Proceedings of the International Conference on Industrial Engineering and Operations Management

Publisher: IEOM Society International, USA DOI: 10.46254/AP07.20260145
Published: March 25, 2026

Deep Progressive Learning for Grain Boundary Segmentation and Grain Size Estimation Using CNN-Based Models

Chaitanya Pramod Bane and Sukhomay Pal
Department of Mechanical Engineering
Indian Institute of Technology Guwahati, Assam, India
c.bane@iitg.ac.in; spal@iitg.ac.in

Abstract

In microstructure analysis, the precise identification of grains and grain boundaries for determining grain size and shape is a tedious and time-consuming manual task that also requires expert knowledge. Existing automated techniques, such as edge detection methods, are considerably faster than manual analysis but often suffer from limited accuracy. Accurate prediction of grain-related properties is essential in several fields, including understanding material properties, predicting material behavior, quality control, and process optimization. To address these challenges, deep learning-based convolutional neural network (CNN) models can be effectively employed. In the present work, a novel method is proposed for detecting grain boundaries and extrapolating incomplete boundaries to form closed grain regions, thereby enabling accurate quantification of grain parameters. This study introduces an ensemble learning–based framework comprising two models: the first employs a unique progressive learning strategy, while the second adopts an iterative approach to close incomplete grain boundaries. The proposed framework is implemented using the U-net architecture. A dataset of microstructure images is acquired using an optical microscope at varied magnifications. The materials investigated are mild steel and Inconel 625, prepared through standard metallographic procedures including polishing and etching prior to image acquisition. The ensemble model is trained on a diverse dataset that also includes artificially generated images. The developed model demonstrates high accuracy in predicting several grain characteristics, including average grain area, total number of grains, grain height, and grain width. The results indicate that the network effectively learns these properties by capturing the curvature and morphology of grain boundaries.

Keywords
Grain boundary prediction, Grain size, CNN, U-Net architecture, Ensemble Learning, Quantitative analysis

1. Introduction
While studying material properties for design, a method for crafting the material involves knowing information right from the atomic level to the macroscopic level. Knowing this information lets the design meet its specific requirements and criteria. This approach has the potential to revolutionize the development of new advanced materials. Metallic materials have distinct crystalline regions which are called grains in metallurgy. The size and shape of these grains can affect physical factors such as strength and strain rate (Figueiredo et al. 2022), fatigue life (Mlikota et al. 2021), ductility (Roberto et al. 2024) and corrosion (Tian et al. 2021). Hence, an accurate quantitative analysis is essential to predict the behaviour of the products to be designed.

Microstructures are traditionally observed under an optical microscope or a scanning electron microscope after carefully polishing and etching a cross-section of metal samples. Since it is difficult to measure the volume of a grain directly, the next best measurement is the grain size or area from a two-dimensional cross-sectional image. Manual methods to find the grain size and shape are Heyn's line intercept method and planimetric method (ASTM standard, 2004). However, the grain size obtained via manual interaction can vary from individual to individual. Moreover, the calculation is a tedious process and thus automation of these techniques is necessary. Computational techniques using image processing methods have risen in the decade, they make use of existing image processing to present numerical details about the microstructure image. Chuang et al. (2012) presented a method to extract grains and determine grain size from images using a traditional watershed algorithm. Li et al. (2022) automated the intercept method using a topological skeleton.

Grain boundary segmentation is an important task as the proper distinction of grain from grain boundary is necessary for the automation of subsequent analysis processes. Manual segmentation of an image with a large number of grains is impractical and needs automatic detection of the grain boundaries. Dengiz et al. (2005) introduced a method to detect grain boundaries for superalloy steel microstructure using fuzzy logic and a neural network. Peregrina-Barreto et al. (2013) gave an automatic method to identify grain boundaries based on the colour difference to threshold the image into grain and grain boundaries. These methods are not easy to implement as the user needs to study information regarding the image and then apply the method. User bias also cannot be avoided as the user needs to assign values to particular parameters. Noise is generated in the output images and is unavoidable. Pre-processing and post-processing with careful observation and modification are needed. In this study, a deep learning-based convolutional neural network (CNN) is experimented to address these issues.

