THE EVOLUTION OF YOLO: "VERSIONS 6-8"

TODAY'S AGENDA

OBJECT DETECTION BEFORE YOLO

The traditional approach to object detection before YOLO was :

- computationally expensive.
- required many iterations
- relied on region-based classifiers to locate areas but more robust classifiers to identify objects.
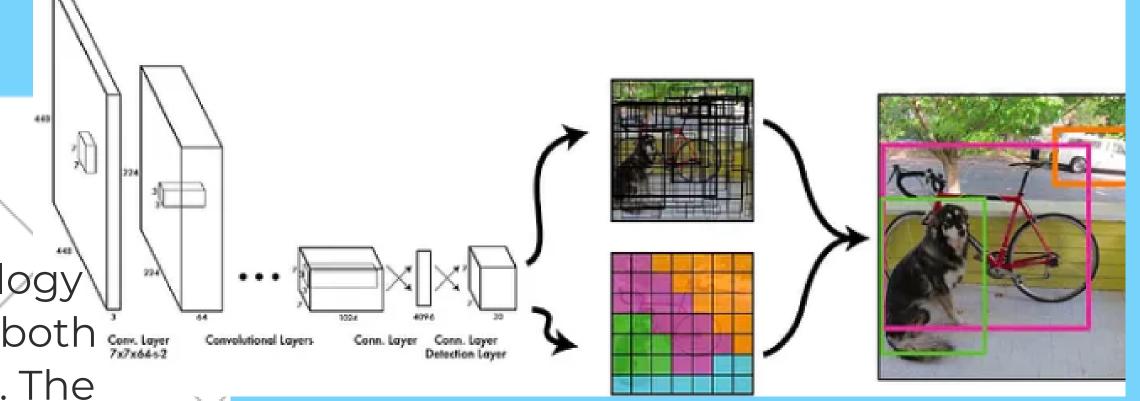
HOW DOES YOLO WORK

YOLO introduced a new methodology for object detection that conducts both stages in the same neural network. The image is divided into cells of equal dimensions, and each cell detects and locates objects within that cell using bounding box coordinates, object labels, and probabilities of the object's presence.

Because YOLO's approach involves each cell processing the grid independently, it reduces the required computing power and time needed for training and inference.

THE EVOLUTION OF YOLO

YOLOV1 (JUN, 2015):

UNIFIED, REAL-TIME OBJECT DETECTION

YOLOV2 (DEC, 2016)

BETTER, FASTER, STRONGER

YOLOV3 (APR, 2018):

AN INCREMENTAL IMPROVEMENT

YOLOV4 (APR, 2020)

OPTIMAL SPEED AND ACCURACY OF OBJECT DETECTION

YOLOV5 (MAY, 2020)

YOLO V6

- IMPROVED ACCURACY: YOLOV6 ACHIEVES HIGHER MEAN AVERAGE PRECISION (MAP) THAN PREVIOUS VERSIONS OF YOLO.

- FASTER INFERENCE TIMES: YOLOV6 IS OPTIMIZED FOR SPEED, WITH SIGNIFICANTLY FASTER INFERENCE TIMES THAN PREVIOUS VERSIONS.

- NEW MODELS: YOLOV6 INCLUDES SEVERAL NEW MODELS, SUCH AS YOLOV6-NANO AND YOLOV6-S, WHICH ACHIEVE HIGH ACCURACY AND FAST INFERENCE TIMES ON DIFFERENT HARDWARE.

- IMPROVED ARCHITECTURE: YOLOV6 INCLUDES ARCHITECTURAL IMPROVEMENTS, SUCH AS CROSS-STAGE PARTIAL CONNECTIONS AND SPP-ATTENTION MODULES, TO IMPROVE ACCURACY AND SPEED.

- NEW LOSS FUNCTION: YOLOV6 INTRODUCES A NEW LOSS FUNCTION, DIOU-NMS, WHICH IMPROVES THE ACCURACY OF OBJECT DETECTION.

- BETTER TRAINING DATA: YOLOV6 IS TRAINED ON A LARGER AND MORE DIVERSE DATASET, WHICH HELPS IMPROVE ITS ACCURACY.

- SUPPORT FOR CUSTOM DATASETS: YOLOV6 INCLUDES TOOLS FOR CREATING AND TRAINING CUSTOM OBJECT DETECTION DATASETS.

