ORIGINAL ARTICLE

NFSDense201: microstructure image classification based on non-fixed size patch division with pre-trained DenseNet201 layers

Prabal Datta Barua[1,2,3,4,5,6,7,8,9] · Sengul Dogan[10] · Gurkan Kavuran[11] · Turker Tuncer[10] ·
Ru-San Tan[12,13] · U. Rajendra Acharya[14]

Abstract

In the field of nanoscience, the scanning electron microscope (SEM) is widely employed to visualize the surface topography and composition of materials. In this study, we present a novel SEM image classification model called NFSDense201, which incorporates several key components. Firstly, we propose a unique nested patch division approach that divides each input image into four patches of varying dimensions. Secondly, we utilize DenseNet201, a deep neural network pretrained on ImageNet1k, to extract 2920 deep features from the last fully connected and global average pooling layers. Thirdly, we introduce an iterative neighborhood component analysis function to select the most discriminative features from the merged feature vector, which is formed by concatenating the four feature vectors extracted per input image. This process results in a final feature vector of optimal length 698. Lastly, we employ a standard shallow support vector machine classifier to perform the actual classification. To evaluate the performance of NFSDense201, we conducted experiments using a large public SEM image dataset. The dataset consists of 972, 162, 326, 4590, 3820, 3925, 4755, 181, 917, and 1624.jpeg images belonging to the following microstructural categories: "biological," "fibers," "film-coated surfaces," "MEMS devices and electrodes," "nanowires," "particles," "pattern surfaces," "porous sponge," "powder," and "tips," respectively. For both four-class and ten-class classification tasks, we evaluated NFSDense201 using subsets of the dataset containing 5080 and 21,272 images, respectively. The results demonstrate the superior performance of NFSDense201, achieving a four-class classification accuracy rate of 99.53% and a ten-class classification accuracy rate of 97.09%. These accuracy rates compare favorably against previously published SEM image classification models. Additionally, we report the performance of NFSDense201 for each class in the dataset.

Keywords Non-fixed size patch division · Deep feature engineering · SEM image classification · DenseNet201

1 Introduction

Nanoscience, a burgeoning research field focused on the exploration of new materials and the characterization of their microscopic properties, has witnessed remarkable growth [1, 2]. This expansion has led to the development of nano-sized particle imaging systems tailored for materials science applications [3]. Among the various tools utilized by scientists in this domain, the scanning electron microscope (SEM) holds a prominent position, enabling the visualization of surface topography and composition for samples of interest [4, 5]. Conventionally, these acquired SEM images are manually annotated by laboratory scientists and subsequently archived in dedicated data repositories [6]. Recognizing the need to facilitate the sharing of the ever-increasing volume of SEM images, the Nanoscience Foundries & Fine Analysis (NFFA)–EUROPE project, a distributed research infrastructure spanning Europe, has established the Information and Data Repository Platform (IDRP) [7]. The IDRP serves as a centralized entry point, ensuring harmonized data policies and facilitating access to SEM data for the scientific community [7]. For effective data sharing, the data must be findable, accessible, interoperable, and reusable [8]. Consequently, it becomes evident that automated classification of SEM images, integrated into the data warehouse as a complementary curation function, is essential for the IDRP or any other significant SEM dataset [6, 9].

Extended author information available on the last page of the article

The proliferation of SEM images capturing various types and sizes of composite materials has presented challenges in their classification, prompting the application of machine learning techniques [6, 10, 11]. In [11], deep learning techniques were employed for nanoparticle detection, while [10] utilized machine learning methods for mineral classification. A U-Net method was adopted to classify scanning electron microscopy-energy-dispersive X-Ray spectroscopy (SEM–EDS) images, achieving an F1-micro score of 88.32% across 12 classes. Ge et al. [12] presented a deep learning model in a review paper that leveraged the computer vision capabilities of convolutional neural networks to extract morphology, distribution, and intensity information from microscopic images. Modarres et al. [9] explored transfer learning in a deep learning-based SEM image classification model, utilizing four pre-trained convolutional neural networks (Inception-slim, Inception-v3, Inception-v4, and ResNet) from the ImageNet1k dataset [13]. These networks were employed to extract features from an SEM dataset comprising 18,577 images distributed across ten classes [6]. The Inception-v3 model achieved approximately 90% accuracy. Transfer learning, in this context, involves fixing the pre-trained initial layers of the convolutional neural networks and training only the final few layers on the target dataset to learn specific features. Consequently, feature extraction becomes less computationally intensive compared to training an entirely new network. Additionally, computer vision techniques have gained widespread adoption for automated image classification tasks in recent years [14]. In [15], a multilayer perceptron with backpropagation training algorithm was employed to automatically segment and classify high-resolution micrographs of cast iron images for non-destructive testing, yielding results comparable to manual human visual classification. Similarly, in [16], a "bag of features" approach was utilized to construct microstructural signatures for classifying 105 micrographs of metallic materials based on similarity matching with local image patterns, achieving an accuracy of 83%. The success of computer vision techniques in automated microstructural analysis has stimulated efforts to explore their applicability in SEM image classification. Osenberg et al. [17] introduced a feature engineering approach for the classification of SEM images. They employed threshold-based models to extract features and subsequently utilized random forest (RF) classifiers to classify the selected features. Their method achieved an impressive classification accuracy of 94%. However, the authors did not provide a comprehensive presentation of their results. In related work, Han et al. [18] proposed a novel synthetic image generation model. They further introduced an attention-based convolutional neural network (CNN) architecture by incorporating two pooling functions and multiplicative

