

```

yolov12.yaml
# YOLOv12 🚀, AGPL-3.0 license
# YOLOv12 object detection model with P3-P5 outputs.

# Parameters
nc: 1 # number of classes
scales: # model compound scaling constants,
# [depth, width, max_channels]
n: [0.50, 0.25, 1024]
s: [0.50, 0.50, 1024]
m: [0.50, 1.00, 512]
l: [1.00, 1.00, 512]
x: [1.00, 1.50, 512]

# YOLO12n backbone
backbone:
# [from, repeats, module, args]
- [-1, 1, Conv, [64, 3, 2]] # 0-P1/2
- [-1, 1, Conv, [128, 3, 2]] # 1-P2/4
- [-1, 2, C3k2, [256, False, 0.25]] #2
- [-1, 1, Conv, [256, 3, 2]] # 3-P3/8
- [-1, 2, C3k2, [512, False, 0.25]] #4
- [-1, 1, Conv, [512, 3, 2]] # 5-P4/16
- [-1, 4, A2C2f, [512, True, 4]] #6
- [-1, 1, Conv, [1024, 3, 2]] # 7-P5/32
- [-1, 4, A2C2f, [1024, True, 1]] # 8

# YOLO12n head
head:
- [-1, 1, nn.Upsample, [None, 2, "nearest"]] #9
- [[-1, 6], 1, Concat, [1]] # cat backbone P4 10
- [-1, 2, A2C2f, [512, False, -1]] # 11

- [-1, 1, nn.Upsample, [None, 2, "nearest"]] #12
- [[-1, 4], 1, Concat, [1]] # cat backbone P3 13
- [-1, 2, A2C2f, [256, False, -1]] # 14

- [-1, 1, Conv, [256, 3, 2]] #15
- [[-1, 11], 1, Concat, [1]] # cat head P4 16
- [-1, 2, A2C2f, [512, False, -1]] # 17

- [-1, 1, Conv, [512, 3, 2]] #18
- [[-1, 8], 1, Concat, [1]] # cat head P5 19
- [-1, 2, C3k2, [1024, True]] # 20 (P5/32-large)

- [[14, 17, 20], 1, Detect, [nc]] # Detect(P3, P4, P5)

```

Yolov12.yaml

```

yolov12_P2.yaml
# YOLOv12-P2 (Drone Optimized)
# YOLOv12 object detection model with P2-P5 outputs.
nc: 1

scales:
n: [0.50, 0.25, 1024]

# ----- BACKBONE -----
backbone:
- [-1, 1, Conv, [64, 3, 2]] # 0-P1/2
- [-1, 1, Conv, [128, 3, 2]] # 1-P2/4
- [-1, 2, C3k2, [256, False, 0.25]] # 2
- [-1, 1, Conv, [256, 3, 2]] # 3-P3/8
- [-1, 2, C3k2, [512, False, 0.25]] # 4
- [-1, 1, Conv, [512, 3, 2]] # 5-P4/16
- [-1, 4, A2C2f, [512, True, 4]] # 6
- [-1, 1, Conv, [1024, 3, 2]] # 7-P5/32
- [-1, 4, A2C2f, [1024, True, 1]] # 8

# ----- NECK -----
head:
# P5 -> P4
- [-1, 1, nn.Upsample, [None, 2, "nearest"]] # 9
- [[-1, 6], 1, Concat, [1]] # 10
- [-1, 2, A2C2f, [512, False, -1]] # 11 (P4)

# P4 -> P3
- [-1, 1, nn.Upsample, [None, 2, "nearest"]] # 12
- [[-1, 4], 1, Concat, [1]] # 13
- [-1, 2, A2C2f, [256, False, -1]] # 14 (P3)

# P3 -> P2 (NEW PART)
- [-1, 1, nn.Upsample, [None, 2, "nearest"]] # 15
- [[-1, 1], 1, Concat, [1]] # 16
- [-1, 2, C3k2, [128, False, 0.25]] # 17 (P2) <-- NEW

# Down P2 -> P3
- [-1, 1, Conv, [256, 3, 2]] # 18
- [[-1, 14], 1, Concat, [1]] # 19
- [-1, 2, A2C2f, [256, False, -1]] # 20 (P3 refined)

# Down P3 -> P4
- [-1, 1, Conv, [512, 3, 2]] # 21
- [[-1, 11], 1, Concat, [1]] # 22
- [-1, 2, A2C2f, [512, False, -1]] # 23 (P4 refined)

# Down P4 -> P5
- [-1, 1, Conv, [1024, 3, 2]] # 24
- [[-1, 8], 1, Concat, [1]] # 25
- [-1, 2, C3k2, [1024, True]] # 26 (P5 refined)

# ----- DETECT -----
- [[17, 20, 23, 26], 1, Detect, [nc]]

