IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received October 18, 2020, accepted November 2, 2020, date of publication November 4, 2020, date of current version November 18, 2020.

*Digital Object Identifier 10.1109/ACCESS.2020.3035884*

# Method for Classifying a Noisy Raman Spectrum Based on a Wavelet Transform and a Deep Neural Network

**LIANGRUI PAN[1], (Student Member, IEEE), PRONTHEP PIPITSUNTHONSAN[1], (Student Member, IEEE), CHALONGRAT DAENGNGAM[2], SITTIPORN CHANNUMSIN[3], SUWAT SREESAWET[3], AND MITCHAI CHONGCHEAWCHAMNAN[1], (Senior Member, IEEE)**

[1]Faculty of Engineering, Prince of Songkla University, Songkhla 90110, Thailand
[2]Faculty of Science, Prince of Songkla University, Songkhla 90110, Thailand
[3]Geo-Informatics and Space Technology Development Agency (GISTDA), Chonburi 20230, Thailand

Corresponding author: Mitchai Chongcheawchamnan (mitchai.c@psu.ac.th)

This work was supported by the Science, Research and Innovation Promotion Fund under Grant 1383848.

**ABSTRACT** Because it is relatively difficult in practice to classify the Raman spectrum under baseline noise and additive white Gaussian noise environments, this paper proposes a new framework based on a wavelet transform and deep neural network for identification of noisy Raman spectra. The framework consists of two main engines. Wavelet transform is proposed as the framework front end for transforming the 1-D noise Raman spectrum to two-dimensional data. The two-dimensional data are fed to the framework back end, which is a classifier. The optimum classifier is chosen by implementing several traditional machine learning (ML) and deep learning (DL) algorithms, and we investigate their classification accuracy and robustness performances. The four chosen MLs are naive Bayes (NB), a support vector machine (SVM), a random forest (RF) and a k-nearest neighbor (KNN), and a deep convolution neural network (DCNN) was chosen as a DL classifier. Noise-free, Gaussian noise, baseline noise, and mixed-noise Raman spectra were applied to train and validate the ML and DCNN models. The optimum back-end classifier was obtained by testing the ML and DCNN models with several noisy Raman spectra (10-30 dB noise power). Based on the simulation, the accuracy of the DCNN classifier is 9% higher than that of the NB classifier, 3.5% higher than the RF classifier, 1% higher than the KNN classifier, and 0.5% higher than the SVM classifier. In terms of robustness to mixed noise scenarios, the framework with the DCNN back end showed superior performance compared with the other ML back ends. The DCNN back end achieved 90% accuracy at 3 dB SNR, while the NB, SVM, RF, and K-NN back ends required 27 dB, 22 dB, 27 dB, and 23 dB SNR, respectively. In addition, in the low-noise test dataset, the F-measure score of the DCNN back end exceeded 99.1%, and the F-measure scores of the other ML engines were below 98.7%.

**INDEX TERMS** Raman spectrum, baseline noise, wavelet transform, deep convolution neural network, accuracy, robustness.

## I. INTRODUCTION

Raman spectroscopy is a material characterization method widely used in industrial process controls, planetary exploration, homeland security, life science, geological field investigation, and laboratory material research [1]. By identifying the Raman spectrum of a small number of substances, an accurate label of the substance can be obtained [2]. For example, in detection of minerals in the field, we might only sample all the minerals and perform experimental analysis on them. It is necessary to perform pre-processing to obtain Raman spectra, such as using Raman spectroscopy to check the composition of chemical substances and implement statistical classification methods.

The associate editor coordinating the review of this manuscript and approving it for publication was Dusmanta Kumar Mohanta.

Preferably, a rapid and accurate classification algorithm is required when dealing with a large Raman spectrum set. Currently, many chemical/biochemical molecular structure databases are available for researchers to access, such as the FT-Raman spectra database [3], an e-VISART database [4], a biomolecule database [5], and an explosive compound database [6]. These databases contain a large amount of raw and processed Raman data for Raman spectroscopy applications.

For practical Raman spectrum data, noise signals in the spectrum can originate from several sources such as the fluorescence process, material density, external light source, charge-coupled device receivers, external charge amplifiers, and environmental noise. In the signal processing aspect, we need to reduce these noises before performing classification. Without a process to reduce such noises, the Raman classification accuracy deteriorates. Hence, several methods were proposed, e.g., baseline correction [7] and surface enhancement [8], to name a few. Although some algorithms have been improved based on these methods, fully automated processing of Raman spectra remains a challenge. Random noise affects the peaks and the sub-peaks of the Raman spectra, which causes difficulty in the extraction of spectrum peaks and subpeaks and finally reduces the classification accuracy.

In recent decades, many automatic baseline correction algorithms have been based on the original Raman spectrum, such as the least square method, asymmetric least square method (AsLS), and the penalty least square method (PLS). Zhang *et al*. proposed an adaptive iterative reweighted penalty least squares (airPLS) algorithm without prior information, such as user intervention and peak detection [7]. He *et al*. proposed a baseline correction method for Raman spectrum correction with the improved asymmetric least squares (IAsLS) [9]. This algorithm estimates the original spectral line by a polynomial fitting method. Compared with AsLS, the root mean-square-error of IAsLs was reduced by 16 times, and the baseline can be automatically subtracted. Beak *et al*. proposed a weighted airPLS method by using the generalized logic function based on the baseline correction of the PLS, which estimates the noise level iteratively and adjusts the weight accordingly [10].

Apart from baseline correction, the noise level also has severe effects on the peak characteristics of the Raman spectrum. Ehrentreich *et al*. proposed a wavelet transform (WT) to identify the peak value through the first level detail coefficient [11]. The position of the spike can be projected from detail to approximation and subsequently to the appropriate position of the original spectrum. After the peak is determined, these regions are replaced by subtraction. Barclays *et al*. proposed a discrete WT for spectrum smoothing and denoising to remove small-amplitude components independent of position in the transform domain [12], and the method has excellent performance in an extended dynamic range. Guo *et al*. proposed a method combining the Mexican-hat wavelet and average algorithm to extract the Raman signal from high and low spectral noise [13]. This method has the characteristics of small relative errors of spectrum intensity and spectrum width. This previous work shows that the WT can effectively be applied to Raman spectroscopy.

Based on ML algorithms, several researchers have proposed better classification and recognition models, from a simple perceptron to a vast artificial neural network. Several researchers aimed to develop ML algorithms for classification problems, including the decision tree (DT), naive Bayes (NB), k-nearest-neighbor (K-NN), and support vector machine (SVM), etc. [1], [14]–[16]. DT (a graphic method of intuition using probability analysis [14]) is a decision analysis method used to calculate the probability that the expected value of the net present value is greater than or equal to zero. The DT is formed according to the probability of various situations. The greedy algorithm is applied to build the DT by considering only the condition of the maximum purity difference as the segmentation point [17]. The construction of the DT is a recursive process. In contrast, SVM aims to find a separate hyperplane in the feature space that divides different data instances into various labels to achieve classification [16]. This algorithm does not make any assumptions about the distribution of the original dataset, and thus, it is widely used in biomedical engineering, chemical materials, and physical spectra. Effendi *et al*. evaluated the ability of near-infrared Raman spectroscopy combined with SVM to improve the classification of different histopathological groups in tissues [18]. Two types of SVM (i.e., C-SVM and v-SVM) with three kernel functions, the linear, polynomial and Gaussian radial basis function (RBF), are used in combination with principal component analysis to develop an effective algorithm for classifying the Raman spectra of different colonic tissues. N. H. Othman *et al.* evaluated the ability to combine near-infrared Raman spectroscopy with SVM to improve the multiclass classification of different histopathological groups in tissues. A diagnostic accuracy of 99.9% was obtained for multiclass classification [4]. The disadvantages are that the efficiency is not very high if there are many observation samples. Other disadvantages are an absence of general solutions for nonlinear problems and sensitivity to missing data.

