

## **Title**

## **Abstract**

# Table of Contents

1	Introduction and background . . . . .	1
1.1	Introduction . . . . .	1
1.1.1	Project Overview . . . . .	1
1.1.2	Structure of the Report . . . . .	1
1.1.3	Assumptions and Technical Specifications . . . . .	2
1.2	背景 . . . . .	2
1.3	项目流程 . . . . .	3
2	Literature review . . . . .	3
2.1	切片机与显微镜的选择 . . . . .	3
2.2	关于切片组织的深度学习 . . . . .	4
3	Methodology and theory . . . . .	4
3.1	计算机视觉-图像分割 . . . . .	4
3.1.1	边缘检测 . . . . .	4
3.1.2	阈值分割 . . . . .	5
3.2	深度学习 . . . . .	6
3.2.1	卷积神经网络 . . . . .	6
3.2.2	迁移学习 . . . . .	6
4	Experimental work/analytical investigation/ design . . . . .	7
4.1	采集数据 . . . . .	7
4.2	模型 1：原始图像 + 简单的 cnn 网络 . . . . .	10
4.3	改进：图片预处理 . . . . .	11
4.3.1	边缘检测 . . . . .	11
4.3.2	阈值分割 . . . . .	14
4.3.3	另一种阈值分割方法-指纹算法 . . . . .	14

4.3.4 小结 . . . . .	15
4.4 模型 2: 预处理图像 + 简单的 cnn 网络 . . . . .	15
4.4.1 小结 . . . . .	15
4.5 模型 3: 原始图像 + 迁移学习 . . . . .	16
4.5.1 小结 . . . . .	17
4.6 模型选择总结 . . . . .	18
5 Presentation of experimental or analytical results/descriptions of final constructed product . .	18
5.1 带入测试集验证准确度 . . . . .	18
5.2 判断最佳切削角度 . . . . .	18
5.3 模型的进一步提高（改变输入分辨率）（数据增强在这里） . . . . .	19
5.4 模型通用性 . . . . .	20
6 Discussion and conclusions . . . . .	21
6.1 讨论 . . . . .	21
6.2 Future work . . . . .	21
6.3 性能提升与优化 . . . . .	22
6.4 总结 . . . . .	22
7 Project management, consideration of sustainability and health and safety . . . . .	22
7.1 Subsection 5.1 . . . . .	22
7.2 Subsection 5.2 . . . . .	22

# 1 Introduction and background

## 1.1 *Introduction*

### 1.1.1 *Project Overview*

这个项目致力于优化生物组织切片机的切割参数，这是生物医学研究和临床诊断中的关键设备。目标是通过确定最佳的切片条件，提高组织样本准备的精度和效率。通过收集在各种切割参数下的组织样本，并进行后续的人工图像分类，本研究采用深度学习技术来分析和预测最有效的切割参数。这项工作不仅有望提高显微检查的组织样本质量，还有助于简化实验室的工作流程，从而推动生物和医学科学的进步。

#### **Objectives:**

1. 收集在不同参数下切割的组织样本的全面数据集。
2. 使用人工图像分类来对这些样本的质量和特征进行分类。
3. 开发和训练一个能够评估组织样本质量的深度学习模型。
4. 利用模型的见解来确定组织切片仪的最佳切割参数。
5. 通过实证测试和改进来验证模型的预测结果。

### 1.1.2 *Structure of the Report*

本项目分为以下几个章节，每个章节都旨在系统地探索研究背景、方法论、实验工作、结果展示、讨论和结论，以及项目管理、可持续性和健康安全方面的考虑：

**引言和背景** - 本章概述了项目的目标、目标和结构安排。它简要介绍了研究的动机和必要性，以及采用的技术协议和规范。

**文献综述** - 对生物组织切片、图像分类和深度学习在生物样本制备中的应用进行深入讨论。本节将当前研究定位于现有研究的背景下。

**方法和理论** - 详细描述了实验方法、理论框架以及数据收集和处理的具体计划。

**实验工作/分析调查/设计** - 描述了实验设计、实施和分析调查的详细步骤。详细阐述了实现项目目标所采用的策略和方法。

**实验或分析结果展示/最终构建产品描述** - 本章展示了实验数据、分析结果或最终设计产品，详细描述了实验或设计结果。

**讨论和结论** - 对结果进行分析，讨论其科学意义和实际价值。本章还提供研究结论，并提出未来研究的潜在方向。

**项目管理、可持续性和健康安全考虑** - 讨论项目管理策略、可持续性问题和健康安全措施，以确保研究工作的高效和安全进行。

**参考文献** - 列出所有引用的文献资料，支持研究并为研究提供基础。

### **1.1.3        *Assumptions and Technical Specifications***

The project is based on several key assumptions and technical protocols, which are:

1. 不同批次之间组织样本性质的一致性。
2. 生物组织切片仪和成像设备的可靠性和精确性。
3. 深度学习模型在解释复杂生物图像数据方面的充分性。

有关组织切片仪设置、图像分类标准和深度学习架构的技术规格详见方法和理论。

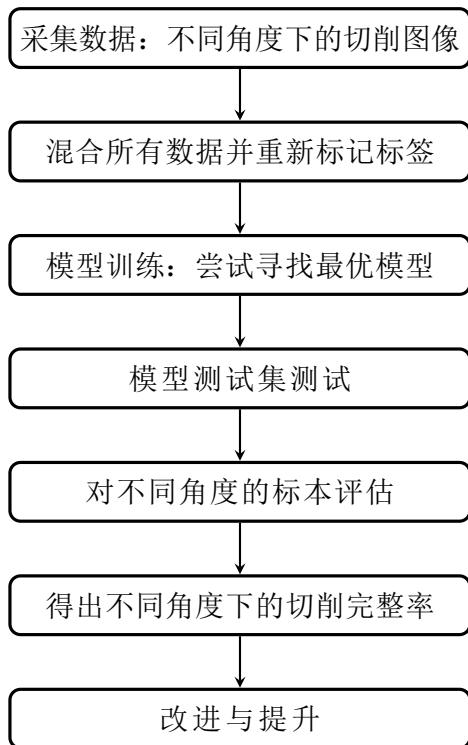
## **1.2              背景**

作为组成生物的最基本单位，人类对于细胞和组织的研究从未停止。作为直观观测细胞形态和结构的重要手段，生物组织切片对于生物医学的研究和临床诊断至关重要。一个完整可用的组织切片对于研究人员和医生来说是非常重要的，因为它可以提供有关细胞结构、组织形态和病理变化的重要信息。而在这之中，切片的质量是至关重要的。

传统的手工切片方法耗时且容易产生变异，因此自动切片机的出现为解决这些问题提供了可能性。对不同的生物组织而言，不同的切削参数会带来不同的效果，有好有坏。因此，为了提高生物切片的利用率和增加高质量标本的产量，针对特定组织确定最佳切割参数将是目标。

机器学习和深度学习在计算机视觉和图像处理领域取得了巨大的成功，机器学习被定义为一系列方法，这些方法可以自动检测数据中的模式，然后采用它们来预测未来的目标或做出决策。（KP Murphy，《机器学习：概率视角》，剑桥：麻省理工学院出版社，第32页，2012年。）在本文中，将通过整合先进的图像分析和机器学习技术来识别切片质量然后对评估在不同切片参数下的组织样本的质量。

