes in the convolutional layer on the performance accuracy. The result is summarized in Table 4. It can be observed that increasing the number of kernel size by 5×5 and 7×7 do not give us a substantial classification improvement, even decreasing the classification rate. The first convolution layer with a kernel size of 7×7 does not detect any additional features, hence the test samples are classified as belonging to the same class.

TABLE 4. Comparison of performance accuracy with respect to various kernel sizes.

Kernel size	Accuracy (%)	
	Cropped image	Full-size image
3×3	99.8	99.9
5×5	98.5	98.9
7×7	24.5	22.2

An advantage of CNN over traditional machine learning algorithms is a feature learning ability, where the representations required for feature detection and classification are proceeded automatically. This takes place a manual feature engineering approach that needs well-designed feature representation to both learn the features and use them to perform classification task. Therefore, it would be beneficial to benchmark the performance of CNN against a machine learning baseline, i.e. support vector machine (SVM). The SVM is one of well-known classification techniques that has been introduced for solving pattern recognition problems. In this classier, the data is projected into a high dimensional space and the classifier builds an optimal separating hyperplane in this space. This requires several kernels and parameters for solving a quadratic problem. As we cope with a large number of features, a linear-based SVM (LIBLINEAR) [38] is used for the purpose of obtaining a classification model with competitive accuracy and somewhat faster training time than other kernels, e.g. LIBSVM [39]. We only carry out a performance benchmark for cropped-image classification, given the fact that it is a computationally infeasible for SVM to classify large size image.

We first extract dense local descriptors of the images using local binary pattern (LBP) approach [40]. After extracting dense local descriptors, we perform classification task using SVM within 1000 iterations. As a result, SVM classifier achieves 72.1% with respect to classification accuracy. This result demonstrates the effectiveness of CNN model in contrast to SVM classifier. Furthermore, in terms of computational time, LBP and SVM take 942.78 seconds for extracting and classifying the degradation images, respectively. This is slightly faster than CNN which needs 6840.04 seconds to train the training set. However, for practical implementation,

the performance of CNN is still acceptable, considering that training task can be done offline and can then be used for an off-line degradation detection/diagnosis tool.

V. CONCLUSION

In this paper, a classification and diagnosis method for material degradation of the T91 steel boiler tubes was presented based on the combination of an adaptive histogram equalization and convolutional neural network (CNN). A raw image data set was firstly preprocessed by histogram equalization algorithm in such a way that the intensity of pixels of each image was distributed uniformly. As we suffered from a small number of data set, we generated more training samples by cutting, flipping, and rotating randomly the images for the input of CNN model.

Our proposed CNN model, which was made up of four convolution layers and two pooling layers had shown a promising performance result in detecting and classifying material degradation image by 99.80% and 99.90% for the cropped image and full-size image prediction, respectively. To the best of our knowledge, this work is the first attempt so far in applying CNN based technique for material degradation image. For future work, it is necessary to consider class activation mapping of CNN with global average pooling which can perform localization within a degradation image.

ACKNOWLEDGMENT

(Woosung Choi and Hyunsuk Huh contributed equally to this work.)

REFERENCES

[1] T. Á. Tejedor, R. Singh, and P. Pilidis, "Maintenance and repair of gas turbine components," in *Modern Gas Turbine Systems*. Amsterdam, The Netherlands: Elsevier, 2013, pp. 565–634.

[2] S. Prakash, "Development of advanced alloys with improved resistance to corrosion and stress corrosion cracking (SCC) in power plants," in *Structural Alloys for Power Plants*. Amsterdam, The Netherlands: Elsevier, 2014, pp. 294–318.

[3] M. Kutz, *Handbook of Environmental Degradation of Materials*. New York, NY, USA: William Andrew, 2018.

[4] M. G. Fontana, *Corrosion Engineering*. New York, NY, USA: McGraw-Hill, 2005.

[5] B. A. Tama and K.-H. Rhee, "An in-depth experimental study of anomaly detection using gradient boosted machine," *Neural Comput. Appl.*, vol. 31, no. 4, pp. 955–965, Apr. 2019.

[6] J. Wang, Y. Ma, L. Zhang, R. X. Gao, and D. Wu, "Deep learning for smart manufacturing: Methods and applications," *J. Manuf. Syst.*, vol. 48, pp. 144–156, Jul. 2018.

[7] J. Davis, T. Edgar, J. Porter, J. Bernaden, and M. Sarli, "Smart manufacturing, manufacturing intelligence and demand-dynamic performance," *Comput. Chem. Eng.*, vol. 47, pp. 145–156, Dec. 2012.

