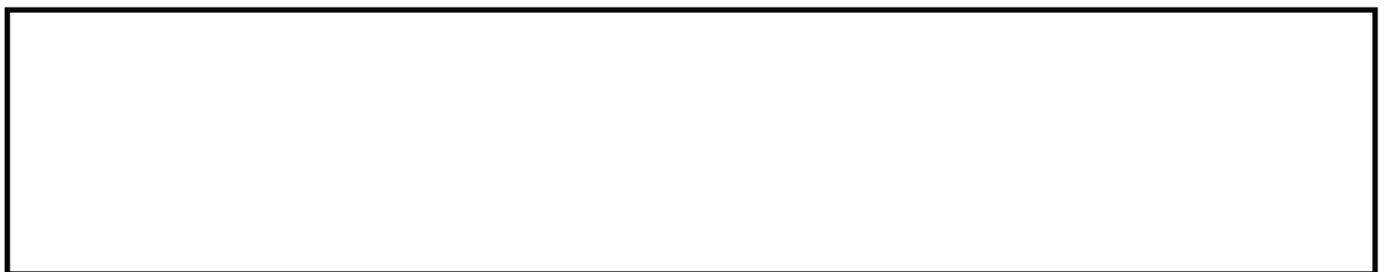


A lightweight deep learning-based approach for concrete crack characterization using acoustic emission signals.

HABIB, M.A., HASAN, M.J. and KIM, J.-M.

2021



Received July 3, 2021, accepted July 18, 2021, date of publication July 26, 2021, date of current version July 30, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3099124

A Lightweight Deep Learning-Based Approach for Concrete Crack Characterization Using Acoustic Emission Signals

MD. ARAFAT HABIB^{ID}, MD. JUNAYED HASAN^{ID}, AND JONG-MYON KIM^{ID}, (Member, IEEE)

Department of Electrical, Electronics and Computer Engineering, University of Ulsan, Ulsan 680749, South Korea

Corresponding author: Jong-Myon Kim (jmkim07@ulsan.ac.kr)

This work was supported by the National Demand Customized Life Safety Research and Development Project through the Korean Ministry of Interior and Safety (MOIS) under Grant 2019-MOIS41-002.

ABSTRACT This paper proposes an acoustic emission (AE) based automated crack characterization method for reinforced concrete (RC) beams using a memory efficient lightweight convolutional neural network named SqueezeNet. The proposed method also includes a signal-to-image technique, which is continuous wavelet transformation (CWT) that decomposes the AE signals over time-frequency scales and extracts the crack/fracture information in both the time and frequency domains. First, AE signals for two types of cracks (minor and severe), along with the normal condition (no crack), are collected from the experimental test bed. Second, the previously mentioned CWT based signal-to-image technique is applied to generate two-dimensional time-frequency images that are then converted to gray scale images for faster computation. These images are supplied to the SqueezeNet for classification of the concrete crack types. We extensively modified the fire module of the SqueezeNet (SQN-MF) by introducing depth-wise convolutional kernels and channel shuffling operations. Not only does the proposed method utilize deep learning-based techniques for crack classification of concrete beams for the first time, but also the CWT-based imaging technique has not yet been explored in this field either. Additionally, this method does not follow the typical AE burst feature (features like AE counts, peak-amplitude, rise time, decay time, etc.) based methods, and as a result, we no longer require extensive human intervention and expertise to get deep understanding of the crack types. SQN-MF achieves AlexNet-level accuracy with fifty times fewer parameters and has an implementable memory size for the field programmable gate array boards. Overall, the method achieves 100% accuracy. It is 20.8% higher than the typical feature extraction and traditional machine learning based methods. We observed a 4% accuracy increase for the proposed SQN-MF compared to the typical SqueezeNet with bypass connections.

INDEX TERMS Concrete crack characterization, continuous wavelet transformation, convolutional neural network, SqueezeNet.

I. INTRODUCTION

Acoustic emission (AE) techniques have become popular lately as a propitious way of monitoring concrete structures [1]–[8]. If any crack occurs in a concrete beam, materials associated with the crack releases energy. This results in wave propagation that can be identified using mounted AE sensors set on the surface of the concrete beam [2], [9]–[14]. Recorded AE signals are used to determine the latest condition of the beam. Crack detection, damage assessment, and crack location identification are meticulously done

The associate editor coordinating the review of this manuscript and approving it for publication was Yu Wang^{ID}.

through the rigorous analysis of the AE signals. AE techniques allow us to listen to the sounds as they form due to plastic deformation and micro-cracking in the materials. AE signals show high sensitivity to crack occurrence and microscopic processes [15]. This helps us greatly to identify stress waves caused due to structural anomalies. AE is particularly distinguished from the other non-destructive techniques since it has the unique ability to perform the detection of debonding in materials and crack propagation recognition. Additionally, AE signals extracted under a heavily loaded concrete beam can facilitate us with efficacious characterization and identification of cracks. As a result, we have a rich set of literature where researchers have

proposed many different methods based on the AE burst features and traditional machine learning algorithms [1]–[3], [5]–[8], [12]–[14], [16]–[22]. These methods propose robust solutions for the characterization of the crack types associated with reinforced concrete (RC) beams, but the following issues should be cautiously addressed to design and build an automated, memory-efficient, and accurate method further:

- Existing methods highly depend on crucial AE burst features like rise time, decay time, peak amplitude, AE energy, counts, etc. These features show non-linear and non-stationary behavior. It is very difficult to sort out the features that have higher sensitivity to crack growth. Certain features can be sensitive to latent cracks, while some might show sensitivity to severe cracks (fractures) at the peak hour of failure. It requires extensive domain knowledge and human expertise to select features that can be effective while performing classification tasks for multiple crack types.
- Some of the works undertook feature selector algorithms for selecting relevant features ([23], [24]). However, we have to keep in mind that the computation complexity is already high since we have to extract AE features using multiple loops for huge chunks of data. Adding a feature selector algorithm makes the whole system expensive.
- Many of the machine learning based algorithms have significantly failed to achieve precision in terms of distinguishing micro-cracks from macro-cracks. Addressing this problem is crucial since any kind of misinterpretation of the macro-cracks as micro may lead to a fatal catastrophe.
- One of the biggest problems of using machine learning algorithms is that we never know if it is applicable in real-time embedded systems. The existing literature does not show any concern about the memory consumption of their proposed methods.

To overcome the limitations mentioned above, it is necessary to develop an automated method that will no longer require any human intervention, get rid of feature selector algorithms, and be memory efficient. Therefore, we propose a deep learning-based algorithm that makes use of a signal-to-image technique named continuous wavelet transformation (CWT) [25] along with a recently developed convolutional neural network (CNN): SqueezeNet [26]. The CWT-based signal to imaging technique supplies SqueezeNet with ample two-dimensional gray scale images, which is used to train the network for corresponding crack type classification. To further decrease the number of parameters and enhance the learning ability, we designed a new fire module for the SqueezeNet classifier that results in 100% classification accuracy. We will address our proposed SqueezeNet with the modified fire module as SQN-MF throughout the paper for our convenience.

A wavelet can be defined as a wave-like oscillation with a zero amplitude that increases and again decreases back

to zero [27]. It can be used as a useful mathematical tool to extract information from variant data types like signals, images, etc. We can divide a function or continuous-time signal into different scale components using a wavelet. A wavelet transform is the representation of a function by wavelets. CWT provides us with a representation of a signal by allowing the translation and scale parameters of the wavelets to vary continuously [28]. CWT is a tool for signal processing that has been widely used for fault diagnosis of mechanical machines, composite materials, etc. Zheng *et al.* proposed a new approach of gear fault diagnosis based on CWT [29]. They argued that CWT can provide a finer scale resolution than orthogonal wavelet transforms. That is why CWT was more appropriate for them to extract mechanical fault information. Kankar *et al.* performed fault diagnosis for rolling element bearings using CWT and two classification techniques: an artificial neural network and a support vector machine [30]. Bhattacharjee *et al.* conducted a study aiming at the estimation of the damping parameters of composite beams formed with glass-fiber-reinforced epoxy, glass-fiber-reinforced polyester, and carbon black filled epoxy [31]. CWT was used to investigate the system response analytically. Even though CWT has been successfully used for many different cases in the field of fault diagnosis, it has not been explored yet as a signal processing tool to determine crack types for RC beams. In this study, we attempt to use CWT as a tool for generating 2D images based on the time domain AE signals to be considered further for crack type classification by SqueezeNet.

Deep learning algorithms are capable of extracting autonomous features from different domains. This allows to eliminate hand-crafted feature extraction and implementation of feature selection algorithms. Even though this state-of-the-art method has often been used for fault classification of bearings [32]–[37], gearboxes ([38]–[41]), rail defects ([42], [43]) etc., we have not found any specific work that focusses on concrete crack classification. According to the best of our knowledge, the proposed work in this paper is going to be the first one driving the trend of crack type classification of RC beams towards deep learning. To practically implement any diagnostic framework for machinal machines or composite materials, we need to build embedded systems to be deployed in real-time. However, deep learning algorithms like CNN are extremely memory consuming and computation power hungry, which make it difficult for us to adopt real-time implementation. Considering these issues, we have chosen SqueezeNet, which has a model size of only 0.5 MB and can be implemented in field programmable gate arrays (FPGA). Even though the architecture is so lightweight and small, it can gain over 90% classification accuracy for concrete crack type classification. Our aim is to classify concrete cracks, which is a way too critical application in which an accuracy of 90% is not enough, since any kind of mistake may lead to catastrophe including the loss of human lives. To increase the accuracy and reduce the number of parameters, we modified the fire module of the SqueezeNet

by changing the *expand* layer. Typical 3×3 filters are replaced by 3×3 depth-wise convolution (DWConv) kernels and a channel shuffle operation. This leads to 100% accuracy of the model.

In our work, first time-domain acoustic emission signals were collected from the test bed where concrete beams were subjected to three-point bending tests. These signals are next processed using CWT. The 2D images formed using CWT contain the information related to the severity of the concrete cracks over different time and frequency scales. These images are further converted to gray-scale images to lessen the computation overload of the SQN-MF classifier. Finally, to identify crack types, the SQN-MF is implemented, which makes use of the 2D grayscale images as input and performs stringent classification of the crack types. The main contribution of this research paper is summarized as follows:

- A reliable signal processing technique, for the first time in the literature of RC beam crack characterization, has been proposed that can precisely capture fault information for minor and severe cracks.
- Unlike most of the existing works, the proposed method is based on deep learning that does not require any hand-crafted AE features. This is important because it helps to ameliorate systems without a high level of human expertise and domain knowledge. Furthermore, fluctuating, unpredictable, and non-linear behavior of the AE features resulting in imprecise crack classification and characterization can be avoided too.
- Since there are no direct feature extraction processes like the methods proposed in [23], [24], no feature selection algorithm is necessary. This helps to get rid of an added layer of complexity.
- The proposed method opens up the horizon of implementing CNN in real-time devices by incorporating SQN-MF, which not only is extremely memory efficient, but classifies crack types with 100 percent accuracy. This kind of accuracy is much higher than traditional classification methods.

The rest of the paper is arranged as follows. In Section II, the works related to the proposed study are discussed in short. Section III presents the technical details of the CWT and SqueezeNet. Section IV describes the proposed methodology elaborately, including the experimental test setup covering the three-point bending tests performed on concrete beams for gathering AE data, validation of the acquired dataset using a state-of-the-art plastic deformation model, modification of the typical SqueezeNet fire module to build a more robust classifier (SQN-MF), and lastly, a step-by-step summary of the crack type classification process via the SQN-MF. The experimental results accompanied by analytical discussion is presented in section V. The conclusion is drawn in Section VI.

II. RELATED WORKS

Carpinteri *et al.* [13] used impact tests and pull out tests to determine the mechanical parameters of a retro-filled concrete beam. Creep effects and micro-cracking were investigated using AE. Aggelis [4] conducted experiments on concrete beams using bending tests with multiple concrete beams. Collected AE signals from the bending tests showed variant patterns at different stages of fracture occurrence. The conducted study was able to differentiate well between micro and macro-cracks. Tsangouri *et al.* [6] argued that observation of the mechanical properties alone cannot characterize macro-cracks taking place in concrete beams. To understand the damage process, they proposed an AE technique that could sort out the damage initiation process by using the cumulative AE hits observed while an increasing load was applied. Load/calm ratios were calculated to comprehend the damage severity. Chen *et al.* [8] performed experiments on the impact of the loading rate on the behavior of the occurred fractures. For on spot fracture monitoring, AE was used. The study shows that AE hits and cumulative ringing count can represent the point of commencement of the concrete boundary effect over time. Also, they figured out that the width of the cracks of the beams is covered accordingly with the increase of the loading rate. Furthermore, as the loading rate increased, the sample specimen's ductility decreased. Lastly, the quantity of shear cracks increased as the loading rate was higher. Sengsiri *et al.* conducted numerical investigations into dynamic modal parameters of fiber-reinforced foamed urethane (FFU) composite beams in railway switches and crossings in [44]. However, they also proposed a method on AE based damage detection in FFU composite railway bearers [45]. AE was used identify the load-deflection curve of the beams. Janeliukstis *et al.* [46] proposed a novel method for AE based condition monitoring of railway prestressed concrete sleepers under flexural tests. Their proposed method enabled preventive, predictive and condition-based track maintenance for the railway industry.

Related works discussed so far do not provide any solutions based on machine learning algorithms. Combining AE features and traditional machine learning models, many of the approaches have performed concrete crack characterization. Das *et al.* [1] designed a framework for the automation of probabilistic classification of the cracks on the basis of the AE signals. They used hand crafted AE features like RA values and average frequency. These waveform parameters were clustered using an unsupervised clustering algorithm. Created clusters intersected among themselves and were later separated by a hyperplane generated using a support vector machine. Based on the understanding of the labeled data, unlabeled data was classified into different crack modes. Yu *et al.* [2] presented a study to observe the slow dynamics of the micro-cracks occurring in polymer concrete (PC) samples. PC samples went through the three-point bending tests and the collected AE data was used to monitor the damage process in real-time. Lastly, the AE data was

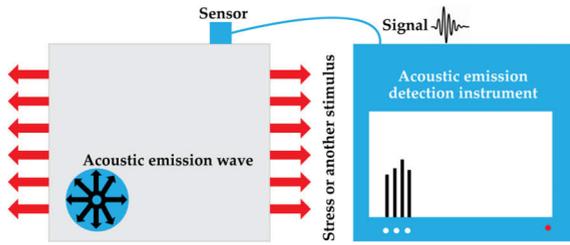


FIGURE 1. Acoustic emission data collection system for concrete beams.

classified using principal component analysis and k-means clustering. Sasmal et al [47] recorded AE signals at different stages over time to identify crack commencement and progression. Acoustic parameters like energy and signal strength are considered to observe their efficiency in determining the crack commencement and propagation. Supervised and unsupervised-pattern recognition algorithms were used for the classification of the AE signals in different damage stages. A great contribution of this work is the validation of the support vector machine-based crack classification via the experimentally observed damage pattern at different stages. Bayer *et al.* [48] observed the crack pattern and propagation in a concrete surface using an artificial intelligence-based technique named voronoi diagrams. It is an image-based technique where a random photo of a concrete crack is used. The crack is next divided into 12 segments for the estimation of the machine learning algorithm's efficiency to predict the crack pattern including its direction. The proposed method in the study was precise, quick, and cheap.

