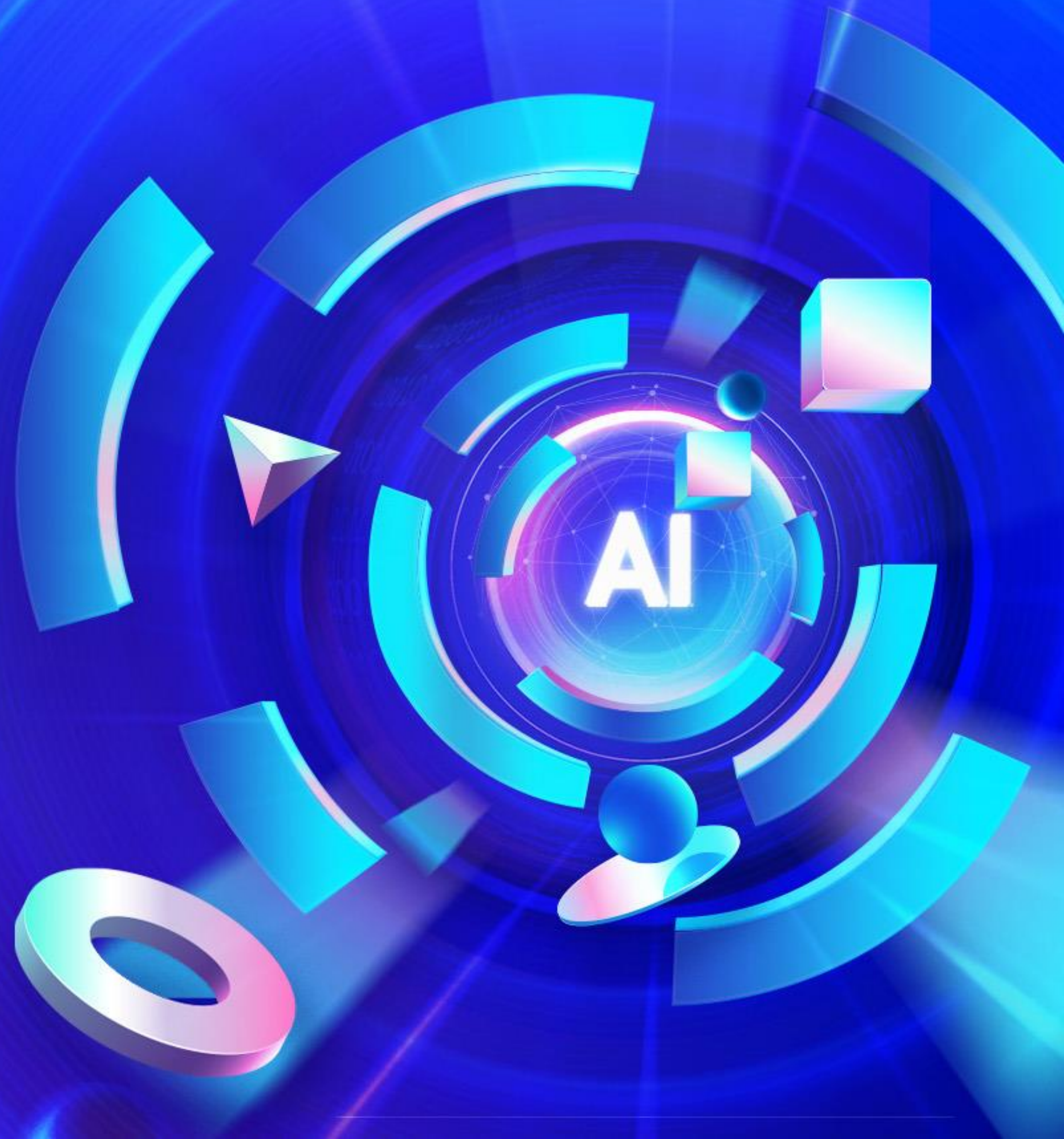


# ECV2023极市计算机视觉 开发者榜单大赛

赛题：仪表读数识别

团队名：MILab

汇报人：高雨浩







- 所在单位：南京理工大学 智能信息处理与系统实验室

- 队伍成员：

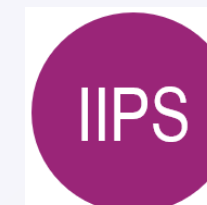
高雨浩：南京理工大学 计算机科学与工程学院 硕士研究生

周 波：南京理工大学 计算机科学与工程学院 硕士研究生

张 旺：南京理工大学 计算机科学与工程学院 硕士研究生

- 指导老师：

姚亚洲：南京理工大学教授、博导，入选国家海外高层次人才计划。



## ● 赛题任务

### 1、仪表识别

对于预测仪表框与真实仪表框，计算f1-score，IoU使用0.7，此部分成绩记为 score1

### 2、刻度点识别

对于预测关键点与真实关键点，计算RMSE，此部分成绩记为 score2

### 3、OCR识别

对于预测的**数值框**，计算f1-score，IoU使用**0.5**，此部分成绩记为 score3

### 4、读数识别

**计算每个仪表的读数得分，并对所有仪表的得分求平均值**，此部分成绩记为 score4

$$final\_score = score_1 \times weight_1 + score_2 \times weight_2 + score_3 \times weight_3 + score_4 \times weight_4$$

