Computer Vision Analysis on Material Characterization Images

Danpeng Cheng, Wuxin Sha, Zuo Xu,* Lixin Huang, Yunpeng Du, Shun Tang, Yaqing Guo, Yuan-Cheng Cao,* and Shijie Cheng

Material characterization has been proved to be the most intuitive approach to understand the chemical composition, structure, and microstructure of materials, which is the basis of material design. One of the most important steps in material design is to extract the characteristics from an image, and find their associations with the material structure and properties. Therefore, in recent years, with the rapid development of machine vision algorithms, characterization images have attracted attention in the field of material characterization. Researchers use computer vision algorithms, such as image denoising and enhancement, to preprocess the representation image, image segmentation and classification to detect and separate each microstructure from the characterization image, and quantitatively analyze the properties of materials. Herein, the application of computer vision algorithms in material image representation is summarized and discussed. The latest and valuable views for experts and scholars in both computer vision and material grounds are presented. Thus, this review provides guidance for material exploration and promotes the developments of artificial intelligence in the field of materials.

1. Introduction

The composition, structure, and morphology that determine the physical and chemical properties of materials are usually tested by a variety of characterization techniques. Common characterization techniques in the field of microscopic imaging include optical microscope (OM), scanning electron microscope (SEM), transmission electron microscope (TEM), scanning transmission electron microscope (STEM), and atomic force microscope (AFM). OM takes advantage of visible-light imaging, whose resolution is at the micrometer level and sample preparation is fast and simple. The electron microscopes use shorter wavelength electron beam imaging, and its resolution can reach the nanometer level. Among them, SEM fine electron beam imaging can obtain the surface morphology and composition information of samples. The high-energy electron beam of TEM can even get the crystal structure and defects through the sample.

In recent years, the latest development of high-end imaging characterization technologies, such as AFM and STEM, makes it possible to observe the microscopic evolution of materials with atomic spatial resolution and subsecond time resolution, which makes it possible to directly observe the atomic-level dynamic phenomena, including defect evolution, dislocation migration, phase transformation, etc. And accumulated a lot of data. Although the ability of high spatial-temporal resolution material data acquisition is constantly improving, the information about process dynamics and thermodynamics inferred from these high-end characterization images is very little. The inherent limitations of manual analysis in volume and speed hinder the deep utilization of high-end characterization technology. These data are simply selected for qualitative research, pure manual analysis is difficult to obtain all the information in the image quickly and accurately, massive data is discarded and wasted. For example, there may be tens of thousands of atoms in a STEM photograph, but only a few have been chosen to quantitatively study the composition and arrangement structure. Due to the existence of instrument noise and image artifacts related to atomic motion, it is difficult to analyze high-throughput and large-capacity microscopic data sets manually. First, it takes a lot of attention to manually determine the exact morphological characteristics and distribution of a single nano-material, identify all particles and track their trajectories to study the dynamics. Second, it is easy to miss crucial information by manual analysis. For instance, it is inefficient to identify overlapped or adjacent nanoparticles images or analyze STEM images with inhomogeneous electron beam transmission shadows.

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/aisy.202100158.

© 2021 The Authors. Advanced Intelligent Systems published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

DOI: 10.1002/aisy.202100158
It can be seen that the success of material science characterization requires not only exploratory research and instrument improvement but also timely and effective processing of high-throughput data generated by these instruments to infer the specific information such as dynamics and thermodynamics describing the microstructure of materials. The processed information is later served as feedback to the experimental design and property optimization. Therefore, it is urgent to develop a visual model for the material characterization image replacing the manual automatic analysis of the material characterization image.

The high-dimensional data processing capability of machine learning brings new opportunities for electron microscope images. Some researchers have constructed material characterization image data sets by collecting and labeling microscopic image data. The neural network model and dynamic statistical model based on deep learning (DL) are constructed to identify and locate the atom or lattice defects. For instance, automatically mark the lattice spacing, classify and count the real morphology of microparticles, quantitatively analyze the tissue dynamics of materials, and realize the automatic high-throughput analysis of material characterization images. Therefore, the material microscopic imaging analysis technology based on machine vision has become a research hotspot in the field of characterization.

Therefore, this article summarizes the recent developments of visual algorithms and analysis methods used in the analysis of material characterization images in recent years, sorts out the automatic analysis technologies of various microcharacterization images at different scales, confirms the challenges existing in the current visual models of material characterization and points out the future development direction of material microimage processing technology. The next part will introduce some commonly used algorithms and models in the field of material image processing.

2. Key Technologies and Latest Hotspots in Computer Vision

Computer vision enjoys the longest research history and the most technology accumulation in the field of artificial intelligence. This section discusses the key technologies and the latest research hotspots in the machine vision area from the three levels of image processing, image analysis, and image understanding shown in Figure 1. In particular, we summarize and make a comparison of the most commonly used vision models: convolutional neural network (CNN), fully convolutional network (FCN), U-net, generative adversarial network (GAN), and variational autoencoder (VAE) in Table 1 to establish the relationship with the processing.