Observing the initial crack formation until the complete failure of the RC beam throughout the time history is a crucial task, which was completely ignored previously. Last year, Kim *et al.* [23] proposed a degradation indicator (DI) based on Mahalanobis-Taguchi system (MTS) that used AE features like peak amplitude, rise time, decay time, etc. MTS was used to fuse all the necessary features and build the DI. Noise associated with the proposed DI was removed using Chebyshev's inequality. The DI provides the ability to observe the development of concrete degradation starting from the crack initiation process until fatal failure. The DI did not require any previous knowledge of the data associated with failure for conducting assessment.

Tra *et al.* used a constant-false-alarm-rate algorithm to detect impulses taking place in the AE signals. AE features including counts, amplitude, rise time, energy, etc. were used in the process. Based on their observation, increasing damage resulted in the change of the properties of these features over time. To classify crack types, they used the k-nearest neighbor classifier.

Habib *et al.* [24] proposed a complete framework for crack characterization of the RC beams that not only provided a reliable crack assessment indicator (CAI), but also conducted precise classification of the normal condition and micro and macro-cracks. They also used a wrapper-based feature selection algorithm name Boruta that provided a robust solution

for crucial feature selection. The proposed CAI could display the time history of the crack occurrence starting from the initial stage to the final failure stage. The CAI plot was lastly passed through a k-NN classifier to classify the normal condition, micro, and macro-cracks. The proposed solution could stringently distinguish between the micro and macro-cracks, which many other previous works failed to achieve.

III. TECHNICAL BACKGROUND

A. MONITORING RC STRUCTURES USING AE

Observing structural degradation is an effective way of AE monitoring. If a crack occurs, a stress wave is produced that acts as an AE source. Stimulus created due to the stress wave acts upon the concrete material leading to local plastic deformation [49]. The stress wave flows through the material from the source to the surface area and gets recorded. AE sensors play a crucial role in the process as it converts the stress wave to an electrical signal [49]. If the sensor is nonintegral, amplification of the signal is necessary before transmission. Integral sensors are free of such processes since those have embedded amplifiers. The wave gets transferred directly to the instrument for the integral sensors. After that, the wave is transmitted for data acquisition for the recording, storing, and analyzing. Fig. 1 shows the basic view of the AE data collection process for a reinforced concrete beam.

The attenuation effect causes difficulty for the AE monitoring of composite materials like RC beams. RC beams are heterogeneous materials made of cement, coarse aggregates, fine aggregates, and steel bars as reinforcement. Also, they have flaws like air voids, pores, etc. Therefore, we have to be careful while choosing the sensor and make sure that the chosen one has a lower resonant frequency and a proper frequency range.

B. ACOUSTIC EMISSION BURST FEATURES

There are some special parameters used to represent AE events. Characterization of AE bursts for the collected AE signals can be done using these parameters. Most of the existing works have used the AE burst features like peak amplitude, AE counts, rise time, decay time, etc. to develop crack assessment indicators and crack classification methods for the RC concrete beams. The aforementioned features are highly related to the maintenance of the concrete health. A pictorial view of these features is shown in Fig. 2. They are also further defined as follows [49], [50]:

Peak amplitude: The voltage that is the highest in an AE waveform is regarded as the peak amplitude.

AE counts: This is the number of times that the AE signal crosses a certain preset threshold in a particular portion of the test.

AE energy: This involves measuring a certain area that is conducted within the rectified signal envelope.

Rise time and decay time: The time duration between the first threshold crossing and the peak amplitude is defined as the rise time. The decay time is the opposite of this. It is the

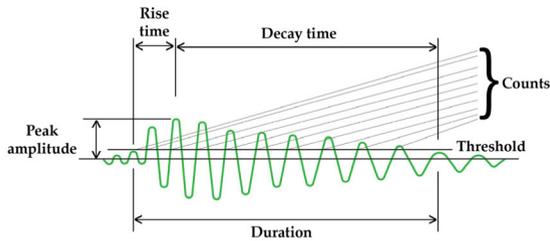


FIGURE 2. A pictorial view of the crucial AE features.

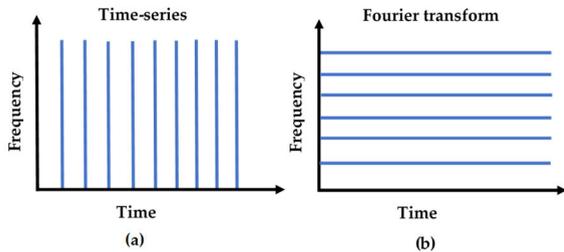


FIGURE 3. Data in two different domains: (a) time and (b) frequency.

time when the peak value of the amplitude decreases while within the threshold.

One of the prime goals of our work is to prove the superiority of the proposed image-based method over the typical AE burst feature-based methods. To do that we have implemented an AE feature-oriented classification method and compared the performance, adopting various parameters. That is why the concepts of the AE burst features are discussed herein.

C. WAVELET ANALYSIS AND CONTINUOUS WAVELET TRANSFORM

It is possible to Fourier transform time domain data and transform it to the frequency domain. Time series data can give us insight regarding exquisite resolution of where the signal lies in time, but it is impossible to comprehend what frequencies are occurring at that moment in time. Similarly, performing a Fourier transform would give us exact information about the frequency components of the data, but there will be no information regarding when those frequencies occurred in time. In Fig. 3(a), it is visible that we have absolutely no way to gain insight about the frequencies that are occurring at a particular time. Also, from Fig. 3(b), there is no way we can gather any information about the time. This crisis has led researchers to develop the spectrogram, where equal weighting to time and frequency is ensured. In spectrograms, we have less time resolution than our original time series and less frequency resolution than our Fourier transform. From spectrograms, we know when individual frequencies turn on and off in time, as they can give us both time and frequency information.

Wavelets have been proposed as a superior method for the time frequency analysis [51]. Spectrograms have a fixed window length. On the other hand, wavelets have longer windows for lower frequencies and shorter windows for

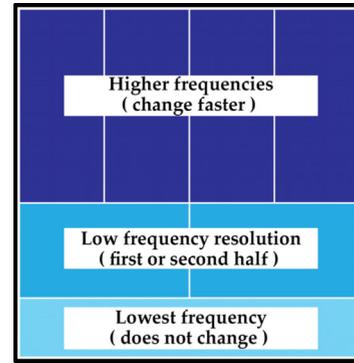


FIGURE 4. Basic wavelet formation criterion and formation.

higher frequencies, making them evidently a better method than spectrograms. In wavelet analysis, there is a hierarchical grading of time and frequency information. Frequencies that are very low tend to last for a long time and do not change much over time. This allows us to have a bar at the lower bottom of Fig. 4. This is basically the lowest frequency that no one cares about when it occurs in time because it is usually assumed that the frequency is on all the time, working as a baseline frequency. Next, the wavelet diagram presented in Fig. 4 would be split in half (colored in sky blue). This second stage will provide us with less frequency resolution, but we would have information if it were in the first half or second half of our time series. In the next level, we are going to break the rest of the diagram (colored in dark blue) into a higher frequency resolution, because at higher frequencies changes are faster in time, but we have less resolution about exactly which of those frequencies are turning on. Different colors have been used for the representation of different layers. They do not carry any technical meaning related to wavelet analysis.

A wavelet is a multiscale time-frequency decomposition. Lower frequencies change more slowly in time, so we do not need much temporal accuracy. For higher frequencies, more temporal accuracy is needed, but we get correspondingly more uncertainty in what exact frequency is turning on and off in that frequency range. Fig.4 is a basic diagram for wavelet analysis. In reality, we might have more than ten levels instead of three for wavelet decomposition. Therefore, wavelet decomposition can be referred to as a really good spectrogram that is tailored to spend as much information in the regions as necessary. For example, low frequencies do not need much information whereas higher frequencies need more temporal resolution, and so forth. For wavelet transforms, we have to take some signal/ time series or spatial data and project that on an orthogonal basis. The orthogonal basis is not going to be just sines and cosines, rather it will be a hierarchy of orthogonal functions that are going to get smaller in time or in space in little windows.

In our proposed method, CWT has been used to process the AE signals acquired from the concrete testbed. CWT utilizes a family of wavelet functions for the purpose of decomposing

the original signal into a spectrum of wavelet coefficients. To obtain this, translation and scaling of the original signal over different time and frequency levels should be performed. As the crack takes place, multiple low to high frequencies are produced into the original signal. Therefore, the energy content differs over different frequency ranges. CWT can be formally written as:

$$\gamma(s, \tau) = \int f(t) \Psi_{s,\tau}^*(t) dt, \tag{1}$$

where $*$ denotes complex conjugation. The values s and τ are the new dimensions named scale and translation after the wavelet transform. This equation displays how a function $f(t)$ can be decomposed into a set of basis functions $\Psi_{s,\tau}(t)$, called wavelets. The wavelets are generated from a single basic wavelet $\Psi(t)$, called the mother wavelet. It can be represented using the following equation:

$$\Psi_{s,\tau}(t) = \frac{1}{\sqrt{s}} \Psi(t - \tau/s). \tag{2}$$

For our implementation, CWT is obtained using a Morse wavelet[52]. The maximum level possible for decomposition was found to be 10 for our case. However, a smaller value is taken for decomposition, not the maximum allowed level. We have chosen the decomposition level to be 6.

D. CONVOLUTIONAL NEURAL NETWORK

A general CNN consists of an input layer, convolutional layers, pulling layers, fully connected layers, and an output layer. Typically, the first layer of a CNN is an input layer, after which a combination of multiple convolutional and pooling layers is added. The last layer is the fully connected layer. Finally, a SoftMax classifier can be used to classify 2D images.

The input layer of a CNN can be used to preprocess the data, and at the same time, it can take raw data as input to the neural network. The preprocessing part may include image conversion as a vector and normalization process to boost the speed of the CNN during the training phase. The convolutional layer is a crucial part of the CNN and is intended to produce a feature map by a convolutional operation of a set of weighted filters. Actually, a convolutional operation is nothing but a point-multiplication summation of two-pixel matrices where one is the input data matrix and the other one is the filter.

It is necessary to use an activation function, the main purpose of which is to make the characteristic map of the output have a nonlinear relationship. Sigmoid, tanh, ReLU [23], etc. can be used as activation functions. ReLU is widely used as an activation function because of the speed of unsaturated nonlinear functions. The function form of ReLU is as follows:

$$f(x) = \begin{cases} 0, & \text{if } x < 0 \\ x, & \text{if } x \geq 0, \end{cases} \tag{3}$$

where x is the input of the activation function.

The pooling layer in the CNN is used to screen the features in the perceptron domain to extract the most representative features in that region. This can be very effective since the output feature scale can be reduced resulting in a reduction of the number of parameters needed for the model and the maintenance of the translation variance. Max pooling [24] is vastly used in most cases in the pooling layer.

The fully connected layer of a CNN summarizes the learned features. Finally, a classifier is used to classify the obtained features. If the length of the input vectors is M and the output vector is N , the number of the parameters in the fully connected layer can be obtained as follows:

$$Q = M \times N + N. \tag{4}$$

For two class classification, the training set has a label sample: $(x^{(1)}, y^{(1)}), \dots, (x^{(k)}, y^{(k)})$. The value of the label y can be either 0 or 1. The logistic regression function can be represented as follows:

$$h_{\theta}(x) = \frac{1}{1 + e^{-\theta^T x}}, \tag{5}$$

where θ is the model parameter of the loss function. The loss function $J(\theta)$ can be presented using the following equation:

$$J(\theta) = \frac{1}{k} \left[\sum_{i=1}^k y^{(i)} \log h_{\theta}(x^{(i)}) + (1 - y^{(i)}) \log(1 - h_{\theta}(x^{(i)})) \right]. \tag{6}$$

If our problem is a multistate classification problem, it is assumed that we have n conditions and n corresponding labels for the SoftMax classifier. In a training set, the label is represented as $y^{(i)} \in \{1, 2, 3, \dots, n\}$. Let us assume that a training sample x is given that has n classes. If the occurrence probability of a state i is $p(y = i|x)$, the output of the SoftMax regression can be represented using the following equation:

$$h_{\theta}(x) = \begin{bmatrix} p(y = 1|x; \theta) \\ p(y = 2|x; \theta) \\ \dots \\ p(y = n|x; \theta) \end{bmatrix} = \frac{1}{\sum_{j=1}^n e^{x\theta_j}} \begin{bmatrix} e^{x\theta_1} \\ e^{x\theta_2} \\ \dots \\ e^{x\theta_n} \end{bmatrix}, \tag{7}$$

where $\theta_1, \theta_2, \dots, \theta_n$ are the model parameters and $(1/\sum_{j=1}^n e^{x\theta_j})$ plays the role of normalization. The loss function, therefore, can be written in the following manner:

$$J'(\theta) = -\frac{1}{k} \left[\sum_{i=1}^k \sum_{j=1}^n \{y^{(i)} = j\} \log \left(\frac{e^{x\theta_j}}{\sum_{j=1}^n e^{x\theta_j}} \right) \right]. \tag{8}$$

In the training process of a CNN, forward propagation of the input signal is performed, and output data can be acquired after it passes through multiple layers. Next, output data and expected labels are compared, and the generated error is backpropagated layer by layer. The corresponding weights are updated, and the expected error lessens as we increase the number of iterations.

E. SQUEEZENET

The focus of recent research works related to deep convolutional neural networks is based on improving accuracy. For a certain accuracy level, it is not much more difficult to sort out multiple CNNs that can achieve that level. With the same accuracy level, CNN models that are small can facilitate us with three advantages. First, smaller CNNs would require less communication across servers while conducting distributed training. They can be trained faster since they need less communication. This is a great advantage for an image-based classification method like ours as communication among servers can be a limiting factor to the scalability of distributed CNN training. Second, using smaller CNNs would ease the process of exporting a new model from the cloud to an autonomous embedded system. A smart way of providing a concrete crack detection solution is deploying an autonomous system that will itself download the required CNN model from the cloud. A lightweight CNN model will be helpful in this regard. Lastly and most importantly, smaller CNN models are feasible for deployment on memory limited hardware like FPGAs. To acquire all these advantages, we have chosen to use SqueezeNet for our proposed solution.

