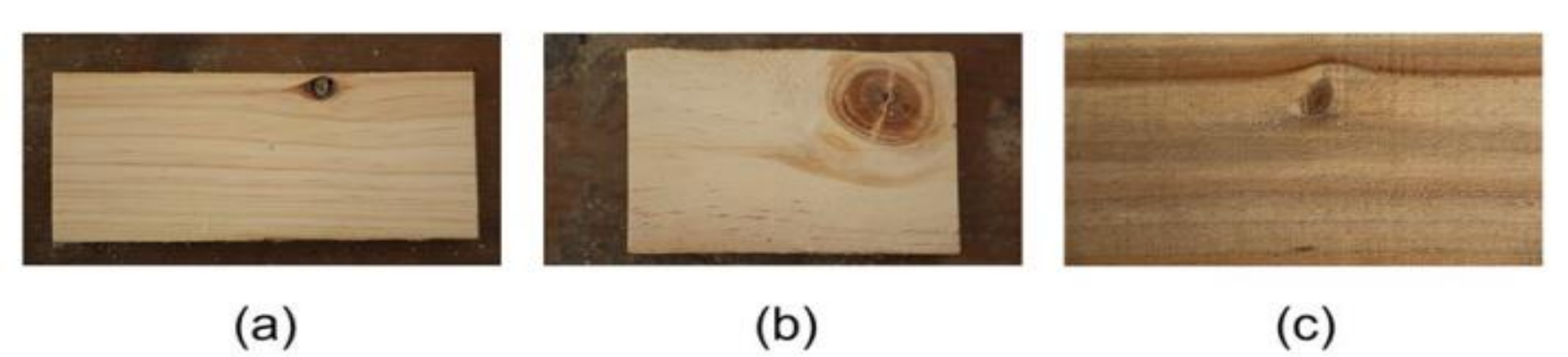
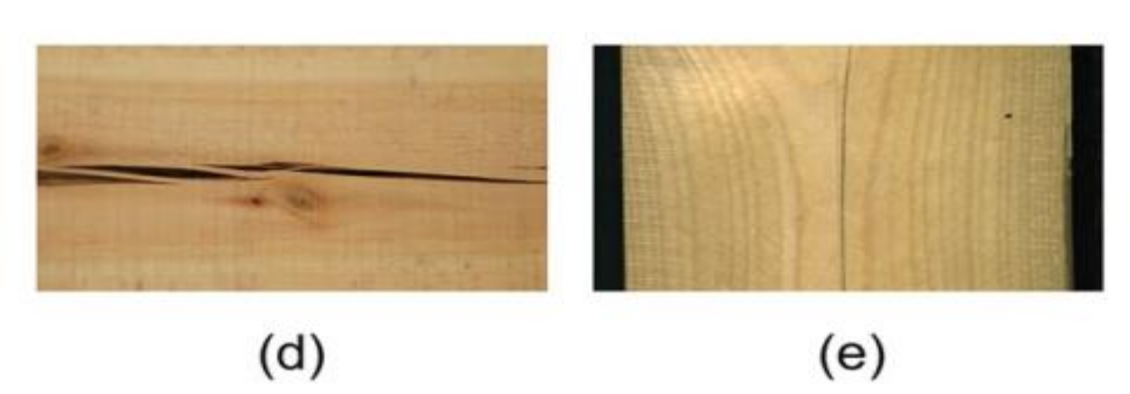
CWB-YOLOV8

