Mechanical Systems and Signal Processing 172 (2022) 108981

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

journal homepage: www.elsevier.com/locate/ymssp

Developing a heterogeneous ensemble learning framework to evaluate Alkali-silica reaction damage in concrete using acoustic emission signals

Li Ai [a], Vafa Soltangharaei [a,*], Paul Ziehl [a,b]

[a] Department of Civil and Environmental Engineering, University of South Carolina, Columbia, SC, USA
[b] Department of Mechanical Engineering, University of South Carolina, Columbia, SC, USA

ARTICLE INFO

Communicated by Simon Laflamme

Keywords:
Acoustic emission
Alkali-silica reaction
Ensemble learning
Convolutional neural network

ABSTRACT

The monitoring and evaluation of Alkali-silica reaction (ASR) damage in concrete structures are required to ensure the serviceability and integrity of concrete infrastructures such as bridges and dams. The innovation of this paper lies in the development of an automatic ASR monitoring and evaluation approach by leveraging acoustic emission (AE) and a heterogeneous ensemble learning framework. In this paper, ASR was monitored by AE sensors attached to a concrete specimen, which was placed in a chamber with high humidity and temperature. The recorded AE signals were filtered and divided by four ASR phases according to signal strength, crack width and expansion strain. A heterogeneous ensemble network including convolutional neural networks (CNN) and random forest models was employed to learn different features from AE signals and classify the AE signals into their corresponding phases. The results suggest that the proposed model has a high performance and classifies the signals into the assigned phases with high accuracy.

1. Introduction

Concrete is the important construction material for infrastructures. It is significant to ensure that the integrity of concrete structures during their service life is maintained. Alkali-silica reaction (ASR) is one of the main mechanisms of damage in concrete [1,2]. The alkali hydroxides (NaOH, KOH) from the cement react with the silica in reactive aggregate and produce hygroscopic gel [3]. The gel tends to absorb water and expands. The expansion of gel imposes pressure on the aggregate and cement matrix and causes cracking [1]. The process of ASR and cracking in concrete is presented in Fig. 1. The ASR expansion and cracking are affected by humidity and temperature. Concrete structures such as bridges, dams, and nuclear power plants exposed to water or high humidity may be affected by ASR [4].

Some conventional approaches have been used for the inspection and monitoring of ASR progress and damage evaluation in concrete structures. Those approaches include visual inspection, coring, and petrographic analysis. Those approaches have been widely utilized for decades, but they have some drawbacks. For instance, a visual inspection cannot be employed to detect early ASR damage because ASR damage usually initiates inside concrete structures and appears on the surface in later stages of the ASR process. Petrographic analysis can detect the early ASR damage inside concrete structures; however, this approach is intrusive, restricting its

* Corresponding author.
E-mail address: vafa@email.sc.edu (V. Soltangharaei).

https://doi.org/10.1016/j.ymssp.2022.108981
Received 20 October 2021; Received in revised form 27 January 2022; Accepted 23 February 2022
Available online 3 March 2022
0888-3270/© 2022 Elsevier Ltd. All rights reserved.

application in sensitive structures such as nuclear power plants. In addition, it is time-consuming and subject to human error [5]. To improve the reliability and efficiency of ASR monitoring, an automatic non-destructive monitoring approach of ASR damage is thereby needed.

Acoustic emission (AE) is a non-destructive structural health monitoring method that could be an option for monitoring ASR. AE is also referred to as a physical phenomenon when stress waves are generated by the rapid release of elastic energy during cracks or damage formation [6,7]. AE signals can be detected and collected by AE sensors attached to structures. The method of recording and processing AE signals to diagnose the health status of structures is AE monitoring [8]. AE is a sensitive method and has a continuous monitoring capability [9–14] and has been utilized for the evaluation of ASR damage in concrete structures [15,16]. Abdelrahman et al. [15] investigated the AE monitoring of ASR damage in three small-scale concrete specimens. The results indicated that the rate of AE activity is correlated to the rate of ASR damage, and AE intensity analysis can be used to classify ASR damage. Soltangharaei et al. [16] evaluated the ASR progress in concrete structures with different internal restraints by using AE. Unsupervised pattern recognition is employed to analyze AE signals. Previous research has proven that AE is a reliable technique for monitoring and evaluating ASR progress in concrete structures. However, traditional analyzing methods of AE signals are usually based on experience and very challenging, especially for complex data sets. Therefore, an intelligent algorithm is needed to assist in analyzing AE data in real-time and alert ASR damage stages.

Machine learning is an intelligent data processing technique [17]. By learning the feature extracted from the data, machine learning can understand the pattern of data and make a decision [18]. Machine learning models such as artificial neural network, support vector machine (SVM), K-nearest neighbor (KNN), AdaBoost and random forest (RF) has been widely utilized for signal processing [19–26]. One of the limitations of machine learning methods is nonautomatic feature extraction. In these methods, feature selection depends on human experience, and some important features may be overlooked. This limitation can be overcome by using architecture selection methods to indicate the feature importance based on statistical sensitivity analysis [27,28]. Albu et al. [27] proposed an architecture selection method for a multilayer feedforward network. The method can indicate the importance of input features by analyzing the partial derivatives of output with regard to the input features. Utilizing deep learning methods can be another way to overcome the limitation. Deep learning methods are improved intelligent techniques which are based on machine learning. It can automatically learn features from complex data sets without feature extractions [29]. CNN is one of the relatively new deep learning algorithms which is widely studied in image recognition [30,31], object detection [32], and semantic segmentation [33]. Combining various machine learning and deep learning models is a strategy to improve model performance. Panigrahy et al. [34] combined SVM and AdaBoost to improve the detection of ventricular fibrillation (VF) rhythm. Taherkhani et al. [35] integrated AdaBoost model with a CNN model to deal with large imbalanced datasets with high accuracy. In recent years, machine learning and deep learning have been applied to AE data in many applications, such as metallic structures [36], aircraft components [37], and rail track monitoring [38]. Ebrahimkhanlou et al. [36] proposed a deep learning framework based on a stacked autoencoder network to locate AE events on the metal structures. Ai et al. [37] developed a passive non-destructive health monitoring system to locate impacts on an aircraft component based on AE, random forest, and stacked autoencoder network. Li et al. [38] utilized a multi-branch CNN model for AE wave classification to obtain a more accurate and comprehensive monitoring system.

Machine learning, deep learning, and AE have also been employed to evaluate the time-dependent damage stages of composite materials and metallic materials [39,40]. Nasiri et al. [39] investigated on online monitoring of degradation process in a composite specimen. Three different stages of degradation: elastic, matrix-driven, and fiber-driven behavior, have been defined. The researchers identify different degradation stages of the specimen by classifying the AE events into one of the three different classes. Chen et al. [40] employed AE for the monitoring of railways. An AE sensor was attached on the rail and the entire monitoring process was divided into four stages according to the mechanism of fatigue growth. An acoustic-homologous transfer learning approach was developed to classify the AE signals into their corresponding stages.

Previous studies indicated using AE and different classification models to evaluate time-dependent damage stages in one specimen. The development process of ASR in concrete can also be divided into several phases since ASR is a time-dependent phenomenon [41]. Therefore, classification models can be trained to identify the ASR damage stages based on received AE data. The authors are currently not aware of similar published studies implementing deep learning and machine learning classification algorithms to recognize the

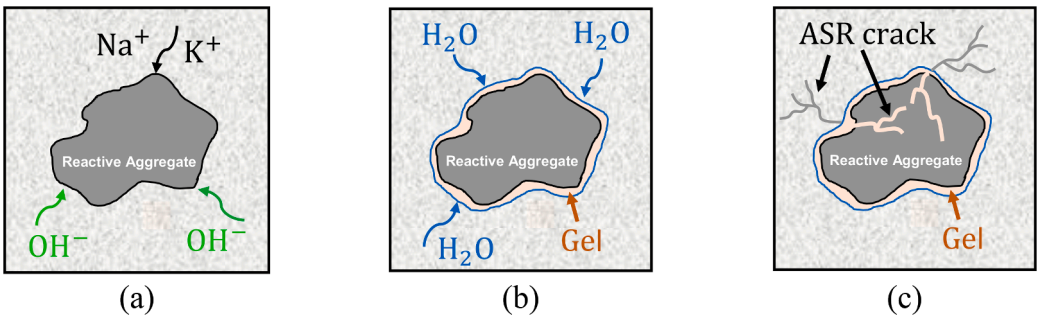

Fig. 1. Mechanism of the ASR in concrete: (a) alkali cement react with reactive aggregates; (b) ASR gels forms around the aggregate and absorb water; (c) ASR gels expansion and crack initiation.

ASR damage phases using AE data. Therefore, this paper investigates a novel ASR monitoring and evaluation approach by employing AE, machine learning, and deep learning algorithms to fill the prementioned gap. The temporal degradation process of a concrete specimen was divided into different stages according to damage mechanisms. The AE signals emitted from different damage mechanisms had different signal signatures. Therefore, the temporal ASR damage evaluation of a specimen was mapped as a classification problem. The input was an AE signal at a given timeslot, and the output was the associated damage stage category. CNN and random forest models were utilized herein to analyze the signatures of AE signals captured during the ASR process and determine the ASR damage stage of concrete based on the AE signals. A heterogeneous ensemble learning framework was designed and proposed to combine CNN and random forest models to consider results from all combined models.

2. Materials and experimental setup

2.1. Specimen and material

In this paper, a concrete block specimen reinforced by steel rebars was prepared. The geometrics of the concrete specimen is presented in Fig. 2b. The dimensions of the block are 305 mm × 305 mm × 1120 mm. The rebar details are shown in Fig. 2c. The specimen had four longitudinal US #7 steel rebars and US #6 steel rebars with 150 mm spacing as transverse reinforcement. All rebars were T-headed to compensate for the short development length.

The concrete has a 0.5 cement to water ratio. The cement used in the mixture was an ASTM C150 [42] Type I/II low-alkali cement with 0.48% Na_2O_{eq}. Crushed greenschist from North Carolina was utilized as a reactive coarse aggregate. Sodium hydroxide (NaOH) was added to the concrete mixture to accelerate the development of ASR. Details of the materials can be found in Table 1.

2.2. Acoustic emission instrument setup

Ten AE sensors were utilized to collect AE events captured during the ASR experiment. The AE sensors were attached to the concrete specimen by using grey double/bubble epoxy. Fig. 2a shows the attached sensors on the concrete specimen (cables were not connected) before conducting the test. The dimension of the specimen and layout of the sensors can be found in Fig. 2b. Sensors 1–2 were attached to the top surface. Sensors 3–4 were attached to the bottom. Sensors 5–7 were attached on the backside surface, and Sensors 8–10 were attached on the front side surface. All AE sensors in this paper were wideband type PKWDI with 26-dB internal preamplification. The sensors have an operating frequency range of 200–850 kHz. An attenuation test has been conducted to verify the sensitivity of PKDWI sensors by performing the Hsu-Nielsen pencil lead break [43]. The results indicated that the sensors were sensitive to receive the signals from the farthest location on the specimen surface. The hardware and software of the AE system were produced by MISTRAS Group, Inc. (Princeton Junction, NJ, USA). AE signals were acquired by a 24-channel Micro-II Express acquisition system. The pre-trigger time, hit definition time (HDT), peak definition time (PDT), and hit lockout time (HLT) were set to 256 μs, 400 μs, 200 μs, and 200 μs, respectively. The sampling rate was set to 5000 kHz. The signal length was set as 1024 μs. Therefore, each signal has 5120 data points. Before the experiment, the level of background noise in the chamber room was tested. Most of the background noise were below 32 dB (ref 1 V/(m/s)). Therefore, the threshold in the experiment was set to 32 dB (ref 1 V/(m/s)) during the monitoring. The AE signals with the amplitude higher than the threshold were recorded. The low and high pass digital filters were set to 20 kHz, and 400 kHz, respectively.