There are three main strategies that the developers of SqueezeNet have adopted to obtain fewer parameters:

- Strategy 1: The first strategy adopted to design the architecture of SqueezeNet is the replacement of 3×3 filters by 1×1 filters. The majority of the filters are 1×1 for which the model has 9X fewer parameters.
- Strategy 2: The next strategy adopted to build the SqueezeNet is to reduce the number of input channels to 3×3 filters. To comprehend this, let us consider a convolutional layer that is entirely formed of 3×3 filters. The quantity of parameters in that very layer will be:

$$Q_p = N_{ic} \times N_f \times (3 \times 3), \tag{9}$$

where Q_p stands for the quantity of parameters, N_{ic} is the number of input channels, and N_f represents the number of filters. Therefore, not only do we need to reduce the number of 3×3 filters as mentioned in the first strategy, but also it is required to decrease the number of input channels to 3×3 filters. To do this, SqueezeNet uses squeeze layers that will be discussed next in this section.

- Strategy 3: The last strategy adopted to make sure that the SqueezeNet has fewer parameters is to down sample late in the network. This is performed to acquire convolution layers with large activation maps. For down sampling, strides are set to greater than one in some of the convolutional and pooling layers. Large activation maps can lead to higher classification accuracy [53]. In short, the first two strategies relate to the reduction of the quantity if parameters in a CNN and the last strategy is regarding the maximization of accuracy on a limited budget of parameters.

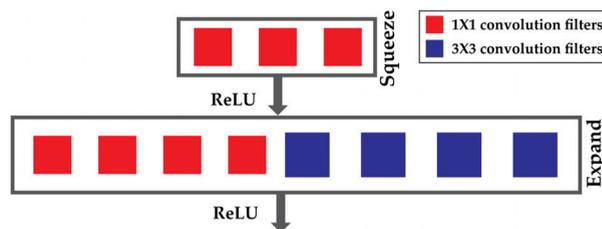


FIGURE 5. The fire module.

Apart from any other typical CNN architecture, the thing that makes SqueezeNet special is the introduction of fire modules. The fire module is described as follows:

- A fire module in a SqueezeNet consists of a squeeze convolutional layer with a 1×1 filter only. It feeds into an expand layer that consists of a mixture of 1×1 and 3×3 convolutional filters. Fig. 5 presents us with an organization of convolutional filters in the fire module. Using a 1×1 filter in fire modules is the part of the first strategy discussed earlier.
- Three hyperparameters are exposed in a fire module: a) $s_{1 \times 1}$: the number of filters in the squeeze layer (all the 1×1), b) $e_{1 \times 1}$: the number of 1×1 filters in the expand layer, and c) $e_{3 \times 3}$: the number of 3×3 filters in the expand layer.
- In the fire module, $s_{1 \times 1} < (e_{1 \times 1} + e_{3 \times 3})$, which helps to keep the number of input channels limited to 3×3 filters according to the second strategy adopted for SqueezeNet, as discussed before.

Fig. 6 represents the SqueezeNet architecture. It starts with a convolution layer, which is named conv1 in the diagram. After that, there are eight fire modules. Finally, it ends with conv10, which is the last convolutional layer. Throughout the network (beginning to end), the number of filters per fire module is gradually increased. After conv1, fire modules 4,8, and conv10, max pooling is performed. The stride is kept to 2 as per strategy 3.

Fig. 6(a) is the typical SqueezeNet (SQN-1) architecture with no bypass, Fig. 6(b) is the same architecture with simple bypass (SQN-2), and Fig. 6(c) is the same architecture with complex bypass (SQN-3). We have implemented all three of these SqueezeNet architectures using our 2D image dataset acquired using CWT from the AE signals. Comparative performance analysis related to the three different types of architecture can be found in section V. For our convenience we will address these networks as SQN-1, SQN-2, and SQN-3, respectively.

IV. PROPOSED METHODOLOGY

In this section of the paper, a description of the experimental test setup is presented first. This facilitates the readers with the idea of our setup, sensor deployment and selection, crack length measurement, etc. Next, we attempt to validate our dataset using the fracture mechanics of the composite materials, like concrete, to make sure that the acquired

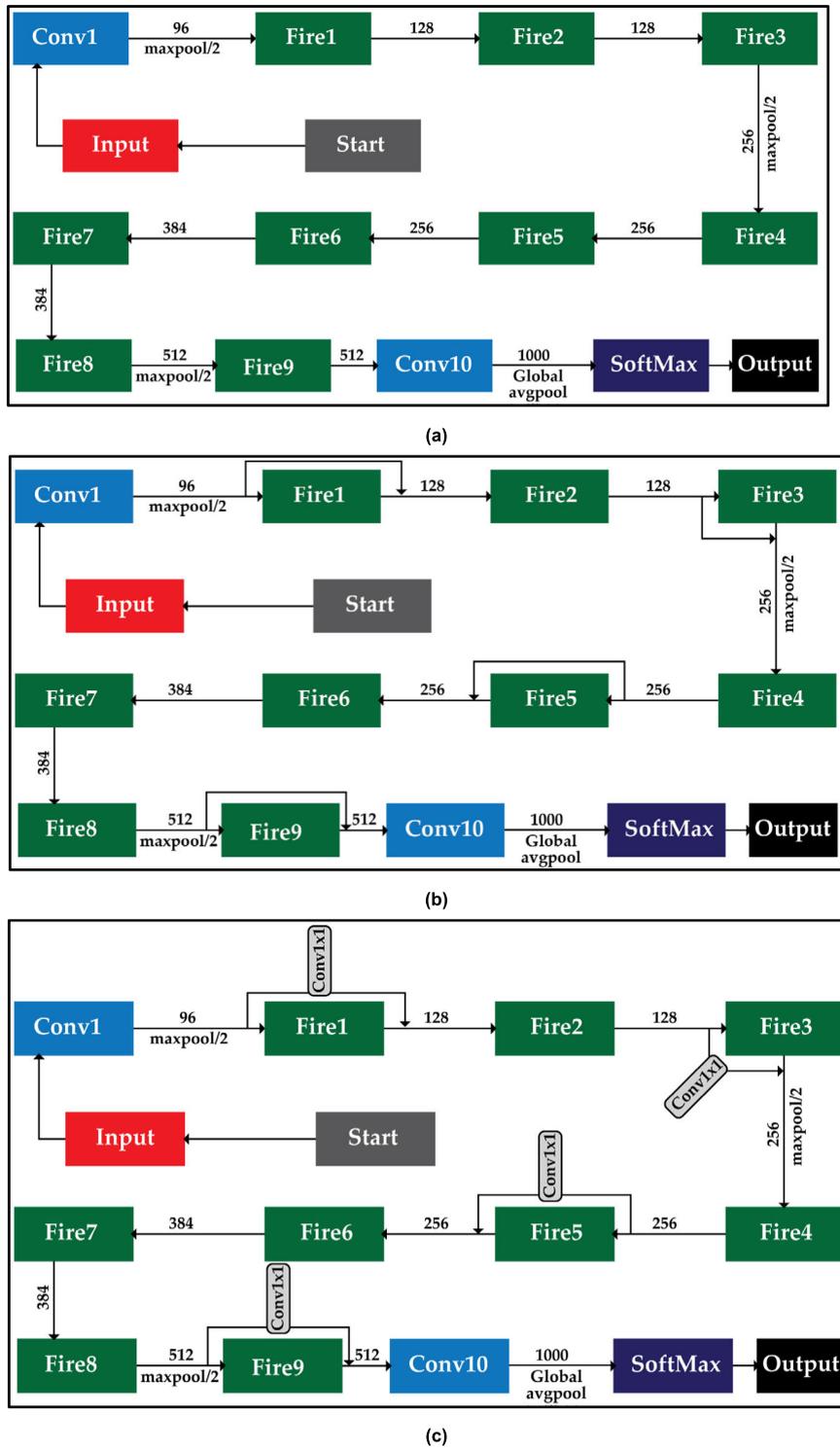
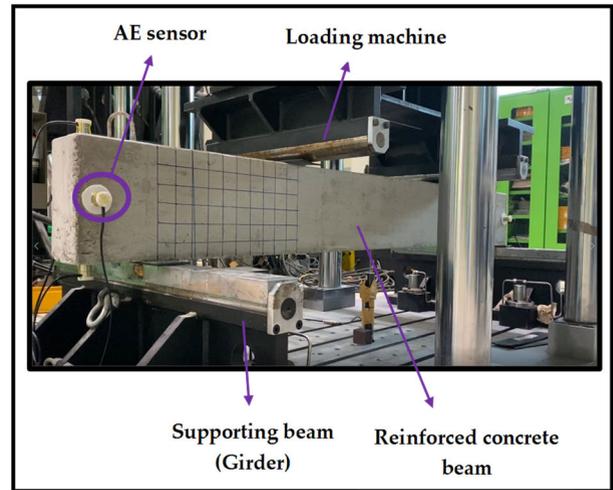
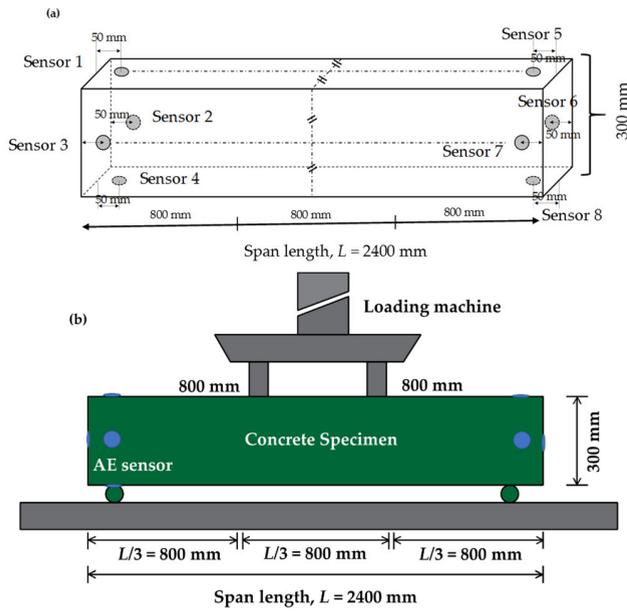


FIGURE 6. SqueezeNet architecture: (a) no bypass, (b) simple bypass, and (c) complex bypass.

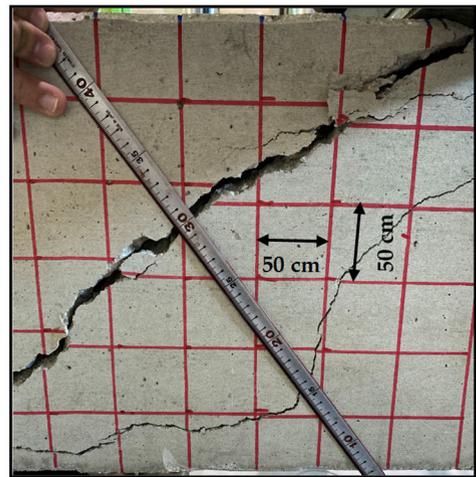
data is good enough for building a concrete crack detection system. Furthermore, our modified SqueezeNet classifier, named SQN-MF, is presented in this section too, followed by a step-by-step description of the entire crack detection method.

A. EXPERIMENTAL TEST BED AND AE DATA COLLECTION

Our testbed was developed at the Ulsan Industrial Artificial Intelligence (UIAI) laboratory in Ulsan, Republic of Korea for performing AE data acquisition. To assess the tensile strength of RC beams, flexural tests are very common. This



(a)



(b)

FIGURE 7. Schematic of the test setup: (a) sensor placement, (b) experimental setup for three-point bending.

kind of test can be conducted to determine the capability of the RC beams to tolerate cracks and fractures that take place because of the bending process. As for the bending test, we can go for four point-bending, three-point bending, or center point bending, the choice of which is mainly dependent on the length of the concrete beam we are using. The three-point bending test can imply half of the total load applied at every one third portion of the beam. On the contrary, the full load is applied at the center point of the beam in center-point bending. As per the standard set by the American society for testing and materials (ASTM), we should choose the length of the RC beam carefully so that it is at least three times the depth [54]. The RC beams used in our test setup had a depth of 300 mm. Conforming to the standard set by ASTM, we have chosen the length of our beams to be 2,400 mm, which is quite large. Considering this large size, we have opted for three-point bending rather than center-point bending. This ensures wide distribution of cracks throughout the entire beam.

Fig. 7 and Fig. 8 represent the experimental test bed for AE data acquisition from the concrete test specimen. The sensor placement and schematic diagram of the performed three-point bending test can be viewed in Fig. 7. Fig. 8 shows the real time setup built at the UIAI laboratory. As the length of the beam was 2,400 mm and the depth was 300 mm, the loads were applied at two different points in the beam, 800 mm apart from one another. The load was applied for a certain amount of time for crack initiation. The load velocity was 2mm/s.

Three different types of AE sensors were used to collect AE data: R3I [55], R15I [56], and WD [57]. For each type of sensor, the three-point bending test was performed three times

FIGURE 8. Real time test setup: (a) the whole setup, (b) experimental setup for three-point bending.

resulting in nine different tests in total. TABLE 1 describes the specifications of the mentioned sensors.

For the reproducibility test, we have conducted all the experiments with different sensors in the same environment. To decide which sensor has the highest sensitivity, sensor mounting was done. We used the pencil lead fracture (PLF) test to imitate the acoustic wave on the surface specimen. To do this, a magnetic pencil with a Teflon shoe was used. Readers are requested to go through [54] if they are interested in the details of this process. Fig. 9 presents us with a pictorial view of the method regarding how to handle a magnetic pencil on the concrete surface. An angle of 30° was used (angle from the plane of the specimen surface). There are two ways to decide if the sensors are significantly coupled, if PLF generates a high amplitude of 99 dB or we have a sensitivity within a ±3 dB margin. Via this process, we choose our

TABLE 1. Name of the sensors and their specifications.

Sensors	Specifications
 R31	Peak sensitivity: 120 dB Operating frequency range: 10-40 kHz Resonant frequency: 25 kHz Temperature range: -35 to 75°C
 R15I	Peak sensitivity: 109 dB Operating frequency range: 50-400 kHz Resonant frequency: 75 kHz Temperature range: -35 to 75°C
 WD	Peak sensitivity: 55 dB Operating frequency range: 100-900 kHz Resonant frequency: 125 kHz Temperature range: -65 to 175°C

TABLE 2. Experimental parameters and specifications.

