

东北大学

# 综合实训报告

接受实习单位: 东北大学

实习日期: 2022 年 5 月 14 日~2022 年 6 月 15 日

计算机学院人工智能专业2019 年级1 班

姓名: 孙翰文 学号 20194943

实习分配班级: 医学影像智能分析班

实习类型: ☐金工实习☐认识实习☒生产实习☐毕业实习

## 实习报告撰写的内容与要求

1. 实习任务：介绍实习的目的、意义、任务及实习单位的概况等内容。通常以前言或引言形式表述，不单列标题及序号。
2. 实习内容：先介绍实习安排概况，包括时间、地点、内容等，然后逐项介绍具体实习流程与实习工作内容，以及专业知识与专业技能在实习过程中的应用。本部分内容应以记叙或白描手法为基调，在完整叙述的基础上，对自己认为有重要意义或需要研究解决的问题进行重点叙述，其它内容则可简述。
3. 实习结果：围绕实习任务要求，对实习中发现的问题进行分析、思考，提出解决问题的对策、建议等。分析问题、解决问题要有依据（如有参考文献可在正文后附录）。分析讨论的内容、推理过程及所提出的对策与建议作为实习报告的重要内容之一，是反映或评价实习报告水平的重要依据。
4. 实习总结或体会：对实习效果进行综合评价，着重介绍自身的收获与体会，内容较多时可列出小标题，逐一列举。总结或体会的最后部分，应针对实习中发现的自身不足，简要地提出今后学习，努力的方向。
5. 将实习日记按照时间顺序以附件形式放在实习报告正文后面。
6. 实习报告正文一律采用计算机排版、A4 纸**双面**打印，正文字体为**小四号**宋体，**1.35 倍**行距，正文页数不低于 15 页，一般不超过 35 页（其中专家报告一共不超过 3 页，**中文字数**不少于 5000 字）。每章标题为小三号黑体字居中。全文要求语句通顺、论述严谨规范。
7. 请注意封面页、扉页、评语页等的打印及装订顺序。分别为①封面页（其背面为扉页“实习报告的内容与要求”）、②目录、③正文、④附录-实习日记、⑤其他附件（如校外实习单位的鉴定材料、专利、获奖证书等）、⑥评语页。

## 目 录

1. 引言.....	1
1.1 实习目的和要求.....	1
1.2 实习内容.....	2
1.3 实习单位概况.....	2
2. 智能眼底影像分析任务.....	4
2.1 任务简介.....	4
2.2 模型介绍.....	5
2.3 数据处理.....	6
2.3.1 数据扩充:.....	6
2.3.2 数据预处理.....	8
2.4 训练过程.....	9
2.4.1 预训练参数:.....	9
2.4.2 调参优化.....	9
2.5 实验总结.....	11
3. 心室分割与心脏病分类.....	12
3.1 任务简介.....	12
3.2 数据分析.....	12
3.3 模型介绍:.....	13
3.4 实验过程.....	14
3.4.1 分割任务.....	14
3.4.2 分类任务.....	15
3.5 实验总结.....	16
4. 低剂量 CT 图像质量优化.....	17
4.1 任务简介.....	17
4.2 数据处理.....	17
4.3 模型尝试:.....	18
4.4 模型简介:.....	18
4.5 模型训练:.....	19
4.6 测试结果与总结.....	20
5. 实习总结.....	21
5.1 实习内容的复杂性评价.....	21
5.2 实习体会、收获与建议.....	22
6. 参考文献.....	23
附录 实习日记.....	24



## 1. 引言

### 1.1 实习目的和要求

#### 1.1.1 实习目的

- (1) 通过计算机系统的设计与开发，验证和扩展本科前三年所学相关课程的理论知识，在实践中深入理解其中的重点和难点，提高基本理论水平，巩固基础知识体系，掌握计算机系统设计与开发的基本技能，培养灵活运用理论知识分析和解决实际问题的能力。
- (2) 通过团队合作开发，培养团结协作的能力。
- (3) 通过专家讲座，了解本学科的科学前沿及进展，拓宽知识面，拓展学术和国际视野，培养创新精神。
- (4) 通过到企业参观与交流，学习课堂以外的实践知识，了解所学专业在社会经济建设中的地位、作用和发展趋势，以及社会对本专业毕业生的素质要求，为最后一学年的课程学习和毕业设计打好基础。
- (5) 通过撰写设计需求文档、实习报告和答辩文档，培养学生查阅中英文文献的能力、锻炼口头和书面的表达能力和沟通能力。

#### 1.1.2 实习要求

##### 一般要求

- (1) 在掌握计算机基本程序设计语言的基础上，利用所学的数学知识、自然科学知识、工程基础知识以及计算机专业知识对所开发的系统进行软硬件分析和建模。
- (2) 在系统的分析与设计过程中，要多方面查阅资料，并利用形式化方法表达所开发系统中涉及的工程元素，采用适当的抽象模型予以合理的解决。
- (3) 要综合考虑系统涉及的各种因素，设计并开发出满足特定要求的软件系统或功能部件。在解决问题时，鼓励采用创造性思维方式解决问题。
- (4) 在设计与开发过程中，应选择主流的开发工具及相关资源，以保证系统的先进性、可兼容性和可扩展性。
- (5) 在完成功能完善、技术复杂的系统时，允许分组分工，但每个学生应努力完成自己的职责并彼此配合。
- (6) 应认真听取专家讲座，了解本专业的前沿知识，并写入生产实习报告。

- (7) 在企业参观与交流时,应严格遵守纪律,保证自身及企业生产安全,认真听取企业专家的介绍,仔细查看了解生产车间各要素的功能,并写入生产实习报告。
- (8) 验收时,在实验环境下回答指导教师的提问,力求准确、全面、流利。
- (9) 撰写依据可靠、方案完整、数据准确、格式规范的实习报告。

### 1.1.3 课题要求

初步掌握 python 语言和 pycharm 开发工具。熟悉 pytorch 等深度学习框架的使用,巩固图像处理和深度学习的理论基础,掌握数据处理的流程,使用计算机视觉的方法完成医学影像分析任务。

要求使用现有的深度学习方法和框架,在给定的具体数据集下,设计一系列模型完成三项医学影像分析任务;任务以小组为单位,互相协作,分工明确;严格出勤;认真完成生产实习报告,并提交程序。

## 1.2 实习内容

- (1) 眼底影像分析: 实现眼底图像病灶检测及辅助疾病诊断。(眼底影像四分类)。
- (2) 心室分割与心脏病分类: 实现 MRI 影像心脏组织分割与分类(语义分割 + 疾病三分类)。
- (3) 低剂量 CT 图像质量优化: 实现低放射剂量的同时保证 CT 成像质量(图像去噪/ 图像风格域迁移)。

