Inclusion Classification by Computer Vision and Machine Learning

Digital technologies are transforming industry at all levels. Steel has the opportunity to lead all heavy industries as an early adopter of specific digital technologies to improve our sustainability and competitiveness. This column is part of AIST's strategy to become the epicenter for steel's digital transformation, by providing a variety of platforms to showcase and disseminate Industry 4.0 knowledge specific for steel manufacturing, from big-picture concepts to specific processes.

Authors

Nan Gao
Department of Materials Science, Carnegie Mellon University, Pittsburgh, Pa., USA
ngao1@andrew.cmu.edu

Mohammad Abdulsalam
Center for Iron & Steelmaking Research, Department of Materials Science, Carnegie Mellon University, Pittsburgh, Pa., USA
mabdulsa@andrew.cmu.edu

Michael Potter
materials scientist, RJ Lee Group Inc., Monroeville, Pa., USA
mpotter@rjleegroup.com

Gary Casuccio
RJ Lee Group Inc., Monroeville, Pa., USA
gcasuccio@rjleegroup.com

Elizabeth Holm
Department of Materials Science, Carnegie Mellon University, Pittsburgh, Pa., USA
eaholm@andrew.cmu.edu

Bryan Webler
assistant professor, Center for Iron & Steelmaking Research, Department of Materials Science, Carnegie Mellon University, Pittsburgh, Pa., USA
webler@cmu.edu

This paper describes the use of computer vision and machine-learning methods to classify non-metallic inclusions in steel based on backscattered electron (BSE) scanning electron microscope (SEM) images obtained during automated inclusion analysis. The use of automated inclusion analysis has produced major contributions to both control of inclusions during steel processing and a mechanistic understanding of inclusion evolution.[1–3] Automated analysis utilizes an SEM equipped with a BSE detector and energy-dispersive x-ray spectroscopy (EDS). Thousands of features can be observed in times on the order of hours, yielding representations of the variable distributions. BSE images provide quantitative information on inclusion amount, size, shape and location, whereas EDS spectra provide information on chemical composition. BSE images also contain information about inclusion composition, since the production of backscattered electrons increases with atomic number. The objective of this work was to create a system that relates BSE images to EDS composition measurements. This required conversion of the BSE images into a numerical representation so that they could be interpreted by a computer.

When humans look at an image, we identify information based on the whole scene. The field of computer vision deals with the extraction of useful features from images using mathematical and statistical models to describe visual information such as edges, corners and blobs contained in images.[4] These features then can be combined to make a numerical representation of the image. Another computer vision approach employs convolutional neural networks (CNN), which are a type of deep machine-learning algorithm that performs very well at image recognition tasks.[5,6] A CNN passes the original image through multiple filter banks to create a multi-scale representation of the image in the form of a high-dimensional vector. The system then uses a classifier that identifies the probability that an image belongs to a given class. Both the filters and the classifier are learned from the training data. Once trained, the model can be used to classify additional images. An advantage of the CNN approach is that it does not require a human to identify the types of features to be considered; instead it learns them from the data.

CNNs are a type of machine-learning algorithm. Machine-learning methods attempt to automate data analysis or make predictions from data without intervention from a human. There are two general classes of machine-learning methods: supervised and unsupervised. The primary difference between them is whether a human assigns labels to the data. For supervised learning, the data are labeled with the ground truth (e.g., for a photo, the ground truth label might be "cat" or "dog"). Functions are built to map the relation between the data and the ground truth. Classification and regression are two common examples of supervised machine learning. Unsupervised learning methods draw inferences from data sets that do not contain labels. The most common unsupervised learning method is cluster analysis, which is used for exploratory data analysis to find hidden patterns or groupings in data.

In this work, a CNN was employed to recognize image features, develop numerical representations of these features and then use these representations for classification. A set of automated inclusion analysis data was used for:

- Training — Images and their classifications are used as inputs. The CNN algorithms optimize the model parameters so that model predictions match the ground truth classifications.
- Validation — A subset of training data used as a check of parameter optimization, as well as for manual parameter adjustment.
- Testing — A set of images (distinct from those used for training and validation) is classified by the CNN and the classifications are compared to the ground truth.