2.3. Experiment procedures

During the experiment, the concrete specimen was sealed in a chamber with dimensions of 243 cm × 243 cm × 122 cm. A steel carrier with wheels (as seen in Fig. 2a) was designed and built to be employed as a support for the concrete specimen. The ASR process accelerated, providing high humidity (95% ± 5%) and a high temperature (37 ± 3 oC) in the chamber. Neoprene pads were employed between the specimen bottom surface and steel carrier to minimize the effect of vibrations from the ground. Pins (as seen in Fig. 2b) were affixed on the surfaces of the specimen using grey double/bubble epoxy for expansion strain measurements. The distance

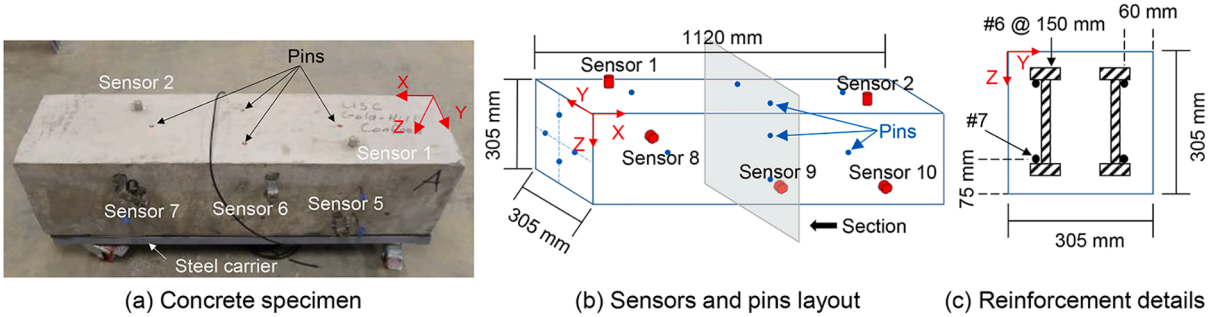

Fig. 2. Structural details of specimen.

Table 1
Concrete mixture proportions.

Mixture components	Quantity (kg/m^3)
Cement	350
w/c ratio	0.5
Reactive coarse aggregate	1050
Nonreactive sand	851
NaOH solution (50% w/w)	9.22

between the pins, along the X-axis, was 500 mm. The distances between the pins along the Y and Z axes were 150 mm. The pin arrangement on the right and back surfaces of the specimen were identical to those on the left and front surfaces. The specimens were monitored for 460 days in the chamber, and AE was recorded in time. The ASR expansions were measured monthly using DEMEC gauges (demountable mechanical strain gauges) on the pin locations. The monthly maintenance of AE sensors was conducted to ensure the strong bonding condition between sensors and the specimen. The first DEMEC gauge measurement was conducted on the 18th day of the experiment. Crack width measurement started on the 146th day (after observing the surface visible cracks) using a Dino-Lite digital microscope with a maximum magnification of 184 X and an ELE crack detection microscope with a magnification of 40 X. More detailed information on the experimental setup and procedures can be found in [16].

3. Methodology

The ASR evaluation method proposed in this paper is based on AE monitoring and a heterogeneous ensemble learning framework. The original AE waveforms acquired by the acquisition system were saved as time series. An image-based dataset was created by converting the AE waveforms to the continuous wavelet transform (CWT) images. A feature-based dataset was obtained by extracting parametric features from the AE waveforms. The detailed descriptions of the two datasets could be found in Section 4.2.

Heterogeneous ensemble learning refers to the ensemble structure composed of different types of models [44]. In this paper, a heterogeneous ensemble learning network includes convolutional neural networks (CNN) and random forest (RF) models and is employed to receive two different AE datasets. The network classifies the AE data into different ASR phases (introduced in section 4). The image-based dataset was divided into training, validation, and testing datasets. The feature-based dataset was divided into training and testing datasets. No validation dataset was prepared for the feature-based data. The reason will be explained in section 3.3.4. Prepared training/validation data was used to train the heterogeneous ensemble learning network. The testing set was used to evaluate

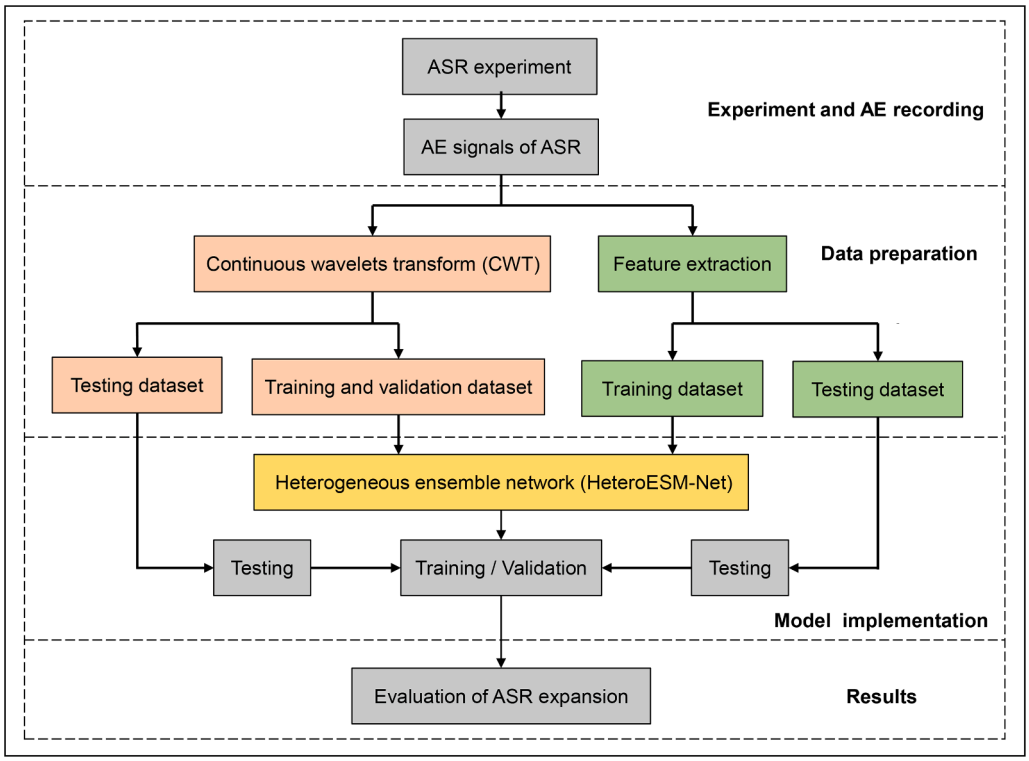

Fig. 3. Workflow of ASR evaluation using ensemble learning framework.

the performance of the trained network. The ratio of the testing dataset over the whole dataset was kept the same in both the image dataset and the feature-based dataset. The procedure of the ASR expansion evaluation is presented in Fig. 3.

3.1. Continuous wavelet transform

CWT is a widely used joint time–frequency analysis approach highlighting time–frequency characteristics from a non-stationary signal such as AE signals [45]. Assume there is a signal: $x(t)$. The CWT of the signal can be expressed as Eq. (1):

$$CWT_{(a,b)} = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} x(t)\psi^{*}\left(\frac{t-b}{a}\right)dt \tag{1}$$

where *CWT* is the continuous wavelet coefficients derived from the signals, a refers to the scale index parameter which controls the scaling of wavelet function and has an inverse relation to frequency,b refers to the translation parameter which controls time-shifting of wavelets. The wavelet coefficients are derived by moving wavelets with different a scale indices through the signals. ψ^{*} refers to the complex conjugate of mother wavelet function ψ. In this paper, Morse wavelet [46] is selected as the mother wavelet function to conduct CWT. The Fourier transform of Morse wavelet is presented in Eq (2):

$$\Psi_{p,\gamma}(x) = U(x)\alpha_{p,\gamma}x^{\frac{p^2}{\gamma}}e^{-x^{\gamma}} \tag{2}$$

where $U(x)$ refers to the unit step, $\alpha_{p,\gamma}$ refers to the normalizing constant, p^2 refers to the time-bandwidth product. γ is the parameter that characterizes the symmetry of the Morse wavelet [47]. In this paper, p^2 was set as 60 and γ was set as 3.

A scalogram image can be used to express the continuous wavelet coefficients. The images of AE waveforms are utilized as an input for the CNN models inside the proposed heterogeneous ensemble learning network.

3.2. Convolutional neural network

CNN is a class of commonly used deep neural networks that are applied for image processing [30]. CNN is composed of three main parts: the input layer, the feature extraction layers, and the fully connected (FC) layer. The input layer is used to input the test and training data. Feature extraction layers are the core of the convolutional neural network, mainly including convolutional layers and pooling layers, which cooperate to derive the features from images and learn potential patterns in the data set.

3.2.1. Convolutional layer

The convolutional layers are utilized to extract the features from images [48]. In the convolutional layer, multiple convolutional kernels are employed to filter the input and generate feature maps [48]. Generally, the output of the j^{th} feature maps of the n^{th} convolutional layer can be obtained by Eq. (3):

$$x_j^n = f\left(\sum_{i=1}^{M} x_i^{n-1} * k_{ij}^n + b_j^n\right) \tag{3}$$

where $f(\hat{A}\cdot)$ refers to the activation function,* refers to the operation of convolutional kernels, k_{ij}^n is the kernel of the n^{th} filter, b_j^n is the corresponding bias matrix, x_i^{n-1} refers to the input feature map transferred from the $(n\text{-}1)^{th}$ convolutional layer.

3.2.2. Pooling layer

The pooling layer is used for the down-sampling of feature maps obtained from the previous convolutional layer [49]. If the image feature maps are directly utilized for the classification without any processing, a great computational complexity will be generated, and the model will be prone to overfitting. Therefore, a further reduction in the dimensionality of feature maps is required, which is the reason for the pooling layer after each convolutional layer. The input feature image is divided into mutually exclusive regions, and the feature information of adjacent image regions is aggregated for analysis. This type of down-sampling method is called pooling [50]. Pooling can be divided into maximum pooling and mean pooling according to different operation modes. The general expression of the pooling layer is provided by Eq. (4):

$$x_j^n = f\left(\beta_j^n s_{down}(x_j^{n-1}) + b_j^n\right) \tag{4}$$

where $f(\hat{A}\cdot)$ refers to the activation function,β_j^n and b_j^n refers to the multiplicative bias and the additive bias, S_{down} refers to the down-sampling function, x_j^{n-1} refers to the input feature maps,x_j^n refers to the output feature map after down-sampling.

3.2.3. Fully connected layer

The FC layer is employed at the end of the CNN model [51]. It converts the feature maps that result from the previous pooling layer to one feature vector. The calculation in the j^{th} FC layer can be expressed by the following equation:

$$x^j = f\left(x^{j-1} w^j + b^j\right) \tag{5}$$

where f refers to the activation function, x^{j-1} is the input feature maps. x^j is the output value. w^j and b^j are the weight and bias.

3.3. Proposed heterogeneous ensemble network

In this paper, two types of AE datasets were prepared, as mentioned before. In the CWT image-based dataset, most of the raw information of each AE waveform was retained. The CNN could learn useful time–frequency features from raw information embedded in CWT images. However, the CNN cannot learn the temporal relationship between individual signals. The AE activity has a relationship with the ASR damage [15,52]. Therefore, an AE feature-based dataset was prepared. In this dataset, a feature named "Hit rate" was calculated based on the number of acquired AE hits in terms of time. The hit rate of an AE signal was calculated by counting the number of hits that occurred in a certain number of hours before and after the current hit (Eq. (6)).

$$HR(h) = Hit_{before}(h/2) + Hit_{after}(h/2) \tag{6}$$

where $HR(h)$ refers to the hit rate of a signal per h hours. $Hit_{before}(h/2)$ refers to the number of AE hits that were received $h/2$ hours before the current hit. $Hit_{after}(h/2)$ refers to the number of hits that were recorded $h/2$ hours after the current hit. In this paper, h was set to 24 h.

AE-based features will be elaborately explained in section 4.2.1. A heterogeneous ensemble network including CNN models and machine learning models was proposed to include both the CWT images, and the AE features. The structure of the proposed ensemble network is presented in Fig. 4. The proposed ensemble network was developed based on the bagging aggregating technique, an ensemble strategy to combine several sub-models and improve accuracy [53]. In a general bagging ensemble model, multiple sub-models will work individually and give their results. The final result will be determined by a majority vote [53].