## 1.3 实习单位概况

东北大学计算机国家级实验教学示范中心 2016 年获批建设,由公共基础实验教学部、专业实验教学部、教学研究和创新实践部、中心办公室四部分组成,实行教学指导委员会领导下的主任负责制,中心实验教学用房面积约 4000 平方米,实验设备 7000 余台套,共计价值超过 2900 万元。中心现有专任教师 132 人,实验技术人员 25 人,合计 157 人,其中教授 38 人,副教授 61 人,高级实验师 11 人。共承担实验课程 60 余门,开设开放性实验课程 8 门,独立设课实验 2 门,平均每年承担实验教学任务超过 60000 人时。

中心坚持以“立德树人、培养满足社会需求的高级人才”为目标,以“重视基础训练,强化专业能力,支持科技创新”为教学理念,以“CDIO 工程教育”为体系模型,结合东北大学计算机学院实际情况,针对计算机科学与技术、物

联网工程、通信工程、电子信息工程四个专业和非电专业学生的差异制定不同的实践能力培养方式和目标。面向非计算机专业学生保证计算机基础实践能力的培养，注重介绍计算机在主修专业上的应用和扩展，鼓励和帮助学生使用计算机技术解决其主修专业的计算问题，并进行深度探究和创新实践；面向计算机类和电子信息类专业，注重各专业理论知识和实际应用相结合，课程实践与课外科技竞赛相结合，加强实践能力培养力的宽度和深度，并通过实验扩展和课程整合提高学生解决复杂工程问题的能力，进行扩展性、开放性和研究性实验教学以提高学生综合实践能力。

2016 年获批国家级实验教学示范中心以来，中心着重在实验教学体系建设、实验教学改革、产学研合作、科技创新实践、人才培养、智能化和信息化建设、社会辐射等方面开展工作。中心在学校和学院专项经费支持下开展实验设备和仪器更新、实验教学内容改进、教学资源建设、教学模式研究、课程体系改革等方面的教学改革项目共 50 余项。与百度、华为、东软等国内知名行业企业深度合作，在课程设计、生产实习、课外科技创新等实践环节引入企业讲师和企业项目案例，建设校外实习基地 6 个，获批教育部产学研协同育人项目 20 余项，引进或开发实验教学综合管理平台、综合实验教学平台、远程自主学习和在线考试平台、虚拟仿真实验平台等信息化系统，更好地开展实验教学和示范中心管理工作。另外，中心平均每年承担辽宁省、沈阳市各类考试、培训及竞赛 20 余次，服务人数超过 13000 人次。

经过三年的建设和发展，在中心师生的共同努力下，中心共建设 MOOC 课程资源 5 门，其中国家级精品课程 2 门，辽宁省精品资源共享课 1 门，出版实验教材 8 部，先后完成了省部级教学改革项目 10 余项，发表教改论文 50 余篇，申请各类专利 40 余项，获得国家级教学成果奖 1 项，省部级教学成果奖 14 项，指导学生获得各类国家级科技竞赛一等奖 40 余人次，孵化创新创业项目 20 余项。

## 2. 智能眼底影像分析任务

### 2.1 任务简介

分级任务主要目的是分析 2D 眼底图像的临床数据。根据视觉特征病灶 信息区分糖尿病视网膜病变的严重程度。一共有四种不同程度的病变。

数据集分析:

我们拿到手的原始数据集内容是一千张眼底图像和对应的边界框标签数据。

眼底分级

|---train (图片)

|---Mess1\_annotation\_train.csv (标注数据)

image	quality	type	Retinopathy_grade	Risk_of_macular_edema
20051019_38557_0100_PP	0	_000	3	1
20051020_43808_0100_PP	1	_101	0	0
20051020_43832_0100_PP	2	_111	1	0
20051020_43882_0100_PP	2	_111	2	0

原始数据集中各等级的图片是混在一起的，我们要根据表格的 retinopathy\_grade 列将它们分别存到四个文件夹中。

将项目和数据文件按照如下路径设置并运行 preparing\_data.py 即可。

本地项目的文件目录部分内容如下（下载的处理好的数据集要放在此文件夹中）：

efficientnetV2\_for\_classification

|---weights

|---runs

|---save\_weights

|---estimate\_info

|---datasets\_for\_efficientnetV2

|---0

|---1

|---2

|---3

|---train.py

|---predict.py

|---preparing\_data.py



```
|---pre_efficientnetv2-s.pth
```

1000 张数据集样本类别情况:

病变级别	类别数量
0	461
1	136
2	207
3	196

## 2.2 模型介绍

我们选用 efficient net v2 来做分类任务。

Efficient Net 的作者通过调整网络的宽度、深度以及输入网络的分辨率来提升网络的性能，最终得到了一系列适用于不同分辨率的模型。但是第一代 Efficient Net 在训练时却很耗费显存并且收敛速度也没有达到理想值，因此作者对第一代的网络做了一些改动，使用 Fused-MBConv 替换掉了浅层的 MBConv，并且设计了渐进式的学习策略来减少训练时间。

论文地址: <https://arxiv.org/abs/2104.00298>

代码: <https://github.com/google/automl/tree/master/efficientnetv2>

网络架构:

Table 4. EfficientNetV2-S architecture – MBConv and Fused-MBConv blocks are described in Figure 2.

Stage	Operator	Stride	#Channels	#Layers
0	Conv3x3	2	24	1
1	Fused-MBConv1, k3x3	1	24	2
2	Fused-MBConv4, k3x3	2	48	4
3	Fused-MBConv4, k3x3	2	64	4
4	MBConv4, k3x3, SE0.25	2	128	6
5	MBConv6, k3x3, SE0.25	1	160	9
6	MBConv6, k3x3, SE0.25	2	272	15
7	Conv1x1 & Pooling & FC	-	1792	1

如图，这里是 EfficientNet 网络的基本框架，主要由 Fused\_MBConv 和 MBConv 构成。

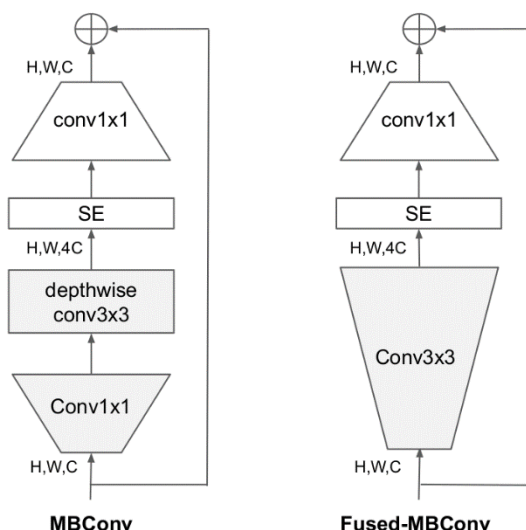
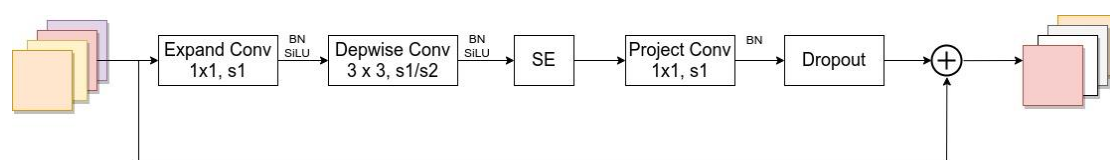
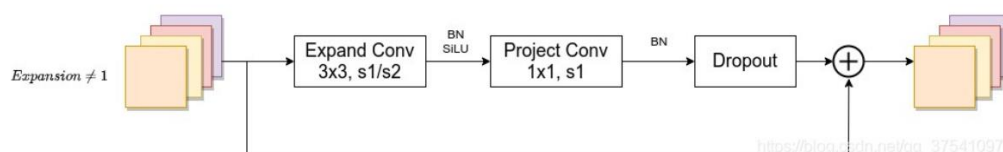


Figure 2. Structure of MBConv and Fused-MBConv.