加权的权重为配置参数，默认值为：

weight\_1 = 0.1          weight\_2 = 0.25

weight\_3 = 0.25        weight\_4 = 0.4

总分 = 算法精度得分\*0.99 + 算法性能得分\*0.01



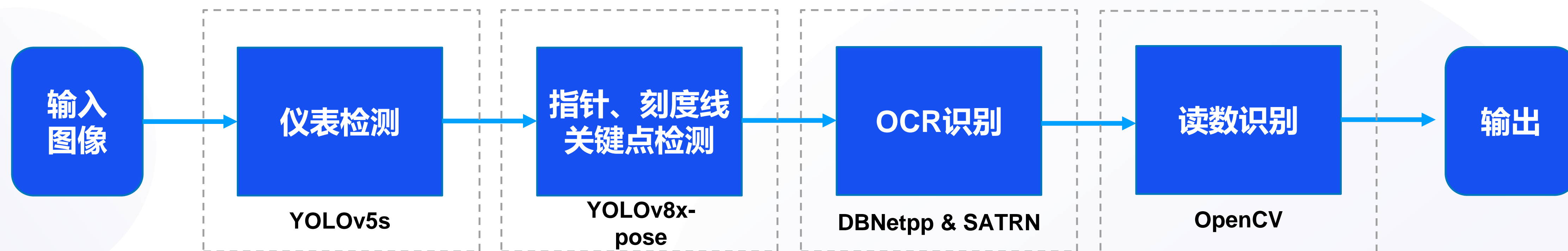
外圈读数：2.886

内圈读数：0.02





## ● 算法流程

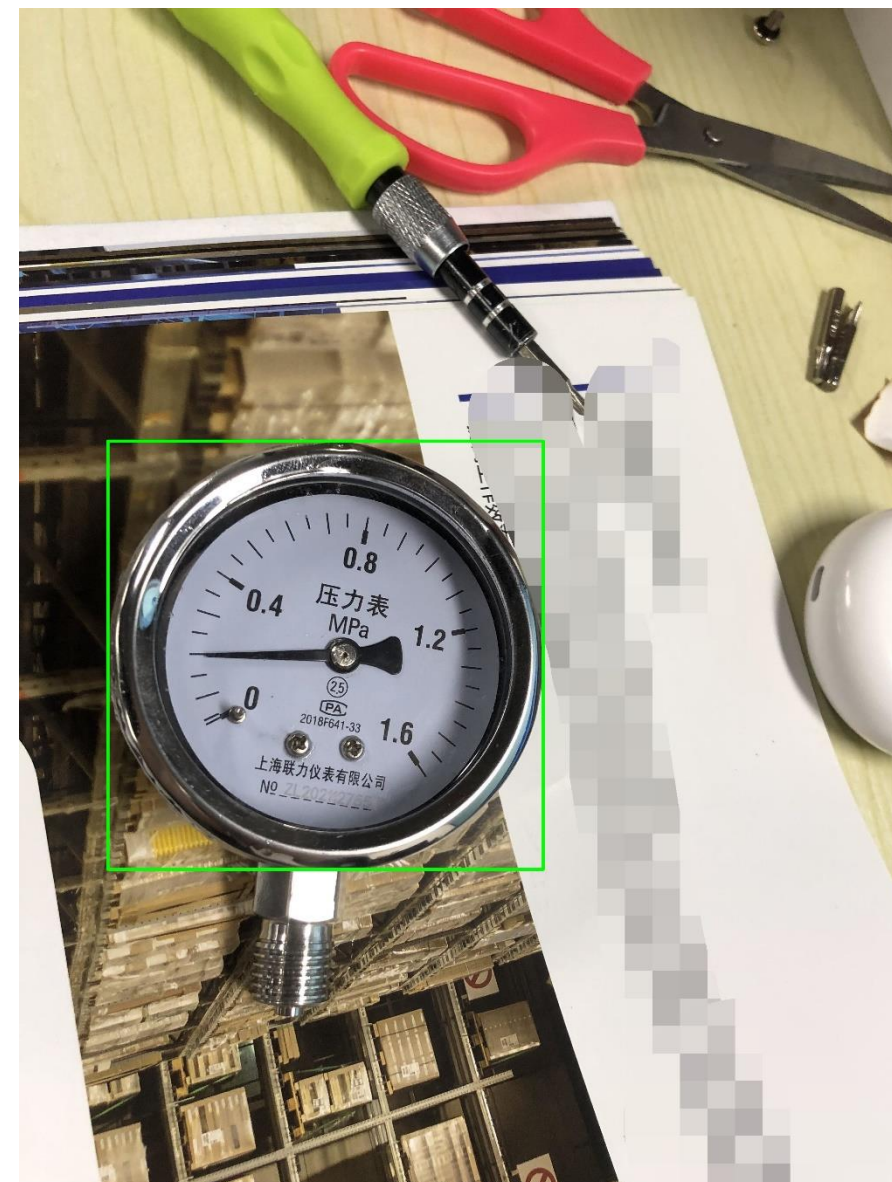
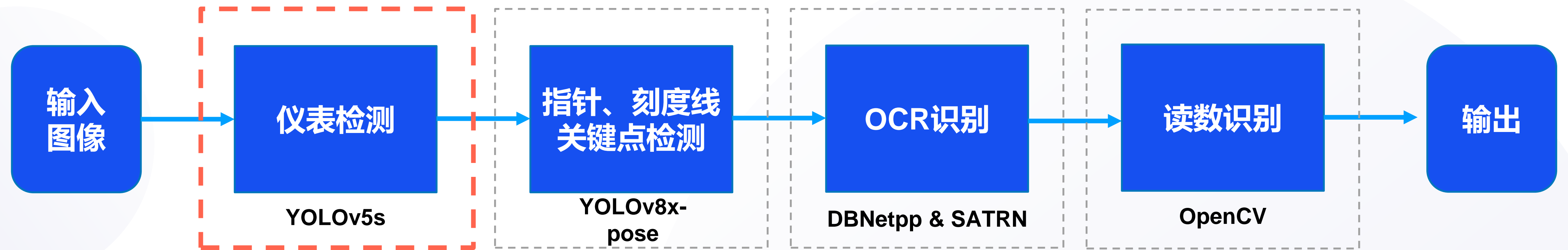


