

# 工业品组装缺陷检测项目汇报

## PRODUCT-DEFECTS-DETECTOR REPORT

数据增强-迁移学习-React-微信小程序-完成检测任务



汇报人：吴学深 殷梓淇 曲志久



校内指导老师：王政

校外指导老师：张建宇、郑艳飞

# 01

## PART ONE

# 项目介绍

本节对项目进行简要介绍，包括项目要求、项目成员及分工、项目结构、技术路线、项目完成情况等。

## 01

## 项目要求

在大规模生产中，工业品组装缺陷的自动化检测能够高效发现不合格产品，提高生产质量。本项目针对工业品组装后可能出现缺陷（缺螺丝）这一问题，训练AI模型进行快速缺陷检测。

## 零件示例



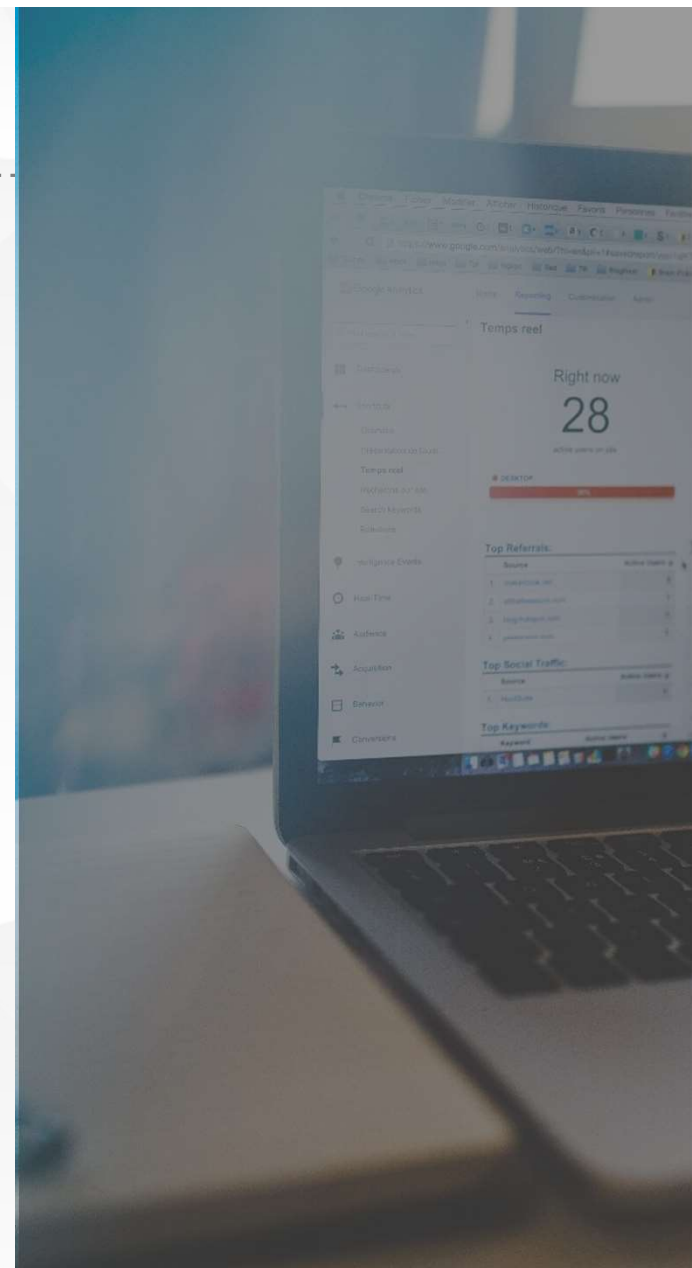
正常零件，能清晰看到四角的4个螺丝



缺陷零件，缺失至少1个螺丝



缺角零件，无法完全看到四角，需要改正拍摄角度



01

## 项目成员分组



### 项目管理组

成员：部文潇 董书玮 李秋慧  
任钰洁 曾泽环 邹蕙蔓 熊炜



### 小程序开发组

成员：殷梓淇 侯广曜



### 算法设计组

成员：吴学深（队长）  
付娆（副队长）  
徐熙（副队长）



### 推理优化组

成员：曲志久 叶郅祺 李鑫鑫