[8] H. S. Kang, J. Y. Lee, S. Choi, H. Kim, J. H. Park, J. Y. Son, B. H. Kim, and S. D. Noh, "Smart manufacturing: Past research, present findings, and future directions," *Int. J. Precis. Eng. Manuf.-Green Technol.*, vol. 3, no. 1, pp. 111–128, Jan. 2016.

[9] F. Tao, Q. Qi, A. Liu, and A. Kusiak, "Data-driven smart manufacturing," *J. Manuf. Syst.*, vol. 48, pp. 157–169, Jul. 2018.

[10] B. A. Tama and K.-H. Rhee, "Tree-based classifier ensembles for early detection method of diabetes: An exploratory study," *Artif. Intell. Rev.*, vol. 51, no. 3, pp. 355–370, Mar. 2019.

[11] P. Wang, R. Yan, and R. X. Gao, "Virtualization and deep recognition for system fault classification," *J. Manuf. Syst.*, vol. 44, pp. 310–316, Jul. 2017.

[12] Z. Chen, C. Li, and R.-V. Sanchez, "Gearbox fault identification and classification with convolutional neural networks," *Shock Vib.*, vol. 2015, Apr. 2015, Art. no. 390134.

[13] R. Liu, B. Yang, E. Zio, and X. Chen, "Artificial intelligence for fault diagnosis of rotating machinery: A review," *Mech. Syst. Signal Process.*, vol. 108, pp. 33–47, Aug. 2018.

[14] Y. Lei, F. Jia, J. Lin, S. Xing, and S. X. Ding, "An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data," *IEEE Trans. Ind. Electron.*, vol. 63, no. 5, pp. 3137–3147, May 2016.

[15] F. Jia, Y. Lei, J. Lin, X. Zhou, and N. Lu, "Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data," *Mech. Syst. Signal Process.*, vol. 72, pp. 303–315, May 2016.

[16] O. Janssens, V. Slavkovikj, B. Vervisch, K. Stockman, M. Loccufier, S. Verstockt, R. Van de Walle, and S. Van Hoecke, "Convolutional neural network based fault detection for rotating machinery," *J. Sound Vib.*, vol. 377, pp. 331–345, Sep. 2016.

[17] X. Ding and Q. He, "Energy-fluctuated multiscale feature learning with deep convnet for intelligent spindle bearing fault diagnosis," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 8, pp. 1926–1935, Aug. 2017.

[18] L. Guo, H. Gao, H. Huang, X. He, and S. Li, "Multifeatures fusion and nonlinear dimension reduction for intelligent bearing condition monitoring," *Shock Vib.*, vol. 2016, Jan. 2016, Art. no. 4632562.

[19] T. Ince, S. Kiranyaz, L. Eren, M. Askar, and M. Gabbouj, "Real-time motor fault detection by 1-D convolutional neural networks," *IEEE Trans. Ind. Electron.*, vol. 63, no. 11, pp. 7067–7075, Nov. 2016.

[20] D. Weimer, B. Scholz-Reiter, and M. Shpitalni, "Design of deep convolutional neural network architectures for automated feature extraction in industrial inspection," *CIRP Ann.-Manuf. Technol.*, vol. 65, no. 1, pp. 417–420, 2016.

[21] O. Abdeljaber, O. Avci, S. Kiranyaz, M. Gabbouj, and D. J. Inman, "Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks," *J. Sound Vib.*, vol. 388, pp. 154–170, Feb. 2017.

[22] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3rd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2008.

[23] H. C. Furtado and I. L. May, "High temperature degradation in power plants and refineries," *Mater. Res.*, vol. 7, no. 1, pp. 103–110, 2004.

[24] B. Neubauer and U. Wedel, "Restlife estimation of creeping components by means of replicas," in *Proc. ASME Int. Conf. Adv. Life Predict. Methods*, 1983, pp. 353–356.

[25] R. Viswanathan, *Damage Mechanisms and Life Assessment of High-Temperature Components*. Novelty, OH, USA: ASM International, 1989.

[26] W. Choi, E. Fleury, B.-S. Kim, and J.-S. Hyun, "Life assessment of steam turbine components based on viscoplastic analysis," *J. Solid Mech. Mater. Eng.*, vol. 2, no. 4, pp. 478–486, Jan. 2008.

[27] W. Choi, K. Fujiyama, B. Kim, and G. Song, "Development of thermal stress concentration factors for life assessment of turbine casings," *Int. J. Pressure Vessels Piping*, vol. 98, pp. 1–7, Oct. 2012.

[28] L. Kachanov, *Introduction to Continuum Damage Mechanics*, vol. 10. Amsterdam, The Netherlands: Springer, 1986.

