Materials Today Communications 27 (2021) 102198

Contents lists available at ScienceDirect

Materials Today Communications

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

SEM-Net: Deep features selections with Binary Particle Swarm Optimization Method for classification of scanning electron microscope images

Gürkan Kavuran

Department of Electrical and Electronics Engineering, Faculty of Engineering and Natural Sciences, Malatya Turgut Özal University, Malatya, Turkey

ARTICLE INFO

Keywords:
Hydrothermal method
Nanoscience
Ag-doped SnO_2
Feature selection
Binary Particle Swarm Optimization
Deep learning

ABSTRACT

Materials Science is increasingly handling artificial intelligence methods to address the complexity in the field of everyday life necessities. Researchers in both academia and industry are interested in imaging techniques used in the characterization of nanomaterial with designed properties to meet the needs of applications in the literature. However, the increase in image size and complexity in its content restricts the use of traditional methods. Recent advances in machine learning have been used to benefit computers' potential to make sense of these images. The approach proposed in this paper aims for the feature reduction with the Binary Particle Swarm Optimization method to execute the classification process on SEM images by concatenating the deeper layers of pre-trained CNN models AlexNet and ResNet-50. The feature vectors were used as input to support vector machine classifier (SVMC) after dimension reduction to obtain the final model. Finally, the trained model's performance was tested using SEM images of Ag-doped SnO_2 nanoparticles, which were prepared by the author using the low-temperature hydrothermal method. To the best of the author knowledge, these images were not available in the databases. The best accuracy value was observed with 3112 features for the SEM dataset with optimized vectors as 99.3 %. An example was illustrated where the feature selection with the BPSO technique could provide novel insight into nanoscience research and test the model with the SEM images of Ag-doped SnO_2 particles that are obtained by the hydrothermal method.

1. Introduction

Nowadays, image classification, image recognition, object tracking, and other algorithms based on machine learning techniques are widely used in many different areas such as; medical imaging, defense technologies, medicine, nanotechnology, electric vehicles, smart cities, etc. The uptrend in materials science with technology development has revealed the need to improve imaging techniques for nano-sized particles. The extraction of feature vectors from different types of microscopy images has increased the variety of deep learning algorithms using artificial neural networks due to the necessity of software development and hardware.

For example, Ge et al. proposed a procedure for applying deep learning to microscopic imaging [1]. In another study, transfer learning techniques were used to process images obtained by scanning electron microscopy (SEM) [2]. In the study by Rezaie et al., the performance of threshold and deep learning methods for detecting crack pixels on images used as inputs for the DIC method were compared [3]. The NN-based image segmentation method was developed for spatter extraction by Tan et al. [4]. Stan et al. suggested that neural networks can be trained to perform highly precise segmentation in large datasets produced by X-ray computed tomography and segmentation [5].

Image processing techniques with deep learning can be a very effective tool in nanoscience, where characterization applications such as image-based optical microscopy, transmission electron microscopy (TEM), scanning electron microscopy (SEM), and scanning probe microscopy (SPM) are widely used. SEM is a tool used in nanoscience and nanotechnology, providing information about surface morphology and composition with a resolution of up to 1 nm to explore the structure of materials [6].

SEM images are obtained in research and development centers and industrial fields by various technical staff for different usage purposes. However, these large-scale data, unfortunately, cannot be classified and stored using appropriate tools. Considering the usage and user diversity of these images, the lack of well-defined labeling systems causes complexity. Also, long-term management of such necessary data is difficult. It is essential to develop applications based on machine learning techniques to overcome all these challenges. The network

E-mail address: gurkan.kavuran@ozal.edu.tr.

https://doi.org/10.1016/j.mtcomm.2021.102198
Received 22 December 2020; Received in revised form 24 February 2021; Accepted 24 February 2021
Available online 28 February 2021

models structured by machine learning techniques will have the following benefits to the scientific areas where SEM images are used;

1 Automatic image classification procedure that can find out which class the new image belongs to, independently of the user;
2 A trained model that allows researchers to determine a specific group of SEM images;

The approach proposed in this paper aims for the feature reduction with the Binary Particle Swarm Optimization method to execute the classification process on SEM images by concatenating the deeper layers of pre-trained CNN models AlexNet (fc6) and ResNet-50 (avg_pool). These models reached successful results in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC), evaluating algorithms for object detection and image classification. The feature vectors were used as input to support vector machine classifier (SVMC) after dimension reduction to obtain the final model. Finally, the performance of our approach was tested using SEM images of Ag-doped SnO_2 nanoparticles.

The contribution of the proposed method to the literature can be explained as follows;

1 The feature selection method based on the metaheuristic optimization algorithm was used to achieve lower computation costs and higher accuracy performance by reducing feature sizes.
2 The trained model's performance was tested using SEM images of Ag-doped SnO_2 nanoparticles, which were not available in the database and were prepared by the author using the low-temperature hydrothermal method. Consequently, it was proved that the model is intended for the end-user by applying the database-independent test procedure and can be embedded in hardware with size reduction.

The remaining part of the paper is organized as follows: Section 2 presents materials and methods, including dataset, description of the proposed approach, and hydrothermal synthesis of nanocomposites. Section 3 is devoted to experiments. At last, the summary of our policy has been given in Section 4.

2. Materials and methods

The manuscript also presents the experimental results succeeded through the feature selection method based on the metaheuristic optimization algorithm we propose. The sub-sections contain the following stages accomplished in this work:

1 In this study, a first public dataset of SEM images that can be used as an instance of imminent deep learning applications in the field of nanoscience was used.
2 The application of the deep learning-based image classification procedure to the SEM dataset can be explained as follows:
   a) The feature extraction process was performed from AlexNet and ResNet-50 by retraining the fully-connected (fc6) layer and the average pooling layer (avg_pool) for the SEM dataset. The extracted feature vectors were used in feature concatenation to form a deep feature stack.
   b) A feature selection method with the Binary Particle Swarm Optimization (BPSO) algorithm was employed to select the most characteristic features from the concatenated in-depth features.
3 SVMC was employed to label the test SEM images with their putative features and compare the classification performance.
4 The final successful model's performance obtained via the proposed method was tested by using the SEM images of pure SnO_2 and Ag-doped SnO_2 nanopowders synthesized by the hydrothermal method in the laboratory.

Table 1
SEM dataset class distributions.

Category	N images
Porous Sponge	181
Powder	916
Nanowires	3820
Fibres	162
Total	5079

2.1. Description of the SEM dataset

This section presents the description of the SEM dataset classified by the author. The nanoscale material structures' diversity also caused the SEM images to be classified into different classes in a wide range. Since the classification applications using image processing techniques are based on supervised learning, the images from the SEM dataset were labeled under different classes by 100 scientists from the CNR-IOM TASC laboratories in Trieste for five years [2,7]. The dataset encompasses a set of ten categories such as porous sponge, particles, nanowires, fibres, films coated surfaces, powder, patterned surfaces, microelectromechanical system (MEMS) devices, pillars, tips, and biological for a total of 18.577 SEM images. In this work, the list of the number of images selected in each category from total categories representing many different nanoscience fields is shown in Table 1. Representative images for each of the categories chosen from the SEM dataset are shown in Fig. 1.

2.2. Description of the proposed approach: deep SEM-Net

Convolutional Neural Networks (CNNs) are the most advanced form of feedforward neural networks, which are frequently used in image processing applications. CNNs are simply neural networks that use convolution in place of general matrix multiplication in at least one of their layers. Multi-layered Convolutional Neural Networks, inspired by the visual centers of mammals, resemble ordinary artificial neural networks. They are also composed of neurons that have learnable preferences and biases. In CNN, data is divided into pieces, and specific filters are applied to each piece. The size of the image decreases according to the applied filter. In traditional neural network layers, each input and each output unit's interaction is multiplied by a weight matrix with a separate parameter that describes it. It means that every output unit interacts with every input unit. However, CNN has less frequent interactions than a conventional artificial neural network (ANN). Convolutional networks have been enormously influential in practical applications due to reductions in data size.