```

Yolov12+P2.yaml

P2 detection head

P2 branch

Bi-directional feature refinement,
top down(FPN) and bottom
up(PAN)

Yolov12n+P2.yaml

```
Yolov12_P2.yaml
# YOLOv12-P2 (Drone Optimized)
# YOLOv12 object detection model with P2-P5 outputs.
nc: 1

scales:
  n: [0.50, 0.25, 1024]

# ----- BACKBONE -----
backbone:
  - [-1, 1, Conv, [64, 3, 2]] # 0-P1/2
  - [-1, 1, Conv, [128, 3, 2]] # 1-P2/4
  - [-1, 2, C3k2, [256, False, 0.25]] # 2
  - [-1, 1, Conv, [256, 3, 2]] # 3-P3/8
  - [-1, 2, C3k2, [512, False, 0.25]] # 4
  - [-1, 1, Conv, [512, 3, 2]] # 5-P4/16
  - [-1, 4, A2C2f, [512, True, 4]] # 6
  - [-1, 1, Conv, [1024, 3, 2]] # 7-P5/32
  - [-1, 4, A2C2f, [1024, True, 1]] # 8

# ----- NECK -----
head:
  # P5 -> P4
  - [-1, 1, nn.Upsample, [None, 2, "nearest"]] # 9
  - [[-1, 6], 1, Concat, [1]] # 10
  - [-1, 2, A2C2f, [512, False, -1]] # 11 (P4)

  # P4 -> P3
  - [-1, 1, nn.Upsample, [None, 2, "nearest"]] # 12
