No human required AI generated datasets

Bc. David Žahour[1]

1 Introduction

The evolution of generative artificial intelligence (AI) has opened new avenues for creating synthetic datasets, a promising solution to limitations in real-world data collection. This study explores the capabilities of Gligen (2), a generative AI platform, in creating a synthetic dataset from textual descriptions of strawberries. The resultant synthetic dataset is then utilized to train the YOLOv8 (1) neural network, a cutting-edge model for object detection, with a focus on comparative analysis of its performance with real-world data.

2 Performance on real world dataset

The Strawberry Digital Images (StrawDI)(3) dataset comprises 8,000 images collected from 20 plantations in Huelva, Spain. Captured during a full picking season. This diverse dataset provides a realistic baseline for training our YOLOv8 models for strawberry recognition and localization tasks.

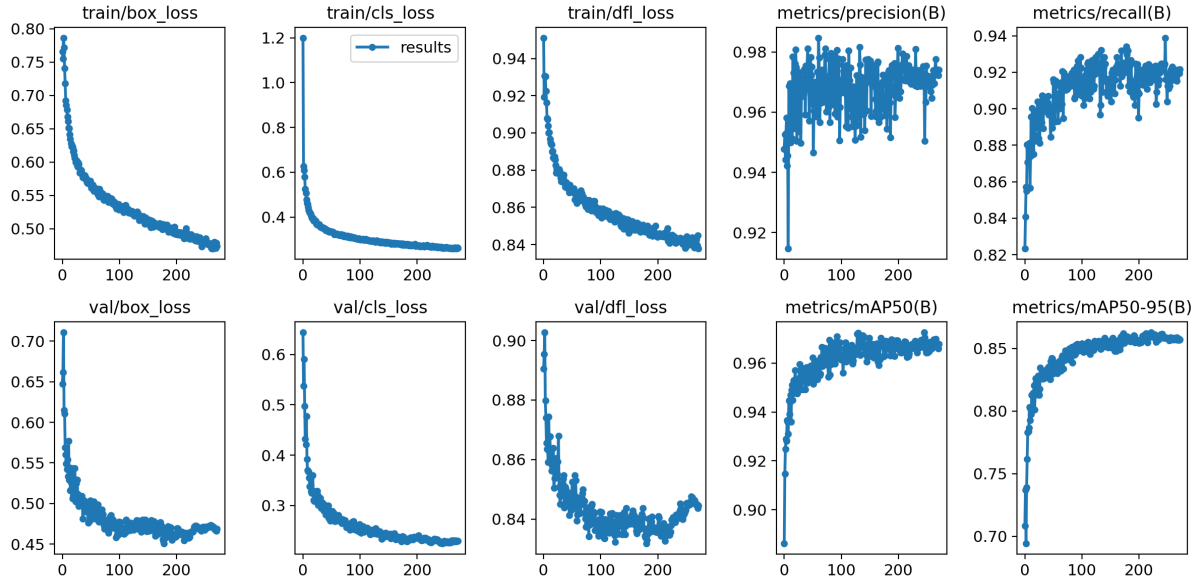

Figure 1: Results of training on StrawDI

The observed results from the YOLOv8n model implementation on the authentic StrawDI dataset set performance baseline.

[1] Student of the master degree program Applied Sciences, field of study cybernetic, e-mail: Zahour@students.zcu.cz

3 Gligen dataset

A corpus comprising 300 distinct phrases was generated via GPT chat, aimed at capturing the diversity of strawberry depictions in various contexts and orientations. With these descriptions as a foundation, the GLIGEN tool was then utilized to generate 563 synthetic images. The positions within these images were randomized, while maintaining a minimum window size of at least 50 pixels. As we can see several frames are hallucinating which will introduce error into the training.

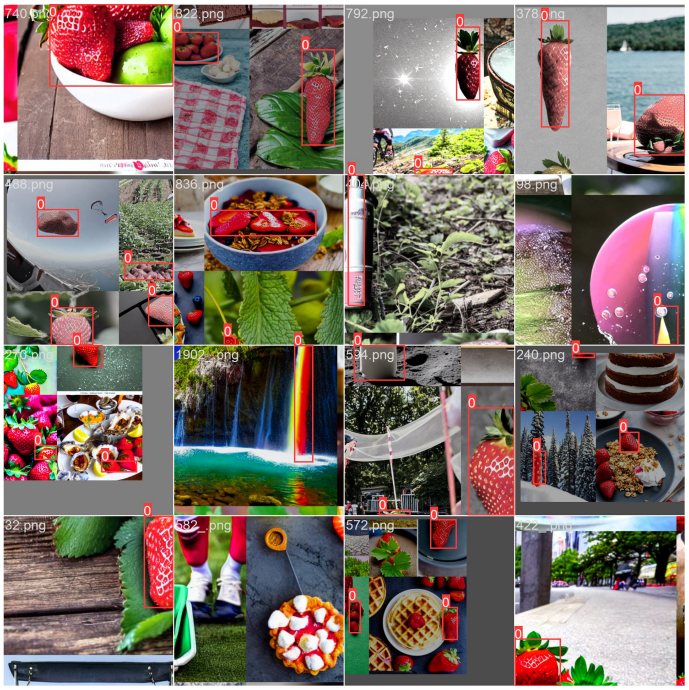

Figure 2: Images with corresponding ground truth labels.

Following a comprehensive training process spanning 468 epochs, we achieved commendable outcomes. The model showcases proficient generalization and is capable of overcoming dataset hallucinations. Regrettably, an evident deficiency in the synthetic dataset is the absence of so-called 'unripe strawberries', where detection seems to fail. In conclusion, the model demonstrated its capability to learn effectively, overcoming associated challenges. This highlights the potential of employing synthetically generated datasets for training alternative models. The primary application is in constructing datasets for specific conditions where obtaining a traditional dataset is either impossible.

References

[1] Jocher, G., Chaurasia, A. & Qiu, J. YOLO by Ultralytics. (2023,1), https://github.com/ultralytics/ultralytics

[2] Li, Y., Liu, H., Wu, Q., Mu, F., Yang, J., Gao, J., Li, C. & Lee, Y. GLIGEN: Open-Set Grounded Text-to-Image Generation. *CVPR*. (2023)

[3] Pérez-Borrero, I., Marín-Santos, D., Gegúndez-Arias, M. & Cortés-Ancos, E. A fast and accurate deep learning method for strawberry instance segmentation. *Computers And Electronics In Agriculture*. **178** pp. 105736 (2020), http://www.sciencedirect.com/science/article/pii/S0168169920300624