![Image 1](https://www.advancedsciencenews.com)

**Figure 1.** Three layers in computer vision. a) Image preprocessing gets data prepared for models to deal with. b) Image analysis aims at extracting features for target tasks. c) Image understanding bridges the gap between image content and linguistic meaning.
of material characterization pictures, hoping to provide technic for material characterization and find new oriented researches.

2.1. Image Preprocessing

As the basis of computer vision, image processing mainly preprocesses image information, such as denoising and enhancement, to improve image quality, and data augmentation aims at enlarging amount of data. Both denoising and augmentation might be considered before training models to get wanted results.

Due to electromagnetic interference, lens jitter and high-speed motion of objects in the scene, the images often contain salt and pepper, Gaussian noise, low contrast, and blurring problems. Therefore, it is necessary to denoise and enhance the image[14–16] to improve the image quality and reduce the difficulty of subsequent image analysis and understanding (Figure 1a). Recently, with the launch of smartphones for night scene shooting, low illumination image denoising[17–19] has become hot spots, targeting dark images.

In addition, data augmentation is considered to alleviate the problems caused by the lack of data. One powerful network for data augmentation is called general modeling (GAN). Most researchers modify the GAN framework through different network architectures, loss functions and evolutionary methods.[20–22] Models have been well developed as approaching users’ demand for one-click make-up and style change from apps, such as TIK TOK and Photoshop.

2.2. Image Analysis

The purpose of image analysis (Figure 1b), the requisite of image understanding, is to make the machine or computer automatically analyze the underlying features and upper structure. One of the key technologies of image analysis is image segmentation, which refers to the extraction of meaningful features in the image, including the edge and region of the object in the image. The computer vision (CV) algorithm classifies the semantics of each region, always achieving more accurate and efficient segment results.[23–27] The pioneering progress as the attention mechanism,[28–32] enabling the machine to selectively process features. Others focus on the real image characteristics, including exposure, contrast, illumination, object shape, and surface texture,[33,34] to solve the segmentation error caused by the differences between synthetic data and real-world images.

2.3. Image Understanding

Image understanding based on image analysis, further studies not only target objects but the relationship between them, for carrying on the natural language expression with practical significance (Figure 1c). Depending on neural networks, image understanding undertakes high-level researches on image combining vision, natural language, and other types of signals, to link disparate pixel areas with linguistic meaning, so that the computer system can automatically understand the semantic information in the image. At present, the research of image understanding mainly focuses on object detection and 3D reconstruction.

Object detection finds out all the interests in the image, generally including location and classification of targets. The mainstream object detection technology is generally divided into two categories: Two-stage detection, first set up candidate regions, containing the general location information of the target, and then classify and finetune the candidate regions.[27,35,36] Another one is one-stage detection, directly generating the probability and coordinate position of object categories.[16–19]
3D reconstruction, in addition, helps the computer understand the global environmental information and achieve real environmental perception. 3D reconstruction technology uses the relationship between image coordinate system and world coordinate system to reconstruct 3D information with the information in multiple 2D images, to obtain the 3D information of objects in the environment. According to the different forms of data processing, it can be divided into voxel-based, point cloud-based, and mesh-based methods.

2.4. Computer Vision Models for Microscopy

Convolution is the essential operation of machine vision models. In essence, each image can be represented as a matrix of pixel values, in which the channel represents some composition of the image, such as the red, green, and blue channels of the image. The convolution operation uses convolutional kernels (matrices) of different sizes to slide on an image and calculate the product of corresponding elements, thus extracting image features from the image and preserving the relationship between pixels. The weights of each convolutional kernel will be updated continuously through training to find the most consistent input–output mapping.

CNN is one of the most representative algorithms of computer vision for its excellent ability to extract information. CNN is built upon a series of convolutional, nonlinear, converging (downsampling), and fully connected layers. An image goes through convolutional and nonlocal layers to obtain feature maps, and downsampling could resize the image meeting different requirements, such as classification and segmentation. The downsampled feature maps finally will be fed to fully connected layers to get the probability of a class or the class that best describes the image.

FCN and U-net both extending from CNN, adjust their structures to balance multiple scales characteristics. FCN replaces the last fast fully connected layers of a basic network into fully convolutional layers. Thus, inputs and outputs of the network can be any size images. At the same time, the structure of jump connection is used to combine the coarse graded rough information with the fine-grained information to produce accurate and fine segmentation. But the segmentation result is not fine enough. Another enchanting network, U-net, links completely symmetric encoder and decoder layers by skip connection. A large number of channels and skip connections are added, compared with FCN, which makes the network spread the context information to a higher resolution and prevents the loss of input information when reducing the image’s resolution. They achieve great success in analyzing and segmenting material images.

Different from the previous cases, generative models, mostly consist of two networks, responsible for producing images different from the original data, such as GAN and VAE. GAN is composed of a generator and discriminator. Generator, which is used to generate images similar to real images in an attempt to cheat discriminator. While, discriminator, as a classifier, is trained to distinguish the generated image from the real image. During training, gradient descent is used to optimize D and G alternately. Through the confrontation between the two, when the generation model can restore the distribution of the training data set, and the discriminator cannot distinguish the original image from the real image, the network can generate a large number of images that are enough to confuse the real image with the false image. Another model, VAE, uses two neural networks to establish two probability distribution models: inference network infers the potential features of the original input image and generates the variational probability distribution of hidden variables, which is called potential space. According to the generated variational probability distribution of hidden variables, the generated network samples images from potential space and restores the approximate probability distribution of the original data.