这个是 MBConv 的具体框架。



这里是 Fused-Conv 的具体框架。

## 2.3 数据处理

### 2.3.1 数据扩充:

对于四分类而言，我们需要更多的数据进行训练。

Kaggle APTOS 2019 Blindness Detection 竞赛数据集:

[APTOS 2019 Blindness Detection | Kaggle](https://www.kaggle.com/c/aptos2019-blindness-detection)

该数据集共有五类病变级别，由于我们进行四分类，我们将该数据集的第 3 类和第 4 类合并为第三类。

病变级别	类别数量
0	1805
1	370
2	999
3	488

测试集：1982 张图片，无标签。

高质量验证集：

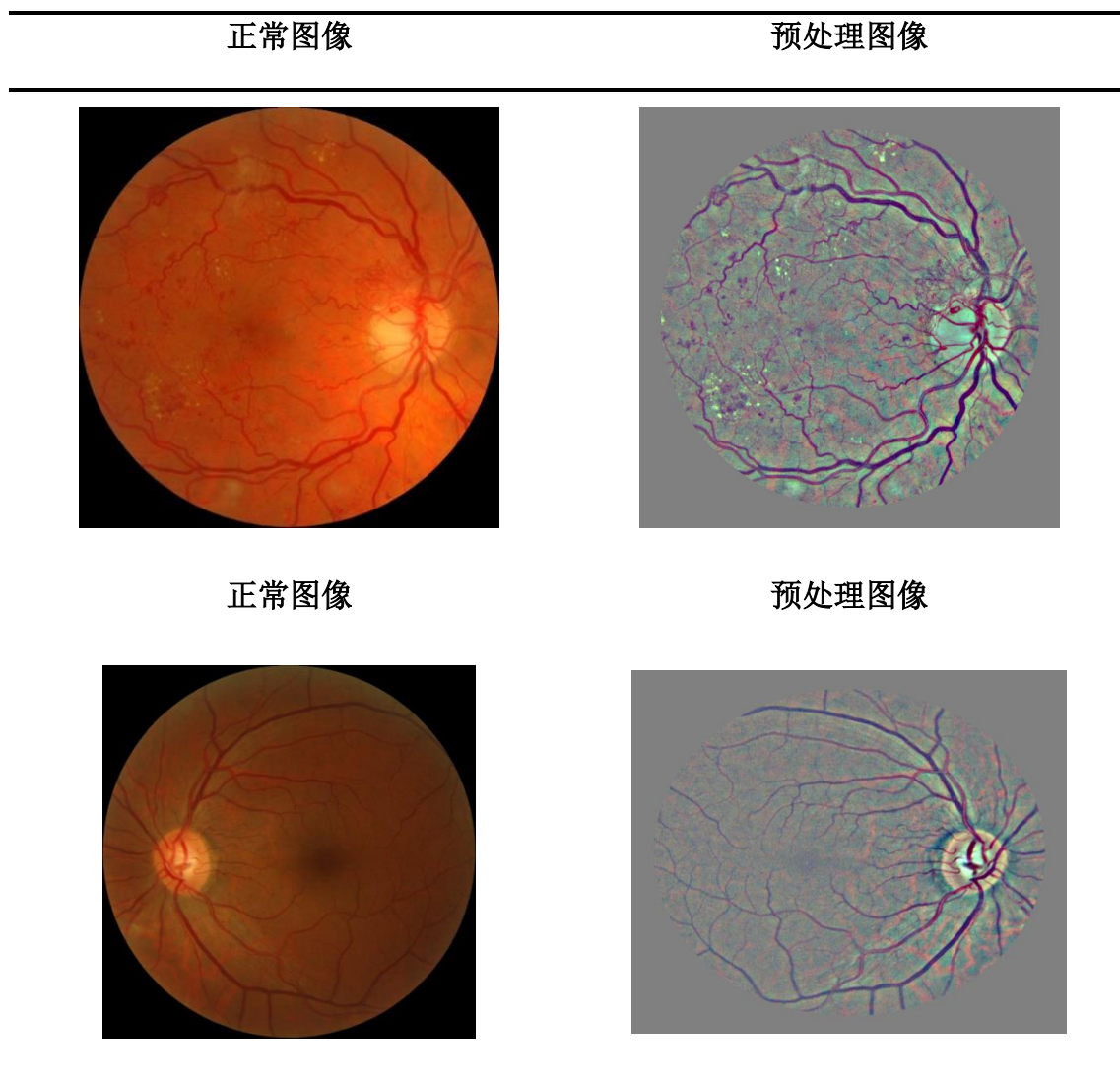
我们找到了一个约 200 张图片的和最初一千张图片相似的四个病变级别的数据集，图片质量较高，我们将其作为检验模型性能的重要测试集。

### 2.3.2 数据预处理

基于眼球的 resize

特征增强

去除眼球周围部分



实验结果表明， 数据预处理几乎没有影响， 说明我们的网络学习能力较强， 成功学习到了预处理凸显的图像特征。

## 2.4 训练过程

### 2.4.1 预训练参数:

EfficientNetv2 主要有-s, -m 两种模型

模型	参数量	性能
Efficient Net V2s	24M	9min/epoch acc>83.7%
Efficient Net V2m	55M	12.5min/epoch acc>84.3%

其中-m 模型比-s 模型参数量大, 训练时间长, 而提升有限, 所以我们都采用-s 模型继续实验。

超参设置:

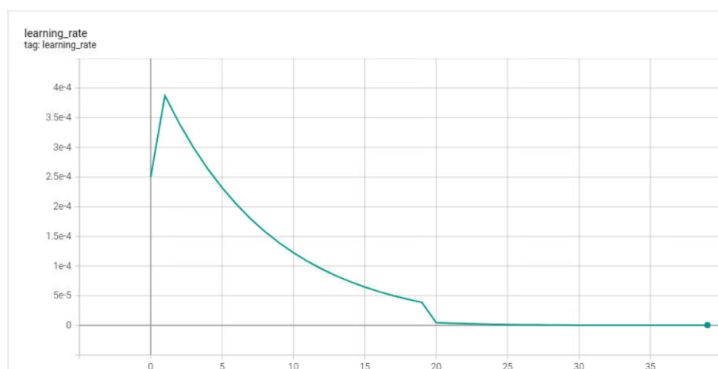
经过一系列的调参实验, 我们确定了一组超参:

参数名	参数值(整体训练/分类模块训练)
Learning rate	5E-4/5E-5
Batch size	32/16
Momentum	使用 SGD 时为 0.9
Warmup	5/2 个 epochs

### 2.4.2 调参优化

我们主要针对优化器 + 调度器的组合进行调整实验, 并确定了两种表现较好的方案:

Warmup 指数衰减 + Adam 优化器:



Warmup 余弦退火 + SGD:

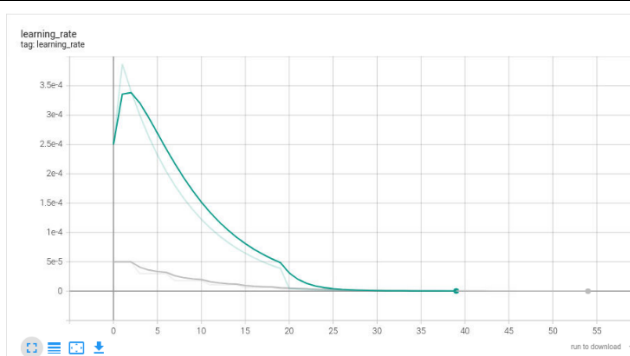


如下是我们保存的最好的一次实验效果:

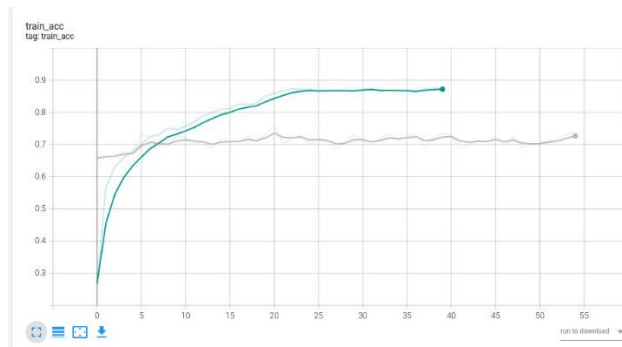
数据

图片

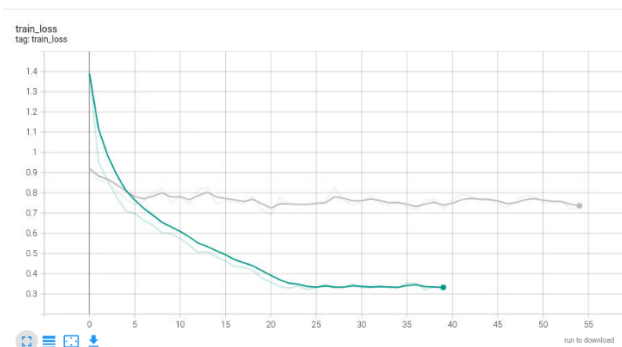
lr



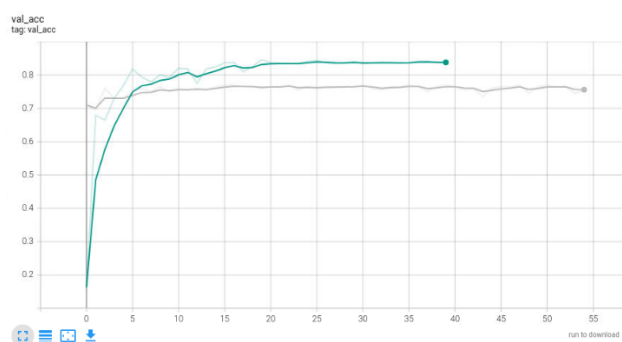
train\_acc



train\_loss



val\_acc



如图，我们本地的 val\_acc 可以达到 84%。

## 2.5 实验总结

在该任务中，我主要负责的是具体的 **efficientnetV2** 模型调参的工作。

从目前的结果来看，模型的性能还有待进一步提升，特别是对于第二级和第三级病变的判断。我们认为这一方面是由于在制定病理分级规则时本身就是将连续的病变过程进行离散化的过程，因此难免存在一些中间状态，因此如果需求允许，让模型给出类似于置信度的指标来辅助诊断或许是更好的选择。

另一方面，我们在扩充数据时发现，网上所能找到的数据集大多都是五级分类的，而我们简单地将其重组为四级或许也会导致第一级和第二级样本的特征混乱，影响模型的训练。

3. 心室分割与心脏病分类

3.1 任务简介

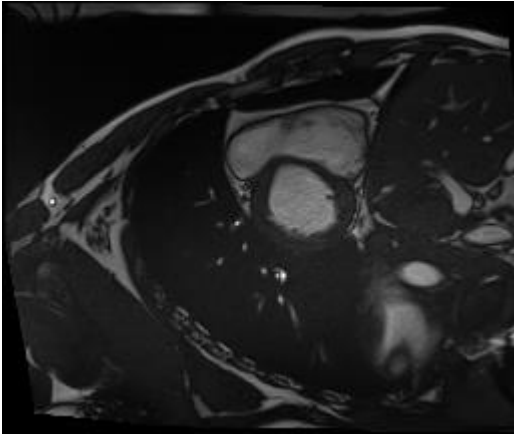
有两个子任务，一是根据心室的原图像做心室分割，二是根据原图像或分割后的图像做分类，即有无心脏病及心脏病类型。一共有三类人群的心脏核磁影像，其中两类分别患有扩张性心肌病 (DCM) 和肥厚性心肌病 (HCM)，另一类是健康人群。

3.2 数据分析

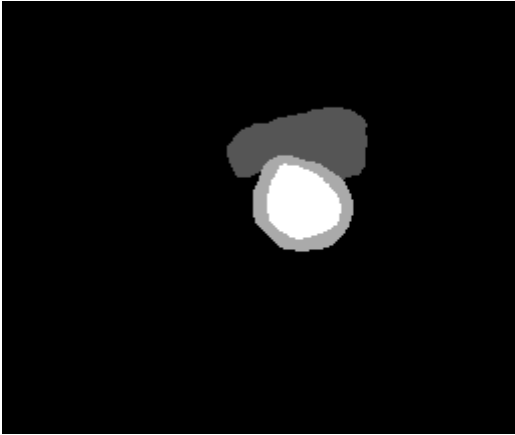
我们的原始图像是 png 格式的灰度图，是核磁影像，图像中包含了三个待分割目标，对应灰度标签为 85，170，255。