### 1.3 项目流程



## 2 Literature review

这篇文献综述探讨了生物组织切片中技术的融合，特别关注图像分类和深度学习在优化切片参数方面的应用。它旨在突出重要的进展，确定当前方法学中存在的差距，并为拟议的项目奠定基础。

### 2.1 切片机与显微镜的选择

近年来，随着科技的发展，自动切片机的出现能够显著简化切片操作和提高切片质量。

M 在《Improved reproducibility in preparing precision-cut liver tissue slices》一文中提出，使用新型徕卡振动刀片切片机可以提高大鼠、小鼠和人体组织切片的准确性和重现性。——

在本次实验中，使用 epredia 提供的 HM355S 机器进行切片处理。该机器是一款热门的用于生物组织切片研究的设备，有不少实验和论文都使用了这款设备进行切片处理。

Elzbieta Klimuszko 使用过 HM355S 机器，以牙齿作为标本进行切片操作，探究牙釉质中的钙镁含量。—— <https://link.springer.com/article/10.1007/s10266-018-0353-6>

Andelko Hrzenjak 使用 HM355S 机器，对病变的子宫内膜组织进行切片操作，研究子宫内膜癌的发生机制。—— <https://www.sciencedirect.com/science/article/pii/S1525157810605685>

同样，对于显微镜的选择也是至关重要的。在本次实验中，使用了来自 Keyence 公司的 VHX7000 显微镜进行图像采集。他不仅能采集生物组织切片的图像（小鼠前列腺细胞 <https://www.frontiersin.org/journals/>），

还能采集无机物(如陶瓷,玻璃)的表面图像。<https://www.sciencedirect.com/science/article/abs/pii/S0109564123002>  
<https://www.taylorfrancis.com/chapters/edit/10.1201/9781003023555-130/looking-foundations-structural-glass-digital-microscope-veer>

实验中将使用 HM355s 切片机和 VHX7000 显微镜进行切片和图像采集。

## 2.2 关于切片组织的深度学习

深度学习技术在生物医学领域的应用已经取得了显著进展。深度学习模型在图像分类、目标检测和分割等任务中表现出色，为生物医学实验室的研究和诊断提供了强大的工具。

Lorena Guachi-Guachi 提出了一种使用 cnn 网络对组织切片进行识别并进行修整。

应用于切片术的卷积神经网络:识别石蜡包埋组织块的修剪末端切割程序 <https://www.sciencedirect.com/science/article/abs/pii/S0109564123002>

在 Biomedical Texture Analysis 一书中，Vincent Andrearczyk 提出了一种专门用于纹理分析的 cnn 架构，相比其它传统架构能够显著提高生物组织的分类准确性。

第 4 章-纹理分析中的深度学习及其在组织图像分类中的应用 <https://www.sciencedirect.com/science/article/abs/pii/S0109564123002>

Yan Xu 提出，从大型自然图像数据库 ImageNet 训练的 CNN 中提取的特征能够转移到组织病理学图像中，这为我们实现迁移学习提供了一种可行的思路。

通过深度卷积激活特征进行大规模组织病理学图像分类、分割和可视化 <https://link.springer.com/article/10.1186/s017-1685-x>

根据以上文献，深度学习技术在组织切片的图像分类和分析中具有广泛的应用前景。通过利用深度学习模型，可以实现对组织样本的高效识别和分类，为优化切片参数提供有力支持。

## 3 Methodology and theory

### 3.1 计算机视觉-图像分割

对于采集的到的图像数据，可以适当进行图像前处理。在保证图像的完整性和质量的前提下，可以对图像进行一定的处理，用以突出图像中我们希望让计算机识别的特征，并且在一定程度上去除图像的无关特征和噪声，以提高后续的深度学习模型的准确性。

图像分割是图像处理中的一个重要步骤，其目的是将图像分割成若干个具有独立意义的区域，以便进一步分析和处理。作为只关注生物组织完整度的模型，在这里需要将生物切片分割成生物组织和石蜡区域两个部分，然后对图像进行分割，突出显示生物组织部分的部分。

常见的图像分割算法有边缘检测，阈值分割等。

#### 3.1.1 边缘检测

对于生物切片组织，一个判断其质量的重要指标是切片的边缘是否清晰。切片边缘的完整性和连续性能够很好的反应该样品是否存在质量问题。

关于边缘检测,有许多的算法,如 sobel 算子,laplacian 算子、Canny 算子等。<https://ieeexplore.ieee.org/abstract/document/4767851/>

**sobel** 算子是一种一阶微分算子, 可用来检测图像边缘。(Learning OpenCV by Bradski and Kaehler.) 假设有一个一维图像  $f(x)$ , 其强度与像素坐标  $x$  之间的关系可以如图 1 表示。可以观察到在Figure 3.1中, 在  $x=2.2$  左右斜率最大, 可见在此处附近图像强度存在突变(存在边缘)。对其求导得到一阶导数  $f'(x)$ , 如Figure 3.2所示, 此时导数的绝对值最大。sobel 算子就是利用这一特性来检测边缘的。

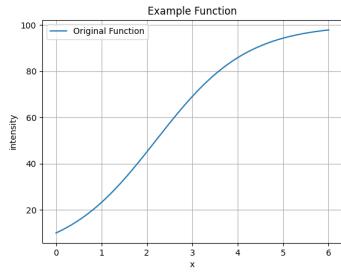


Figure 3.1  $f(x)$

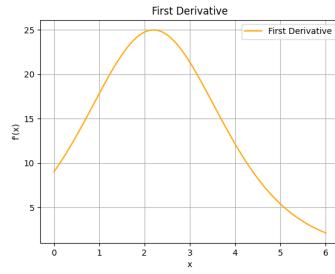


Figure 3.2  $f'(x)$

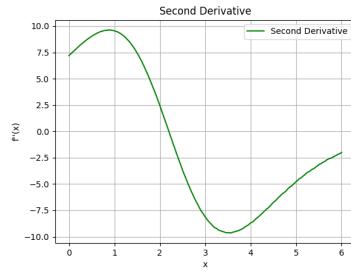


Figure 3.3  $f''(x)$

**Laplacian** 算子是一种二阶微分算子, 其对图像的边缘检测效果较好。它是对 sobel 算子再进行一次求导得出。在 2D 图像中, Laplacian 算子的定义如下:

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \quad (3.1)$$

如上图所示, 对一阶导数再次求导得到二阶导数  $f''(x)$ , 如Figure 3.3所示, 可以看到在  $x=2.2$  左右, 二阶导数为 0, 即说明当 laplacian 算子  $\nabla^2 f$  的值为 0 时, 说明图像强度存在突变, 即存在边缘。

**Canny** 算子是一种多阶微分算子, 他在 sobel 算子计算后的基础上加入了对噪声的抑制。他由 John F. Canny 于 1986 年提出。<https://ieeexplore.ieee.org/document/4767851/authors#authors>) 简而言之, 其在 sobel 算子计算后, 通过非极大值抑制, 滞后阈值等步骤, 设置了阈值, 排除图像中的假边缘, 得到了更加准确的边缘检测结果。

在 Experimental work/analytical investigation/ design 这一章节中将会对采集到的图像数据进行三种边缘检测算法的实验, 对比其效果。

### 3.1.2 阈值分割

除了边缘检测, 还有一种方法是阈值分割。阈值分割是将图像中的像素点分为两类, 一类是大于阈值的像素点, 另一类是小于阈值的像素点。这种方法适用于图像中的目标和背景的灰度差异较大的情况。