Models, with pool generalization ability, may suffer from underfitting or overfitting problems. Underfitting models do not capture the data characteristics well, limited by simple model structure. These problems can be solved by introducing feature items and expert knowledge such as enhancing certain weights. Overfitting, another annoying issue, cannot successfully express other data except training one due to the restricted amount of datasets and intricate models. Data problems are commonly shown in small data sets, where the model does not have enough data for training to grasp the distribution and law of real data. Under such circumstances, data augmentation becomes a valuable solution, as discussed in Section 2.1. Analogously, for the model problem, we need some tips to adjust the model, so that the model can find the mapping not only for the training set. The regularization of weights and random dropout of some neurons can make all weights not indispensable, which increases the difficulty of model training. In addition, batch standardization makes inputs of the next layer close to Gaussian distribution to avoid biases caused by the current input. The aforementioned methods can adjust the generalization ability of the model to a certain extent.

In addition, there are many differences between the microscopic images and traditional visual images, such as noise type, number of channels, and extracted information. Better applying these models to the field of materials requires us to understand the imaging principle and image features in the field of materials. On the one hand, domain features served as a priori knowledge, another annoying issue, cannot successfully express other data except training one due to the restricted amount of datasets and intricate models. Data problems are commonly shown in small data sets, where the model does not have enough data for training to grasp the distribution and law of real data. Under such circumstances, data augmentation becomes a valuable solution, as discussed in Section 2.1. Analogously, for the model problem, we need some tips to adjust the model, so that the model can find the mapping not only for the training set. The regularization of weights and random dropout of some neurons can make all weights not indispensable, which increases the difficulty of model training. In addition, batch standardization makes inputs of the next layer close to Gaussian distribution to avoid biases caused by the current input. The aforementioned methods can adjust the generalization ability of the model to a certain extent.

3. Imaging Instruments and Vision-Based Framework

The internal components, structure, and morphology that determine the physical and chemical properties of materials are usually tested by a variety of characterization techniques. With the latest development of microscopy technologies, human beings observe atomic spatial resolution and subsecond temporal resolution of the structure of materials. Here, we review the latest publications on the applications of OM, AFM, TEM, and scanning transmission electron microscopy (STEM) in the field of material imaging. Understanding the principle of microscopic...
imaging is helpful for us to deal with the information of image extraction.

3.1. Imaging Instruments

The optical system of OM uses the visible light and lens coefficient to magnify and image tiny objects. The object passes through the objective lens to form a magnified real image, and then passes through the eyepiece to form a magnified virtual image. OM observe selected cross sections of transparent materials without slicing the samples. For example, in the field of biology, biological activities can be observed by tracking the fluorescence of specific atomic or molecular markers. It can be real-time and dynamic observation, and occupies a dominant position in the field of biology. However, the diffraction limit of an OM is limited to 1000× amplification and 200 nm resolution.

AFM composed of a tip sensitive to the weak force, scans the sample due to the interaction between the atoms on the surface of the sample and the atoms on the tip of the probe. Therefore, the position changes of each scanning point of the microcantilever can be measured, and the surface morphology information of the sample with nanometer resolution can be obtained. AFM has three obvious advantages. First, AFM can provide a real 3D surface profile. AFM can image almost any object surface, providing qualitative and quantitative information of physical properties and statistical information. Second, AFM does not need to pretreat the sample, such as plating conductive film, to prevent irreversible damage. Third, it can work under normal pressure or even in a liquid environment, which provides opportunities to study biological macromolecules or living biological tissues. However, AFM is sensitive to the probe, resulting in slow imaging speed and small range.

A transmission electron microscope (TEM) uses a very short wavelength electron beam as the illumination source and uses an electromagnetic perspective mirror to focus the image, which is used to analyze micro- to nanolevel samples. TEM transmits the electron beam which has been accelerated and focused to the very thin sample and collides with the atoms in the sample, resulting in scattering. The brightness of the generated image is related to the atomic number, crystal structure, electron density, and thickness of the sample. TEM has a strong scattering ability, so it can realize the processing of micro- and nanoareas and the study of structural composition. TEM has high resolution and can directly observe the image of heavy metal atoms. Moreover, the bright-field and dark-field images are beneficial to the analysis of structural defects, and the phase information can be deduced. TEM mainly suffers from the high precision and price of the instrument, the complex manufacturing process of the sample and the need to image in a vacuum.

SEM uses physical signals, such as secondary electron and backscattered electron imaging, to observe the surface morphology and composition of the sample and the structure of the cleavage surface. SEM sample is easy to make without slicing and can be rotated in 3D space, which is conducive to multangle observation. However, due to the charge effect and the irregular deflection of the electron beam caused by the electrostatic field, the uneven brightness and darkness of the image, image deformity and image drift will be caused. In addition, the irregular discharge of charged samples may cause bright spots and lines in the image.