operations. Moreover, they addressed the issue of vanishing gradients by employing residual blocks. Their model achieved a higher accuracy of 95% and was compared against well-known CNN architectures such as MobileNet, VGG16, and ResNet50. Nonetheless, the authors did not explore the utilization of high-performing CNN models, such as DenseNets or EfficientNets. Dahy et al. [19] proposed a feature selection model utilizing a metaheuristic optimization technique. They applied their feature selector to the deep features extracted from SEM images. Unfortunately, their model lacked novelty, as they solely focused on evaluating the performance of their proposed feature selector. Moreover, it is worth noting that metaheuristic optimization-based feature selectors tend to exhibit high time complexity. Scott- Fordsmand and Amorim's paper, [20], extensively discussed the profound impact of leveraging machine learning models for the automatic classification of nanomaterials. The authors emphasized the intrinsic significance of this field, as it directly influences various aspects of human life. However, the paper lacked the presentation of any proposed models or the provision of classification results.

1.1 Motivation and the proposed model

SEM images play a crucial role in the field of material sciences. To reduce classification costs, machine learning models have gained popularity for automating the classification of SEM images [6, 10, 11]. In this study, we propose a feature engineering model that combines patch division techniques inspired by computer vision approaches [21] with transfer learning using a pre-trained convolutional neural network.

Traditional fixed-sized patch division models, such as the vision transformer [21] and multilevel perceptron-mixer [22], have demonstrated impressive classification performance. However, their utility has been limited by the high dimensionality of the extracted feature vectors. To address this limitation, we introduce a novel nested patch division method that divides input images into non-fixed size patches. This approach reduces the number of patches required to cover the entire image and enhances computational efficiency compared to standard fixed size patch division methods.

Our inspiration for feature extraction stems from the successes achieved by pre-trained convolutional neural networks and patch-based models. However, standard fixed-size patch-based models often impose significant computational burdens. To mitigate this challenge, we employ a nested patch division model that necessitates fewer patches to encompass the entire input image. Specifically, we utilize DenseNet201 as our feature generator, which is a 201-layer convolutional neural network

pre-trained on ImageNet1K [13]. Accordingly, our proposed model is named NFSDense201.

In our approach, we employ iterative neighborhood component analysis (INCA) as the feature selector, followed by a support vector machine (SVM) as the classifier. This combination allows us to effectively extract discriminative features from the merged feature vector obtained by concatenating the features extracted from the nested patches. The SVM then performs the final classification task.

By integrating these components, our NFSDense201 model offers a comprehensive solution for SEM image classification. The proposed nested patch division, feature extraction using DenseNet201, and the combination of INCA and SVM collectively contribute to the model's effectiveness in classifying SEM images.

1.2 Novelties and Contributions

The contributions of this work are outlined below, highlighting the key aspects of our approach:

- *Local Image Feature Extraction:* In computer vision, fixed-size patches are commonly employed for extracting local image features. However, this often leads to high-dimensional feature vectors. To address this, we introduce a novel nested patch division method, enabling comprehensive coverage of the input image with fewer non-fixed size patches. This approach effectively reduces the dimensionality of the extracted feature vectors while preserving the necessary information.
- *Transfer Learning with DenseNet201:* Training image classification models from scratch using unseen images can be computationally demanding. In our study, we leverage the pre-trained DenseNet201 architecture, which was originally trained on the ImageNet1k dataset [23]. By utilizing the already learned features in the initial layers of DenseNet201, we focus on training the last layers specifically for SEM image feature extraction. This strategy significantly reduces the training time required for the model.
- *Efficient Feature Selection with INCA:* NFSDense201 generates redundant features due to the parallel extraction of features from multiple overlapping nested non-fixed size patches. To address this issue, we employ iterative neighborhood component analysis (INCA) to efficiently filter out redundant features. This process results in a highly condensed selected feature vector that contains the most discriminative features [24]. By iterating the feature selection function during the learning process, we determine the optimal length of the selected feature vector specific to our study dataset.

These contributions collectively make NFSDense201 a highly efficient SEM image classification model. We validate our model through extensive training and testing on a large-scale SEM image dataset. Leveraging a standard shallow support vector machine (SVM) classifier, our model achieves excellent classification performance. These results provide strong justification for our design decisions, specifically incorporating the learning of non-fixed size nested patches using a pre-trained deep network into our SEM image classification model.

2 Dataset

The open-access SEM image dataset [6], comprising a total of 21,272 images, was utilized in our study. These images were previously annotated by domain experts and categorized into ten distinct classes: "biological," "fibers," "film-coated surfaces," "MEMS devices and electrodes," "nanowires," "particles," "pattern surfaces," "porous sponge," "powder," and "tips," encompassing 972, 162, 326, 4590, 3820, 3925, 4755, 181, 917, and 1624.jpeg images, respectively. To ensure consistency, we uniformly resized all images to dimensions of 1024×768 pixels.

## 3 NFSDense201 model for SEM image classification

Our classification model comprised the following sequence of operations: image resizing, nested patch division, hybrid deep feature extraction, feature vector concatenation, feature selection, and classification (Fig. 1). The steps are detailed as follows:

Step 1 Resize SEM image to 224×224 sized images.
Step 2 Apply nested patch division to resize images.
Step 3 Extract features by using two layers of the pre-trained DenseNet201.
Step 4 Merge the generated feature vectors.
Step 5 Choose the best features by applying INCA.
Step 6 Classify the generated feature using SVM with a 75:25 split ratio.

