MetaScope: A Model for Microstructural Analysis of Carbon Steels Using Machine Learning Techniques

Lucian Fernando Bellini[1], Éric Cassiano Drey dos Santos[2], Yeun Haur Kang[1], William Haupt[3], Carlos Amaral Hölbig[1,4]

[1]Curso de Ciência da Computação

[2]Curso de Análise e Desenvolvimento de Sistemas

[3]Programa de Pós-Graduação em Projeto e Processos de Fabricação

[4] Programa de Pós-Graduação em Computação Aplicada

Universidade de Passo Fundo (UPF) – Passo Fundo, RS – Brazil

{192558, 202957, 193593, williamhaupt, holbig}@upf.br

Abstract. Microstructural characterization in metallography often relies on manual analysis, which introduces bias and reduces reproducibility. To overcome this, we propose MetaScope, a deep learning model for pixel-level segmentation of ferrite and pearlite in SAE 1020 and 1045 steels. Using 2,200 images with masks generated by a Gaussian Mixture Model, the model combines an U-Net architecture with an EfficientNetB0 encoder. This approach demonstrates the potential of using machine learning techniques to support metallographic workflows and future expansion to other materials.

1. Introduction

Microstructural characterization of carbon steels is critical to ensuring that their final mechanical properties meet the intended design specifications. For instance, steels classified as SAE 1020 exhibit tensile strength and malleability characteristics that make them more suitable for certain applications compared to SAE 1045 steels [Ishtiaq et al. 2022]. However, in traditional design of experiments, particularly in metallographic workflows, the reliance on manual image interpretation makes the process susceptible to human bias and inconsistencies across laboratories. These limitations become even more pronounced when working with large datasets or when rapid feedback is required.

Considering this, it is relevant to cite the advancements in machine learning techniques applied to scientific problems across multiple knowledge areas. Bibliometric studies [Gargiulo et al. 2022] show that nearly half of AI-related publications now come from outside its traditional domains, with materials science among the fastest adopters. This shift underscores the versatility of AI in addressing complex, data-heavy challenges. In microstructural characterization, such methods extract quantitative data from metallographic images, reducing bias and improving reproducibility. By detecting subtle patterns and handling large datasets, machine learning delivers rapid, consistent, and highly accurate results [Amano et al. 2025].

In response to these challenges, which were a severe problem in the Institute of Technology at UPF, we developed MetaScope, a model designed to automate the microstructural characterization of carbon steels through deep learning. At its core, MetaScope addresses the segmentation problem by employing a U-Net convolutional neural network architecture [Motyl and Madej 2022], enhanced with EfficientNetB0 [Luengo et al. 2022] as a pretrained feature extractor to improve accuracy in identifying microstructural patterns. The system is capable of pixel-level discrimination between ferrite and pearlite phases in metallographic images, enabling precise quantification of their respective fractions.

2. Methods

2.1. Sample data

The dataset used in this study consisted of microstructural images of SAE 1020 and SAE 1045 steels, captured at the Institute of Technology of the University of Passo Fundo specifically for the development of the MetaScope model. A total of 2.200 images were available, evenly distributed between the two steel kinds (1.100 images each). Every image was provided in high resolution (3840×2748 pixels) and saved in the TIFF format.

2.2. Segmentation masks (labels)

In essence, our semantic segmentation model learns to classify each pixel by comparing its predictions to a ground truth mask. This mask is an image of the steels we talked about earlier, with each pixel representing a different part. Considering this, the initial stage of data preparation involved generating pixel-level segmentation masks for each microstructural image (Fig. 1). For preprocessing, all images were resized to 256×256 pixels to normalize their dimensions to model training and avoid overloading.

Since the ground-truth masks were not available, we adopted an unsupervised approach for label generation. Based on the accuracy obtained in preliminary tests, the Gaussian Mixture Model (GMM) algorithm was selected to create the masks, as it provided better results, particularly in handling overlapping regions. This algorithm assigns pixels to classes according to similarities in their color and intensity distributions.

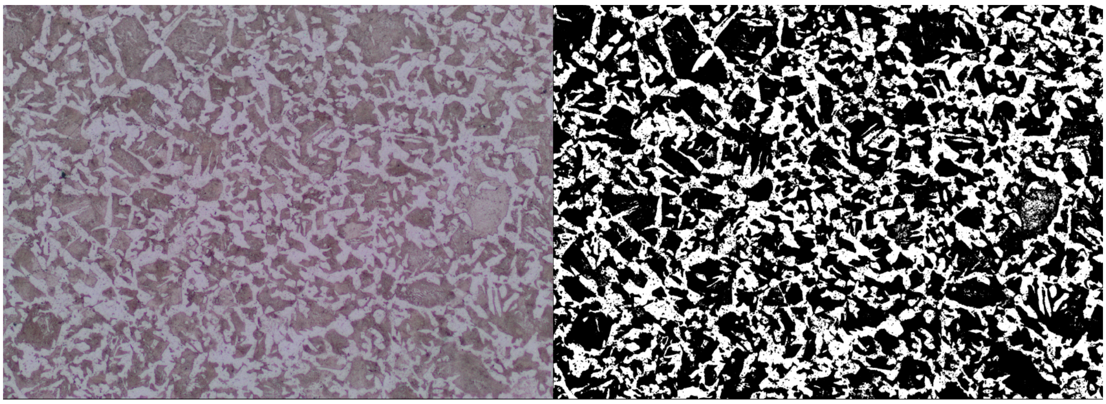

Figure 1. Original microstructural image (left) and segmentation mask (right).