Scanning transmission electron microscope (STEM) is to configure the transmission accessory on the scanning mirror, so that it has both scanning function and transmission function, and can obtain the internal structure of the material. Compared with TEM, STEM has a lower accelerating voltage, which can significantly reduce the damage of electron beam to the sample and improve the image contrast. It is suitable for the microstructure characterization of organic polymer, biological, and other soft materials. Second, STEM can generate scanning secondary electron image and transmission image simultaneously, which can obtain the surface morphology and internal structure information at the same position. STEM technology is highly demanding and requires an extremely harsh vacuum environment.

Although these tools reflect different material properties, such as scale, morphology, structure, and composition, due to different imaging principles, they all have some common challenges. First of all, it is difficult to obtain material-specific information from images quickly. Speed and comprehensiveness of information are crucial when high-throughput datasets accumulate. The second challenge is the accuracy of information extraction, limited by noise, sample preparation technology and inevitable damage to instruments. The third is impossible to deal with the multidimensional problem, such as particles overlapping, only by manpower. Fourth, but not the last, is how to reconstruct new materials and predict their performance. Computer vision algorithms, however, shed light on those challenges. We hope to find a universal process by ideals in vision ground that can extract information from any micrograph, even though the materials are in different physical systems.

3.2. Vision-Based Framework

With the advantage of vision algorithms, we develop a DL framework for material characterization images, which can be used to realize microimaging research in material science based on machine vision. It consists of five parts: task analysis, data preparation, model design, and result analysis (Figure 2).

Task analysis needs to divide the task into classification or regression catalogs, and data collected from humans or simulation should be prepared before being fed into the model, for example, denoising, labeling, and splitting into training/testing/validation datasets. The third step is to design the model according to the task-type and data characteristics. Model design is of importance in the whole task considering functions and predicted results of the goal. For instance, FCN, U-net, and visual algorithms, however, shed light on those challenges. We hope to find a universal process by ideals in vision ground that can extract information from any micrograph, even though the materials are in different physical systems.
4. Computer Vision in Microscopy

Computer vision framework provides an efficient and accurate automatic means for microscopic image characterization of materials, which has been confirmed in many fields. Algorithms have been used in the production of microscope simulation datasets, defect detection, morphological feature analysis, composition research, and material design. Here, we point out the specific problems in each material characterization task accompanying with key solutions of vision-based networks suitable for target material characterization technologies in recent years.

4.1. Production of Simulation Dataset

In the traditional computational material design task, the cost of collecting real data is very high. We may only have limited samples of material microstructure and properties, such as due to stem, TEM and other means of electron microscopy characterization, expensive price, and complex sample preparation process. In addition, simulation calculation, such as DFT calculation, involves a variety of approximations and exchange-related functions, and its results need to be verified by other tests. For large-scale systems with multiple time scales and disordered structures, the calculation will be too intensive, and the calculation cost is usually too high. However, in recent years, any number of material samples can be generated by combining machine vision with the fabrication of material simulation data sets, with negligible computational cost. As a branch of machine vision networks, the generation of GAN and VAE has great potential. The algorithm easily generates a large number of simulation data, while maintaining reasonable accuracy to study the shape distribution of real samples.

For example, Ma et al. Generated a large number of section optical images of polycrystal iron based on the GAN model for grain segmentation. The generator is composed of a U-net network with encoding and decoding, which is used to capture the small noise of real data in the actual test. Then, the image transmission model mixes the realistic features into the simulated image, and finally obtains the composite image. Liu et al. used the DCGAN network to generate a constant temperature diagram for nondestructive testing (Figure 3a). The network size of DCGAN can be adjusted according to the original constant temperature map, and CNN is used to replace the multilayer sensing in the original GAN. DCGAN network learns abundant information from thermal images, and the generated high-dimensional images make up for the shortage of thermal images in pulsed thermal imaging and eliminate the noise to a certain extent. In fact, VAE also has a similar function served as a generator. Cang et al. used an unsupervised method based on VAE to extract hidden features from heterogeneous alloy materials and generate morphological constraints of images (Figure 3b). Compared with the Markov random field model, the material properties of the simulated samples match the real samples better. Mamun et al. used the generation model of VAE to create synthetic alloy samples, helped models to make a reliable prediction of creep life, and assisted alloy reverse design.

4.2. Grain and Texture Segmentation

The application of the image segmentation algorithm to microscope images shows great potential, which greatly improves the grain boundary segmentation, particle detection, and counting technology, and is conducive to the recognition of material microstructure and morphological research. Although the traditional threshold segmentation can distinguish the foreground and background of a large number of images, the high resolution of micrographs brings complex data structure, pollutants generated in the process of sample preparation, and the variability of the microstructure of material samples all bring great challenges to the image segmentation task in micrographs. Thanks to the computer vision task, the segmentation problem is transformed into a classification problem, which can alleviate the aforementioned problems that cannot be solved by traditional methods.
Many more frequently, one of the slices of a microscopic image can have a resolution of \(3200 \times 3200\), which may be too large for limited GPU training. Most articles use pooling to compress the image size. Pooling is a nonlinear downsampling, and it is also the most common structure in neural networks. Pooling uses the maximum or average value for each subregion of the input data, so as to reduce the resolution and retain the most important information of the image. In addition, some researchers cut the photos, and expanded the segmentation data set while solving the problem of too large pictures. Maksov et al. only used the first frame of film to complete the network training, and segmented the defects in Mo mixed WS2 image (Figure 4a).