Explanations of the individual steps are detailed in the following subsections.

Figure 1 illustrates the architecture of the NFSDense201 SEM image classification model. The input image was divided into four nested patches, each serving as a distinct input to the pre-trained DenseNet201. Through the utilization of the network's global average pooling and fully connected layers, two deep feature vectors were extracted from each patch, resulting in a total of eight (= 4 × 2)

Fig. 1 Schema of the NFSDense201 SEM image classification model. Here, p: patches, f: feature vectors
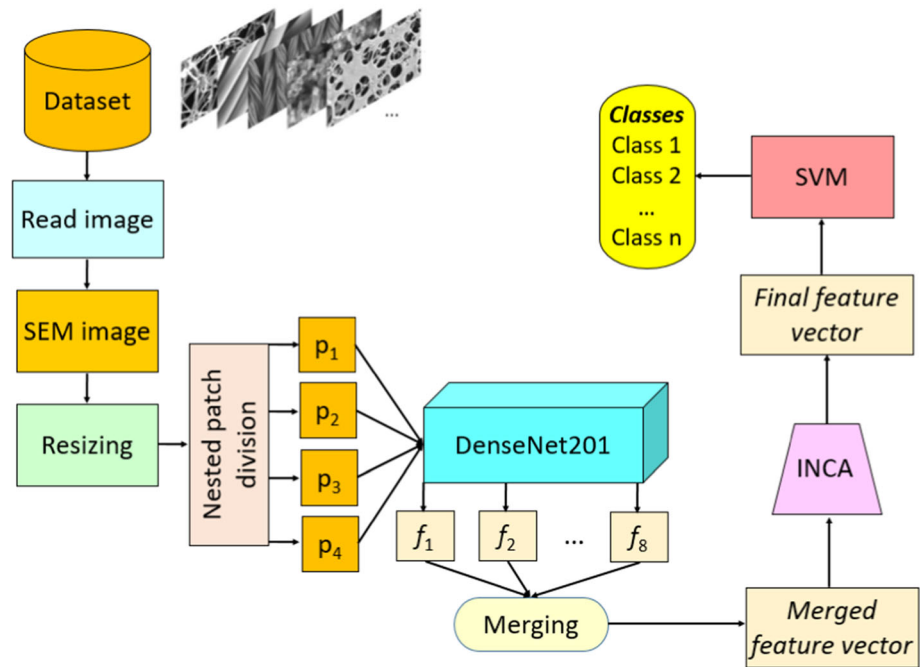
feature vectors per input image. These feature vectors were subsequently concatenated to form a merged feature vector.

To enhance discriminative power, the iterative neighborhood component analysis (INCA) technique was employed. INCA iteratively evaluated the loss value and selected the most significant features during the learning process. This iterative feature selection process enabled the generation of a final feature vector of optimal length that was specific to the dataset under consideration.

The resulting final feature vector was then fed into a standard support vector machine (SVM) classifier for classification. The SVM utilized the extracted features to assign the input image to one of the predefined classes. This sequential process of feature extraction, iterative feature selection, and classification formed the core of the NFSDense201 model's functionality.

3.1 Feature extraction

To preprocess the SEM images from the dataset, a resizing operation was performed, resulting in images of size 224×224, which aligns with the dimensions used in the vision transformer [21]. This image size was also chosen as the input dimension for the DenseNet201 model. A nested patch division approach was employed to create four patches with incremental dimensions: 56×56, 112×112, 168×168, and 224×224 (the last patch being identical to the input image) (refer to Fig. 2). Each patch was then fed into the DenseNet201 model for inductive-based feature extraction.

The DenseNet201 architecture, which had been pretrained on the ImageNet1K database, containing approximately one million images across 1000 classes, was utilized to extract local features from each patch. Specifically, the last fully connected layer (fc1000) and the global average pooling layer (avg_pool) of the DenseNet201 network were employed to generate two deep feature vectors of lengths 1000 and 1920, respectively. The global features were extracted from the last patch, which was identical to the input image, while the local features were extracted from the first three smaller patches.

The resulting feature vectors from the four patches per SEM image were concatenated to form a merged feature vector with a length of 11,680 (= $[1000 + 1920] \times 4$). This merged feature vector captured the combined information from the different patches. The steps involved in this process are summarized as follows:

1. Perform image resizing to obtain images of size 224×224.
2. Utilize a nested patch division algorithm (Algorithm 1) to create non-fixed size patches. The initial patch size was fixed at 56×56, resulting in the generation of four patches per input image (as shown in Fig. 2).

By following these steps, the input SEM images were appropriately processed, and the necessary local and global features were extracted using the DenseNet201 model. Algorithm 1 depicted the presented nested patch division.

Algorithm 1. Pseudocode for nested patch division.

Input: Image (Img) with size 224 × 224. (ins) represent the incremental sizes of the patches.
For this work, ins is set initially as 56.

Output: Patches ($ptch$).

01: $c = \frac{ins}{2}$; // c is a counter.

02: $ce = \frac{224}{2} = 112$; // ce is the center coordinate of the image

03: **for** i=1 to NoP **do** // Herein, NoP is a number of patches. Here, four patches are created

04:    $ptch_i = Img(ce - c + 1: ce + c, ce - c + 1: ce + c, :)$; // Patch dividing.