## ● 模型选择

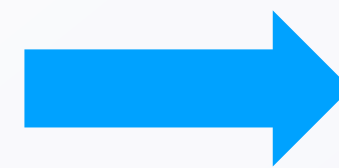
- 仪表检测：YOLOv5s
- 关键点检测：YOLOv8x-pose
- 文本检测：DBNetpp
- 文本识别：SATRN

# 算法设计

## 算法流程



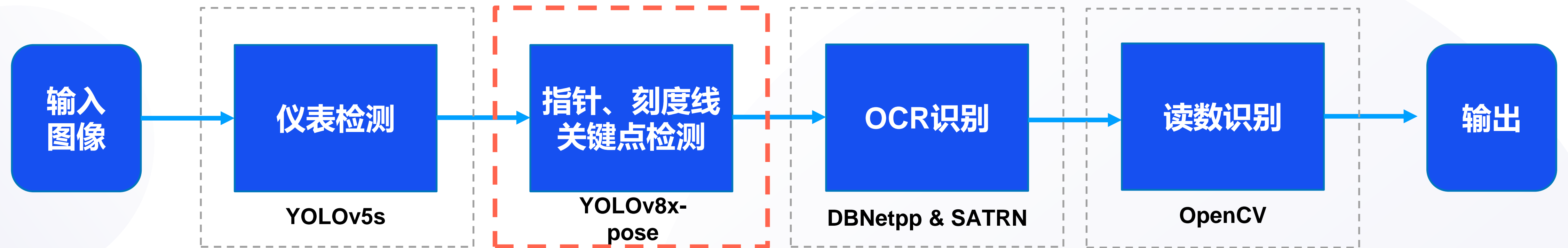
仪表检测



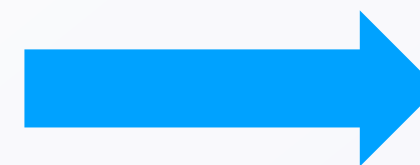


# 算法设计

## 算法流程



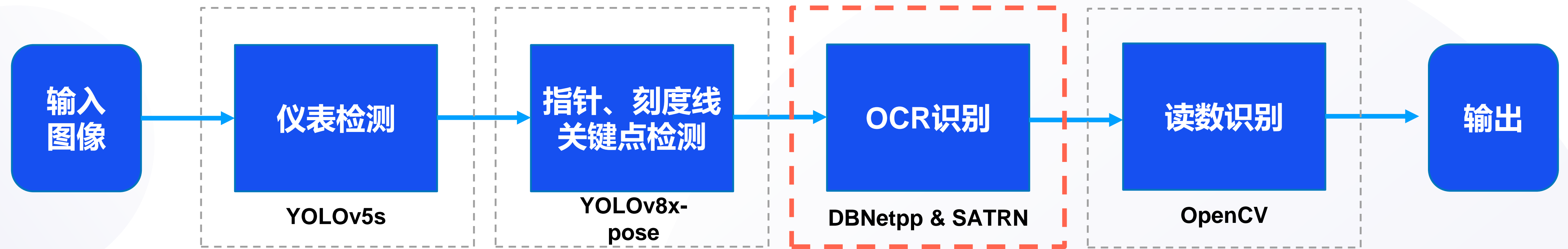
关键点检测