The proposed ensemble network contains two branches and four blocks. Branch 1 contains blocks 1–3. Branch 2 contains block 4. Each block is a homogeneous ensemble network composed of five sub-models with the same structures. Previous research has reported that CNN structures such as VGG, ResNet, and GoogLeNet work well on the fault diagnosis of bearing using vibration signals [54–57] and the identification of acoustic emission signals in the applications such as the monitoring of wear and bridge [58–60]. Therefore, the CNNs with ResNet-18, GoogLeNet, and VGG-19 structures are used as the sub-models in blocks 1–3 to obtain a good performance in the classification of AE signals. Random forest models are employed as machine learning sub-models in block 4. The proposed heterogeneous ensemble network model was named RGVF-HeteroESM-Net in this paper, in which RGVF consists of the first letters of ResNet-18 (R), GoogLeNet (G), VGG-19 (V), and random forest (F). RGVF-HeteroESM-Net is referred to as a heterogeneous ensemble network, which combines ResNet-18, GoogLeNet, VGG-19, and random forest models. The CWT image-based dataset is used as the input of blocks 1–3, and the AE feature-based dataset is utilized as the input for block 4. A bootstrapping [61] process is utilized to increase the diversity of each sub-model during the training. Each sub-model randomly selects 90% of samples from the training dataset as the true training data. All 20 sub-models were trained individually and tested on the test dataset. The final results were obtained through majority voting.

3.3.1. Block 1: VGG-19

VGG is developed based on the AlexNet model [62], which is a commonly used CNN structure. The number of layers was extended

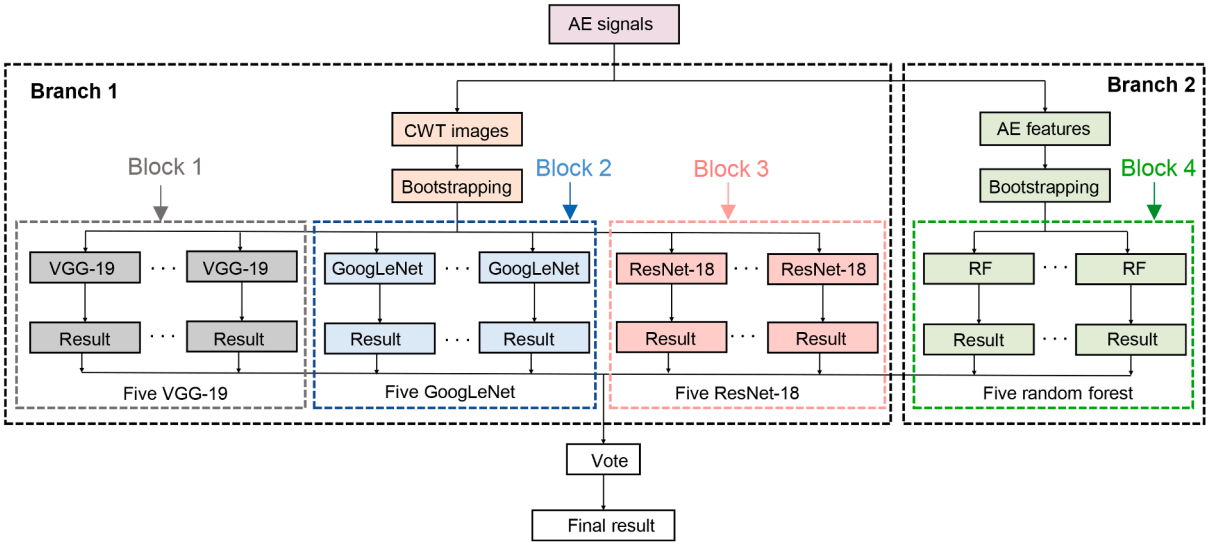

Fig. 4. Structure of the proposed RGVF-HeteroESM-Net.

up to 19. The advantage of VGG-19 compared to AlexNet lies in replacing larger convolution kernels in AlexNet (11 × 11, 7 × 7, and 5 × 5 convolution kernels) by stacking 3 × 3 convolution kernels [30]. Using the stacked small convolution kernels, the number of layers and the nonlinearity of the network increases, which gives the network the ability to learn more complex features.

The VGG-19 structure is used in block 1 of the proposed ensemble network. The last FC layer of the VGG-19 structure is modified to have the class number consistent with the number of ASR phases. Fig. 5 shows the main structure of the modified VGG-19.

3.3.2. Block 2: GoogLeNet

GoogLeNet is another improved model based on AlexNet [63]. The number of layers of GoogLeNet is extended up to 22. The innovation of GoogLeNet is the use of inception modules in the network [63]. Generally, in a layer of CNN, there is only one convolution kernel. However, the inception module uses multiple convolution kernels of different scales in a single layer. Therefore, some complex features in the images can be extracted by convolution kernels of different sizes. The feature extraction ability of each layer is thereby enhanced. GoogLeNet was utilized as the CNN structures in block 2 of the proposed ensemble network. Similar to the VGG-19 structure, the last FC layer of the GoogLeNet structure is modified. Fig. 6 shows the main structure of the modified GoogLeNet.

3.3.3. Block 3: ResNet-18

The ResNet-18 structure is a CNN structure with the idea of residuals [64]. Sometimes, the network's performance becomes less reliable when using a deeper structure due to gradient vanishing/explosion problems hindering network convergence [64]. A residual block module was developed and applied in the ResNet-18 structure to overcome this problem by introducing skip connections that enable gradients to flow across several layers [64]. The skip connections cause the outputs to learn a residual mapping.

In this paper, ResNet-18 was employed as one CNN structure in block 3 of the proposed ensemble network. As in the VGG-19 and GoogLeNet, the last FC layer is modified. Fig. 7 shows the main structure of the modified ResNet-18.

3.3.4. Block 4: Random forest

Random forest is a machine learning model using a bagging algorithm [65]. The sub-model inside the random forest is a decision tree. Bootstrapping [61] is used to randomly selected data from the training dataset and create an input dataset for each decision tree. Some samples may be selected repeatedly, while some samples may not be selected. The data in the dataset that may not be chosen (even one time) is called out-of-bag (OOB) data, which can be used to test the generalization performance of the decision tree [66]. The function of OOB data is similar to the validation dataset of a general machine learning method. Therefore, there is no need to have a specific validation dataset, and the data for the random forest is divided by training and testing data. The decision tree will derive the results individually after bootstrapping. The final result of the random forest is voted or averaged from all the results. The main structure of a random forest with n decision trees is presented in Fig. 8.

4. Results and discussion

4.1. Analysis of AE signals and ASR phase definition

During the ASR experiment, the AE signals were captured by ten AE sensors and recorded by the acquisition system. The data was filtered before conducting data analyses. A filtering procedure based on AE event definition was developed to remove extraneous data. An AE event refers to a set of AE hits captured by multiple AE sensors in a specific time interval [16]. The time interval was defined based on the velocity of AE wave and the dimensions of the specimen. In this paper, the AE events recorded by at least four AE sensors were kept. The rest of the signals were considered as noise and were thereby removed. After filtering, 4413 AE signals remained for the analyses. As mentioned in Section 2.2, each AE signal has 5120 data points. Therefore, the filtered AE dataset used in deep learning

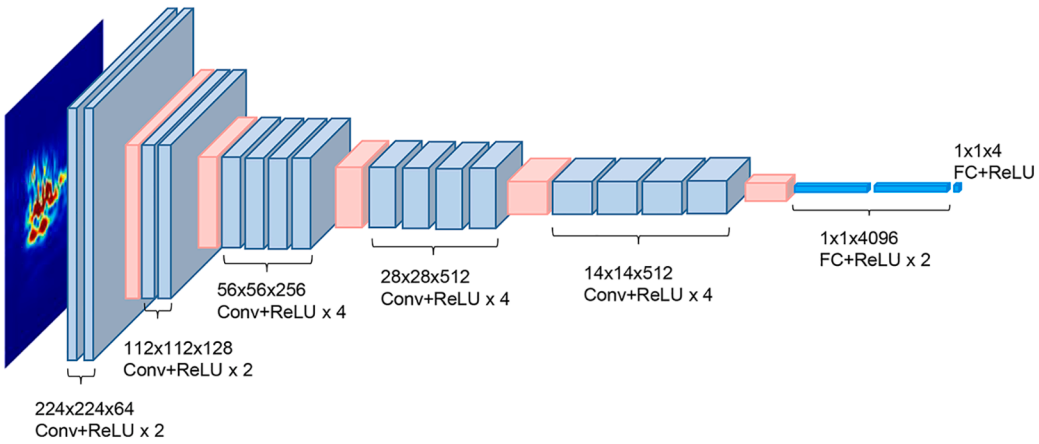

Fig. 5. Main structure of the modified VGG-19.

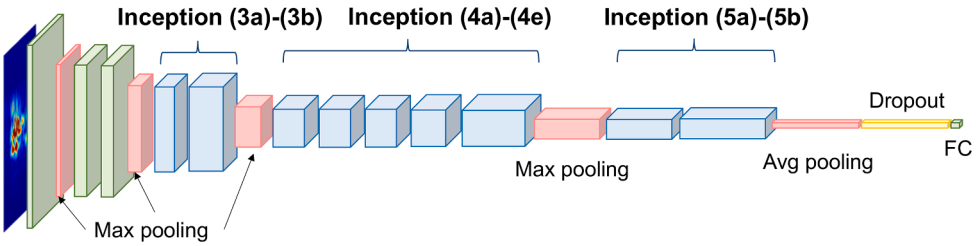

Fig. 6. Main structure of the modified GoogLeNet.

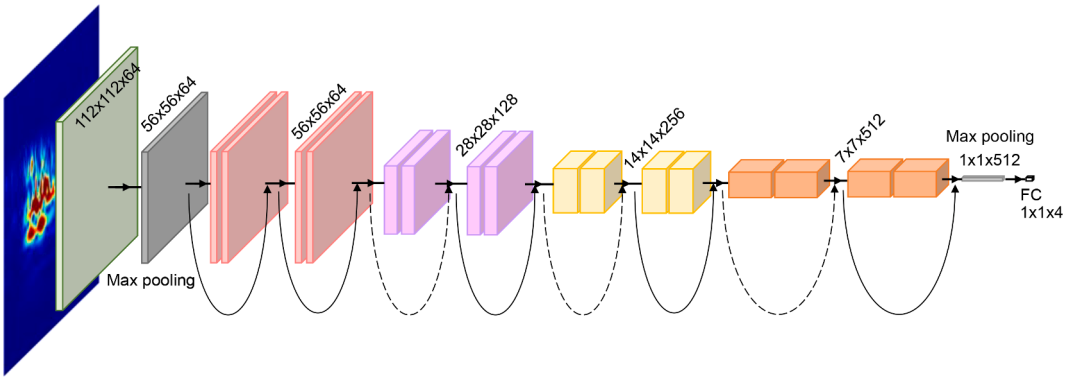

Fig.7. Main structure of the modified ResNet-18.

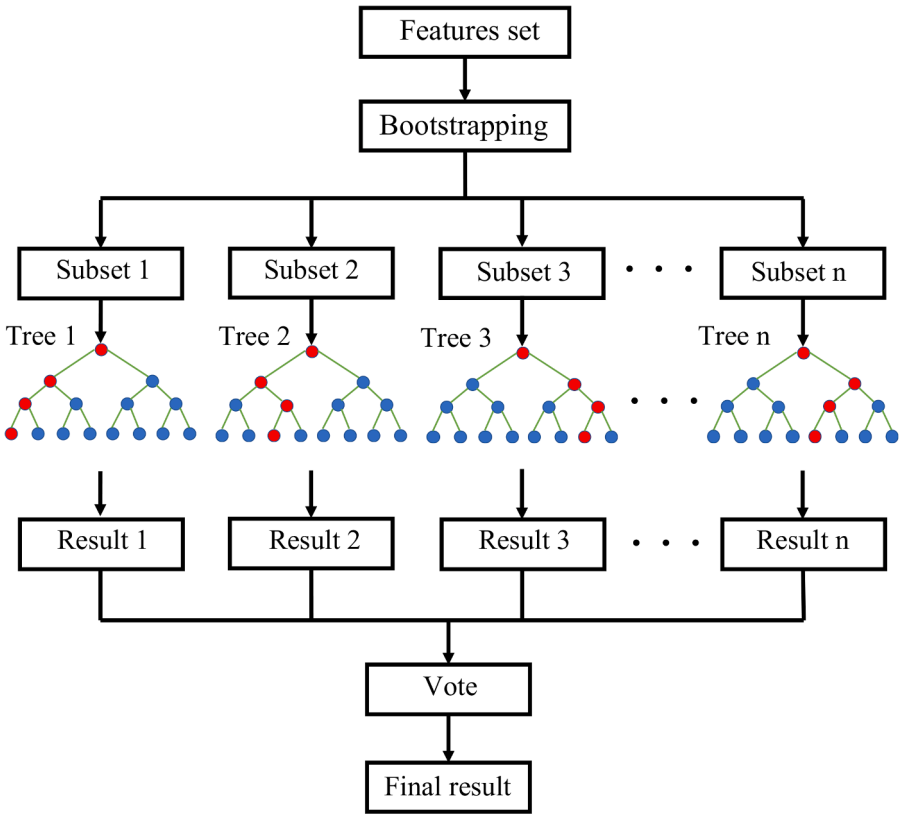

Fig. 8. Structure of random forest.

models is a matrix with a size of 5120 × 4413.