05:    $c = c + \frac{ins}{2}$;

06: **end for i**

3. Extract deep features using the fc1000 and global avg_pool layers, respectively, of the pre-trained DenseNet201.where fv_i represents the ith feature vector; $\zeta(.)$, fc1000 layer; $\varrho(.)$, global avg_pool layer; and $\gamma(.,.)$, merging function. From each patch, two deep feature vectors of lengths 1000 and 1920 were generated using the fc1000 and global avg_pool layers of DenseNet201. There were concatenated into a feature vector of length 2920.

$$fv_i = \gamma(\zeta(ptch_i), \varrho(ptch_i)), i \in \{1, 2, \ldots, 4\} \quad (1)$$

4. Merge the four feature vectors generated from the four patches to obtain one merged feature vector per input image.

$$feat(j + 2920 \times (i - 1)) = fv_i(j), j \in \{1, 2, \ldots, 2920\} \quad (2)$$

Fig. 2 Nested (non-fixed size) patch division of a sample input SEM image that had been resized to 224 × 224 (top panel). Defining the initializing unit for patch division as 56, four (= 224/56) non-fixed sized patches that were centered on the input image were created with the following incremental dimensions: 56 × 56, 112 × 112, 168 × 168, and 224 × 224 (bottom panel). The fourth patch had the same dimensions as the input image, which allowed for global feature extraction, in addition to local feature extraction from the first three smaller patches
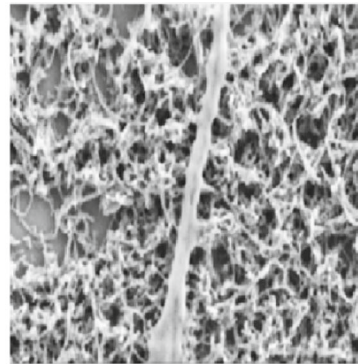
(a) Original image
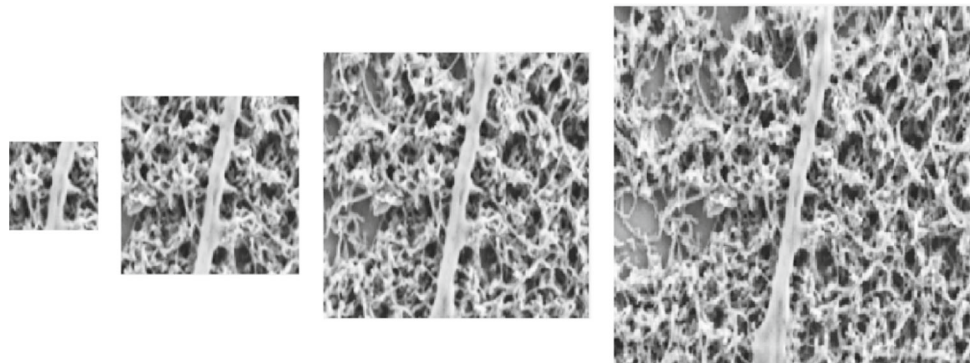
(b) Patches with sizes of 56 × 56, 112 × 112, 168 × 168, and 224 × 224.

where *feat* represents the merged feature vector of length 11,680 (= 2920 × 4).

3.2 Feature selection

The usage of overlapping nested patch-based feature extraction inherently leads to the generation of redundant features in the central region of the input image, as depicted in Fig. 2. To address this issue, our proposed model incorporates INCA, a straightforward yet highly effective feature selection mechanism. INCA operates by iteratively selecting the most discriminative features based on computed loss values [24], thereby filtering out redundant and non-informative features. In our experiments, we set the parameters of INCA as follows: the iteration range was defined from 500 to 1000, and the loss function calculator employed was SVM. Remarkably, when applied to our extensive study dataset consisting of 21,272 images, INCA successfully generated a final feature vector of optimal length 698.

3.3 Classification

For classification purposes, we employed cubic SVM, a widely recognized and efficient shallow classifier [25, 26]. To evaluate the performance of our model, we adopted a 75:25 training-to-test split hold-out validation strategy. The SVM parameters were configured as follows: the kernel utilized was a third-degree polynomial function, the coding scheme employed was one-vs-all, and the box-constraint parameter was set to 1.

4 Experiment

In this section, we have presented our experimental results. Moreover, we have defined two cases to get generalizable results.

4.1 Setup

The SEM image dataset utilized in this study is publicly available. Initially, the dataset was obtained by downloading it from the relevant source. Our model has been implemented within the MATLAB (2021b) environment, making use of a modestly configured personal computer equipped with 16 gigabytes of main memory, a 1 terabyte hard disk, a central processing unit operating at 3.60 gigahertz, and the Windows 11 operating system. To facilitate our implementation, we acquired the pre-trained DenseNet201 network. The proposed model was coded using m files and functions. Furthermore, we employed the

MATLAB classification learner toolbox to generate the SVM code. Moreover, our proposed NFS-DenseNet201 constitutes a parametric deep feature engineering model, and the specific parameters employed are provided as follows.

4.1.1 Deep feature extraction

We conducted fully connected and global average pooling of the pre-trained DenseNet201. The DenseNet201 architecture was used with default settings. For feature extraction, we resized the images to dimensions of 224 × 224. Additionally, we employed four nested patches with sizes of 56 × 56, 112 × 112, 168 × 168, and 224 × 224.

4.1.2 Feature selection

INCA was utilized during the feature selection phase, and the parameters of this selector are presented as follows. We defined the iteration range from 500 to 1000, and the loss function calculator employed was SVM. Notably, the number of iterations of the NCA corresponds to half of the total number of observations. By using a greedy algorithm, the feature vector with the lowest misclassification rate.