YOLO V7

Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors
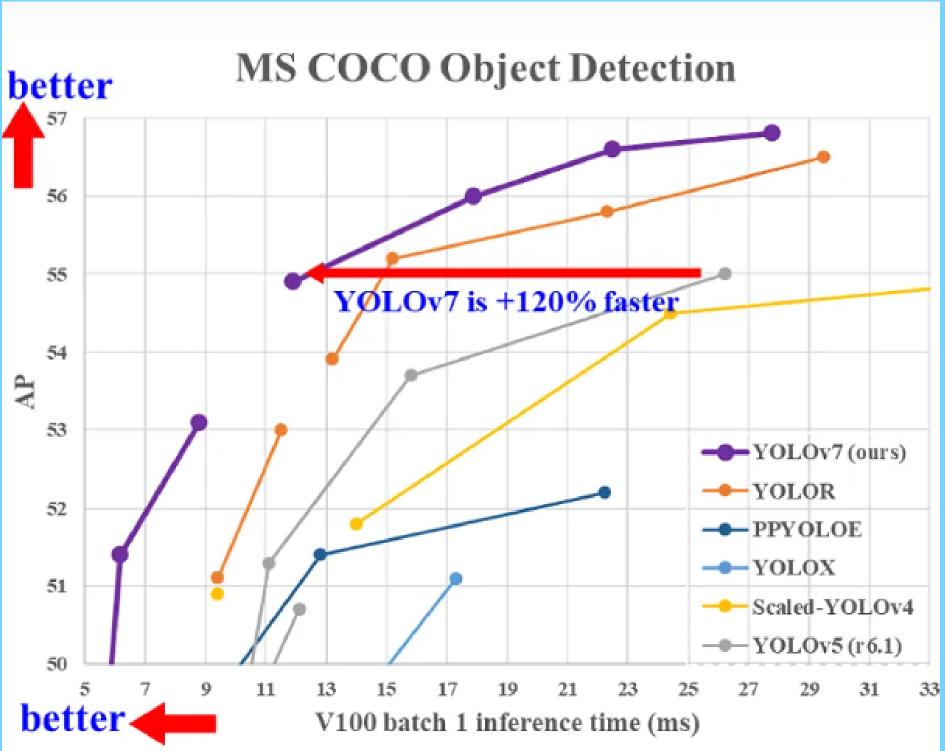
Significance of YOLOv7

surpasses all known object detectors in both speed and accuracy

Asma Ahmd

5 FPS to 160 FPS

56.8% AP

50% REDUCTION IN COST

75% fewer parameters and 36% lesser computional time and 1.5 times higher AP than YOLO v4

YOLO V7 COMPARISON

- YOLOR
- YOLOX
- Scaled-YOLOv4
- YOLOv5
- DETR
- Deformable DETR
- DINO-5scale-R50
- ViT-Adapter-B

MS COCO Object Detection

better

AP

V100 batch 1 inference time (ms)

better

- YOLOv7 (ours)
- YOLOR
- PPYOLOE
- YOLOX
- Scaled-YOLOv4
- YOLOv5 (r6.1)