NOV 2024

DATASET



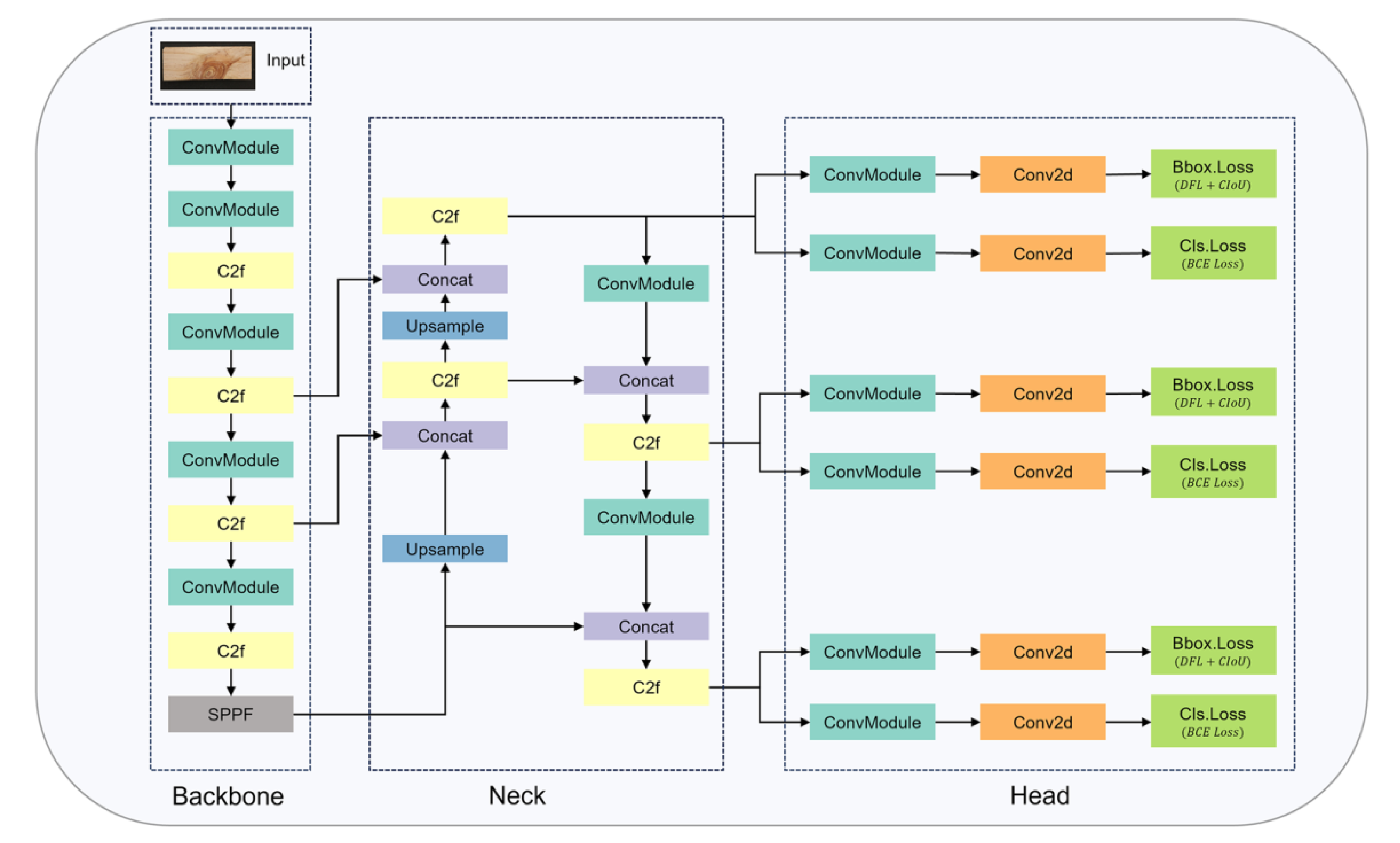


DATASET

- Number of samples
  - Around 7300
- Number of classes: 6

Table 1 Defect distribution

Defect category	Number of defective images	Number of defects	Defect proportion (%)
Live_knot	2431	3487	30.9
Dead_knot	2005	2980	26.4
Knot_with _crack	871	1079	9.6
Crack	728	1254	11.1
Resin	476	1186	10.5
Marrow	756	1301	11.%



Input
ConvModule
ConvModule
C2f
ConvModule
C2f
ConvModule
C2f
ConvModule
C2f
SPPF
C2f
Concat
Upsample
C2f
Concat
Upsample
ConvModule
Concat
C2f
ConvModule
Concat
C2f
ConvModule
Conv2d
Bbox.Loss
(DFL + CIoU)
ConvModule
Conv2d
Cls.Loss
(BCE Loss)
ConvModule
Conv2d
Bbox.Loss
(DFL + CIoU)
ConvModule
Conv2d
Cls.Loss
(BCE Loss)
ConvModule
Conv2d
Bbox.Loss
(DFL + CIoU)
ConvModule
Conv2d
Cls.Loss
(BCE Loss)
Backbone
Neck
Head