The CNN models are seen to capture salient features at different levels. The output accuracy of these models depends on the training data. Once trained these CNN models can predict similar outputs for new images rapidly and without any operator bias (Ronneberger et al. 2015). As these supervised training models largely depend upon the availability of training data, they are most of the time limited to definite tasks. The generation of a robust model requires a varied dataset. As no standard dataset is available, the creation of such a dataset is a primary focus of this study. For segmentation tasks, CNN-based models such as U-Net (Ronneberger et al. 2015), DeepLab (Chen et al. 2018), and Mask R-CNN (He et al. 2017) have been shown to deliver state-of-the-art performance. U-Net is used for this task, it features an architecture that consists of a compressive path which goes convoluting the input image and extracts the necessary feature at every layer, and an expansive path that tries to rebuild the original features along with concatenating the cropped features extracted in the compressive path. Similar works have been done in the past but they require optimal post-processing with morphological operations to have complete grain boundaries (Patrick et al. 2023). High-quality images from optical microscope are necessary for this along with careful manual segmentation for training the model. The U-net model has proved to be very effective in fields where limited data is present such as biological samples, this makes the model better suitable for the present work. It becomes prudent to check the performance of the existing detection methods to that of the CNN models. These existing techniques are manual segmentation, manual thresholding, gradient-based approach (Sobel and Feldman 1990, Scharr 2004, Canny 1986, Prewitt 1970) and Holistically nested edge detection technology (Xie and Tu 2015).

With proper segmentation of grain and grain boundary, the task moves onto the grain size calculation. In this study, a slightly modified version of the convolutional network described by Manit et al. (2017) is employed. The segmented image is fed as the input to this model and the output layer is used to predict the values for size and shape of the grains in the image. The hidden layer consists of convolution layers, pooling layers, dropout layer and a final fully connected layer. Rectified linear unit (ReLU) activation function is used for faster training. Mean squared error (MSE) with Adaptive moment estimation (Adam) optimizer (Kingma and Ba 2015) is used for training the models. The line intercept method by Heyn (1925) involves drawing a line across the image and counting the number of grains intersected by the line. The average grain size is then determined by dividing the total number of grains by the line length. By recording these intercepts, researchers can estimate grain size distribution within the sample. While the grain intercept method provides a substantial number of measurements, it is inherently limited as a one-dimensional approximation of a complex three-dimensional structure. Nonetheless, its continued utilization in current research reinforces its relevance in this study. The planimetric method, a 2-dimensional approach (ASTM Standard 2004, Saltikov 1967), is utilized for extracting the size and shape of grains. The planimetric size data includes four measurements. The first measurement, termed the total grain metric, represents the overall count of grains across all images. The second measurement is the average individual grain area. The third and fourth measurements correspond to the total grain area per image and the total grain boundary area per image. The planimetric shape data includes three distinct measurements. The first measurement is the average width of the grains, termed the average X-diameter. The next measurement is the average diameter in the Y-direction, also known as the average height. The final value is the aspect ratio, calculated by dividing the average X-diameter by the average Y-diameter, providing a measure of the width-to-height relationship.

In this study, a method is presented to measure grain size for metallic samples using deep learning-based CNN techniques with no user interaction, in accordance with ASTM grain testing standards. Two CNN models are implemented with different architecture simultaneously, the first takes inspiration from the U-net architecture (Ronneberger et al. 2015), which gives an image with distinctly segmented grain and grain boundaries. This image is used as the input for the next model to give out the grain size. Training of the models is done using images obtained from visualising a polished and etched sample of mild steel and Inconel-625 under an optical microscope at 10X, 50X and 100X magnifications. Comparison of accuracy with existing methods is done to show the capability of the proposed method.

2. Methods

2.1. Grain Boundary Detection Model

Figure 1 illustrates the operation of a single convolution block. A convolutional block is a building block used in a convolutional neural network (CNN) for image recognition. This convolution block is used to extract features and can have more than one convolution layer. First, a convolution operation is performed on the input image using a specified number of filters. Subsequently, batch normalization is applied, followed by the specified activation function.