  - [[-1, 4], 1, Concat, [1]] # 13
  - [-1, 2, A2C2f, [256, False, -1]] # 14 (P3)

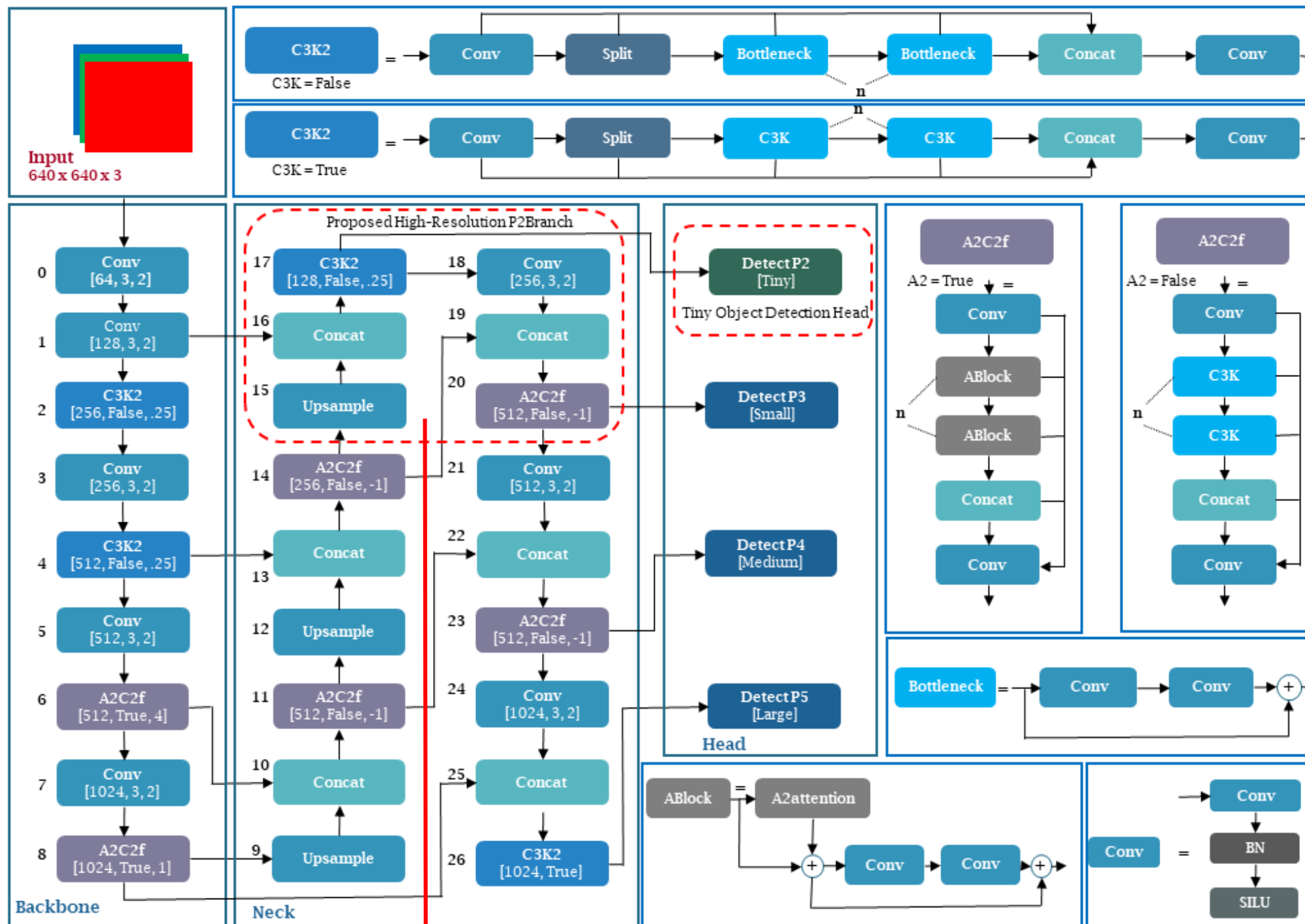
  # P3 -> P2 (NEW PART)
  - [-1, 1, nn.Upsample, [None, 2, "nearest"]] # 15
  - [[-1, 1], 1, Concat, [1]] # 16
  - [-1, 2, C3k2, [128, False, 0.25]] # 17 (P2) <-- NEW

  # Down P2 -> P3
  - [-1, 1, Conv, [256, 3, 2]] # 18
  - [[-1, 14], 1, Concat, [1]] # 19
  - [-1, 2, A2C2f, [256, False, -1]] # 20 (P3 refined)

  # Down P3 -> P4
  - [-1, 1, Conv, [512, 3, 2]] # 21
  - [[-1, 11], 1, Concat, [1]] # 22
  - [-1, 2, A2C2f, [512, False, -1]] # 23 (P4 refined)

  # Down P4 -> P5
  - [-1, 1, Conv, [1024, 3, 2]] # 24
  - [[-1, 8], 1, Concat, [1]] # 25
  - [-1, 2, C3k2, [1024, True]] # 26 (P5 refined)

# ----- DETECT -----
- [[17, 20, 23, 26], 1, Detect, [nc]]
```



Architecture of the proposed YOLOv12+P2 detection framework

Architecture of the proposed P2 detection framework