对于样品来说, 一个很简单的方法就是将石蜡区域和生物组织区域(样品在制备时已染色)的颜色进行对比, 然后通过阈值分割的方法将其分割开来。假定生物组织为黄色, 石蜡为白色, 那么可以通过设置一个阈值, 将图像中的白色部分分割出来, 那么剩下的就是生物组织部分。

此外, 关于阈值分割还有更多的方法, 比如下面就是一个基于 Otsu 方法的指纹提取算法。将其用在此处能够显著提高生物组织的分割效果。Yue Yaru 和 Zhu Jialin 在《Algorithm of fingerprint

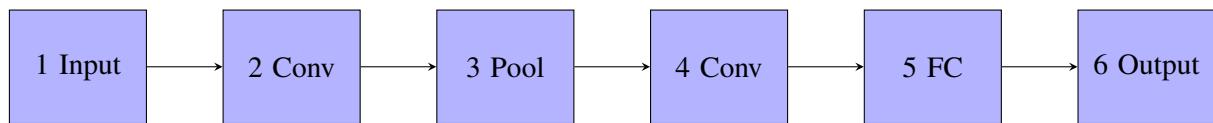
extraction and implementation based on OpenCV》一文中提出了一种基于 OpenCV 的指纹提取算法。该算法对 Otsu 方法进行了改进，特别是在光照不均匀、图像模糊的情况下能够实现准确、简单、运行时间短的指纹提取。<https://ieeexplore.ieee.org/document/7984539>

相关的对比和实验将在 Experimental work/analytical investigation/ design 这一章节中进行。

## 3.2 深度学习

### 3.2.1 卷积神经网络

卷积神经网络（CNN）是一种深度学习模型，尤其擅长处理图像数据。它通过一系列卷积层自动学习空间层次的特征，无需手动特征提取。一个典型的 CNN 模型包括卷积层、池化层、全连接层等。（<https://iopscience.iop.org/article/10.1088/1755-1315/440/4/042055/meta>）其架构如下所示。



where Conv is convolutional layer，是 cnn 的核心层，用于提取图像的特征。Pool 是池化层，用于减小特征图的尺寸，减少计算量。FC 是全连接层，用于将卷积层和池化层提取的特征进行分类或回归分析，最后输出结果。

对于一个典型的训练 cnn 的方法，包括前向传播、损失计算、反向传播和权重更新的过程。

1. 前向传播：输入数据通过网络的每一层，直到输出层。
2. 计算损失：使用损失函数（如交叉熵损失）计算网络输出和实际标签之间的差异。
3. 反向传播：计算损失函数关于网络权重的梯度。
4. 权重更新：使用梯度下降算法或其变种（如 Adam 或 RMSprop）来更新网络权重，目标是减少损失函数的值。

训练完成后，CNN 可以用来预测新的、未见过的图像的标签。CNN 的特点就是在于能够自动、有效地学习空间层次结构的特征。

### 3.2.2 迁移学习

显然，对于复杂的图像来说，搭建简单的 cnn 网络明显不够。这时候就需要迁移学习。迁移学习是一种机器学习方法，通过将一个模型训练的知识迁移到另一个模型上，从而加速训练过程。迁移学习的核心思想是利用源领域的知识来帮助目标领域的学习。

对于 cnn 模型来说，迁移学习的方法有很多，比如微调、特征提取等。微调是指在一个预训练的模型上，通过调整网络的参数，使其适应新的任务。特征提取是指在一个预训练的模型上，固定网络的参数，只训练全连接层，以适应新的任务。

常用的已经训练好的模型有 VGG、Inception 等。这些模型在 ImageNet 等数据集上已经训练好，即模型内部各层的权重已经确定，可以用来进行迁移学习。

表2展示了VGG16,VGG19(<https://arxiv.org/abs/1409.1556>),InceptionV3,(<https://arxiv.org/abs/1512.00567>)Xception (<https://arxiv.org/abs/1610.02357>), 等模型的参数数量。可以看到,这些模型的参数数量都很大,可以较为准确地提取复杂图像的特征。

**Table 3.1 Comparison of CNN Models**

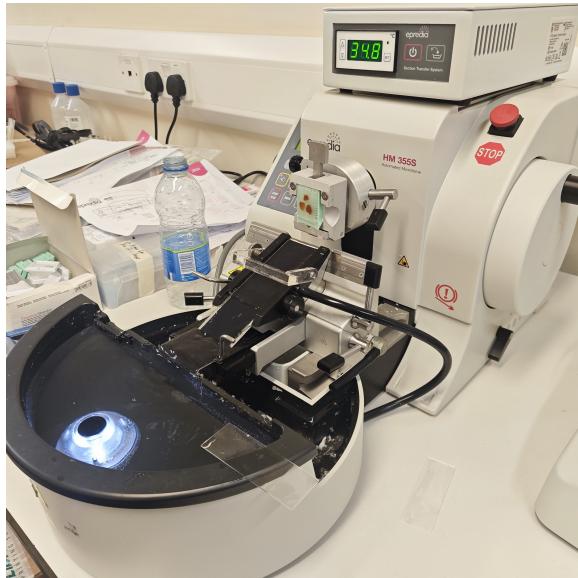
Model	VGG16	VGG19	InceptionV3	Xception
<b>Number of Parameters</b>	138,357,544	143,667,240	23,851,784	22,910,480

## 4 Experimental work/analytical investigation/ design

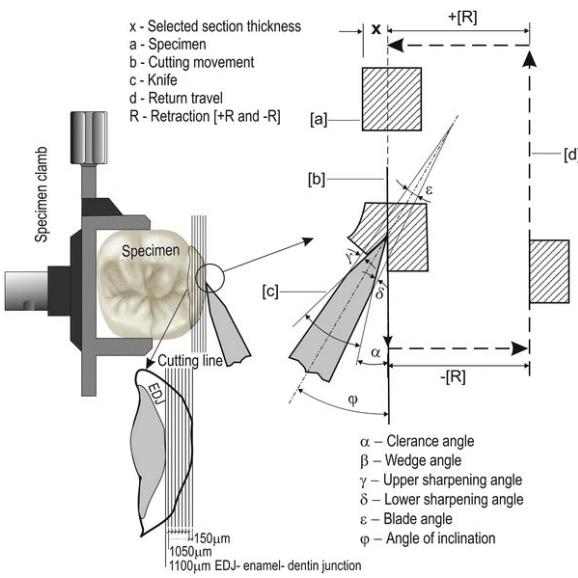
### 4.1 采集数据

要进行深度学习,所需要的第一步就是采集数据。在本实验中,我们使用了预先从生物实验室制备好的石蜡包埋好的组织切片(鱼的卵巢组织),将其放在HM355s自动切片机上依据切片机的使用手册,以不同的切削角度执行切片操作。记录切削数据。

其中切片机(Figure 4.1)的切片示意图(以牙齿为例)如Figure 4.2所示



**Figure 4.1** 切片机



**Figure 4.2** 切片机示意图

用于切片的生物组织(示例)如Figure 4.3所示

在切削过程中,从切角为8度开始(如Figure 4.1中的angle of inclination),每次增加0.5度,直到切角为12度。切片机在切片过程中保持给进速度为25,厚度为1。