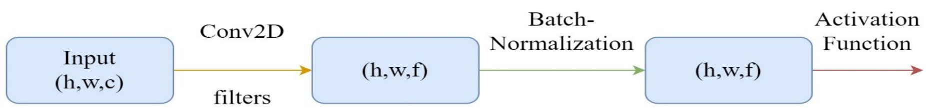

Figure 1. Convolution block used in grain boundary detection model

An improved U-Net architecture is proposed, here the convolution block is used instead of a simple convolution layer. Adam is chosen as the optimization algorithm, which could maximize the performance of the model and achieve the highest score in all evaluation metrics. The loss considered is binary cross-entropy as the given problem will be of binary segmentation. The F1 score or Dice coefficient is used as a similarity metric. The structure of the proposed segmentation model is shown in Figure 2. The proposed architecture has 3 down-sampling levels and 3 up-sampling levels with a bottleneck layer of 128 × 128. Each blue box corresponds to the feature map. The number of channels is denoted on top of the box. The x-y-size is provided inside the box. Yellow boxes represent copied feature maps. The arrows denote the different operations.

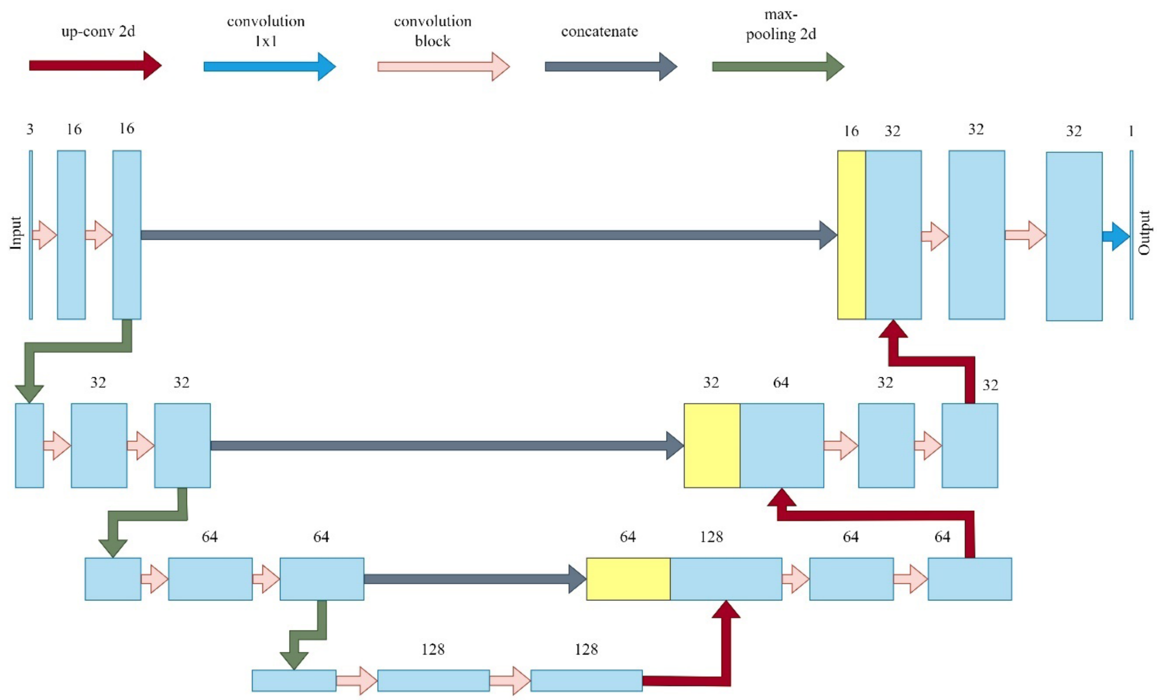

Figure 2. Proposed Grain Boundary Detection CNN architecture

The model is trained using input images acquired from micrographs and corresponding output masks generated by four different methods. Different training sets are created using combinations of each method. This is done to test whether it would train a more robust model and to see the effects it can have on the accuracy. The mask generation in this specific case is classifying the pixels as either grains or grain boundaries. The masks used are generated manually by humans. The dataset is split into 95% for training and 5% for testing. Two crucial aspects of the training process are determining the training data and defining the desired outcome, referred to as the mask.