The amplitude of the filtered AE signals (shown as red dots) and the cumulative signal strength (CSS) are shown in Fig. 9. A rapid rise of CSS can be observed around 130 days. Two additional significant steps can be observed around 200 days and 350 days. The jumps in the CSS curve can be attributed to the crack formations or the propagation of existing cracks. The crack width was monthly measured for the surface cracks. The maximum crack widths were not necessarily attributed to the same crack and same location during the ASR process. The maximum crack widths in terms of time are illustrated in Fig. 9. A curve was also fitted to the crack width data, as seen in the figure. The first visual crack was observed at 146 days. The crack width expands rapidly from 0.06 mm at 146 days to 0.32 mm at 331 days. A crack with a width of 0.66 mm was observed at 438 days. Microscopic photos of cracks at 269, 331, and 438 days are also presented in Fig. 9. The magnification in the pictures is 184X. In addition to the crack width, expansion strains were measured monthly on the specimen surfaces. The volumetric strain refers to the accumulation of average strains along the X, Y, and Z axes. The fitted curve of the volumetric strain is also illustrated in Fig. 9. The fitted volumetric strain versus time and crack width versus time have a similar trend.

The development process of ASR in concrete can usually be divided into several phases [41]. In this paper, the entire ASR process was divided into four phases, considering the crack widths, CSS trends, and the volumetric strains. The first phase is between 1 and 130 days, before the first rapid rise of CSS (microcrack phase). No visible cracks were observed during this phase on the concrete surfaces,

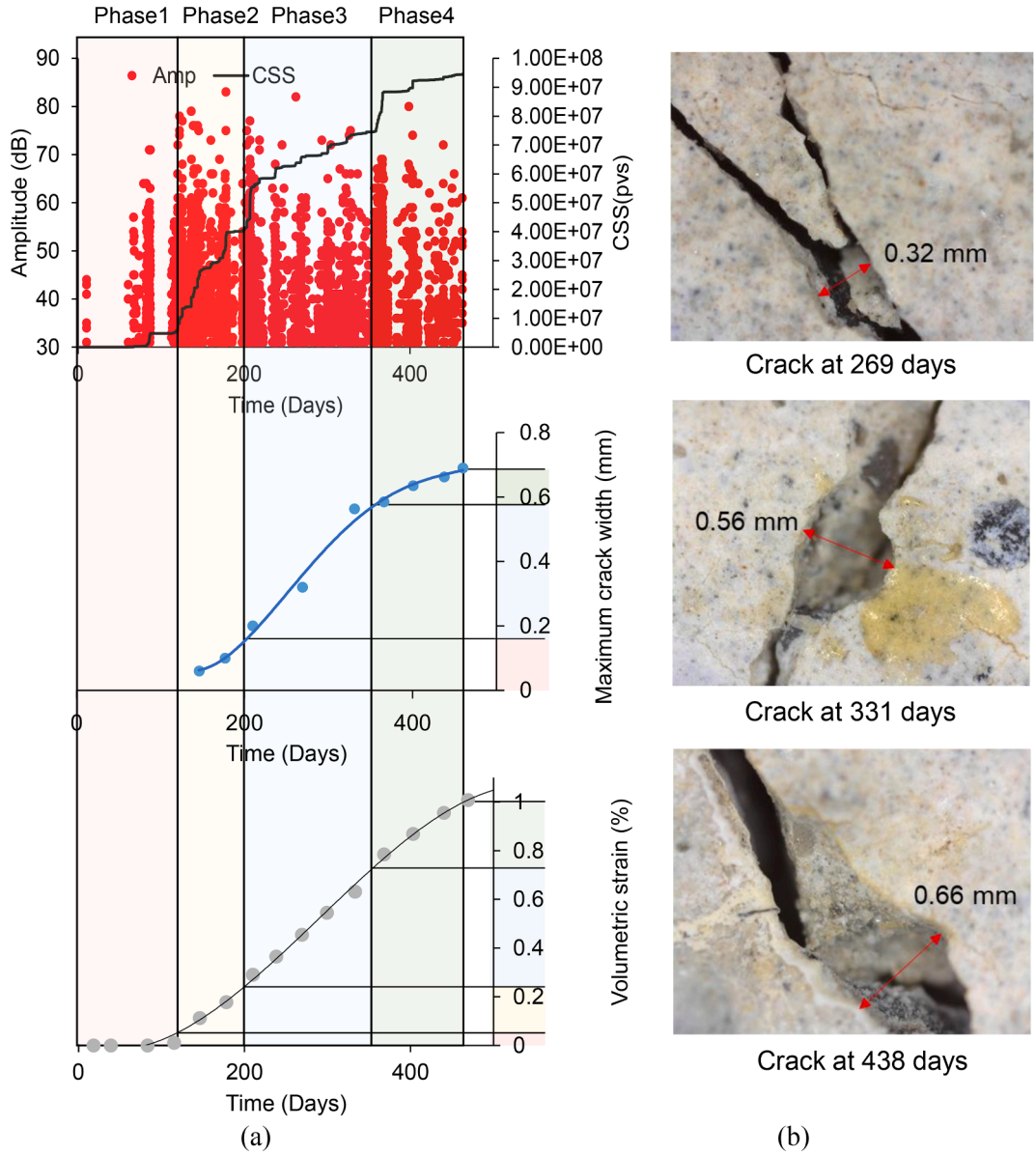

Fig. 9. ASR phase definition: (a) CSS of AE, crack width, and volumetric strain; (b) microscopic photos of cracks.

and the volumetric strain ranged from 0% to 0.07%. The days between the first and the second significant steps of CSS (days 130–200) were defined as phase 2 (macrocrack initiation phase). In phase 2, the microcracks were merged and formed macrocracks. The first visible crack was observed in this phase, and the maximum crack width increased from 0 mm to 0.15 mm. The strain changed from 0.07% to 0.24% in phase 2. The days between the second and third steps of CSS (day 200–350) were defined as phase3 (macrocrack extension and dilatation phase), where the crack width increased rapidly from 0.15 to 0.57 mm, and the strain changed from 0.24% to 0.74%. The last phase (between 350 and 460 days) was assigned to phase 4 (macrocrack rate decrease). In this phase, existing cracks were widened, but the rate of crack width widening was decreased. The maximum crack width increased from 0.57 mm to 0.69 mm in phase 4. The strain changed from 0.74% to 1%. This paper aims to automatically classify AE signals recorded during the ASR process according to the assigned phases (phase 1, phase 2, phase 3, and phase 4) by using the proposed heterogeneous ensemble learning framework.

4.2. Data preparation

4.2.1. Feature-based dataset

The feature-based dataset served as the input of the random forest model. Twenty-five features (15 parametric features and 10 energy-frequency features) plus hit rate (introduced in section 3.3) were considered for the model.

AE parametric features are used to represent the specific characteristic of AE waveform. Some AE parametric features (amplitude, rise time, duration, counts to peak, counts) are shown in Fig. 10a. Definitions of parametric features are provided in Table 2.

The energy-frequency features refer to the energies in the frequency bands of the signals. The AE time domain waveforms were converted to the frequency domain spectrum by conducting a fast Fourier transform (FFT). The frequency range of the AE signals is 0–400 kHz and was divided into ten bands with a width of 40 kHz. The area under the FFT spectrum in each frequency band is the energy enclosed by that frequency band [16]. Therefore, ten energy-frequency features could be derived from each AE signal. The procedure to extract the energy-frequency feature in the frequency band of 80–120 kHz is presented in Fig. 10b. The energies in the frequency bands were normalized to the total energy of the signal.

4.2.2. Image-based dataset

All the recorded 4413 AE signals were transformed to CWT coefficients to form the image-based dataset. Fig. 11 illustrates example CWT images of signals for four phases. The amplitudes of time-domain waveforms were normalized from −1 to 1. The FFT spectra were normalized to a range of 0 to 1, and the wavelet coefficients were scaled between 0 and 1. The signals from different phases have different frequency contents, as seen in Fig. 11.

The overall procedures for the data preparation are shown in Fig. 12. AE parametric features were extracted from the original AE waveform. The energy-frequency features were extracted by the FFT spectrums, which were transformed from the waveforms. AE hit rate for each signal was calculated by the AE hit number recorded during the ASR experiment. The three parts (26 features) form the AE features-based dataset. In other words, the feature-based dataset has 4413 samples; each sample has 26 features. Furthermore, the AE waveforms were utilized to generate CWT coefficients. The Y-axis of the CWT coefficient was then converted to a logarithmic coordinate to present the time–frequency component more clearly. The coefficients were saved as RGB images with the size of 224 × 224 × 3 pixels. The CWT image-based dataset was composed of the derived RGB images and employed as the input of the CNN models.

4.3. ASR evaluation using single CNN model

The performance of ResNet-18, GoogLeNet, and VGG-19 models was evaluated in this section. 80% of the image-based dataset (3531 images) was randomly selected as a training set. 10% of the image-based dataset (441 images) was randomly selected as a validation set. Finally, a test set was constructed by the remaining 10% (441 images) of the CWT images. The training, validating, and

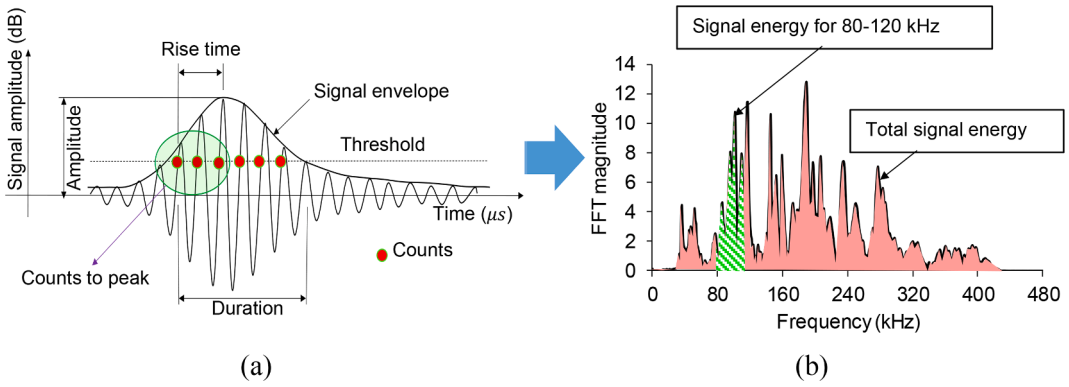

Fig. 10. Extraction of energy-frequency features: (a) AE time-domain waveform and some typical features; (b) energy-frequency features extraction procedure for 80–120 kHz.

Table 2
Descriptions of the AE parametric features.

