

# A JOURNEY THROUGH YOLO VERSIONS FOR AIRCRAFT DETECTION IN SATELLITE IMAGERY

## GROUP 8

Nguyễn Tiến Đạt	20225481
Phan Minh Hòa	20225495
Nguyễn Võ Ngọc Khuê	20225504
Nguyễn Hoàng Quân	20225519
Nguyễn Hải Đăng	20225478

# TABLE OF CONTENTS

01

INTRODUCTION

02

METHODOLOGY

03

RESULTS

04

CONCLUSION

**01**

# **INTRODUCTION**

# PROBLEM DESCRIPTION

**Aircraft detection** from satellite imagery is a **critical task** in various fields

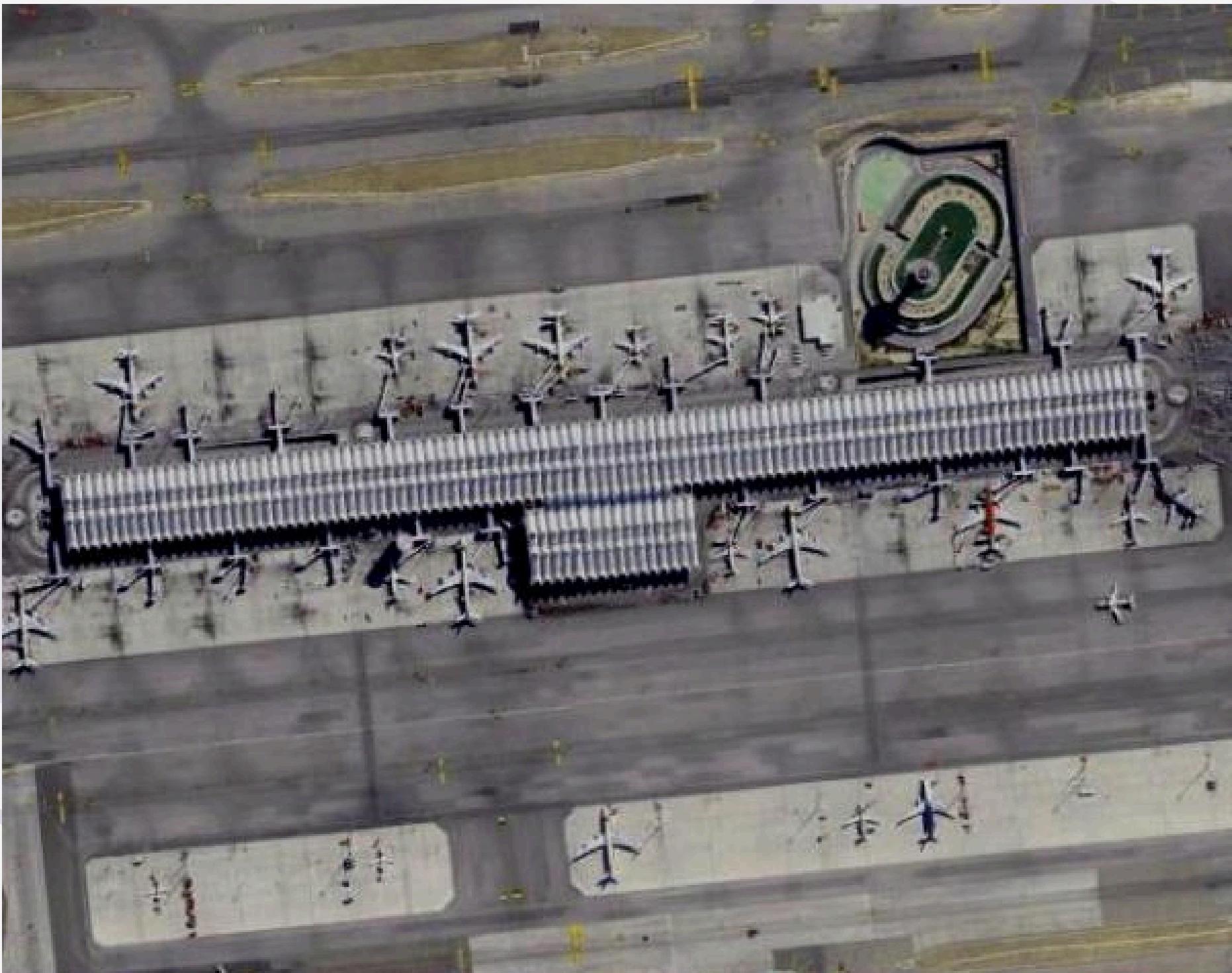
**Traditional methods** has **limitations**

**Challenge:** Aircrafts in satellite imagery are:

- small
- densely packed
- overlap with backgrounds

**Main objective of the project:**

- evaluate the performance of YOLO on the GDIT dataset
- highlight advancements in detection accuracy and efficiency



# GDIT DATASET

## **GDIT Aerial Airport Dataset:**

- 708 images
- 600px x 600px
- 14,335 labelled aircrafts

**One class only:** "aircraft"

## **Dataset complexity:**

- various aircraft types
- various sizes
- density
- overlapping
- partial obscure