YOLOv7 is +120% faster

YOLOV7 ARCHITECTURE

- THE YOLOV7 ARCHITECTURE IS BASED ON PREVIOUS YOLO MODEL
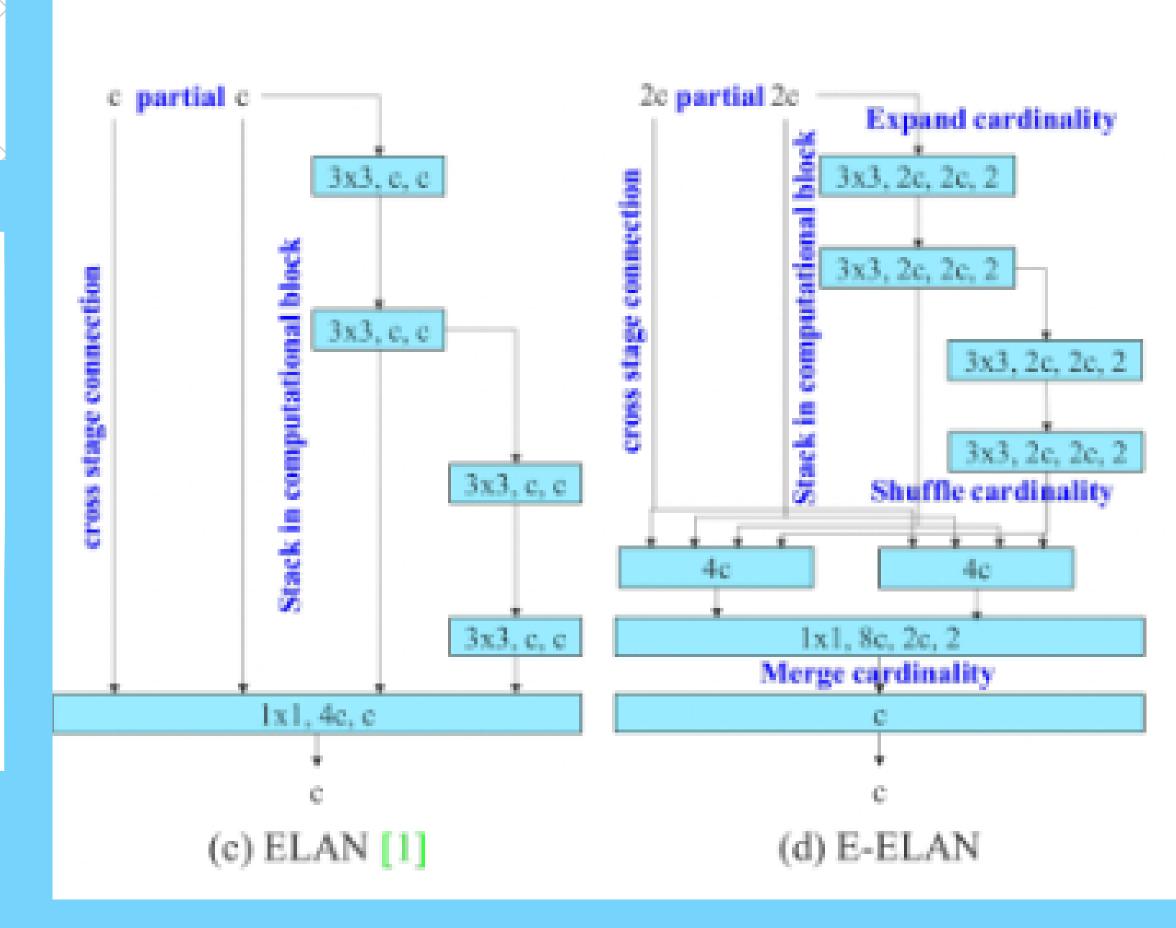
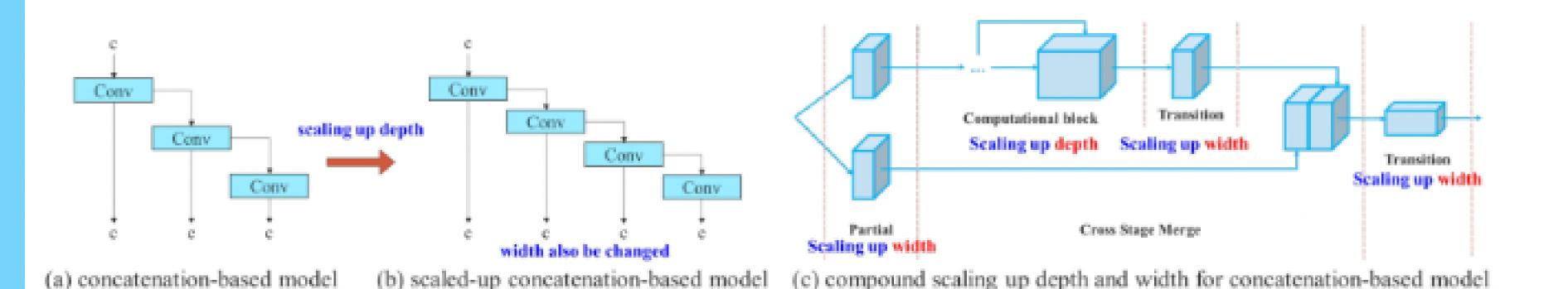ARCHITECTURES, NAMELY YOLOV4, SCALED YOLOV4, AND

YOLO-R.

.

EXTENDED EFFICIENT LAYER AGGREGATION NETWORK (E-ELAN)

- **The computational block in the YOLOv7 backbone.**

- **that uses "expand, shuffle, merge cardinality" to improve the model's learning ability without damaging the original gradient path.**

(c) ELAN [1]