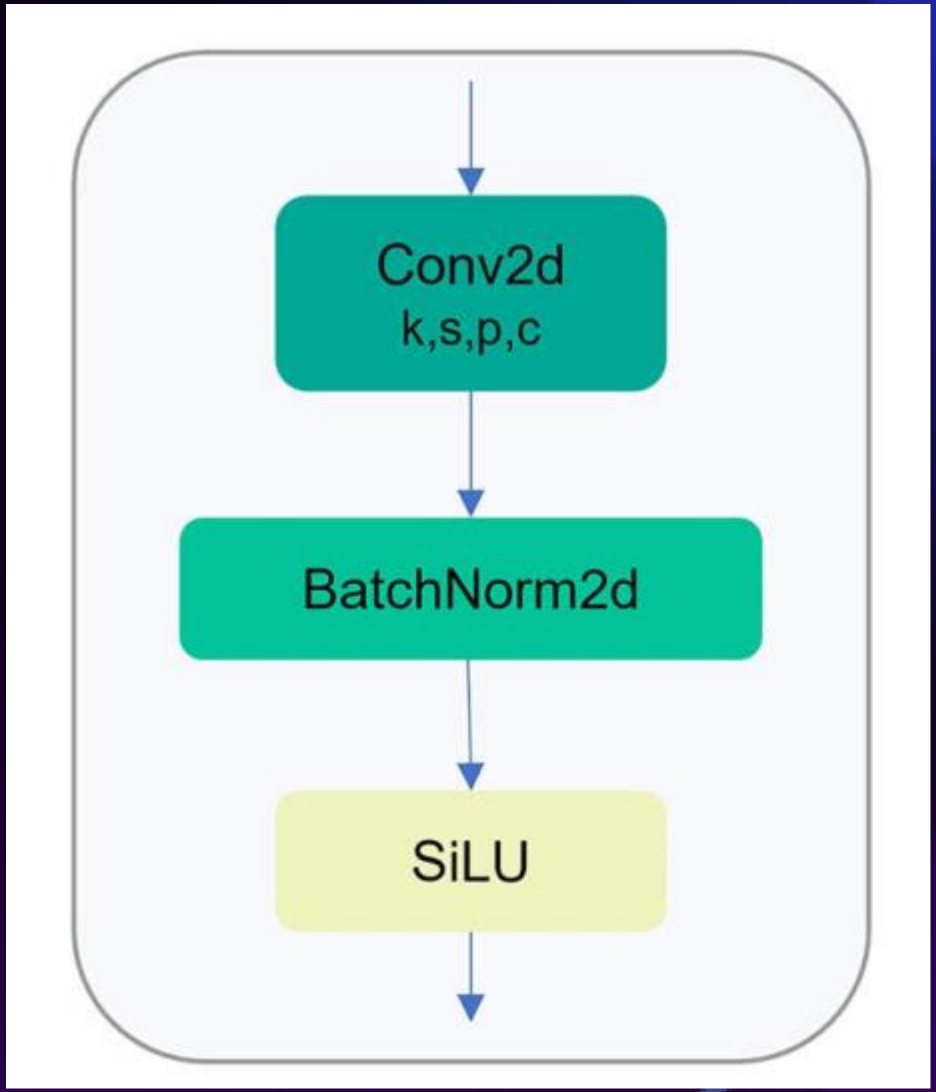
FIRST IMPROVMENT

- ConvModule drawbacks
  - Cannot handle the complex and changeable appearances of wood defects

- Replace ConvModule with CondConv

- CondConv adaptively changes the weights of the convolution kernel based on input conditions.

- Enhances the extraction of features for different types of wood.