an example from GDIT dataset

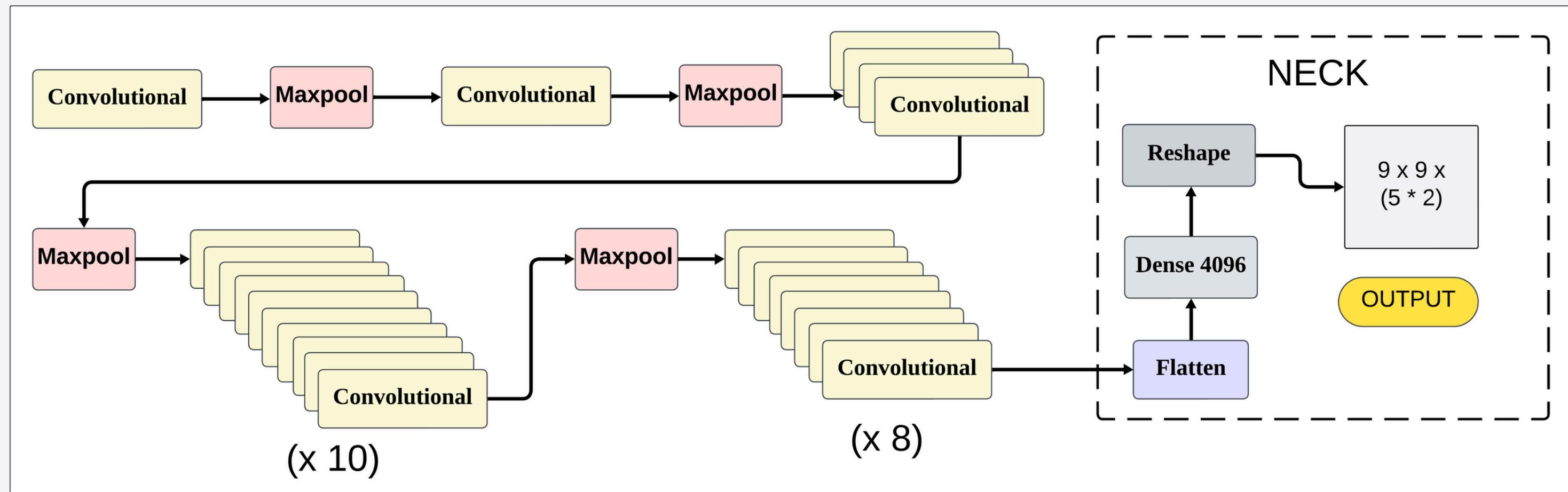
**02**

# **METHODOLOGY**

# 1. YOLOV1

## Key Innovations:

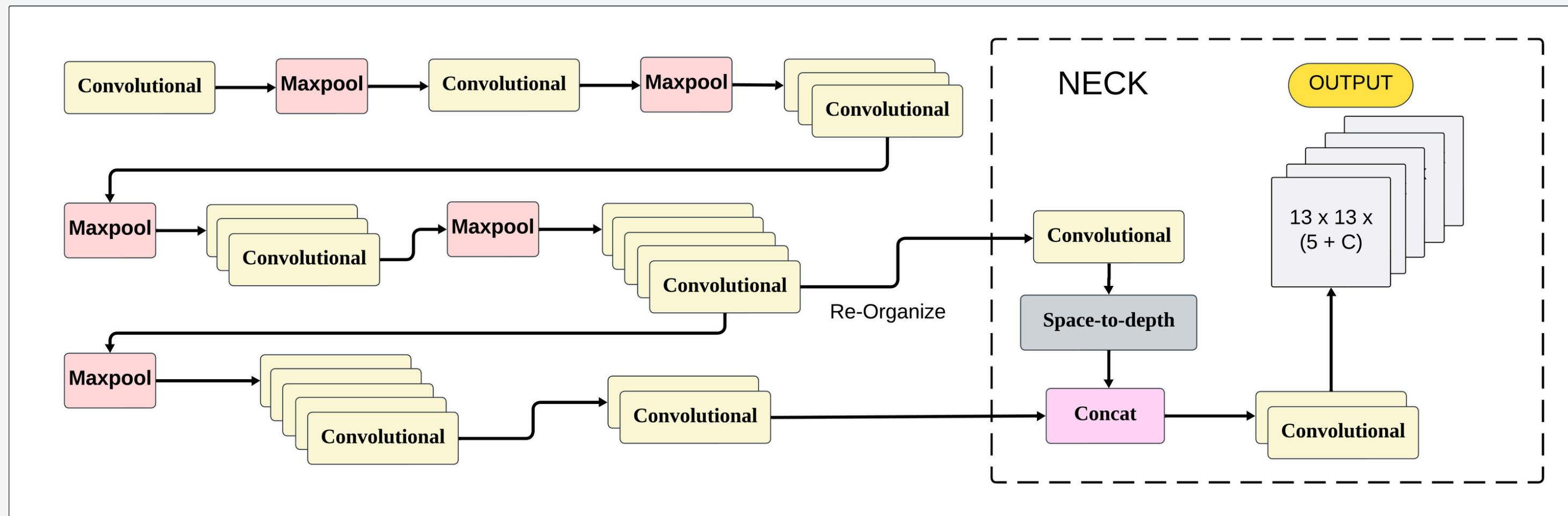
- **Pioneer** the use of a single neural network for **real-time object detection**
- Combine **localization** and **classification** in one streamlined process.
- Achieved **45 FPS** on a single GPU, setting a benchmark.



# 2. YOLOV2

## Key Innovations:

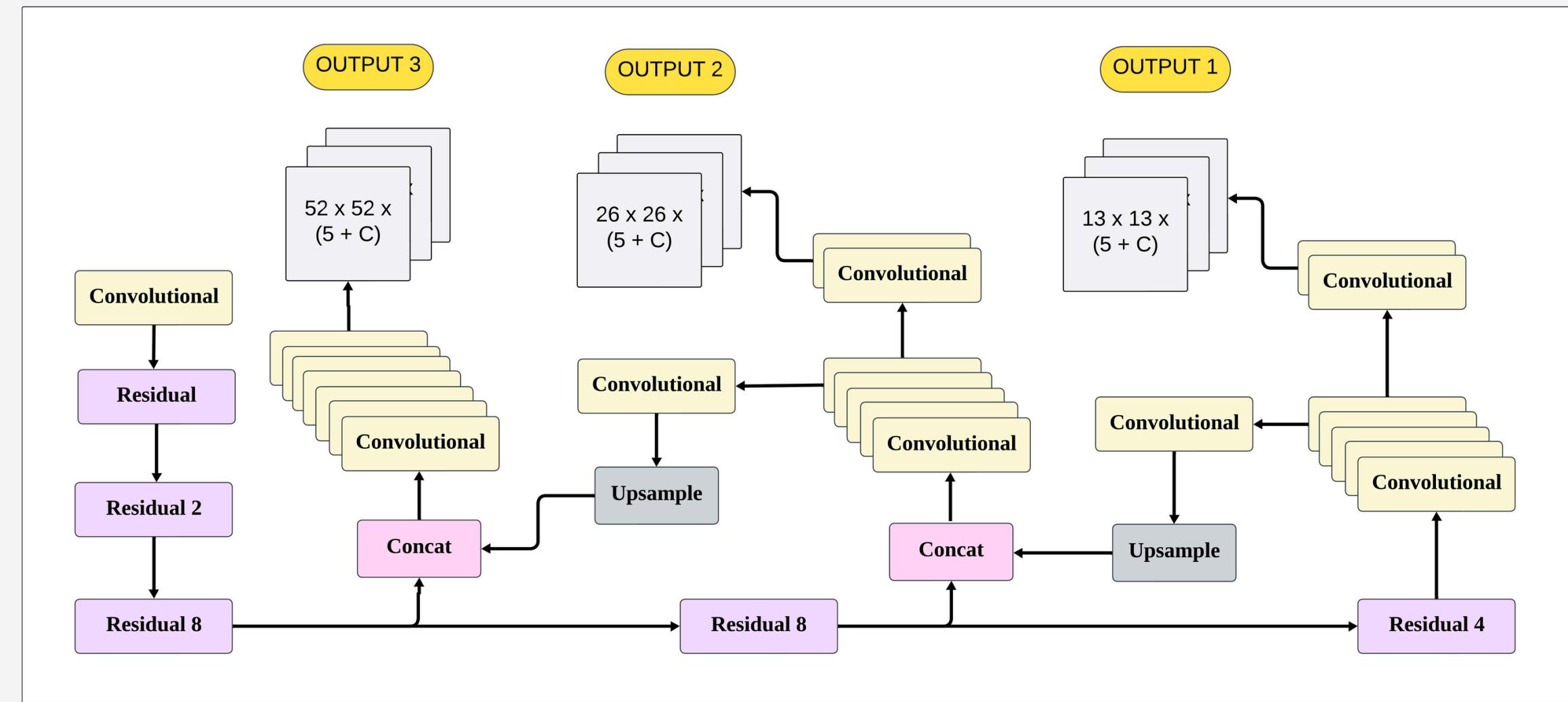
- **Anchor boxes** for handling objects of varying sizes and aspect ratios.
- **Multi-scale training** enabling different input resolutions.
- **Hierarchical classification system** highlights relationships between classes.