Experts manually mark the features in the image that need to be segmented, and feed neural network training, which can greatly avoid the segmentation trouble caused by the sample pollutants, and is suitable for any material. The most commonly used networks are FCN and U-net based on encoder–decoders. FCN is used to analyze the distribution of carbon fibers in cement-based composite improving the electrical properties of CFRC. Azimi et al. added a max voting scheme to FCN for the classification of low carbon steel SEM images (Figure 4b). U-net, also has been successfully used in the field of materials, such as atom segmentation and texture segmentation. Chen et al. designed a self-tuning semi-supervised framework with pseudolabel prediction based on U-net. Using both labeled and unlabeled metal image training models, good microstructure recognition results can be obtained with a small number of labeled images. Similar work has also been used in defect detection of carbon fiber-reinforced plastics and composition analysis of metallographic images.

High-quality image tagging takes a lot of time and manpower. For segmentation tasks, which is why researchers try to avoid human tagging. For example, Mei et al. used an unsupervised learning method to train defect-free samples to realize automatic segmentation of texture surface defects (Figure 4c). zhao et al. also trained only defect-free samples and used GAN network to reconstruct defect images.

Then, the features of the reconstructed image and the original image are extracted, and the part with large feature difference can accurately locate the defect.
location. Maksov et al. used the periodicity of the atomic arrangement and Fourier transforms to automatically find out the defect location, which was used as the ground truth of network training.\cite{13}

4.3. Particle Tracking

Machine vision algorithm provides an important point of view for particle tracking technology. Because the composition and motion of particles have a great influence on the function of particles, fast and accurate particle tracking technology is very important. Traditional particle tracking technology is very difficult. Due to the strict signal-to-noise ratio, fixed particle size and number, and complex diffusion algorithm, the risk of worse tracking results is caused. In this region, machine vision algorithms combined with the use of simulated data sets can effectively overcome these limitations. Ziatdinov et al. proved that image processing can denoise the image (Figure 5a).\cite{66} They use the convolutional neural network based on U-net architecture to avoid regional pollution and accurately identify the type and location of atoms. After that, by studying the beam-induced translation of silicon atoms at the edge of graphene and the weight of graphene, the comprehensive analysis of the electron beam-induced reversible process is completed.
Later, Ziatdinov further combined DL with hybrid modeling, successfully transformed the time-dependent coordinates and directions of particle surface in AFM photos with instrument noise and artifacts, and studied the dynamic process of protein self-assembly on inorganic surfaces and the patterns formed.

DL-based algorithm greatly improves the positioning accuracy.

Helgadottir et al. realized the accurate tracking of multiparticle and nonspherical bodies under the condition of the unstable light source (Figure 5b). The network is composed of three convolution layers and two dense layers. By introducing the radial distance between the particle and the image center into networks with convolution and dense layers, improving the accuracy of particle tracking.

Atomic tracking is also used in 3D space. For example, Newby et al. realized particle tracking in 3D space by simulating how...
objects moving in 3D space map in 2D space.\textsuperscript{[70]} For the problem of overlapping particles in space, Franchini et al. used an encoder–decoder network to decode the depth information of particles (Figure 5c).\textsuperscript{[71]} In this article, a semisynthetic data set is created, which connects the particle image with its 3D position, including all the subtle differences of particle shape and the spherical distortion that may occur in the process of moving. Through training, even if two particles share the same center position, they can still be detected. At the same time, it is mentioned that the range of detectable depth of particle image can be increased by 67\%, compared with the traditional threshold method.

Moreover, DL is also used to quantify Brownian motion characteristics of nanoparticles in surface plasma resonance microscopy images,\textsuperscript{[72]} track small and dense particles,\textsuperscript{[73]} track a single particle in liquid-cell transmission electron microscopy (LCTEM),\textsuperscript{[74]} and predict particle motion.\textsuperscript{[75,76]}

### 4.4. Structural Reconstruction

Reconstruction, as a supplement of representation means, provides morphological features at different depth scales and contains quantitative structural and functional information of spatial distribution. In most material systems, the corresponding microstructure reflects a certain degree of randomness, such as particle size distribution, number density or surface area. However, the characteristics must be statistical. The purpose of reconstruction is to generate new microstructures according to the statistical characteristics of the input microstructures, so as to increase the existing imaging data and even guide the design of future imaging experiments, such as determining the required imaging scale and resolution. Machine vision provides help for reconstruction from two aspects: speed and flexibility. The general steps are divided into 1) dimensionality reduction of complex microstructure images; 2) obtain the characteristic information of each microstructure to improve the accuracy of reconstruction; 3) end-to-end reconstruction.