Specification	Value
Number of sensors used (per experiment)	8
Used sensors' type	R31, WD and R15I AE sensors
Chosen sensor	R31
Final number of beams used	9
Duration of signal acquisition	Around 15 minutes
Concrete type	24 MPa
Reinforcement (Steel bars)	Korean Standard Reinforced Steel Bar, D16 (SD400)
Load (Initial and final)	1.03 kN, 111.46 kN
Load velocity	2 mm/second
Displacement	In-plane
Measurement of the displacement	Mid-span

adaptive threshold was used. This threshold was dependent on the maximum amplitude of the acquired signals.

The acquired signals from the three-point bending test can be divided into three types. The following figure (Fig. 10) presents those signals for different conditions. Fig. 10(a) is the plot representing the normal condition in the beam (with no crack). Fig. 10(b) and Fig. 10(c) represent the AE signals for micro and macro-cracks, respectively. In our study, we often refer to macro-cracks as fractures. We used a linear variable differential transformer (LVDT) to measure the in-plane distance of the cracks at the mid span. The LVDT was fixed under the concrete specimen. A concrete beam with fractures can be viewed from Fig. 11. The fractures in the beam took place after a gradually increasing load was applied to the beam using a loading machine. A summary of the experimental specifications is presented in TABLE 2.

B. DATASET VALIDATION

Before using a certain AE based data set, we need to make sure the data has been collected properly and it follows some of the inert properties of fracture mechanics.

One of the main sources of AE is plastic deformation. Plastic deformation refers to the state of a material when it cannot return back to its original shape when applied force is removed. The starting point of the plasticity, specially at or near the yield stress, results in the highest level of AE. This can be observed on a curve of stress versus strain. Our stress versus strain plot derived from the experimental setup follows the standard one as described in [58]. Fig. 12 presents us with a typical standard stress-strain curve where we can particularly observe the changes in the curve at the yield strength and the ultimate strength points. A similar type of curve obtained from our dataset can be viewed in Fig. 13.

According to the model presented by Palmer *et al.* [59] that makes use of linear elastic fracture mechanics for analysis, the total AE count is proportional to the area of the elastic-to-plastic boundary ahead of the crack. Therefore, AE is linked with the discontinuity or applied stress at the fracture. Based

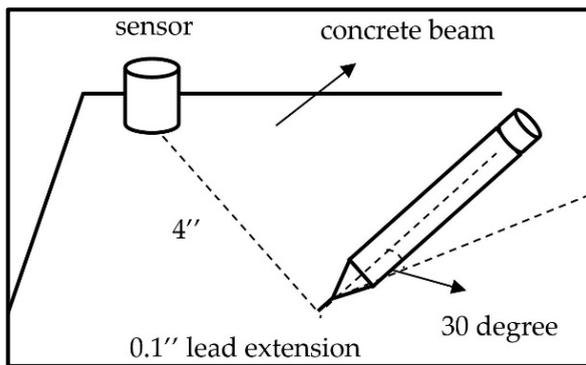


FIGURE 9. Pencil lead fracture test to select the appropriate sensor.

required sensor to be R31. The total number of sensors used in each experiment was eight.

The RC beams we used had six dimensions. We tried to cover all the surfaces using eight sensors. Glue gel and mounting tape were used to attach the sensors in the surface area. We put the sensors as distant as possible from the loading point to avoid any kind of problem regarding sensor placement that may occur while cracks are propagating during flexural tests. The AE signal acquisition process was undertaken for around 15 minutes at a 10 MHz sampling rate. The R31 sensors used in the process had an integrated preamplifier that has a gain of 40 dB. We did not set any kind of AE threshold before conducting the experiments, rather an

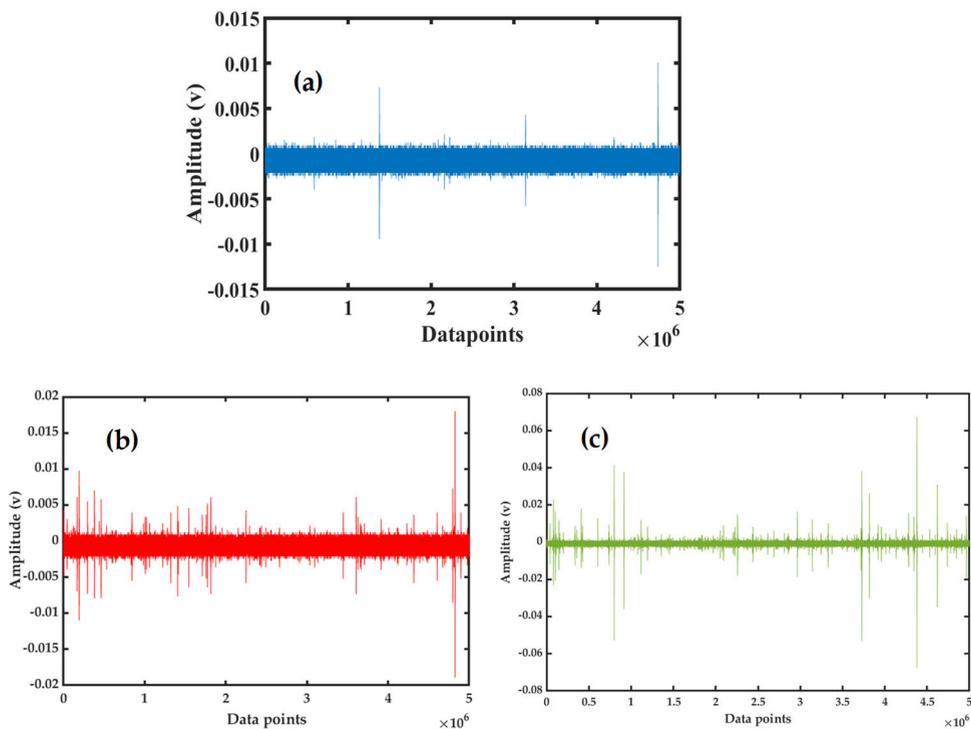


FIGURE 10. AE signals of variant conditions: (a) No crack (normal condition) (b) micro-cracks, and (c) fractures/macro-cracks.

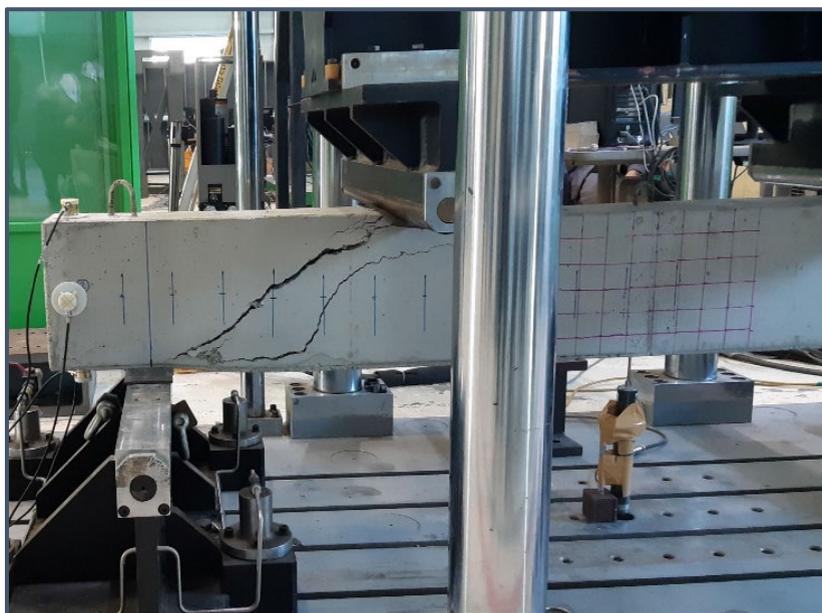


FIGURE 11. Fractures occurring in the beam due to an increasing load in the three-point bending test.

on the linear elastic fracture model, we can write:

$$N = D.S, \tag{10}$$

where D is the proportionality constant, N is the total acoustic emission count, and S is the plastic zone ahead of the crack. D is dependent on strain rate, thickness of the concrete spec-

imen, and temperature. The commencing crack length, and the applied stress (σ) vastly affect the plastic zone size. S can be written as:

$$S = C \left[\sec \left(\frac{\pi \sigma}{2\sigma_1} \right) - 1 \right], \tag{11}$$

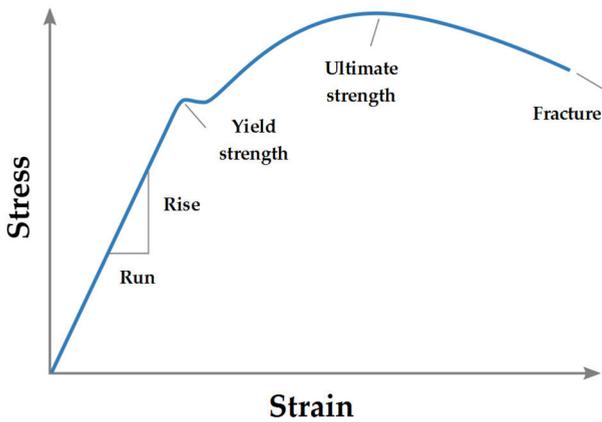


FIGURE 12. A typical stress-strain curve.

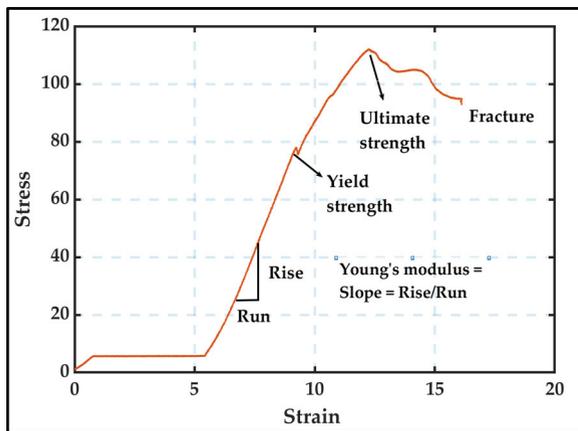


FIGURE 13. Stress-strain curve derived from our dataset.

where C is half of the crack length and σ_1 is the characteristic stress for the material. Since we are dealing with linear fracture mechanics, the characteristic stress for the material is equivalent to the yield stress. From (10) and (11), we can write N as:

$$N = DC \left[\sec \left(\frac{\pi \sigma}{2\sigma_1} \right) - 1 \right]. \quad (12)$$

According to [60], the fracture stress σ_f can be formulated as follows:

$$\sigma_f = \frac{2}{\pi} \sigma_1 \text{arc sec} \left[\exp \left(\frac{\pi K_{IC}^2}{8\sigma_1^2 C} \right) \right], \quad (13)$$

where K_I is the stress intensity factor. If we combine (12) and (13), we can obtain counts for failure N_f .

$$N_f = D.C \left[\exp \left(\frac{\pi K_{IC}^2}{8\sigma_1^2 C} \right) - 1 \right]. \quad (14)$$

For small stress, (14) can be reduced as follows:

$$N_f = D \frac{\pi K_{IC}^2}{8\sigma_1^2} \quad (15)$$

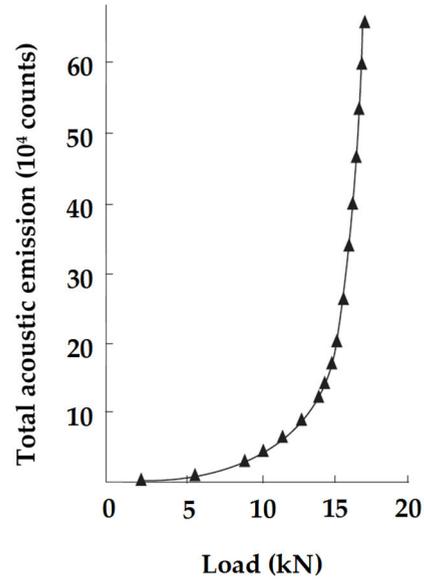


FIGURE 14. Reference Acoustic emission counts vs. load graph.

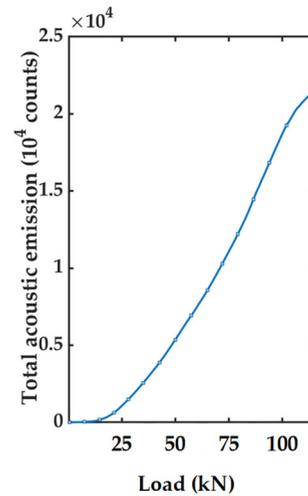


FIGURE 15. Acoustic emission counts vs. load graph derived from our own dataset.

Fig. 14 presents an example graph from [59] that shows us the acoustic emission count versus load obtained using the formulations made in this section. The same formulations have been used to create a similar type of plot using our dataset in Fig. 15. Like Fig. 14, as the load was increased, we found the AE counts increase as well.

C. MODIFIED SqueezeNet FOR BETTER ACCURACY

The classical version of SqueezeNet has a simple model that is built upon a basic building block named the fire module. As discussed in section III, convolution kernels small in size have been used to decrease the parameter size and memory demand in SqueezeNet while simultaneously maintaining good accuracy. The 3×3 filters used in the fire module produce a good number of parameters despite the fact

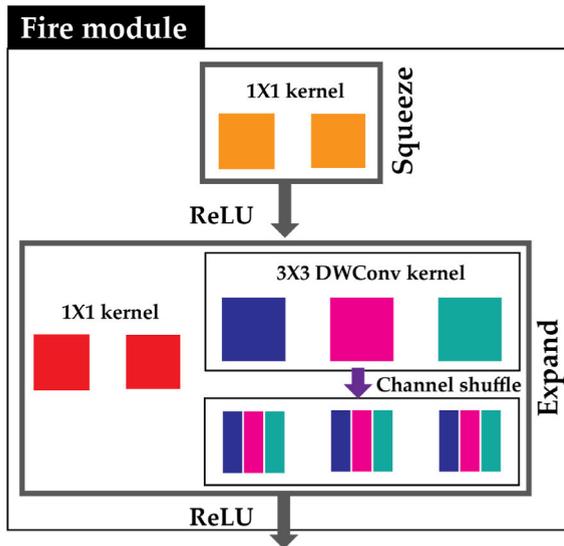


FIGURE 16. Modified fire module.

that fire modules perform some basic functions. To decrease the number of parameters and enhance the learning ability, we designed a new fire module for our SqueezeNet classifier (SQN-MF). Fig. 16 presents our modified fire module as described in this section. The modified fire module contains the typical structure of the original one with the 1×1 kernel that gets trained alongside the original 3×3 filters, but a change has been made in the expand layer. The typical 3×3 filters are replaced by a 3×3 DWConv kernel and a channel shuffle operation. The reasons behind choosing the DWConv and channel shuffling are:

- The depth-wise separable convolution is a divided convolution that divides a standard convolution into a pointwise and depth-wise convolution of 1×1 kernel size [61]. Using this, it is possible to decrease the complex nature of the network, while at the same time we can maintain fair precision by keeping a low number of training weight parameters. Furthermore, depth-wise separable convolution can facilitate us with separate channels from the convolution region. Also, it can connect the input and output feature maps one-to-one through a convolution operation.
- The channel shuffle can address the fact that the output channel is solely related to its corresponding input. It allows information exchange among channels and can facilitate with enhanced feature description. This process of channel shuffling divides the channels in each group into a couple of sub-groups and feeds every single group with variant sub-groups in the next layer. This kind of information exchange among groups of channels can improve accuracy.