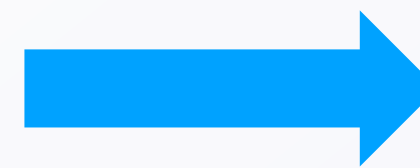


# 算法设计

## 算法流程



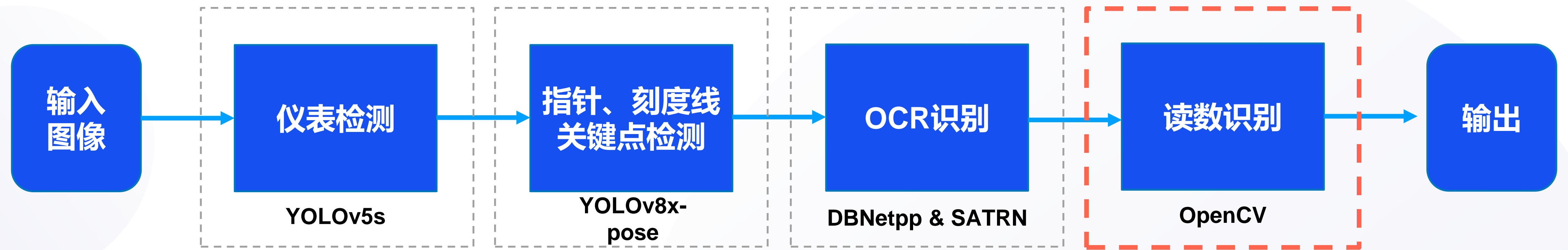
OCR识别



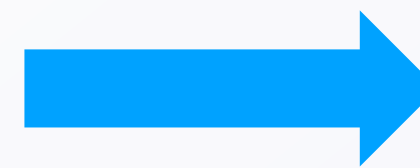


# 算法设计

## 算法流程



读数识别



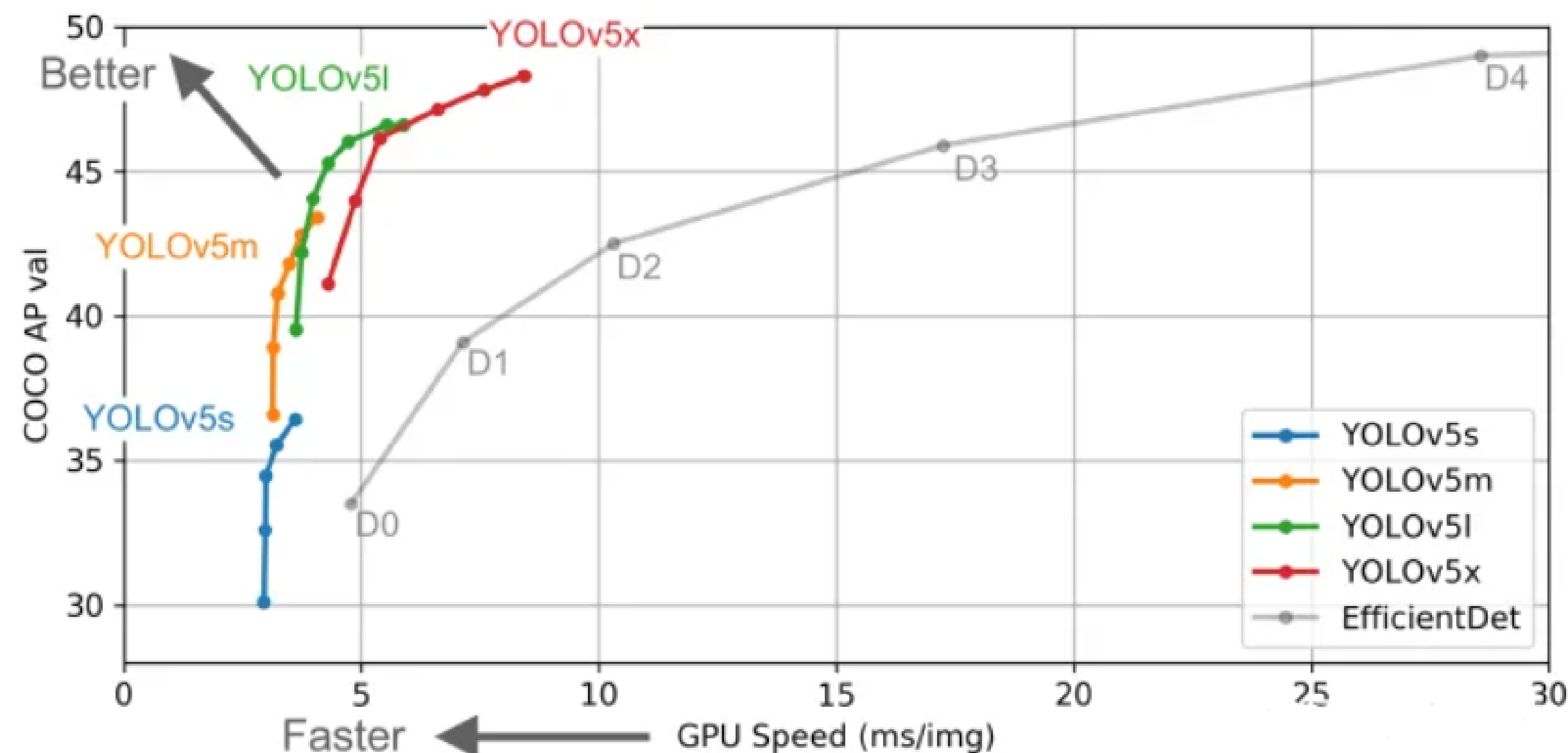


## ● 基于YOLOv5s的仪表检测

### ➤ 选型理由

Yolov5s网络小，速度快。虽然AP精度低，但检测的是仪表这种大目标也是完全够用的。

