***Chọn 3 phiên bản (ngẫu nhiên) YOLO và so sánh kiến trúc, ưu điểm và nhược điểm.***

| **Phiên bản** | **Kiến trúc** | **Ưu điểm** | **Nhược điểm** |
| --- | --- | --- | --- |
| **YOLOv3** | - Backbone: Darknet-53 với 53 lớp convolutional, kết hợp residual connections để tránh vanishing gradient. - Neck và Head: Multi-scale predictions tại 3 scales khác nhau (sử dụng Feature Pyramid Network - FPN), anchor-based với 3 anchor boxes mỗi cell. - Dự đoán bounding box sử dụng logistic regression cho objectness và binary cross-entropy cho class. | - Tốc độ thực thời cao (khoảng 20 FPS trên MS COCO với AP 36.2%). - Cải thiện phát hiện vật nhỏ nhờ multi-scale predictions. - Đơn giản, chỉ một neural network duy nhất cho toàn bộ quy trình. | - Đã lỗi thời, accuracy thấp hơn so với các phiên bản mới (ví dụ: kém YOLOv5 và YOLOv8). - Khó train (dùng framework Darknet ít hỗ trợ), yêu cầu bộ nhớ cao và thời gian train dài. - Khó phát hiện vật rất nhỏ do stride lớn và cấu hình anchor. |
| **YOLOv5** | - Backbone: CSPDarknet53 được chỉnh sửa, bắt đầu bằng Stem (strided convolution), kết hợp SPPF (Spatial Pyramid Pooling Fast) để tăng tốc tính toán. - Neck: Modified CSP-PAN (Path Aggregation Network). - Head: Anchor-based, decoupled cho objectness, classification và regression; hỗ trợ AutoAnchor để tự động điều chỉnh anchors. - Có 5 kích cỡ model (nano, small, medium, large, extra-large). | - Dễ sử dụng và train với PyTorch, tốc độ inference cao (lên đến 200 FPS trên NVIDIA V100). - Cộng đồng hỗ trợ mạnh, tích hợp tốt cho mobile (iOS/Android) và edge devices. - Accuracy tốt (AP 50.7% trên COCO), với augmentations nâng cao như Mosaic và MixUp. | - Accuracy thấp hơn YOLOv8 (ví dụ: mAP 37.4% cho model small so với 44.9% của YOLOv8). - Cần tune anchors thủ công cho dataset đặc biệt. - Trade-off giữa tốc độ và accuracy ở model nhỏ. |
| **YOLOv8** | - Backbone: Modified CSPDarknet53 với C2f modules (cross-stage partial bottleneck) để fusion features hiệu quả hơn. - Neck: PANet với SPPF. - Head: Anchor-free, decoupled cho objectness, classification và regression; hỗ trợ multi-task (detection, segmentation, pose estimation, classification). - Có 5 kích cỡ model tương tự YOLOv5. | - Accuracy vượt trội (AP 53.9% trên COCO, cao hơn YOLOv5 3.2%), tốc độ cạnh tranh (280 FPS trên NVIDIA A100). - Anchor-free giảm complexity và hyperparameters, dễ train hơn. - Hỗ trợ nhiều task, tích hợp CLI/Python API và Ultralytics HUB cho no-code training. | - Model lớn (large/extra-large) yêu cầu tài nguyên tính toán cao, không phù hợp edge devices hạn chế. - Tăng complexity do multi-task, có thể làm chậm inference ở một số trường hợp. |