

Fig. 1 Major mango producing districts in Uttar Pradesh, India.

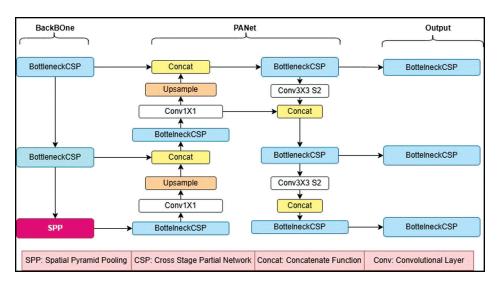


Fig. 2 YOLOv5 architecture.

June 2025] 741

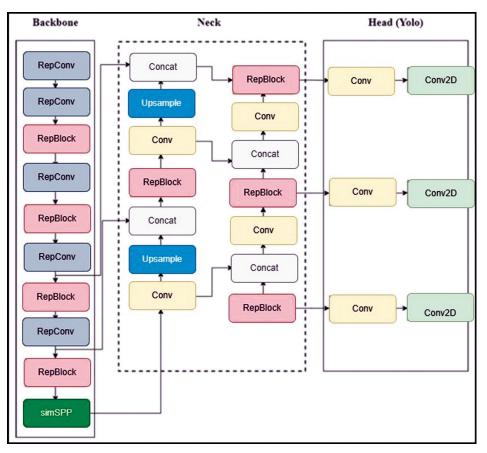


Fig. 3 YOLOv6 architecture.

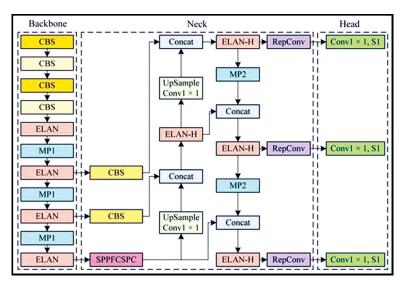


Fig. 4 YOLOv7 architecture

Note: It enhances object detection through efficient feature extraction, refined feature fusion, and optimized inference. Its backbone employs CBS (Conv-BN-SiLU) for effective feature learning, while ELAN (Efficient Layer Aggregation Network) improves gradient flow and computational efficiency by structuring feature aggregation. The inclusion of MP1 (Max Pooling) aids in spatial downsampling, and SPPFCSPC (Spatial Pyramid Pooling - Fast CSPC) enhances multi-scale feature extraction. As shown in Figure 4, the neck leverages concatenation and UpSample Conv1×1 layers for seamless multi-scale feature fusion, with ELAN-H further refining feature aggregation. MP2 (Max Pooling) ensures efficient downsampling while preserving spatial integrity. The detection head integrates RepConv (Reparameterized Convolution), streamlining multiple convolutions into a single operation for faster inference, and the final Conv1×1, S1 layer ensures accurate class predictions. Compared to YOLOv5, YOLOv7 offers superior speed and accuracy through ELAN-based feature extraction, RepConv-optimized inference, and enhanced multi-scale fusion, making it a highly efficient choice for real-time object detection.

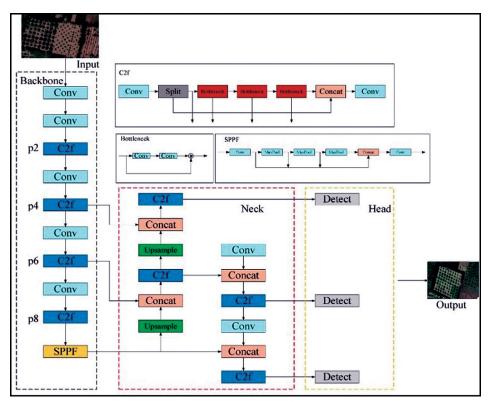


Fig. 5 YOLOv8 architecture.