Adapting OpenAI's CLIP Model for Few-Shot Image Inspection in Manufacturing Quality Control: An Expository Case Study with Multiple Application Examples

F. M. Megahed[1], ***Y. Chen***[2], ***B. M. Colosimo***[3], ***M. L. Grasso***[3], ***L. Jones-Farmer***[4], ***S. Knoth***[5], ***H. Sun***[6], ***I. Zwetsloot***[7]

[1]Department of Farmer School of Business, Miami University, USA
[2]University of Dayton, USA
[3]Politecnico di Milano, Italy
[4]Miami University, USA
[5]Helmut Schmidt University, Germany
[6]University of Georgia, USA
[7]University of Amsterdam, Netherlands

This expository paper introduces a simplified approach to image-based quality inspection in manufacturing using OpenAI's CLIP (Contrastive Language-Image Pretraining) model adapted for few-shot learning. While CLIP has demonstrated impressive capabilities in general computer vision tasks, its direct application to manufacturing inspection presents challenges due to the domain gap between its training data and industrial applications. We evaluate CLIP's effectiveness through five case studies: metallic pan surface inspection, 3D printing extrusion profile analysis, stochastic textured surface evaluation, automotive assembly inspection, and microstructure image classification. Our results show that CLIP can achieve high classification accuracy with relatively small learning sets (50-100 examples per class) for single-component and texture-based applications. However, the performance degrades with complex multi-component scenes. We provide a practical implementation framework that enables quality engineers to quickly assess CLIP's suitability for their specific applications before pursuing more complex solutions. This work establishes CLIP-based few-shot learning as an effective baseline approach that balances implementation simplicity with robust performance, demonstrated in several manufacturing quality control applications.

Keywords: Computer vision, Industry 4.0, Supervised fault detection, Vision transformer, And visual inspection

Acknowledgements
The work is partially supported by the NSF Grant No. FM-2134409 and CMMI-2412678.