At the beginning of the classification process, the number of the data was expanded by applying image augmentation methods as online throughout the learning period. The imbalance problem in the number of images for the classes may not be suitable for building a robust image classifier. Besides, image augmentation methods have been used to eliminate the overfitting problems in the literature. Therefore, deep networks need large amounts of training data to achieve good performance. In this work, the image data augmentation library of MATLAB was used to perform augmentation during training progress. The augmented image data store applies a combination of multiple transformations, such as; random rotation (in the range [−15 15] degrees), vertically and horizontally reflection, and shear (in the range of [−0.5 0.5] horizontal and vertical) to the training data. Some examples of the augmented images are given in Fig. 2a. The classification process proposed in this study consists of three stages: feature extraction using pre-trained deep network structures, reducing the size of the obtained feature vectors by the BPSO method, and finally determining the labels of the SEM images by utilizing the SVMC. For this purpose, fc6 and avg_pool layers were preferred for deep feature extraction by considering AlexNet and ResNet-50 models. The extracted feature vectors were

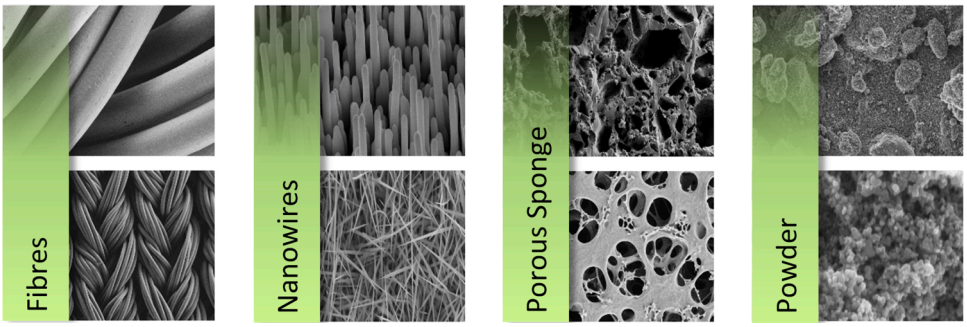

Fig. 1. Categories are chosen for the SEM images dataset.

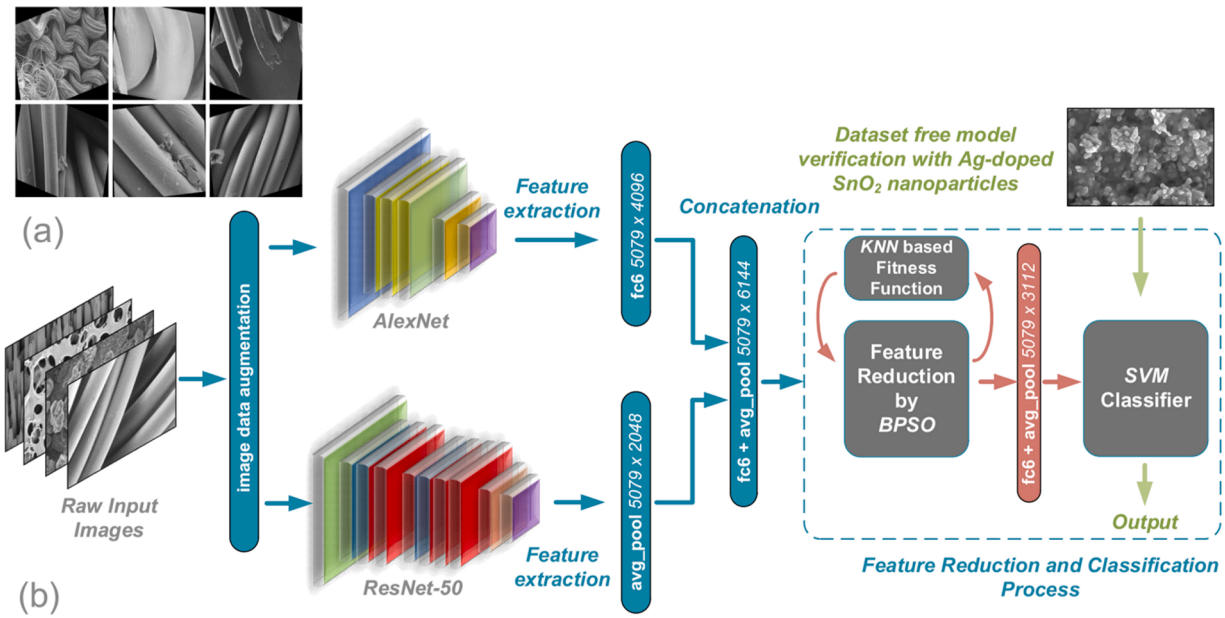

Fig. 2. a) Representative images after augmentation. b) Illustration of the proposed approach.

used in feature concatenation to form a deep feature stack. A feature selection method with the BPSO algorithm was employed to select the most characteristic features from the concatenated deep features. A support vector machine classifier was employed to label the test SEM images with the selected features. The proposed approach is demonstrated in Fig. 2b.

2.2.1. Feature extraction from pre-trained CNN models

In computer vision and image processing, a feature is typically the building block that provides information about whether a particular region of an image has certain properties. Features may be specific structures such as pixels, edges, motifs, pieces, or objects in the image. As mentioned in the previous sections, the input of the SVMC was fed to classify SEM images using different combinations of features obtained from the fully connected and average pooling layers of AlexNet and ResNet-50. AlexNet, which was developed by Alex Krizhevsky, Ilya Sutskever, and Geoff Hinton, was the first study in the field of computer vision and is still popular as a convolutional network [8]. Although the network structure bears similarities with LeNet, it contains convolution layers that are deeper, larger, and stacked on top of each other. The CNN model designed by Alex et al. obtained impressive results in the ImageNet challenge. It has an architecture consisting of 25 layers, eight blocks at the same level. Convolution filters are 11×11, 5×5, and 3×3 in size, and the maximum pooling layer (filter size 3×3 and 2-step shift), ReLU layer, three fully-connected layers, two of which are 4096 and the other 1000 size, and a dropout layer are used. Besides, a normalization process was performed at the output of the convolution layers to speed up the model's performance. The best error result in the competition was 16.4 % with the AlexNet model. After the fully-connected layer, the classification process was completed with the softmax layer. We had a ReLU layer and maximum pooling layer after each convolution operation. Besides, a dropout layer was used between fc6-fc7 and fc7-fc8 layers. The ResNet model proposed by He et al. got the best rank in the ILSVRC competition with an error result of 3.57 %. Three different models were created in the ResNet structure, such as ResNet-18, ResNet-50, and ResNet-101 [9].

Generally, fully-connected layers followed by a softmax activation layer that gives the probability distribution over each class are used for feature extraction in image classification problems. It has been observed that in convolutional neural networks, more intricate details are captured in deep layers than in lower layers. The fc6 layer was chosen for deep feature extraction in AlexNet to test the optimization algorithm's efficiency. When an image is convoluted, the feature matrix's size increases, which contains the activation values for the weights and biases. The matrices' size will be even more significant as convolution layers are added together and the network deepens. Hence the pooling layers are used to extract features from the feature matrices and control the size of weights and biases. They also prevent the overfitting of the model. Therefore, the average pooling layer was preferred in ResNet-50. Besides, the SGDM optimization technique was used to optimize parameters in the AlexNet and ResNet-50 models.

2.2.2. Feature selection with Binary Particle Swarm Optimization

Since the concatenated deep features contain many avoidable components in their structure, the necessity of feature selection methods arose. The feature selection method aims to obtain the best features that define the target class during the classification process by removing unrelated characteristics. In this way, classification quality is improved, and temporal problems are eliminated. Feature selection methods are generally examined under three subtitles: the filter method, the embedded method and the wrapper method. As compared to the other approaches, the wrapper can usually achieve the best subset of features vector and time consumption [10,11]. In this study, it is aimed to solve the feature reduction problem by using Binary Particle Swarm Optimization (BPSO), which is one of the wrapper feature selection methods.

PSO and BPSO, which are metaheuristic optimization algorithms, are some of the most preferred wrapper methods for feature selection. Generally, these algorithms are preferred in the literature to avoid the computational complexity in other wrapper methods that include temporal problems [12,13]. PSO is a metaheuristic optimization algorithm based on the search mechanism for the best solution by navigating the search space of the swarm members [14]. PSO models the movement of particles representing the candidate solution towards the best solutions in the search space with the social swarm behavior, aiming for the particles to arrive at an optimal solution [15].