吴学深：前端逻辑与样式；后端开发、推理加速对接；小程序对接；  
算法模型开发

徐熙：可解释算法模型；算法模型开发；后端对接；小程序对接

付娆：前端设计与开发；数据预处理；数据增强

曲志久：Intel推理加速；模型压缩

殷梓淇：微信小程序开发

部文潇：模型训练；项目文档编写；答辩PPT制作



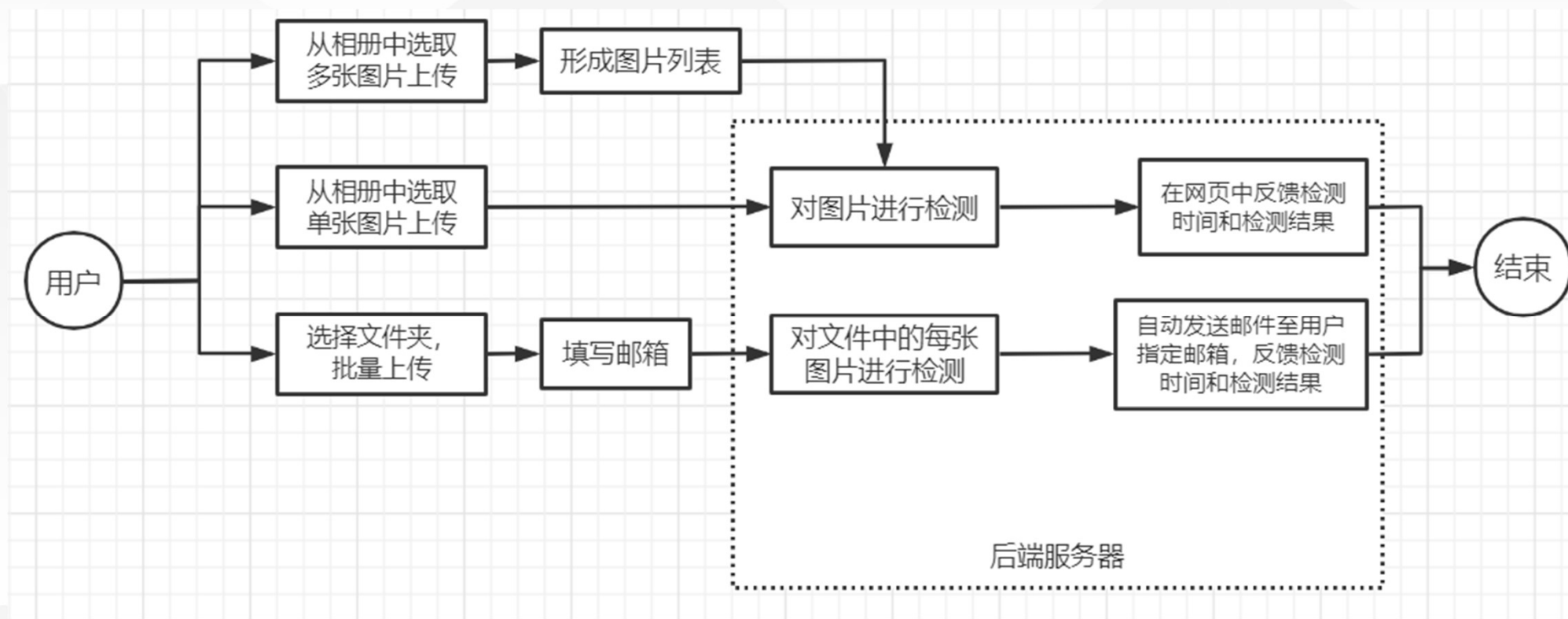


# 02

## PART TWO

### 项目演示

目前，项目已成功在web端和小程序端运行。  
请看实机演示。



Web端服务流程图





# 03

## PART THREE

### 开发历程

21天的工作中，针对缺陷检测这一核心任务，我们不断提出新思路、尝试新方案，提高检测正确率，增加项目新功能，优化用户使用体验。



由于提供的数据较少，只有一百多张，其中的好件（无缺陷零件）只有16张，为避免过拟合等问题，我们采用数据增强方法来扩充数据集。



原图



旋转90°



旋转180°



旋转270°



左右翻转



上下翻转

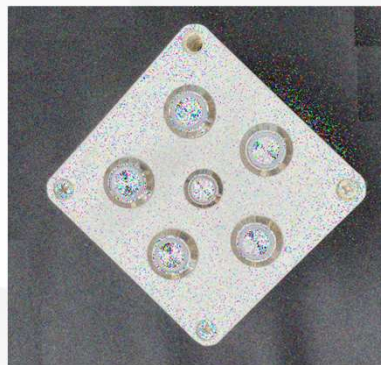
考虑到实际检测中，由于天气、时间、拍摄工具等因素影响，图片质量可能与测试集中给出的图片不同，增加了下述数据。