Parametric features	Feature descriptions
Amplitude (dB)	The peak amplitude of AE waveform
Energy (arbitrary units. = $10^{-14}V^2$ s)	The measure of the electrical energy measured for an AE signal
Count	The number of threshold crossings
Counts to peak (PCNTS)	The number of threshold crossings from the first crossing to the peak
Rise time (μs)	The time interval between first threshold crossing and peak
Duration (μs)	The time between the first and last threshold crossing
Average frequency (Hz)	Counts divided by duration
Frequency centroid (Hz)	A parameter to characterize the overall frequency content of an AE signal
Absolute energy (Atto-Joule)	The absolute measure of the electrical energy measured for an AE signal
Peak frequency (Hz)	Frequency of maximum signal contribution
Reverberation frequency (Hz)	Frequency after the peak
Initial frequency (Hz)	Frequency before the peak
Signal strength (pVs)	Integral of the rectified voltage signal over waveform duration
Root mean square (RMS) (mV)	The effective voltage with a characteristic time T_{RMS} for an average ranging between 10 and 1000 *ms*
Average signal level (ASL) (dB)	The effective voltage with a characteristic time T_{ASL} for an average ranging from 10 to 1000 *ms*

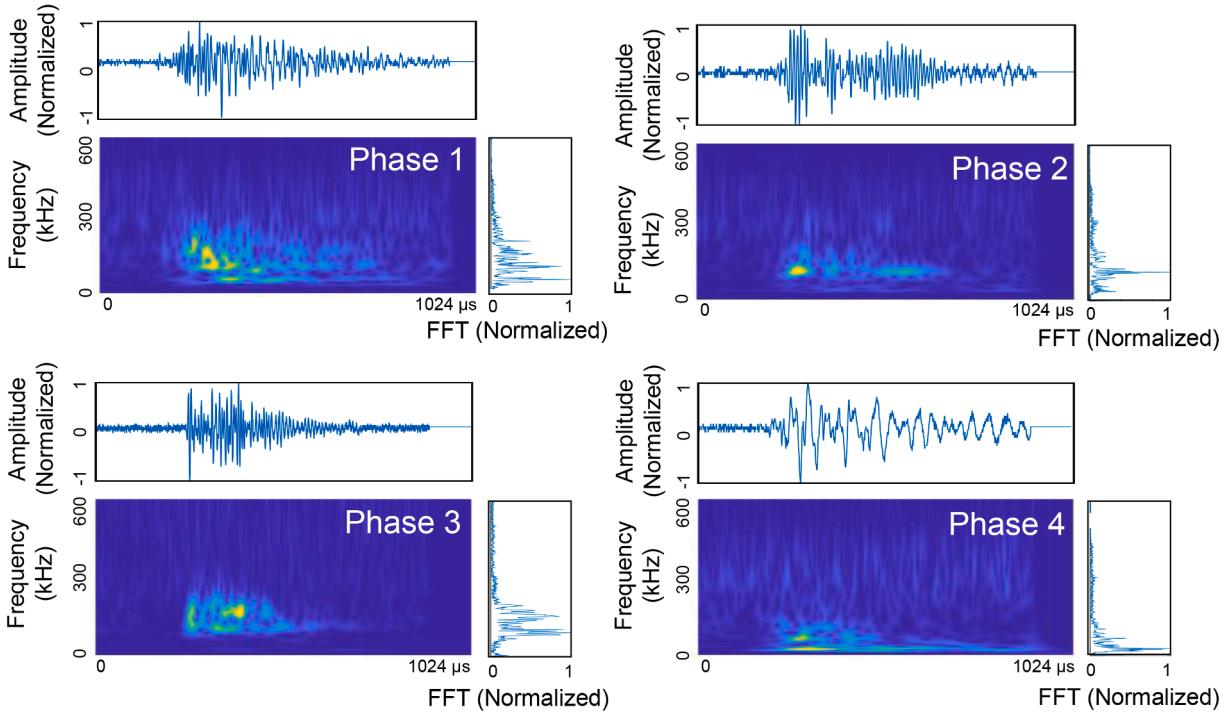

Fig. 11. CWT images of AE signals in phase 1–4.

testing processes for the CNN models were conducted on a workstation with a CPU-Intel i7-6700 3.40 GHz, 32 GB RAM, and an Nvidia GPU-GTX1080. For all the CNN models, the gradient descent optimization was conducted using the Adaptive moment estimation (Adam) method. The optimized hyperparameters were set as follows: The minibatch size was 35, the learning rate was 0.0001, and the maximum number of epochs was 25. The three models were separately trained and tested by five trials. The bootstrapping process was applied to each CNN model. In each training process, 90% of the images (3178 images) were randomly selected from the training set as an input set to train the CNN models. The minibatch size was set as 35 so that there were 90 mini-batches in the input set. The maximum number of epochs was 25. In total, 2250 iterations were conducted. Fig. 13a presents the validation accuracy curves of the three models from iteration 1 to 2250. ResNet-18 converges faster (around iteration 400) than GoogLeNet and VGG-19 (around iteration 800), and the curve of ResNet-18 is more stable than the other two models after converging. The validation curves of different models are close to each other. Fig. 13b shows that ResNet-18 has the highest validation accuracy among the three models. However, the differences in validation accuracies between the models are not significant.

Testing accuracies of CNN models are presented in Table 3. The ResNet-18 model has the best performance. The average testing accuracy is 80.4% for VGG-19, 82.1% for GoogLeNet, and 82.8% for ResNet-18. The average accuracy for the three models ranges between 80% and 83%, indicating the acceptable performance of models.

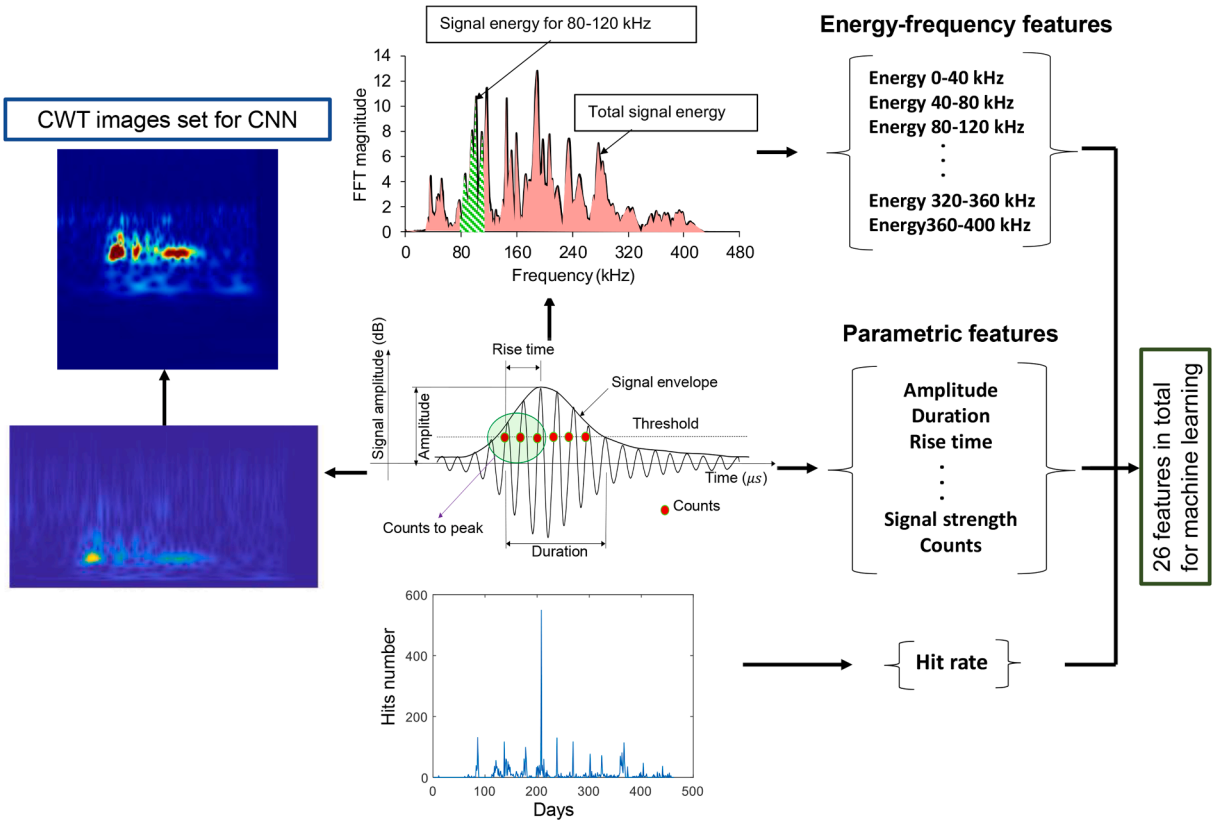

Fig. 12. Overview of the data preparation.

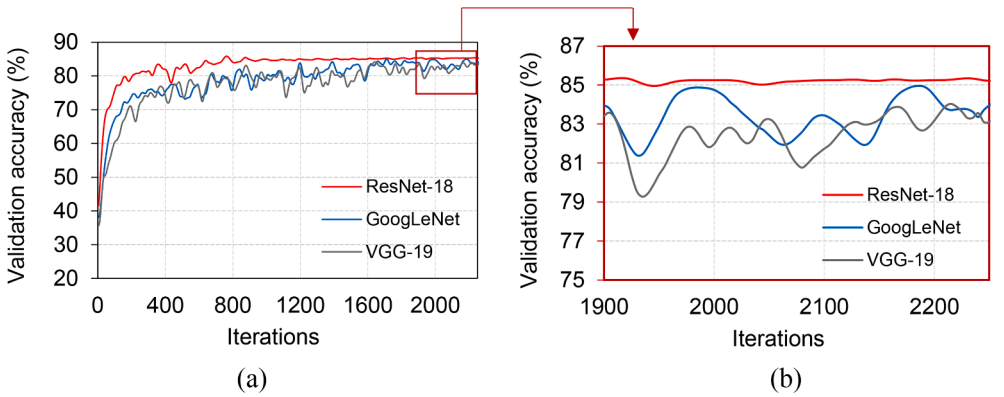

Fig. 13. Validation accuracy curves: (a) iteration 1–2250; (b) iteration 1900–2250.

Table 3
Testing accuracies of the three CNN models.

Trials	VGG-19	GoogLeNet	ResNet-18
Trial 1	80.2%	81.4%	84.8%
Trial 2	82.3%	78.9%	82.3%
Trial 3	79.4.%	84.6%	83.7%
Trial 4	83.4%	83.2%	83.0%
Trial 5	76.9%	82.3%	80.3%
Average Accuracy	80.4%	82.1%	82.8%

The accuracy distributions of CNN models (ResNet-18, VGG-19, and GoogLeNet) were estimated by testing five trials for each model, as shown in Fig. 14. The boxes illustrate the interquartile ranges in the figure. The interquartile range describes the middle 50% of values when ranking from low to high. The VGG-19 model has the highest interquartile range, and ResNet-18 has the lowest interquartile range. The standard deviations of the testing accuracies were also calculated for three models: 0.168 for ResNet-18, 0.213 for GoogLeNet, and 0.254 for VGG-19. The ResNet-18 has the highest average testing accuracy, the lowest standard deviation. Therefore, it could be an optimal option among the models if a single CNN model is used to evaluate the ASR progression in concrete structures.

The feature extraction process of ResNet-18 was visualized by showing the activations of feature maps in hidden layers when the testing accuracy was the highest (84.8%) among the five trials. In Fig. 15, the first convolutional layer (Conv1), two convolutional layers in the middle (Res2b_branch2b, and Res4a_branch2b), the last convolutional layer (Res5b_branch2b), and the last FC layer were selected to show the feature maps for the four ASR phases. Each feature map includes several smaller parts, named tiles in this paper, which are the output of different channels. The tiles with the strongest channel are zoomed and highlighted by red boxes in Fig. 15. White pixels in the feature map indicate strong positive activations. Black pixels are strong negative activations. The position of a pixel in the feature map corresponds to the same position in the input image. A white pixel at some locations in a channel indicates that the channel was strongly activated at that position. The first convolution layer learned basic features such as the outlines of the time–frequency components in the CWT images. More complicated features were learned in the deeper layers. The 4×1 vectors at the last FC layer show different color combinations for different phases. The vectors were transferred to a SoftMax layer for the classification.

There were 441 CWT images in the testing dataset. Therefore, 441 high-dimensional feature maps were generated by each layer. The feature maps were selected from the first convolutional layer, two convolutional layers in the middle, the last convolutional layer, and the last FC layer of the three CNN models. The selected feature maps were reduced to two dimensions by using t-distributed stochastic neighbor embedding (t-SNE). The t-distributed stochastic neighbor embedding (t-SNE) is an unsupervised dimension reducing technique that could embed high-dimensional data to a low-dimensional space to visualize the data in a cartesian coordinates system [67]. The reduced-dimensional data are presented in Fig. 16. Each data point in the figure is referred to as an embedding feature map dataset and was colored according to the corresponding ASR phase. In the first convolution layer, data points corresponding to four phases were scattered and mixed. The data points became more concentrated as the features were extracted in the deeper layers (shown from left to right in Fig. 16). Finally, the data points were divided into four classes (phases) in the last layer (FC layer). The scatter graphs are consistent with the testing accuracy results, as seen in Fig. 16. The four final clusters were clearly more separated for ResNet-18 model compared to the other models.