At the beginning of the algorithm, the particles are randomly distributed over the search space. Each particle tends to continually move towards better solutions, guided by local and global best particles. The PSO algorithm updates the velocities of the particles according to the following Equation [16,17],

$$V_i^n(t+1) = \omega(t)V_i^n(t) + c_1 r_1\left(X_{Lbest}^n(t) - X_i^n(t)\right) + c_2 r_2\left(X_{Gbest}^n(t) - X_i^n(t)\right) \tag{1}$$

In the n-dimensional search space, the i-th particle's velocity and position can be represented by $V_i^n(t)$ and $X_i^n(t)$, respectively. Then, the particle position is updated as shown in Eq. (3) by using the probability value via Eq. (2).

$$S\left(V_i^n(t+1)\right) = \frac{1}{1+\exp\left(-V_i^n(t+1)\right)} \tag{2}$$

$$X_i^n(t+1) = \begin{cases} 1, & if\ rand < S\left(V_i^n(t+1)\right) \\ 0, & otherwise \end{cases} \tag{3}$$

Here, each particle's position and velocity in an n-dimensional search space are represented by an n-dimensional vector. In Eq. (1), the local best solution is X_{Lbest}^n, and the global best solution is X_{Gbest}^n. The local learning coefficient c_1 sets the degree of local best orientation, and c_2 is the global learning coefficient that sets the degree of global best orientation. Here, r_1 and r_2 are random numbers and give an arbitrary degree of freedom to particle motions [15]. This degree of freedom allows the algorithm to find new solutions in every iteration. The acceleration parameter $\omega(t)$ causes the particle to slow down, which is updated at each iteration step. Decreased acceleration allows particles to settle into the correct location during the progress of iterations.

In this study, the solution is represented in binary form to randomly generate a population of initial solutions from the d-dimensional feature set, which can be bit 0 or 1 [12]. The fitness function, which takes into account both classification performance and feature size, can be defined as follows [18]:

$$F = \alpha\frac{|S|}{|T|} + (1-\alpha)e \tag{4}$$

$$e = \sum_{j=1}^{n} \omega_j I\{\hat{y}_j \neq y_j\} \tag{5}$$

where α is the control parameter in the range of [0,1], T is the total number of features in each dataset, S is the feature subset's length, and e is the classification error value [19,20]. Classification error is the weighted ratio of misclassified features. Here $\hat{y}_j$ is the class label, and I {x} is the indicator function. The error function is computed by using the k-nearest neighbor that offers simplicity and ease of application with Euclidean distance [20].

2.3. Support vector machine classification

A Support Vector Machine (SVM), whose theoretical foundations were laid by Vapnik in the late 1960s, is a pattern classification method based on the statistical learning theorem used in the solution of binary classification problems [21,22]. SVMs aim to find the most appropriate separator plane that classifies the dataset as much as possible by determining the situation where the distance between the two classes is the greatest. This goal is achieved by finding the largest boundary between different samples after transferring the nonlinear sample space to a high dimension where it can be linearly separated [23].

Algorithmically, SVM is a data partitioning-based learning method. For this purpose, it uses the most suitable separator hyperplane to split the data. The kernel function in the structure of the SVM is simply a product of the input space. SVMs aim to find the most suitable separator plane defined by the linear decision function given by Eq. (6) [24].

$$f(x) = \omega^T x + b \tag{6}$$

where x with an n-dimensional space, ω is weight vector, and b is the bias. The decision function for each training sample x_i is given by Eq. (7) [25].

$$\begin{array}{l} \omega^T x_i + b \geq 0 \quad for\ all\ y_i = 1 \\ \omega^T x_i + b < 0 \quad for\ all\ \ y_i = -1 \end{array} \tag{7}$$

These constraints can be expressed together as given in Eq. (8).

$$y_i\left(\omega^T x_i + b\right) \geq 1, \quad i = 1, \ldots l \tag{8}$$

In this case, the most suitable hyperplane with the widest limit can be found by making the $J(\omega)$ function the smallest under the constraints given by Eq. (9).

$$J(\omega) = \frac{1}{2}\omega^T\omega = \frac{1}{2}\|\omega\|^2 \tag{9}$$

In the case of classifications that cannot be separated linearly, it adds a positive artificial ξ_i variable and C regulation parameter.

$$J(\omega) = \left(\frac{\|\omega\|^2}{2} + C\sum_{i=1}^{l}\xi_i\right) \tag{10}$$

The expression given by Eq. (10) is a quadratic optimization problem. The Lagrange multipliers method is used to make the J (w) function the smallest [22,23,26,27]. The Lagrange function is given by Eq. (11).

$$L(\omega, b, \alpha) = \frac{1}{2}\left(\omega^T\omega\right) - \sum_{i=1}^{l}\alpha_i\left\{\left[y_i\left(\omega^T x_i + b\right)\right] - 1\right\} \tag{11}$$

Since this equation solution is quite complex, Karush-Kuhn-Tucker (KKT) conditions should be used to find the appropriate solution. The result of the decision function that cannot be linearly separated uses a kernel function as follows:

$$f(x) = sign\left(\sum_{i=1}^{l}\alpha_i y_i(x)(x_i) + b\right) \tag{12}$$

2.4. Hydrothermal synthesis of SnO_2 and Ag-SnO_2 nanocomposites

Tin (II) chloride dihydrate ($SnCl_2.2H_2O$, 98 %), silver nitrate ($AgNO_3$, 95 %), ethanol (CH_3CH_2OH, 95 %), and ammonium hydroxide (NH_4OH, 25 %) were purchased from Sigma Aldrich. The distilled water

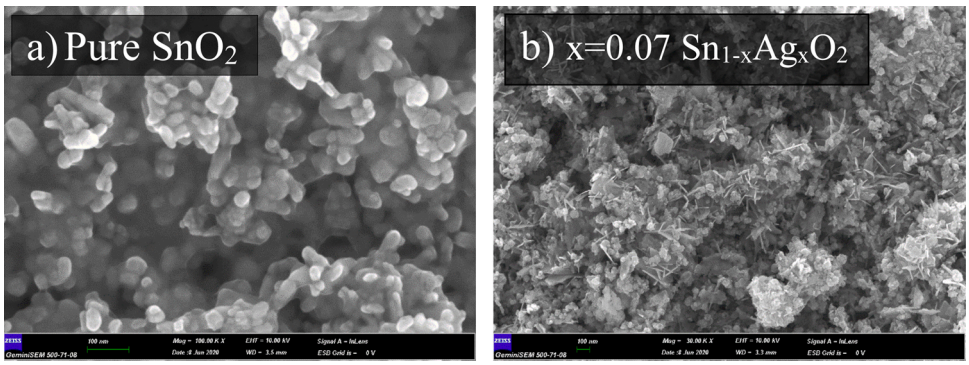

Fig. 3. SEM images of $Sn_{1-x}Ag_xO_2$; a) x = 0.0 b) x = 0.07.

was prepared by a Millipore Milli-Q UF plus lab ultra-pure water purification system and used as the solvent. All reagents used were of analytical grade without further purification.

Pure and Ag-doped SnO_2 nanoparticles were prepared by a low-temperature hydrothermal method. 0.12 M solution of $SnCl_2.2H_2O$ was dissolved in 60 mL of distilled water after stirring vigorously with a magnetic stirrer for 30 min. The solution's pH was adjusted to ~ 10 by adding NH_4OH dropwise and stirred on a hot plate with a magnetic stirrer for 15 min. The homogeneous solution was sonicated in an ultrasonic bath for 5 min and transferred to a 60 mL Teflon-linked autoclave to react 24 h under 200 °C. After the hydrothermal reactor was cooled back to room temperature, the solution was filtrated, washed several times with distilled water, and dried at 600 °C for 2 h in a muffle furnace to obtain pure SnO_2.