2.3. Model development and training

For model development, the dataset was randomly split into training and testing sets using a fixed random seed (42) to ensure reproducibility, following an 80/20 proportion. Before the initial training and evaluation stages, a validation subset was derived from the training set, comprising 15% of its images, to support hyperparameter tuning and monitor the model's generalization performance during the training process.

The U-Net architecture was selected for the segmentation task due to its ability to perform pixel-wise classification while preserving spatial context. To improve accuracy, the network employed an EfficientNetB0 encoder pre-trained on ImageNet [Deng et al. 2009]. This approach leveraged transfer learning to accelerate convergence and reduce the training time while maintaining the segmentation quality.

3. Results and discussion

The EfficientNetB0 U-Net architecture achieved a mean Intersection over Union (mIoU) of 95.11% and Pixel Precision of 97.86% on the test set, demonstrating high segmentation performance. This shows the model's ability to delineate ferrite and pearlite regions with high precision, even in areas where the phases partially overlap and the contrast between microstructural features is subtle.

Despite the positive results, certain limitations must be acknowledged. The ground truth masks were generated using a GMM segmentation process, which, while effective for most cases, may introduce bias in images containing atypical textures or illumination patterns. Furthermore, the dataset is restricted to two steel types, limiting the model's applicability to broader metallographic contexts. Future work should explore expanding the dataset to include different alloys and microstructural configurations, as well as refining annotation methods to further improve model robustness.

In summary, the EfficientNetB0 U-Net demonstrated strong segmentation performance, with stable convergence and high precision. These results emphasize the advantages of combining a U-Net architecture with a pre-trained backbone, particularly in scenarios demanding high accuracy. Although the model successfully met the internal requirements of the School of Engineering, further improvements are still needed to enhance its generalization capability, especially to better handle external factors such as lighting variations and visual noise.

4. Conclusion

The developed model was capable of segmenting pearlite and ferrite phases in SAE 1020 and SAE 1045 steels, producing results that were consistent with the expected microstructural patterns. The combination of a U-Net architecture with an EfficientNetB0 encoder enabled stable training and satisfactory segmentation quality. The current dataset, however, only includes two steel kinds, which may restrict the model's application to other materials. Future work will focus on expanding the dataset to include a wider range of alloys and microstructural properties, as well as possible model architecture changes to improve accuracy and robustness.

Acknowledgements

This research is part of the "UPF Connection 2030: unfolding interdisciplinary boundaries between research and extension toward the goals of sustainable development" project, which is funded by CAPES PROEXT PG - PROCESS No. 88881.926842/2023-01.

References

Amano, N., Lei, B., Müller, M., Mücklich, F., and Holm, E.A. (2025) "Microscopy modality transfer of steel microstructures: Inferring scanning electron micrographs from optical microscopy using generative AI", Materials Characterization, Elsevier, Netherlands. doi:10.1016/j.matchar.2024.114600.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L. (2009) "ImageNet: A large-scale hierarchical image database", IEEE Conference on Computer Vision and Pattern Recognition, IEEE, USA. doi:10.1109/CVPR.2009.5206848.

Gargiulo, F., Fontaine, S., Dubois, M., and Tubaro, P. (2022) "A meso-scale cartography of the AI ecosystem", arXiv preprint, arXiv:2212.12263, Cornell University, USA.

Ishtiaq, M., Inam, A., Tiwari, S., and Seol, J.B. (2022) "Microstructural, mechanical, and electrochemical analysis of carbon doped AISI carbon steels", Applied Microscopy, 52:10. SpringerOpen, Republic of Korea. doi:10.1186/s42649-022-00079-w.

Motyl, M., Madej, Ł. (2022) "Supervised pearlitic–ferritic steel microstructure segmentation by U-Net convolutional neural network", Archives of Civil and Mechanical Engineering. doi:10.1007/s43452-022-00531-4.

Luengo, J., Moreno, R., Sevillano, I., Charte, D., Peláez-Vegas, A., Fernández-Moreno, M., Mesejo, P., Herrera, F. (2022) "A tutorial on the segmentation of metallographic images: Taxonomy, new MetalDAM dataset, deep learning-based ensemble model, experimental analysis and challenges", Information Fusion. doi:10.1016/j.inffus.2021.09.018.