

方法部分

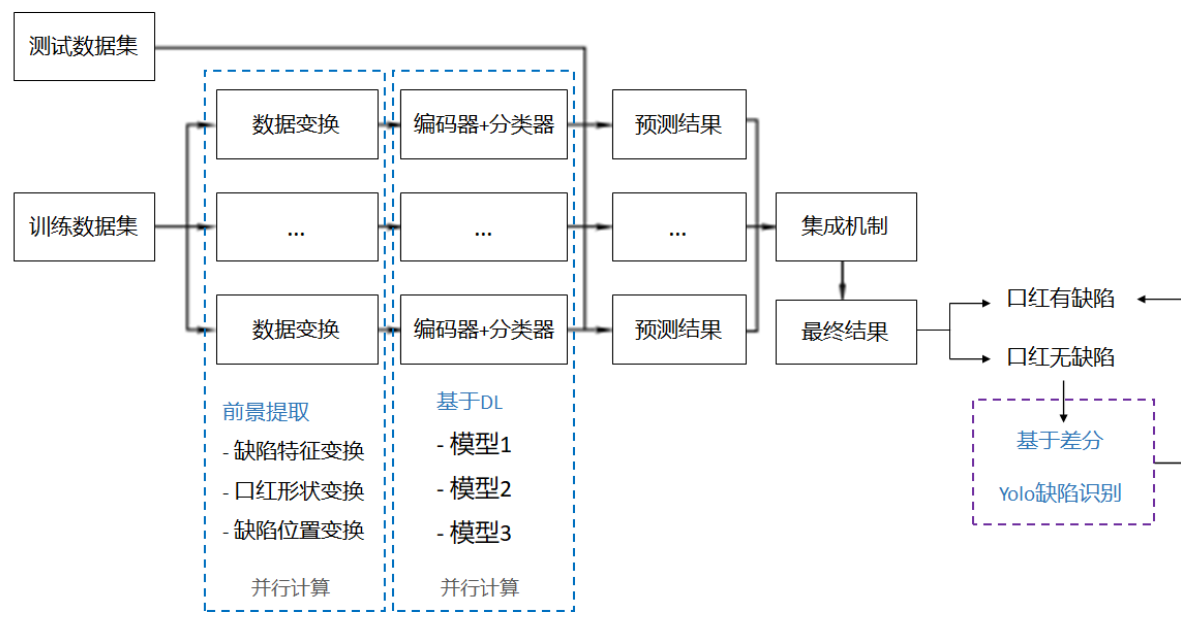
1. 总体框架

在机器学习的有监督学习算法中，我们的目标是学习出一个稳定的且在各个方面表现都较好的模型，但实际情况往往不这么理想，有时我们只能得到多个有偏好的模型。**集成学习的概念**：通过构建并结合多个学习期来完成学习任务，通过集成策略（如投票等）产生。

如下图所示，其中本方案中的异质集成 [1]（包含数据变换，编码器+分类器，预测结果，集成机制几部分），由多个不同类型算法的组件学习器（component learner）集成，因口红数据在二分类下的模型数据敏感性不同，适合采用异质集成 [2]（个体分类器算法类型不同）。通过数据增强，进一步扩充样本空间，除以提升模型泛化能力为目标，也是为后续个体学习器构造训练数据子集做准备。

此外，由于口红缺陷识别任务以召回率为红线，因此在二分类的识别基础上对判定为口红无缺陷的样本进一步进行识别，拟采用基于差分的方法及基于yolo v5的缺陷识别方法（当前仅采用yolo）。而基于差分和yolo的方法，在进一步提升召回率以确保算法能够找回所有缺陷样本的同时，支持不影响召回率的多分类识别。这种两阶段的设计思路也是为后续在保证模型召回率的前提下进一步提升准确率做铺垫。