One of the advantages of neural networks is dimension reduction, which can extract the hidden features of low-dimensional anisotropic micrographs from multiscale and high-dimensional 3D structures, and then reconstruct more accurate microstructure according to the extracted features, such as VAE network. Kim et al. used a VAE network to generate more continuous microstructure pictures based on 4000 microstructure pictures of dual-phase steels to explore the microstructure under the best mechanical properties (Figure 6a).\textsuperscript{[77]} The output of the VAE network is used to study simulation and explore the relationship between structure and attribute. Finally, Gaussian process regression is used to link potential blank points and ferrite particle size with mechanical energy. Girard et al. constructed a structure combining vector quantizing VAE with a histogram to classify pre-explored nuclear materials and fine-grained process parameters.\textsuperscript{[78]} The encoder quantizes the input 3D micrograph by histogram and encodes it into a 1D feature map, which is called a feature vector, while the decoder reconstructs the original image by the feature vector. The feature vectors obtained from the index histogram can provide a new idea for quantitative analysis of microstructure images.

Another major advantage of introducing a neural network into stochastic reconstruction lies in the speed of generating 3D structures. Among them, the commonly used generation model is GAN. GAN can use a small number of samples to generate realistic 3D structures of different scales, avoid large-scale acquisition processes with a high speed. Feng et al. established a bicycle GAN-based network framework, which maps a single 2D image to different 3D images of porous media with...
Bicycle GAN consists of three parts: generator G, discriminator D, and encoder E. G receives 2D slice images and random noise and generates corresponding 3D structure, D distinguishes the true and false of input 3D structure and E encodes the received 3D structure to the distribution. The training set consists of pairs of 2D slices and 3D structures, in which the 2D image is the bottom slice of the 3D structure. Once the model is trained and a new 2D image is an input, the conversion of the 3D structure can be completed on a common CPU, which only takes 1 s. compared with 10 h by classical method. Valsecchi et al. also developed a 2D–3D reconstruction method for porous sandstone (Figure 6b). The difference is that the GAN network distinguishes 2D images. Using the similar idea of microcomputed tomography, the algorithm randomly extracts a group of 2D sections from the 3D structure generated by the generator, inputs the discriminator, and evaluates the quality of the reconstructed structure by checking the cross section of the reconstructed structure, so as to train the network. The network generated 3D structure is also extremely fast (20–30 ms).

In addition, the commonly used method in machine vision, transfer learning, solves the problem of insufficient data of specific material microstructure training set in reconstruction. Transfer learning takes a task as a pretraining model and applies the learned knowledge and experience to different but related problems. Since then, the model is no longer limited to some types of materials, but runs through the whole microstructure property system, showing great flexibility. Bostanabad uses depth network VGG-19 to pretrain 2D images to obtain structural features such as edges and particles (Figure 6c). With target features, the network transforms and optimizes the dimensions of initial stochastic 3D images in batches, so that the reduced features and 2D image features are in the same distribution interval. The 3D reconstruction of composite, alloy, porous, and polycrystalline microstructures is successfully realized. The pretrained VGG-19 network was also used in the research of Li et al. to find three statistically identical structures from the microstructure of any material. The training can reduce the difference between the target microstructure and the target microstructure, and generate the reconstruction microstructure similar to the marker material. Moreover, the author tries to prune the migration network and find out the influence of different models on network initialization.

5. Conclusion

In this review, we summarize the latest development of machine vision and its application in the imaging analysis of the microstructure of different materials. According to a large number of published articles, we propose the total process of microscopic image analysis based on machine vision, including task analysis, data processing, model design, feature analysis, and result verification, which extends the main problems of machine vision in the field of materials.

The main tasks of machine learning are focused on material structure extraction, dynamic analysis, and accelerating simulation calculation. The obvious advantages of machine vision are its ability to capture features accurately, strong generalization, and high automation. First, the microstructure image is directly used as the input without artificial filtering data, which makes CNN-based model hopeful to extract hidden information from massive pixels. Second, the method is not limited to a single physical or mathematical model, transfer learning makes different kinds of microstructure achieve linkage, which can be highly repetitive. Finally, the trained model can deal with massive multidimensional data easily and reduce a lot of labor. In addition, the input microstructure image can be encoded to obtain multichannel representation (multidimensional data), which is expected to separate multiple phases of materials.

In addition to using convolution layer, pooling layer and activation function stack, network structures such as graph neural network (GNN) and long short memory network (LSTM) are introduced in the field of computer vision. Semantic, time, and other information of different structures are added in the image feature extraction, so that the vision system can combine hearing, text and even taste to complete more accurate classification or regression tasks. This kind of method can also be used in the field of materials, such as the fusion of results obtained by different testing techniques. Different analysis and test technologies have different length scales and complementary information. Excessive dependence on a single analytical test often leads to a lack of objective understanding of material laws and conflicts with other types of test results. Therefore, analyzing the data from multiple sources at the same time can be closer to the essence of the internal law of materials, and form comprehensive and effective analysis results to feedback to the experimental design. The machine learning method has shown a broad prospect in the field of multicharacterization data fusion analysis, which enhances the relationship between imaging testing technology and material properties.

Some chemistry scholars worry that artificial intelligence algorithms, especially DL, cannot summarize what physical or mathematical objective laws are from the values of weights in a pile of networks. They believe that there are coincidences in the coupling process of certain data, and there will be various shifts between the test datasets and the actual data, but this does not mean that most DL algorithms will lead to such biased conclusions. On the contrary, the results summarized from a large number of data may be more practical to multiple real-world domains. Facts have proved that, as the amount of data accumulates, DL are beneficial to make accurate predictions on unknown inputs in chemistry such as prediction properties of materials, generating electronic noses, and evaluate DNA damage. Moreover, experts in various fields use DL as a tool, adding their domain knowledge for judgment could avoid lots of biased conclusions. In addition, some so-called deviations may bring researchers new lights to rethink some problems.