我们采用极市官方提供的yolov5s训练套件进行训练，上手简单，且仪表识别准确率达到了99.7%，效果满足实际需求。





## ● 基于YOLOv8x-pose的指针和刻度关键点检测

### ➤ 选型理由

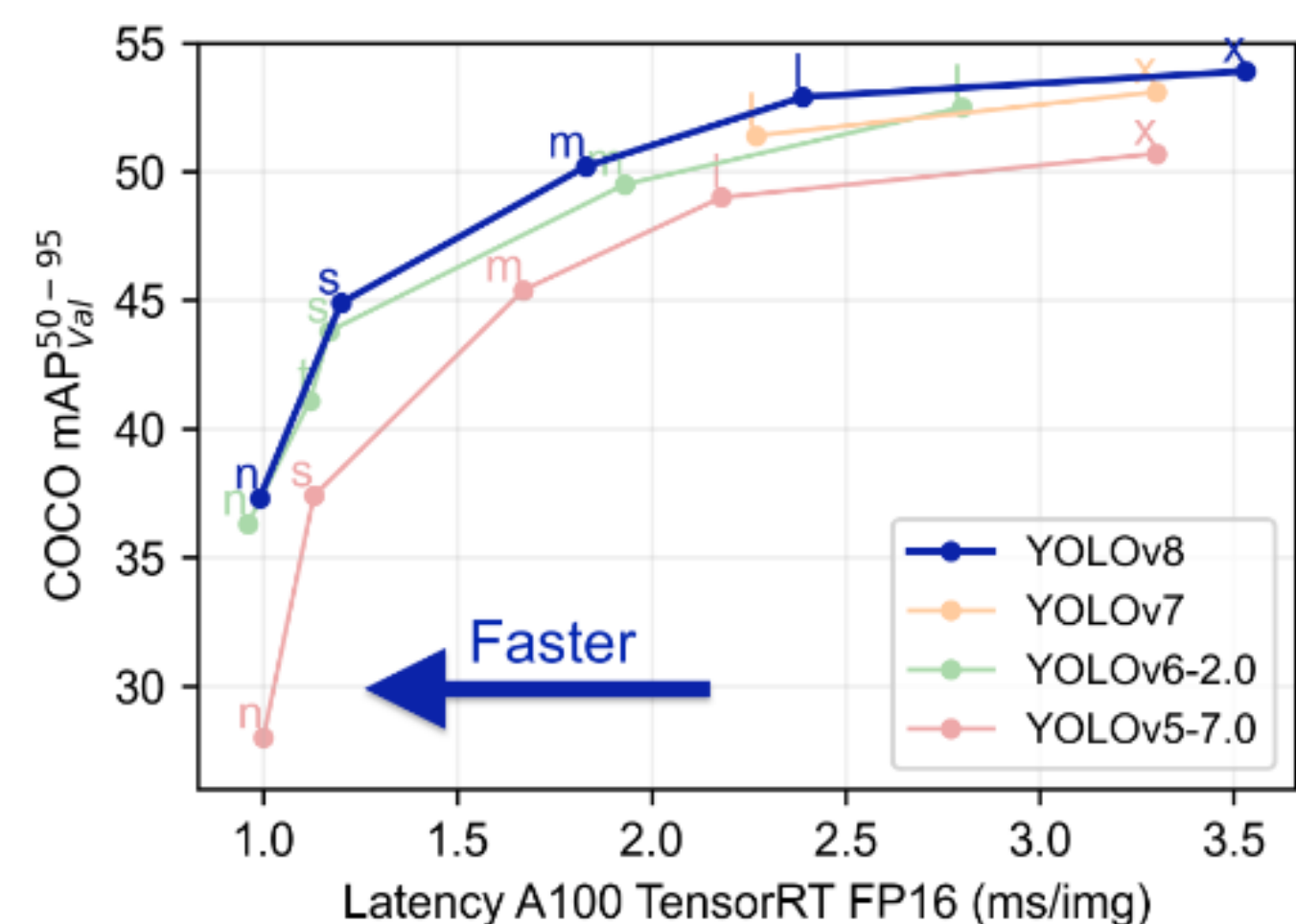
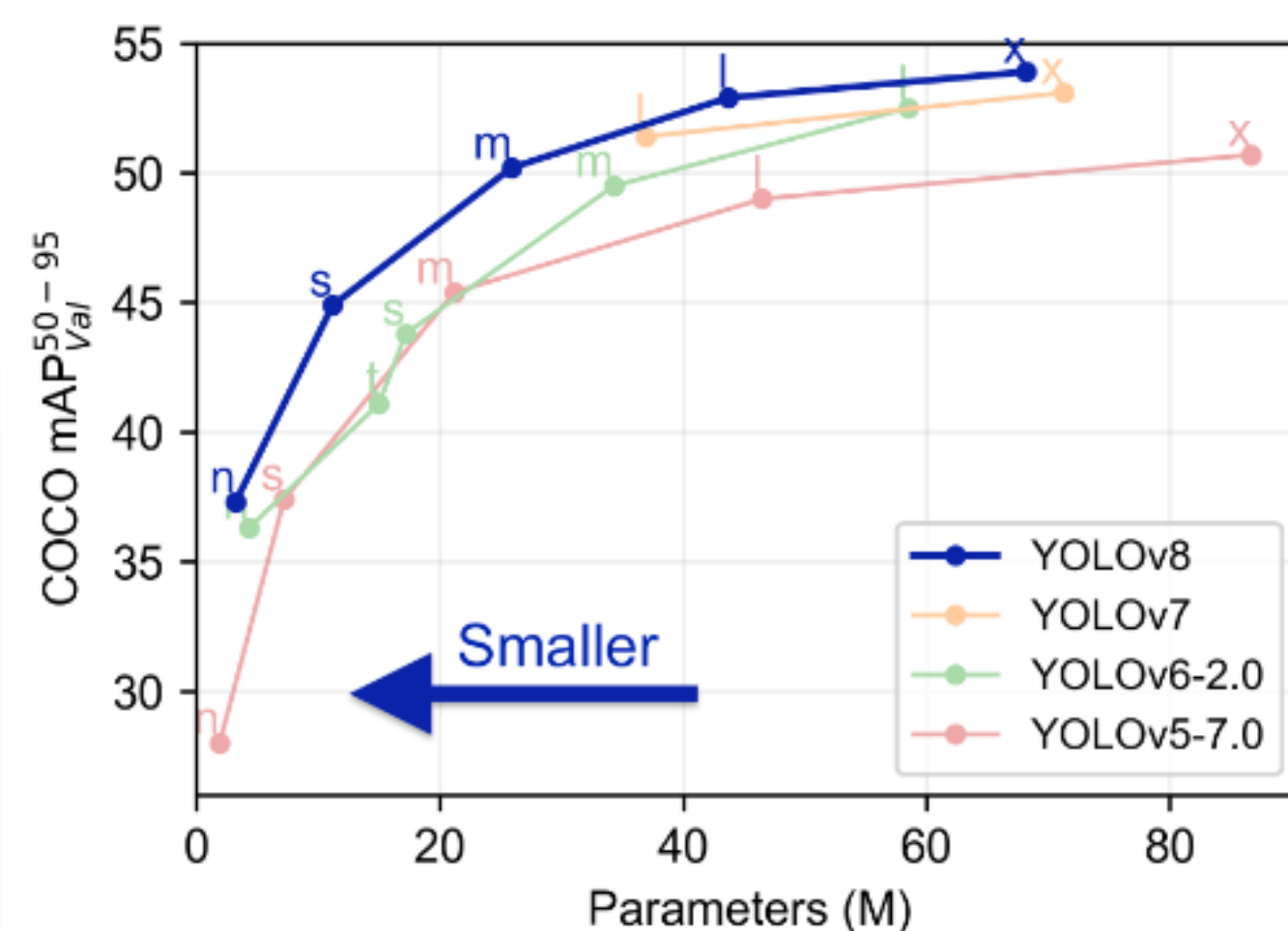
YOLOv8是YOLOv5团队ultralytics在今年开源的SOTA模型，引入了新的改进，进一步提升了性能和灵活性。YOLOv8 设计快速、准确且易于使用，使其成为目标检测、实例分割、图像分类和姿态估计任务的绝佳选择。