For the preparation of Ag-doped SnO_2 (1.0 wt%, 3.0 wt%, 5.0 wt%, 7.0 wt%), the typical synthesis process was carried out as follows: 1.6244 g of $SnCl_2.2H_2O$ was dissolved in 50 ml distilled water after being stirred for 15 min. to form solution X. Solution Y was prepared by dissolving the appropriate amount of $AgNO_3$ in 10 ml distilled water and stirred on a hot plate with a magnetic stirrer for 15 min. After that, solution Y was added to solution X and stirred for 10 min. Then the appropriate amount of NH_4OH was added to the mixture for adjusting the pH to ~ 10. The solution was transferred into a Teflon-linked autoclave and maintained at 200 °C for 24 h. The rest of the procedure was the same as the synthesis of pure SnO_2. The sample morphology was analyzed using SEM, ZEISS model LS-10 Life Science equipment. Fig. 3 shows the SEM images of pure and Ag-doped SnO_2 nanostructures. When the addition of Ag increased in the main structure, the porous structure and the sizes of the microspheres increased.

3. Experiments

In this section, the experimental setup and the evaluation of the proposed method are presented. The input images were resized to 227 × 227 and 224 × 224 to be compatible with AlexNet and ResNet-50 models, respectively. The deep CNN features were extracted from fc6 and avg_pool activations from AlexNet and ResNet-50 models. Three different feature sets were constructed by the output of the fc6 and avg_pool layers, such as AlexNet (fc6 5079 × 4096), ResNet-50 (avg_pool 5079 × 2048), and AlexNet (fc6) + ResNet-50 (avg_pool). Consequently, 5079 × 6144 dimensional concatenated feature vectors were obtained. All the feature sets were normalized according to the zero mean. The SVM classifier was used in the classification stage with 75 % training and 25 % testing data partition. The impact of the proposed method on its accuracy and evaluation metrics are discussed with the computational efficiency. To evaluate the quantitative performance of the proposed method, evaluation metrics such as Accuracy (ACC), True Positive Rate (TPR), True Negative Rate (TNR), Positive Predictive Value (PPV), F1 score (F1), and Matthew Correlation Coefficient (MCC) were statistically computed from the confusion matrix. All the experiments were performed in a MATLAB environment running on a PC with AMD Ryzen 5 2600 3.4 GHz CPU, 16 GB memory, and 4 GB NVIDIA GeForce GTX 1650 GPU. The selected evaluation metrics are defined as:

$$ACC = \frac{N_{TP} + N_{TN}}{N_{TP} + N_{TN} + N_{FP} + N_{FN}} \tag{13}$$

$$TPR = \frac{N_{TP}}{N_{TP} + N_{FN}} \tag{14}$$

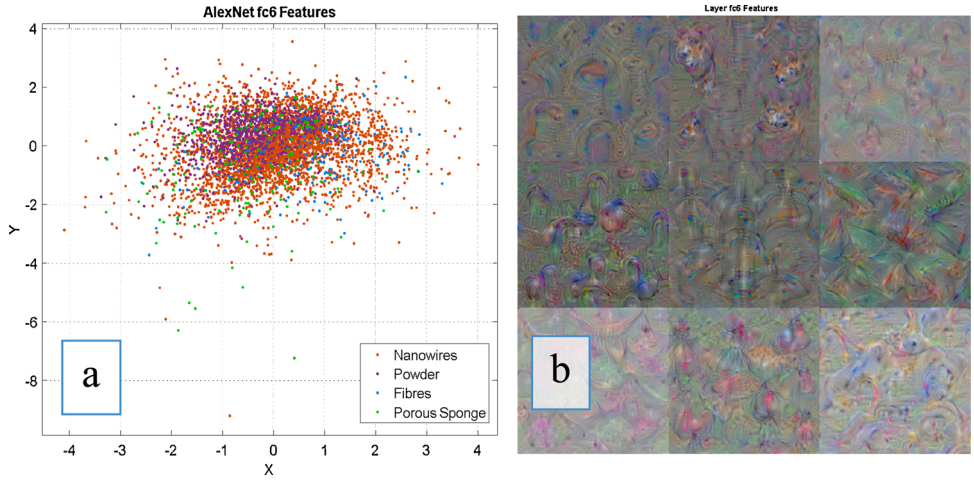

Fig. 4. a) The scatter plot of the features from AlexNet fc6 layer. b) The visualization of the first 9 features learned by the fc6 layer.

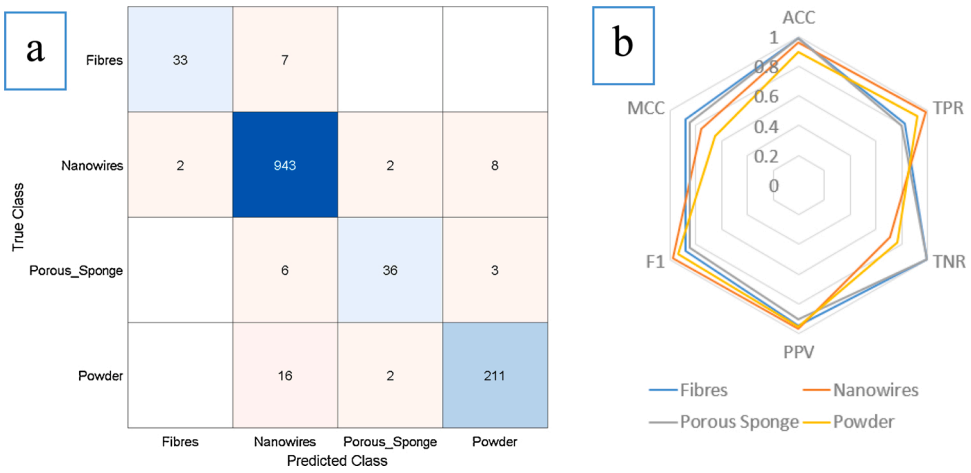

Fig. 5. a) The multi-class confusion matrix of the SVMC with fc6 features. b) Radar chart of the classification scores of the SVMC for fc6 features.

Table 2
Classification scores of the SVMC with deep features from the fc6 layer.

Class	ACC	TPR	TNR	PPV	F1	MCC
Fibres	**0.99**	0.825	**0.997**	0.942	0.880	**0.877**
Nanowires	0.961	**0.987**	0.704	**0.97**	**0.978**	0.753
Porous Sponge	0.989	0.8	0.996	0.9	0.847	0.843
Powder	0.894	0.921	0.766	0.95	0.935	0.651
Total	**0.959**	**0.883**	**0.866**	**0.941**	**0.910**	**0.781**

$$TNR = \frac{N_{TN}}{N_{TN} + N_{FP}} \tag{15}$$

$$PPV = \frac{N_{TP}}{N_{TP} + N_{FP}} \tag{16}$$

$$F_1 = \frac{2N_{TP}}{2N_{TP} + N_{FP} + N_{FN}} \tag{17}$$

$$MCC = \frac{N_{TP}N_{TN} - N_{FP}N_{FN}}{\sqrt{(N_{TP} + N_{FP})(N_{TP} + N_{FN})(N_{TN} + N_{FP})(N_{TN} + N_{FN})}} \tag{18}$$

Here, N_{TP}, N_{TN}, N_{FP}, and N_{FN} define the number of correctly classified classes, the number of correctly classified opposite classes, the number of incorrectly classified classes, and the number of the misclassified classes, respectively.

Stage 1: In this stage, the SVMC was fed by the deep features obtained from the retrained AlexNet's fc6 layer. Fig. 4a illustrated the distribution of features obtained from the fc6 layer in the 2D feature space to emphasize the difficulty of the classification process. We visualized the scatter plot of the features in the only first two columns of the 5079 × 4096 matrix. Here X represents the 1 st column, and Y represents the 2nd column. Fig. 4b allows us to see how AlexNet, trained on the SEM dataset, builds up its understanding of images over the fc6 layer. It is the same as how neurons along the way to the vision center of living beings detect complex patterns.