# 3. YOLOV3

## Key Innovations:

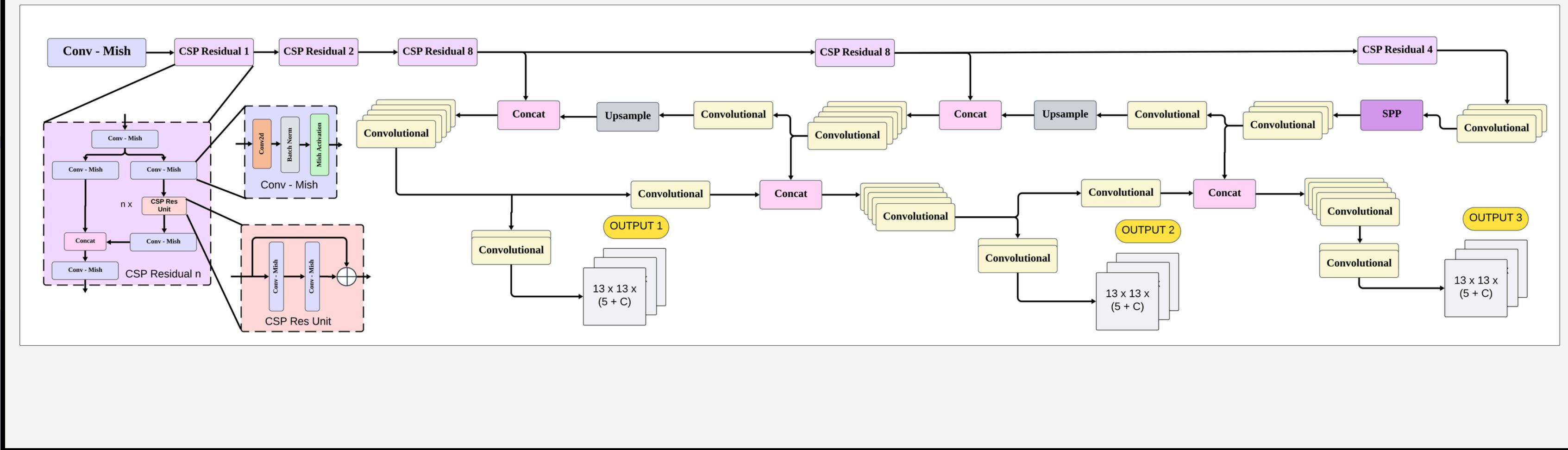
- **Multi-scale predictions** at three levels improved performance across diverse object sizes.
- **Upgraded backbone**, Darknet-53, combined with enhanced feature extraction, increased accuracy.



# 4. YOLOV4

## Key Innovations:

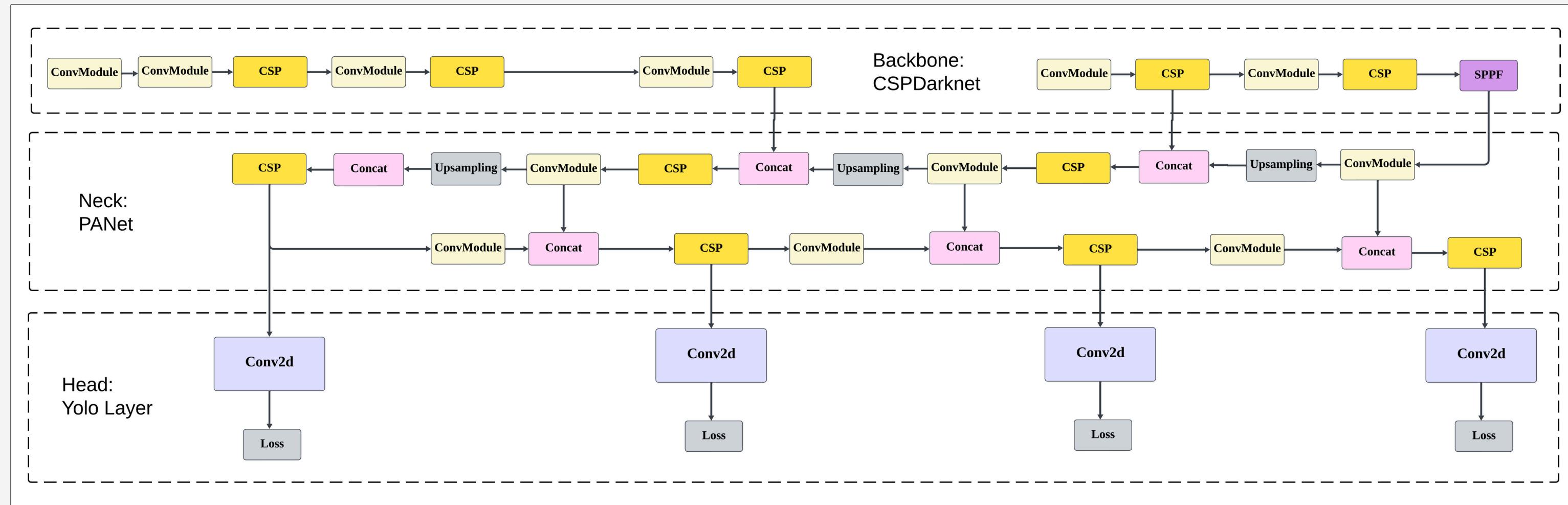
- Stronger backbone
- SPP/PANet structure for feature extraction.
- Data augmentation (Mosaic)
- New loss function (CIoU)