Traditional ML classifiers such as NB, RF, K-NN, and SVM were applied to Raman spectroscopy [18]–[22], and it was shown that NB has fast convergence. Julio *et al.* used Raman spectroscopy and Bayesian classifiers to classify breast biopsies of healthy and cancerous tissues with 100% accuracy [19]. The main disadvantage of the NB algorithm is its feature redundancy. The RF algorithm is simple to understand and interpretable. This algorithm was applied to analyze complex surface enhanced Raman scattering (SERS) data to obtain accurate and complex interpretations based on previous knowledge of the available SERS signals [20]. One of the disadvantages is that this method does not support online learning, and thus after the arrival of new samples, the decision tree, which is the RF algorithm structure, needs to be rebuilt. Another drawback of the RF algorithm is its

overfitting problem. On the other hand, the K-NN algorithm is suitable for automatic classification of class domains with a large sample size. In [21], K-NN was used to classify SERS data was able to supply early detection of dengue fever with a classification accuracy of 82.14%. The main disadvantage of this algorithm is that it easily generates misclassification for those class domains with small sample sizes where the output is also not interpretable. SVM can solve high-dimensional problems such as large-scale feature spaces, small sample sizes and interaction of nonlinear features. N. H. Othman *et al.* evaluated the ability to combine near-infrared Raman spectroscopy with SVM to improve the multiclass classification of different histopathological groups in tissues [18], and a diagnostic accuracy of 99.9% for multiclass classification was obtained [18]. The disadvantages of this approach are that its efficiency is not high if there are many observation samples. Other disadvantages include the absence of a general solution for nonlinear problems and sensitivity to missing data.

Many scholars have also proposed the use of a machine learning algorithm combined with other methods to construct a novel hybrid model or framework and obtain higher accuracy in practice. Considering the characteristics of attitude data, Jianyu *et al.* adopted a hybrid sparse auto-encoder (SAE) and support vector machine (SVM) approach and proposed construction of an intelligent fault diagnosis model by learning from the attitude dataset with multiple fault information [22]. Feng *et al.* proposed a continuous naive Bayesian learning framework for sentiment classification of product reviews on large-scale and multidomain e-commerce platforms and fine-tuned the learning distribution method based on three hypotheses to better adapt to different fields [23]. Mwaffaq *et al.* collected real-time user symptom data based on an Internet of Things framework and used machine learning (SVM, NB, KNN, RF, etc.) algorithms to accurately identify potential COVID-19 cases. The accuracy of the five algorithms was greater than 90% [24]. Wei *et al.* proposed a new unbalanced fault diagnosis framework based on cluster MWMOTE and the LS-SVM classifier based on moth flame optimization (MFO) and proved that the framework provides a higher fault diagnosis recognition rate and algorithm robustness [25].

Unlike the ML algorithms, the goal of the neural network (NN) method is to learn the feature level of high-level features composed of low-level feature groups [26]. Jianyu *et al.* designed the Sean method with sparse automatic encoder (SAE) and echo state network (ESN) for fault diagnosis, and its superiority was proved by comparison with other intelligent fault diagnosis technologies in experiments [27]. However, various depth neural networks (DNN) have been proposed to increase the depth of NN, and the accuracy is greater than 90%, which is better than that of the traditional NN. Weng *et al.* applied the migration network framework to biomedical engineering and used a coherent anti-Stokes Raman scattering image to diagnose lung cancer automatically [28]. When using this model to analyze other cancer cells, the accuracy of cancer cell image recognition is 89.2%, confirming that the DNN is a powerful image processing technology. Natalia *et al*. constructed a multilevel DL framework, the core of which is the unsupervised NN and a group of supervised NNs [29]. The accuracy of this particular DNN is 85% compared with a convolution neural network (CNN) in classification of land cover and crop types in multiband and multisource satellite images.

The main steps of traditional noise classification are data collection, data preprocessing, and model classification. In data preprocessing, baseline correction is the most common method for Raman spectroscopy. However, application of baseline correction algorithms might affect the desired Raman fingerprints, although baseline noise is reduced. Therefore, a baseline correction algorithm is not the best preprocessing algorithm for Raman spectroscopy. From the signal theory aspect, the Raman spectrum is a one-dimensional (1-D) signal, and hence 1-D CNN was proposed to identify a spectrum peak from a noisy Raman spectrum. M. Fukuhara *et al.* used a digitally generated Lorentz spectrum to determine the optimal filter size (close to line width) and the number of filters to extract Raman peaks. However, this method has many steps, and the extracted peaks were partially missing. The recognition accuracy of 1-D CNN is quite low in a relatively large noise environment (noise close to the sub-peak) [30]. Consequently, we propose a Raman spectrum classification algorithm based on a two-dimensional deep convolution neural network (2-D DCNN). The WT is proposed to transform 1-D noisy Raman spectra to a 2-D scape map [19]. All spectrum information and noise information on the noisy Raman spectrum are retained in the scale graph without loss. The 2-D data in the scale map domain are related to Raman shift and intensity. This 2-D DCNN model is trained with several datasets and is subsequently validated with other datasets for testing.

In the proposed framework, WT and DCNN should overcome the deficiencies of Raman spectroscopy classification in complex noise environments. The main contributions of this paper can be summarized as follows:

1) The Raman spectrum noise in real environments is simulated, and the method of data preprocessing by wavelet transform is proposed. This method can simultaneously extract the characteristics of the Raman shift domain and intensity domain in the Raman spectrum signals and transform the original image data into a $224 \times 224 \times 3$ multiresolution scale map.
2) A DCNN is proposed as the back-end classification framework. The DCNN extracts features from multiresolution scale maps, accelerates the training of neural networks and generates end-to-end DCNN classifiers without gradient explosion and overfitting. In this framework, the performance of the ML classifiers and DCNN classifiers is evaluated according to the precision, recall rate and test accuracy.
3) The proposed framework is applied to the classification of a mixed noise Raman spectrum. The performance

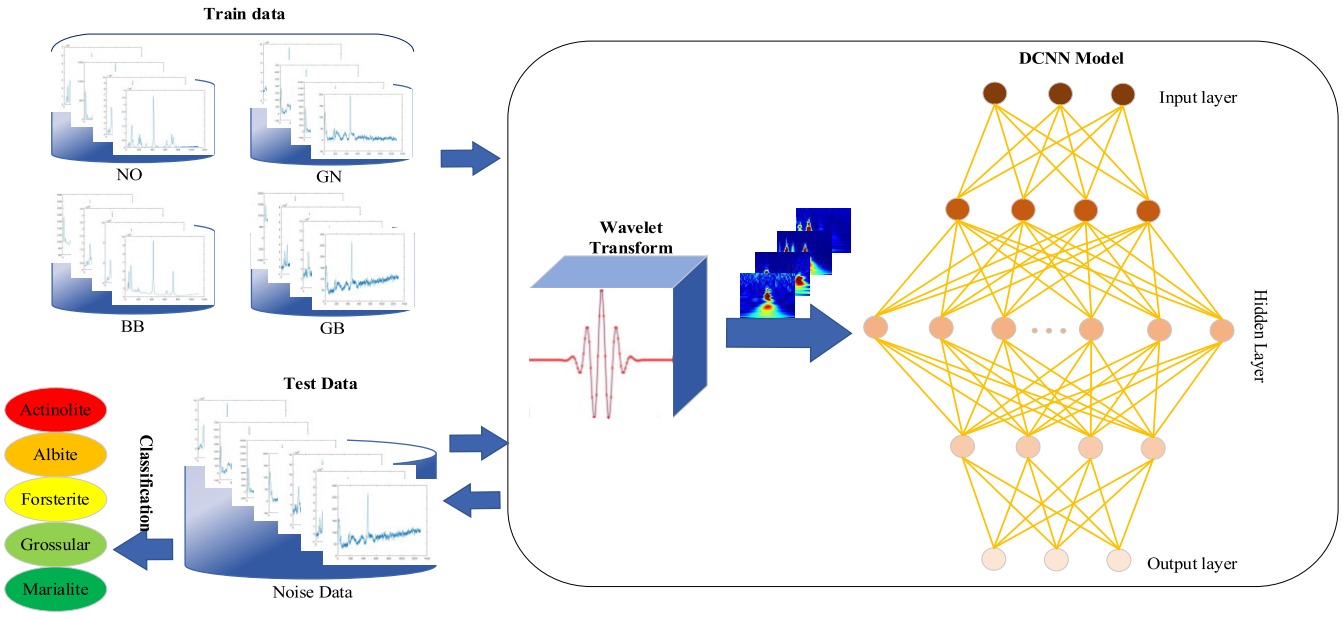


**FIGURE 1.** Proposed framework based on wavelet transform and DCNN.

of the proposed framework is verified on a mixed high noise Raman spectrum dataset. The accuracy, precision and robustness of the proposed framework are tested on the dataset. The experimental results show that compared with NB, SVM, RF, and K-NN, the proposed framework has better classification accuracy and stronger robustness.