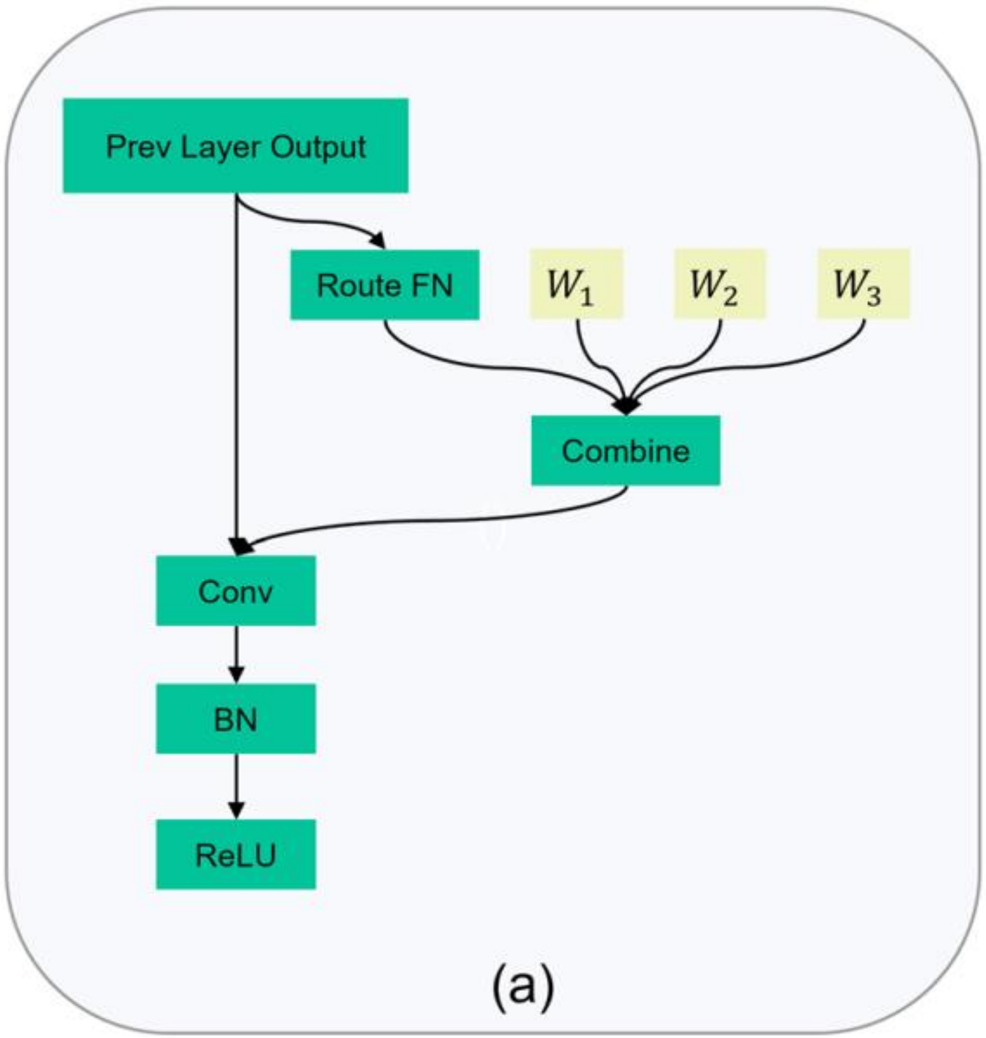


- W
  - Convolution kernel
- ROUTR FN
  - Calculate the weights of different convolution kernels

$$r(x) = \mathrm{Sigmoid}\big(\mathrm{GlobalAveragePool}(x)R\big),$$

- Output: generates customized weight coefcients for the input features

$$\mathrm{Output}(x) = \sigma((\alpha_1 W_1 + \cdots + \alpha_n W_n) \cdot \boldsymbol{x}),$$



(a)

SECOND IMPROVMENT

MULTIHEAD ATTENTION

- Multihead handle medium- and long-distance information
- BUT requiring large amounts of computing resources

- Use BiFormer, use dynamic sparse attention mechanism
  - Four-layer pyramid
    - 3 patch merging layers
    - 1 patch embedding layer

