Training computer vision models without labelled data using physics informed domain randomisation

Basil Jose[1], Fabian Hampp[1]

[1]) Institute of Combustion Technology for Aerospace Engineering (IVLR), Pfaffenwaldring 38-40, University of Stuttgart

Data driven deep learning models have always been identified with extensive success rate in the field of computer vision. However, an important bottleneck is the generation of labelled trainings data in order to create accurate instance segmentation models for the segregation of the desired signal from its background. Various tools are available to ease the data annotation process, yet it still requires significant manual intervention. Domain randomisation [1] is a novel strategy to generate synthetic trainings data that can sufficiently substitute the necessity for conventional data annotation. In this context, AI-models are trained on data generated using physics informed domain randomisation and used to segment instances of soot filaments and spray droplets and ligaments from experimental images acquired via shadowgraphy and stereoscopic particle image velocimetry. Automated image segmentation process identifies and spatially segregates instantaneous soot filaments or spray characteristics like droplet size and distribution. In addition, segmenting temporally resolved experimental images further enables the delineation of soot evolution, flow field and spray mixture formation, etc. Recently, these temporally resolved measurements have been proven highly insightful. [2, 3] in understanding the involved material interactions and their evolution properties. The advanced fundamental understanding aids, for example, the development of alternate jet fuels and fuel flexible combustion systems. Models generated using the above strategy are subsequently combined with conventional data analysis tools which supplements the latter with contextual detections, modularity and performance enhancements.

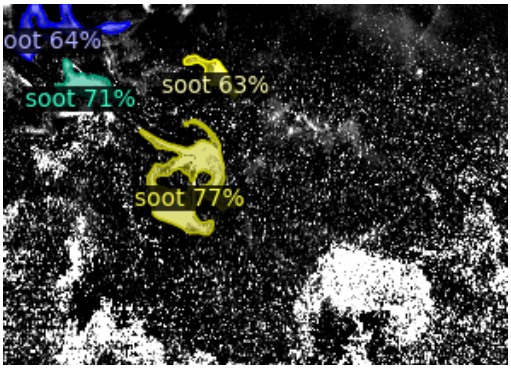

Figure 1: Inference on a real Mie scattering image using AI-model trained on synthetic data. The model can generate contextual detections from randomly generated training data.

References

[1] J. Tobin, R. Fong, A. Ray, J. Schneider, W. Zaremba, and P. Abbeel. 2017. Domain randomisation for transferring deep neural networks from simulation to the real world. *CoRR* abs/1703.06907.

[2] V. Narayanaswamy and N. T. Clemens. 2013. Simultaneous LII and PIV measurements in the soot formation region of turbulent non-premixed jet flames. *Proceedings of the Combustion Institute* Volume 34, Issue 1. pp. 1455-1463.

[3] M. Stöhr, K. P. Geigle, R. Hadef, I. Boxx, C. D. Carter, M. Grader, and P. Gerlinger. 2019. Time-resolved study of transient soot formation in an aero-engine model combustor at elevated pressure. *Proceedings of the Combustion Institute* 37(4):5421–5428.