(d) E-ELAN

MODEL SCALING

OBJECT DETECTION MODELS COME IN DIFFERENT SIZES TO MEET THE REQUIREMENTS OF VARIOUS APPLICATIONS. **YOLOV7 SCALES** NETWORK **DEPTH AND WIDTH TOGETHER** WHILE CONCATENATING LAYERS TO MAINTAIN OPTIMAL ARCHITECTURE FOR DIFFERENT SIZES.
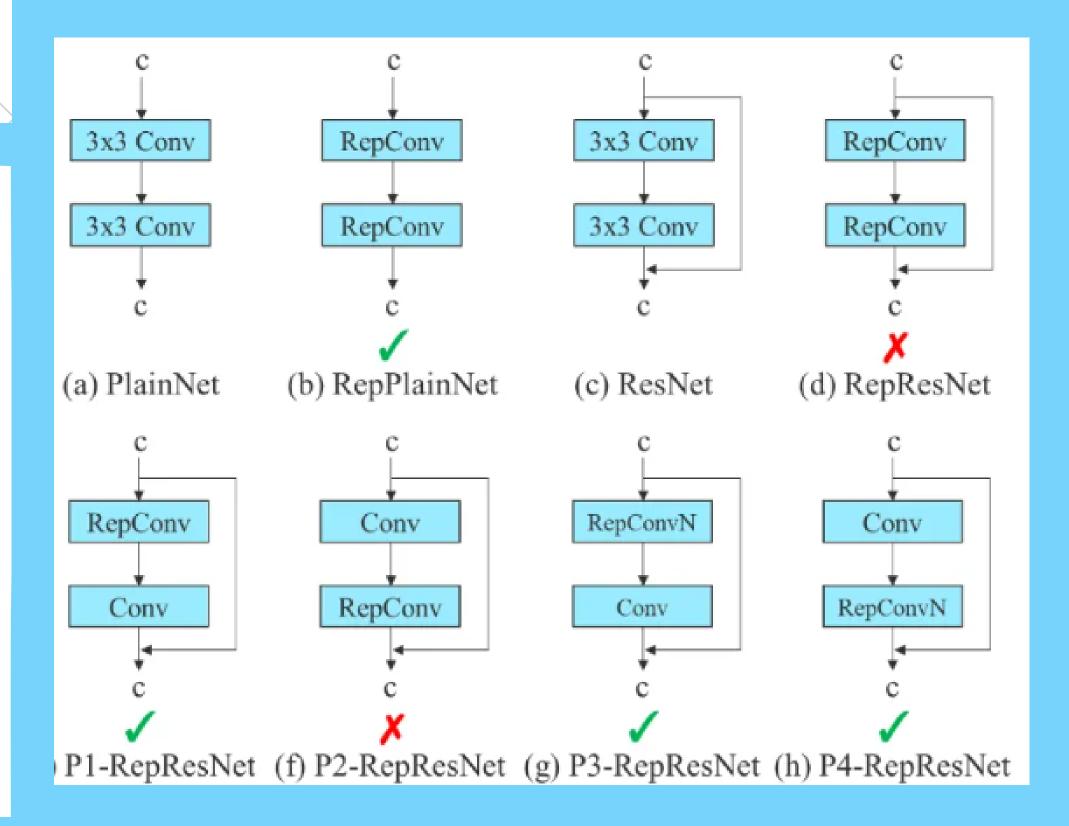

Compound scaling in YOLOv7 model sizes

SCALING FACTOR

1. Resolution
2. Depth
3. Width
4. Stage

MODEL RE-PARAMETERIZATION

RE-PARAMETERIZATION TECHNIQUES INVOLVE AVERAGING MODEL WEIGHTS TO IMPROVE THE MODEL'S ROBUSTNESS.
YOLOV7 USES GRADIENT FLOW PROPAGATION PATHS TO DETERMINE WHICH NETWORK MODULES SHOULD USE RE-PARAMETERIZATION STRATEGIES.

(a) PlainNet (b) RepPlainNet (c) ResNet (d) RepResNet

P1-RepResNet (f) P2-RepResNet (g) P3-RepResNet (h) P4-RepResNet

YOLOV8

YOLOV8 IS THE LATEST STATE-OF-THE-ART YOLO MODEL DEVELOPED BY ULTRALYTICS FOR OBJECT DETECTION, IMAGE CLASSIFICATION, AND INSTANCE SEGMENTATION TASKS. IT INCLUDES ARCHITECTURAL AND DEVELOPER EXPERIENCE IMPROVEMENTS OVER YOLOV5 AND IS ACTIVELY DEVELOPED WITH COMMUNITY FEEDBACK TO MAKE IT THE BEST IT CAN BE. ULTRALYTICS PROVIDES LONG-TERM SUPPORT FOR THEIR MODELS, MAKING THEM ACCESSIBLE TO A WIDE RANGE OF DEVELOPERS.

WHY SHOULD I USE YOLOV8?

1.YOLOv8 has a high rate of accuracy measured by COCO and Roboflow 100.

2. YOLOv8 comes with a lot of developer-convenience features, from an easy-to-use CLI to a well-structured Python package.

3.There is a large community around YOLO and a growing community around the YOLOv8 model, meaning there are many people in computer vision circles who may be able to assist you when you need guidance.

' THERE IS CURRENTLY NO PUBLISHED PAPER FOR YOLOV8, SO INFORMATION ON ITS RESEARCH METHODOLOGY AND ABLATION STUDIES IS LIMITED. '