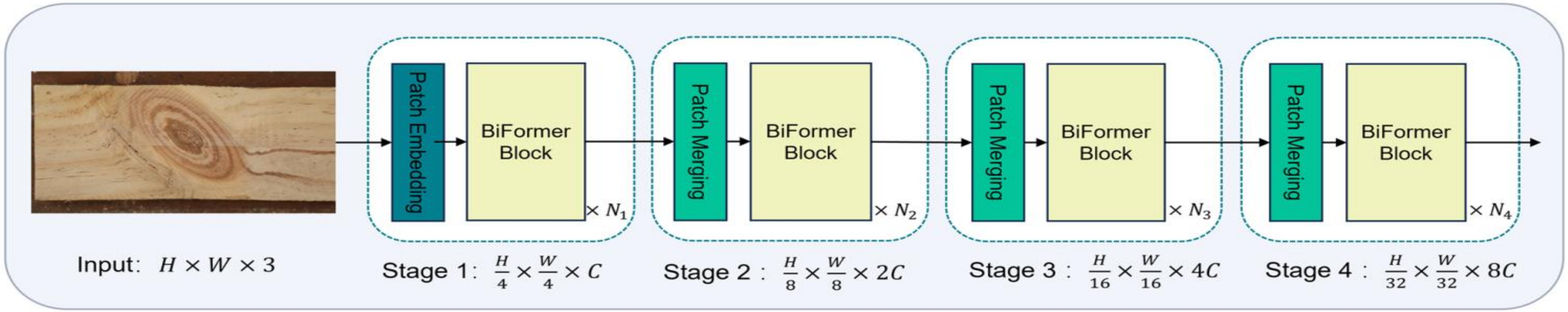
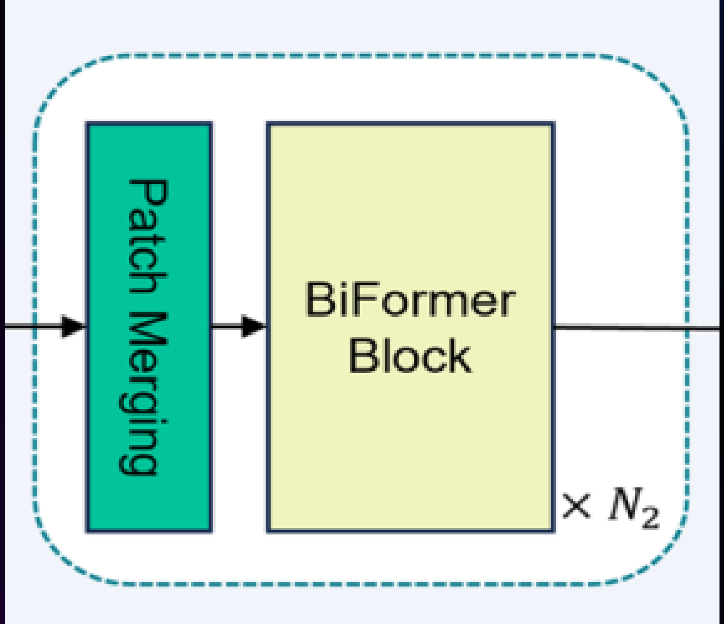


Fig. 5 BiFormer structure

BIFORMER

- Stage 2 – 4
  - Patch Merging:
    - Input feature downsample, spatial patches are merged, and their channels are concatenated
    - Channel depth increase
  - BiFormer Blocks
    - Depthwise Convolution
      - Captures local spatial information and encodes positional relationships.
    - Bi-Level Routing Attention
      - Filters out irrelevant regions
    - Multilayer Perceptron (MLP)



THIRD IMPROVMENT

LOSS FUNCTION

- Complete Intersection over Union (CIoU) Loss is bounding box regression loss
- improve the alignment of predicted bounding boxes with ground truth

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

D: Euclidean distance between predicted and ground-truth boxe

C: Diagonal length of the smallest enclosing box that contains both the predicted and ground-truth boxes.

• IoU represents the intersection-over-union ratio between the anticipated bounding box and the actual bounding box

C-IOU DRAWBACKS

•The penalty weights in CIoU are not dynamically adjusted based on sample quality.
•CIoU applies a fixed penalty regardless of sample quality, which reduces the generalizability of the model.
•Wood defect images often result in low-quality samples.
•in the bounding box regression branch replace **CIoU** by **Wise-IoU**