4.4. ASR evaluation using single random forest model

The performance of the single random forest model was evaluated in this section. The number of decision trees was initially set to 500. Testing data was 10% of the feature-based dataset (441 samples). The signal in the testing sets for both the random forest model and the CWT model were kept the same for the comparison purpose between the models. The remaining data (90% of the feature-based dataset, 3972 samples) were employed as the training set. The resulting OOB accuracy was 72.8%. The reason for the low accuracy might be due to some unrelated features among the 26 features. Therefore, a feature selection was employed to remove the unrelated features. One of the advantages of the random forest is that the importance of features can be calculated during the training [65]. The rank of features according to the importance could thereby be obtained. Fig. 17a presents the rank of all the 26 features in the feature-based dataset. The importance of the features "Hit rate" is significantly higher than the rest, indicating the importance of the feature. The feature "Energy", and "PCNTs" hold the lowest importance. Based on the feature ranking, 25 subsets of features were created by using backward elimination. The number of features in the subsets varies from 1 to 25. For example, the subset with one feature has

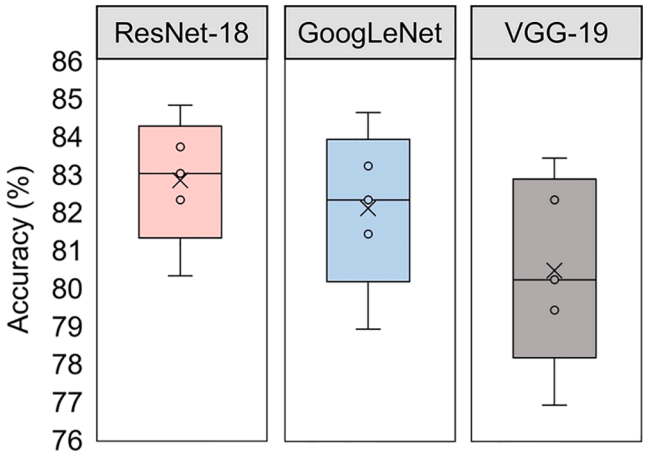

Fig. 14. Boxplot of the testing accuracies of three CNN models.

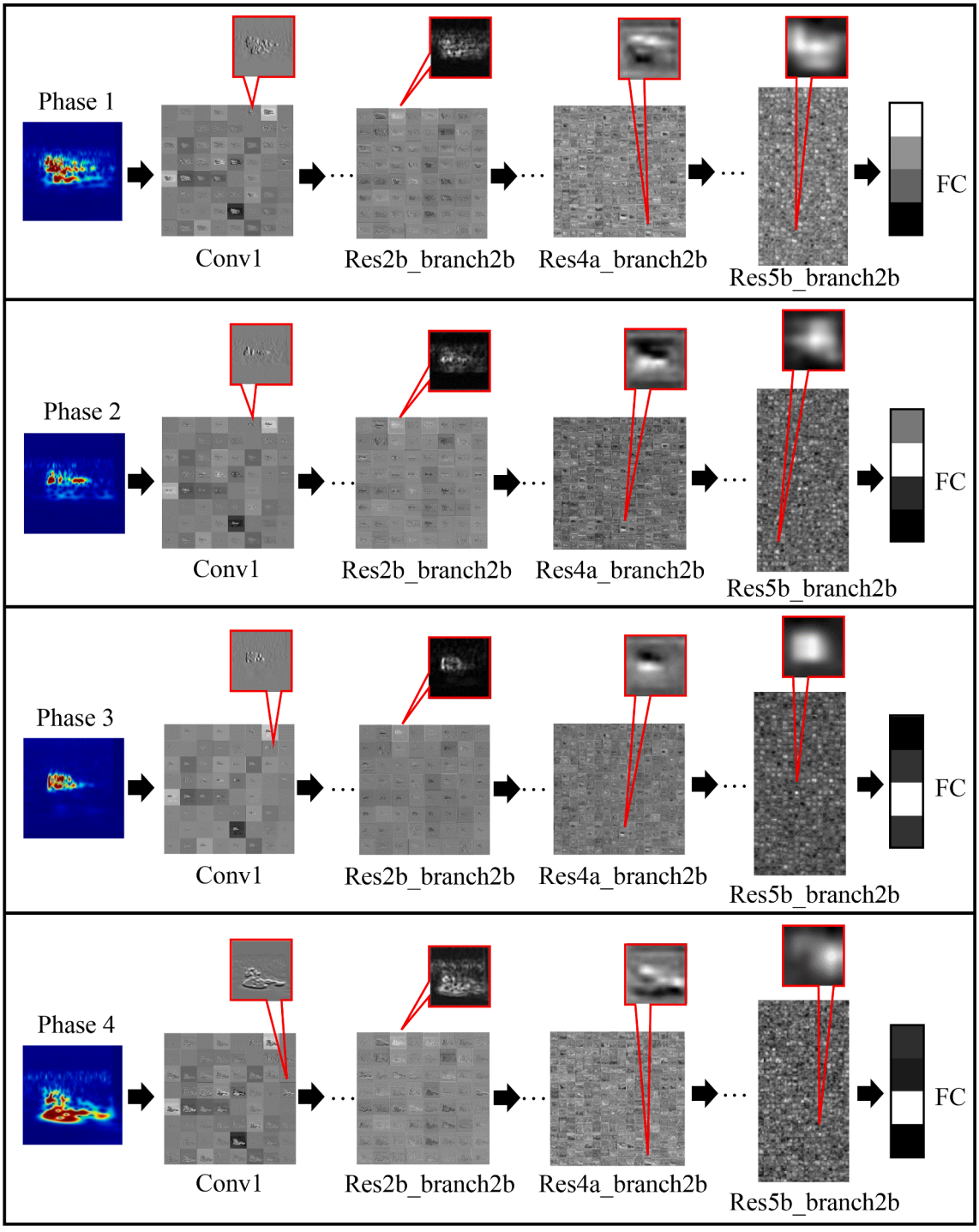

Fig. 15. Visualization of feature extraction process of ResNet-18 in four phases.

"Hit rate", the subset with two features has "Hit rate" and "Duration".

The optimum number of the decision tree was determined before evaluating the performance of the random forest model. Fig. 17b presents the OOB errors of all 25 subsets when the number of decision trees increases from 1 to 500. The OOB errors of all the models trained by the 25 subsets decrease rapidly when the number of trees increases. The errors remain almost constant as the number of decision trees increases to 200. Applying too many decision trees does not reduce the error but increases the computing time. Therefore, an appropriate number of decision trees was set as 200 in the final model.

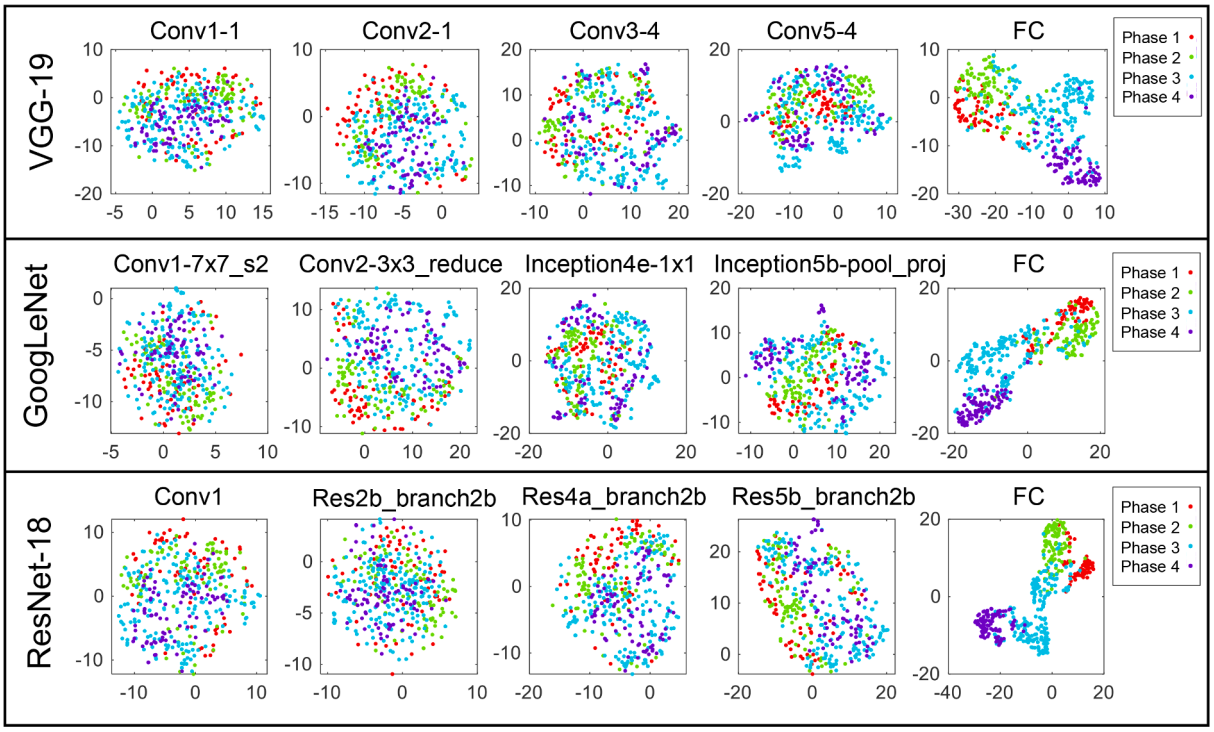

Fig. 16. Visualization of the activation distribution through t-SNE.

The results of training are presented in Fig. 18. The highest OOB accuracy (82.3%) is obtained using the first five features. Therefore, the subset with "Hit rate", "Duration", "Energy 0–40 kHz", "Frequency centroid", and "Energy 80–120 kHz" was utilized as the selected features for the final model. A random forest with 200 decision trees was trained using the first five features. The model was employed to evaluate the performance in the testing set. The model was trained and tested with five trials. Bootstrapping was implemented before training, 90% of the samples (3178 samples) were randomly selected from the training set as the true input set to train the random forest model. The testing accuracies for the five trials were: 83.9%, 81%, 82.8%, 82.8%, and 84.1%. The average accuracy was 82.9%, and the standard deviation was 0.0123. The performance of this random forest model is close to ResNet-18. A lower standard deviation was observed because the random forest is a bagging ensemble algorithm that can decrease the standard deviation.

According to the rank of features, the hit rate has been proven to be a significant feature for the model to classify AE signals recorded during ASR. However, the hit rate of AE signals can only be calculated if the AE monitoring of concrete is not less than 24 h. The accuracy of random forest with and without using the hit rate is presented in Table 4. The testing accuracy decreased from 82.9% to 50.3% when the hit rate was deleted from the features.

The two approaches have similar performance in the AE signal classification and can be used for the ASR evaluation. However, the random forest model with the hit rate feature has a limitation of 24-hours monitoring intervals contrary to the CNN models.

4.5. ASR evaluation using the proposed heterogeneous ensemble network

The performance of the proposed RGVF-HeteroESM-Net was evaluated in this section. All 20 sub-models in blocks 1, 2, 3, and 4 work independently. As mentioned before, the diversity of the sub-models was ensured by applying the bootstrapping method. After training all models, the ensemble network was tested. The testing result is presented as a confusion matrix in Fig. 19. The overall accuracy was 93.0%. The number of AE signals correctly classified to their corresponding ASR phase is shown in the main diagonal of the matrix. Precision and recall rates are employed as parameters to evaluate the classification performance in each phase. Generally, the precision rate can be calculated by Eq. (7):

$$Precision = \frac{TP}{TP + FP} \tag{7}$$

where TP, is the true positive, referring to the number of samples correctly classified to the attributed class. FP is the false positive, which refers to the number of samples that were misclassified into a class. The classification precision rate of phases 1 to 4 are 93.8%, 92.2%, 91.2%, and 96.7%, respectively.