# 5. YOLOV5

## Key Innovations:

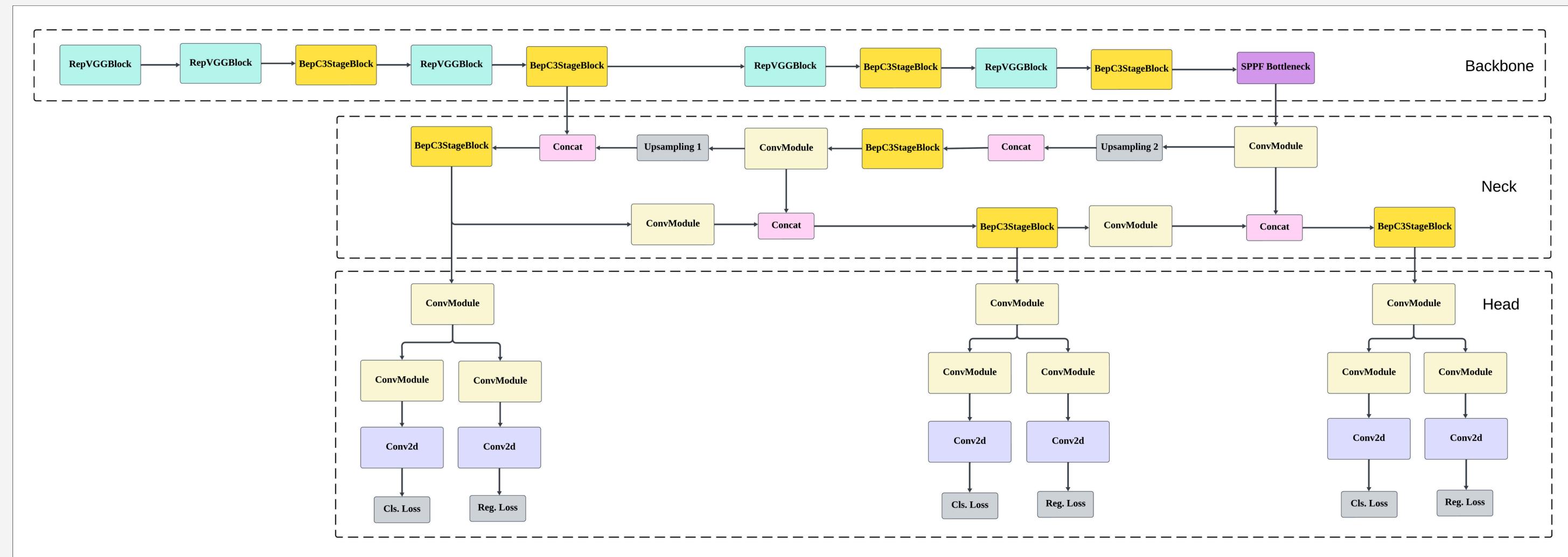
- **Lightweight and modular design:** C3 modules and the SPPF block.
- **Enhanced data augmentation techniques and anchor optimization.**
- **Loss scaling combined with EMA weighting**



# 6. YOLOV6

## Key Innovations:

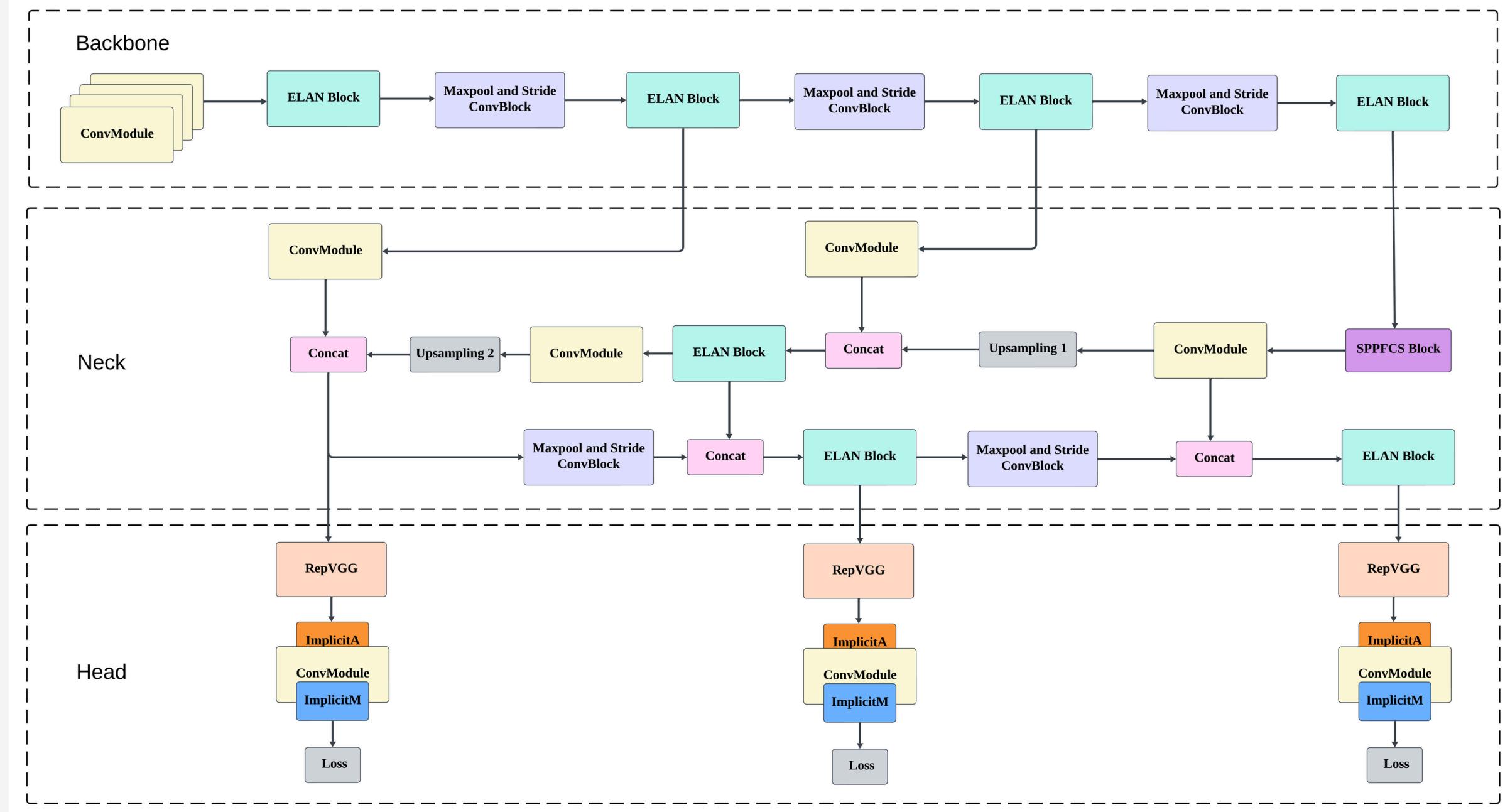
- **Advanced Label Assignment** aligns classification and regression tasks.
- **Anchorless design** enhances flexibility and inference speed.
- **Further optimizations** in the pipeline's complexity.