WISE-IOU ALGORITHM

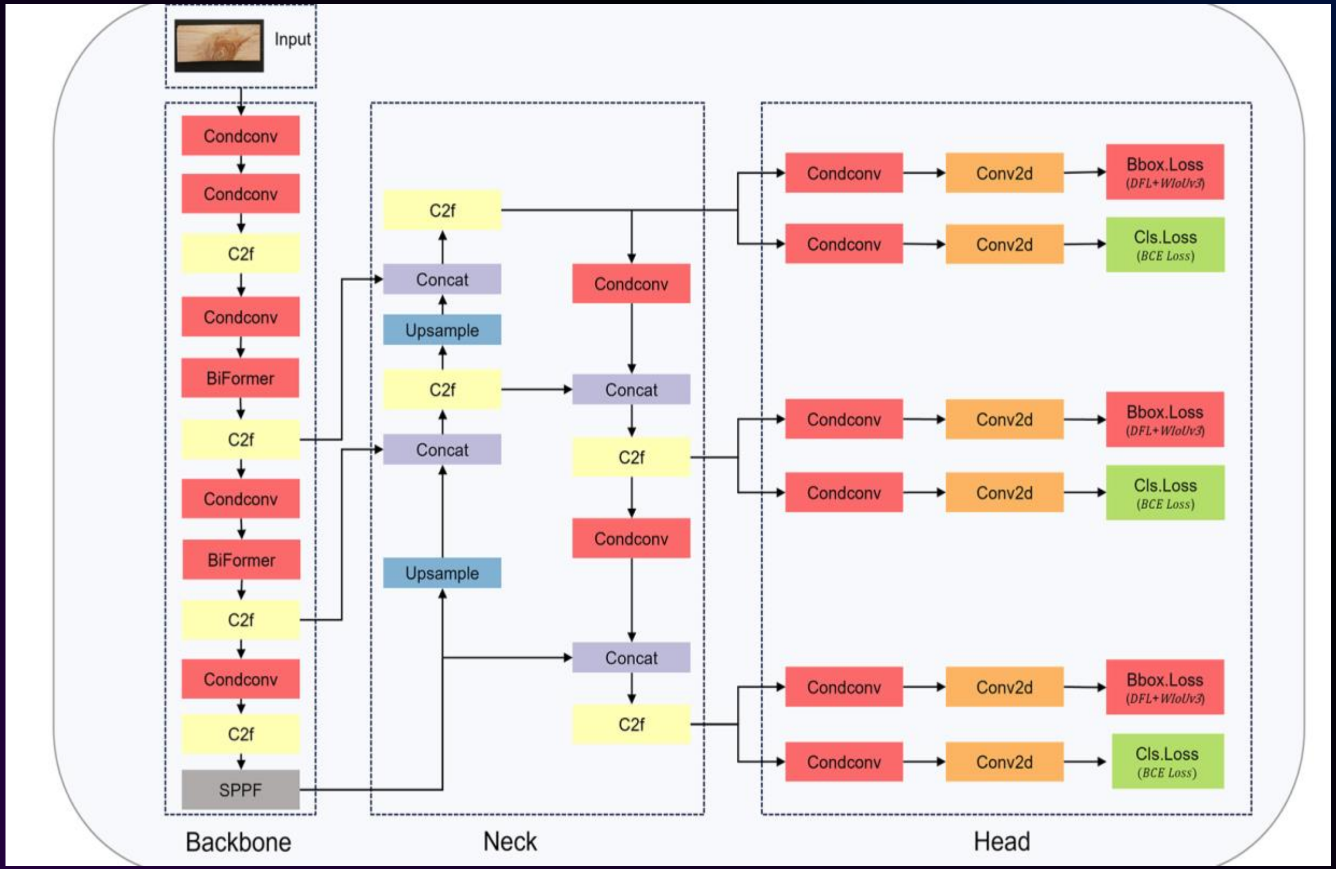
- Wise-IoU incorporates a distance-based attention mechanism.
- Low-quality samples contribute less to the gradient updates.
- High-quality samples, on the other hand, receive greater attention.