**精度上YOLOv8相比YOLOv5高出一大截，但速度略有下降。**

**yolo-pose是估计人体姿态的一个模型，它将人体姿态分为17个关键点。**

为了尽可能地提高关键点检测精度，我们使用了规模最大的YOLOv8x-pose模型。

<https://github.com/ultralytics/ultralytics>





- 基于YOLOv8x-pose的指针和刻度关键点检测

- 指针关键点数据集制作





## ● 基于YOLOv8x-pose的指针和刻度关键点检测

### ➤ 刻度关键点数据集制作



### ➤ 训练策略

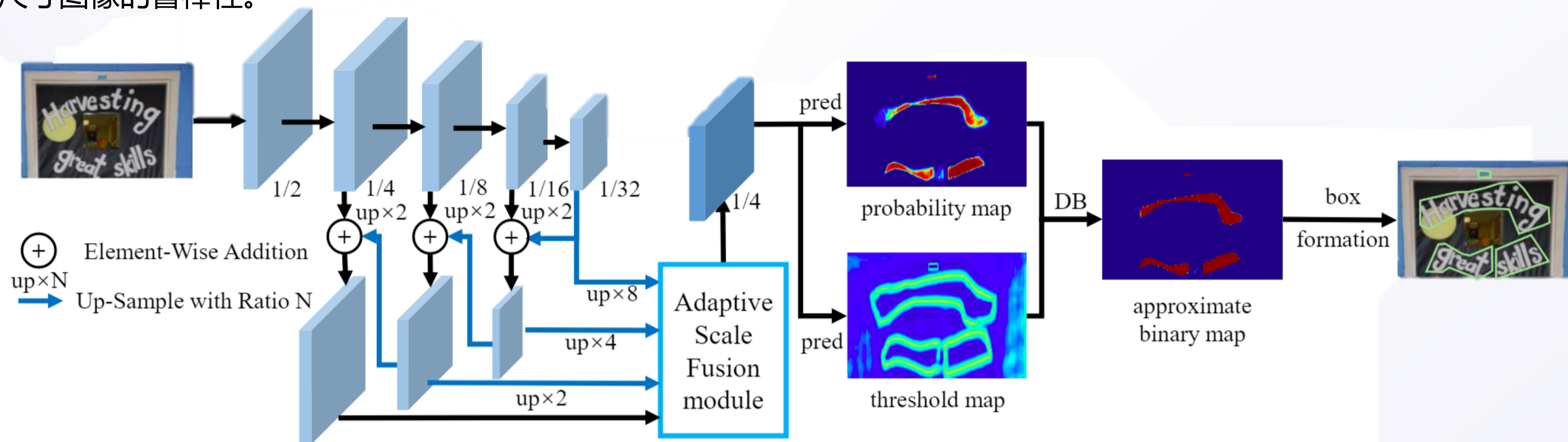
使用MSCOCO2017预训练模型在仪表数据集上微调。我们将**关键点权重从12提高到20**。



## ● 基于DBNetpp的文本检测

### 选型理由

由于仪表尺寸大小不一，我们选用DBNetpp模型，它在DBNet模型基础上引入了多级特征聚合模块(Adaptive Scale Fusion, ASF)，ASF模块由阶段注意力和空间注意力子模块构成，加强了不同尺度特征的融合，提高了处理不同尺寸图像的鲁棒性。