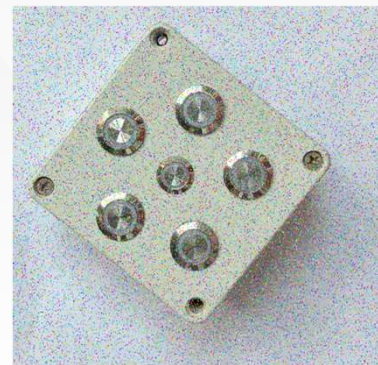
欠曝光



过曝光



高斯噪声



椒盐噪声

反复训练发现VGG16的迁移学习效果最好。

VGG16/19+MLP

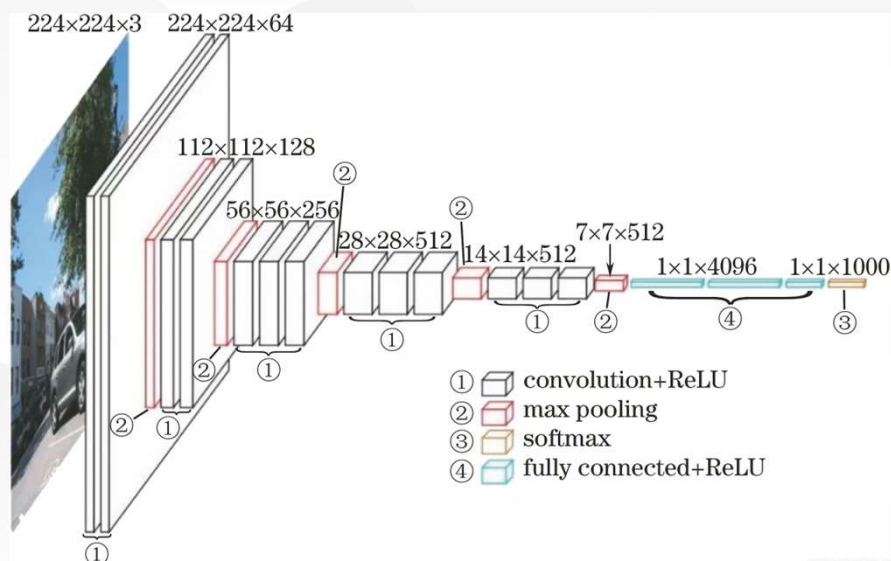
VGG16+CNN+weighted BCE

VGG16+Global Average Pooling

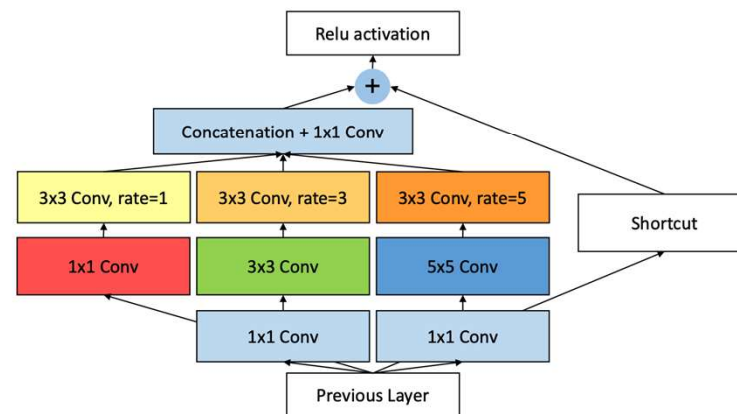
VGG16+RFB+focal loss

VGG16+ViT+focal loss

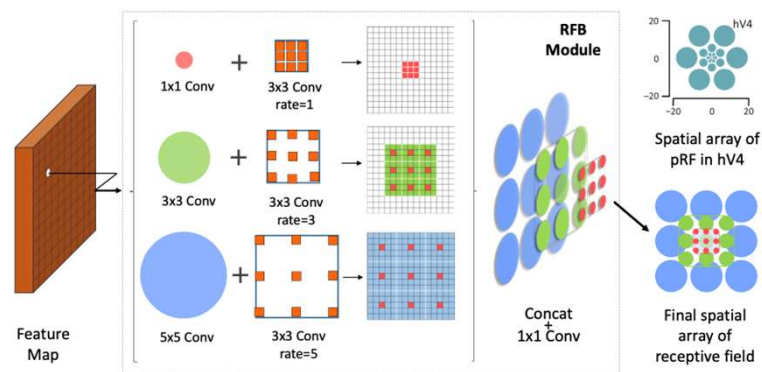
ResNet50+MLP



- 对Inception的改进
- 有效对人类视觉感受野进行模拟，能够提取深层视觉语义信息
- 采用了多分支卷积获取不同尺寸的感受野(Reception Fields)