# 7. YOLOV7

## Key Innovations:

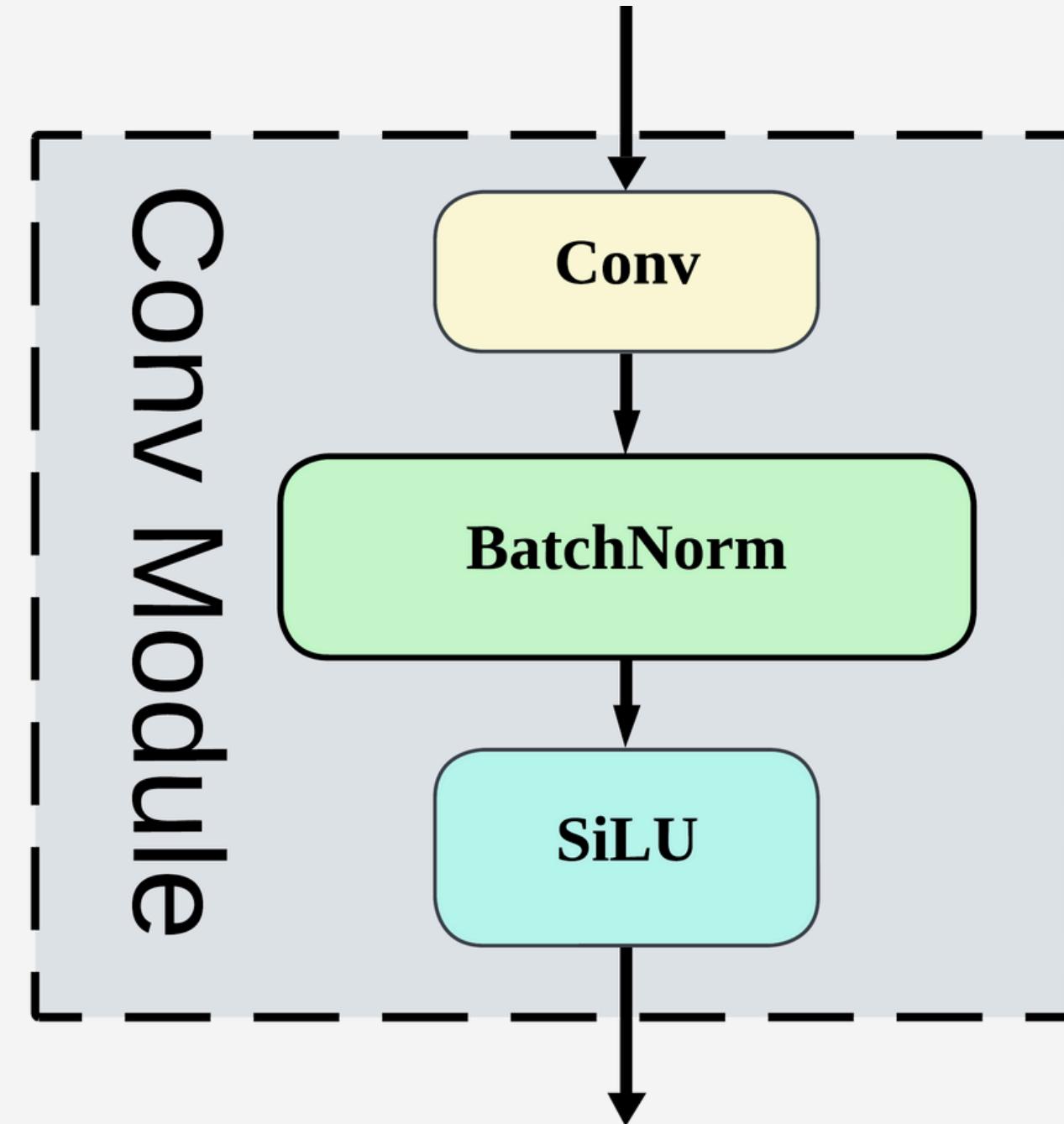
- **Innovative architectures:**
  - E-ELAN
  - Auxiliary supervision
  - SimOTA label assignment
  - RepConv blocks.
- **Unprecedented** efficiency and accuracy in object detection, segmentation, and pose estimation tasks.



# 8. YOLOV8

## Key Components:

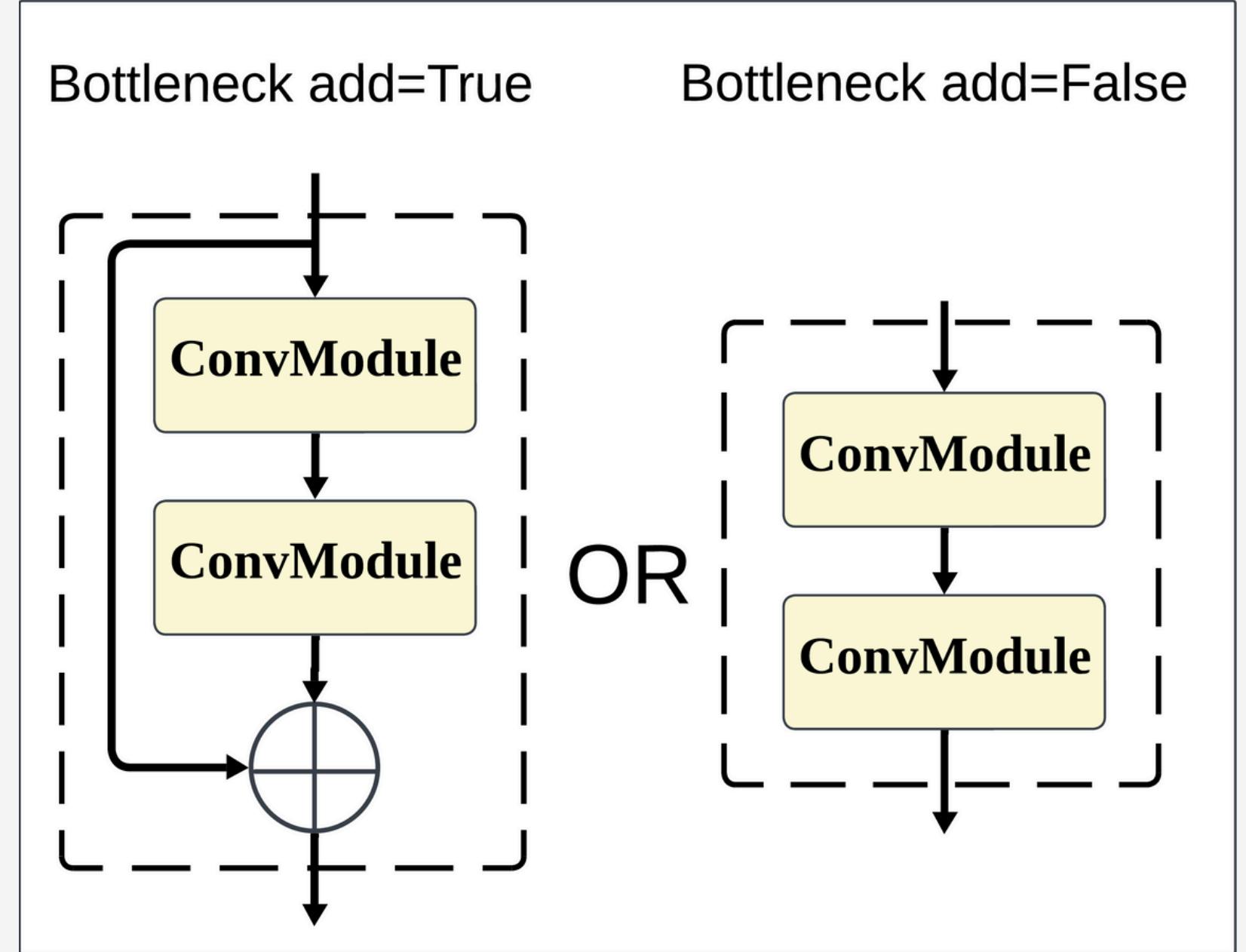
- **Convolutional Block:**
  - Convolution
  - Batch normalization
  - SiLU Activation