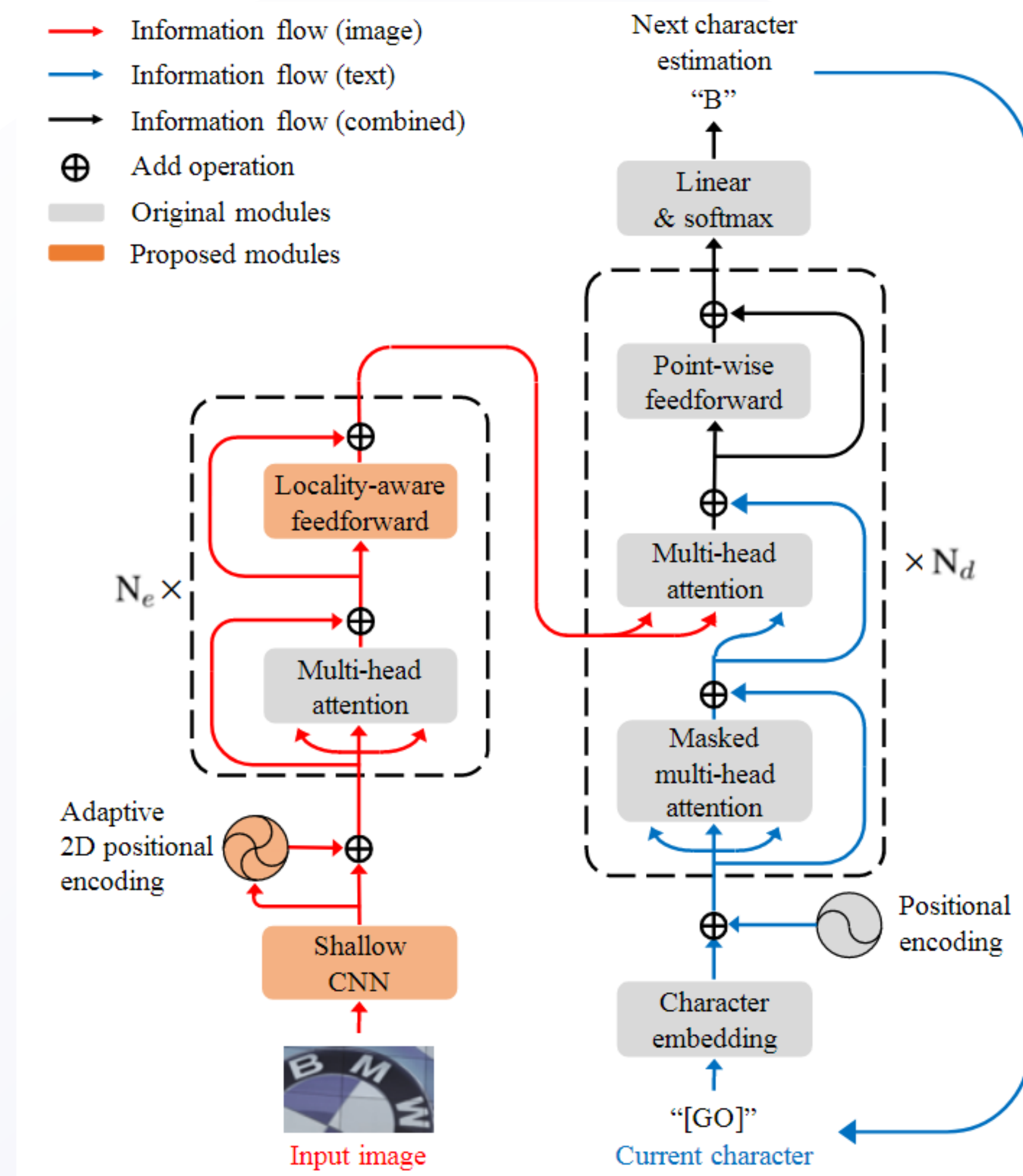


## ● 基于SATRN模型的文本识别

### 选型理由

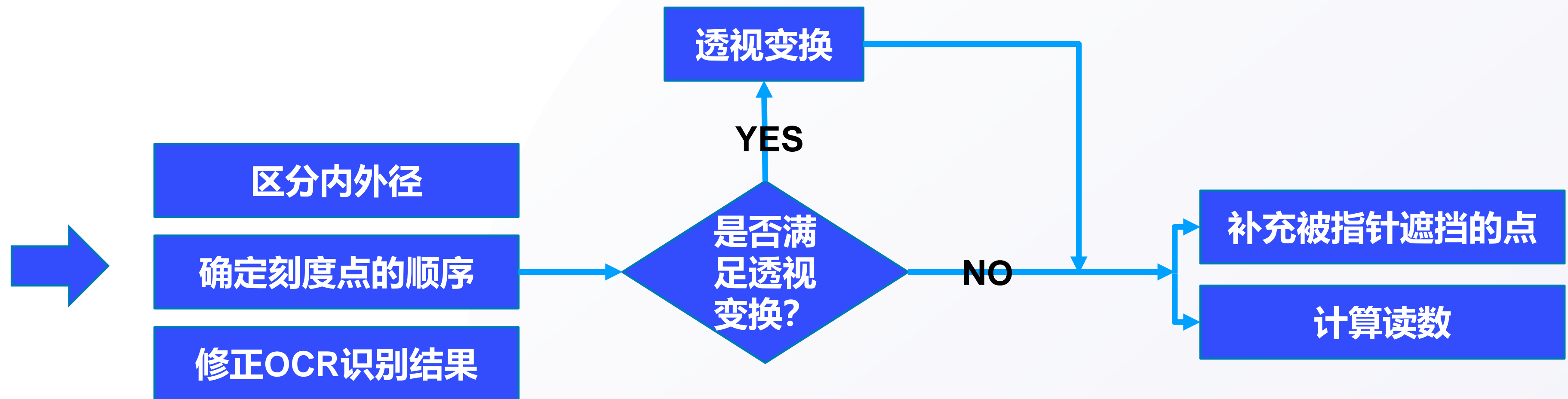
SATRN模型利用self-attention机制对字符的2D空间关系进行建模，并且在FFN模块中引入卷积层，增强了模型对全局和局部特征的捕捉能力。

SATRN模型对于大曲率弯曲、大角度旋转文本依然具备足够的识别能力，在多个不规则文本数据集上达到SOTA。





## ● 读数识别流程





## • 区分内外径



根据刻度点和数值框距离指针原点的距离区分内径和外径，分别进行后续处理。



## ● 确定刻度点的顺序



先将刻度点坐标转成以指针原点为中心的极坐标，根据角度从小到大排序[16,0,4,8,12]

然后计算相邻刻度点两两间的角度差值，选择最大差值的下一个刻度点作为起点，调整顺序后[0,4,8,12,16]