- 采用残差连接方式缓解梯度消失

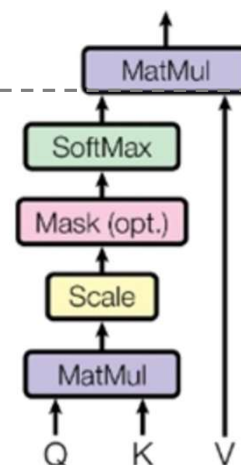


Reception Field Block

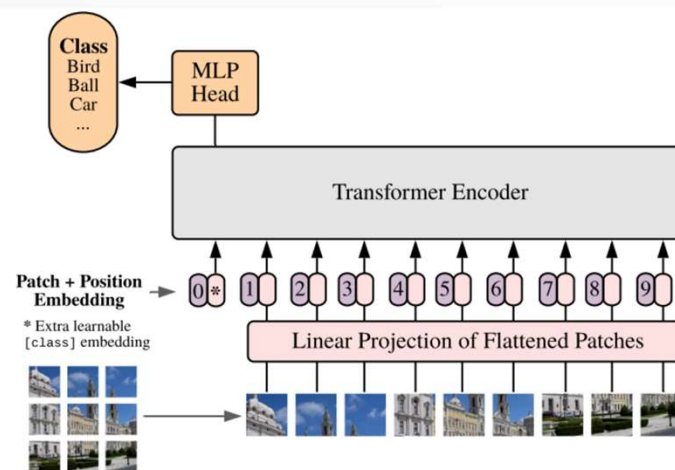


Mixed Dilated Convolution

- 将图像网格化以得到具有语义的Image Patch
- 基于多头注意力机制，能够有效对感兴趣的像素区域加以关注
- 使用Position Embedding获取时序依赖信息
- 使用与Transformer相似的编码器-解码器结构



Self- product Attention



Vision Transformer

- 取出Conv5-3 layer
- 512 个 feature maps (14\*14)
- 乘以Dense layer对应的weight
- Upsample heatmap 到原图片
- Bilinear Upsampling