The grain boundaries are assigned a black pixel (pixel value to zero) and the grains are assigned a white pixel (pixel value to 1). Existing segmentation methods: (i) Manual segmentation, (ii) Manual thresholding, (iii) Gradient-based methods, and (iv) Holistically-Nested Edge Detection (HED) method are used to generate pre-processed images for the real grains. This means that for every single manually segmented image there are 4 images that can be considered as its input namely, HED pre-processed image, Canny-Edge processed image, Manual thresholding and the Real grain itself.

2.2 Grain Size Prediction Model

The proposed neural network architecture is a CNN designed for grayscale image-based regression tasks. It begins with a 300×300 single-channel input and processes the data through a series of convolution and pooling layers, gradually increasing the number of feature maps to capture hierarchical image features. These convolutional layers are followed by ReLU activation, introducing non-linearity, and max-pooling to downscale the feature maps. A dropout layer with a 50% drop rate is utilized to prevent overfitting. Subsequently, a fully connected layer with 1024 neurons and ReLU activation is employed to learn high-level representations from the flattened features. The model concludes with a linear output layer, well-suited for regression, providing a continuous numeric prediction. During training, the model optimizes its parameters using the Adam optimizer and minimizes the MSE loss function. Furthermore, the network's performance is evaluated based on mean absolute error (MAE).

3. Data Collection

3.1. Sample Preparation and Microstructure Image Acquisition

The dataset required is obtained by polishing and etching mild steel and IN-625 samples of face dimensions as 5 mm × 30 mm. The specimens are polished by different grades of emery papers and velvet cloth with fine alumina powder (0.3 μm) for mirror polish. The samples are etched with a suitable chemical reagent to enhance grain boundary visibility. Images are taken from every sample at different magnifications and in various regions across the surface. A resolution of 5 MP is considered, as higher quality images can lead to better results. Figure 3 shows one of the obtained images.

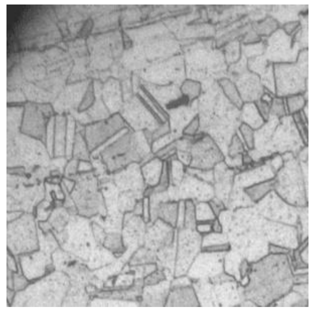

Figure 3. A high resolution image of mild steel microstructure

3.2. Image segmentation

The various methods and techniques employed in creating the dataset for training the developed models include manual segmentation, manual thresholding, gradient-based methods, and the Holistically-Nested Edge Detection (HED) method. Figure 4 illustrating the manual segmentation method.

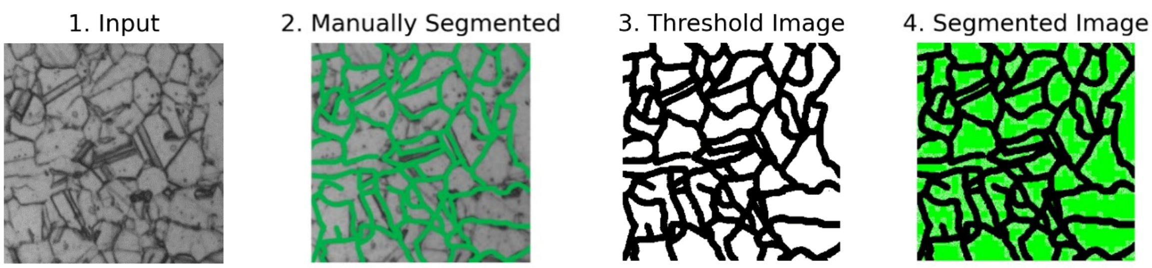

Figure 4. Steps involved in manual segmentation method

4. Results and Discussion

Numerous tests are carried out to find the optimal parameters using hyperparameter adjustment to minimize MSE. By varying various parameters, including learning rates, the number of neurons in a fully connected layer, and the number of epochs for training, different values of loss were obtained. Google Colab, along with the TensorFlow/Keras library, was utilized for the development, analysis, and optimization of the proposed CNN framework. The models are trained until the training and validation losses start to deviate from each other. Figure 5 shows the variation of accuracy with number of iterations.

