

Title: Agri-vision Bangladesh: A multi-crop augmented image dataset for automated disease diagnosis in Bottle Gourd, Zucchini, Papaya, and Tomato

Author Names: Md Masum Billah , Md. Anisur Rahman , Saifuddin Sagor , Sanzida Parvin , Mohammad Shorif Uddin

Abstract: This article introduces Agri-Vision Bangladesh, a comprehensive, augmented image dataset designed to advance automated disease diagnosis in four economically vital agricultural crops: Bottle Gourd (*Lagenaria siceraria*), Zucchini (*Cucurbita pepo*), Papaya (*Carica papaya*), and Tomato (*Solanum lycopersicum*). Addressing the scarcity of region-specific agricultural data, a total of 5266 original images were acquired directly from diverse agricultural fields in Bangladesh using a SONY ALPHA 7 II full-frame camera under natural lighting conditions. The dataset encompasses 28 distinct classes, covering a wide spectrum of biotic stressors including viral (Mosaic Virus, Leaf Curl), fungal (Downy Mildew, Anthracnose, Alternaria Blight), bacterial (Bacterial Blight, Xanthomonas), and pest-induced damage (Insect Hole, White Spot), alongside Healthy samples. To ensure scientific reliability, each image underwent a rigorous two-stage validation process by senior agronomists. To tackle class imbalance and facilitate the training of data-intensive Deep Learning models, the dataset was expanded using a Python-based augmentation pipeline incorporating geometric transformations (rotation, flipping) and photometric adjustments (noise, brightness) resulting in a final repository of 28,000 images (5266 original and 22,734 augmented). All files are standardized to 512×512 pixels in JPG format. This expert-validated resource serves as a critical benchmark for developing robust computer vision algorithms (e.g., CNNs, Vision Transformers) for precision agriculture, enabling research into fine-grained classification, object detection, and cross-crop transfer learning in subtropical farming environments.

Key words: Artificial intelligence Computer vision Data science Machine learning Bottle gourd leaves Papaya Leaves Zucchini Leaves Tomato Leaves

<https://doi.org/10.1016/j.dib.2026.112528https://www.sciencedirect.com/science/article/pii/S2352340926000818>

<https://www.sciencedirect.com/science/article/pii/S2352340926000818>