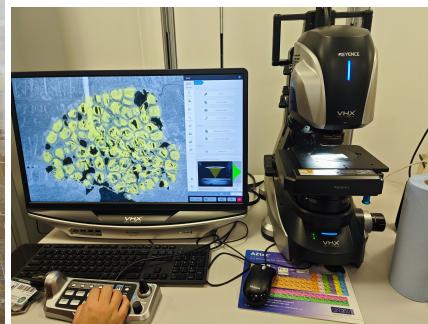
在切片完成之后,将切好的不同类型的组织切片放在载玻片上(如Figure 4.4)所示,待其晾干后转移至VHX7000显微镜下,通过显微镜对每份样品进行拍照,获取到每份样品的电子图像数据(如Figure 4.5)。



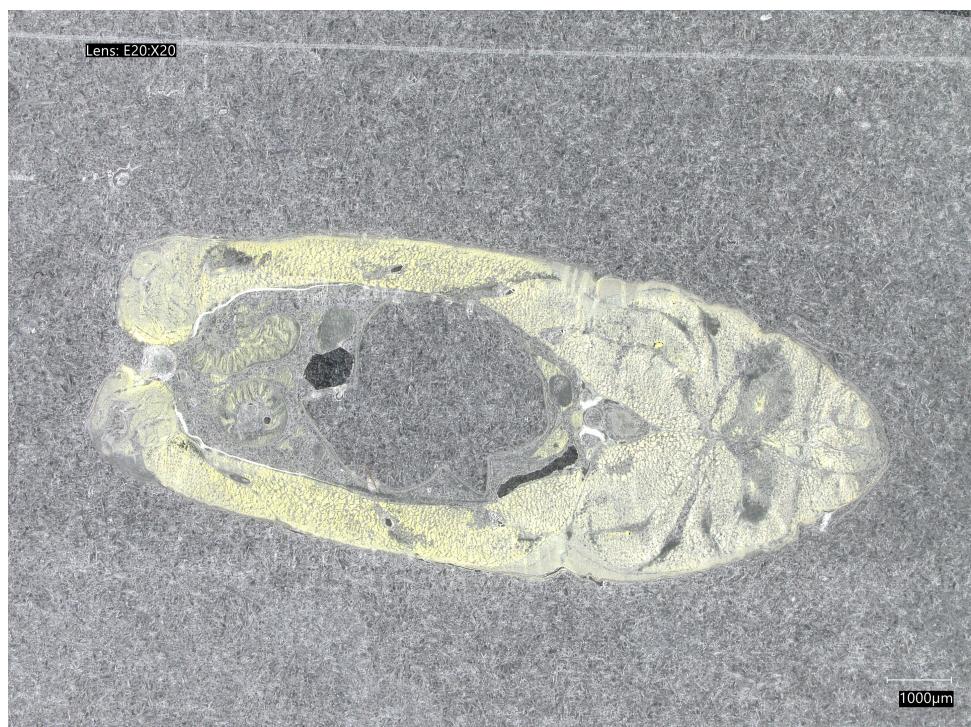
**Figure 4.3** 生物组织切片



**Figure 4.4** 采集样本



**Figure 4.5** 显微镜



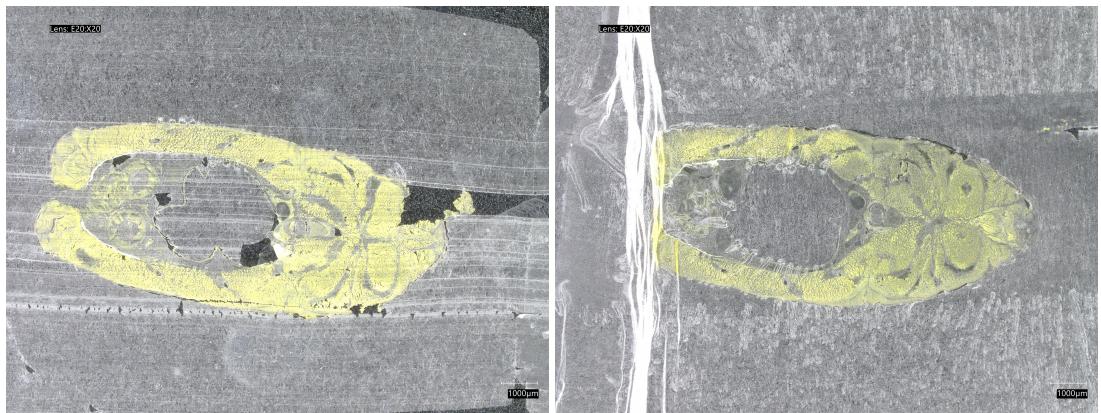
**Figure 4.6** 切角 9.5 度的样本

据此，一共得到 9 组数据，代表了从 8 到 12 每 0.5 度切角的数据。一共得到约为 200 张图片，每张图片的分辨率为 2880\*2160。其中一张（切角 9.5 度）如Figure 4.6所示。

对于这 9 组数据，需要找到的是在何种切角下，生物组织的完整度最高（质量最好）。因此现在需要将这 9 组数据根据生物组织的完整度进行重新标注。

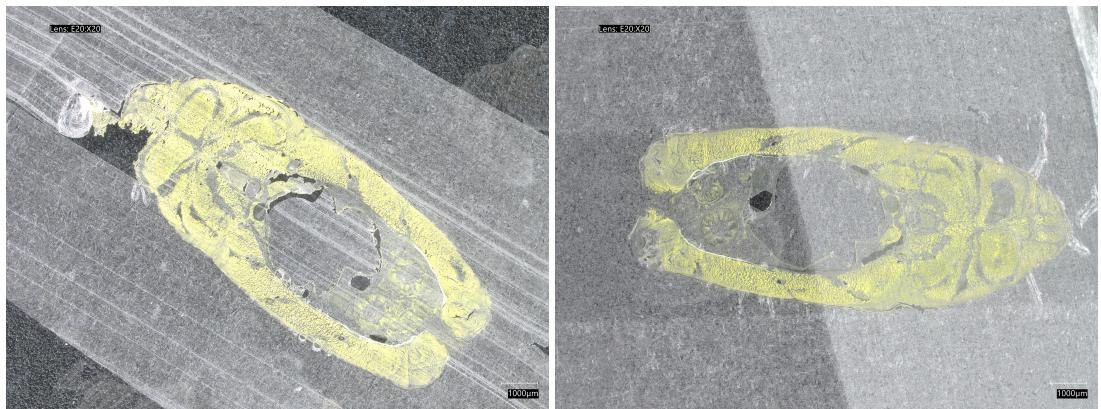
根据刀片切割对生物组织的影响，我们将生物组织的完整度分为五个类别，分别为：horizontal line, vertical line, slope, other, normal。其中 horizontal line 代表显著的以水平白线为瑕疵的样本，vertical line 代表显著的以竖直白线为瑕疵的样本，slope 代表图像采集中存在明显的角度（不利于观察），other 代表其他瑕疵（如采集中存在明暗显著差别的色块），normal 代表正常的生物组织样本。这五类标签的示例图片如下所示。