4.1.3 Classification

To classify the selected features, we employed a 3rd-degree polynomial order SVM. The settings of the classifier are as follows: the kernel employed was a third-degree polynomial function, the coding scheme utilized was one-vs-all, the box-constraint parameter was set to 1 and validation is 75:25 split ratio.

4.2 Classification tasks

To evaluate the classification performance in a general setting, we devised two distinct cases, each encompassing a different number of classes. The specific details of these cases are outlined as follows:

Case 1: In this case, we utilized a subset of the dataset consisting of 5,080 SEM images. These images were drawn from four categories: "fibers," "nanowires," "porous sponge," and "powder" (used for training and testing).

Case 2: For this case, we employed the entire dataset, comprising a total of 21,272 images spanning all ten categories.

It is worth noting that the optimal number of features selected by INCA in Case 2, namely 698, was applied to both cases. This consistent feature selection approach

allowed us to use the same set of 698 features to evaluate the classification performance in both scenarios.

4.3 Model performance evaluation

The model's performance was evaluated using standard performance metrics [27, 28], namely accuracy, recall, and precision. We calculated both class-wise and overall performance metrics. The mathematical explanations of these performance evaluation parameters are illustrated as follows:

$$accuracy = \frac{tp + tn}{tp + tn + fp + fn} \quad (3)$$

$$recall = \frac{tp}{tp + fn} \quad (4)$$

$$precision = \frac{tp}{tp + fp} \quad (5)$$

$$f1 = \frac{2tp}{2tp + fn + fp} \quad (6)$$

where tp, tp, fn, and fp represent the numbers of true positives, true negatives, false negatives, and false positives, consecutively.

4.4 Results

For both Case 1 and Case 2, the performance of NFSDense201 was truly remarkable, as demonstrated by the excellent overall results presented in Table 1. Additionally, the class-wise classification performance, depicted in Fig. 3, was commendable. In particular, NFSDense201 achieved a remarkable accuracy rate of 99.53% for the four-class classification task, while obtaining a respectable accuracy rate of 97.09% for the ten-class classification task.

Regarding Case 1, the category-wise outcomes varied. The "porous sponge" category exhibited a flawless performance, achieving a perfect classification rate of 100%. On the other hand, the "fibres" category, which constituted the smallest group within the dataset, had a slightly lower classification rate of 97.56%.
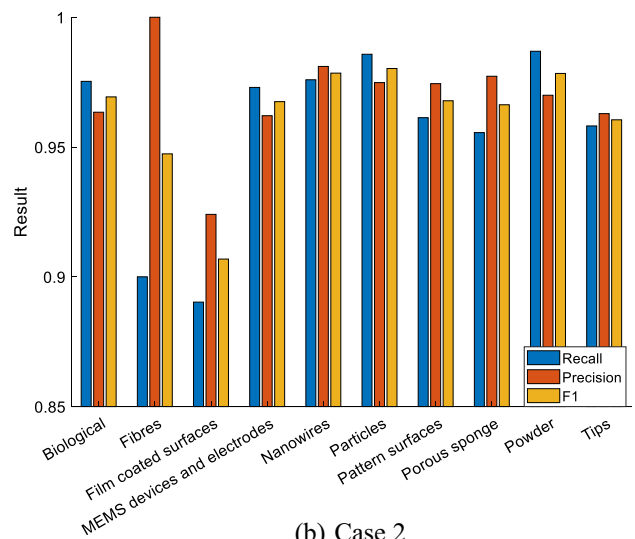
In the context of Case 2, the overall class-wise F1 scores displayed a range of performance levels. The "particles"

Table 1 Overall classification performances of NFSDense201

Performance metrics	Case 1	Case 2
Accuracy (%)	99.53	97.09
Overall precision (%)	99.01	96.90
Overall recall (%)	99.18	95.62
Overall F1 (%)	99.09	96.23
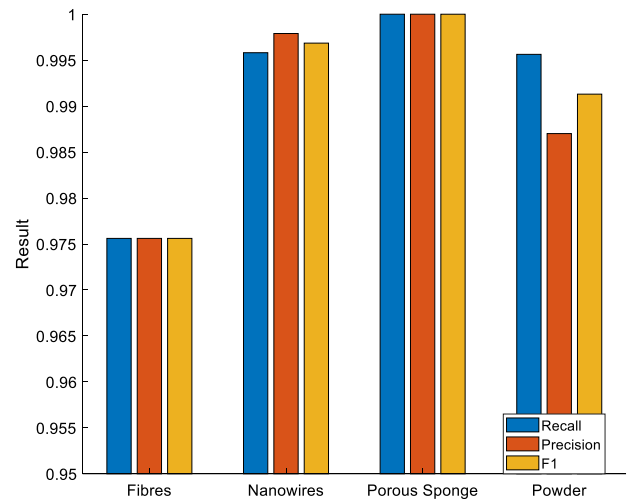
(a) Case 1

(b) Case 2

Fig. 3 Class-wise recall, precision, and F1 scores of NFSDense201 for Case 1 (**a**) and Case 2 (**b**)

category showcased the highest F1 score of 98.03%, signifying a notable classification accuracy. Conversely, the "film-coated surfaces" category demonstrated a comparatively lower F1 score of 90.68%, indicating some room for improvement.