This procedure is an example of supervised machine learning. In this work, observations were either classified as "inclusion" or "not inclusion," i.e., a binary classifier was constructed. In the future, additional classes could be added so that better relationships between inclusion BSE image and chemical composition could be constructed. This approach could reduce the need for EDS analysis during automated inclusion analysis.

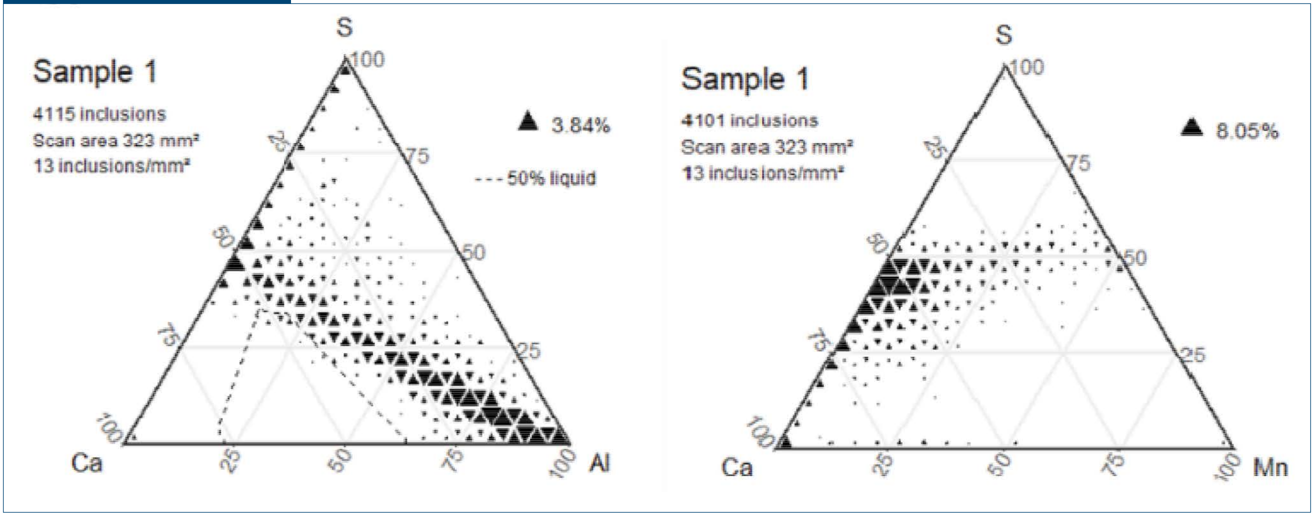
Materials and Methods

Sample and Automated Inclusion Analysis Details — Automated inclusion analysis data was provided from four samples. Data was collected from the Automated Steel Cleanliness Analysis Tool (ASCAT) system. Each sample was taken from different heats of plate product. Both BSE images (128 x 128 resolution) and EDS data were provided. A total of 49,372 observations were made for the four samples. From this data set, 15,924 observations classified as "inclusion" and 3,128 observations classified as "not inclusion" by internal ASCAT rules were selected. The remaining observations were also classified as "not inclusion" by the rules but were not selected for this analysis because they consisted of blank fields of view or very small particles with low EDS counts. The ASCAT classifications were taken to be the ground truth for the computer vision analysis.

Computer Vision Methods — The data set was first balanced, i.e., made to contain equal numbers of "inclusion" and "not inclusion" observations. This was performed by randomly sampling 3,128 observations from the "inclusion" class. The total data set size was therefore 6,256 images. From this data set, 4,003 observations were used to train the model, 1,001 observations were used to validate the model during training, and 1,252 observations were used to test the model after training was completed.