Using computer vision tools to solve material problems will become an inevitable trend in the future, which means that experts from artificial intelligence and material areas must work together. The prior knowledge provided by material experts can improve the accuracy of program operation results, and strive to find the intuitive mapping relationship between hidden features and material parameters with physical meaning, so as to enhance the interpretability of the model. Artificial intelligence experts can develop more friendly software/platform for material analysis, which extends the main problems of machine vision in the field of materials.
experts, aiming at a different microscope and experimental conditions, reduce the time-consuming caused by artificial statistics, and provide guidance for material exploration, to truly promote the development of the material field.

Acknowledgements

This work was supported by National Natural Science Foundation of China (52077096), and Dicastal-HUST Cooperation in Alloy Wheels Intelligent Manufacturing.

Conflict of Interest

The authors declare no conflict of interest.

Keywords

artificial intelligence, computer vision, guide imaging, machine learning, material characterization, microscopy

Received: August 17, 2021
Revised: September 23, 2021
Published online: October 27, 2021

[1] Y. J. Shin, W. Shin, T. Taniguchi, K. Watanabe, P. Kim, S.-H. Bae, 2d Mater. 2021, 8, 035017.
[2] G. Tao, X. Xu, R. S. Li, F. Liu, N. Li, AcS Sensors 2021, 6, 1321.
[3] H. Zhou, C. Qin, S. Han, L. Zhang, R. Chen, G. Zhang, Y. Liu, Z. Wu, S. Li, L. Xiao, D. Jia, Nano Lett. 2021, 21, 1477.
[4] H. Roghani, S. A. Tayebifar, K. Kasraee, M. S. Asl, Ceram. Int. 2020, 46, 28922.
[5] X. Zhou, Z. Jia, A. Feng, K. Wang, X. Liu, L. Chen, H. Cao, G. Wu, Compos. Commun. 2020, 21, 100404.
[6] T. A. Saleh, Chem. Eng. J. 2021, 404, 126987.
[7] Z. Liu, Z. Zhang, Z. Wang, B. Jin, D. Li, J. Tao, R. Tang, J. D. Youre, Proc. Natl. Acad. Sci. U. S. A. 2020, 117, 3397.
[8] M. R. Rosenberger, H.-J. Chuang, M. Phillips, V. P. Oleshko, K. M. McCreary, S. V. Sivaram, C. S. Hellberg, B. T. Jonker, AcS Nano 2020, 14, 14240.
[9] Y. T. Chen, T. Zhang, W. Zhang, D. H. Ping, K. Hono, A. Inoue, T. Sakurai, Mater. Trans. 2002, 43, 2647.
[10] J.-H. Tseng, C.-H. Weng, C.-C. Wang, M.-S. Ho, L.-T. Yen, J.-Y. Chen, G. Gaybullaev, C. Poongpadee, Y.-T. Lin, Chin. Eng. J. 2021, 421, 127880.
[11] Q. Sun, W. Peng, X. Hai, Appl. Surf. Sci. 2021, 562, 149976.
[12] Y. Jiang, Z. Chen, Y. Hang, P. H. Gao, S. Xie, P. Purohit, M. W. Tate, J. Park, S. M. Gruner, V. Elser, D. A. Muller, Nature 2018, 559, 343.
[13] A. Maksov, O. Dyck, K. Wang, K. Xiao, D. B. Geoghan, B. G. Sumpter, R. K. Vasudevan, S. Jesse, S. V. Kalinin, M. Ziatdinov, npj Comput. Mater. 2019, 5.
[14] P. Liu, H. Zhang, K. Zhang, L. Lin, W. Zuo, in Proc. 2018 IEEE/CVF Conf. on Computer Vision and Pattern Recognition Workshops, IEEE, Piscataway, NJ 2018, pp. 886–895.
[15] K. Zhang, W. Zuo, Y. Chen, D. Meng, L. Zhang, IEEE Trans. Image Process. 2017, 26, 3142.
[16] K. Zhang, W. Zuo, L. Zhang, IEEE Trans. Image Process. 2018, 27, 4608.

