

Towards Interpretable and Efficient Agricultural Image Classification: A Review of Autoencoder-Enhanced YOLOv8 Architectures with Spatial Attention and Feature Compression

¹Ms. Deepak B. Patil and ²Mr. Prathmesh S. Powar

¹Research Scholar, ²Asst. Professor, ¹Computer Science and Engineering,

²Computer Science and Engineering,

¹Ashokrao Mane Group of Institutions, Vathar, Kolhapur, India

²Ashokrao Mane Group of Institutions Kolhapur, Vathar, India

¹deepakbpatil87@gmail.com, ²psp@amgoi.edu.in

Abstract: This review presents a comprehensive analysis of recent advancements in agricultural image classification using deep learning, emphasizing three key innovations: feature compression via autoencoders, spatial attention mechanisms, and model interpretability. Autoencoders efficiently reduce high-dimensional agricultural imagery, while attention modules like CBAM and PAM enhance spatial focus and feature refinement. YOLOv8, known for its lightweight design, is explored for crop classification tasks, with added interpretability through Grad-CAM and SHAP. Through an extensive literature survey, we compare model architectures, datasets, and performance outcomes across tasks like disease detection, crop type classification, and weed identification. The review identifies current research gaps, including the limited integration of compression and interpretability in unified frameworks. We conclude by proposing future directions toward efficient, interpretable, and real-time deployable deep learning systems for precision agriculture.

Keywords: Grad-CAM, Image classification, Agricultural classification, autoencoder