D. PROPOSED METHODOLOGY FOR CRACK CLASSIFICATION

The proposed methodology consists of building a 2D image dataset based on the three different conditions (normal/

micro-cracks/ macro-cracks) and precise crack type classification using our modified SqueezeNet classifier. It can be described in the following steps:

1. AE signals are collected from the experimental setup where the RC beams are subjected to three-point bending tests.
2. The acquired AE signals can be divided into three types as described in Fig. 10. At this step, the CWT of the respective AE signals representative of the normal condition, micro-cracks and macro-cracks are obtained.
3. To reduce the computation overhead for our classifier, 2D color images are changed to gray-scale images. A full-scale image dataset is built using 900 images in total, each class containing 300 images.
4. Next, we implement our modified SqueezeNet classifier for distinguishing three different conditions associated with the RC beam under flexural tests.

Fig. 17 presents an elaborate pictorial view of the crack classification process as described in the above steps.

V. EXPERIMENTAL RESULTS AND DISCUSSION

The proposed concrete crack detection framework (CWT + SQN-MF) is applied on the AE data gathered from the three-point bending test. In this study, AE signal acquisition is performed from our self-designed testbed with a gradually increasing load from the loading machine, increased up to 111.46 kN. Based on our visual inspection and time domain feature analysis, we could divide the whole duration of the test into three parts. The first one is the normal condition in which no crack on the RC beam was observed. The second division corresponds to the occurring micro-cracks in the RC beam due to the gradually increasing load. Lastly, the third one consists of signals representing macro-cracks/ fractures. We have taken 300 samples from each type and built our image dataset applying CWT. A well-balanced dataset has been formed for our analysis. Fig. 18 represents the original time-domain AE signals and their corresponding 2D images derived using CWT. In the time-domain signals, we can see that it is not only very difficult, but almost impossible to get insights on three different conditions (normal, micro-cracks, and macro-cracks) due to too much noise and fluctuating points. However, the images obtained using CWT, have distinguishable patterns for the mentioned conditions.

To obtain the fault features visually, we applied CWT, but still the 2D color images do not have significant meaning, which is associated with the change of amplitude of the signal along with the time. That is why we have taken the contour plot of the CWT images acquired for different crack conditions of the concrete specimen. Fig. 19 presents the contour plot, which provides better visualization of the changes in pattern for different conditions. From both Fig. 18 and 19, we can realize that the acquired CWT images have distinguishable patterns across different time-frequency scales for each crack condition.

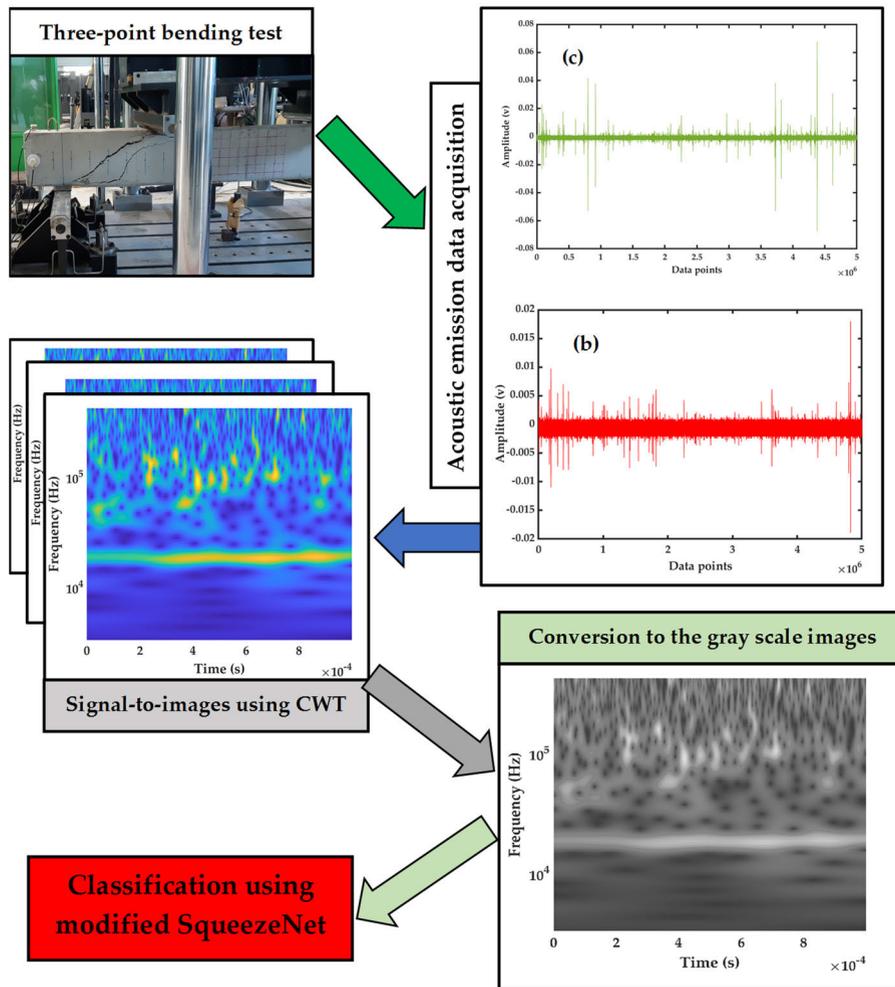


FIGURE 17. Pictorial view of the proposed methodology.

The main goal of our study is to prove our proposed method in three aspects. First, we would like to compare the proposed method with two recently proposed AE feature based typical ML algorithm-oriented crack detection methods [24], [62]. Second, we applied crack classification tasks with multiple variants of SqueezeNet along with the typical one. To prove the superiority of the proposed SQN-MF, we will present a performance comparison of our method against these versions of SqueezeNet and two other popular deep learning architectures. Lastly, we use the same SQN-MF classifier with the 2D wavelet images acquired from three different sensors so that we can prove that our method is robust enough to acquire good accuracy even though the data gathered from the R15I and WD sensors are not suitable enough.

A. COMPARISON OF THE PROPOSED METHOD WITH THE TYPICAL AE FEATURE BASED ML ALGORITHMS

Until today, the existing literature regarding concrete crack detection methods are ruled by the crucial AE feature extraction methods, and often researchers have associated conven-

tional machine learning algorithms for crack identification. The process is expensive and tedious for many reasons. Not only do we have to extract a good number of AE features (requires extensive amount of time and human expertise), but also, if the number of the features are high, we never know which features are truly useful. Alongside a machine learning algorithm, we also need a feature selection algorithm. Our first approach is to compare our study with the mentioned method. Fig. 20 presents us with some of the typical AE features described in section III(B). We used the data obtained from the AE signals acquired from our testbed (R3I sensor).

As we can see from the above figure, we need a significant amount of human expertise to comprehend the inert meaning of the features such as when the incipient crack is taking place, when exactly the fracture is taking place, etc. Not only this, but we also implied a wrapper based feature selection algorithm named Boruta [63] to rank the features before passing them through a classifier, since bad features can end up giving us poor accuracy. Fig. 21 presents the feature ranking obtained using Boruta.

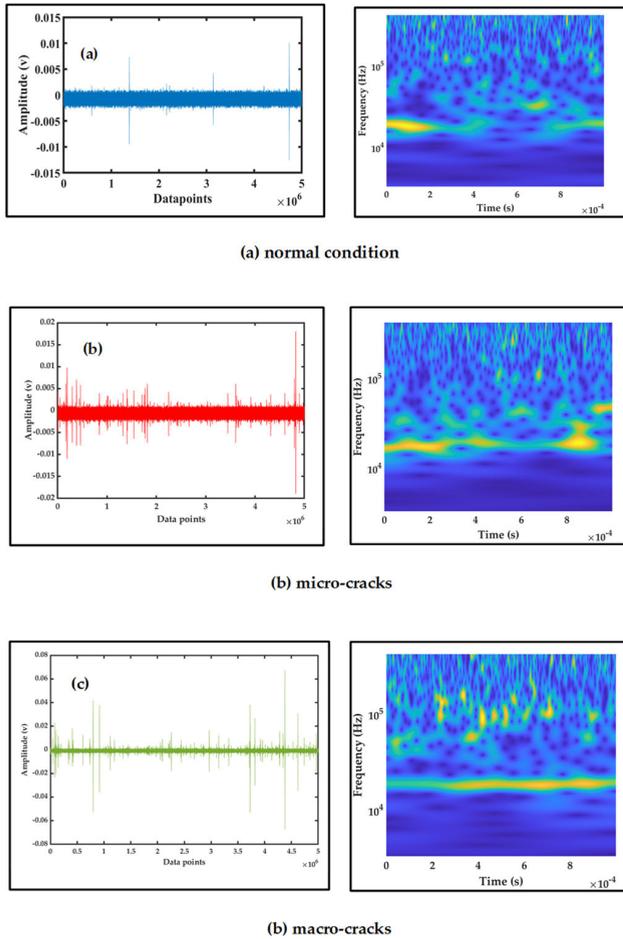


FIGURE 18. Time domain acoustic emission signals and their corresponding CWT images.

As we can see from Fig. 21, we have to put even more effort towards choosing which features to use. This process adds more complexity. In many cases, the k-nearest neighbor algorithm (k-NN) has been used in the literature for crack type detection. We have done the same for our acquired AE data and compared the performance of our algorithm with k-NN based fault detection models in terms of average classification accuracy and F1-score. For the k-NN algorithm used in [24], the data set consists of direct feature values like AE counts, decay time, etc. The number of samples for each class: normal, micro-cracks, and macro-cracks were 300 in number. The crack detection method (k-NN+CFAR) proposed in [62] used 1,030 feature vectors representing different conditions of the RC beam. However, to have a balanced comparison we took 900 samples for this method too. In both cases (k-NN+AE features and k-NN+CFAR), datasets consisted of 1D feature values processed from the raw AE signals. On the other hand, we used the CWT image dataset (total of 900 images, 300 for each class type) for our proposed SQN-MF. For all three cases, the datasets were divided into two parts: 70% of the data for training and the remaining 30% for testing. Description of the datasets used

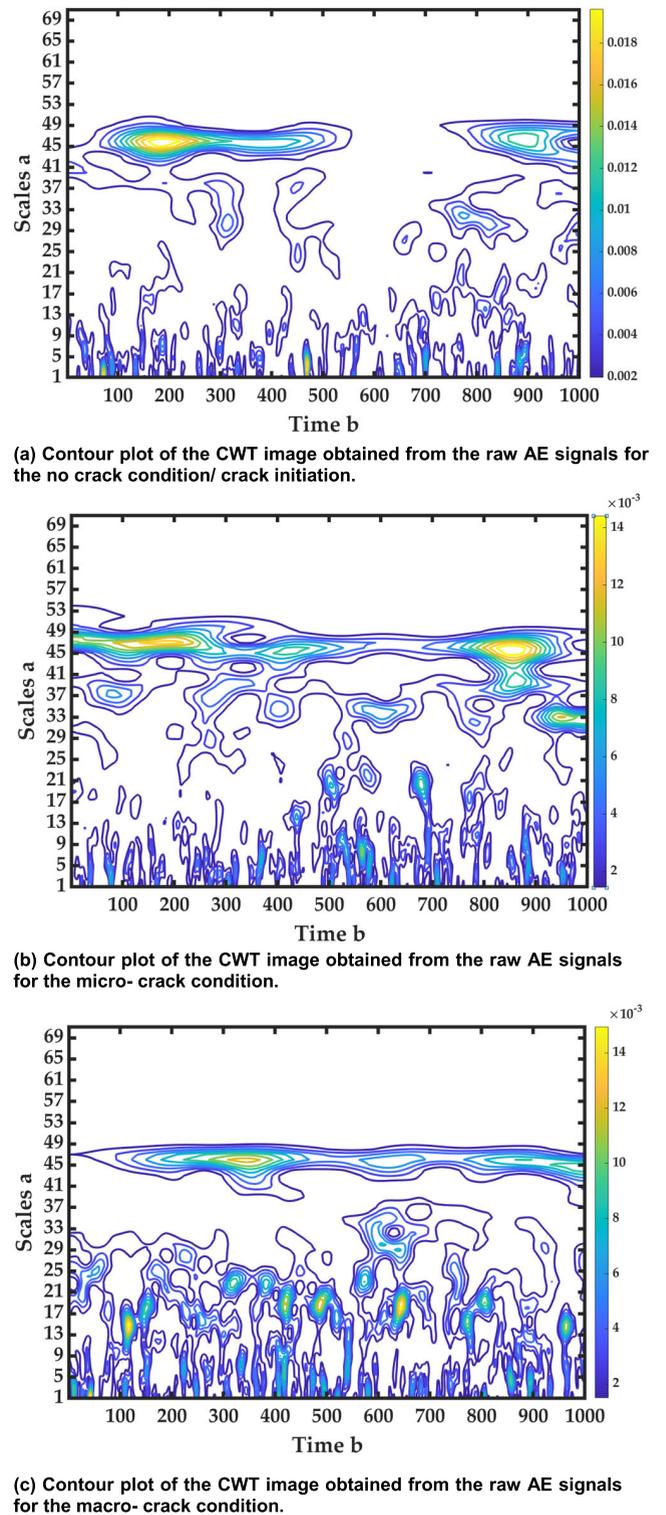


FIGURE 19. Contour plots of the three types of beam condition: (a) normal (b) micro-crack, and (c) macro-crack.

for the k-NN, and deep learning-based methods are presented in TABLE 3 and 4. We used two parameters for performance comparison: average classification accuracy and F1-score.