Analyzing the overall class-wise recall for Case 2, it is evident that the "powder" category performed exceptionally well, achieving a recall rate of 98.96%. Conversely, the "fibers" category exhibited a lower recall rate of 89.02%. It is worth noting that despite the relatively lower recall rate for the "fibers" category, the precision was remarkably high at 99.79% for Case 2. This implies that while approximately 11% of the "fibers" images may have been misclassified (as indicated by the recall rate of 89.02%), if an SEM image was classified as "fibers," it

Table 2 State-of-the-art SEM image classification methods

References	Method	Images	Classes	Validation	Results (%)
Kavuran [29]	Transfer learning; metaheuristic optimization; support vector machine	5079	4	75:25 split	Acc 99.30
Tian et al. [33]	Pre-trained VGG16	100	2	70:30 split	Acc 88.00
Yin et al. [34]	Convolutional neural network with attention	3469	4	70:30 split	Acc 98.56
Li et al. [30]	U-Net	15,600*	13	fourfold	Acc 87.00
Leracitano et al. [31]	Multi-layer perceptron	1280	2	Leave-one-out	Acc 92.50
Tsutsui et al. [32]	Gray-level co-occurrence matrix; two classifiers	480	8	80:20 split	Acc 85.00
Our method	NFSDense201	Case 1: 5080	4	75:25 split	Acc 99.53 Pre 99.01 Rec 99.18 F1 99.09
		Case 2: 21,272	10	75:25 split	Acc 97.09 Pre 96.90 Rec 95.62 F1 96.23

*Microscopic images of minerals. Acc, accuracy; F1, F1 score; Pre, precision; Rec, recall
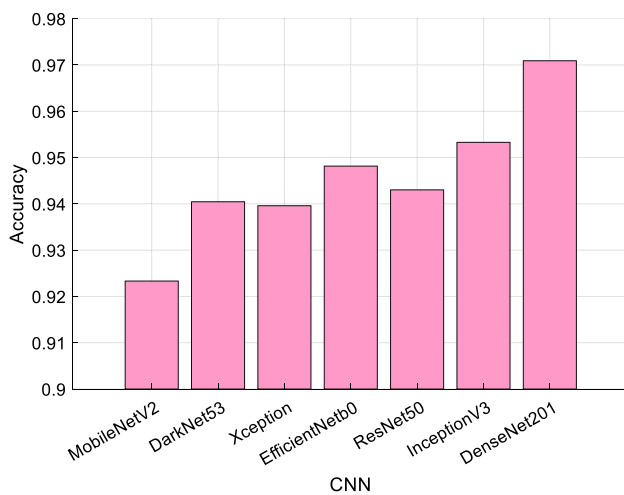

Fig. 4 Comparative analysis of pre-trained CNNs for the NFS architecture using Case 2

was highly likely to be correct, considering the significantly high precision rate.

5 Discussion

NFSDense201 introduces a novel approach to feature generation by leveraging the last fc1000 and global avg_pool layers of the pre-trained DenseNet201. This model aims to extract deep features from non-fixed size patches of incremental dimensions obtained through a unique nested patch division technique applied to resized input SEM images. By incorporating an identical patch that matches the input image, the model efficiently generates comprehensive global and local features. To address feature redundancy resulting from overlapping nested patches centered around the image, we employ INCA feature selection. Through this process, we identify an optimal feature vector length of 698 for the study dataset. For classification, we utilize a standard shallow cubic SVM with a 75:25 split. Our NFSDense201 model achieves remarkable accuracy rates of 99.53% for Case 1 and 97.09% for Case 2.

In conducting a nonsystematic review of existing models for SEM image classification (see Table 2), we find that the NFSDense201 model outperforms its counterparts. Notably, our model utilizes the largest SEM image dataset to date. We specifically compare our results to those of Kavuran et al. [29] who employed transfer learning, feature reduction with metaheuristic optimization, and an SVM classifier for a four-class Case 1 classification task using the same dataset. Interestingly, their optimized model achieved a similar classification accuracy of 99.30% [29], despite our model not utilizing any optimization method. In a separate study, Li et al. [30] proposed deep and machine learning models for classifying minerals in microscopic images into 13 categories. While their end-to-end deep learning model achieved the highest F1 score of 92% [30], it incurred substantial time complexity. In contrast, our NFSDense201 model, based on transfer learning, achieves excellent classification results with minimal time costs. Furthermore, Leracitano et al. [31] obtained an accuracy of

92.50% using a multilevel perceptron-based image classification model, albeit with a relatively small dataset. Tsutsui et al. [32] employed a gray-level co-occurrence matrix to extract textural features from SEM images and achieved 85% accuracy using two shallow classifiers on a limited dataset. Similarly, Tian et al. [33] reported an accuracy of 88% with their pre-trained VGG16-based model on a small SEM image dataset. Lastly, Yin et al. [34] proposed an attention-convolutional neural network model, attaining an impressive accuracy of 98.56% on a sizable four-class dataset.

Within the existing literature, we found no substantial model that specifically addresses Case 2, as far as our knowledge extends. As a consequence, we are unable to provide any comparative findings pertaining to the aforementioned dataset containing 10 distinct classes. To address this gap, we designed a comparative scenario to assess the performance of our proposed model against other prevalent CNN architectures, including (i) MobileNetV2, (ii) DarkNet53, (iii) Xception, (iv) EfficientNetb0, (v) ResNet50, and (vi) InceptionV3. Employing our NFS-based deep feature extraction architecture, we applied it to these pre-trained CNNs and obtained results using Case 2, involving a substantial image dataset.

To illustrate the outcomes, both the calculated results and the performance of our proposed NFDenseNet201 are depicted in Fig. 4.