## ● 修正OCR识别结果

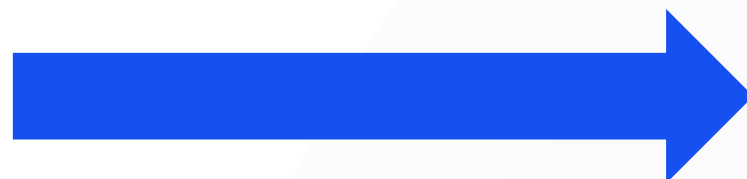


内径确定刻度点顺序: [0,500,1000,**500**,2000,2300]

修正OCR识别结果: [0,500,1000,**1500**,2000,2300]



## ● 透视变换



1、确定最后一个点的坐标：已知最后一个点在透视变换后的极坐标角度为 $45^\circ$ ，假设所有刻度点距离圆心距离为 $R$ ，则最后一个点的坐标可以表示为 $(45^\circ, R)$ 可以根据该极坐标得出笛卡尔坐标系下的坐标。

2、确定其他三个点的坐标：根据每个刻度点之间的夹角为 $45^\circ$ ，可以依次计算出其他三个点在变换后的坐标。

这里一共有6个间隔平分整个 $270^\circ$ 量程，所以每个刻度夹角为 $270/6=45^\circ$

通过这种方式，可以确定四个刻度点在透视变换后图像中的坐标，并用这些坐标来进行透视变换。



## ● 透视变换



```
nums_list = [ [0, 1, 2, 3, 4],  
               [0, 4, 8, 12, 16],  
               [0, 0.4, 0.8, 1.2, 1.6],  
               [0, 1, 2, 3, 4, 5],  
               [0, 20, 40, 60, 80, 100],  
               [0, 2, 4, 6, 8, 10],  
               [0, 0.2, 0.4, 0.6, 0.8, 1.0],  
               [0, 0.02, 0.04, 0.06, 0.08, 0.1],  
               [0, 1, 2, 3, 4, 5, 6],  
               [0, 10, 20, 30, 40, 50, 60],  
               [0, 5, 10, 15, 20, 25],  
               [0, 0.5, 1, 1.5, 2, 2.5] ]
```

通过观察数据集，预先定义一系列可以进行透视变换的刻度模式，并将其存储在 `nums_list` 中，当OCR 识别的刻度序列存在于这个列表中时，即可进行透视变换。



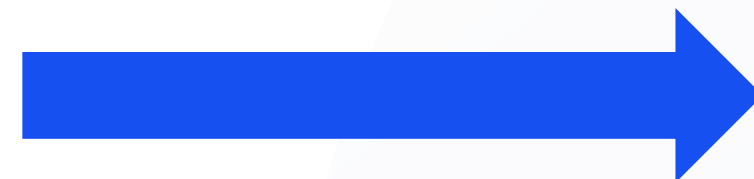
## ● 补充被指针遮挡的点



- 1、首先，确定缺失点的前后点，以及它们与圆心的平均距离。这些信息将用于推算缺失点的极坐标。
- 2、根据前后点的角度和平均距离，计算出缺失点的极坐标。极坐标由极径和极角组成，极径表示点到圆心的距离，极角表示点在极坐标系中的角度。将缺失点的极坐标转换为笛卡尔坐标。
- 3、如果进行了透视变换，需要将转换后的坐标应用于透视矩阵，以获得原图中的坐标。



## ● 计算读数



确定指针的前后刻度点位置，并记录它们在极坐标系中的角度值。假设前刻度点角度为 $\theta_1$ ，后刻度点角度为 $\theta_2$ 。指针角度为 $\theta_r$ 。假设指针前刻度点的读数为 $r_1$ ，后刻度点读数为 $r_2$ ，则最终的读数为 $r = r_1 + \Delta r * \Delta \theta$ 。

$$\text{其中 } \Delta r = r_2 - r_1, \Delta \theta = \frac{\theta_r - \theta_1}{\theta_2 - \theta_1}$$



- 存在问题



仪表的数字文本被边框遮挡导致无法识别。



## ● 改进方法



参考透视变换所用的查表方法，为内外径各建立一个list，以外径举例：

```
outer_list = [  
    [32, 50, 100, 150, 200],  
    [0, 4, 8, 12, 16],  
    [0, 2, 4, 6, 8, 10],  
    [0, 1, 2, 3, 4, 5, 6],  
    [0, 0.2, 0.4, 0.6, 0.8, 1],  
    [0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6],  
    [0, 2, 4, 6, 8, 10, 12, 14],  
    [0, 10, 20, 30, 40, 50, 60],  
    [0, 1000, 2000, 3000, 4000, 5000, 6000],  
    [0, 100, 200, 300, 400, 500, 600, 700, 800, 860]  
]
```

将左图中识别的外径[0,100,200,300,400,500,600]与该outer\_list进行匹配，得到最接近的[0,100,200,300,400,500,600,700,800,860]。即可对照模板将遗漏的刻度点补全。



# 方案总结

## 最终成绩

排名	团队名	final_score	性能分	score
	MILab	0.9063	0.0926	0.8973

## 优点

- 1、目标检测算法和ocr识别算法均源自开源算法库中SOTA模型，实现简单高效，且具备丰富的工业部署落地支持。
- 2、没有采用传统直线检测的方式检测关键点，通过类似姿态估计的方式提高了关键点检测精度。
- 3、提出了一种基于查表的鲁棒仪表读数方法，对于遮挡严重的情况也能较好地处理。
- 4、通过透视变换，将形变的仪表图像修正，使得读数更加准确。

## 可改进点：

- 1、将模型文件转换为TensorRT格式，加速推理。
- 2、更改ocr识别模型，提升效率。







南京理工大学  
NANJING UNIVERSITY OF SCIENCE & TECHNOLOGY



× 感谢观看 ×

团结 献身 求是 创新