The recall rate can be calculated as follow:

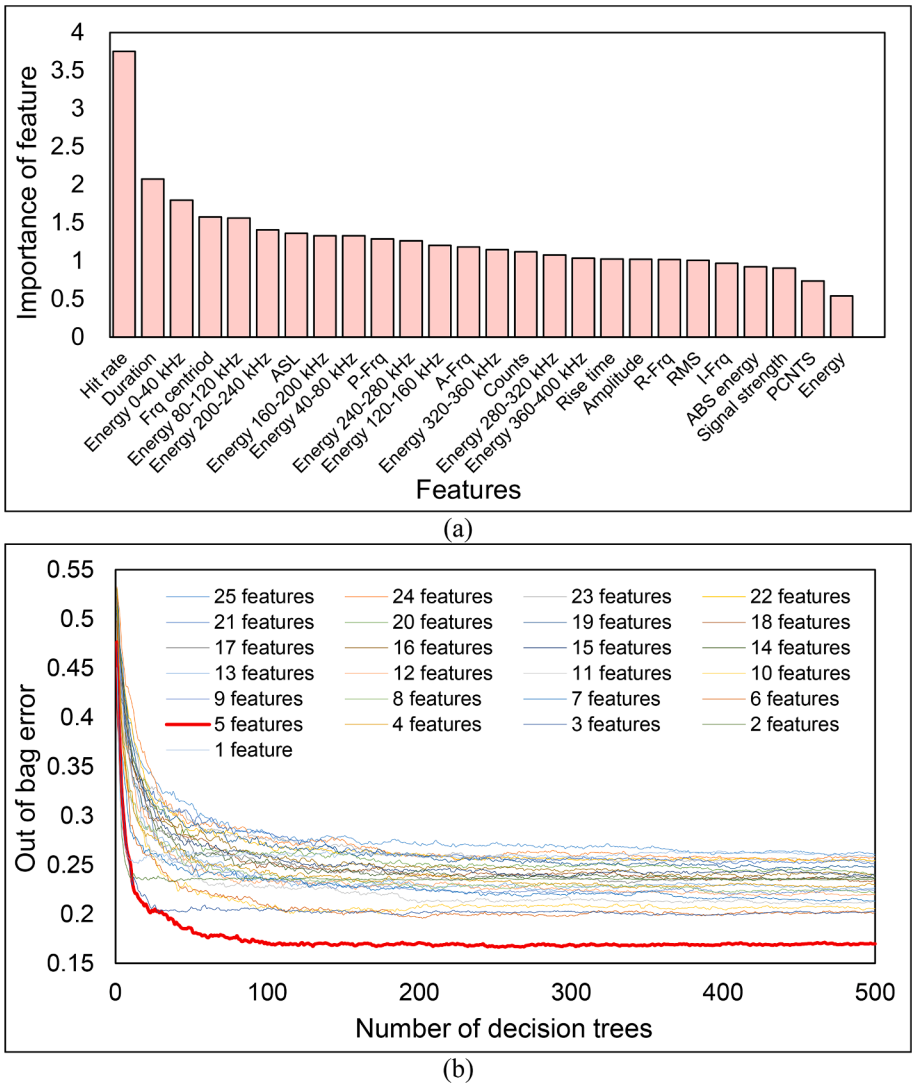

Fig. 17. Evaluation of random forest. (a) rank of features; (b) OOB versus the number of decision trees.

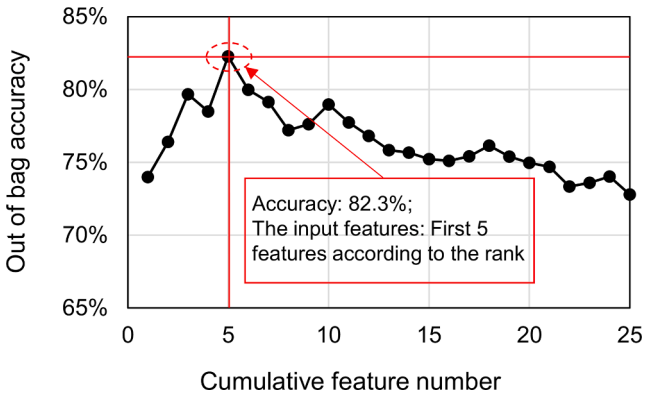

Fig. 18. Testing results of random forest using 25 subsets.

Table 4
Testing accuracies of the random forest.

Model	Testing accuracy
RF with hit rate input	82.9%
RF without hit rate input	50.3%

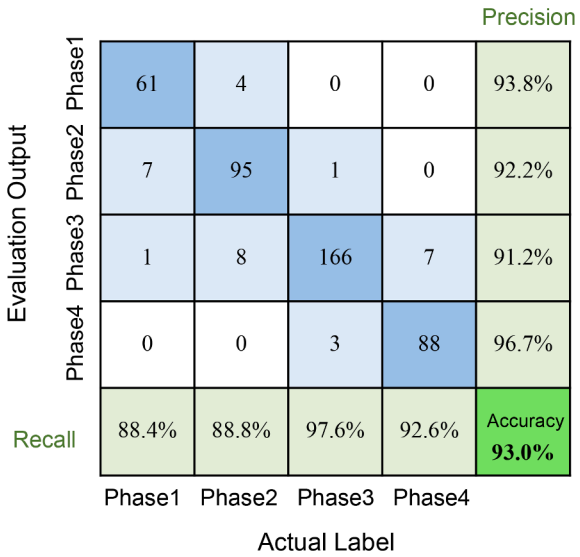

Fig. 19. The confusion matrix of RGVF-HeteroESM-Net.

$$Recall = \frac{TP}{TP + FN} \tag{8}$$

where *FN* is the false negative, which is the number of samples that belong to a class but are misclassified in other classes. The classification recall rate of phase 1 to 4 are 88.4%, 88.8%, 97.6%, and 92.6%, respectively.

The minimum AE monitoring time required for the RGVR-HeteroESM-Net is 24 h because the random forest with the hit rate feature was included in the model. An additional heterogeneous ensemble network without random forest was created, and the performance was tested. This ensemble net was composed of five resNet-18, five GoogLeNet, and five VGG-19 models, which was named "RGV-HeteroESM-Net". The testing accuracy was 89.6 % for this model (Table 5).

Four ensemble networks included: five ResNet-18 in block 3 (ResNet-18-ESM), five GoogLeNet in block 2 (GoogLeNet-ESM), five VGG-19 in block 1 (VGG-19-ESM), and five random forests in block 4 (RF-ESM) were also created and tested. Their testing accuracies were 87.5%, 86.2%, 84.4, and 85.3%, respectively (Table 5). The accuracies of the ensemble networks, the single CNN models, and the single random forest model are presented in Table 5. As seen in the table, the accuracies of ensemble networks are more than the corresponding single models, indicating the performance improvement of models using the heterogeneous ensemble strategy. The RGVF-HeteroESM-Net had the highest accuracy (93.0%) among the developed models.

Two shallow machine learning methods, SVM and KNN were also tested and compared with the models proposed. The accuracy

Table 5
Testing accuracies and computing times of all the models.

Model	Accuracy	Training time (s)	Testing time (s)
RGVF-HeteroESM-Net	93.0%	30111.14	0.97
RGV-HeteroESM-Net	89.6%	30083.81	0.91
ResNet-18-ESM	87.5%	4498.87	0.23
GoogLeNet-ESM	86.2%	7641.29	0.27
VGG-19-ESM	84.4%	17943.65	0.41
RF-ESM	85.3%	33.17	0.06
ResNet-18	82.8%	899.77	0.05
GoogLeNet	82.1%	1528.26	0.05
VGG-19	80.4%	3588.73	0.08
RF	82.9%	6.63	0.01
SVM	63.7%	3.39	0.01
KNN	57.1%	2.47	0.02

results are also presented in Table 5. The input data of the machine learning models are the AE feature-based dataset (26 features). The RBF was selected as the kernel function of SVM [68]. The optimized number of neighborhood "K" was set to 6 for the KNN model after several trials. The accuracy of SVM (63.7%) and KNN (57.1%) were much lower than the single models and the ensemble networks.

The high accuracy (93%) of the classification demonstrates that the proposed ensemble model can effectively distinguish the phases of ASR damage by using AE signals. This model is expected to apply to the concrete structures affected (or will be affected) by ASR. Several sensors are attached to the structure at the region with a higher risk of degradation, and AE signals are acquired for a period of time. Then the signals are used as an input of the model, and the model is expected to determine the current ASR damage phase with relatively high reliability, although the model should be validated and improved for the data collected from the different specimens before field application. Computation time is also a point of great concern in practical applications. It determines whether a model is capable of running within a reasonable time frame. The computational times of all the models are also compared and presented in Table 5. The testing time in Table 5 refers to the time to classify a single signal in the test dataset. According to the results, the training times for the ensemble networks were higher than the single models. The training time for the CNN models (ResNet-18, GoogLeNet, VGG-19) is more than the training times for the machine learning models (random forest, SVM, KNN). Among the three CNN models, ResNet-18 has the lowest training time while VGG-19 has the highest. In the field application, the models will be trained offline with historical AE signals, and therefore, the training time will not be a primary concern. The classification of ASR phases in the field would be conducted based on the trained model, meaning that the testing time is the primary factor to be considered. The differences between times for the trained models to classify a single AE signal are negligible. All the trained models can finish the classification within 1 s. Therefore, considering the accuracy and computational time, the RGVF-HeteroESM-Net had the highest accuracy (93.0%) and an acceptable testing time (0.97 s).

5. Summary and conclusions

This paper proposes a temporal ASR evaluation method based on AE monitoring and an ensemble learning framework. A concrete block with reactive aggregates was used as an experimental specimen. The ASR process was accelerated by providing high temperature and humidity in the chamber. The ASR expansion was measured using DEMEC gauge on a regular basis, and the crack width was measured by a Dino-Lite digital microscope. The AE data were continuously acquired for 460 days and were divided into four phases, defined in terms of time. Two types of input datasets, including the CWT images and AE features, were created based on the recorded AE signals. A heterogeneous ensemble learning network composed of ResNet-18, GoogLeNet, VGG-19, and random forest was developed to classify the data into the ASR phases with high accuracy. The proposed methodology showed an acceptable classification performance and computational time on the testing dataset. This indicates that the ASR damage phase can be potentially determined for the concrete structures in the field. The main conclusions of the paper are summarized as follows:

The classification performance of single models such as ResNet-18, GoogLeNet, VGG-19, and random forest were tested. The results depicted that ResNet-18 and the random forest model were more accurate than other models.

The hit rate is an important feature to be considered in the machine learning methods for the ASR AE data. However, this feature was calculated based on a monitoring time not less than 24 h. Low classification accuracy was obtained when the random forest model was trained without the hit rate.

The classification performance of the ensemble networks in this paper was investigated. The results suggest that the ensemble networks are more accurate than the single models. Among all the ensemble networks, RGVF-HeteroESM-Net has the best performance.

Effect of source/sensor distance on the AE model result can be a potential future work to improve the model. The propagation distance between the sources and the sensor is not always the same, which can influence the frequency content and other signal features [69,70]. This can be more significant when the developed method is used for large-scale structures. One of the challenges in the application of the proposed method in the field can be the result generalization. Future research should focus on the study of the result generalization by testing the trained model on different concrete specimens such as specimens with different or same materials and boundary conditions.

6. Data availability

The raw/processed data required to reproduce these findings can be made available upon request and with the written permission of the sponsor.

Declaration of Competing Interest

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

Acknowledgments

This research was partially supported by the U.S Department of Energy-Nuclear Energy University Program (NEUP) under the contract DE-NE0008544.