Figure 4 demonstrates that the optimal deep feature extraction model for our proposed deep feature engineering architecture is DenseNet201, achieving an impressive accuracy of 97.09%. Following closely, InceptionV3 attains a respectable accuracy of 95.33%. Conversely, among the pre-trained CNNs, MobileNetV2 exhibits the lowest performance for this particular problem, with an accuracy of 92.33%.

Key characteristics of the NFSDense201 model are outlined as follows:

- We introduce a novel nested patch division method, which enables effective feature extraction from SEM images.
- This method is combined with downstream deep feature extraction using a pre-trained deep network, resulting in a novel deep feature engineering model.
- The model is trained and evaluated on an extensive dataset comprising 21,272 SEM images.
- Remarkably, NFSDense201 achieves classification accuracy rates of 97.09% and 99.53% for ten- and four-class SEM image classification tasks, respectively, using a standard shallow SVM classifier without any optimization. These results compare favorably with existing literature.

- NFSDense201 demonstrates computational efficiency, making it highly practical for implementation.
- The presented architecture can be readily adapted to address various classification problems.

Nevertheless, certain limitations should be acknowledged. In this study, we employ a cubic SVM classifier without hyperparameter tuning, potentially hindering the model's performance. Exploring optimization methods may yield improved classification results. Additionally, alternative advanced classifiers could have been examined. However, the primary focus of this work is to showcase the discriminative capabilities of the features generated by the main upstream model components, namely the novel nested patch division and the pre-trained DenseNet201. Given this objective, a robust yet shallow classifier like SVM adequately serves our present research goals.

6 Conclusions

In this study, we have successfully demonstrated the feasibility and practicality of the NFSDense201 model for accurate SEM image classification. By integrating the innovative nested patch division technique and efficient deep feature extraction using the pre-trained DenseNet201, we have achieved outstanding classification results. Leveraging the INCA feature selector in combination with a cubic SVM classifier, our model achieved remarkable accuracy rates of 97.09% and 99.53% for ten- and four-class classification tasks, respectively. Notably, these results were obtained using the largest publicly available SEM image dataset.

Our proposed model exhibits several desirable qualities. Firstly, it is computationally lightweight, enabling efficient processing of images. Secondly, its implementation is straightforward, ensuring ease of use for researchers and practitioners. Moreover, we believe that the NFSDense201 model holds promise beyond SEM image classification and can be readily applied to other computer vision tasks with minimal modifications.

As future research directions, we suggest exploring alternative methods and networks to enhance the nested patch division and/or replace the pre-trained DenseNet201 model. By integrating these new components, we can develop next-generation feature engineering models tailored to diverse image classification applications. This avenue of investigation has the potential to further improve the performance and versatility of our approach.

Declarations

References

1. Adams FC, Barbante C (2013) Nanoscience, nanotechnology and spectrometry. Spectrochim Acta Part B 86:3–13
2. Nasrollahzadeh M, Sajadi SM, Sajjadi M, Issaabadi Z (2019) An introduction to nanotechnology. Interface science and technology, vol 28. Elsevier, Amsterdam, pp 1–27
3. Hanus MJ, Harris AT (2013) Nanotechnology innovations for the construction industry. Prog Mater Sci 58(7):1056–1102
4. Ul-Hamid A (2018) A beginners' guide to scanning electron microscopy, vol 1. Springer, Berlin
5. Rydz J, Šišková A, Andicsová Eckstein A (2019) Scanning electron microscopy and atomic force microscopy: Topographic and dynamical surface studies of blends, composites, and hybrid functional materials for sustainable future. Adv Mater Sci Eng. https://doi.org/10.1155/2019/6871785
6. Aversa R, Modarres MH, Cozzini S, Ciancio R, Chiusole A (2018) The first annotated set of scanning electron microscopy images for nanoscience. Sci Data 5(1):1–10
7. NFFA-EUROPE (2016) Draft metadata standard for nanoscience data. NFFA project deliverable D11.2, http://www.nffa.eu/media/124786/d112-draft-metadata-standard-for-nanoscience-data_20160225-v1.pdf
8. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J-W, da Silva Santos LB, Bourne PE (2016) The FAIR guiding principles for scientific data management and stewardship. Sci Data 3(1):1–9
9. Modarres MH, Aversa R, Cozzini S, Ciancio R, Leto A, Brandino GP (2017) Neural network for nanoscience scanning electron microscope image recognition. Sci Rep 7(1):1–12
10. Li C, Wang D, Kong L (2021) Application of machine learning techniques in mineral classification for scanning electron microscopy-energy dispersive X-ray spectroscopy (SEM-EDS) images. J Petrol Sci Eng 200:108178
11. Kharin AY (2020) Deep learning for scanning electron microscopy: Synthetic data for the nanoparticles detection. Ultramicroscopy 219:113125
12. Ge M, Su F, Zhao Z, Su D (2020) Deep learning analysis on microscopic imaging in materials science. Materials Today Nano 11:100087
13. Krizhevsky A, Sutskever I, Hinton GE (2017) Imagenet classification with deep convolutional neural networks. Commun ACM 60(6):84–90
14. Redmon J, Farhadi A (2017) YOLO9000: better, faster, stronger. In: proceedings of the IEEE conference on computer vision and pattern recognition, pp 7263–7271
15. DeCost BL, Holm EA (2015) A computer vision approach for automated analysis and classification of microstructural image data. Comput Mater Sci 110:126–133
16. de Albuquerque VHC, Cortez PC, de Alexandria AR, Tavares JMR (2008) A new solution for automatic microstructures analysis from images based on a backpropagation artificial neural network. Nondestruct Test Eval 23(4):273–283
17. Osenberg M, Hilger A, Neumann M, Wagner A, Bohn N, Binder JR, Schmidt V, Banhart J, Manke I (2023) Classification of FIB/SEM-tomography images for highly porous multiphase materials using random forest classifiers. J Power Sources 570:233030
18. Han Y, Liu Y, Chen Q (2023) Data augmentation in material images using the improved HP-VAE-GAN. Comput Mater Sci 226:112250. https://doi.org/10.1016/j.commatsci.2023.112250
19. Dahy G, Soliman MM, Alshater H, Slowik A, Hassanien AE (2023) Optimized deep networks for the classification of nanoparticles in scanning electron microscopy imaging. Comput Mater Sci 223:112135
20. Scott-Fordsmand JJ, Amorim MJ (2023) Using Machine Learning to make nanomaterials sustainable. Sci Total Environ 859:160303
21. Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, Dehghani M, Minderer M, Heigold G, Gelly S (2020) An image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint https://arxiv.org/abs/2010.11929
22. Tolstikhin I, Houlsby N, Kolesnikov A, Beyer L, Zhai X, Unterthiner T, Yung J, Keysers D, Uszkoreit J, Lucic M (2021) MLP-Mixer: An all-MLP Architecture for Vision. arXiv preprint https://arxiv.org/abs/2105.01601
23. Huang G, Liu Z, Van Der Maaten L (2017) Weinberger KQ Densely connected convolutional networks. In: proceedings of the IEEE conference on computer vision and pattern recognition. pp 4700–4708
24. Tuncer T, Dogan S, Özyurt F, Belhaouari SB, Bensmail H (2020) Novel multi center and threshold ternary pattern based method for disease detection method using voice. IEEE Access 8:84532–84540
25. Vapnik V (1998) The support vector method of function estimation. Nonlinear Modeling. Springer, Berlin, pp 55–85
26. Vapnik V (2013) The nature of statistical learning theory. Springer, Berlin
27. Powers DM (2020) Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. arXiv preprint https://arxiv.org/abs/2010.16061
28. Warrens MJ (2008) On the equivalence of Cohen's kappa and the hubert-arabie adjusted rand index. J Classif 25(2):177–183
29. Kavuran G (2021) SEM-Net: deep features selections with binary particle swarm optimization Method for classification of scanning electron microscope images. Mater Today Commun 27:102198
30. Zhang J, Liu M, Xiong P, Du H, Zhang H, Lin F, Hou Z, Liu X (2021) A multi-dimensional association information analysis approach to automated detection and localization of myocardial infarction. Eng Appl Artif Intell 97:104092