## II. MATERIALS AND METHOD

This section describes a new framework based on a WT and a 2-D DCNN. Fig. 1 shows a diagram of the proposed framework, which consists of three stages. The first stage is data preparation. The input of the one-dimensional Raman spectrum is pre-processed by adding noise. Noisy Raman datasets were created and grouped into training and test groups and converted into 2D scale map data using the WT. The second stage is the training of the classifier. The training dataset is applied to a ML and 2-D DCNN algorithm, and the classifier based on ML and 2-D DCNN is obtained from this stage. Finally, the noisy Raman datasets with different noise levels are tested with different classifiers and compared with the traditional ML classifiers. Currently, the primary performance of DCNN is based on the research of a two-dimensional classifier. The following content is explained in detail from the dataset generation, wavelet transform and deep neural network.

### A. MATERIAL

RRUFF is a complete collection of high-quality Raman spectroscopy databases composed of 4051 well-defined minerals [1]. Raman spectrometers obtained these Raman spectra with laser wavelengths of 532 nm and 780 nm. In this paper, we chose the Raman spectrums of Actinolite, Albite, Forsterite, Grossular, and Marialite as the noiseless spectrum datasets. In the RRUFF database, there are at most 13 original Raman spectra of each material. In a practical environment, two noise types, i.e., baseline background noise (BBN) and additive white Gaussian noise (AWGN), are unavoidable. Hence we developed additional Raman datasets by adding BBNs and AWGNs to these original Raman spectra. These datasets are prepared for training and verification of the accuracy and robustness of the classifiers.

For BBNs, several noise patterns were created by formulating BBN spectra with a summation of multiple sinusoidal functions. The number of valleys in the BBN-contaminated Raman spectrum depends on the number of sinusoidal functions. The BBN patterns were randomly determined by the number of sinusoidal functions, the position and amplitude of the valley peak, and the width of each valley. Different baselines are equivalent to the fluorescence noise and the other types of shot noise from electronic devices. For AWGN, noise signals for different noise powers were created. Both noise signals, BBN and AWGN, are added to the noiseless Raman spectrum. The signal-to-noise ratio (SNR) is used as a parameter to quantify the noise. Theoretically, SNR is defined by:

$$\mathrm{SNR} = 10\log_{10}\frac{\mathrm{P_s}}{\mathrm{P_n}} \tag{1}$$

where $\mathrm{P_s}$ is the signal power, and $\mathrm{P_n}$ is the noise power. In this paper, noisy Raman signal datasets with SNR of 30 dB to 80 dB are used as the training datasets, and the noisy Raman signals with SNR of 1 dB to 30 dB are used as the test datasets.

Fig. 2 shows five noiseless Raman spectra of Actinolite retrieved from the RRUFF database. As shown in Fig. 2, the

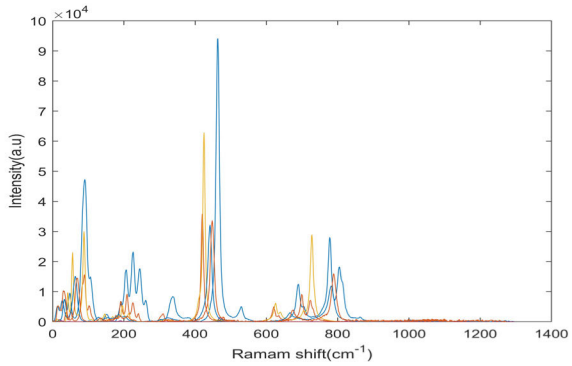


**FIGURE 2.** Original Raman spectra of Actinolite.

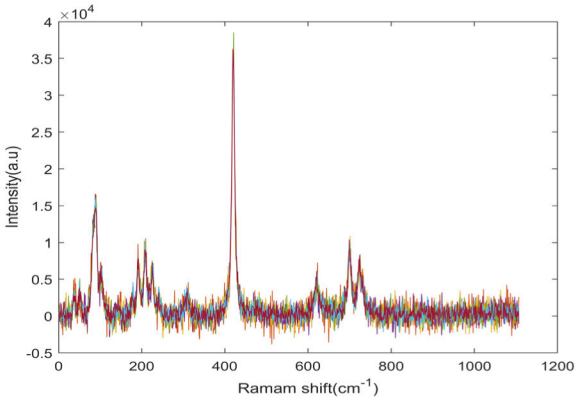


**FIGURE 3.** AWGN-contaminated Raman spectra of Actinolite.

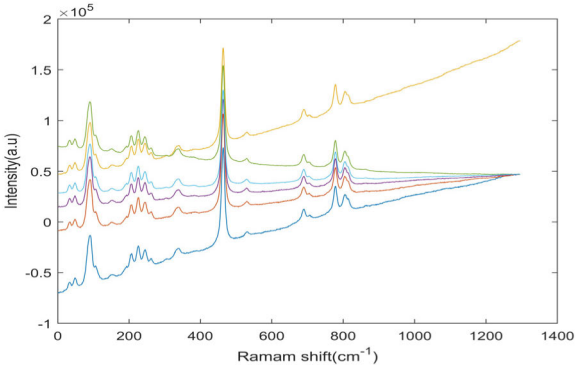


**FIGURE 4.** BBN-contaminated Raman spectra of Actinolite.

spectrum footprints are quite different, even though they were obtained from a single material. This phenomenon occurs naturally for spectrum patterns of all other materials. For Raman spectroscopy application, it is necessary to train a classifier with a sufficient number of input datasets for each material. In practice, the environment for Raman spectrum sensing is noisy, which poses a challenge for developing a classifier. Fig. 3 shows 1-D Raman spectra of Actinolite contaminated with AWGNs. Fig. 4 shows 1-D Raman spectra of Actinolite corrupted with different BBN patterns, and Fig. 5 shows 1-D Raman spectra contaminated with both BBNs and AWGNs.

Table 1 shows the number of spectrum datasets applied in this paper. Five materials of Actinolite, Albite, Forsterite,

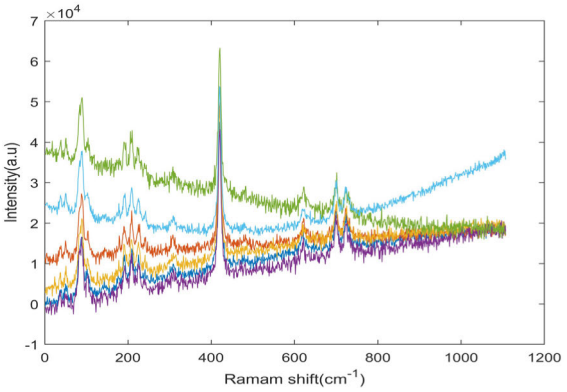


**FIGURE 5.** BBN- and AWGN-contaminated Raman spectra of Actinolite.

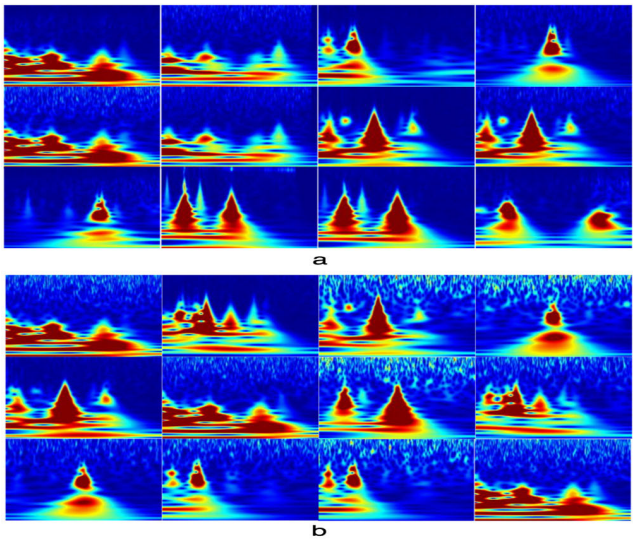


**FIGURE 6.** Scale map of the Raman spectrum after WT.

**TABLE 1.** Composition of the Raman spectrum dataset.

| Class | Type | Number of spectrum dataset | |
|---|---|---|---|
| | | Original | Noisy |
| 1 | Actinolite | 11 | 2534 |
| 2 | Albite | 13 | 2993 |
| 3 | Forsterite | 13 | 2993 |
| 4 | Grossular | 13 | 2993 |
| 5 | Marialite | 10 | 2381 |

Grossular, and Marialite are listed. A total of 60 original Raman spectra were retrieved from the RRUFF database. In this paper, we consider these original spectra as the noiseless Raman datasets. The details of the original spectrum of

each material are listed in Table 1. To train the algorithm, we created more datasets by adding BBNs and AWGNs to the noiseless spectra, leading to 13,894 datasets of noisy spectra. The number of datasets of each noisy spectrum class is shown in Table 1.