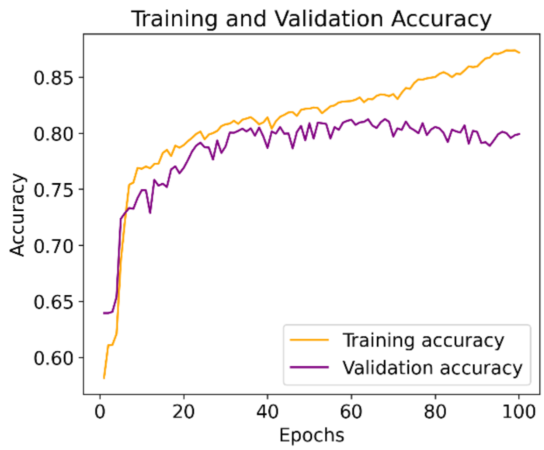

Figure 5. Variation of F1-Score with number of iterations

In this study, the F1-score (Terven et al. 2023) is used to assess the prediction accuracy of the model in identifying grain boundaries. Figure 6 presents the F1-scores obtained for the existing segmentation methods. Manually segmented images are used as the ground truth for evaluation; therefore, they achieve an F1-score of 1.0, indicating perfect overlap. It is seen that the traditional methods used to segment show very less F1-scores. The least being that of HED and manual thresholding and the gradient-based canny edge detection scoring almost equally. However, these datasets with low scores can be used to increase the variety in the dataset making it to be more robust and effective on unseen data. This proves that existing methods have many challenges and cannot accurately capture complex features such as grain boundaries. Moreover, it is a time-consuming process with inputs needed from the user. Deep learning-based techniques such as the U-net model can improve this process.

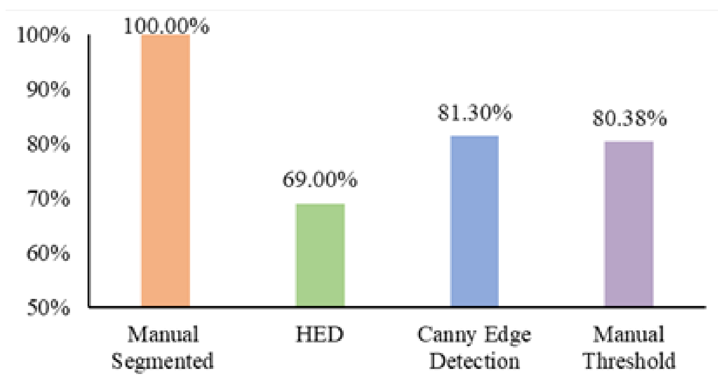

Figure 6. F1-score comparison for three existing methods

Figure 7 illustrates the F1-scores for the generated training sets with existing methods and after training on U-net. The F1-score increases greatly in all of the cases except Case 1 where it is slightly less. This can be expected as Case 1 consists of the manually segmented images which are considered as the ground truth while calculating the F1-scores. Even with only 25% of real grain images in Case 5 it performs considerably well. The cases with varied inputs performed better than the ones that had less variation. The prediction of microstructure of the developed U-net model is shown in Figure 8.