对于每一张图片，我们需要将其标注为以上五个类别中的一个。这将作为我们的数据集，用于



**Figure 4.7 horizontal line**

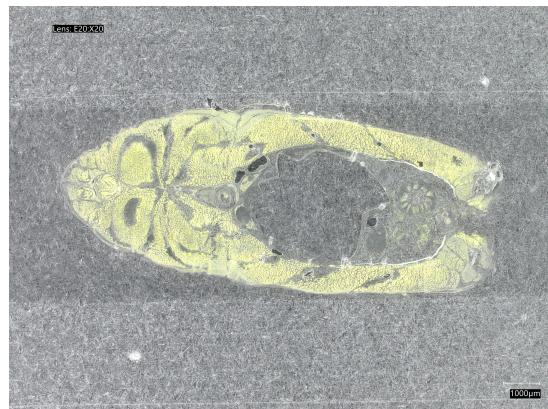
**Figure 4.8 vertical line**



**Figure 4.9 slope**

**Figure 4.10 other**

训练模型。



**Figure 4.11 normal**

## 4.2 模型 1：原始图像 + 简单的 *cnn* 网络

对于一个全新的数据集，在不确定图像复杂度对应的何种模型之前，首先尝试一个简单的 *cnn* 网络（架构如下），以了解数据集的特点和图像复杂度。

Table 4.1 显示的三个初始模型，分别为 a, b, c。这三个模型的区别在于卷积层的数量和大小，全连接层的大小。a 和 b 相比修改了卷积层的神经元数量，c 相比 a 修改了全连接层的神经元数量。

数据的预处理部分，先将数据集分为训练集和测试集，其中训练集占 80%，测试集占 20%。之后将图像的长宽缩小一倍（即从 2880\*2160->1440\*1080）并归一化数据。在训练过程中，我们使用了 Adam 优化器，交叉熵损失函数，使用早停。

下面图组展示了模型 1a, 1b, 1c 的准确度和损失随着训练次数的变化。

在图中，观察到模型 1a、1b 和 1c 在不同训练周期（Epochs）的准确度与损失的变化情况。模型 1a、1b 和 1c 的训练准确度随着时间逐步提高，趋向于稳定，而训练损失则呈下降趋势，接近于零，这表明模型在学习训练数据方面表现得相对良好。然而，对于验证集，三个模型的准确度似乎在约 80% 到 85% 的区间内稳定，而验证损失在某些情况下较高，特别是模型 1a 的验证损失在后期趋近于 2.5，表现出较大波动。这表明存在一定程度的过拟合，即模型在未见过的数据上的表现不如在训练集上。特别值得注意的是，模型 1c 相对于其他模型而言，在验证损失方面表现最佳，这可能意味着其结构或参数调整对于泛化能力的提升更为有效。在这里过拟合的原因推测可能是模型的复杂度不够，数据集的复杂度过高，模型无法很好的提取特征。这些结果指出虽然模型在训练集上能够实现高准确度和低损失，但在验证集上的泛化能力还有待提高。

Table 4.1 Configuration of the simple CNN model

Layer Type	Configuration 1a	Configuration 1b	Configuration 1c
Input Layer	-	-	-
Conv Layer 1	Conv3-32 (relu)	Conv3-16 (relu)	Conv3-32 (relu)
Pooling Layer 1	MaxPooling	MaxPooling	MaxPooling
Conv Layer 2	Conv3-32 (relu)	Conv3-32 (relu)	Conv3-32 (relu)
Pooling Layer 2	MaxPooling	MaxPooling	MaxPooling
Conv Layer 3	Conv3-32 (relu)	Conv3-64 (relu)	Conv3-32 (relu)
Pooling Layer 3	MaxPooling	MaxPooling	MaxPooling
Flattening Layer	Flatten()	Flatten()	Flatten()
FC(Full connect)	Dense(128, relu)	Dense(128, relu)	Dense(256, relu)
Output Layer	-	-	-

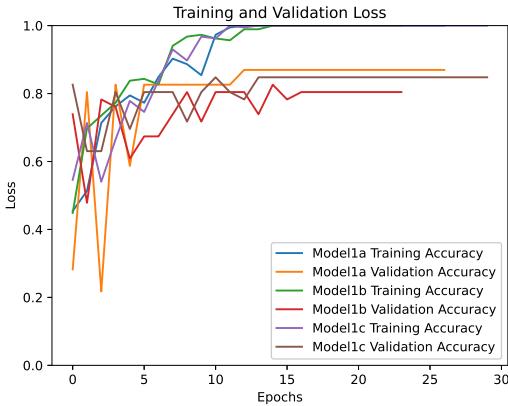


Figure 4.12 Model 1 accuracy

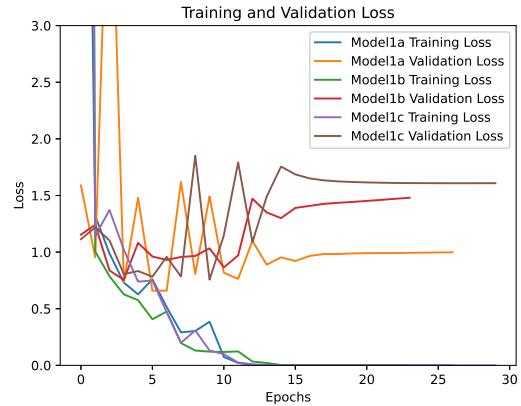


Figure 4.13 Model 1 loss

### 4.3 改进：图片预处理

在模型表现能力欠佳的情况下，我们考虑是否是图像过于复杂导致模型难以提取出显著特征。因此我们考虑对图像进行预处理，以突出图像中我们希望让计算机识别的特征，并且在一定程度上去除图像的无关特征和噪声，以提高后续的深度学习模型的准确性。

在这里采用边缘检测，阈值分割两种方法对图像进行预处理。

#### 4.3.1 边缘检测

正如在 3.1.1 中所提到的，边缘检测的原理是通过检测像素点的灰度值的变化（梯度）来确定图像中的边缘。假定原始图像是Figure 4.6.

在进行边缘检测之前，还需要进行一步前处理-高斯模糊。这么做的原因是，高斯模糊可以减少图像中的噪声，平滑图像的梯度，减小识别假边缘的几率，使得边缘检测更加准确。（<https://ieeexplore.ieee.org/abstract/>）在高斯模糊核的选择上，选择高斯核分别为 21, 41, 61, 81（图像宽度的 1%, 2%, 3%, 4%）。高斯模糊后的图像如下所示。为了方便更直观的展示高斯模糊核对边缘检测的影响，这里采用 sobel 算子计算经过高斯模糊后的边缘并手动提亮 50。