[29] R. K. Penny and D. L. Marriott, *Design for Creep*. Amsterdam, The Netherlands: Springer, 1995.

[30] G. A. Webster and R. A. Ainsworth, *High Temperature Component Life Assessment*. Dordrecht, The Netherlands: Springer, 1994.

[31] S. E. Umbaugh, *Digital Image Processing and Analysis: Applications with MATLAB and CVIPtools*. Boca Raton, FL, USA: CRC Press, 2017.

[32] K. Zuiderveld, "Contrast limited adaptive histogram equalization," in *Graphics Gems*. San Diego, CA, USA: Academic Press Professional, 1994, pp. 474–485.

[33] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, p. 436, 2015.

[34] C. Farabet, Y. LeCun, K. Kavukcuoglu, E. Culurciello, B. Martini, P. Akselrod, and S. Talay, "Large-scale FPGA-based convolutional networks," in *Scaling up Machine Learning: Parallel and Distributed Approaches*. New York, NY, USA: Cambridge Univ. Press, 2011, pp. 399–419.

[35] F. Chollet and J. J. Allaire, *Deep Learning with R*. Shelter Island, NY, USA: Manning Publications Company, 2018.

[36] L. Perez and J. Wang, "The effectiveness of data augmentation in image classification using deep learning," 2017, *arXiv:1712.04621*. [Online]. Available: https://arxiv.org/abs/1712.04621

[37] A. Mikołajczyk and M. Grochowski, "Data augmentation for improving deep learning in image classification problem," in *Proc. Int. Interdiscipl. PhD Workshop (IIPhDW)*, May 2018, pp. 117–122.
[38] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin, "LIBLINEAR: A library for large linear classification," *J. Mach. Learn. Res.*, vol. 9, pp. 1871–1874, Jun. 2008.
[39] C.-C. Chang and C.-J. Lin, "Libsvm: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, p. 27, Apr. 2011.
[40] X. Wang, T. X. Han, and S. Yan, "An HOG-LBP human detector with partial occlusion handling," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Sep./Oct. 2009, pp. 32–39.

WOOSUNG CHOI received the B.S. degree from Inha University, Incheon, South Korea, in 2003, the M.S. degree from KAIST, Daejeon, South Korea, in 2005, and the Ph.D. degree from the Department of Mechanical and Aerospace Engineering, Seoul National University, Seoul, South Korea, in 2018. He is currently a Senior Researcher with the KEPCO Research Institute, Daejeon, where he is in charge of life and risk assessment for power plant and the development of national standard for power generation industry (Korea Electric Power Industry Code). His research interests include prognostics and health management, multi-fidelity analysis with FEA, and deep learning analysis.

HYUNSUK HUH received the B.S. degree from the Ulsan National Institute of Science and Technology, Ulsan, South Korea, in 2018. He is currently pursuing the M.S./Ph.D. degree with POSTECH. His research interests include artificial intelligence in mechanical systems and applications of artificial intelligence in robotics.

BAYU ADHI TAMA received the Ph.D. degree from Pukyong National University, South Korea. He was a Postdoctoral Researcher with the School of Management Engineering, Ulsan National Institute of Science and Technology, South Korea. He is currently a Postdoctoral Researcher with the Industrial Artificial Intelligence Laboratory POSTECH, South Korea. He has published more than 20 papers in academic conferences and journals. His research interests include machine learning and artificial intelligence techniques applied for cyber security, medical informatics, and industrial applications. He received a full scholarship for his doctoral study from the Korean Government and an award for his excellent academic achievement from the Ministry of Education, South Korea.

GYUSANG PARK received the B.S. and M.S. degrees in aerospace engineering from Chonbuk National University, Jeonju, South Korea, in 2015 and 2017, respectively. He is currently a Researcher with the Korea Electric Power Research Institute, Daejeon, South Korea. His research interests experimental approach include power plant and industrial gas turbine performance.

SEUNGCHUL LEE received the M.S. and Ph.D. degrees from the University of Michigan, Ann Arbor, USA, in 2008 and 2010, respectively. He was an Assistant Professor with the Ulsan National Institute of Science and Technology, South Korea, and a Postdoctoral Research Fellow with the University of Michigan, Ann Arbor, USA. He has been an Assistant Professor with the Department of Mechanical Engineering, POSTECH, since 2018. His research interests include industrial artificial intelligence with mechanical systems, deep learning for machine healthcare, the IoT-based smart manufacturing, and manufacturing data visualization. He extends his research work to developing self-sustainable systems via an intelligent, informatics, and the industrial IoT system design at POSTECH.