References

[1] E. Garcia-Diaz, J. Riche, D. Bulteel, C. Vernet, Mechanism of damage for the alkali–silica reaction, Cem. Concr. Res. 36 (2) (2006) 395–400.
[2] T. Ichikawa, M. Miura, Modified model of alkali-silica reaction, Cem. Concr. Res. 37 (9) (2007) 1291–1297.
[3] M. Thomas, The effect of supplementary cementing materials on alkali-silica reaction: A review, Cem. Concr. Res. 41 (12) (2011) 1224–1231.
[4] V. Soltangharaei, R. Anay, N. Hayes, L. Assi, Y. Le Pape, Z. Ma, P. Ziehl, Damage mechanism evaluation of large-scale concrete structures affected by alkali-silica reaction using acoustic emission, Applied Sciences 8 (11) (2018) 2148, https://doi.org/10.3390/app8112148.
[5] F. Rajabipour, E. Giannini, C. Dunant, J.H. Ideker, M.D.A. Thomas, Alkali–silica reaction: current understanding of the reaction mechanisms and the knowledge gaps, Cem. Concr. Res. 76 (2015) 130–146.
[6] H.N.G. Wadley, R. Mehrabian, Acoustic emission for materials processing: a review, Materials Science and Engineering 65 (2) (1984) 245–263.
[7] C.B. Scruby, An introduction to acoustic emission, J. Phys. E: Sci. Instrum. 20 (8) (1987) 946–953.
[8] C. Grosse, M. Ohtsu (Eds.), Acoustic Emission Testing, Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.
[9] V. Soltangharaei, J. Hill, L. Ai, R. Anay, B. Greer, M. Bayat, P. Ziehl, Acoustic emission technique to identify stress corrosion cracking damage, Structural Engineering and Mechanics 75 (2020) 723–736.
[10] D. Li, K.S.C. Kuang, C.G. Koh, Rail crack monitoring based on Tsallis synchrosqueezed wavelet entropy of acoustic emission signals: A field study, Structural Health Monitoring 17 (6) (2018) 1410–1424.
[11] M. Abdelrahman, M. ElBatanouny, K. Dixon, M. Serrato, P. Ziehl, Remote Monitoring and Evaluation of Damage at a Decommissioned Nuclear Facility Using Acoustic Emission, Applied Sciences 8 (9) (2018) 1663, https://doi.org/10.3390/app8091663.
[12] Q. Hao, Y.i. Shen, Y. Wang, J. Liu, An adaptive extraction method for rail crack acoustic emission signal under strong wheel-rail rolling noise of high-speed railway, Mech. Syst. Sig. Process. 154 (2021) 107546, https://doi.org/10.1016/j.ymssp.2020.107546.
[13] T.B. Quy, J.-M. Kim, Crack detection and localization in a fluid pipeline based on acoustic emission signals, Mech. Syst. Sig. Process. 150 (2021) 107254, https://doi.org/10.1016/j.ymssp.2020.107254.
[14] J. Tai, X. Liu, X. Wang, Y. Shan, T. He, An adaptive localization method of simultaneous two acoustic emission sources based on energy filtering algorithm for coupled array signal, Mech. Syst. Sig. Process. 154 (2021) 107557, https://doi.org/10.1016/j.ymssp.2020.107557.
[15] M. Abdelrahman, M.K. ElBatanouny, P. Ziehl, J. Fasl, C.J. Larosche, J. Fraczek, Classification of alkali–silica reaction damage using acoustic emission: A proof-of-concept study, Constr. Build. Mater. 95 (2015) 406–413.
[16] V. Soltangharaei, R. Anay, L.i. Ai, E.R. Giannini, J. Zhu, P. Ziehl, Temporal Evaluation of ASR Cracking in Concrete Specimens Using Acoustic Emission, J. Mater. Civ. Eng. 32 (10) (2020) 04020285, https://doi.org/10.1061/(ASCE)MT.1943-5533.0003353.
[17] M.I. Jordan, T.M. Mitchell, Machine learning: Trends, perspectives, and prospects, Science 349 (6245) (2015) 255–260.
[18] I. Goodfellow, Y. Bengio, A. Courville, Machine learning basics, Deep learning 1 (2016) 98–164.
[19] V. Cherkassky, Y. Ma, Practical selection of SVM parameters and noise estimation for SVM regression, Neural networks 17 (1) (2004) 113–126.
[20] I.W. Tsang, J.T. Kwok, P.-M. Cheung, N. Cristianini, Core vector machines: Fast SVM training on very large data sets, Journal of Machine Learning Research 6 (2005).
[21] H. Zhang, A.C. Berg, M. Maire, J. Malik, SVM-KNN: Discriminative nearest neighbor classification for visual category recognition, IEEE (2006) 2126–2136.
[22] M. Belgiu, L. Drăguţ, Random forest in remote sensing: A review of applications and future directions, ISPRS J. Photogramm. Remote Sens. 114 (2016) 24–31.
[23] R. Sun, Y. Chen, A. Dubey, P. Pugliese, Hybrid electric buses fuel consumption prediction based on real-world driving data, Transportation Research Part D: Transport and Environment 91 (2021) 102637, https://doi.org/10.1016/j.trd.2020.102637.
[24] L. Ai, V. Soltangharaei, R. Anay, M.J. van Tooren, P. Ziehl, Data-Driven Source Localization of Impact on Aircraft Control Surfaces, IEEE (2020) 1–10.
[25] L.i. Ai, V. Soltangharaei, M. Bayat, B. Greer, P. Ziehl, Source localization on large-scale canisters for used nuclear fuel storage using optimal number of acoustic emission sensors, Nucl. Eng. Des. 375 (2021) 111097, https://doi.org/10.1016/j.nucengdes.2021.111097.
[26] L. Ai, V. Soltangharaei, W.d. Backer, P. Zieh, M.v. Tooren,2020. A Minimally Intrusive Impact Detection System for Aircraft Moveable using Random Forest, CAMX 2020.
[27] F. Albu, A. Mateescu, N. Dumitriu, 1997, Architecture selection for a multilayer feedforward network, 131-134.
[28] B. Dorizzi, J. Mota, F. Albu, 1997, A step towards equalization for radiomobile channel: Neural Networks and Variable Selection.
[29] I. Goodfellow, Y. Bengio, A. Courville, Y. Bengio, Deep learning, MIT press Cambridge, 2016.
[30] A. Krizhevsky, I. Sutskever, G.E. Hinton, 2012, Imagenet classification with deep convolutional neural networks, 1097-1105.
[31] L. Ai, V. Soltangharaei, P. Ziehl,2021. Evaluation of ASR in concrete using acoustic emission and deep learning, Nuclear Engineering and Design 380 111328.
[32] F. Guo, Y.u. Qian, Y. Shi, Real-time railroad track components inspection based on the improved YOLOv4 framework, Autom. Constr. 125 (2021) 103596, https://doi.org/10.1016/j.autcon.2021.103596.
[33] F. Guo, Y.u. Qian, Y. Wu, Z. Leng, H. Yu, Automatic railroad track components inspection using real-time instance segmentation, Comput.-Aided Civ. Infrastruct. Eng. 36 (3) (2021) 362–377.
[34] D. Panigrahy, P.K. Sahu, F. Albu, Detection of ventricular fibrillation rhythm by using boosted support vector machine with an optimal variable combination, Comput. Electr. Eng. 91 (2021) 107035, https://doi.org/10.1016/j.compeleceng.2021.107035.
[35] A. Taherkhani, G. Cosma, T.M. McGinnity, AdaBoost-CNN: An adaptive boosting algorithm for convolutional neural networks to classify multi-class imbalanced datasets using transfer learning, Neurocomputing 404 (2020) 351–366.
[36] A. Ebrahimkhanlou, B. Dubuc, S. Salamone, A generalizable deep learning framework for localizing and characterizing acoustic emission sources in riveted metallic panels, Mech. Syst. Sig. Process. 130 (2019) 248–272.
[37] L.i. Ai, V. Soltangharaei, M. Bayat, M. Van Tooren, P. Ziehl, Detection of impact on aircraft composite structure using machine learning techniques, Meas. Sci. Technol. 32 (8) (2021) 084013, https://doi.org/10.1088/1361-6501/abe790.
[38] D. Li, Y. Wang, W.-J. Yan, W.-X. Ren, Acoustic emission wave classification for rail crack monitoring based on synchrosqueezed wavelet transform and multi-branch convolutional neural network, Structural Health Monitoring 20 (4) (2021) 1563–1582.
[39] A. Nasiri, J. Bao, D. Mccleeary, S.-Y. Louis, X. Huang, J. Hu, Online Damage Monitoring of SiC f-SiC m Composite Materials Using Acoustic Emission and Deep Learning, IEEE Access 7 (2019) 140534–140541.
[40] S.-X. Chen, L.u. Zhou, Y.-Q. Ni, X.-Z. Liu, An acoustic-homologous transfer learning approach for acoustic emission–based rail condition evaluation, Structural Health Monitoring 20 (4) (2021) 2161–2181.
[41] W. Wallau, S. Pirskawetz, K. Voland, B. Meng, Continuous expansion measurement in accelerated concrete prism testing for verifying ASR-expansion models, Mater. Struct. 51 (2018) 1–15.
[42] Retrieved May 09 from (2021). https://www.astm.org/Standards/C150.
[43] N.N. Hsu, 1977, Acoustic emissions simulator, Google Patents.
[44] J.N. van Rijn, G. Holmes, B. Pfahringer, J. Vanschoren, The online performance estimation framework: heterogeneous ensemble learning for data streams, Machine Learning 107 (1) (2018) 149–176.
[45] A. Ebrahimkhanlou, S. Salamone, Acoustic emission source localization in thin metallic plates: A single-sensor approach based on multimodal edge reflections, Ultrasonics 78 (2017) 134–145.
[46] J.M. Lilly, S.C. Olhede, Generalized Morse wavelets as a superfamily of analytic wavelets, IEEE Trans. Signal Process. 60 (11) (2012) 6036–6041.
[47] J.M. Lilly, S.C. Olhede, Higher-order properties of analytic wavelets, IEEE Trans. Signal Process. 57 (1) (2009) 146–160.
[48] S. Albawi, T.A. Mohammed, S. Al-Zawi, Understanding of a convolutional neural network, Ieee (2017) 1–6.
[49] M. Sun, Z. Song, X. Jiang, J. Pan, Y. Pang, Learning pooling for convolutional neural network, Neurocomputing 224 (2017) 96–104.
[50] B. Graham,2014. Fractional max-pooling, arXiv preprint arXiv:1412.6071.
[51] H. Nakahara, T. Fujii, S. Sato, A fully connected layer elimination for a binarizec convolutional neural network on an FPGA, IEEE (2017) 1–4.

[52] T. Shin, V. Soltangharaei, P. Ziehl, Y. Zhang,2019. Prediction of Volumetric Strain in Concrete Due to ASR Reactions Using Acoustic Emission Technique and Artificial Neural Network, Structural Health Monitoring 2019.
[53] N.C. Oza, S.J. Russell, Online bagging and boosting, PMLR (2001) 229–236.
[54] L. Wen, X. Li, X. Li, L. Gao, A new transfer learning based on VGG-19 network for fault diagnosis, IEEE (2019) 205–209.
[55] Z. Ullah, B.A. Lodhi, J. Hur, Detection and identification of demagnetization and bearing faults in PMSM using transfer learning-based VGG, Energies 13 (15) (2020) 3834, https://doi.org/10.3390/en13153834.
[56] G. Fan, J. Li, H. Hao, Vibration signal denoising for structural health monitoring by residual convolutional neural networks, Measurement 157 (2020) 107651, https://doi.org/10.1016/j.measurement.2020.107651.
[57] G. Cao, K. Zhang, K. Zhou, H. Pan, Y. Xu, J. Liu, A Feature Transferring Fault Diagnosis based on WPDR, FSWT and GoogLeNet, IEEE, 2020, pp. 1–6.
[58] H. Xin, L. Cheng, R. Diender, M. Veljkovic, Fracture acoustic emission signals identification of stay cables in bridge engineering application using deep transfer learning and wavelet analysis, Advances in Bridge Engineering 1 (2020) 1–16.
[59] F. König, G. Jacobs, A. Stratmann, D. Cornel, Fault detection for sliding bearings using acoustic emission signals and machine learning methods, IOP Publishing 1097 (1) (2021) 012013, https://doi.org/10.1088/1757-899X/1097/1/012013.
[60] F. König, C. Sous, A. Ouald Chaib, G. Jacobs, Machine learning based anomaly detection and classification of acoustic emission events for wear monitoring in sliding bearing systems, Tribol. Int. 155 (2021) 106811, https://doi.org/10.1016/j.triboint.2020.106811.
[61] S. Abney, 2002, Bootstrapping, 360-367.
[62] K. Simonyan, A. Zisserman,2014. Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv:1409.1556.
[63] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, 2015, Going deeper with convolutions, 1-9.
[64] K. He, X. Zhang, S. Ren, Deep residual learning for image recognition, J. Sun (2016) 770–778.
[65] L. Breiman, Random forests, Machine learning 45 (2001) 5–32.
[66] L. Breiman,1996. Out-of-bag estimation.
[67] L. Van der Maaten, G. Hinton,2008. Visualizing data using t-SNE, Journal of machine learning research 9.
[68] B. Scholkopf, K.-K. Sung, C.J. Burges, F. Girosi, P. Niyogi, T. Poggio, V. Vapnik, Comparing support vector machines with Gaussian kernels to radial basis function classifiers, IEEE Trans. Signal Process. 45 (1997) 2758–2765.
[69] E. Maillet, C. Baker, G.N. Morscher, V.V. Pujar, J.R. Lemanski, Feasibility and limitations of damage identification in composite materials using acoustic emission, Compos. A Appl. Sci. Manuf. 75 (2015) 77–83.
[70] A. Farhidzadeh, A.C. Mpalaskas, T.E. Matikas, H. Farhidzadeh, D.G. Aggelis, Fracture mode identification in cementitious materials using supervised pattern recognition of acoustic emission features, Constr. Build. Mater. 67 (2014) 129–138.