Fig. 5a demonstrates the multi-class confusion matrix of the SVM classifier test process. To calculate the performance criteria of the classification algorithm, N_{TP}, N_{TN}, N_{FP}, and N_{FN} values were obtained by using the test samples shown in the confusion matrix. While the diagonal values give N_{TP}, the sum of the row data of each N_{TP} value gives N_{FN}, and the sum of the column data gives N_{FP}. Also, the N_{TN} value for each class is found by the sum of the cornerstones of the 3 × 3 matrix formed by the center N_{TP}. As shown in Fig. 5a, the false-negative rate appears majorly within Nanowires and Powder. Nanowires class achieved nearly the perfect classification, whereas the Powder class had an accuracy of 89.4 % within 40 and 229 test samples, respectively. In the Powder class distribution, 16 samples were misclassified as Nanowires, and 2 samples were misclassified as Porous Sponge.

In Table 2, we can see the detailed classification results of the AlexNet. The obtained results show that the Nanowires class had higher TPR values, PPV, and F_1 as 98.7 %, 97 %, and 97.8 %, respectively. While the Fibres class presented the highest ACC, TNR, and MCC, the

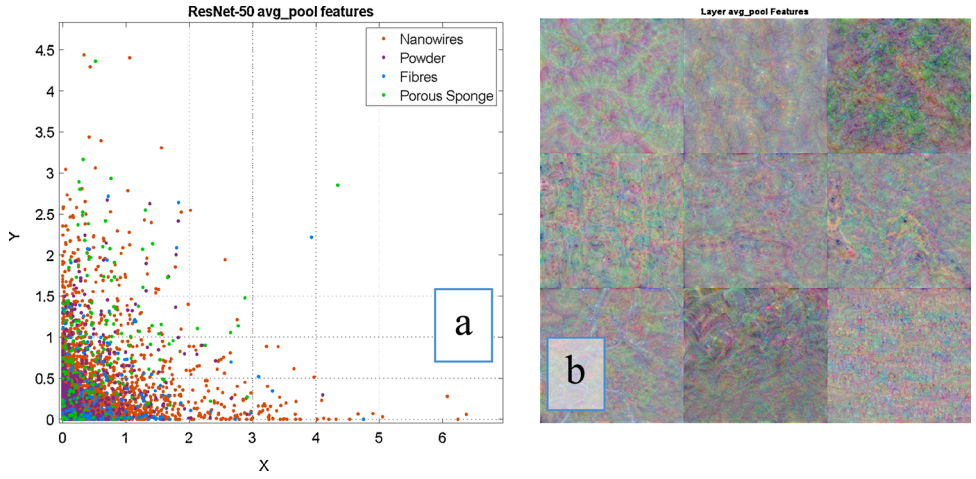

Fig. 6. a) The scatter plot of the features from ResNet-50 avg_pool layer. b) The visualization of the first 9 features learned by the avg_pool layer.

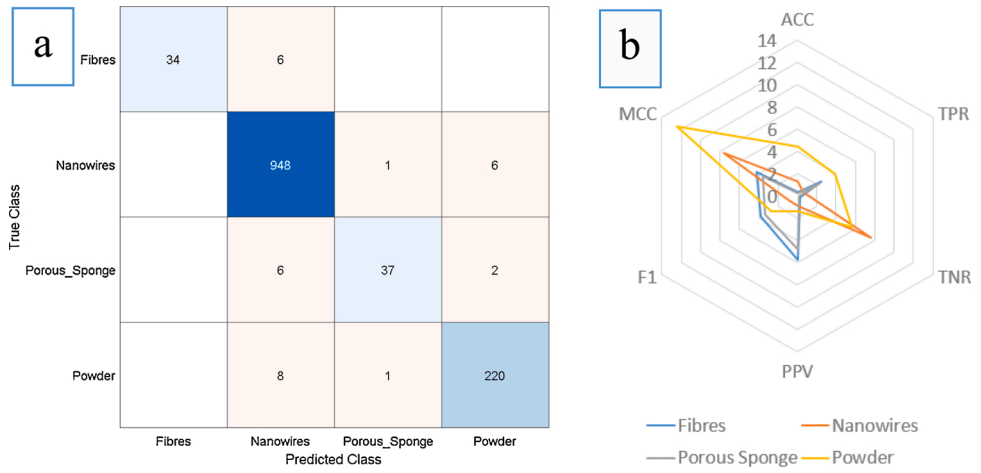

Fig. 7. a) The multi-class confusion matrix of the SVMC with avg_pool features. b) Demonstration of the performance alteration of avg_pool relative to fc6.

Table 3
Classification scores of the SVMC with deep features from the avg_pool layer.

Class	ACC	TPR	TNR	PPV	F1	MCC
Fibres	**0.993**	0.850	**1**	**1**	0.918	**0.919**
Nanowires	0.974	**0.993**	0.780	0.979	**0.986**	0.829
Porous Sponge	0.991	0.822	0.998	0.948	0.881	0.879
Powder	0.938	0.960	0.822	0.964	0.962	0.776
Total	**0.974**	**0.906**	**0.900**	**0.973**	**0.937**	**0.851**

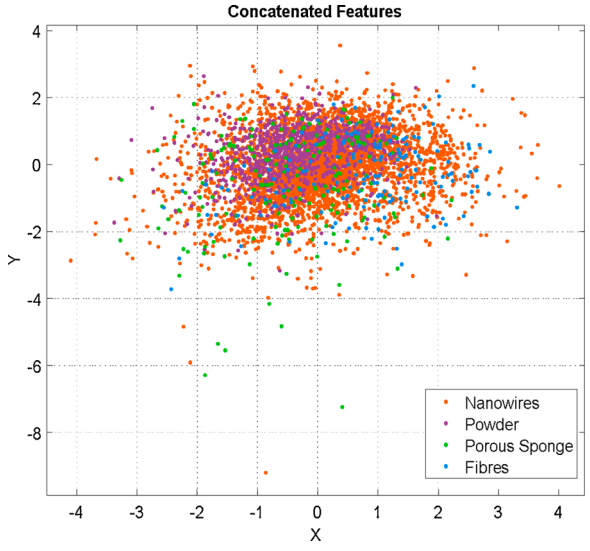

Fig. 8. The scatter plot of the features from AlexNet + ResNet-50.

model with fc6 features reached an overall accuracy of 95.9 %.

Stage 2: The scatter plot of the features obtained with the retrained ResNet-50 avg_pool layer is shown in Fig. 6a. Also, the first 9 features learned by the avg_pool layer are visualized in Fig. 6b. Fig. 7a presents the confusion matrix of the classifier.

According to Fig. 7a, there are 37 misclassified samples among 1279 test samples. Nanowires presented the most misclassification samples with 27. In Table 3, we can see the detailed classification results of the ResNet-50. The obtained results show that the Fibres class had higher ACC, TNR, PPV, and MCC values as 99.3 %, 100 %, 100 %, and 91.9 %, respectively. While the Nanowires class presented the highest TPR and F1, the model with avg_pool features reached an overall accuracy of 97.4 %. When AlexNet and ResNet-50 performance criteria are compared, it is seen that the highest increases in PPV and F1 values were in the Fibres class with 5.8 % and 3.8 %, respectively. It was calculated that the highest increase in the TNR value was in Nanowires with 7.6 %. The highest increases in ACC, TPR, and MCC were seen in Powder with 4.4 %, 3.9 %, and 12.5 %, respectively. The radar chart where all these comparisons are visualized is shown in Fig. 7b.

Stage 3: Fig. 8 depicts the scatter plot of the new concatenated vector obtained by linking the feature vectors of both networks. Hence, all of the positive and negative aspects of both networks were combined. This total deep feature vector, which feeds the SVM classifier's input, also contains many excessive elements. Although the improvement in system performance seems positive, the high-dimensional fused vector and the long processing time also have negative effects. The feature reduction method was used to eliminate this problem in the next step. While the number of misclassified samples was 37 in ResNet-50, it was 13 in the new model obtained using the concatenated feature vector. The most misclassification samples dropped from 27 to 4 in the Nanowires class. Here, we can see a remarkable performance enhancement in all classes and overall criteria.

According to the chart given in Fig. 9b, it is seen that the highest increases in TPR, PPV, and F1 values were in the Porous Sponge class with 8.9 %, 5.2 %, and 7.2 %, respectively when ResNet-50 and AlexNet + ResNet-50 performance criteria are compared. It was calculated that the highest increase in the TNR value was in Nanowires with 11.6 %. The highest increase in ACC and MCC was seen in Powder as 4.3 % and 15.5 %, respectively. The detailed classification performance criteria are given in Table 4.