[17] C. Guo, C. Li, J. Guo, C. C. Loy, J. Hou, S. Kwong, R. Cong, in 2020 IEEE/CVF Conf. on Computer Vision and Pattern Recognition, IEEE, Piscataway, NJ 2020, pp. 1777–1786.
[18] S. Moran, P. Marza, S. McDonagh, S. Parisot, G. Slabaugh, in CVPR Seattle, IEEE, Piscataway, NJ 2020.
[19] C. Gong, D. Wang, M. Li, Y. Chandra, Q. Liu, in CVPR, Virtual, IEEE, Piscataway, NJ 2021.
[20] A. Radford, L. Metz, S. Chintala, in ICLR, Caribe Hilton, San Juan, Puerto Rico, 2016.
[21] T. De Bel, J.-M. Bokhorst, J. Van Der Laak, G. Litjens, Med. Image Anal. 2021, 70, 102004.
[22] A. Siorohin, S. Lathuiliere, E. Sangineto, N. Sebe, IEEE Trans. Pattern Analysis Mach. Intell. 2021, 43, 1156.
[23] O. Ronneberger, P. Fischer, T. Brox, in Medical Image Computing and Computer-Assisted Intervention, Pt III (Eds.: N. Navab, J. Hornegger, W. M. Wells, A. F. Frangi), Springer-Verlag, Munich 2015, pp. 234–241.
[24] J. Long, E. Shelhamer, T. Darrell, in 2015 IEEE Conf. on Computer Vision and Pattern Recognition, IEEE, Piscataway, NJ 2015, pp. 3341–3440.
[25] H. Zhao, J. Shi, X. Qi, X. Wang, J. Jia, in 30th IEEE Conf. on Computer Vision and Pattern Recognition, IEEE, Piscataway, NJ 2017, pp 6230–6239.
[26] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A. L. Yuille, IEEE Trans. Pattern Anal. Mach. Intell. 2018, 40, 834.
[27] K. He, G. Gkioxari, P. Dollar, R. Girshick, in 2017 IEEE Int. Conf. on Computer Vision, IEEE, Piscataway, NJ 2017, pp 2980–2988.
[28] J. Fu, J. Liu, H. Tian, Y. Li, Y. Bao, Z. Fang, H. Lu, in 2019 IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR), IEEE, Piscataway, NJ 2019, pp. 3141–3149.
[29] M. Feng, L. Zhang, X. Lin, S. Z. Gilani, A. Mian, Pattern Recognition 2020, 107, 107446.
[30] C. Peng, J. Ma, Pattern Recognition 2020, 107, 107498.
[31] H. Hu, Q. Li, Y. Zhao, Y. Zhang, IEEE Trans. Ind. Inform. 2021, 17, 2880.
[32] Q. Jin, H. Cui, C. Sun, Z. Meng, L. Wei, R. Su, Expert Syst. Applic. 2021, 176, 114848.
[33] Y. Yang, S. Soatto, in 2020 IEEE/CVF Conf. on Computer Vision and Pattern Recognition, IEEE, Piscataway, NJ 2020, pp. 4084–4094.
[34] H. Ma, X. Lin, Z. Wu, Y. Yu, in CVPR, 2021.
[35] R. Girshick, J. Donahue, T. Darrell, J. Malik, in 2014 IEEE Conf. on Computer Vision and Pattern Recognition, IEEE, Piscataway, NJ 2014, pp. 580–587.
[36] S. Ren, K. He, R. Girshick, J. Sun, in Advances in Neural Information Processing Systems, Vol. 28 (Eds.: C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, R. Garnett), IEEE Computer Soc, Los Alamos, CA 2015.
[37] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, in 2016 IEEE Conf. on Computer Vision and Pattern Recognition, IEEE, Piscataway, NJ 2016, pp. 779–788.
[38] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, A. C. Berg, in Computer Vision—Eccv 2016, Pt I, Vol. 9905 (Eds.: B. Leibe, J. Matas, N. Sebe, M. Welling), 2016, pp. 21–37.
[39] H. Law, J. Deng, in Computer Vision—ECCV 2018, Pt X, Vol. 11218 (Eds.: V. Ferrari, M. Hebert, C. Sminchisescu, Y. Weiss), Springer-Verlag, Berlin 2018, pp. 765–781.
[40] B. Yang, S. Wang, A. Markham, N. Trigoni, Int. J. Comput. Vision 2020, 128, 53.
[41] Q. Wang, M.-K. Kim, Adv. Eng. Inform. 2019, 39, 306.
[42] F. Hu, J. Zhao, Y. Huang, H. Li, Comput.-Aid. Civil Infrastruct. Eng. 2021, 36, 89.
[43] Z. Zhou, Y. Zhang, H. Foroosh, in CVPR (virtual event) 2021.
Danpeng Cheng received her master’s degree in School of Electrical Engineering from University of Rochester, in 2019. She is currently pursuing her Ph.D. degree in School of Electrical and Electronic Engineering in Huazhong University of Science and Technology(HUST). Her research interests focus on artificial intelligent (AI)-assisted material characterization, micrograph, and computer vision (CV).
Zuo Xu is a research professor of CITIC Dicastal Co., Ltd., he received bachelor's degree in metallurgy and heat treatment from Kunming Institute of technology, in 1987 and master's degree from Renmin University of China in 2009. He carries out basic and applied research in the lightweight of automobile components. Recently, he has made a series of important progress in the field of green and intelligent manufacturing, which has been successfully applied in the intelligent production line, making the company grow into first “Lighthouse Factory” in the automobile components in the world.

Shijie Cheng is the member of Chinese Academy of Sciences and professor of Huazhong University of Science and Technology. He received his bachelor’s degree from Xi’an Jiaotong University, in 1967, his master's degree from HUST, in 1981, and his Ph.D. from the University of Calgary (Canada), in 1986, respectively, all in electrical engineering. In 2007, he was elected as the member of the Chinese Academy of Sciences. He is currently engaged in the research on energy storage systems for electric power system stability and advanced materials for electrical engineering.