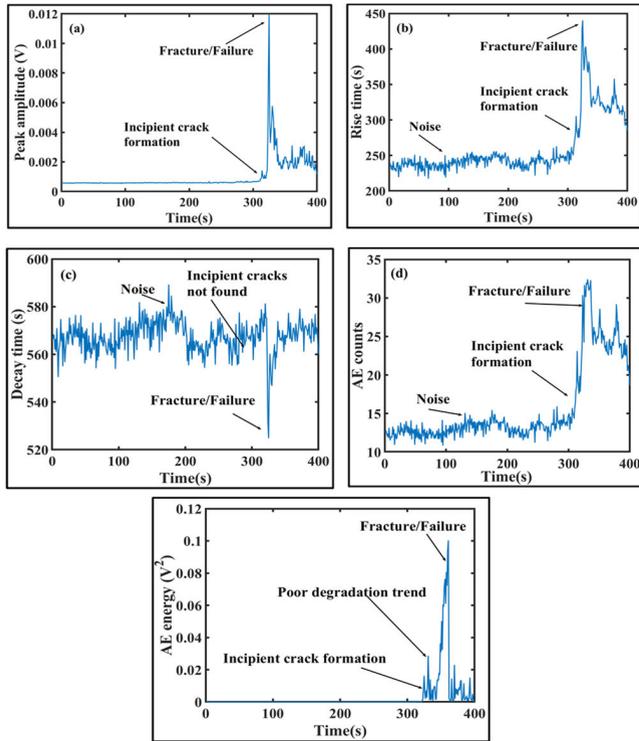


FIGURE 20. Time history of different AE features: (a) peak amplitude (b) rise time (c) decay time (d) AE counts (e) AE energy.

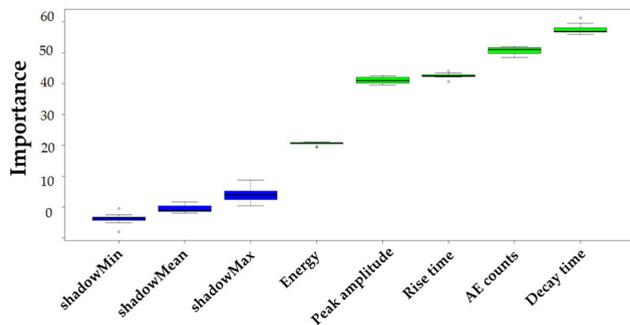


FIGURE 21. Feature ranking using Boruta.

F1-score can be calculated as follows:

$$F1 - score = 2 \times \frac{Recall \times Precision}{Recall + Precision}, \quad (16)$$

where precision is the ratio of correctly predicted positive observations to the total predicted positive observations, and recall is the ratio of correctly predicted positive observations to all observations.

Fig. 22 presents us with the confusion matrices of the k-NN based methods and our proposed method. The k-NN along with the feature selection algorithm (Fig. 22(a)) performs very poorly. Many of the samples corresponding to the normal condition have been misclassified as micro-cracks. Micro-cracks have been detected as macro-cracks as well, and vice versa. The k-NN along with the CFAR algorithm (Fig. 22(b)) has misclassified macro-cracks too but performs better than

TABLE 3. Dataset description of the methods implied in this study.

Methods	Total number of samples	Number of samples for training	Number of samples for testing	Data type
k-NN based methods	900	630	270	1D feature vector
Deep learning-based methods	900	630	270	2D gray scale images

TABLE 4. Number of samples for each class in different methods.

Methods	Number of classes	Class 1: Normal	Class 2: Micro	Class 3: Macro
k-NN based methods	3	300	300	300
Deep learning-based methods	3	300	300	300

TABLE 5. Performance comparison (k-NN based methods vs. the proposed method).

Methods	Accuracy (%)	F1-score
k-NN + selected AE features	70.5	55.8
k-NN + CFAR	79.2	68.9
Proposed method	100	100

the previous method. The proposed method in this study can classify the normal condition, micro-cracks, and macro-cracks perfectly with 100 percent accuracy. From Fig. 22 (c), it is visible that none of the samples have been misclassified. TABLE 5 presents the performance comparison of the three methods in terms of average classification accuracy and F1-score. It can be clearly seen that the proposed method's accuracy and F1-score are way higher than the typical k-NN based AE feature-oriented methods.

B. PERFORMANCE COMPARISON WITH THE OTHER VARIANTS OF THE SqueezeNet

We have implemented SQN-1, SQN-2, and SQN-3 to classify our CWT images obtained for different fault conditions. As stated in the main paper by the authors of SqueezeNet, SQN-2 performs better than the typical one and SQN-3. Therefore, we decided to compare our proposed SQN-MF with the version of the typical SqueezeNet with simple bypass connections, which is SQN-2. To have a comparative view of the performance of these two methods, we have put the accuracy and loss curves in Fig. 23. From Fig. 23(a) and (b), we can see that the accuracy for SQN-2 is 96% and the lowest amount of loss is 0.0566. The performance is quite good, but considering our application, i.e., building a robust crack characterization method, even such high accuracy is

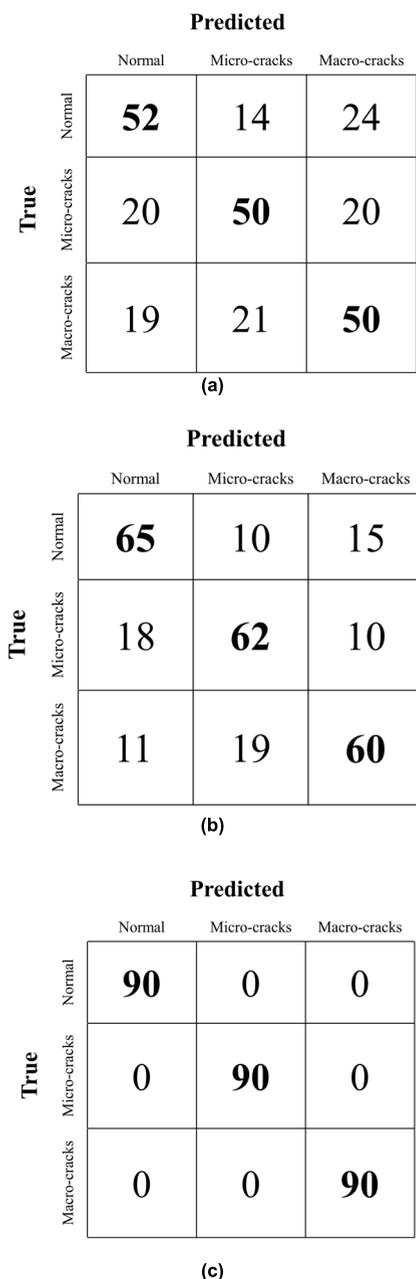


FIGURE 22. Confusion matrices: (a) k-NN + feature selection algorithm, (b) k-NN + CIFAR, and (c) Modified SqueezeNet algorithm (SQN-MF).

not adequate as it involves huge financial aspects and human lives.

Fig. 23(c) and (d) presents the loss and accuracy curve for our proposed SQN-MF. We can see that it has 100% accuracy and almost zero loss. It can be concluded that SQN-MF performs better than SQN-2. The achieved 100% accuracy is very robust and highly necessary for a critical application like concrete crack detection.

C. EVALUATION OF THE PROSPED METHOD WITH THE OTHER DATASETS

We used three different sensors (R3I, R15I, and WD) for AE data collection. Via a PLF test, we chose R3I sensors. This

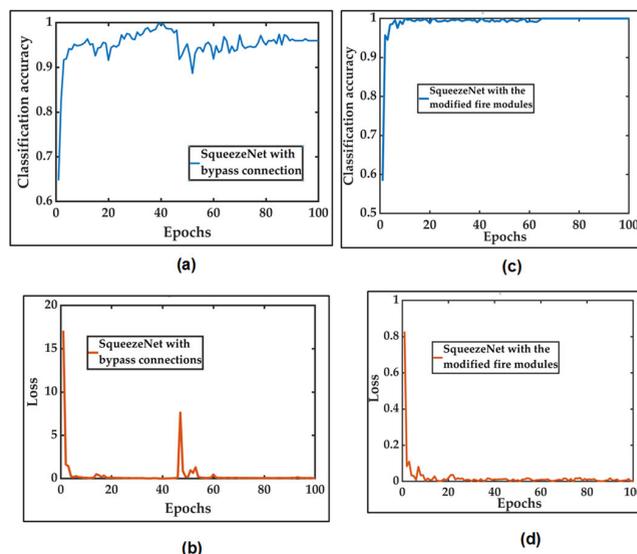


FIGURE 23. (a) accuracy curve for SqueezeNet with bypass connection, (b) loss curve for SqueezeNet with bypass connection, (c) accuracy curve for SqueezeNet with the modified fire modules, and (d) loss curve for SqueezeNet with the modified fire modules.

particular sensor is usually chosen for structural health monitoring ranging from small to medium concrete and geological structures. It can reject high background noise, and the distances between sensors can be relatively close. The proposed method in our study has gained 100% accuracy using the wavelet image dataset acquired from the AE signals collected from this very sensor. However, we wanted to apply SQN-MF to the dataset obtained from the other two sensors too. In a similar way, we first took the CWT of the AE signals corresponding to R15I and WD sensors and classified them with the SQN-MF with modified fire modules. It was expected that we would not get very good results, but the results were not as bad as expected. For R15I and WD we achieved 87.2% and 69.3% classification accuracy respectively, which are still decent as these two sensors are not usually used for structural health monitoring. R15I sensors are extensively used in monitoring structures like pipelines, vessels, bridges, and storage tanks. WD sensors are general-purpose sensors that are useful where frequency analysis of the AE signal is needed in determining the predominant frequency band of AE sources for noise discrimination. Comparative analysis of the results not only proves the effectiveness of the sensor selection method we used (PLF test), but also it shows the robustness of our proposed method since it gains a fair amount of accuracy in poor data sets too. Fig. 24 presents us with the confusion matrices of the applied method proposed in our study for the three different sensors.

From Fig. 24, we can see that not a single sample is misclassified by the proposed method when it uses the data obtained from the R3I sensor. For the other two cases, there are a good number of samples that were misclassified. In some cases, the proposed method could not distinguish between macro and micro-cracks when the data obtained

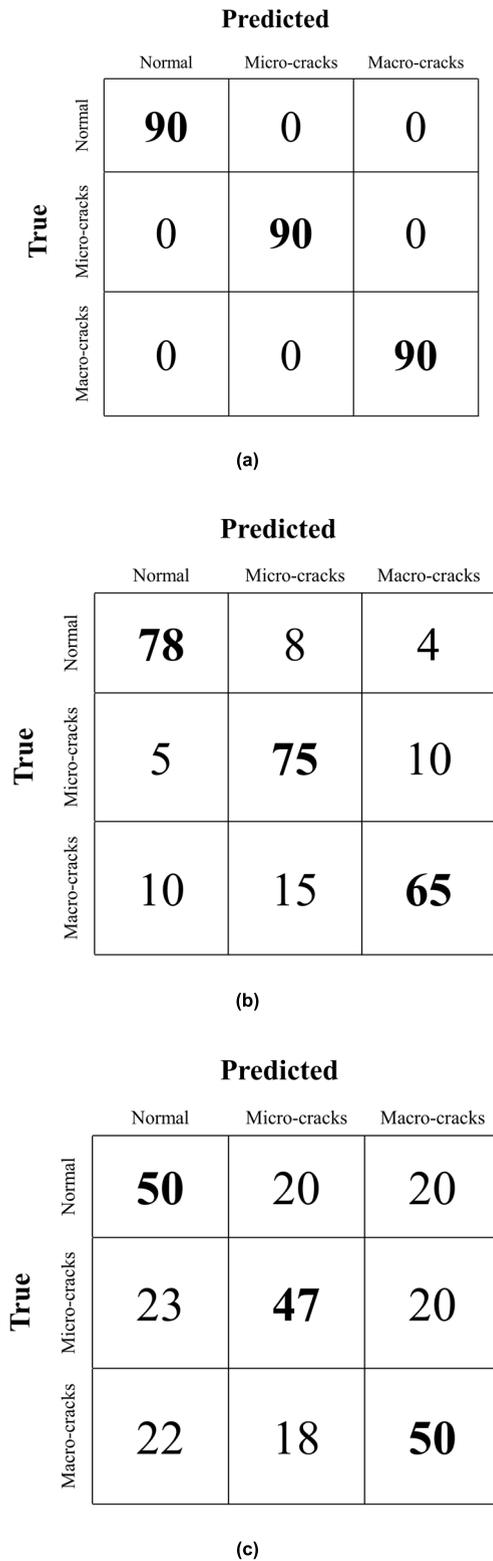


FIGURE 24. Confusion matrices: (a) R31 sensor, (b) R15I sensor and (c) WD sensor.

from the other two sensors was used. Still, for both cases, the proposed method can gain at least 70% accuracy. The comparison of the performance of the proposed method can be better understood from the following table.

TABLE 6. Performance comparison of the proposed method getting applied for different sensor (R31, R15I, and WD) data.

Used sensor	Accuracy (%)	F1-score
WD	69.3	53.9
R15I	87.2	80.7
Proposed method	100	100

TABLE 7. Performance comparison of the proposed method getting applied for different sensor (R31, R15I, and WD) data.

Deep learning architecture	Accuracy (%)	F1-score
AlexNet	82.6	73.9
VGG16	91.1	86.7
Proposed method	100	100

D. COMPARISON OF THE PROPOSED METHOD WITH OTHER DEEP LEARNING ARCHITECTURES

We have considered other deep learning architectures to compare our proposed method with. The main paper of SqueezeNet has extensively claimed that they can reach AlexNet [64] level accuracy with lesser parameters and only with a model size of 0.5MB. We also wanted to compare our proposed method with AlexNet to see if we could reach the desired level of accuracy. We have used the same data set built using the 2D wavelet images acquired from the time domain signals and applied AlexNet. We have achieved accuracy similar to [26] for AlexNet which is 82.6%. We have used another deep learning architecture randomly which is VGG16 [65]. It gained 91.1% accuracy. Fig. 25 presents the confusion matrices derived for the three classes in our system. TABLE 7 presents us with the performance comparison among the proposed method, VGG16 and AlexNet.

E. DISCUSSION

k-NN based algorithms we have implemented are based on 1D feature vectors. Time domain AE signals were used to extract crucial features. Even though we have tried our best to collect the data as precisely as possible, feature plots of these vectors had some outliers which were difficult to overcome. This is the reason why it is not possible to get high accuracy using the k-NN algorithms. Fig. 26 presents the plot of the feature vectors for three different conditions. As we can see that some feature points are not exactly in the coordinates that they are supposed to be which has led to the misclassification of some of the samples.