# 8. YOLOV8

## Key Components:

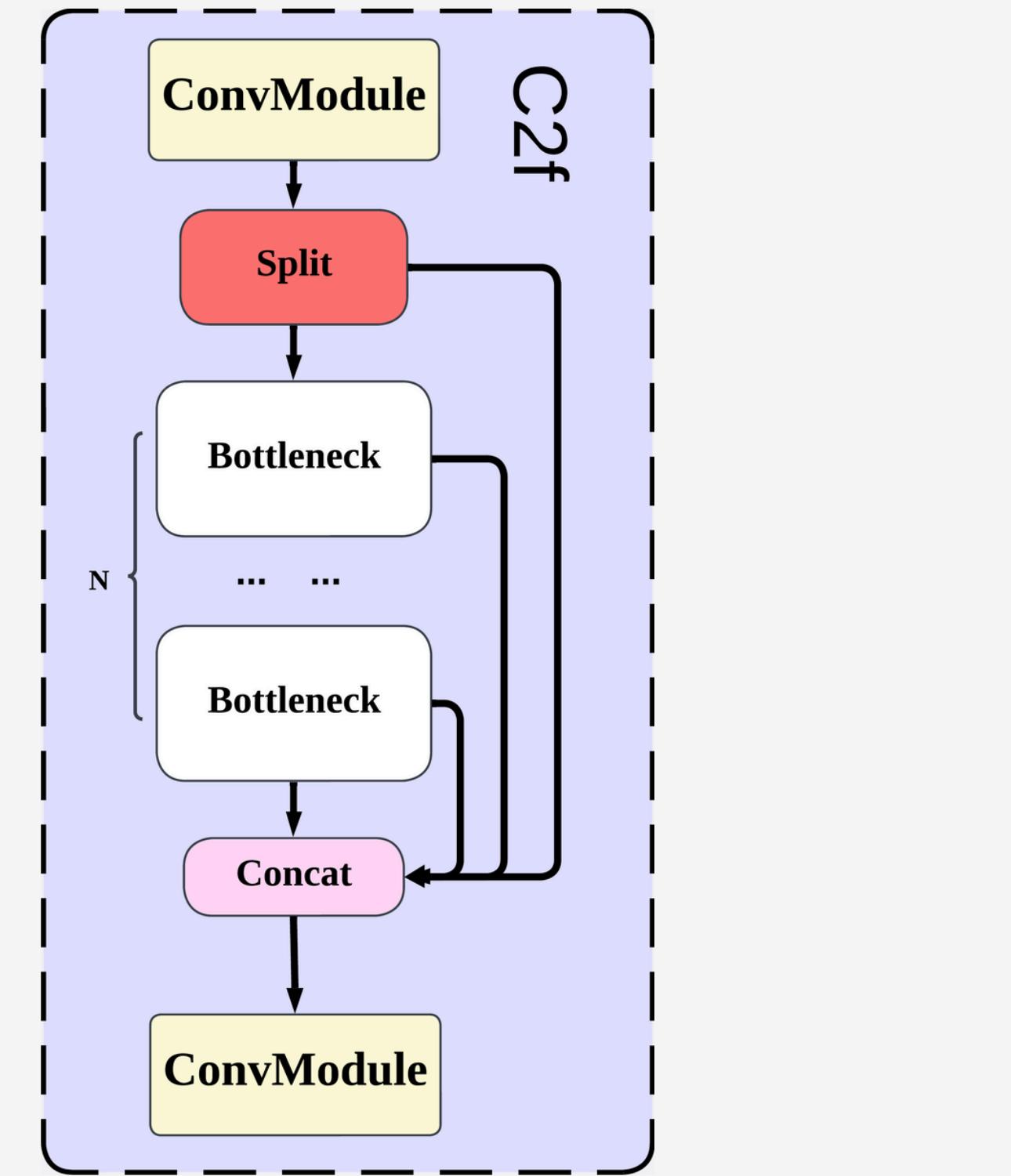
- **Convolutional Block:**
  - Convolution
  - Batch normalization
  - SiLU Activation
- **Bottleneck Block:** shortcut connections



# 8. YOLOV8

## Key Components:

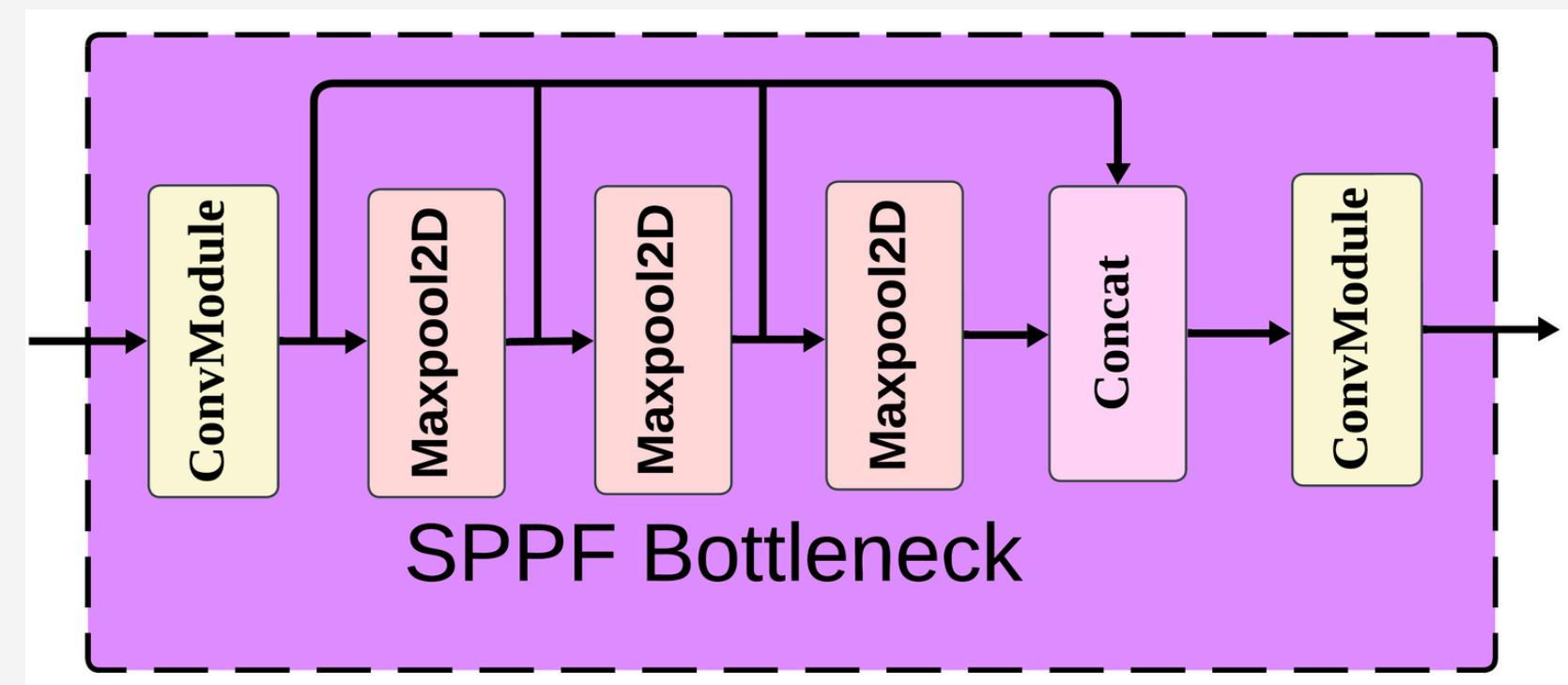
- **Convolutional Block:**
  - Convolution
  - Batch normalization
  - SiLU Activation
- **Bottleneck Block:** shortcut connections
- **C2f Block:** Combines bottleneck blocks and convolution layers