Stage 4: In this step, a feature selection method with the BPSO algorithm was employed to select the most characteristic features from the concatenated deep features. After that, an SVM classifier was employed to label the test SEM images with the selected features. The efficient deep features were determined with a simple BPSO by considering a feature range and their achievements. Dimension reduction in the concatenated feature vector was completed by taking feedback from the KNN objective function. Thus, a feature set with the highest accuracy could be identified. The parameters of the optimization algorithm are given in Table 5;

Fig. 10a illustrates the convergence curve of BPSO methods for the concatenated deep features. In Fig. 10b, the scatter plot of optimized

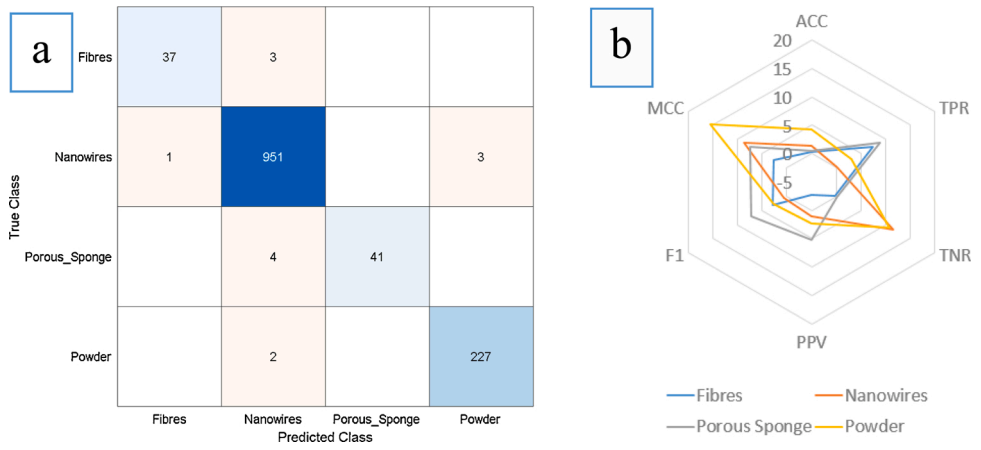

Fig. 9. a) The multi-class confusion matrix of the SVMC with concatenated features. b) Demonstration of the performance alteration of concatenated features relative to avg_pool.

Table 4
Classification scores of the SVMC with deep features from fc6+avg_pool layers.

Class	ACC	TPR	TNR	PPV	F1	MCC
Fibres	0.996	0.925	0.998	0.973	0.948	0.947
Nanowires	**0.987**	**0.995**	0.896	0.99	**0.993**	0.917
Porous Sponge	0.996	0.911	**1**	**1**	0.953	**0.953**
Powder	0.981	0.991	0.931	0.987	0.989	0.931
Total	**0.990**	**0.956**	**0.956**	**0.988**	**0.971**	**0.937**

Table 5
The parameters of the BPSO algorithm.

Parameters		Value
N	Number of particles	10
T	Maximum number of iterations	30
C_1	Cognitive factor	1
C_2	Social factor	1
V_{max}	Maximum velocity	5
W_{max}	Maximum bound on inertia weight	0.8
W_{min}	Minimum bound on inertia weight	0.3

features is given. According to Table 5, the total feature vector's size decreased from 5096 × 6144 to 5096 × 3112 after optimization, while the total accuracy value increased to 99.3 %. The most significant improvement was seen in Nanowires with a 4.4 % TNR value, while a performance improvement was observed in all classes. While the number of misclassified samples was 13 in AlexNet + ResNet-50, this value was 9 in the new model obtained using the optimized feature vector. The misclassified samples in the Porous Sponge class completely disappeared.

According to the radar chart given in Fig. 11b, it is seen that the highest increase in PPV and F1 values were in the Fibres class with 2.7 % and 1.3 %, respectively, when AlexNet + ResNet-50 and optimized AlexNet + ResNet-50 performance criteria are compared. It was calculated that the highest increase in TNR and MCC values was in Nanowires with 4.4 % and 3.1 %, respectively. Besides, the highest increase in ACC was seen in Nanowires with 0.5 %. The class with the highest increase in TPR was Porous Sponge. According to the detailed classification performance criteria given in Table 6, F1 and MCC values also prove that a stable classification was obtained.

One of the most widely used metrics to evaluate machine learning algorithms' performance is the Receiver Operating Characteristic (ROC) curve, which provides the true-positive rate as a function of the false-positive rate. It is a critical evaluation tool to check any classification model's performance on balanced and imbalanced binary prediction problems. The ROC curve of our four classes is shown in Fig. 12. All classes had a similar performance according to the AUC with the optimized AlexNet + ResNet-50 features. The area of the ROC curve was obtained as 1 for all classes. The success rate of the obtained model was high; furthermore, it should be taken into consideration that the area under the curve was nearly perfect.

Table 7 shows the evaluation metrics of SVMC for all of the methods in the study. The results in the table indicate that the proposed method can provide an improvement in MCC. When there is an imbalance

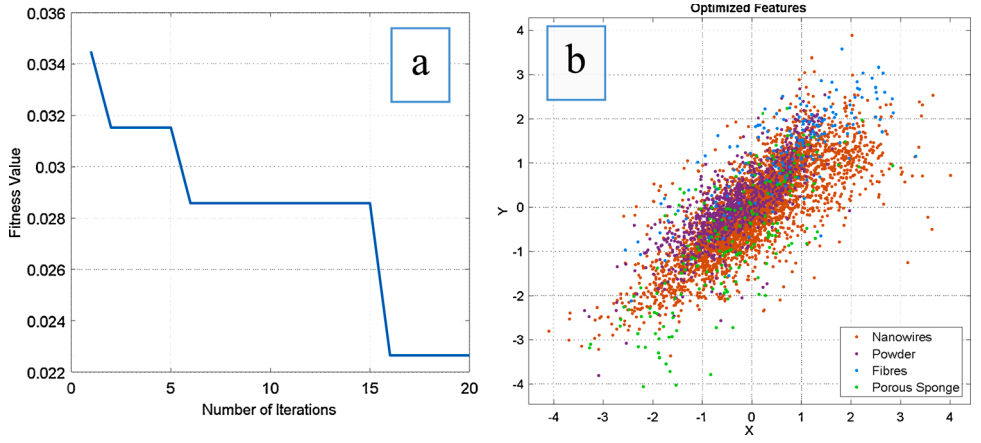

Fig. 10. a) The convergence curve of the BPSO method. b) The scatter plot of the optimized features from AlexNet + ResNet-50.

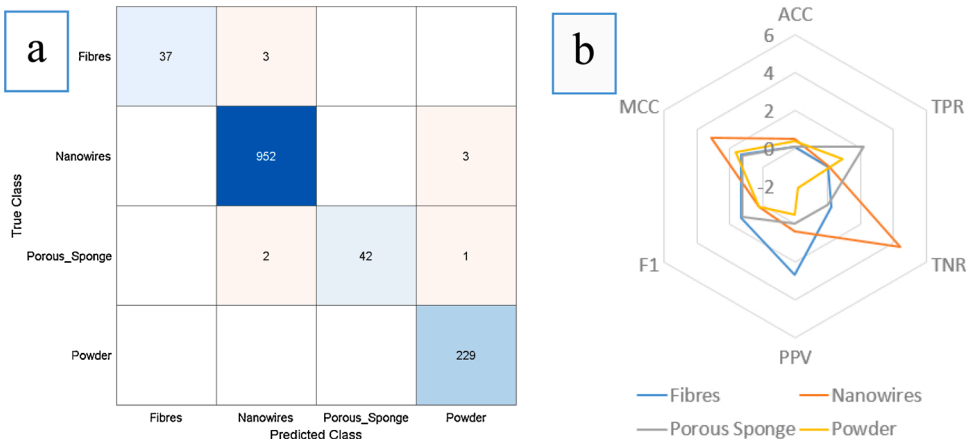

Fig. 11. a) The multi-class confusion matrix of the SVMC with optimized features.
b) Demonstration of the performance alteration of optimized features relative to concatenated ones.