### B. WAVELET TRANSFORM

Many methods can be used to extract signal features. The Fourier transform (FT) is a well-known method for extracting signal information and describing it in a spectrum domain. Because FT can effectively extract information for a stationary signal [31], [32], it is not suitable for some cases, e.g., nonstationary signals, short-intervals or transient signals. Nonstationary signals such as electroencephalography and electrocardiography signals [33], [34] can be analyzed using short-time FT [35], which divides a whole time-domain signal into several short time windows and performs FT. For time-varying nonstationary signals, high frequency is suitable for a small window and low frequency is suitable for a large window. However, the window size should be chosen carefully. Time-varying nonstationary signals processed by fast FT can only obtain the frequency components of a signal, but it does not know when the components appear. Therefore, two signals with a massive time-domain difference might have the same spectrum. In other words, the FT method is not suitable for a nonstationary signal such as the Raman spectrum. For Raman spectroscopy applications, the measured Raman spectrum usually fluctuates, and spectrum shifting and spectrum peak oscillation occur due to environmental noise.

A WT was proposed to deal with a nonstationary signal and a noisy environment. Unlike the basis function of FT, which uses a trigonometric function, WT provides a new set of mathematic functions as a basis function. These basis functions are localized in both the time and spectrum domain, thus avoiding the Gibbs effect and achieving orthogonalization. Hence better reliable and detailed time-scaled signal information is obtained than with the FT [11], [13], [36]. Theoretically, the WT for a signal x (t) is performed by,

$$X(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} \psi \left( \frac{t - b}{a} \right) x(t)\, dt, \tag{2}$$

where $\psi(\cdot)$ is the wavelet basis function, and $a$ and $b$ are the scale and translation variables of the WT. It is shown in (2) that $a$ and $b$ control the dilation and the translation of $\psi(\cdot)$.

We propose use of the WT for Raman spectroscopy application. A Raman signal is transformed into a wavelet domain. In this paper, the Morlet wavelet with a center frequency of 1 is used to transform the signal. A large number of center frequencies are obtained by a scale transformation in which a series of basic functions in different intervals are obtained by Raman displacement. The basis function is integrated with the product of a particular segment (corresponding to the interval of the basis function) of the original signal. The frequency corresponding to the extreme value is the frequency contained in this region of the original signal. The WT can avoid the Gibbs effect and also realize orthogonalization. We divide the simulated Raman spectral data into training and testing sets and subsequently analyze them with wavelet multiresolution [37]. After performing the WT, the noise generates a white fringe area, and the higher the noise in the signal, the denser the white stripe area. The noise of the signal is positively correlated with the range of the stripe area.

### C. DEEP CONVOLUTION NEURAL NETWORK (DCNN)

Fig. 7 shows the proposed 2-D DCNN [38]. With the signal input size of 2-D DCNN of 224 × 224 pixels, the WT output is 224 × 224 pixels. The hidden layer in the proposed DCNN consists of a set of activation functions, a full connection layer and a pooling layer. A 7 × 7 convolution kernel and 3 × 3 matrix convolution are applied [39]. The kernel convolution size is chosen such that the speed and accuracy of the feature extraction process are obtained. Due to the large-scale convolution kernel in the extraction of features, noise is inevitably introduced. Use of a small convolution kernel in feature learning can reduce the error probability and also avoid errors caused by a large number of calculations. The convolution operations are defined by,

$$\mathbf{Y^i} = \mathbf{f}\left(\mathbf{b^i} + \sum\nolimits_{\mathbf{j}} \mathbf{k^{ji}} * \mathbf{x^i}\right) \tag{3}$$

where $x^i$ is the input Raman scale map, $Y^i$ is the output characteristic diagram, * is the convolution symbol, $k^{ji}$ is the convolution kernel between the characteristic graphs i and j, and $b^i$ is the $i^{th}$ weight bias. The function f(·) represents the activation function.

In this paper, 64 filters of size 7 × 7 are used for analysis of the Raman spectrogram. The ReLU function was chosen as the activation function of each convolution layer [39] for its handling of overfitting in the DCNN model. The function reduces the interdependence of parameters and produces a sparse neural network model, which in turn reduces the overfitting problem [40]. The ReLU function introduces the nonlinear relationship to the input of the neural node x, which is defined by,

$$f(x) = \begin{pmatrix} x & \text{if } x \geq 0 \\ 0 & \text{if } x < 0. \end{pmatrix} \tag{4}$$

The pooling layer down-samples subsequently use the pooling filter to obtain the maximum value from the input. Each neuron pool in the output $Y_i^j$ map is on the nonoverlapping region of the input $x_i$. Maximum pooling is defined by,

$$Y_i^j = \max\left\{x_{i.m}^j\right\}. \tag{5}$$

With the increase in the depth of NN shown in Fig. 7, the parameter change of the former layer in the process of training affects the change of the latter layer (because the output of the former layer is the input of the latter). This effect increases as the network depth increases, which causes the problem of difficult network training and fitting. To solve this problem, a batch normalization (NB) layer is added to improve the distribution of the output features in the hidden

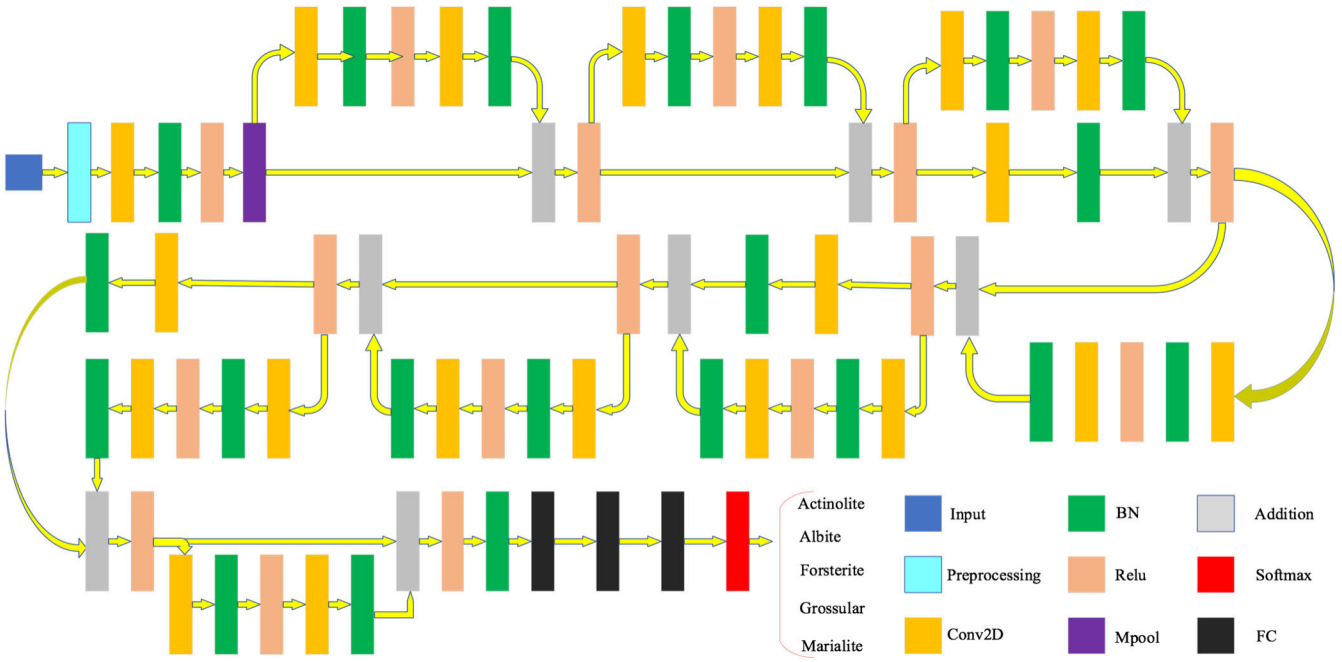


**FIGURE 7.** A simple two-dimensional deep convolution neural network model.

layer and the convergence speed during the training state [41]. This process is defined by,

$$\hat{x}^{(i)} = \frac{x^{(i)} - E\left[x^{(i)}\right]}{\sqrt{Var\left[x^{(i)}\right]}} \tag{6}$$

From (6), $\hat{x}^{(i)}$ represents the output of the BN layer, and i is the dimension of the feature scale map and is equal to 2, $E\left[x^{(i)}\right]$ is the mean, and $Var\left[x^{(i)}\right]$ is the variance of the input. The additional layer shown in Fig. 7 obtains the output features from the ReLU layer and the feature map of the BN layer. This layer combines and passes to the following NN.