三种不同病症的诊断和分割的效果密切相关，因为他们和心室心肌的形状大小有较大的关系。

心脏 MRI 影像



分割结果



分类任务指标

label	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142
51	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
52	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
53	0	0	0	0	0	0	0	0	0	0	0	0	0	85	85	85	85	85	85	85
54	0	0	0	0	0	0	85	85	85	85	85	85	85	85	85	85	85	85	85	85
55	0	0	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85
56	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85
57	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85
58	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85
59	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85
60	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85
61	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85
62	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85
63	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85
64	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85
65	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85
66	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85
67	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85
68	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85
69	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85
70	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85
71	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85
72	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85	85
73	85	85	85	85	85	85	85	170	170	170	170	170	170	170	170	170	170	170	170	170
74	85	85	85	85	170	170	170	170	170	170	170	170	170	170	170	170	170	170	170	170
75	85	85	170	170	170	170	170	170	170	170	170	170	170	170	170	170	170	170	170	170
76	85	170	170	170	170	170	170	170	170	170	170	170	170	170	170	170	170	170	170	170
77	170	170	170	170	170	170	170	170	170	170	170	170	170	170	170	170	170	170	170	170
78	170	170	170	170	170	170	170	170	170	170	170	170	170	170	170	170	170	170	170	170
79	170	170	170	170	170	170	170	170	170	170	170	170	170	170	170	170	170	170	170	170
80	170	170	170	170	170	170	170	170	170	255	255	255	255	255	255	170	170	170	170	170
81	170	170	170	170	170	170	170	170	170	255	255	255	255	255	255	255	255	255	255	170
82	170	170	170	170	170	170	170	170	170	255	255	255	255	255	255	255	255	255	255	255
83	170	170	170	170	170	170	170	255	255	255	255	255	255	255	255	255	255	255	255	255
84	170	170	170	170	170	170	255	255	255	255	255	255	255	255	255	255	255	255	255	255
85	170	170	170	170	170	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255
86	170	170	170	170	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255
87	170	170	170	170	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255
88	170	170	170	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255
89	170	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255
90	170	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255
91	170	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255	255



**数据集介绍:****数据集结构:**

Heart Data

```
|-- Image_DCM
    |-- png
        |-- image
            |-- 不同患者的心室核磁图像;
        |-- label
            |-- 不同患者的心室分割结果;
    |-- Image_HCM
        |-- png
            |-- image
                |-- 不同患者的心室核磁图像;
            |-- label
                |-- 不同患者的心室分割结果;
    |-- Image_NOR
        |-- png
            |-- image
                |-- 不同患者的心室核磁图像;
            |-- label
                |-- 不同患者的心室分割结果;
```

**3.3 模型介绍:****分割任务:****U-net++分割网络:**

医学影像的分割任务尤其注重宏观和微观层面特征的结合：既要注重整体的结构完整，也要在边界区域精准划分界线。而且由于医学图像本身难以采集且质量不如自然图片好的特点，所使用的模型也不宜参数量过大 UNet++的优势是可以抓取不同层次的特征，将它们通过特征叠加的方式整合，加入更浅的 U-Net 结构，使得融合时的特征图尺度差异更小。

**分类任务:**

Efficient Net V2 分类网络， 在任务一中已有介绍。

### 3.4 实验过程

特别注意：该任务仅可使用老师给的数据集，即每一类大约有 150 图片和对应分割标签。

#### 3.4.1 分割任务

对于分割任务，我们采用 pytorch 开源的分割模型实现框架：segmentation\_models\_pytorch  
该库的主要功能有：

- 1、高级 API（只需两行即可创建神经网络）；
- 2、用于二分类和多类分割的 9 种模型架构（Unet、Unet++、MAnet、Linknet、FPN、PSPNet、PAN、DeepLabV3、DeepLabV3+）；
- 3、每种架构有 113 种可用的编码器；
- 4、所有编码器均具有预训练的权重，以实现更快更好的收敛。

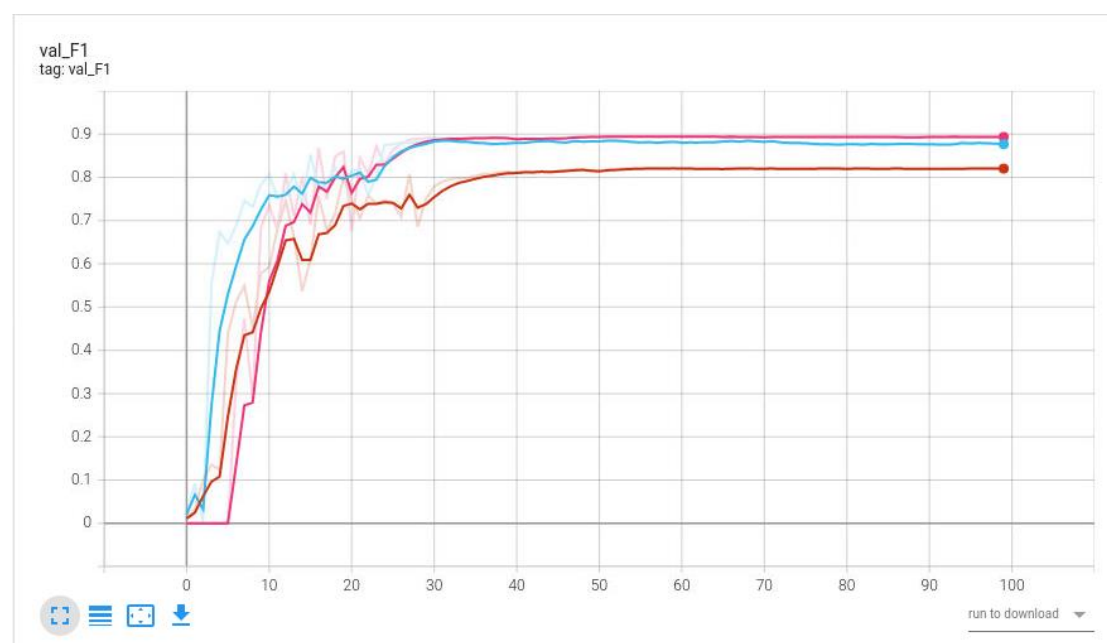
我们选取了一个可以更换很多降采样 backbone 的 Unet++ 实现方法，前期先通过实验确定最佳的 encoder 来替代原有论文中简单的降采样方法，然后再进行分割网络的微调。

如下是一些基础超参：

参数名	参数值
Learning rate	5E-4
Batch size	16
Warmup	5 个 epochs
调度策略	lr_step = [1, 0.1, 0.05, 0.002, 0.0005] / 20 轮

接着，我们针对不同的 backbone 做了一些尝试，最终选择了 resnet18 作为 backbone。

模型	参数量	性能
Resnet18 + Unet++(紫色)	11M	F1>89%
xception + Unet++(蓝色)	22M	F1>87%
Efficientnet b3 + Unet++(棕色)	10M	F1>81%



如图所示，我们本地分割的 F1\_score 可以达到 90%。

### 3.4.2 分类任务

我们沿用了任务一的 Efficientnet v2 分类网络，首先，我们考虑仅利用原图像或分割后的标签图像。实验结果表明，仅用原图像效果较差，网络不能很好的学习图像的特征，而仅使用标签图像分类效果尚可，不过可能由于缺少图像的全局信息不能达到很好的效果。所以，我们最终决定将分割后的图像和原图像按一定比例混合后(灰度值按比例混合)进行分类。