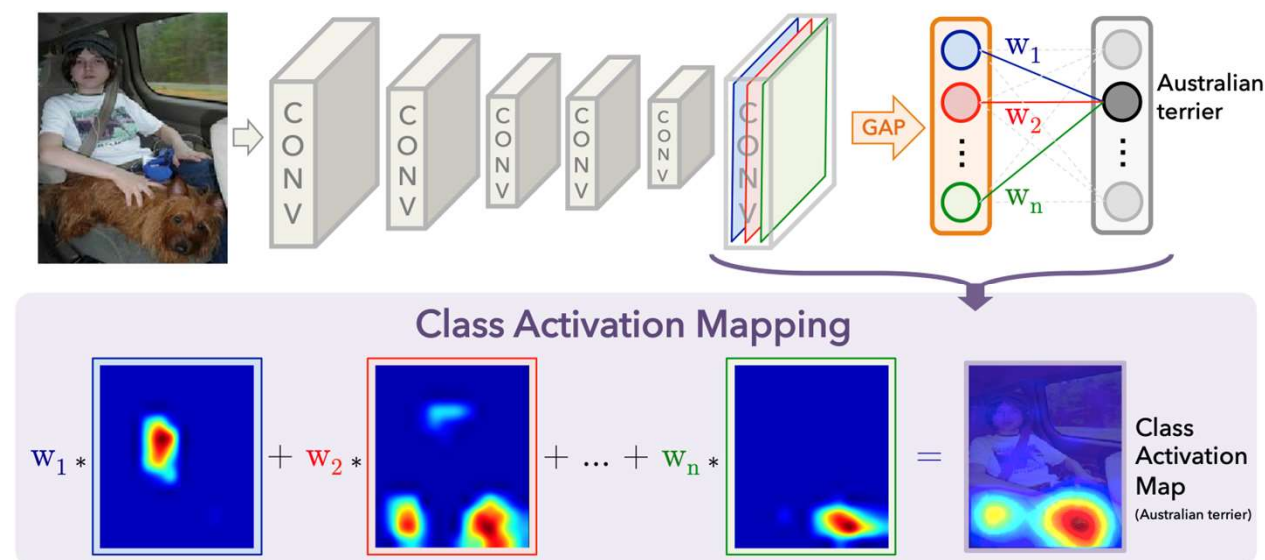
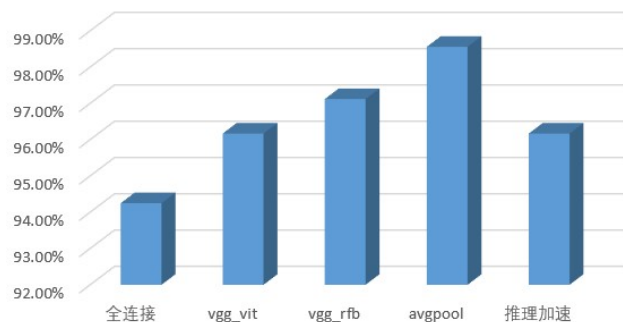


Figure 2. Class Activation Mapping: the predicted class score is mapped back to the previous convolutional layer to generate the class activation maps (CAMs). The CAM highlights the class-specific discriminative regions.

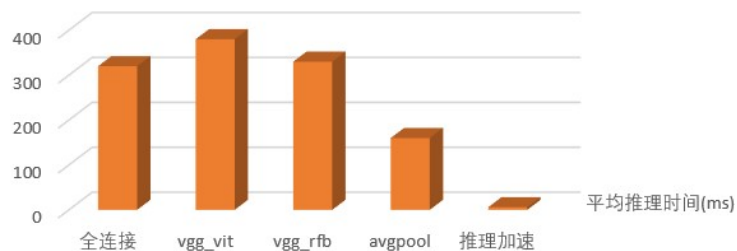


	正确率	平均推理时间(ms)	模型大小(M)
全连接	94.25%	320	60.2
vgg_vit	96.17%	380	60.8
vgg_rfb	97.12%	330	60.3
avgpool	98.56%	160	59.9
推理加速	96.17%	6	15.05