简而言之，整体方案可看做两阶段的模型集成策略，其中第一阶段包含数据变换及异质集成，目的是在提升召回率（通过异质集成 [3]）及提升模型泛化性（通过数据变换实现 [4-5]，在1.1中介绍）。而第二阶段的yolo缺陷检测及差分算法，则是在进一步提升召回率的同时，通过两阶段的思路实现不以召回率损失为代价，而进一步提升模型分类准确率。



1.1. 数据变换

(1) 口红前景提取

该部分为数据变换预处理工作：

为了减少其他无关因素对缺陷检测的影响，本文采用Unet网络对口红膏体进行前景提取，步骤如下：

步骤1：采用labelme标注工具对口红膏体轮廓进行标注,转换json格式标签为png格式图片后得到image和label图片，以CCD1机位为例，如图1所示。

步骤2：采用Unet网络对口红膏体进行分割训练，Unet网络的结构如图2所示。

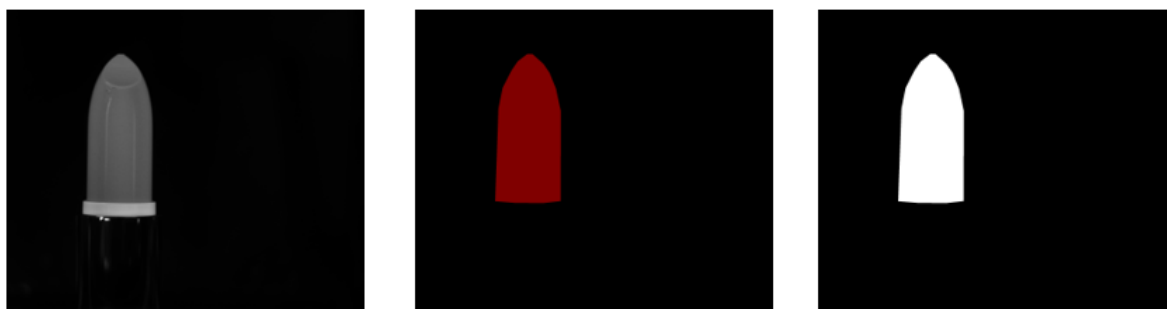


图1 image图片(左)、label图片 (中) 、颜色变换后的label图片(右)