# 8. YOLOV8

## Key Components:

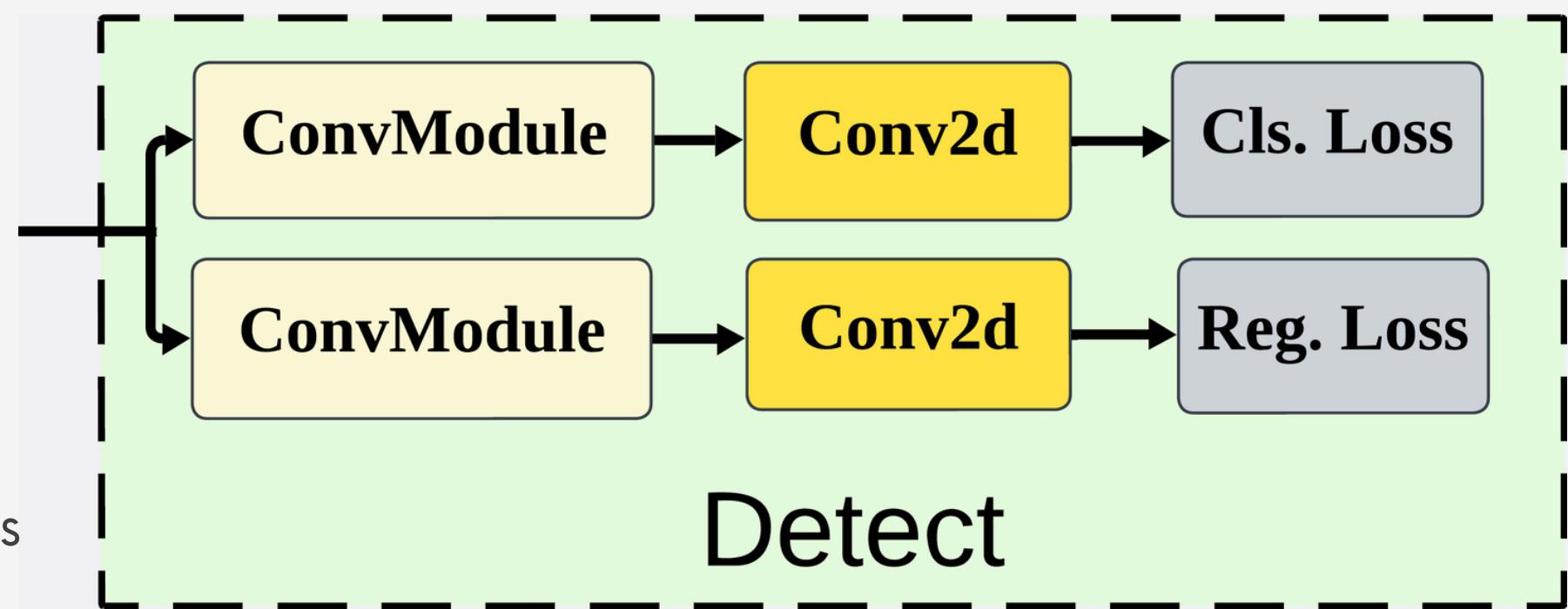
- **Convolutional Block:**
  - Convolution
  - Batch normalization
  - SiLU Activation
- **Bottleneck Block:** shortcut connections
- **C2f Block:** Combines bottleneck blocks and convolution layers
- **Spatial Pyramid Pooling Fast (SPPF):** Enables multi-scale feature extraction



# 8. YOLOV8

## Key Components:

- **Convolutional Block:**
  - Convolution
  - Batch normalization
  - SiLU Activation
- **Bottleneck Block:** shortcut connections
- **C2f Block:** Combines bottleneck blocks and convolution layers
- **Spatial Pyramid Pooling Fast (SPPF):** Enables multi-scale feature extraction
- **Detect Block:** revolutionary anchor-free design



# 8. YOLOV8

## Key Innovations:

- **Mosaic Augmentation:** improve robustness stitches multiple images into one
- **Task-Aligned Assigner (TAA):** better matching between predicted and ground truth boxes by balancing classification and regression scores.



# 8. YOLOV8

## Key Innovations:

- **Loss Functions:**

- **Distribution Focal Loss (DFL):**

$$DFL(S_n, S_{n+1}) = -[(y_{n+1} - y) \log S_n + (y - y_n) \log S_{n+1}]$$

where  $S_n = \frac{y_{n+1} - y}{y_{n+1} - y_n}$  and  $S_{n+1} = \frac{y - y_n}{y_{n+1} - y_n}$

- **Complete IoU (CIoU) Loss:**

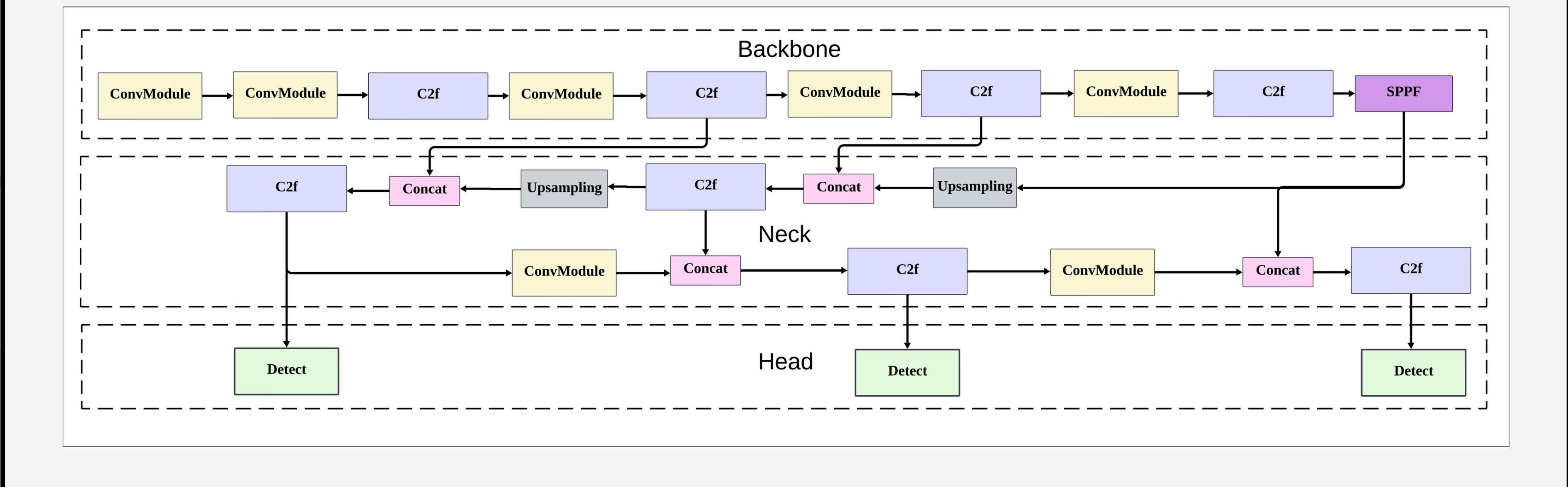
$$CIoU = 1 - IoU + \frac{d^2}{c^2} + \alpha \frac{v}{1 - IoU + v}$$