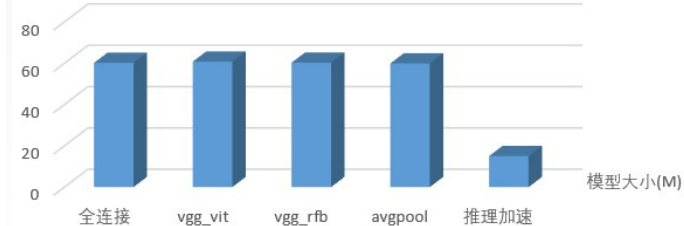
测试集正确率



平均推理时间(ms)

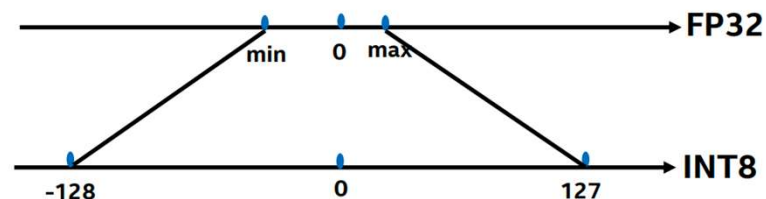
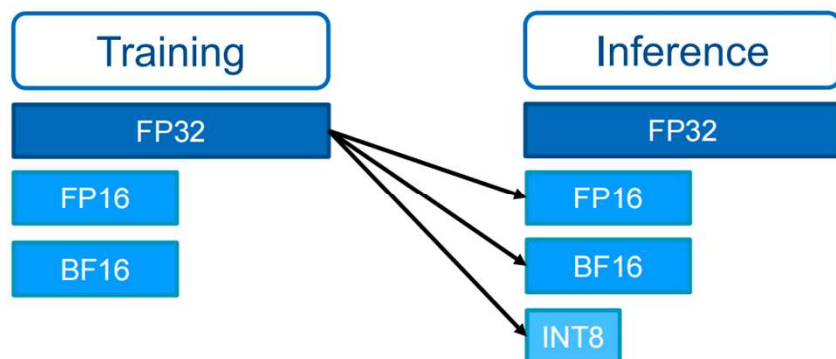


模型大小(M)