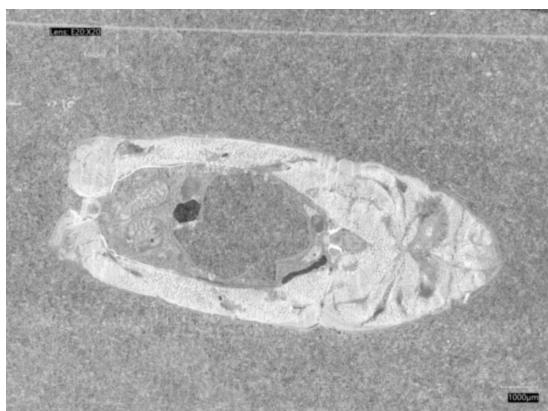


Figure 4.14 blurred k=21



Figure 4.15 sobel k=21

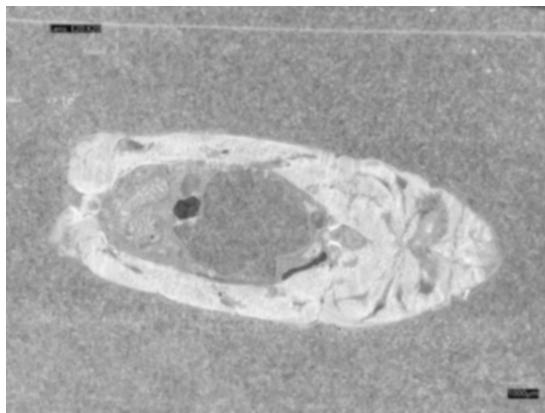


Figure 4.16 blurred k=41



Figure 4.17 sobel k=41

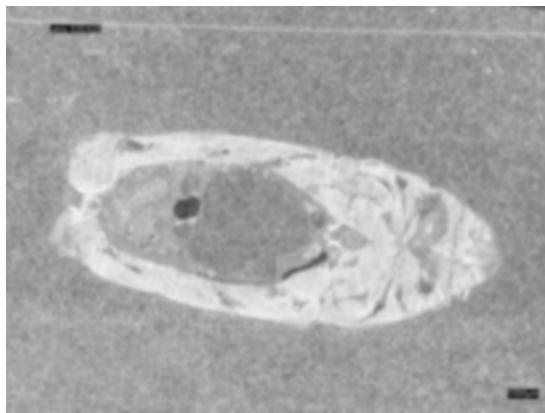


Figure 4.18 blurred k=61



Figure 4.19 sobel k=61

从Figure 4.14到Figure 4.20可以看到，随着高斯模糊核的增大，图像的细节逐渐模糊，边缘也逐渐变得模糊。从Figure 4.15到Figure 4.21可以看到，随着高斯模糊核的增大，边缘检测的效果逐渐减弱，边缘变得不明显。考虑到图像边缘的清晰度和底噪的对比，我们选择高斯模糊核为 61。以下是在高斯模糊（k=61）后使用 python 的 opencv 库执行 laplacian 算子的结果。

canny 算法相对于 sobel 算法稍显复杂-引入了阈值检测，非极大值抑制等步骤。canny 算法引入了两个阈值，分别为低阈值和高阈值。其中，当图像的梯度值大于高阈值时，被认为是边缘；当

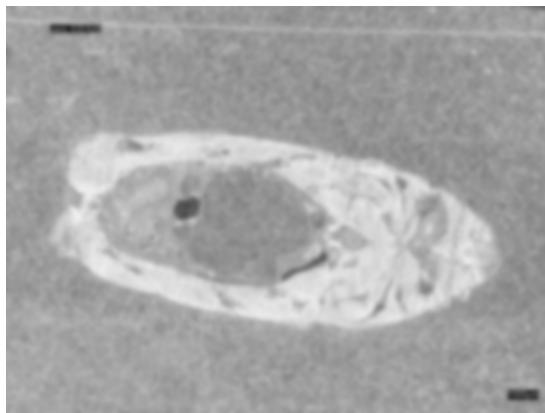


Figure 4.20 blurred k=81

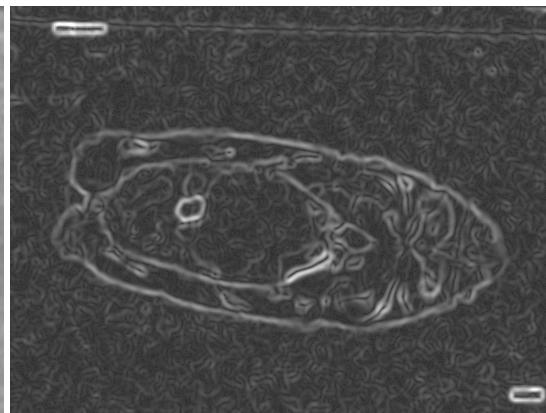
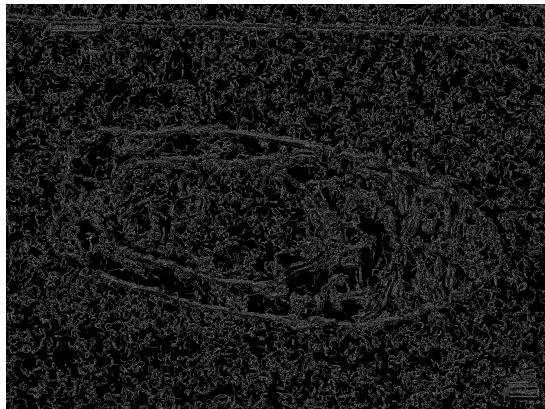


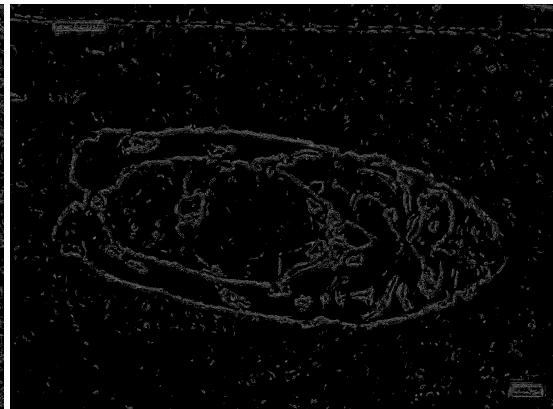
Figure 4.21 sobel k=81



**Figure 4.22 laplacian**



**Figure 4.23 canny 2 5**



**Figure 4.24 canny 4 10**

图像的梯度值小于低阈值时，被认为不是边缘；当图像的梯度值在两者之间时，如果与高阈值的边缘相连，则被认为是边缘，否则被认为不是边缘。这样的处理可以有效的去除图像中的噪声，得到更加准确的边缘检测结果。

通常情况下，高阈值和低阈值的比值在 2:1 到 3:1 之间。在这里我们选择阈值比为 2.5 : 1，探究不同阈值对边缘检测的影响。

取低阈值为 2 4 6，此时对应的高阈值为 5 10 15。canny 算法的结果如下所示。



**Figure 4.25 canny 6 15**

在三张 canny 算法的结果中，可见Figure 4.24的效果最好，其能在保证边缘细节得到大部分保留的情况下，去除了大部分的噪声。因此我们选择 canny 算法的阈值为 4 10。

### 总结

对比 sobel, laplacian 和 canny 算法的结果，sobel 算法的效果一般，对于底噪不是能很好的去除，边缘检测效果还算显著。laplacian 算法最差，边缘甚至已经不可见，这可能是因为该算法对噪声最敏感。canny 算法的效果最好，能够在保证边缘细节的情况下，去除大部分的噪声。因此我们选择 canny 算法作为图像预处理的方法。

### 4.3.2 阈值分割

考虑到生物组织样本的主体是黄色，石蜡是白色，我们可以通过设置一个阈值，将图像中的白色部分分割出来，那么剩下的就是生物组织部分。在这里使用 python 的 opencv 库中进行操作。首先将图像进行对比度增强，增加饱和度，更好的凸显出生物组织的颜色 (Figure 4.26)。之后读取图像的每个像素点，将黄色周围半径 15 左右的像素点进行保留（约为图像宽的百分之一），其他的色块进行去除。(如Figure 4.27)。

但是观察发现，这种方法对于生物组织和石蜡的分割效果并不好，因为生物组织在切片过程中会掉落碎片组织，出现在标本各处，进而影响黄色像素点的识别。此时还需要进一步的处理，去除黑色色块。此时只需要将黑色色块进行掩码反转，使其变为白色即可。结果如图Figure 4.28所示。