where  $v = \frac{4}{\pi^2} \left( \arctan \frac{w_{gt}}{h_{gt}} - \arctan \frac{w_p}{h_p} \right)^2$

# 8. YOLOV8

## Key Innovations:

- **Anchor-Free Design**
- **Mosaic Augmentation:** stitches multiple images into one
- **Task-Aligned Assigner (TAA):** balances classification and regression scores.
- Loss Functions: **Distribution Focal Loss (DFL)** and **Complete IoU (CIoU) Loss**





# **03 RESULT**

# 1. EVALUATION METRICS

- **Precision:**  $P = \frac{TP}{TP + FP}$

- **Recall:**  $R = \frac{TP}{TP + FN}$

- **Average Precision:**

$$AP = \int_0^1 P(R) dR$$

- **Mean Average Precision:**

$$mAP = \frac{1}{n} \sum_{i=1}^n AP_i$$

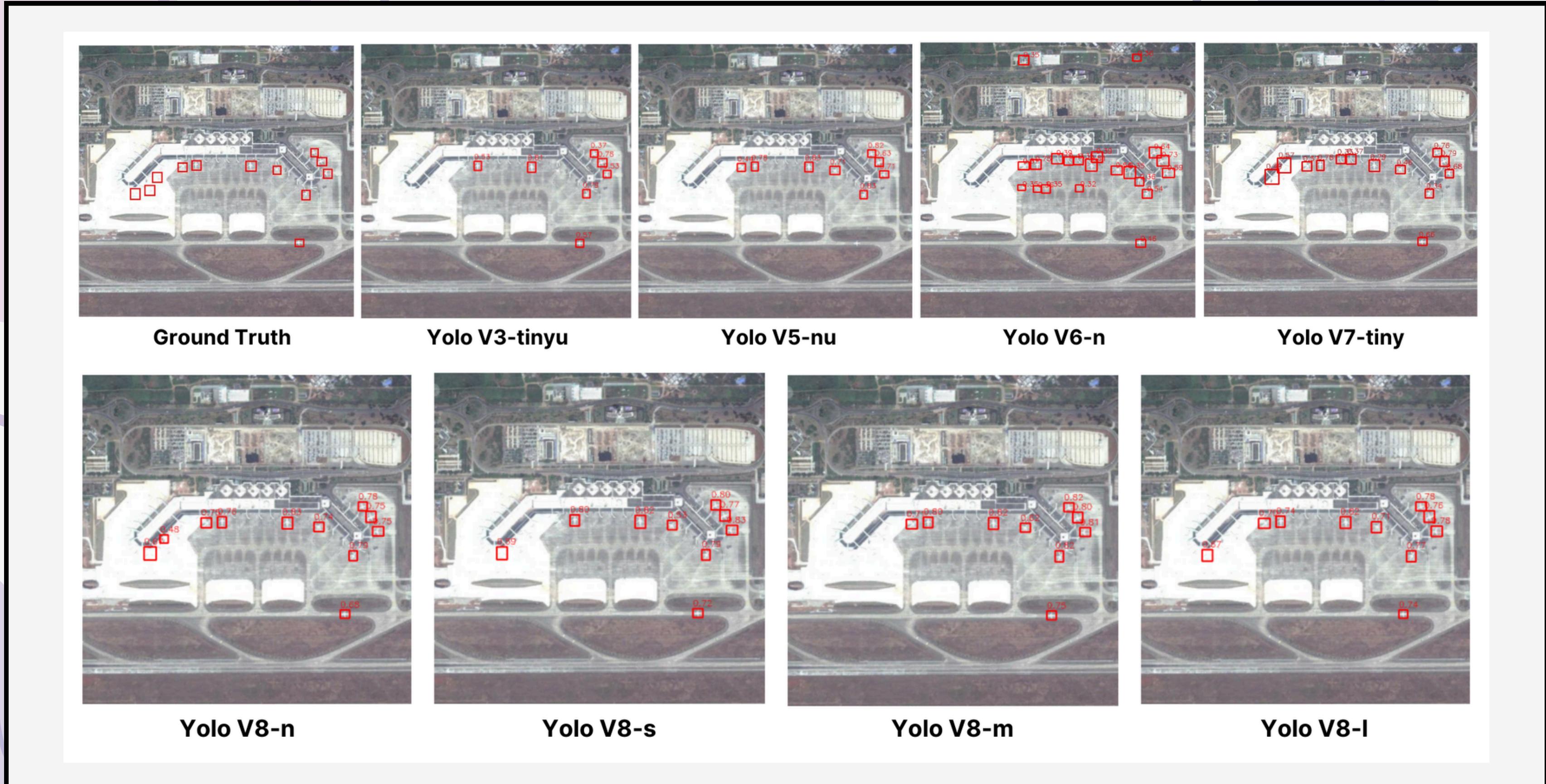
**mAP@0.5**

**mAP@0.5:0.95**

# 2. RESULT & COMPARISON

	Parameter	Precision	Recall	mAP@0.5	mAP@0.5:0.95
<b>YOLOv3-tiny</b>	8.7M	0.898	0.807	0.877	0.462
<b>YOLOv5-n</b>	2.5M	0.917	0.860	0.922	0.513
<b>YOLOv6-n</b>	4.3M	0.877	0.79	0.854	0.448
<b>YOLOv7-tiny</b>	6M	0.918	0.854	0.893	0.458
<b>YOLOv8-n</b>	3.2M	0.925	0.863	0.932	0.519
<b>YOLOv8-s</b>	11.2M	0.923	0.889	0.934	0.538
<b>YOLOv8-m</b>	25.9M	0.933	0.900	0.945	0.545
<b>YOLOv8-l</b>	43.7M	0.944	0.914	0.950	0.552

# 2. RESULT & COMPARISON



**04**

# **CONCLUSION**



**THANK YOU**