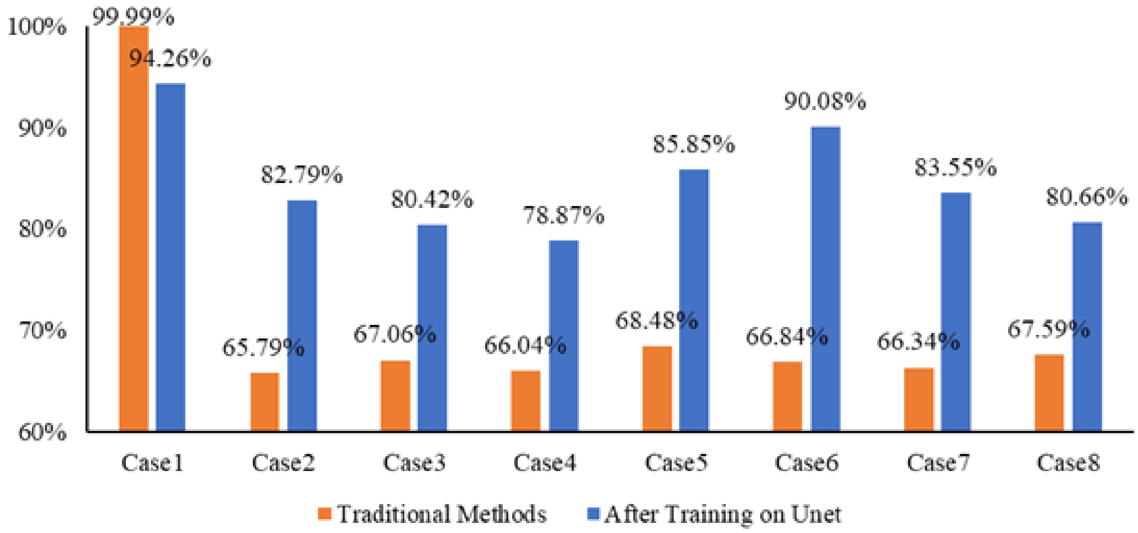

Figure 7. Comparison of F-scores across eight training sets obtained using traditional methods and the trained U-net model.

(a) (b) (c)

Figure 8. (a) Original image, (b) corresponding ground truth mask, and (c) CNN-predicted segmentation

6. Conclusion

Using edge detection as the primary focus, a U-net model is trained to detect grain boundaries. Conventional edge detection techniques are computationally time-consuming and highly sensitive to parameter selection. To overcome these limitations, an automated deep learning–based approach is explored in this work. The application of various pre-processing techniques enables data augmentation, leading to a more robust and reliable prediction model. To further enhance model autonomy and generalization, a large dataset of microstructure images is acquired. The model faces challenges in accurately identifying grain boundaries for extremely fine grains. This limitation can be addressed by implementing a deeper network architecture. Overall, the U-net model outperforms existing edge detection methods, as once trained, it does not require manual parameter tuning for grain boundary detection.

References

ASTM, ASTM E112-96: Standard test methods for determining average grain size, ASTM International, 2004.

Canny, J., A computational approach to edge detection, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6, pp. 679–698, 1986.

Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K. and Yuille, A. L., DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 4, pp. 834–848, 2018.

Chuang, H. S. and Lin, C. H., Automated grain sizing using mark-based watershed algorithm, *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, Munich, Germany, 2012, pp. 2332–2335.

Figueiredo, R. B., Kawasaki, M. and Langdon, T. G., The role of grain size in achieving excellent properties in structural materials, *Journal of Materials Research and Technology*, vol. 30, pp. 3448–3462, 2024.

Figueiredo, R. B., Wolfa, W. and Langdon, T. G., Effect of grain size on strength and strain rate sensitivity in the CrMnFeCoNi high-entropy alloy, *Journal of Materials Research and Technology*, vol. 20, pp. 2358–2368, 2022.

He, K., Gkioxari, G., Dollár, P. and Girshick, R., Mask R-CNN, *Proceedings of the IEEE International Conference on Computer Vision*, Venice, Italy, 2017, pp. 2980–2988.

Heyn, E., *Physical Metallography*, Wiley, 1925.

Kingma, D. P. and Ba, J. L., Adam: A method for stochastic optimization, *Proceedings of the International Conference on Learning Representations*, San Diego, USA, 2015.

Li, X., Cui, L., Li, J., Chen, Y., Han, W., Shonkwiler, S. and McMains, S., Automation of intercept method for grain size measurement: A topological skeleton approach, *Materials & Design*, vol. 224, p. 111358, 2022.