### 4.3.3 另一种阈值分割方法-指纹算法

在进行文献综述的时候，发现有一篇论文是基于 otsu 算法改进的分割方法用于进行指纹分割。考虑到切片样本和指纹都属于生物组织，因此我们尝试使用论文中提到的算法进行分割。结果如图Figure 4.29所示。



Figure 4.26 enhanced image

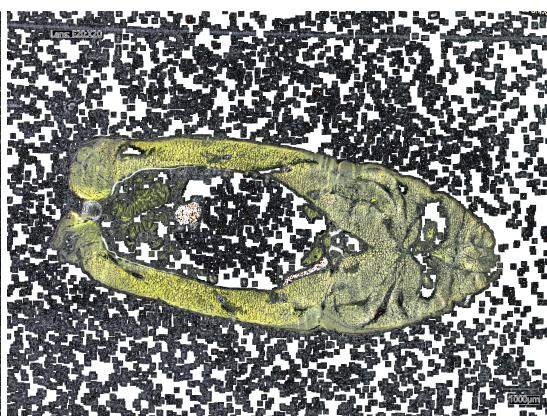


Figure 4.27 yellow picture

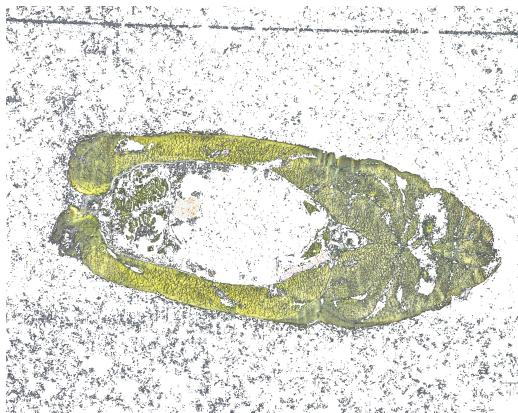


Figure 4.28 final

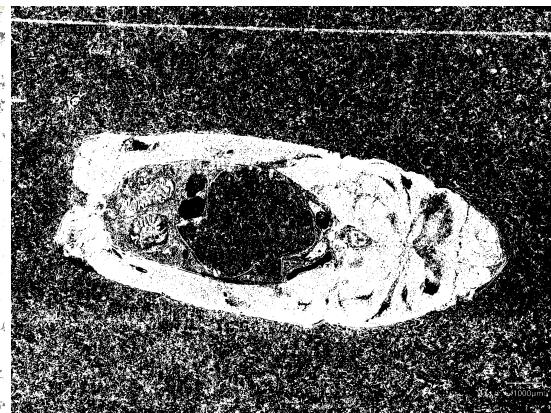


Figure 4.29 fingerprint

#### 4.3.4 小结

根据上文提到的图像预处理方法，我们可以看到，边缘检测和阈值分割的效果都不错，都能够很好的突出生物组织的特征，去除石蜡的干扰。对此我们可以设置三组数据集，分别是经过边缘检测后的图像，经过阈值分割后的图像和经过指纹算法分割后的图像。这三组数据集将作为我们的训练集，用于下一节的模型训练。

### 4.4 模型 2：预处理图像 + 简单的 *cnn* 网络

在这里基础模型选用在上一节模型 1 中表现最好的模型 1c。在这里我们将模型 1c 的输入改为经过预处理后的图像，即经过边缘检测后，阈值分割和指纹算法分割后的图像。在这里模型的架构不变，只是输入的数据发生了变化。

所有的模型 2 采用和模型 1c 同样的架构构成，分别由三个包含 32 个特征图，卷积核为  $3 \times 3$  的卷积层和最大池化层，一个包含 256 个神经元的全连接层组成。模型 2a 的输入为经过 **canny** 边缘检测后的图像。模型 2b 采输入为经过阈值分割后的图像。模型 2c 输入为经过指纹算法分割后的图像。

结果如下：

#### 4.4.1 小结

在图中，对比了模型 2a、2b 和 2c 的训练和验证准确度以及损失的变化情况。模型 2a 和模型 2c 的训练和验证准确度在经过约 8 个训练周期后开始趋于稳定，其中训练准确度接近于 100%，而验证准确度稳定在 65% 和 75% 左右。尽管准确度较高，两者的验证损失仍旧较高，都在 1 以上。这可能指示了模型对训练数据过拟合，而对未见数据的泛化能力有限。

对于模型 2b，其在约 10 个训练周期后开始收敛，与模型 2a 和 2c 相比，模型 2b 拥有最高的的验证准确度，约为 82% 左右，但是其损失显著低于模型 2a 和 2c，在 1-1.2 波动。这表明模型 2b 在泛化到验证集上时的性能更优，损失更低，反映了模型更好的适应性和鲁棒性。

可能原因是模型 2a 和 2c 可能处理的是灰度图像，而模型 2b 处理的是彩色图像。彩色图像包含的 RGB 通道信息可以提供更丰富的特征，从而可能增强了模型的特征提取和泛化能力。然

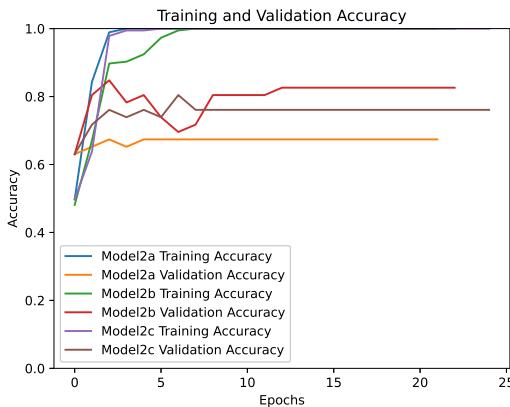


Figure 4.30 Model 2 accuracy

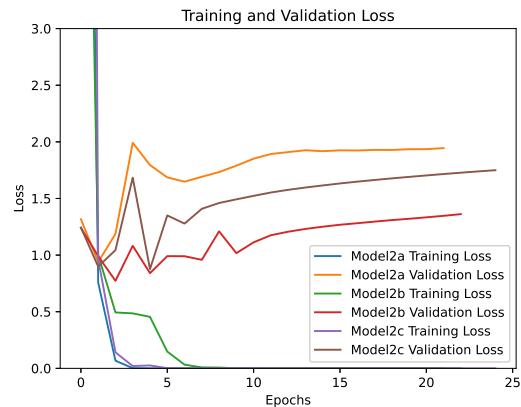


Figure 4.31 Model 2 loss



Figure 4.32 origin

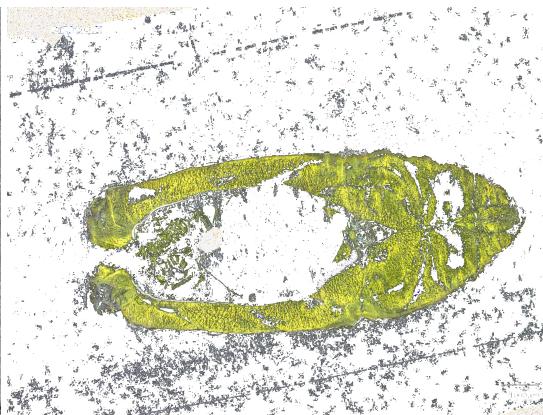


Figure 4.33 yellow

而，即使彩色图像提供了额外信息，前处理步骤，尤其是模型 2b 的阈值分割，可能会导致重要细节的丢失，这反过来可能会影响到模型在特定图像上的表现。这种情况下，模型的预处理步骤需要仔细检查，以确保不会因过于激进的图像简化而丢失关键信息。