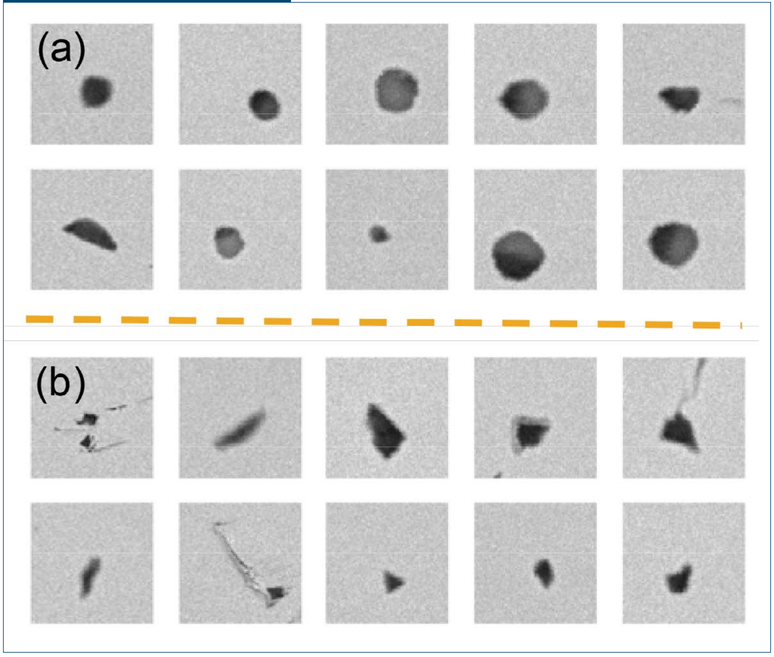
In this work, a publicly available CNN, VGG16, was used. This CNN has been developed for image classification tasks on a wide variety of images. The initial model parameters were taken for the VGG16 CNN pre-trained on the ImageNet database. ImageNet is a database of more than 1 million natural images of various everyday objects (but no micrographs). Despite having little relation to inclusions, features generated from the transfer learning process represent high-level conceptual information that can be utilized for problems well outside the scope of the original training database.[7,8] This is an example of transfer

Figure 1



Ternary diagrams showing representative inclusion compositions for the four samples studied in this work. The axes are in mole fraction of Ca, Al, S or Ca, Mn, S. The symbol size is proportional to the number density of inclusions of a particular composition.

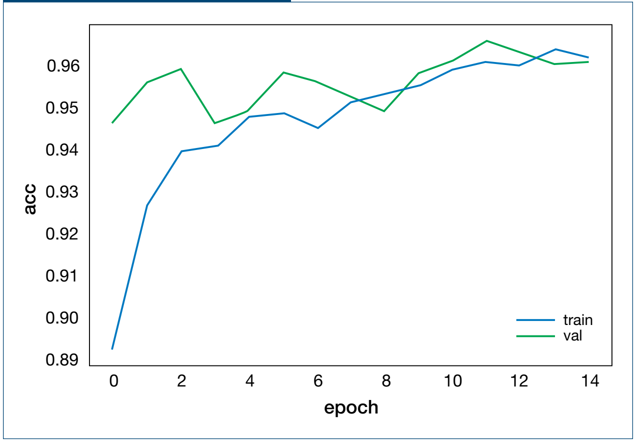
Figure 2



Representative backscattered electron (BSE) scanning electron microscope (SEM) images of features classified as (a) inclusion and (b) "not inclusion" by the internal ASCAT rules. These images are examples of those that were used for the CV analysis.

learning, a method in machine learning whereby knowledge from one problem can be applied to a different problem. The rationale of transfer learning is that all objects actually share some characteristics. The characteristics learned from one database can be used to extract the features from another similar database even if some of images have never been seen by the model before. In this problem, the investigators wanted to see if the parameters or features learned from a large number of natural images (ImageNet Data set) still apply to learned features from the inclusion data.[9,10] To do this, the VGG16 front end is applied, with parameters learned on the ImageNet database, but the classifier is retrained. That is, instead of identifying images as belonging to one of the ImageNet classes (cat, airplane, etc.), the system classifies them as "inclusion" or "not inclusion."