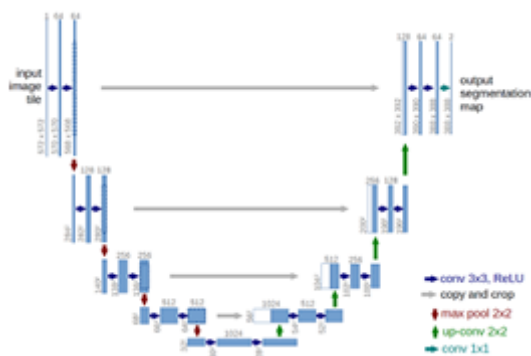


图2 Unet结构示意图

步骤3：将分割后得到的图片与原图像处理后最终得到前景提取后的图像，如图3所示。

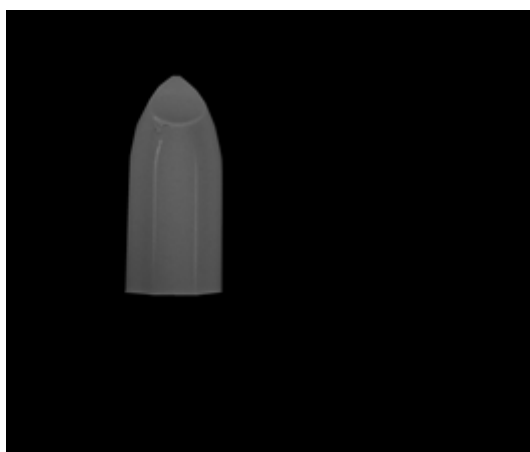


图3 前景提取后得到的图像

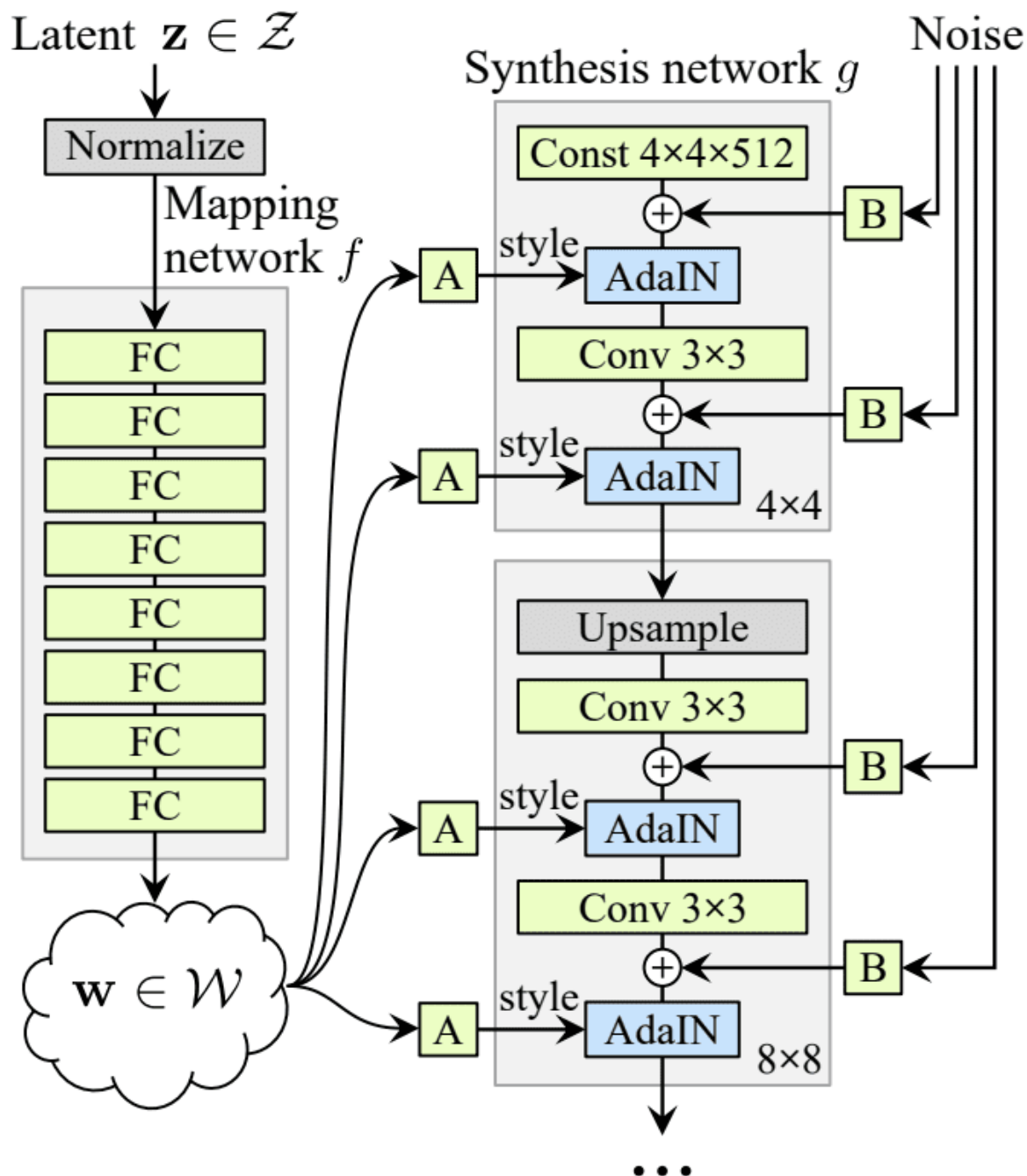
(2) 缺陷特征变换

基于标注信息，对区域内缺陷进行旋转，仿射变换，添加高斯噪声等方式，模拟当前数据未出现，但未来可能出现的缺陷实例。每类缺陷需针对其不变特征进行定制化选择变换方法，以确保生成缺陷特征的质量。