Shown in Fig. 7, a cascaded network of three full connection layers (FC) obtains the output from the BN layer. Each FC layer is defined by,

$$Y_i = \sum\nolimits_i w_{ii'} x_i + b_{i'} \tag{7}$$

where $x_i$ represents the input of the FC layer, $w_{ii'}$ represents the weight matrix, $b_{i'}$ represents the bias, and $Y_i$ represents the output of each FC layer. In short, the feature scale-map datasets are placed on one dimension, and the nonlinear problem is solved by the multilayer FC connection [26].

The output of the multilayer is applied to the Softmax function to produce the classification output. This function provides a value range between 0.0-1.0.

## III. DATA PREPARATION AND EXPERIMENT

In this work, we focus on multiple methods. Traditional ML classifiers and DCNN classifiers have notable similarities in data preprocessing and feature extraction. The ML algorithm has achieved superior accuracy in the field of Raman spectrum classification [18]–[21]. To choose the best classifier to conserve time in future applications, the typical traditional ML classifiers of NB, RF, k-NN, and SVM were selected for comparison with the DCNN classifiers [1]. The performances of these ML algorithms were compared with that of the DCNN network proposed in Fig. 7. The concept of cross-entropy was introduced to evaluate the loss function of the proposed DCNN network. Let L represent the loss mean of output, which is defined by

$$L = -\frac{1}{N}\sum\nolimits_{i=1}^{N} y^{(i)} \log \hat{y}^{(i)} + (1 - y^{(i)}) \log(1 - \hat{y}^{(i)}), \tag{8}$$

where $\hat{y}^{(i)}$ represents the feature of the Raman spectrum scale map of the neuron output, $y^{(i)}$ is the corresponding target output, N is the total number of training data, and the summation operator is performed on all training inputs. The DNN framework uses the Adam optimizer, and the initial learning rate is 0.001. Each epoch trains 15 data, and the training ends after 200 epochs.

## IV. RESULTS AND DISCUSSION

### A. RESULTS

The first step tests and compares the performances among the chosen ML classifiers. In the second step, the experiment focuses on the performance analysis of the 2-D DCNN classifier.

#### 1) TRADITIONAL ML CLASSIFIERS

Fig. 8 evaluates the performance of four ML classifiers using three sets of datasets, namely, GN, BB and GB, according to

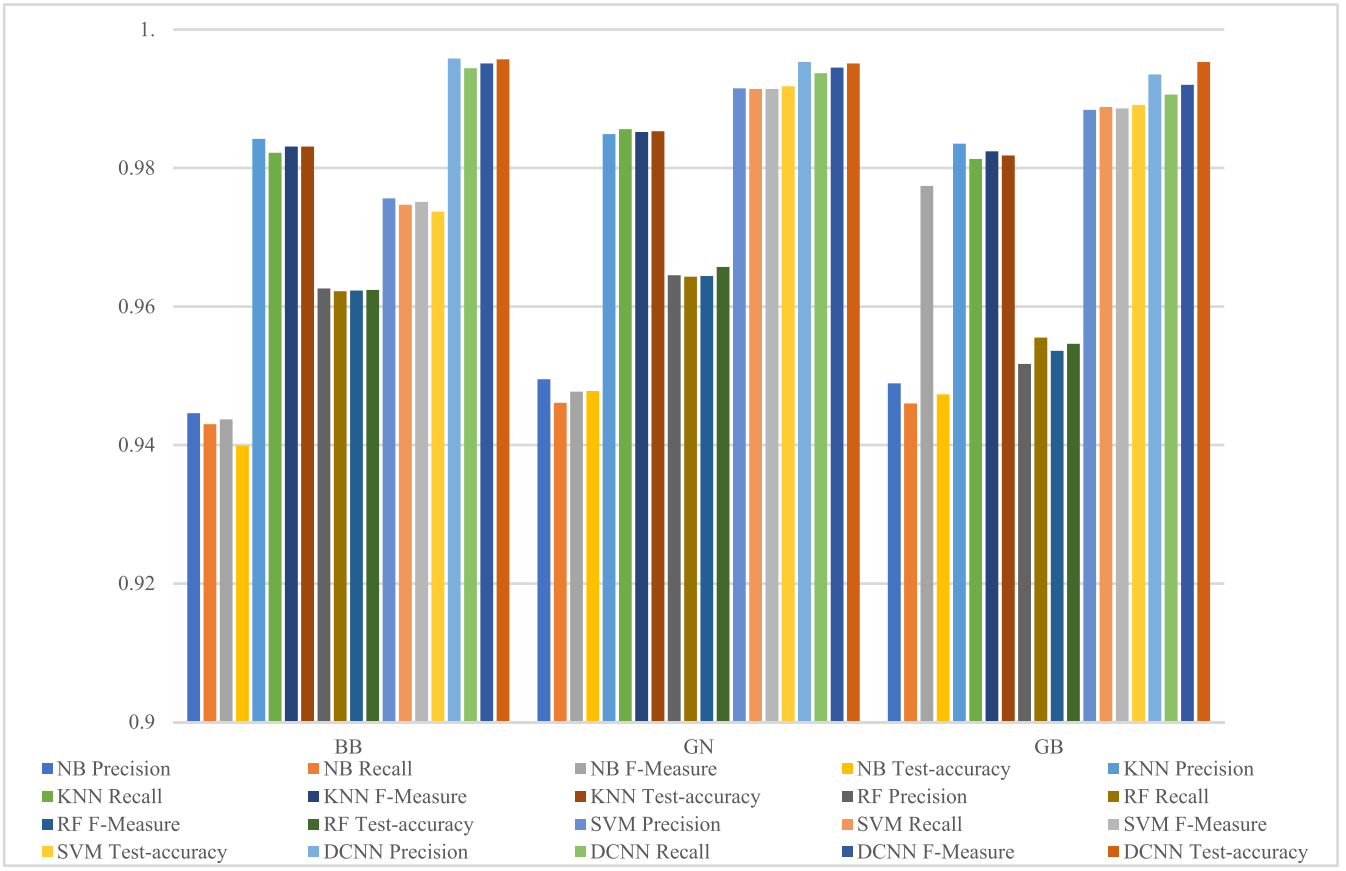


**FIGURE 8.** **Proposed framework for analyses of the classification results of different datasets.**

the precision, recall, F-measure, and test accuracy [42], [43]. The accuracy of the SVM classifier in all scenarios (BB, GN and GB) was 4.5%, 2.5% and 0.5% higher than other ML classifiers, respectively. Secondly, in the BB scenarios, the accuracy of these classifiers is better than 95%. This result shows that these ML classifiers are robust to BB noise. Finally, these ML classifiers are highly useful for GN and GB scenarios.

#### 2) PROPOSED CLASSIFIER

The DCNN network extracts the features of the scale graph of the Raman spectrum. Due to the backpropagation error in the whole training process, the weight and deviation of the DCNN model tend to a stable range, thus improving the training accuracy. To evaluate the model, we assessed it using the following measures:

- Training and validation accuracies.
- Precision, recall, F-measure, and classification statistics of the category forecasts [42], [43].
- Testing accuracy
- Confusion matrix as a holistic measure of a classifier.