Figure 3



Accuracies of training and validation data during the VGG16 model training process. The highest training accuracy obtained was approximately 96%.

Results and Discussion

Ternary diagrams representing the inclusion compositions measured by EDS are given in Fig. 1 (shown for Sample 1 only, but inclusion compositions in other samples were very similar). Representative BSE images of observations classified by ASCAT rules as "inclusion" and "not inclusion" are shown in Fig. 2.

The VGG16 classifier was trained for 15 epochs (i.e., the number of times the training data is passed through the model and its parameters tuned). The accuracies, expressed as the percentage of correctly classified observations relative to the total number of observations, for the training and validation data during the training process are shown in Fig. 3.

The results from the test data set are shown in the form of a confusion matrix in Fig. 4. In this representation on diagonal entries represent correct predictions for the "inclusion" and "not inclusion" class. The overall accuracy of model predictions for the test data was 98%.

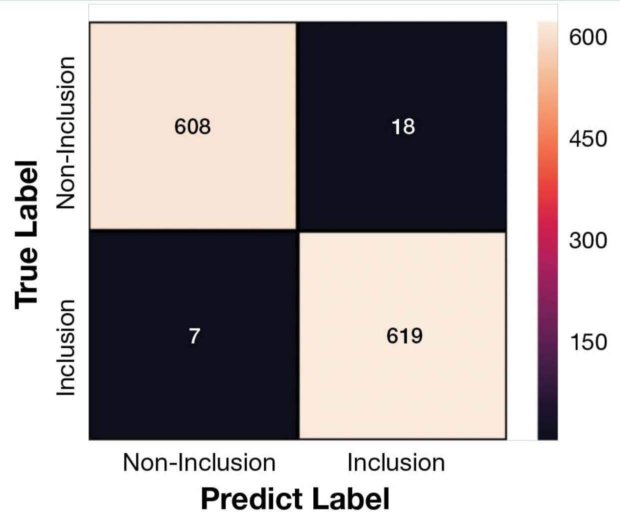
The CNN training process required 280 seconds for 4,003 images (approximately ~70 ms per observation). The average time spent analyzing test images was 69 ms per image. Based on the analysis setup procedures, the EDS scans required 1,000 ms per feature.

A previous study[11] applied a similar CNN approach to a different inclusion data set that was collected on a different SEM. In that study accuracies were 72% for the test data. Fig. 5 shows representative images illustrating the differences between the two data sets. The higher-contrast images of the current data set appeared to improve accuracy of the CNN approach.

Conclusions

This work showed that a computer vision approach could be used to predict whether an observation was an inclusion or not an inclusion with high accuracy.

Figure 4



Confusion matrix of CV model predictions from the test data set. On diagonal entries represent correct predictions for the "inclusion" and "not inclusion" class.