(3) 基于StyleGAN2的口红样式生成

StyleGAN中的“**Style**”是指数据集中口红的主要属性，比如口红的形状大小等信息，而不是风格转换中的图像风格。该模型通过隐空间对于图像属性的解耦，来可控的生成指定样式的口红图像。

StyleGAN 的网络结构包含两个部分，第一个是**Mapping network**，即下图 (b)中的左部分，由隐藏变量 z 生成中间隐藏变量 w 的过程，这个 w 就是用来控制生成图像的style，即风格，为什么要多此一举将 z 变成 w 呢，后面会详细讲到。第二个是**Synthesis network**，它的作用是生成图像，创新之处在于给每一层子网络都喂了 A 和 B ， A 是由 w 转换得到的仿射变换，用于控制生成图像的风格， B 是转换后的随机噪声，用于丰富生成图像的细节，即每个卷积层都能根据输入的 A 来调整“style”。整个网络结构还是保持了 **PG-GAN (progressive growing GAN)** 的结构。



(3) 缺陷位置变换

通过对生成的缺陷特征及口红样式，进行叠加组合，构造增强后的口红数据集。并供后续模型学习（实际情况根据实验选择采样子集训练或整体训练）。

1.2. 基于DL的口红图像二分类

口红二分类是简单地将口红划分为以下两类，no_defect(无缺陷口红图像)和with_defect(有缺陷口红图像)。口红二分是只针对于有无缺陷而不对口红具体缺陷进行检测的一种分类方案。本项目采用VGG网络对口红图片进行二分类检测。

VGG由Oxford的Visual Geometry Group提出。VGGNet的结构非常简洁，构建较为容易，因此本项目采用VGG19模型对口红照片进行二分类以区分缺陷口红和无缺陷口红，二分类步骤如下：

步骤1：将前景提取后得到的图像根据四个机位以及有无缺陷按照一定比例随机划分训练集和测试集。

步骤2：将训练集分别放入no_defect(无缺陷口红图像)和with_defect(有缺陷口红图像)文件夹中，并送入VGG19模型进行训练。

步骤3：训练一定周期后，采用训练好的模型对测试集进行检测，并得到对于测试集的查准率和查全率。本项目的识别结果都以验证集为基础，以口红支数为单位，其中查准率的公式为：AI识别结果中包含人工识别结果的数量÷AI识别结果，查全率的公式为：AI识别结果中包含人工识别结果的数量÷人工识别结果。查准率和查全率的计算示意图如图4所示。

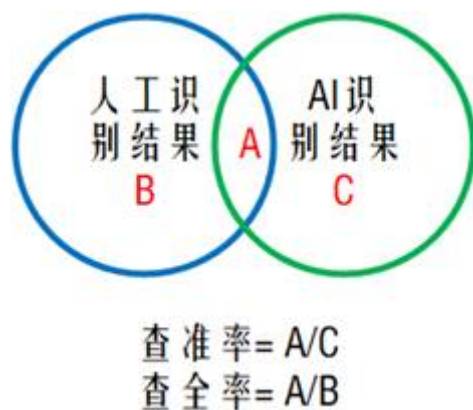


图4 查准查全计算示意图

总体框架部分参考文献：

- [1] Dong, Xibin, et al. "A survey on ensemble learning." *Frontiers of Computer Science* 14.2 (2020): 241-258.
- [2] Wang, Feng, et al. "An ensemble learning based prediction strategy for dynamic multi-objective optimization." *Applied Soft Computing* 96 (2020): 106592.
- [3] Yang, Xiyun, et al. "Image recognition of wind turbine blade damage based on a deep learning model with transfer learning and an ensemble learning classifier." *Renewable Energy* 163 (2021): 386-397.
- [4] Wu, Sen, et al. "On the generalization effects of linear transformations in data augmentation." *International Conference on Machine Learning*. PMLR, 2020.
- [5] Munoz-Bulnes, Jesús, et al. "Deep fully convolutional networks with random data augmentation for enhanced generalization in road detection." *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2017.