Fig. 8 details the parameters of the three evaluation indicators obtained from 2-D DCNN. The precision, recall and F-measure values reached 99.4%, 99.23% and 99.3%, respectively. The training and verification accuracy is better than 99.6%, and the testing accuracy is better than 99.5%. The confusion matrix clearly shows the classifier performances and pinpoints the error occurring from each dataset group. All types of predicted tags are similar to real tags, with an accuracy of 99.2%, as shown in Fig. 9.

### *B. DISCUSSION*

According to Fig. 8, the experiment counted the use of different classifiers to test the Raman spectrum scale map under different noise scenarios. The experiment compares the ML classifiers and the DCNN classifier according to the evaluation indicators (precision, recall, F-measure, and test-accuracy). The results shown in Fig. 8 reveal that the best ML classifiers under the BB, GN and GB environments are K-NN, SVM and SVM, respectively. However, the DCNN classifier outperforms the ML classifiers for all noise conditions. In the BB noise scenario, the precision, recall, F-measure, and test-accuracy of the DCNN classifier were 1.1%, 1.2%, 2%, and 1.2% higher, respectively, than those of K-NN. In the GN noise scenario, the precision, recall, F-measure, and test-accuracy of the

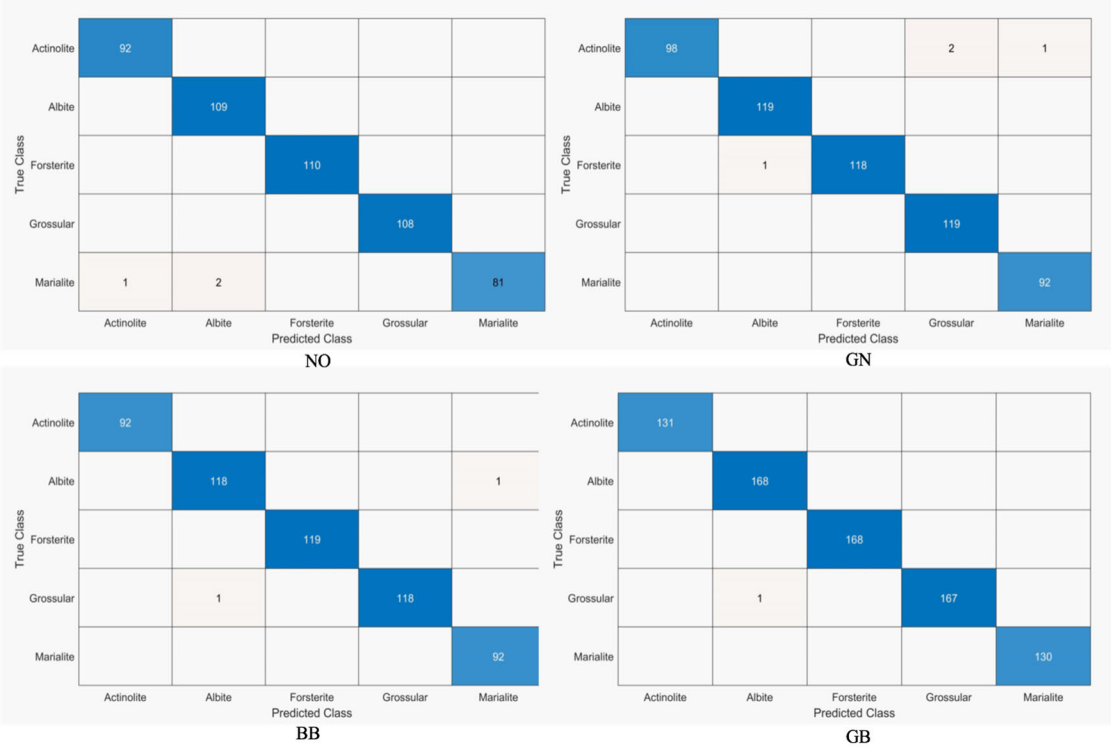


**FIGURE 9.** Confusion matrix obtained by testing different datasets. The horizontal axis represents the predicted label, and the vertical axis represents the actual label.

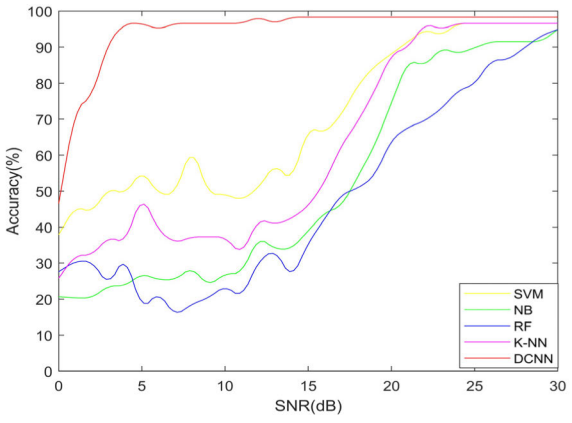


**FIGURE 10.** Classification accuracy of baseline free datasets of mixed Gaussian noise by various methods.

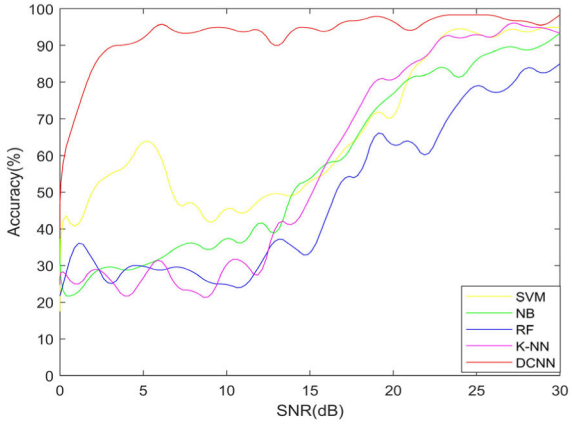


**FIGURE 11.** Classification accuracy of the baseline dataset of mixed Gaussian noise by various methods.

DCNN were 0.4%, 0.2%, 0.3%, and 0.4% higher than those of SVM. Finally, the precision, recall, F-measure, and test-accuracy of the DCNN classifier were 0.5%, 0.2%, 0.4%, and 0.6% higher than those of SVM. To verify the accuracy of the classifier again, we tested the Raman spectrum signal (GB) with 10-30 dB SNR from the simulated dataset and tested the two classifiers. The accuracy results in Table 2 show that the proposed method is 9%, 3.5%, 1%, and 0.5% better than the NB, RF, K-NN, and SVM classifiers, respectively.

**TABLE 2.** ML classifier and DCNN classifier test for the Raman spectrum signal (GB) under 10-30 dB SNR.

| | NB | RF | K-NN | SVM | DNN |
|---|---|---|---|---|---|
| Accuracy | 0.8638 | 0.9217 | 0.9419 | 0.9516 | 0.9559 |

To verify the robustness of the classifier, we tested the classifiers with a noisy processed Raman scale map in the GN

condition of 0-30 dB SNR noise. The accuracy performances of the 4 ML and 2-D DCCN classifiers are shown in Fig. 10. The accuracy rate of the ML classifiers is shallow in the high noise environment, and it remains below 65% in the 0-14 dB noise environment. In contrast, the DCNN classifier starts to obtain more than 90% accuracy in a 4 dB noise environment. This result clearly shows that the accuracy of the DCNN classifier in the high noise environment is much better than that of the ML classifiers. With the increase of SNR, the accuracy of the ML classifiers starts to improve. However, the accuracy of a DCNN classifier already exceeds 97% at 14 dB SNR.

Fig. 11 shows the accuracy performance of both types of classifiers for the GB scenario. Among the ML classifiers, the SVM classifier model is better than other ML classifiers in this scenario. However, the DCNN classifier outperforms all ML classifiers. Even in a low SNR region, the accuracy of the DCNN classifier is better than 90%. It can be concluded that the proposed DCNN classifier is more robust than the previous ML classifiers in all noise conditions.