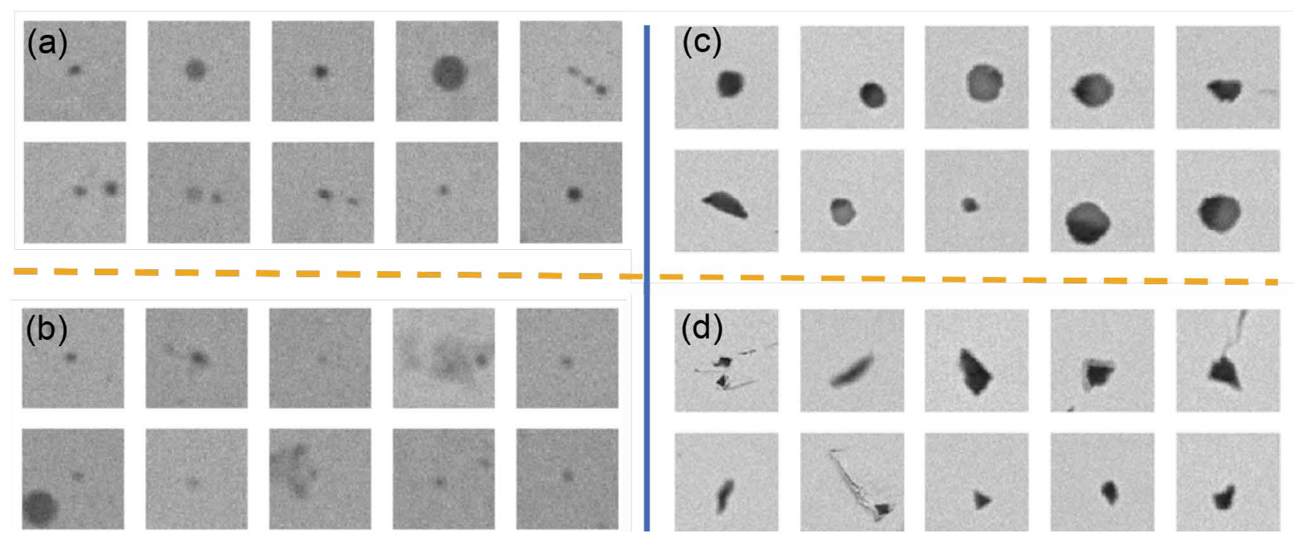
The prediction was made based on only BSE images. The next step in this work is to classify inclusions by chemical composition based only on BSE images. The influence of SEM setup parameters and the resulting BSE images was also shown to be important. Methods to accommodate variability in BSE images must be developed. However, with appropriate control of microscope setup and with sufficient data for model training, the CNN approach has the potential to aid filtering (i.e., identifying observations that are not inclusions before EDS measurement) and also to reduce the need for EDS measurements during automated inclusion analysis.

References

1. S.R. Story, S.M. Smith, R.J. Fruehan, G.S. Casuccio, M.S. Potter and T.L. Lersch, "Application of Rapid Inclusion Identification and Analysis," *Iron & Steel Technology,* September 2005, pp. 41–49.
2. E.B. Pretorius, H.G. Oltmann and B.T. Schart, "An Overview of Steel Cleanliness From an Industry Perspective," *AISTech 2013 Conference Proceedings,* Vol. I, 2013, pp. 993–1026.
3. P. Kaushik, J. Lehmann and M. Nadif, "State of the Art in Control of Inclusions, Their Characterization, and Future Requirements," *Metall. Mater. Trans. B,* Vol. 43, No. 4, 2012, pp. 710–725.
4. R. Szeliski, *Computer Vision,* Springer, London, U.K., 2011.
5. Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu and M.S. Lew, "Deep Learning for Visual Understanding: A Review," *Neurocomputing,* Vol. 187, 2016, pp. 27–48.
6. Y. LeCun, Y. Bengio and G. Hinton, "Deep Learning," *Nature,* Vol. 521, No. 7553, 2015, pp. 436–444.
7. K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," September 2014.
8. T. Mueller, A.G. Kusne and R. Ramprasad, *Machine Learning in Materials Science,* John Wiley & Sons Ltd., 2016, pp. 186–273.
9. K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," September 2014.
10. O. Russakovsky et al., "ImageNet Large-Scale Visual Recognition Challenge," *Int. J. Comput. Vis.,* Vol. 115, No. 3, 2015, pp. 211–252.
11. M. Abdulsalam, N. Gao, E.A. Holm and B.A. Webler, "Classification of Non-Metallic Inclusions Via Computer Vision and Machine Learning," *Clean Steel 9 Proceedings,* Budapest, Hungary, September 2018. ✦

This paper was presented at AISTech 2019 — The Iron & Steel Technology Conference and Exposition, Pittsburgh, Pa., USA, and published in the Conference Proceedings.

Figure 5



Representative BSE images from two automated inclusion analysis data sets, from [x], classified as "inclusion" (a), from [x], classified as "not inclusion" (b), from this work, classified as "inclusion" (c), and from this work, classified as "not inclusion" (d).