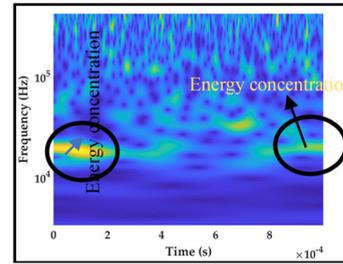
The 2D CWT images acquired using continuous wavelet transform have been used for the deep learning architectures. CWT images have distinguishable patterns across different time-frequency scales for each type of condition: normal, micro-cracks, and macro-cracks. For no crack, which is the normal condition, we can see that energy of the AE signals is slightly concentrated at the two-opposite end of the image. For the images representing micro-cracks have deep energy concentration at the two opposite ends, however the pattern starts to fade as we move towards center. For the images

		Predicted		
		Normal	Micro-cracks	Macro-cracks
True	Normal	70	10	10
	Micro-cracks	15	67	8
	Macro-cracks	10	20	70

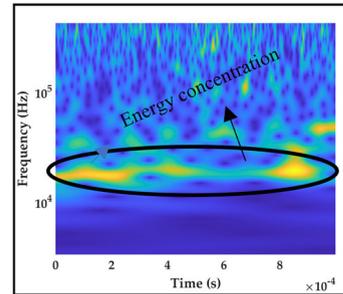
(a)

		Predicted		
		Normal	Micro-cracks	Macro-cracks
True	Normal	80	6	4
	Micro-cracks	7	79	4
	Macro-cracks	5	10	75

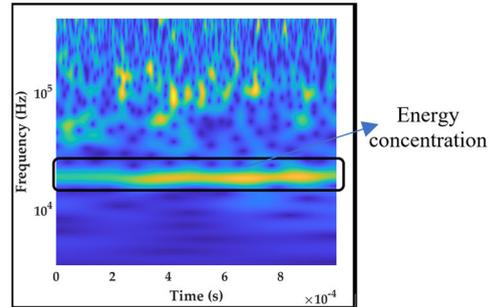
(b)



(a)



(b)



(c)

FIGURE 25. Confusion matrices: (a) AlexNet and (b) VGG16.

FIGURE 27. Variation of the energy distribution in wavelet images: (a) energy concentration at the ends, (b) energy concentration at the end spreading towards center, (c) energy concentration almost creating a distinguishable prominent line at the center.

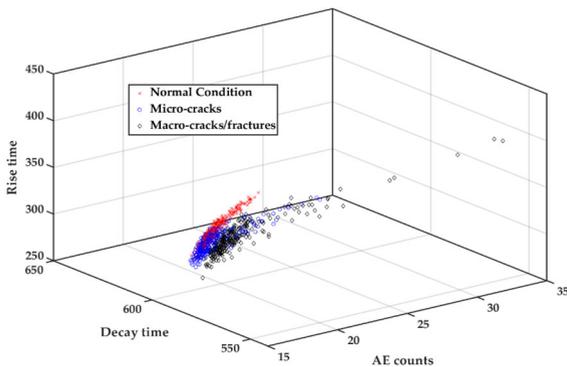


FIGURE 26. 3D plot of the feature space associated with the k-nearest neighbor algorithms.

associated with the macro-cracks, we have deep energy concentration and prominent pattern which almost forms a line at the center of the image. Such distinguishable patterns are very much feasible to be used for classification by deep learning algorithms. Fig. 27 shows us the variant energy concentration of the wavelet images vividly.

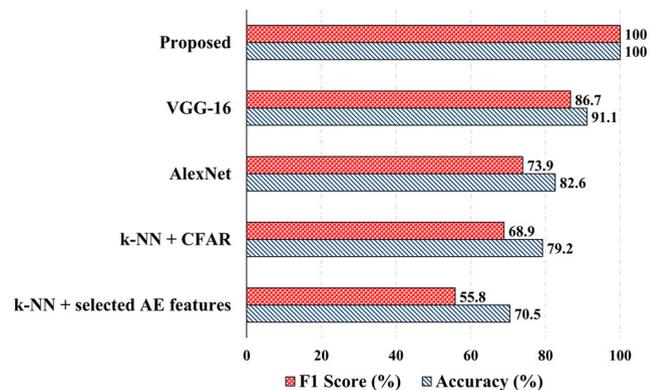


FIGURE 28. Performance comparison of all the algorithms in one figure.

Two primary reasons behind SQN-MF (the proposed method) getting such high are the introduction of the depth-wise convolution and channel shuffle operation. The depth-wise separable convolution is a divided convolution

that divides a standard convolution into a pointwise and depth-wise convolution of 1×1 kernel size [61]. Using this, it was possible to decrease the complex nature of the network, while at the same time we could maintain fair precision by keeping a low number of training weight parameters. Furthermore, depth-wise separable convolution facilitated us with separate channels from the convolution region. Also, it could connect the input and output feature maps one-to-one through a convolution operation. The channel shuffle can address the fact that the output channel is solely related to its corresponding input. It allowed information exchange among channels and could facilitate with enhanced feature description. This process of channel shuffling was used to divide the channels in each group into a couple of sub-groups and fed every single group with variant sub-groups in the next layer. This kind of information exchange among groups of channels improved accuracy. Fig. 28 summarizes the performance comparison of all the algorithms including the proposed one in one figure. It can be observed that the proposed method has vividly outsmarted the others.

VI. CONCLUSION

This paper introduced a continuous wavelet transform (CWT) based imaging technique accompanied by SqueezeNet with modified fire modules (SQN-MF) to monitor the crack formation process in reinforced concrete (RC) beams. First, acoustic emission (AE) signals were acquired from the three-point bending test. After that, using CWT we built an image-based dataset corresponding to the collected AE signals. The proposed lightweight SQN-MF contains our newly designed fire module with a depth-wise convolutional kernel and a channel shuffle operation that reduces the number of parameters and increases the learning ability. It makes use of the CWT based grayscale images representing three different conditions: normal, micro-crack, and macro-crack, with an overall 100% average classification accuracy. It is 20.8% higher than the traditional AE feature-based machine learning models. The proposed SQN-MF outsmarts the typical SqueezeNet (SQN-1) and a variant of it with simple bypass connections (SQN-2). Lastly, SQN-MF achieves a fair amount of classification accuracy for the data collected using R15I and WD sensors (87.2 and 69.3), even though these sensors are not used for structural health monitoring and are supposed to provide us with poor data points. The proposed study is unique because it not only is the first attempt to implement an advanced and sophisticated deep learning algorithm, but also, we make use of the lightweight architecture of the SqueezeNet (< 0.5 MB), which paves the way for developing FPGA based systems that are impossible to imagine for other deep learning architectures due to their overwhelming size. The small size of the model will facilitate us with less training time and communication overhead that will result in easier distributed training. Additionally, the proposed model will have significant advantages for any kind of future autonomous embedded system designed to detect

concrete cracks, since the size of the model will allow for faster exporting from the cloud.

REFERENCES

- [1] A. K. Das, D. Suthar, and C. K. Y. Leung, "Machine learning based crack mode classification from unlabeled acoustic emission waveform features," *Cement Concrete Res.*, vol. 121, pp. 42–57, Jul. 2019, doi: [10.1016/j.cemconres.2019.03.001](https://doi.org/10.1016/j.cemconres.2019.03.001).
- [2] X. Yu, M. Bentahar, C. Mechri, and S. Montrésor, "Passive monitoring of nonlinear relaxation of cracked polymer concrete samples using acoustic emission," *J. Acoust. Soc. Amer.*, vol. 146, no. 4, pp. 323–328, 2019, doi: [10.1121/1.5127519](https://doi.org/10.1121/1.5127519).
- [3] D. G. Aggelis, S. De Sutter, S. Verbruggen, E. Tsangouri, and T. Tysmans, "Acoustic emission characterization of damage sources of lightweight hybrid concrete beams," *Eng. Fract. Mech.*, vol. 210, pp. 181–188, Apr. 2019, doi: [10.1016/j.engfracmech.2018.04.019](https://doi.org/10.1016/j.engfracmech.2018.04.019).
- [4] D. G. Aggelis, "Classification of cracking mode in concrete by acoustic emission parameters," *Mech. Res. Commun.*, vol. 38, pp. 153–157, Apr. 2011, doi: [10.1016/j.mechrescom.2011.03.007](https://doi.org/10.1016/j.mechrescom.2011.03.007).
- [5] N. K. Banjara, S. Sasmal, and V. Srinivas, "Investigations on acoustic emission parameters during damage progression in shear deficient and GFRP strengthened reinforced concrete components," *Measurement*, vol. 137, pp. 501–514, Apr. 2019, doi: [10.1016/j.measurement.2019.01.099](https://doi.org/10.1016/j.measurement.2019.01.099).
- [6] E. Tsangouri, O. Remy, F. Boulpaep, S. Verbruggen, G. Livitsanos, and D. G. Aggelis, "Structural health assessment of prefabricated concrete elements using acoustic emission: Towards an optimized damage sensing tool," *Construct. Building Mater.*, vol. 206, pp. 261–269, May 2019, doi: [10.1016/j.conbuildmat.2019.02.035](https://doi.org/10.1016/j.conbuildmat.2019.02.035).
- [7] J. G. Yue, S. K. Kunnath, and Y. Xiao, "Uniaxial concrete tension damage evolution using acoustic emission monitoring," *Construct. Building Mater.*, vol. 232, Jan. 2020, Art. no. 117281, doi: [10.1016/j.conbuildmat.2019.117281](https://doi.org/10.1016/j.conbuildmat.2019.117281).
- [8] C. Chen, X. Fan, and X. Chen, "Experimental investigation of concrete fracture behavior with different loading rates based on acoustic emission," *Construct. Building Mater.*, vol. 237, Mar. 2020, Art. no. 117472, doi: [10.1016/j.conbuildmat.2019.117472](https://doi.org/10.1016/j.conbuildmat.2019.117472).
- [9] T.-C. Hou and J. P. Lynch, "Electrical impedance tomographic methods for sensing strain fields and crack damage in cementitious structures," *J. Intell. Mater. Syst. Struct.*, vol. 20, no. 11, pp. 1363–1379, Jul. 2009, doi: [10.1177/1045389X08096052](https://doi.org/10.1177/1045389X08096052).
- [10] M. A. A. Aldahdooh and N. M. Bunnori, "Crack classification in reinforced concrete beams with varying thicknesses by mean of acoustic emission signal features," *Construct. Building Mater.*, vol. 45, pp. 282–288, Aug. 2013, doi: [10.1016/j.conbuildmat.2013.03.090](https://doi.org/10.1016/j.conbuildmat.2013.03.090).
- [11] K. Otsuka and H. Date, "Otsuka2000.Pdf," *Eng. Fract. Mech.*, vol. 65, pp. 1–13, 2000.
- [12] S. J. Pantazopoulou and M. Zanganeh, "Assessing damage in corroded reinforced concrete using acoustic emission," *J. Mater. Civ. Eng.*, vol. 13, no. 5, pp. 340–348, 2001.
- [13] A. Carpinteri, G. Lacidogna, and M. Paggi, "Acoustic emission monitoring and numerical modeling of FRP delamination in RC beams with non-rectangular cross-section," *Mater. Struct.*, vol. 40, no. 6, pp. 553–566, Feb. 2007, doi: [10.1617/s11527-006-9162-4](https://doi.org/10.1617/s11527-006-9162-4).
- [14] S. K. Verma, S. S. Bhadauria, and S. Akhtar, "Review of nondestructive testing methods for condition monitoring of concrete structures," *J. Construct. Eng.*, vol. 2013, pp. 1–11, Apr. 2013, doi: [10.1155/2013/834572](https://doi.org/10.1155/2013/834572).
- [15] K. M. Holford, A. W. Davies, R. Pullin, and D. C. Carter, "Damage location in steel bridges by acoustic emission," *J. Intell. Mater. Syst. Struct.*, vol. 12, no. 8, pp. 567–576, 2001, doi: [10.1106/KDNY-AJOU-KP2B-P52R](https://doi.org/10.1106/KDNY-AJOU-KP2B-P52R).
- [16] D. G. Aggelis, D. V. Soulioti, N. Sapouridis, N. M. Barkoula, A. S. Paipetis, and T. E. Matikas, "Acoustic emission characterization of the fracture process in fibre reinforced concrete," *Construct. Building Mater.*, vol. 25, no. 11, pp. 4126–4131, Nov. 2011, doi: [10.1016/j.conbuildmat.2011.04.049](https://doi.org/10.1016/j.conbuildmat.2011.04.049).
- [17] M. K. ElBatanouny, A. Larosche, P. Mazzoleni, P. H. Ziehl, F. Matta, and E. Zappa, "Identification of cracking mechanisms in scaled FRP reinforced concrete beams using acoustic emission," *Experim. Mech.*, vol. 54, no. 1, pp. 69–82, Jan. 2014, doi: [10.1007/s11340-012-9692-3](https://doi.org/10.1007/s11340-012-9692-3).
- [18] P. G. Ranjith, D. Jasinge, J. Y. Song, and S. K. Choi, "A study of the effect of displacement rate and moisture content on the mechanical properties of concrete: Use of acoustic emission," *Mech. Mater.*, vol. 40, no. 6, pp. 453–469, Jun. 2008, doi: [10.1016/j.mechmat.2007.11.002](https://doi.org/10.1016/j.mechmat.2007.11.002).