31. Ieracitano C, Paviglianiti A, Campolo M, Hussain A, Pasero E, Morabito FC (2020) A novel automatic classification system based on hybrid unsupervised and supervised machine learning for electrospun nanofibers. IEEE/CAA J Autom Sinica 8(1):64–76

32. Tsutsui K, Terasaki H, Uto K, Maemura T, Hiramatsu S, Hayashi K, Moriguchi K, Morito S (2020) A methodology of steel microstructure recognition using SEM images by machine learning based on textural analysis. Mater Today Commun 25:101514

33. Tian X, Daigle H, Jiang H (2018) Feature detection for digital images using machine learning algorithms and image processing. In: SPE/AAPG/SEG Unconventional Resources Technology Conference, OnePetro

34. Yin C, Cheng X, Liu X, Zhao M (2020) Identification and classification of atmospheric particles based on SEM images using convolutional neural network with attention mechanism. Complexity. https://doi.org/10.1155/2020/9673724

Authors and Affiliations

Prabal Datta Barua[1,2,3,4,5,6,7,8,9] ⬤ · Sengul Dogan[10] · Gurkan Kavuran[11] · Turker Tuncer[10] · Ru-San Tan[12,13] · U. Rajendra Acharya[14]

✉ Prabal Datta Barua
  Prabal.Barua@usq.edu.au

  Sengul Dogan
  sdogan@firat.edu.tr

  Gurkan Kavuran
  gurkan.kavuran@ozal.edu.tr

  Turker Tuncer
  turkertuncer@firat.edu.tr

  Ru-San Tan
  tanrsnhc@gmail.com

  U. Rajendra Acharya
  Rajendra.Acharya@usq.edu.au

[1] Cogninet Australia, Sydney, NSW 2010, Australia

[2] School of Business (Information System), University of Southern Queensland, Darling Heights, Australia

[3] Faculty of Engineering and Information Technology, University of Technology Sydney, Sydney, NSW 2007, Australia

[4] Australian International Institute of Higher Education, Sydney, NSW 2000, Australia

[5] School of Science & Technology, University of New England, Armidale, Australia

[6] School of Biosciences, Taylor's University, Subang Jaya, Malaysia

[7] School of Computing, SRM Institute of Science and Technology, Chennai, India

[8] School of Science and Technology, Kumamoto University, Kumamoto, Japan

[9] Sydney School of Education and Social Work, University of Sydney, Camperdown, Australia

[10] Department of Digital Forensics Engineering, Technology Faculty, Firat University, Elazig, Turkey

[11] Department of Electrical and Electronics Engineering, Faculty of Engineering and Natural Sciences, Malatya Turgut Ozal University, Malatya, Turkey

[12] Department of Cardiology, National Heart Centre Singapore, Singapore, Singapore

[13] Duke-NUS Medical School, Singapore, Singapore

[14] School of Mathematics, Physics and Computing, University of Southern Queensland, Springfield, Australia