## V. CONCLUSION

We proposed a new framework based on the wavelet transform and deep neural network for Raman spectroscopy. The framework consists of two main engines, a wavelet transform front end and a classifier back end. The wavelet engine extracts the characteristic information of the Raman shift and intensity domains in the multiresolution scale map dataset. This dataset is fed into the classifier back end. Under three noise scenarios of baseline background noise, Gaussian noise, and background baseline and Gaussian noises, the DCNN classifier was found to be the optimum classifier for the proposed framework among the four machine learning classifiers of NB, RF, K-NN, and SVM. The key statistical measures of precision, recall, test-accuracy, and F-measure obtained from the deep learning network were excellent and worked better than those obtained from the machine learning algorithms. In terms of noise robustness, the framework with the DCNN classifier shows superior performance compared with the others. An accuracy of 90% at 3 dB SNR was achieved with the deep learning classifier, while the NB, SVM, RF, and K-NN require SNR of 27 dB, 22 dB, 27 dB, and 23 dB SNR, respectively. We believe that the proposed classification framework has great potential in practical Raman spectroscopy applications and also in other classification applications such as EEG, ECG and especially mixed noise signals. Our future work will implement the framework in a low-cost portable Raman spectroscopy device.

## REFERENCES


[1] J. Liu, M. Osadchy, L. Ashton, M. Foster, C. J. Solomon, and S. J. Gibson, "Deep convolutional neural networks for Raman spectrum recognition: A unified solution," *Analyst*, vol. 142, no. 21, pp. 4067–4074, 2017, doi: 10.1039/C7AN01371J.

[2] F. Tuinstra and J. L. Koenig, "Raman spectrum of graphite," *J. Chem. Phys.*, vol. 53, no. 3, pp. 1126–1130, Aug. 1970, doi: 10.1063/1.1674108.

[3] R. Withnall, B. Z. Chowdhry, J. Silver, H. G. M. Edwards, and L. F. C. de Oliveira, "Raman spectra of carotenoids in natural products," *Spectrochimica Acta A, Mol. Biomolecular Spectrosc.*, vol. 59, no. 10, pp. 2207–2212, Aug. 2003, doi: 10.1016/S1386-1425(03)00064-7.

[4] K. Castro, M. Pérez-Alonso, M. D. Rodríguez-Laso, L. A. Fernández, and J. M. Madariaga, "On-line FT-Raman and dispersive Raman spectra database of artists' materials (e-VISART database)," *Anal. Bioanal. Chem.*, vol. 382, no. 2, pp. 248–258, May 2005, doi: 10.1007/s00216-005-3072-0.

[5] J. De Gelder, K. De Gussem, P. Vandenabeele, and L. Moens, "Reference database of Raman spectra of biological molecules," *J. Raman Spectrosc.*, vol. 38, no. 9, pp. 1133–1147, Sep. 2007, doi: 10.1002/jrs.1734.

[6] J. Hwang, N. Choi, A. Park, J.-Q. Park, J. H. Chung, S. Baek, S. G. Cho, S.-J. Baek, and J. Choo, "Fast and sensitive recognition of various explosive compounds using Raman spectroscopy and principal component analysis," *J. Mol. Struct.*, vol. 1039, pp. 130–136, May 2013, doi: 10.1016/j.molstruc.2013.01.079.

[7] Z.-M. Zhang, S. Chen, and Y.-Z. Liang, "Baseline correction using adaptive iteratively reweighted penalized least squares," *Analyst*, vol. 135, no. 5, p. 1138, 2010, doi: 10.1039/b922045c.

[8] D. A. Stuart, K. B. Biggs, and R. P. Van Duyne, "Surface-enhanced Raman spectroscopy of half-mustard agent," *Analyst*, vol. 131, no. 4, p. 568, 2006, doi: 10.1039/b513326b.

[9] S. He, W. Zhang, L. Liu, Y. Huang, J. He, W. Xie, P. Wu, and C. Du, "Baseline correction for Raman spectra using an improved asymmetric least squares method," *Anal. Methods*, vol. 6, no. 12, pp. 4402–4407, 2014, doi: 10.1039/C4AY00068D.

[10] S.-J. Baek, A. Park, Y.-J. Ahn, and J. Choo, "Baseline correction using asymmetrically reweighted penalized least squares smoothing," *Analyst*, vol. 140, no. 1, pp. 250–257, 2015, doi: 10.1039/C4AN01061B.

[11] F. Ehrentreich and L. Sämmchen, "Spike removal and denoising of Raman spectra by wavelet transform methods," *Anal. Chem.*, vol. 73, no. 17, pp. 4364–4373, Sep. 2001, doi: 10.1021/ac0013756.

[12] V. J. Barclay, R. F. Bonner, and I. P. Hamilton, "Application of wavelet transforms to experimental spectra: Smoothing, denoising, and data set compression," *Anal. Chem.*, vol. 69, no. 1, pp. 78–90, Jan. 1997, doi: 10.1021/ac960638m.

[13] H. Guo, Q. He, and B. Jiang, "The application of mexican hat wavelet filtering and averaging algorithm in Raman spectra denoising," in *Proc. Congr. Image Signal Process.*, 2008, pp. 321–326, doi: 10.1109/CISP.2008.191.

[14] S. R. Safavian and D. Landgrebe, "A survey of decision tree classifier methodology," *IEEE Trans. Syst., Man, Cybern.*, vol. 21, no. 3, pp. 660–674, Jun. 1991, doi: 10.1109/21.97458.

[15] S. Sunny and D. S. Peter, "Combined feature extraction techniques and naive Bayes classifier for speech recognition," in *Proc. Conf. Comput. Sci. Inf. Technol.*, Jul. 2013, pp. 155–163, doi: 10.5121/csit.2013.3416.

[16] S. Luo, Z. Dai, T. Chen, H. Chen, and L. Jian, "A weighted SVM ensemble predictor based on AdaBoost for blast furnace ironmaking process," *Int. J. Speech Technol.*, vol. 50, no. 7, pp. 1997–2008, Jul. 2020, doi: 10.1007/s10489-020-01662-y.

[17] J. Rabcan, V. Levashenko, E. Zaitseva, M. Kvassay, and S. Subbotin, "Application of fuzzy decision tree for signal classification," *IEEE Trans. Ind. Informat.*, vol. 15, no. 10, pp. 5425–5434, Oct. 2019, doi: 10.1109/TII.2019.2904845.

[18] E. Widjaja, W. Zheng, and Z. Huang, "Classification of colonic tissues using near-infrared Raman spectroscopy and support vector machines," *Int. J. Oncol.*, Mar. 2008, pp. 1–8, doi: 10.3892/ijo.32.3.653.

[19] J. C. Martínez Romo, F. J. Luna-Rosas, R. Mendoza-González, A. Padilla-Díaz, M. Mora-González, and E. Martínez-Cano, "Improving sensitivity and specificity in breast cancer detection using Raman spectroscopy and Bayesian classification," *Spectrosc. Lett.*, vol. 48, no. 1, pp. 40–52, Jan. 2015, doi: 10.1080/00387010.2013.855640.

[20] S. Seifert, "Application of random forest based approaches to surface-enhanced Raman scattering data," *Sci. Rep.*, vol. 10, no. 1, Dec. 2020, Art. no. 5436, doi: 10.1038/s41598-020-62338-8.

[21] N. H. Othman, A. R. M. Radzol, K. Y. Lee, and W. Mansor, "Reduced featured k-NN classifier model optimal for classification of dengue fever from salivary Raman spectra," in *Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2019, pp. 471–474, doi: 10.1109/EMBC.2019.8856427.

[22] J. Long, J. Mou, L. Zhang, S. Zhang, and C. Li, "Attitude data-based deep hybrid learning architecture for intelligent fault diagnosis of multi-joint industrial robots," *J. Manuf. Syst.*, Sep. 2020, pp. 1–8, doi: 10.1016/j.jmsy.2020.08.010.

[23] F. Xu, Z. Pan, and R. Xia, "E-commerce product review sentiment classification based on a naïve bayes continuous learning framework," *Inf. Process. Manage.*, vol. 57, no. 5, Sep. 2020, Art. no. 102221, doi: 10.1016/j.ipm.2020.102221.

[24] M. Otoom, N. Otoum, M. A. Alzubaidi, Y. Etoom, and R. Banihani, "An IoT-based framework for early identification and monitoring of COVID-19 cases," *Biomed. Signal Process. Control*, vol. 62, Sep. 2020, Art. no. 102149, doi: 10.1016/j.bspc.2020.102149.