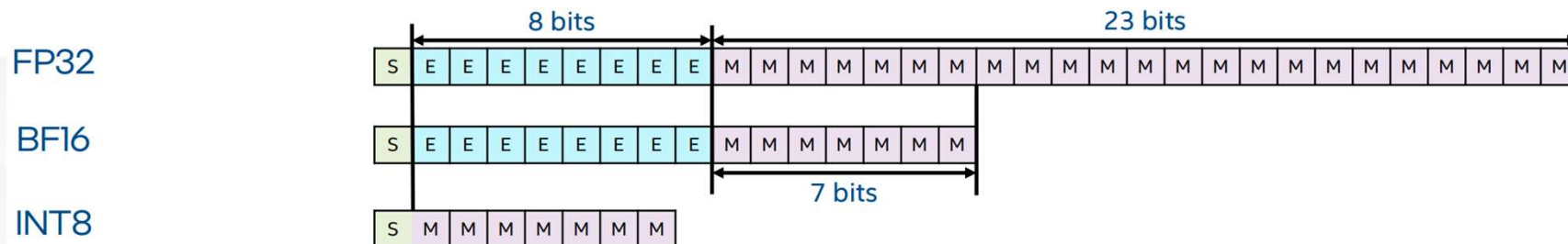
经过推理加速后的模型在推理时间、模型大小等方面有显著优势，更符合应用需求。

## 低精度推理和量化



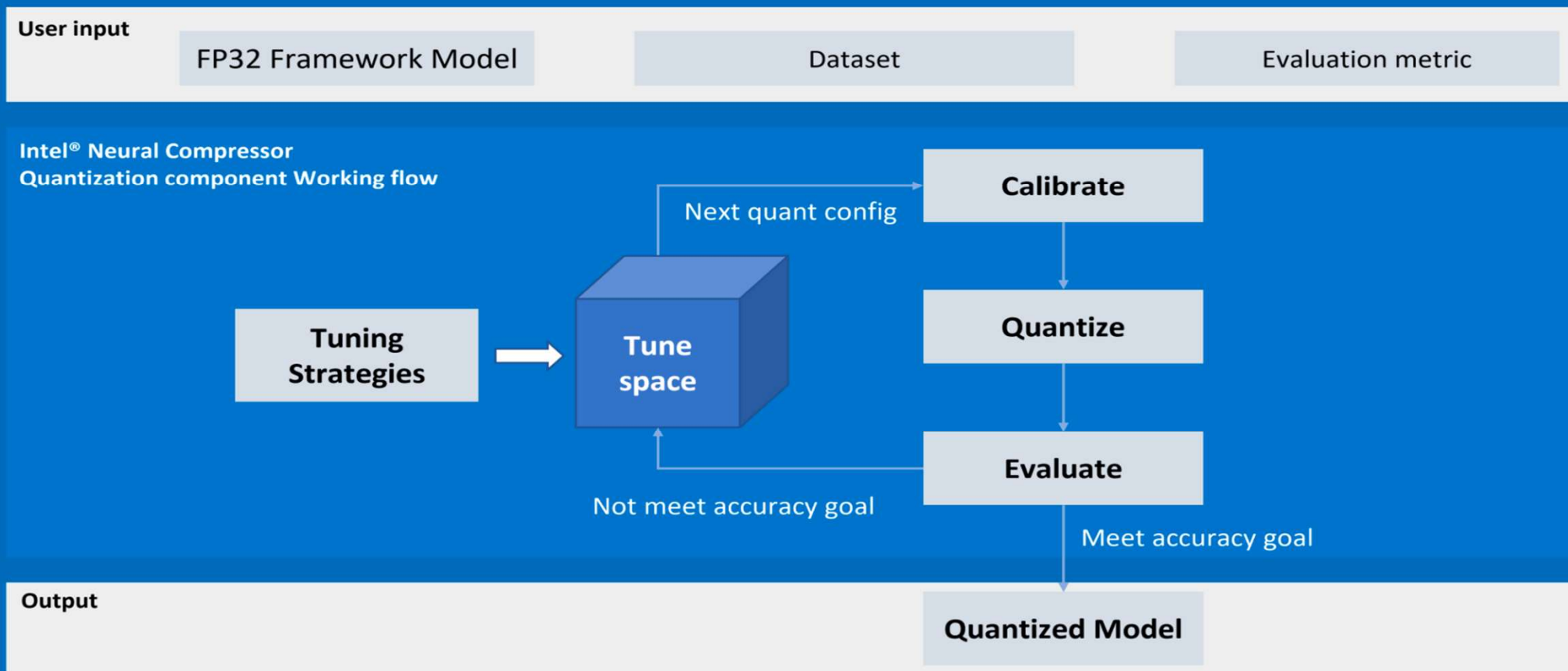
FP32 - INT8

	FP32 - 32 bits	INT8 - 8 bits	Benefit (Theory)
Memory and Bandwidth	4	1	-75%
Computing (MulAdd)	4	1	-75%
Precision	1	<1	-x%



# 工作流程

- High efficiency for quantization with automatic accuracy-driven tuning strategy.
  - Less time to achieve the quantization goals comparing with handcraft work.
  - Tuning strategy: basic, Bayesian, random, sigopt.
- Flexible workflow.
  - Dataset & metric & evaluation function not only support built-in but also customized one.



# 量化方式

Launcher

Model



Dummy Dataset



INC Quantization without Tune

Launcher

Model



Calibration Dataset



INC Quantization without Tune

Launcher

Model



Calibration Dataset



INC Quantization with Tune



Evaluation Dataset

Build-In MES Metric



INC Quantization with Tune

Launcher

Model



Calibration Dataset



INC Quantization with Tune



Evaluation Dataset

Evaluation Metric/Func

Accuracy    Performance



High

High

校内指导老师

# 感谢观看

THANK YOU FOR LISTENING

鸣谢：王政老师、张建宇老师、郑艳飞老师