Table 6
Classification scores of the SVMC with optimized features from fc6+avg_pool layers.

Class	ACC	TPR	TNR	PPV	F1	MCC
Fibres	0.997	0.925	**1**	**1**	0.961	0.960
Nanowires	**0.992**	0.996	0.94	0.994	**0.995**	0.948
Porous Sponge	0.997	0.933	**1**	**1**	0.965	**0.965**
Powder	0.985	**1**	0.913	0.982	0.991	0.947
Total	**0.993**	**0.964**	**0.963**	**0.994**	**0.978**	**0.955**

problem between classes in the dataset, ACC and F1 score cannot be considered reliable metrics since it does not take into account the ratio between positive and negative elements. These evaluation metrics can lead to inflated, overly optimistic results. A practical solution to the class imbalance issue comes from the MCC through its mathematical properties that incorporate the dataset imbalance and its invariability for class swapping [28].

Fig. 13 depicts a comparison of performances of the proposed method and other features set. The proposed method provided 2%, 10 %, and 17 % enhancements over Concatenated, ResNet-50, and AlexNet, respectively. It has been demonstrated that a more stable separation is performed by discarding unnecessary features in the decision-making process for the classification mechanism.

A comparison of the SVMC performance with preferred feature sets was given in Table 8. While the first column of Table 8 shows the pre-trained net, the second column shows the constructed feature sets. The rest of the columns show the feature sizes, ACC values, prediction speed, training time, and size of the features file in megabytes. The best results for the SEM dataset were achieved by the optimized features from AlexNet (fc6) + ResNet-50 (avg_pool). The obtained ACC, TPR, TNR, PPV, F1, and MCC scores were 99.3 %, 96.4 %, 96.3 %, 99.4 %, 97.8 %, and 95.5 %, respectively. Even though the accuracy rates were close to each other in both networks, the optimized fc6+ avg_pool was ahead in terms of training time and data size.

When the proposed approach's performance is compared between the state-of-the-art methods, it has been observed that the structure provides superiority in most of the evaluation metrics. The first implementation of the classification process with SEM images is presented by Modarres et al. [2]. It was preferred to use more subclasses with the same dataset as ours. The features were obtained by retraining the Inception-v3 model using the transfer learning method. The final model achieved 90 % accuracy, 80 % precision, and 90 % sensitivity. In another study, Chowdhury et al. proposed feature extraction and dimensionality reduction methods to classify microstructural image data [29]. The light optical microscopy was used instead of the SEM for imaging the dendritic microstructure. Maximum classification accuracies of 91.85 and 97.37 for different tasks were achieved. However, the dataset and imaging techniques are entirely different from ours.

Similarly, microstructural image data for the seven microstructure classes were used in the classification process in [30]. The features were obtained by the histogram-based method, and the classification process was performed via the SVMC. The classifier system achieves a 5-fold cross-validation accuracy of 83 %.

The absence of such a large-scale dataset has restricted the variety of studies in the literature. The proposed approach has a pioneering structure in this particular field since deep neural networks have been considered in a limited number of applications.

4. Conclusions

In this study, a first public dataset of SEM images that can be used as a reference for future deep learning applications in the field of nanoscience was used. We proposed a method for feature extraction from an AlexNet and ResNet-50 by retraining the fully-connected (fc6) layer and the average pooling layer (avg_pool) on the SEM dataset. The extracted feature vectors were then used in feature concatenation to form a single deep feature representation of the input SEM images. A feature selection method with a Binary Particle Swarm Optimization algorithm was employed to select the most characteristic features from the concatenated deep features. A support vector machine classifier was employed to label the test SEM images with the selected features. The performance of the final successful model obtained by the proposed method was tested by using SEM images of pure SnO_2 and Ag-doped SnO_2 nanopowders synthesized by the hydrothermal method in the laboratory. The best accuracy value was observed with 3112 features for the SEM dataset with optimized vectors. The obtained ACC, TPR, TNR, PPV, F1, and MCC scores were 99.3 %, 96.4 %, 96.3 %, 99.4 %, 97.8 %, and 95.5 %, respectively.

1 The feature selection method based on the metaheuristic optimization algorithm was used to achieve lower computation costs and higher accuracy performance by reducing feature sizes.
2 The performance of the trained model was tested using SEM images of Ag-doped SnO_2 nanoparticles, which were not available in the database and were prepared by the author using the low-temperature hydrothermal method. Consequently, it was proved that the model is intended for the end-user by applying the database-independent test procedure and can be embedded in hardware with size reduction.
3 An example was illustrated where the feature selection with the BPSO technique could provide novel insight into nanoscience

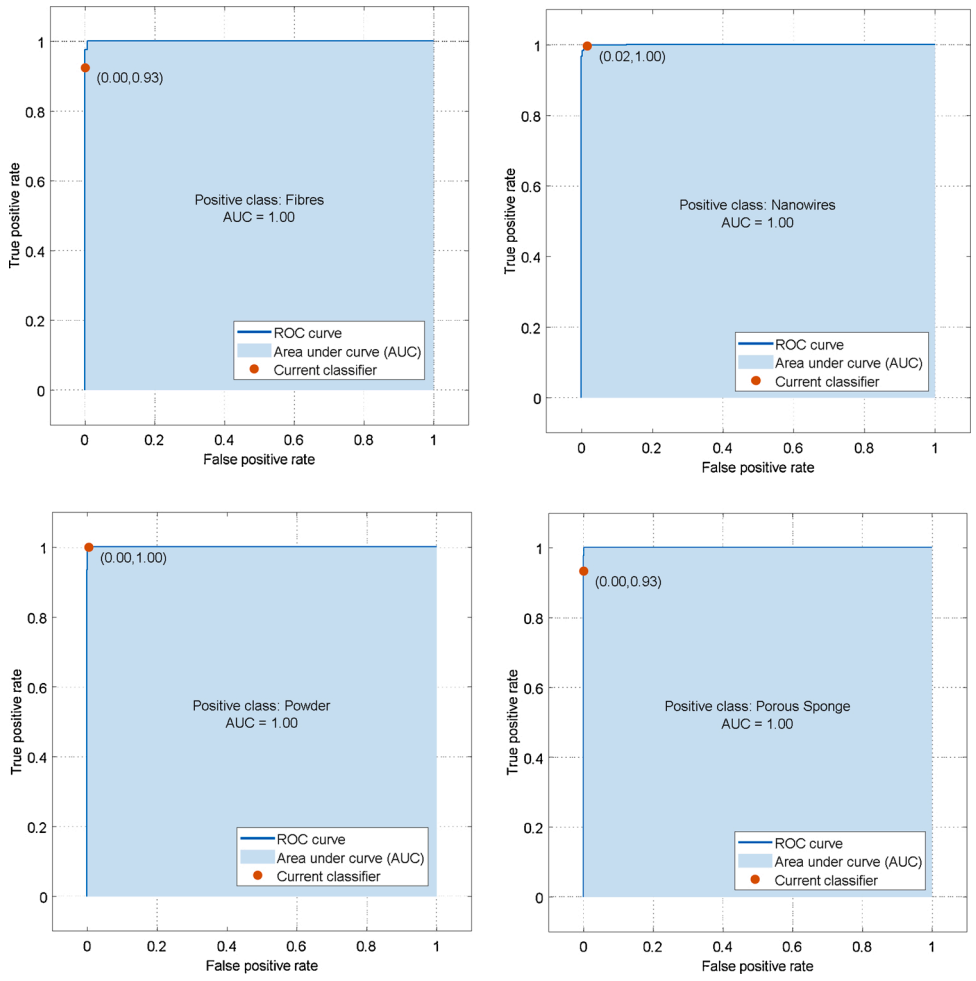

Fig. 12. The ROC curves of our four classes.

Table 7
Performance comparison between the proposed method and the other pre-trained nets.