- Outlier Handling: allows the model to handle diverse and noisy datasets better
- Better habdle unseen or challenging defect samples

$$L_{WIoUv1} = (1 - IoU)\exp\left(\frac{(x - x_{gt})^2 + (y - y_{gt})^2}{(W_g^2 + H_g^2)^*}\right)$$

CWB-YOLO8

1. use the BiFormer attention mechanism
2. use Wise-IoUv3 to replace the original CIoU loss function
3. use the CondConv module to replace the original convolution module



IMPROVMENT

- improve the ability of the model to understand the complex features of wood(BiFormer)
- improving the bounding box regression process(Wise-IoUv3)
- Dynamically adjusts the weight of the convolution kernel to improve the extraction of different types of wood defect features(CondConv)

RESULTS

Table 3 Results of an ablation experiment

Algorithm	Module			Results			
	CondConv	Wise-IoU	BiFormer	mAP@0.5	mAP@0.5:0.95	Precision	Recall
YOLOv8				0.857	0.530	0.826	0.807
	√			0.866	0.547	0.829	0.817
	√	√		0.871	0.542	0.835	0.821
	√	√	√	0.892	0.588	0.870	0.833

RESULTS

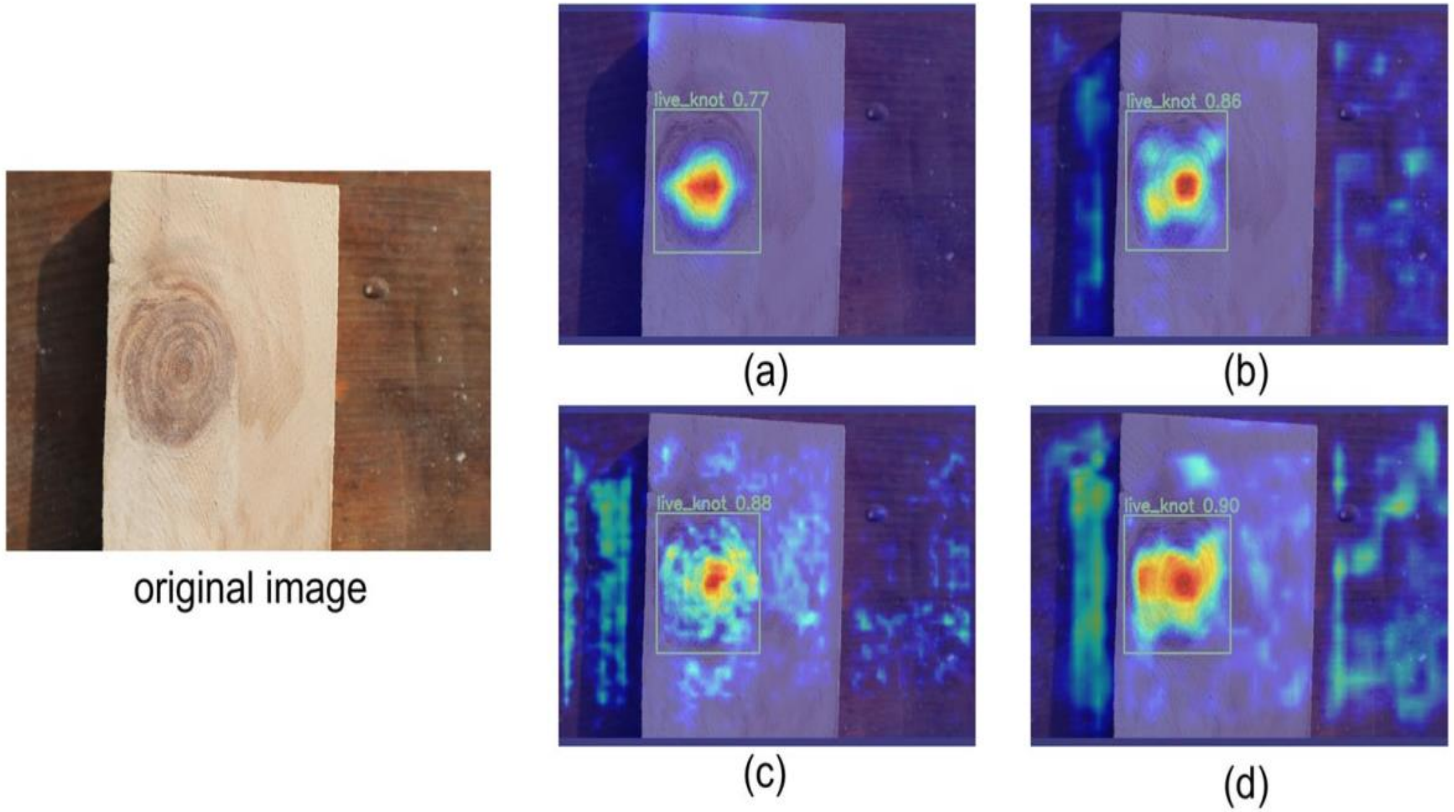
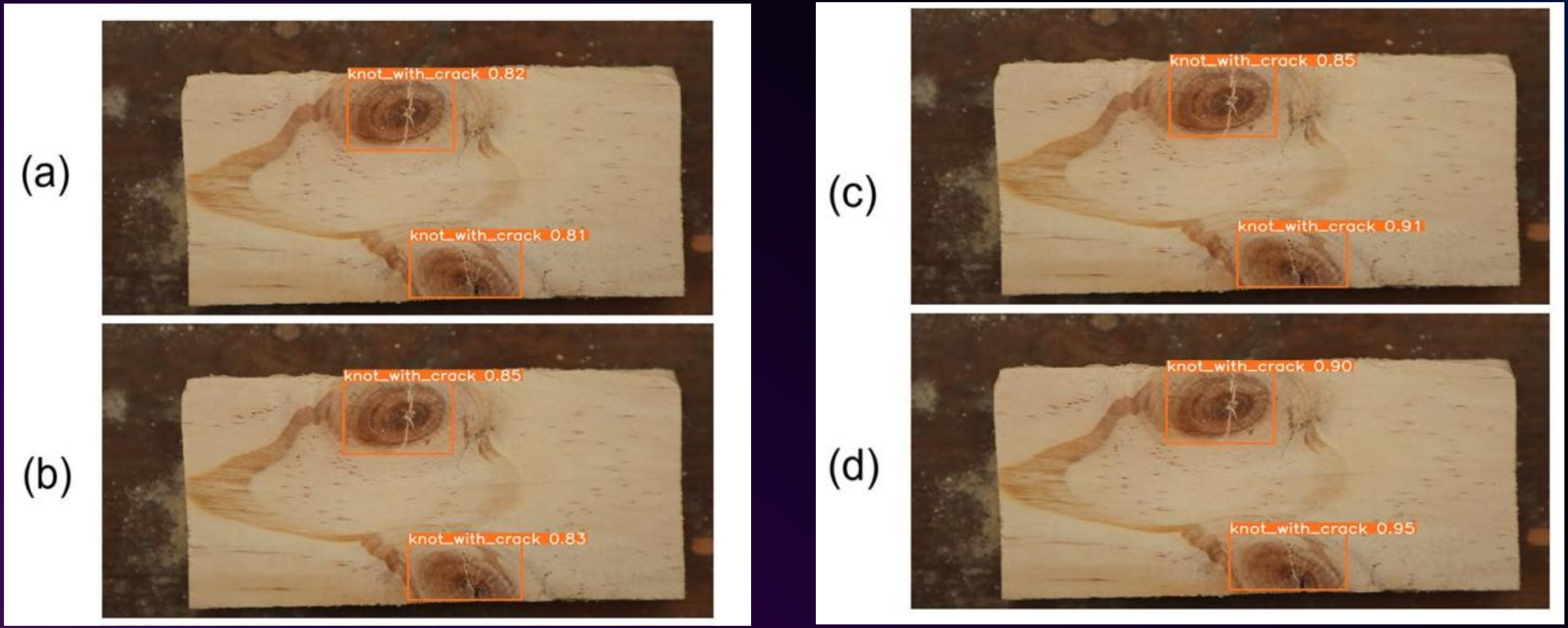


Fig. 13 Original image and heatmaps: **a** YOLOv8, **b** YOLOv8 + CondConv, **c** YOLOv8 + CondConv + Wise-IoU, and **d** CWB-YOLOv8

RESULTS


(a)
knot_with_crack 0.82
knot_with_crack 0.81
(b)
knot_with_crack 0.85
knot_with_crack 0.83
(c)
knot_with_crack 0.85
knot_with_crack 0.91
(d)
knot_with_crack 0.90
knot_with_crack 0.95

FUTURE WORKS

- EfficientAD
- PatchCore