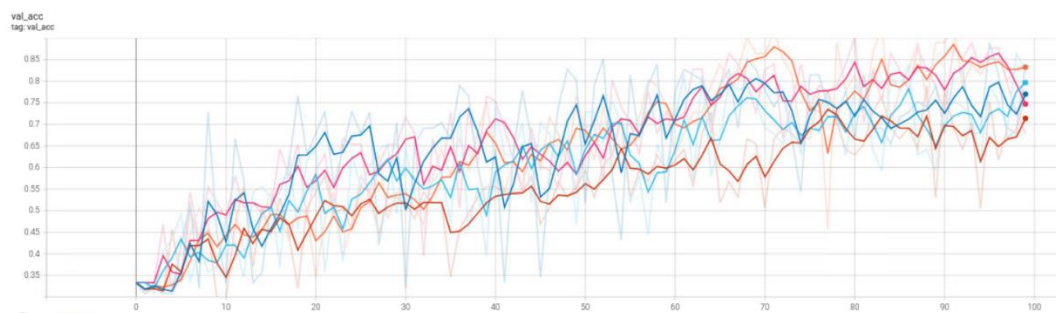
模型大小选择：对于 Efficientnet v2，我们主要尝试 -s，-m 两种模型，其中-s 是轻量级模型，-m 是中量模型。经过简单的实验，我们发现在同样的参数下，-s 的训练时间小于-m 且-s 模型的训练效果明显好于-m 模型。我们认为可能是因为混合过的模型本身较为简单，适合轻量级的模型学习。

接下来,我们调参设定了一些基础超参：

参数名	参数值
Learning rate	5E-4
Batch size	16
Warmup	5 个 epochs
指数衰减倍率	0.97/epoch

接着，我们用该组超参，在不同的混合比例下做实验：

混合比例 (Image:Label)	性能
10:0 (棕色)	F1>68%
5:5 (蓝色)	F1>80%
3:7 (紫色)	F1>86%
2:8 (橙色)	F1>88%
0:10 (深蓝)	F1>76%



可以发现，Image 和 Label 2:8 混合时，分类效果最好，在本地测试集的准确率可以达到 88%。

### 3.5 实验总结

在该实验中，我主要负责是分类任务。

针对心室分割与分类任务，我们分别采用 Unet++和 Efffcientnet v2 网络。我主要负责的是分类任务，分类任务最重要的是比较好的提取图像的特征，这里我需要考虑同时考虑原图像与分割图像的特征，并通过实验调整最佳的比例。

实验也存在一些问题。一是分割的泛化性能不足，。二是分类的思想过于简单，图像融合可以有更多的方法，比如局部融合而不是单纯按比例融合。

## 4. 低剂量 CT 图像质量优化

### 4.1 任务简介

CT(Computed Tomography), 即计算机断层成像, 能够实现快速的高空间分辨率、高密度分辨率三维成像, 被广泛应用疾病的临床诊断。CT 扫描过程会对人体造成一定剂量的电离辐射, 对患者的健康造成潜在的危害。通过低电流或电压达到降低辐射剂量, 直接导致图像噪声增加, 降低图像质量。如何实现低放射剂量的同时保证 CT 成像质量是 CT 成像领域研究的重点。

### 4.2 数据处理

数据集包含 10 组临床低剂量扫描图像和常规剂量扫描数据, 低剂量和常规剂量图像通过扫描不同的患者, 其中 8 组为训练集, 2 组数据作为测试集。图像数据由二进制文件 “\*.raw” 和信息文件 “\*.mhd” 组成。“\*.mhd” 文件包含了图像基本信息。

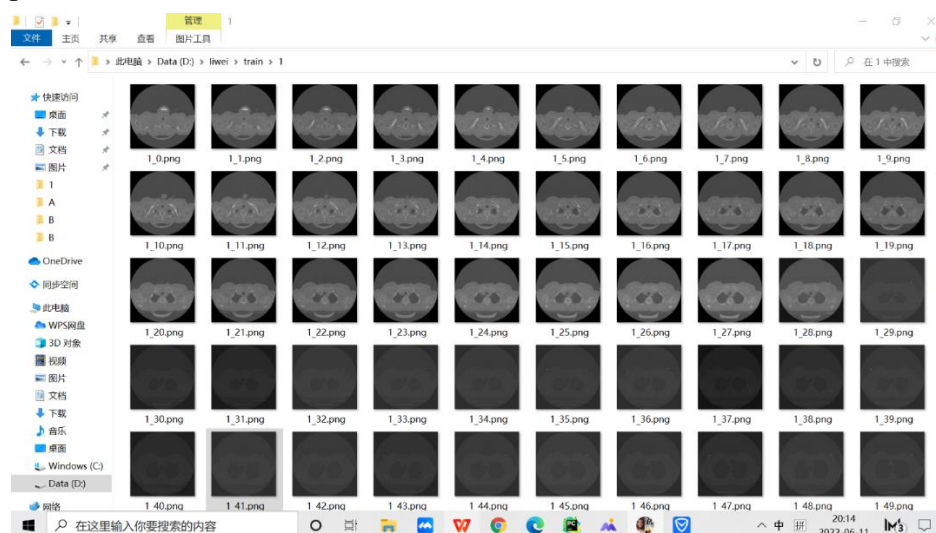
CT 图像和正常像素值在  $[-256, 256]$  之前正常图像不同, 他的 CT 值可以在  $[-3000, 3000]$ , 它和图像灰度值之间有如下映射关系:

$$HU = \text{pixel value} * \text{Rescale slope} + \text{Rescale intercep}$$

其中 Hu 值就是 CT 值, pixel value 代表灰度值。

同时, CT 图像一般是三维图像, 它的图像维度一般为:  $(300 * 512 * 512)$ , 不同维度代表在不同深度截面上扫描的人体图像。我们需要将每一维度的 CT 图像提取出来并保存为一张二维 png 图片。

具体来说, 我们使用了 python 的 Simple ITK 库读取 CT 文件, 并直接保存为 png 图片。



如图, 保存后有一组的亮度过暗, 对于正常剂量的样本, 我们通过调整对比度将图片调亮, 对于低剂量的样本我们则不处理。

### 4.3 模型尝试:

在初期，我们基于图像去噪的思路尝试，并参考了如下两篇 cvpr 论文：

Restormer: Efficient Transformer for High-Resolution Image Restoration

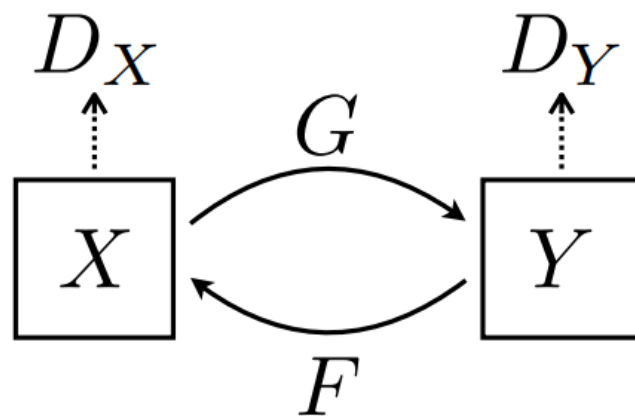
Robust Equivariant Imaging: a fully unsupervised framework for learning to image from noisy and partial measurements