- [19] K. Ohno and M. Ohtsu, "Crack classification in concrete based on acoustic emission," *Construct. Building Mater.*, vol. 24, no. 12, pp. 2339–2346, Dec. 2010, doi: [10.1016/j.conbuildmat.2010.05.004](https://doi.org/10.1016/j.conbuildmat.2010.05.004).
- [20] H.-D. Yun, W.-C. Choi, and S.-Y. Seo, "Acoustic emission activities and damage evaluation of reinforced concrete beams strengthened with CFRP sheets," *NDT E Int.*, vol. 43, no. 7, pp. 615–628, 2010, doi: [10.1016/j.ndteint.2010.06.006](https://doi.org/10.1016/j.ndteint.2010.06.006).
- [21] R. V. Sagar, B. K. R. Prasad, and R. Sharma, "Evaluation of damage in reinforced concrete bridge beams using acoustic emission technique," *Nondestruct. Test. Eval.*, vol. 27, no. 2, pp. 95–108, Jun. 2012, doi: [10.1080/10589759.2011.610452](https://doi.org/10.1080/10589759.2011.610452).
- [22] S. F. Huang, M. M. Li, D. Y. Xu, M. J. Zhou, S. H. Xie, and X. Cheng, "Investigation on a kind of embedded AE sensor for concrete health monitoring," *Res. Nondestruct. Eval.*, vol. 24, no. 4, pp. 202–210, Oct. 2013, doi: [10.1080/09349847.2013.789949](https://doi.org/10.1080/09349847.2013.789949).
- [23] M. A. Habib, A. Rai, and J.-M. Kim, "Performance degradation assessment of concrete beams based on acoustic emission burst features and Mahalanobis—Taguchi system," *Sensors*, vol. 20, no. 12, p. 3402, Jun. 2020, doi: [10.3390/s20123402](https://doi.org/10.3390/s20123402).
- [24] A. Habib, C. H. Kim, and J. M. Kim, "A crack characterization method for reinforced concrete beams using an acoustic emission technique," *Appl. Sci.*, vol. 10, no. 21, pp. 1–21, 2020, doi: [10.3390/app10217918](https://doi.org/10.3390/app10217918).
- [25] M. Yamada, *Wavelets: Applications*, J.-P. Françoise, G. L. Naber, and T. S. B. T.-E. M. P. Tsun, Eds. Oxford, U.K.: Academic, 2006, pp. 420–426.
- [26] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <1MB model size," 2016, *arXiv:1602.07360*. [Online]. Available: <http://arxiv.org/abs/1602.07360>
- [27] S. P. Nanavati and P. K. PaniGRAHI, "Wavelet transform," *Resonance*, vol. 9, no. 3, pp. 50–64, Mar. 2004, doi: [10.1007/BF02834988](https://doi.org/10.1007/BF02834988).
- [28] A. Grossmann, R. Kronland-Martinet, and J. Morlet, "Reading and understanding continuous wavelet transforms," in *Wavelets*. Berlin, Germany: Springer, 1990, pp. 2–20.
- [29] H. Zheng, Z. Li, and X. Chen, "Gear fault diagnosis based on continuous wavelet transform," *Mech. Syst. Signal Process.*, vol. 16, nos. 2–3, pp. 447–457, Mar. 2002, doi: [10.1006/mssp.2002.1482](https://doi.org/10.1006/mssp.2002.1482).
- [30] P. Kankar, S. C. Sharma, and S. Harsha, "Rolling element bearing fault diagnosis using autocorrelation and continuous wavelet transform," *J. Vibrot. Control*, vol. 17, no. 14, pp. 2081–2094, Mar. 2011, doi: [10.1177/1077546310395970](https://doi.org/10.1177/1077546310395970).
- [31] A. Bhattacharjee and B. K. Nanda, "Damping study of composites using wavelet analysis," *J. Vib. Control*, vol. 24, no. 21, pp. 5141–5151, Dec. 2017, doi: [10.1177/1077546317745256](https://doi.org/10.1177/1077546317745256).
- [32] S. Wu, X.-Y. Jing, Q. Zhang, F. Wu, H. Zhao, and Y. Dong, "Prediction consistency guided convolutional neural networks for cross-domain bearing fault diagnosis," *IEEE Access*, vol. 8, pp. 120089–120103, 2020, doi: [10.1109/ACCESS.2020.3005422](https://doi.org/10.1109/ACCESS.2020.3005422).
- [33] G. Jin, T. Zhu, M. W. Akram, Y. Jin, and C. Zhu, "An adaptive anti-noise neural network for bearing fault diagnosis under noise and varying load conditions," *IEEE Access*, vol. 8, pp. 74793–74807, 2020, doi: [10.1109/ACCESS.2020.2989371](https://doi.org/10.1109/ACCESS.2020.2989371).
- [34] Q. Jiang, F. Chang, and B. Sheng, "Bearing fault classification based on convolutional neural network in noise environment," *IEEE Access*, vol. 7, pp. 69795–69807, 2019, doi: [10.1109/ACCESS.2019.2919126](https://doi.org/10.1109/ACCESS.2019.2919126).
- [35] F.-P. An, "Rolling bearing fault diagnosis algorithm based on FMCNN-sparse representation," *IEEE Access*, vol. 7, pp. 102249–102263, 2019, doi: [10.1109/ACCESS.2019.2931616](https://doi.org/10.1109/ACCESS.2019.2931616).
- [36] L. Guo, Y. Lei, S. Xing, T. Yan, and N. Li, "Deep convolutional transfer learning network: A new method for intelligent fault diagnosis of machines with unlabeled data," *IEEE Trans. Ind. Electron.*, vol. 66, no. 9, pp. 7316–7325, Sep. 2018, doi: [10.1109/TIE.2018.2877090](https://doi.org/10.1109/TIE.2018.2877090).
- [37] A. Amini, M. Entezami, Z. Huang, H. Rowshandel, and M. Papaalias, "Wayside detection of faults in railway axle bearings using time spectral kurtosis analysis on high-frequency acoustic emission signals," *Adv. Mech. Eng.*, vol. 8, no. 11, Nov. 2016, Art. no. 1687814016676000, doi: [10.1177/1687814016676000](https://doi.org/10.1177/1687814016676000).
- [38] P. Cao, S. Zhang, and J. Tang, "Preprocessing-free gear fault diagnosis using small datasets with deep convolutional neural network-based transfer learning," *IEEE Access*, vol. 6, pp. 26241–26253, 2018, doi: [10.1109/ACCESS.2018.2837621](https://doi.org/10.1109/ACCESS.2018.2837621).
- [39] J. Jiao, M. Zhao, J. Lin, and C. Ding, "Deep coupled dense convolutional network with complementary data for intelligent fault diagnosis," *IEEE Trans. Ind. Electron.*, vol. 66, no. 12, pp. 9858–9867, Dec. 2019, doi: [10.1109/TIE.2019.2902817](https://doi.org/10.1109/TIE.2019.2902817).
- [40] M. Xia, T. Li, L. Xu, L. Liu, and C. W. de Silva, "Fault diagnosis for rotating machinery using multiple sensors and convolutional neural networks," *IEEE/ASME Trans. Mechatronics*, vol. 23, no. 1, pp. 101–110, Feb. 2018, doi: [10.1109/TMECH.2017.2728371](https://doi.org/10.1109/TMECH.2017.2728371).
- [41] W. You, C. Shen, D. Wang, L. Chen, X. Jiang, and Z. Zhu, "An intelligent deep feature learning method with improved activation functions for machine fault diagnosis," *IEEE Access*, vol. 8, pp. 1975–1985, 2020, doi: [10.1109/ACCESS.2019.2962734](https://doi.org/10.1109/ACCESS.2019.2962734).
- [42] J. Sresakoolchai and S. Kaewunruen, "Detection and severity evaluation of combined rail defects using deep learning," *Vibration*, vol. 4, no. 2, pp. 341–356, Apr. 2021, doi: [10.3390/vibration4020022](https://doi.org/10.3390/vibration4020022).
- [43] H. Alawad, S. Kaewunruen, and M. An, "A deep learning approach towards railway safety risk assessment," *IEEE Access*, vol. 8, pp. 102811–102832, 2020, doi: [10.1109/ACCESS.2020.2997946](https://doi.org/10.1109/ACCESS.2020.2997946).
- [44] P. Sengsri, C. Ngamkhanong, A. L. O. de Melo, and S. Kaewunruen, "Experimental and numerical investigations into dynamic modal parameters of fiber-reinforced foamed urethane composite beams in railway switches and crossings," *Vibration*, vol. 3, no. 3, pp. 174–188, Jul. 2020, doi: [10.3390/vibration3030014](https://doi.org/10.3390/vibration3030014).
- [45] P. Sengsri, C. Ngamkhanong, A. L. O. D. Melo, M. Papaalias, and S. Kaewunruen, "Damage detection in fiber-reinforced foamed urethane composite railway bearers using acoustic emissions," *Infrastructures*, vol. 5, no. 6, p. 50, Jun. 2020, doi: [10.3390/infrastructures5060050](https://doi.org/10.3390/infrastructures5060050).
- [46] R. Janeliukstis, A. Clark, M. Papaalias, and S. Kaewunruen, "Flexural cracking-induced acoustic emission peak frequency shift in railway prestressed concrete sleepers," *Eng. Struct.*, vol. 178, pp. 493–505, Jan. 2019, doi: [10.1016/j.engstruct.2018.10.058](https://doi.org/10.1016/j.engstruct.2018.10.058).
- [47] A. Thirumalaiselvi and S. Sasmal, "Pattern recognition enabled acoustic emission signatures for crack characterization during damage progression in large concrete structures," *Appl. Acoust.*, vol. 175, Apr. 2021, Art. no. 107797, doi: [10.1016/j.apacoust.2020.107797](https://doi.org/10.1016/j.apacoust.2020.107797).
- [48] G. Bayar and T. Bilir, "A novel study for the estimation of crack propagation in concrete using machine learning algorithms," *Construct. Building Mater.*, vol. 215, pp. 670–685, Aug. 2019, doi: [10.1016/j.conbuildmat.2019.04.227](https://doi.org/10.1016/j.conbuildmat.2019.04.227).
- [49] M. N. Noorsuhada, "An overview on fatigue damage assessment of reinforced concrete structures with the aid of acoustic emission technique," *Construct. Building Mater.*, vol. 112, pp. 424–439, Jun. 2016, doi: [10.1016/j.conbuildmat.2016.02.206](https://doi.org/10.1016/j.conbuildmat.2016.02.206).
- [50] S. M. Ali, K. H. Hui, L. M. Hee, M. S. Leong, and M. A. Al-Obaidi, "Observations of changes in acoustic emission parameters for varying corrosion defect in reciprocating compressor valves," *Ain Shams Eng. J.*, vol. 10, pp. 253–265, Feb. 2019, doi: [10.1016/j.asej.2019.01.003](https://doi.org/10.1016/j.asej.2019.01.003).
- [51] P. Loughlin and L. Cohen, "Wavelets: A comparison with the spectrogram and other methods for time-frequency analysis," *J. Acoust. Soc. Amer.*, vol. 127, no. 3, p. 1936, Mar. 2010, doi: [10.1121/1.3384871](https://doi.org/10.1121/1.3384871).
- [52] S. C. Olhede and A. T. Walden, "Generalized Morse wavelets," *IEEE Trans. Signal Process.*, vol. 50, no. 11, pp. 2661–2670, Nov. 2002, doi: [10.1109/TSP.2002.804066](https://doi.org/10.1109/TSP.2002.804066).
- [53] C. Szegegy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," 2014, *arXiv:1409.4842*. [Online]. Available: <http://arxiv.org/abs/1409.4842>
- [54] *Standard Guide for Determining the Reproducibility of Acoustic Emission Sensor Response*, Standard ASTM E976-10, 2010.
- [55] *R31-AST Sensor*. Accessed: Apr. 7, 2020. [Online]. Available: http://pacjapan.com/pacjapan_products/AESensor/PDF/R031-AST.pdf
- [56] *R151-AST Sensor*. Accessed: May 29, 2020. [Online]. Available: <http://www.pacndt.com/downloads/Sensors/IntegralPreamp/R151-AST.pdf>
- [57] *WD-100-900 KHZ Wideband Differential AE Sensor*. Accessed: May 29, 2020. [Online]. Available: <https://www.physicalacoustics.com/by-product/sensors/WD-100-900-kHz-Wideband-Differential-AE-Sensor>
- [58] B. Kato, H. Aoki, and H. Yamanouchi, "Standardized mathematical expression for stress-strain relations of structural steel under monotonic and uniaxial tension loading," *Mater. Struct.*, vol. 23, no. 1, pp. 47–58, Jan. 1990, doi: [10.1007/BF02472997](https://doi.org/10.1007/BF02472997).
- [59] I. G. Palmer and P. T. Heald, "The application of acoustic emission measurements to fracture mechanics," *Mater. Sci. Eng.*, vol. 11, no. 4, pp. 181–184, 1973, doi: [10.1016/0025-5416\(73\)90076-1](https://doi.org/10.1016/0025-5416(73)90076-1).
- [60] J. Masounave, J. Lanteigne, M. N. Bassim, and D. R. Hay, "Acoustic emission and fracture of ductile materials," *Eng. Fract. Mech.*, vol. 8, no. 4, pp. 701–709, 1976, doi: [10.1016/0013-7944\(76\)90043-6](https://doi.org/10.1016/0013-7944(76)90043-6).
- [61] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient convolutional neural networks for mobile vision applications," 2017, *arXiv:1704.04861*. [Online]. Available: <http://arxiv.org/abs/1704.04861>

- [62] V. Tra, J.-Y. Kim, I. Jeong, and J.-M. Kim, "An acoustic emission technique for crack modes classification in concrete structures," *Sustainability*, vol. 12, no. 17, p. 6724, Aug. 2020, doi: [10.3390/su12176724](https://doi.org/10.3390/su12176724).
- [63] M. B. Kursu, A. Jankowski, and W. R. Rudnicki, "Boruta—A system for feature selection," *Fundamenta Informaticae*, vol. 101, no. 4, pp. 271–285, 2010, doi: [10.3233/FI-2010-288](https://doi.org/10.3233/FI-2010-288).
- [64] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017, doi: [10.1145/3065386](https://doi.org/10.1145/3065386).
- [65] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. 3rd Int. Conf. Learn. Represent. (ICLR)*, San Diego, CA, USA, May 2015, pp. 1–14. [Online]. Available: <http://arxiv.org/abs/1409.1556>.



MD. ARAFAT HABIB was born in Rangpur, Bangladesh. He received the B.Sc. degree from BRAC University, Bangladesh, and the master's degree in computer engineering with Chosun University, South Korea.

He is currently working with the University of Ulsan, South Korea. His program was funded in part by Chosun University and the National Research Foundation of Korea. His current research interests include artificial intelligence, signal processing, fault diagnosis, and condition monitoring of industrial machinery. During his master's program, he has received Chosun University Excellent Student Award based on academic performance and the Brain Korea Scholarship.



MD. JUNAYED HASAN received the B.Sc. degree in computer engineering from BRAC University, Dhaka, Bangladesh, in 2016. He is currently pursuing the Ph.D. degree in computer engineering with the University of Ulsan (UOU), Ulsan, South Korea.

He worked on the development of advanced algorithms for explainable condition monitoring systems. Since 2017, he has been working as a Graduate Research Assistant with the Ulsan Industrial Artificial Intelligence (UIAI) Laboratory, Department of Electrical, Electronics and Computer Engineering, UOU. He is the author of more than ten scientific articles and three book chapters. His research interests include data-driven fault diagnosis and anomaly detection of complex engineering systems using explainable artificial intelligence, machine learning, advanced signal processing, fault feature extraction, and feature engineering.



JONG-MYON KIM (Member, IEEE) received the B.S. degree in electrical engineering from Myongji University, Yongin, South Korea, in 1995, the M.S. degree in electrical and computer engineering from the University of Florida, Gainesville, FL, USA, in 2000, and the Ph.D. degree in electrical and computer engineering from Georgia Institute of Technology, Atlanta, GA, USA, in 2005.

He is currently a Professor with the School of IT Convergence, University of Ulsan, Ulsan, South Korea. His research interests include fault diagnosis and condition monitoring, parallel processing, and embedded systems.

Dr. Kim is a member of the IEEE Industrial Electronics Society.

...