	ACC	TPR	TNR	PPV	F1	MCC
AlexNet	95.85 %	88.33 %	86.58 %	94.05 %	91.00 %	78.10 %
ResNet-50	97.40%	90.63 %	90.00 %	97.28 %	93.68 %	85.08 %
Concatenated	99.00 %	95.55 %	95.63 %	98.75 %	97.08 %	93.70 %
Optimized	99.28 %	96.35 %	96.33 %	99.40 %	97.80 %	95.50 %

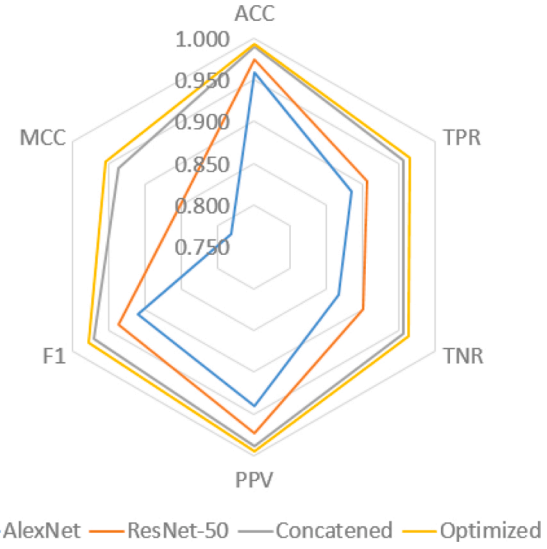

Fig. 13. Radar chart of the performance in terms of all feature sets.

research and test the model with the SEM images of Ag-doped SnO_2 particles obtained by the hydrothermal method.

CRediT authorship contribution statement

Gürkan Kavuran: Conceptualization, Data curation, Methodology, Software, Visualization, Investigation, Validation, Writing - review & editing.

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.

Table 8
Comparison of the SVMC performance with preferred feature sets.

Pretrained Net	Features	Feature Sizes	Accuracy	Prediction Speed (obs/sec)	Training Time (sec)	Size (MB)
AlexNet	fc6	5079 × 4096	0.959	500	45.594	80
ResNet-50	avg_pool	5079 × 2048	0.974	1400	20.267	40
AlexNet + ResNet-50	fc6+ avg_pool	5079 × 6144	0.990	270	65.147	120
AlexNet + ResNet-50	Optimized fc6+ avg_pool	5079 × 3112	0.993	1100	27.22	60

References

[1] M. Ge, F. Su, Z. Zhao, D. Su, Deep learning analysis on microscopic imaging in materials science, Mater. Today Nano. 11 (2020), 100087, https://doi.org/10.1016/j.mtnano.2020.100087.

[2] M.H. Modarres, R. Aversa, S. Cozzini, R. Ciancio, A. Leto, G.P. Brandino, Neural network for nanoscience scanning electron microscope image recognition, Sci. Rep. (2017), https://doi.org/10.1038/s41598-017-13565-z.

[3] A. Rezaie, R. Achanta, M. Godio, K. Beyer, Comparison of crack segmentation using digital image correlation measurements and deep learning, Constr. Build. Mater. [Accepted] (2020), 120474, https://doi.org/10.1016/j.conbuildmat.2020.120474.

[4] Z. Tan, Q. Fang, H. Li, S. Liu, W. Zhu, D. Yang, Neural network based image segmentation for spatter extraction during laser-based powder bed fusion processing, Opt. Laser Technol. 130 (2020), https://doi.org/10.1016/j.optlastec.2020.106347.

[5] T. Stan, Z.T. Thompson, P.W. Voorhees, Optimizing convolutional neural networks to perform semantic segmentation on large materials imaging datasets: X-ray tomography and serial sectioning, Mater. Charact. 160 (2020), https://doi.org/10.1016/j.matchar.2020.110119.

[6] D. McMullan, Scanning electron microscopy 1928-1965, Scanning 17 (2006) 175–185, https://doi.org/10.1002/sca.4950170309.

[7] R. Aversa, M.H. Modarres, S. Cozzini, R. Ciancio, A. Chiusole, Data descriptor: the first annotated set of scanning electron microscopy images for nanoscience, Sci. Data (2018), https://doi.org/10.1038/sdata.2018.172.

[8] A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks. Adv. Neural Inf. Process. Syst., 2012, https://doi.org/10.1061/(ASCE)GT.1943-5606.0001284.

[9] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (2016), https://doi.org/10.1109/CVPR.2016.90.

[10] B. Tran, B. Xue, M. Zhang, Variable-length particle swarm optimization for feature selection on high-dimensional classification, IEEE Trans. Evol. Comput. (2019), https://doi.org/10.1109/TEVC.2018.2869405.

[11] J. Too, A.R. Abdullah, N.M. Saad, W. Tee, EMG feature selection and classification using a Pbest-guide binary particle swarm optimization, Computation 7 (2019), https://doi.org/10.3390/computation7010012.

[12] L.Y. Chuang, H.W. Chang, C.J. Tu, C.H. Yang, Improved binary PSO for feature selection using gene expression data, Comput. Biol. Chem. (2008), https://doi.org/10.1016/j.compbiolchem.2007.09.005.

[13] B. Xue, M. Zhang, W.N. Browne, Particle swarm optimisation for feature selection in classification: novel initialisation and updating mechanisms, Appl. Soft Comput. J. (2014), https://doi.org/10.1016/j.asoc.2013.09.018.

[14] J. Kennedy, R. Eberhart, Particle swarm optimization, IEEE Int. Conf. Neural Networks - Conf. Proc. (1995), https://doi.org/10.4018/ijmfmp.2015010104.

[15] R. Eberhart, J. Kennedy, New optimizer using particle swarm theory, Proc. Int. Symp. Micro Mach. Hum. Sci. (1995), https://doi.org/10.1109/mhs.1995.494215.

[16] Y. Shi, R. Eberhart, Modified particle swarm optimizer, Proc. IEEE Conf. Evol. Comput. ICEC (1998), https://doi.org/10.1109/icec.1998.699146.

[17] Y. Shi, R.C. Eberhart, Parameter selection in particle swarm optimization. Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), Springer Verlag, 1998, pp. 591–600, https://doi.org/10.1007/bfb0040810.

[18] J. Too, A.R. Abdullah, N.M. Saad, A new co-evolution binary particle swarm optimization with multiple inertia weight strategy for feature selection, Informatics 6 (2019), https://doi.org/10.3390/informatics6020021.

[19] N. Al-Madi, H. Faris, S. Mirjalili, Binary multi-verse optimization algorithm for global optimization and discrete problems, Int. J. Mach. Learn. Cybern. (2019), https://doi.org/10.1007/s13042-019-00931-8.

[20] E. Emary, H.M. Zawbaa, A.E. Hassanien, Binary ant lion approaches for feature selection, Neurocomputing (2016), https://doi.org/10.1016/j.neucom.2016.03.101.

[21] C. Cortes, V. Vapnik, Support-vector networks, Mach. Learn. (1995), https://doi.org/10.1023/A:1022627411411.

[22] V. Vapnik, The support vector method of function estimation. Nonlinear Model., 1998, https://doi.org/10.1007/978-1-4615-5703-6_3.

[23] C.J.C. Burges, A tutorial on support vector machines for pattern recognition, Data Min. Knowl. Discov. (1998), https://doi.org/10.1023/A:1009715923555.

[24] E. Osuna, R. Freund, F. Girosi, Support Vector Machines: Training and Applications, 1997.

[25] B. Schölkopf, Learning with kernels, Proc. 2002 Int. Conf. Mach. Learn. Cybern. (2002), https://doi.org/10.7551/mitpress/4175.001.0001.

[26] S. Abe, Support Vector Machines for Pattern Classification, Springer, London, London, 2010, https://doi.org/10.1007/978-1-84996-098-4.

[27] S.R. Gunn, Support vector machines for classification and regression. Technical Report, 1998.

[28] D. Chicco, G. Jurman, The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation, BMC Genomics 21 (2020) 1–13, https://doi.org/10.1186/s12864-019-6413-7.

[29] A. Chowdhury, E. Kautz, B. Yener, D. Lewis, Image driven machine learning methods for microstructure recognition, Comput. Mater. Sci. (2016), https://doi.org/10.1016/j.commatsci.2016.05.034.

[30] B.L. Decost, E.A. Holm, A computer vision approach for automated analysis and classification of microstructural image data, Comput. Mater. Sci. (2015), https://doi.org/10.1016/j.commatsci.2015.08.011.