最终，我们选择基于 GAN 模型的图像风格域迁移方法：

我们对比了 attentionGAN, RegGAN, CycleGAN 等主流 GAN 模型，最终选择 CycleGAN 模型进一步实验。

### 4.4 模型简介:

CycleGAN 可以在缺少成对数据的情况下，学习从源数据域  $X$  到目标数据域  $Y$  的方法(低剂量  $\rightarrow$  高剂量)。



与正常的 GAN 模型相同，定义了对抗损失：

$$\mathcal{L}_{GAN}(G, D_Y, X, Y) = \mathbb{E}_{y \sim p_{data}(y)} [\log(D_Y(Y))] + \mathbb{E}_{x \sim p_{data}(x)} [\log(1 - D_Y(G(x)))]$$

$$\min_G \max_{D_Y} \mathcal{L}_{GAN}(G, D_Y, X, Y)$$

$$\min_F \max_{D_X} \mathcal{L}_{GAN}(F, D_X, Y, X)$$

同时，定义了循环一致性损失：

$$x \rightarrow G(x) \rightarrow F(G(x)) \approx x$$

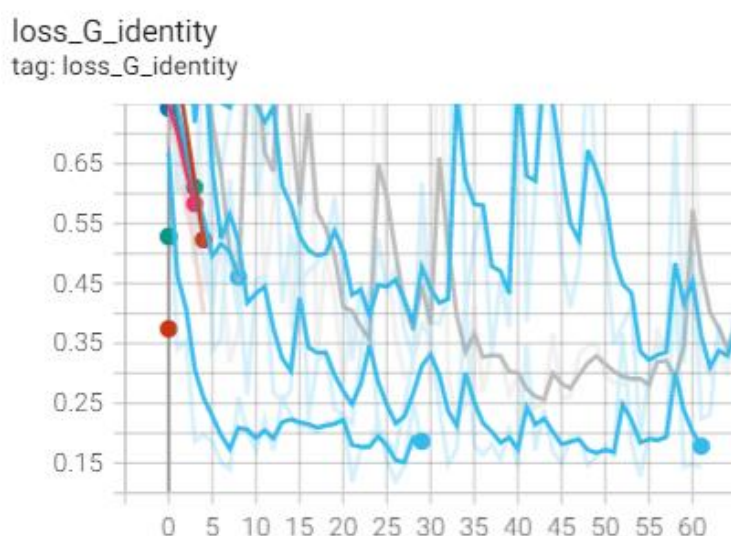
$$y \rightarrow F(y) \rightarrow G(F(y)) \approx y$$

$$\mathcal{L}_{cyc}(G, F) = \mathbb{E}_{x \sim p_{data}(x)} [\|F(G(x)) - x\|_1] + \mathbb{E}_{y \sim p_{data}(y)} [\|G(F(y)) - y\|_1]$$

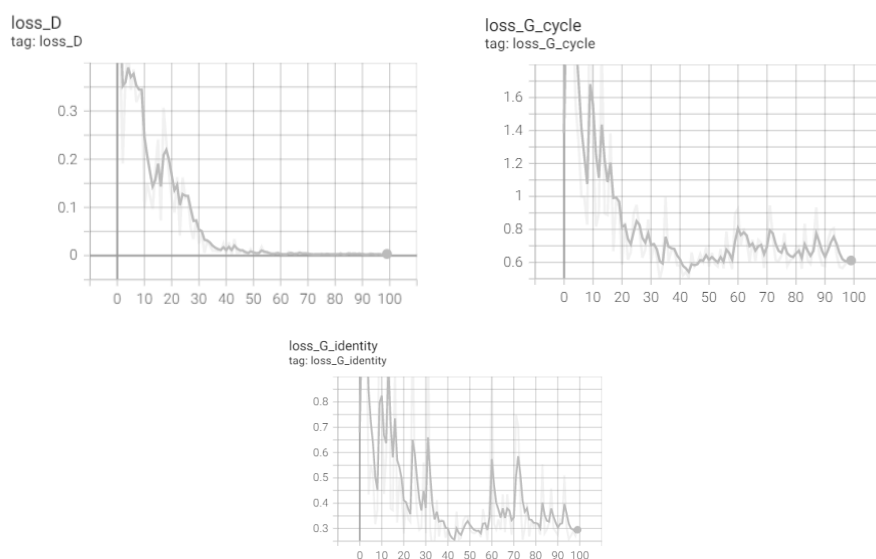
循环一致性损失是强引导信息，可以加速模型收敛，同时让模型在迁移的同时保留图像固有属性。

## 4.5 模型训练:

练策略:由于图片数量巨大,一个 epoch 约需 14 到 20 分钟,且 batchsize 最大值只能取 2 (24GB 显存), 因此我们观察前 30 个 epoch (八到九个小时) 的 Loss 下降情况和图片的迁移效果, 并以此决定是否继续训练。

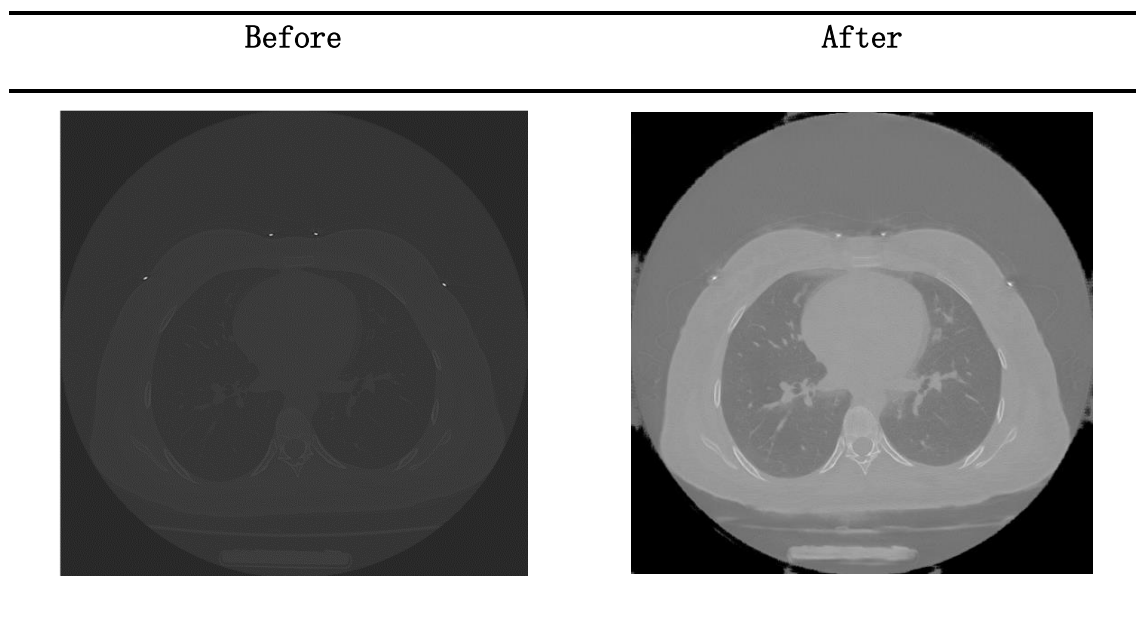


如图, 我们会观察梯度的下降情况并提前终止。  
在做了不同的优化器与调度器组合后, 我们选择了一组效果最好的:  
优化器:Adam betas=(0.5, 0.999)  
规划器:LambdaLR (50epoch 后学习率随时间减小)