Mahmud, B. U. and Hong, G. Y., Semantic image segmentation using CNN-based technique, *Proceedings of the IEEE World Conference on Applied Intelligence and Computing*, Sonbhadra, India, 2022, pp. 210–214.

Manit, J., Schweikard, A. and Ernst, F., Deep convolutional neural network approach for forehead tissue thickness estimation, *Current Directions in Biomedical Engineering*, vol. 3, no. 2, pp. 103–107, 2017.

Mlikota, M., Dogahe, K., Schmauder, S. and Božić, Ž., Influence of the grain size on the fatigue initiation life curve, *International Journal of Fatigue*, vol. 158, p. 106562, 2021.

Patrick, M. J., Eckstein, J. K., Lopez, J. R., Toderase, et al., Automated grain boundary detection for bright-field transmission electron microscopy images via U-Net, *Microscopy and Microanalysis*, vol. 29, no. 6, pp. 1968–1979, 2023.

Peregrina-Barreto, H., Terol-Villalobos, I., Rangel-Magdaleno, J., Herrera-Navarro, A., Morales-Hernández, L. and Manríquez-Guerrero, F., Automatic grain size determination in microstructures using image processing, *Measurement*, vol. 46, no. 1, pp. 249–258, 2013.

Prewitt, J., Object enhancement and extraction, *Picture Processing and Psychopictorics*, pp. 75–149, 1970.

Ronneberger, O., Fischer, P. and Brox, T., U-Net: Convolutional networks for biomedical image segmentation, *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, Munich, Germany, 2015.

Saltikov, S. A., The determination of the size distribution of particles in an opaque material from a measurement of the size distribution of their sections, in Elias, H. (Ed.), *Stereology*, Springer, Berlin, Heidelberg, 1967.

Scharr, R., Optimal edge detection in the presence of noise, *Proceedings of the International Conference on Computer Vision Theory and Applications*, 2004.

Smith, A., Dengiz, O. and Nettleship, I., Grain boundary detection in microstructure images using computational intelligence, *Computers in Industry*, vol. 56, no. 8–9, pp. 855–866, 2005.

Sobel, I. and Feldman, G., An isotropic 3×3 image gradient operator, *Machine Vision for Three-Dimensional Scenes*, pp. 376–379, 1990.

Terven, J., Cordova-Esparza, D.-M., Ramirez-Pedraza, A. and Chávez Urbiola, E., Loss functions and metrics in deep learning, 2023.

Tian, L., Zheng, R., Yuan, C., Yang, G., Shi, C., Zhang, B. and Zhang, Z., Effect of grain size on the corrosion behavior of fully recrystallized ultra-fine grained 316L stainless steel fabricated by high-energy ball milling and hot isostatic pressing sintering, *Materials Characterization*, vol. 174, p. 110995, 2021.

Xie, S. and Tu, Z., Holistically-nested edge detection, *Proceedings of the IEEE International Conference on Computer Vision*, Santiago, Chile, 2015.

Biographies

Mr. Chaitanya Pramod Bane was a student of the Department of Mechanical Engineering, Indian Institute of Technology Guwahati, Assam, India. He completed his Master of Technology degree in Manufacturing Science and Engineering from the Department of Mechanical Engineering.

Dr. Sukhomay Pal is a Professor in the Department of Mechanical Engineering at the Indian Institute of Technology Guwahati, India. Prof. Pal's research interests span welding process monitoring and control, tool condition monitoring, and the application of Artificial intelligence in manufacturing processes. He has contributed extensively to both academic research and industry-oriented projects. He has led and participated in several sponsored and consultancy projects supported by agencies including SERB (India), DRDO (Naval Research Board), NEEPCO Ltd., etc. Prof. Pal has supervised numerous Ph.D. and M.Tech students in topics related to welding process modeling, and manufacturing systems. An active researcher and prolific author, Prof. Pal has published numerous articles in international journals and conferences. He has contributed to edited books and chapters, and serves as a reviewer for reputed international journals and conferences.