[25] Z. Elamrani Abou Elassad, H. Mousannif, and H. Al Moatassime, "A real-time crash prediction fusion framework: An imbalance-aware strategy for collision avoidance systems," *Transp. Res. C, Emerg. Technol.*, vol. 118, Sep. 2020, Art. no. 102708, doi: 10.1016/j.trc.2020.102708.

[26] D. Erhan, Y. Bengio, A. Courville, P.-A. Manzagol, P. Vincent, and S. Bengio, "Why does unsupervised pre-training help deep learning?" in *Proc. Int. Conf. Artif. Intell. Statist.*, 2010, pp. 201–208.

[27] J. Long, Z. Sun, C. Li, Y. Hong, Y. Bai, and S. Zhang, "A novel sparse echo autoencoder network for data-driven fault diagnosis of delta 3-D printers," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 3, pp. 683–692, Mar. 2020, doi: 10.1109/TIM.2019.2905752.

[28] S. Weng, X. Xu, J. Li, and S. T. C. Wong, "Combining deep learning and coherent anti-Stokes Raman scattering imaging for automated differential diagnosis of lung cancer," *J. Biomed. Opt.*, vol. 22, no. 10, p. 1, Oct. 2017, doi: 10.1117/1.JBO.22.10.106017.

[29] N. Kussul, M. Lavreniuk, S. Skakun, and A. Shelestov, "Deep learning classification of land cover and crop types using remote sensing data," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 5, pp. 778–782, May 2017, doi: 10.1109/LGRS.2017.2681128.

[30] M. Fukuhara, K. Fujiwara, Y. Maruyama, and H. Itoh, "Feature visualization of Raman spectrum analysis with deep convolutional neural network," *Anal. Chim. Acta*, vol. 1087, pp. 11–19, Dec. 2019, doi: 10.1016/j.aca.2019.08.064.

[31] M. R. Almeida, K. D. S. Oliveira, R. Stephani, and L. F. C. de Oliveira, "Fourier-transform Raman analysis of milk powder: A potential method for rapid quality screening," *J. Raman Spectrosc.*, vol. 42, no. 7, pp. 1548–1552, Jul. 2011, doi: 10.1002/jrs.2893.

[32] J. Binoy, J. P. Abraham, I. H. Joe, V. George, V. S. Jayakumar, J. Aubard, and O. F. Nielsen, "Near-infrared Fourier transform Raman, surface-enhanced Raman scattering and Fourier transform infrared spectra Andab Initio calculations of the natural product nodakenetin angelate," *J. Raman Spectrosc.*, vol. 36, no. 1, pp. 63–72, Jan. 2005, doi: 10.1002/jrs.1272.

[33] S. U. Amin, M. Alsulaiman, G. Muhammad, M. A. Mekhtiche, and M. S. Hossain, "Deep learning for EEG motor imagery classification based on multi-layer CNNs feature fusion," *Future Gener. Comput. Syst.*, vol. 101, pp. 542–554, Dec. 2019, doi: 10.1016/j.future.2019.06.027.

[34] N. Feng, S. Xu, Y. Liang, and K. Liu, "A probabilistic process neural network and its application in ECG classification," *IEEE Access*, vol. 7, pp. 50431–50439, 2019, doi: 10.1109/ACCESS.2019.2910880.

[35] J. Huang, B. Chen, B. Yao, and W. He, "ECG arrhythmia classification using STFT-based spectrogram and convolutional neural network," *IEEE Access*, vol. 7, pp. 92871–92880, 2019, doi: 10.1109/ACCESS.2019.2928017.

[36] A. Parmar, S. Gulia, S. Bajaj, V. Gambhir, R. Sharma, and M. N. Reddy, "Signal processing of Raman signatures and realtime identification of hazardous molecules using continuous wavelet transformation (CWT)," in *Proc. Int. Conf. Signal Process. Commun. Eng. Syst.*, Jan. 2015, pp. 323–325, doi: 10.1109/SPACES.2015.7058275.

[37] Ü. Lepik, "Numerical solution of differential equations using Haar wavelets," *Math. Comput. Simul.*, vol. 68, no. 2, pp. 127–143, Apr. 2005, doi: 10.1016/j.matcom.2004.10.005.

[38] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.

[39] J. Dong, M. Hong, Y. Xu, and X. Zheng, "A practical convolutional neural network model for discriminating Raman spectra of human and animal blood," *J. Chemometrics*, vol. 33, no. 11, Nov. 2019, doi: 10.1002/cem.3184.

[40] A. Fred Agarap, "Deep learning using rectified linear units (ReLU)," 2018, *arXiv:1803.08375*. [Online]. Available: http://arxiv.org/abs/1803.08375

[41] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," 2015, *arXiv:1502.03167*. [Online]. Available: http://arxiv.org/abs/1502.03167

[42] S. Ghosh, A. Rana, and V. Kansal, "A nonlinear manifold detection based model for software defect prediction," *Procedia Comput. Sci.*, vol. 132, pp. 581–594, 2018, doi: 10.1016/j.procs.2018.05.012.

[43] S. Ghosh, A. Rana, and V. Kansal, "A benchmarking framework using nonlinear manifold detection techniques for software defect prediction," *Int. J. Comput. Sci. Eng.*, vol. 21, no. 4, pp. 593–614, 2020, doi: 10.1504/IJCSE.2020.106871.

**LIANGRUI PAN** (Student Member, IEEE) was born in Anhui, China, in 1997. He received the bachelor's degree from Anhui Polytechnic University, in 2019. He is currently pursuing the master's degree in electrical engineering with the Prince of Songkla University, Thailand. His research interests include machine learning, deep learning, and pattern recognition. He is also a member of the Chinese Society of Electrical Engineering.

**PRONTHEP PIPITSUNTHONSAN** (Student Member, IEEE) received the bachelor's degree from the Prince of Songkla University, in 2010, and the master's degree, in 2017. He is currently pursuing the Ph.D. degree in computer engineering. Since 2015, he has been working as a Programmer with GISTDA. His research interests include deep learning and big data.

**CHALONGRAT DAENGNGAM** received the B.S. degree in physics from the Prince of Songkla University, Songkhla, Thailand, in 2005, the M.Sc. degree in nanoelectronics and nanomechanics from the University of Leeds, U.K., in 2006, and the Ph.D. degree in physics from Virginia Tech, USA, in 2012. He is currently working as an Assistant Professor with the Department of Physics, Faculty of Science, Prince of Songkla University. His research interests include nonlinear optical properties of nanomaterials, photonics, and standoff Raman spectroscopy.

**SITTIPORN CHANNUMSIN** received the bachelor's degree in electronics engineering from the King Mongkut's Institute of Technology Ladkrabang (KMITL), Thailand, in 2006, the M.Sc. degree in space technology and planetary exploration from the University of Surrey, in 2011, and the Ph.D. degree in aerospace engineering from the University of Glasgow, U.K., in 2016. He is currently a Supervisor of the Astrodynamics Research Laboratory (AstroLab), Geo-Informatics and Space Technology Development Agency (GISTDA). He also serves as a Co-Editor and a Technical Reviewer. His research and publication interests include spaceflight dynamics, space debris mitigation, machine learning, and orbit and attitude control and optimization.

**SUWAT SREESAWET** received the bachelor's degree in electronical mechanical and manufacturing engineering from Kasetsart University, Thailand, and the master's and Ph.D. degrees in aerospace engineering from Wichita State University, Wichita, KS, USA, in 2014 and 2018, respectively. He is currently working as a Researcher with the Astrodynamics Research Laboratory (AstroLab) in Geo-Informatics and Space Technology Development Agency (GISTDA). His research and publication interests include spaceflight dynamics, trajectory optimization, and altitude control.

**MITCHAI CHONGCHEAWCHAMNAN** (Senior Member, IEEE) was born in Bangkok, Thailand. He received the B.Eng. degree in telecommunication from the King Mongkut's Institute of Technology Ladkrabang, Bangkok, in 1992, the M.Sc. degree in communication and signal processing from Imperial College London, London, U.K., in 1995, and the Ph.D. degree in electrical engineering from the University of Surrey, Guildford, U.K., in 2001. He joined the Mahanakorn University of Technology, Bangkok, as a Lecturer, in 1992. In 2008, he joined the Faculty of Engineering, Prince of Songkla University, Songkhla, Thailand, as an Associate Professor. His current research interests include deep learning, microwave circuit design and microwave techniques for agricultural applications.