如上图，是训练最好的一次结果。

#### 4.6 测试结果与总结



如图，这是在老师下发的测试集上的测试结果，效果明显比较差。

在该任务中，我主要负责的是数据处理和部分的模型实验工作。

我们在该任务中表现较差，一是 cycleGAN 可能不是非常适合这个任务，二是我们仅使用了老师给的数据，对于这种无监督学习的任务，如果自己去寻找 CT 图像加入数据集扩充可能效果会更好，三是我们调参做的不好，loss 曲线并没有讲的很好，可以在更多的硬件条件下探索更多的调参方法。



## 5. 实习总结

### 5.1 实习内容的复杂性评价

生产实习作为一种综合性的实践活动，应当验证、巩固和充实所学理论知识，加深对相关内容的理解，接触课堂以外的实践知识，拓宽知识面，培养创新精神和实践能力。其中一个重要的标准是“深入运用工程原理，解决复杂工程问题”，这其中所提到的“复杂工程问题”必须具备下述特征（1），并具备下述特征（2）-（7）的部分或全部：

- （1）必须运用深入的工程原理，经过分析才可能得到解决；
- （2）涉及多方面的技术、工程和其它因素，并可能相互有一定冲突；
- （3）需要通过建立合适的抽象模型才能解决，在建模过程中需要体现出创造性；
- （4）不是仅靠常用方法就可以完全解决的；
- （5）问题中涉及的因素可能没有完全包含在专业工程实践的标准和规范中；
- （6）问题相关各方利益不完全一致；
- （7）具有较高的综合性，包含多个相互关联的子问题。

本次实习的工程复杂性评价如表 4.1 所示：

表 4.1 工程复杂性评价表

	复杂工程问题特征	问题描述及解决方案（方法）
1	必须运用深入的工程原理，经过分析才可能得到解决	项目的三个任务都涉及了深度学习的理论和编程知识，需要具备对于神经网络模型一定程度的理解并熟悉深度学习框架。
2	涉及多方面的技术、工程和其它因素，并可能相互有一定冲突	项目涉及了深度学习的理论知识，python 编程和 pytorch 框架的使用，linux 系统的使用，还有各种任务的保存与恢复的方法。
3	需要通过建立合适的抽象模型才能解决，在建模过程中需要体现出创造性	我们首先需要针对各个不同的任务选择合适的神经网络架构，并且针对具体的数据修改不同网络模型的细节参数。
4	不是仅靠常用方法就可以完全解决的	深度学习的泛化能力往往不是很好，常常出现过拟合等问题，我们需要针对不同的数据使用一些 trick。

5	问题中涉及的因素可能没有完全包含在专业工程实践的标准和规范中	正常
6	问题相关各方利益不完全一致	无
7	具有较高的综合性，包含多个相互关联的子问题	数据处理 -> 模型选择 -> 模型调优 -> 模型泛化

## 5.2 实习体会、收获与建议

在本次实习之前，我个人对 AI 的认识和理解主要停留于理论层面，包括各种神经网络的结构。本次实践给了我一个将理论和实践相结合的途径。主要收获如下：

一 利用 `pytorch` 框架实现具体的深度学习流程，包括数据的读入，数据的预处理，神经网络的搭建等。这一方面是锻炼了我的工程实现能力，二是让我对理论知识的细节有了更深入的理解，比如神经网络的维度对齐等。

二 锻炼了对于 AI 任务的团队协作能力，我们采用 `github` 进行团队协作，将数据集与预训练权重放在谷歌云盘上。AI 任务的协作主要体现在对模型细节的调优，关键是我们规定一个模型参数的恢复与保存机制。

三 认识到了工程层面的许多不足，缺少模型调优手段，同时对不同特征，不同类型的数据处理经验十分少。我们也深刻认识到了数据对于 AI 模型的重要性。

同时，我对于 AI 相关应用也有了一些体会，我体会到了 AI 在图像领域的应用前景十分广阔，同时也意识到了数据对于现今的 AI 模型的重要性，希望今后能有更过机会接触 AI 的项目，在医学图像领域继续深造。

## 6. 参考文献

- [1]Tan, M. and Le, Q. V., “EfficientNetV2: Smaller Models and Faster Training”, <i>arXiv e-prints</i>, 2021.
- [2] @misc{Yakubovskiy:2019,  
Author = {Pavel Yakubovskiy},  
Title = {Segmentation Models Pytorch},  
Year = {2020},  
Publisher = {GitHub},  
Journal = {GitHub repository},  
Howpublished = {\url{https://github.com/qubvel/segmentation\_models.pytorch}}}
- [3]Chen, D., Tachella, J., and Davies, M. E., “Robust Equivariant Imaging: a fully un supervised framework for learning to image from noisy and partial measurements”, <i>arXiv e-prints</i>, 2021.
- [4]Waqas Zamir, S., Arora, A., Khan, S., Hayat, M., Shahbaz Khan, F., and Yang, M. -H., “Restormer: Efficient Transformer for High-Resolution Image Restoration”, <i>arXiv e-prints</i>, 2021.
- [5] @inproceedings{CycleGAN2017,  
title={Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networkss}, author={Zhu, Jun-Yan and Park, Taesung and Isola, Phillip and Efros, Alexei A}, booktitle={Computer Vision (ICCV), 2017 IEEE International Conference on}, year={2017}

## 附录 实习日记

5 月 14 日	开题， 介绍各个子任务
5 月 15 日	分析第一个任务， 数据处理。
5 月 16 日	确定第一个任务的模型和方案。
5 月 17 日	做实验。
5 月 18 日	做实验。
5 月 19 日	调参。
5 月 20 日	调参。
5 月 21 日	数据扩充。
5 月 22 日	模型优化。
5 月 24 日	确定最终模型， 准备验收材料。
5 月 27 日	第一个子任务验收。
5 月 28 日	公布第二个子任务。
5 月 29 日	处理分析数据。
5 月 30 日	确定模型方案。
5 月 31 日	做实验。
6 月 1 日	优化模型。
6 月 2 日	优化模型。
6 月 3 日	调参， 配置测试环境。
6 月 4 日	微调分割模型。

6 月 6 日	第二个子任务验收
6 月 7 日	第三个子任务验收
6 月 8 日	尝试模型去噪方案。
6 月 9 日	调整方案，尝试基于 GNN 的模型。
6 月 10 日	确定最终方案， 模型调优。
6 月 13 日	对接验收标准。
6 月 14 日	准备验收材料。
6 月 15 日	第三个子任务验收。