一个丢失关键信息的例子如下所示：

Figure 4.33是模型 2b 训练集(经过黄色阈值分割后的图像)中的一张图片，对比原图(Figure 4.32)可以观察发现，原本切片中能够被接受的平行白线被阈值分割算法显著增强了，这有极有可能会影响模型的训练效果，即模型会在一定程度上与 horizontal line 混淆。

由此可以看出，对于图像预处理，其实并不能显著的提高模型的训练效果，反而可能会因为过于激进的预处理而丢失关键信息，导致模型的训练效果下降。在后面将会尝试使用迁移学习的方法，使用预训练好的大规模深度学习模型，将其迁移到我们的数据集上，以提高模型的训练效果。

## 4.5 模型 3：原始图像 + 迁移学习

现在我们已经尝试过了简单的 cnn 网络，以及对图像进行预处理后的 cnn 网络。既然训练结果不是很理想，那我们为什么不去尝试更大更深的模型？在这一节，我们尝试使用迁移学习的方

法，使用预训练好的大规模深度学习模型，将其迁移到我们的数据集上，以提高模型的训练效果。

正如在第三节 methodology 里提到的，在这里将使用 VGG16，VGG19 和 InceptionV3 三个模型进行迁移学习。这三个模型都是在 ImageNet 数据集上训练好的模型，具有已经训练好的权重。

在这里为了避免迁移学习过拟合，不仅使用了原有的早停法，还限制了模型的学习率为 1e-5（对于 inceptionV3 模型，学习率为 1e-4）。

model3a 是使用 VGG16 模型进行迁移学习的模型。model3b 是使用 VGG19 模型进行迁移学习的模型，VGG19 和 VGG16 相比只是在中间增加了 3 个额外的卷积层，其他则与 VGG16 相同。model3c 是使用 InceptionV3 模型进行迁移学习的模型，其中 InceptionV3 是一个相对于 VGG16 和 VGG19 更加复杂的模型，其在训练过程中引入了 Inception 模块，能够更好的提取图像的特征。

在这里统一将模型的输入调整为 224\*224，因为 VGG16 和 VGG19 模型在预训练时的输入层是 224\*224 的图像，而 InceptionV3 的默认值为 299\*299 的图像。

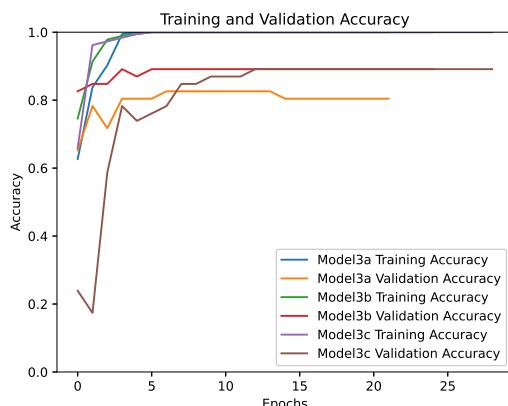


Figure 4.34 Model 3 accuracy

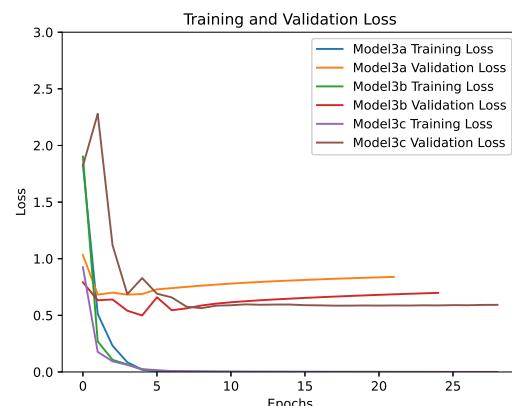


Figure 4.35 Model 3 loss

对比三个模型，可以观察得到 model3b 和 model3c 的验证集的准确度显著高于 model3a，约为 90% 左右。观察损失图，可以得出 model3c 在这三个模型中验证集的损失是最低的，model3b 次之，model3a 表现最差。这可能是因为 InceptionV3 模型的复杂度更高，能够更好的提取图像的特征。而 VGG16 和 VGG19 模型相对于 InceptionV3 模型而言，更加简单，可能在提取图像特征上存在一定的局限性。此外，model3b 性能显著高于 model3a 可以说明，VGG19 模型相对于 VGG16 模型而言，多出的三个卷积层能够更好的提取图像的特征。符合模型的复杂度越高，其训练效果越好的规律。

#### 4.5.1 小结

对比 VGG16，VGG19 和 InceptionV3 三个模型，可以发现 InceptionV3 的训练效果最好，其训练准确度和验证准确度收敛于 1 和 0.9 左右，损失收敛于 0.6 左右。这说明 InceptionV3 模型的训练效果最好，其泛化能力最强。

**Table 5.1 model accuracy on test set**

	normal	horizontal_line	vertical_line	slope	other
accuracy(%)	98.4	95.6	80.0	96.1	95.2

**Table 5.2 normal accuracy on different angle**

	8	8.5	9	9.5	10	10.5	11	11.5	12
accuracy(%)	80	81.5	83.5	93.3	96.6	88.8	84.2	66.6	62.2

## 4.6 模型选择总结

横向对比模型系列，模型 1，模型 2 和模型 3，可以发现模型 3 的训练效果最好。特别是模型 3c。究其原因，可能是因为模型系列 3 是基于大规模图像识别的超深卷积网络，其在训练过程中能够更好的提取图像的特征，构建自己的特征空间。值得注意的是，模型 3c，属于 InceptionV3，在架构上具有模块化的设计，包括了多个“inception 模块”。其包含了多尺度的卷积层，同时在同一层内并行运行。在特征提取上，Inception 模块可以在同一层内捕捉不同尺度的特征，使得网络能够自适应地选择更合适的特征表示。在处理深度上，InceptionV3 利用批量标准化和残差连接来帮助训练深层网络，可以显著解决梯度消失的问题。

因此，我们选择模型 3c 作为我们的最终模型，用于后续的进一步应用和测试。

# 5 Presentation of experimental or analytical results/descriptions of final constructed product

在这一节我们主要讨论模型的测试结果和模型进一步改进的空间。

## 5.1 带入测试集验证准确度

在这里我们将训练好的模型应用到测试集上，计算模型的准确度。

Table 5.1是模型在测试集上的准确度：（准确度定义为标签与模型预测一致的样本数占总样本数的比例）

可见，模型预测结果在 normal 上表现较好，而在 vertical\_line 上表现较差。这可能是由于用于测试的样本量较少，导致模型学习不足，难以判断。

## 5.2 判断最佳切削角度

在这里采用模型 3d (inceptionv3) 来判断最佳切削角度。我们将模型应用到测试集上，计算在每个切削角度下可用样本的占比。找到可用样本占比最多的区间，即为最佳切削角度。

由Table 5.2可知，最佳切削角度为 10 度。

若要在保证切削质量为百分之 80 的情况下，切削角度应在 9 度到 10.5 度之间。

**Table 5.3 model accuracy on test set**

	normal	horizontal_line	vertical_line	slope	other
accuracy(%)	98.4	96.7	85.6	96.5	96.5

### 5.3 模型的进一步提高（改变输入分辨率）（数据增强在这里）

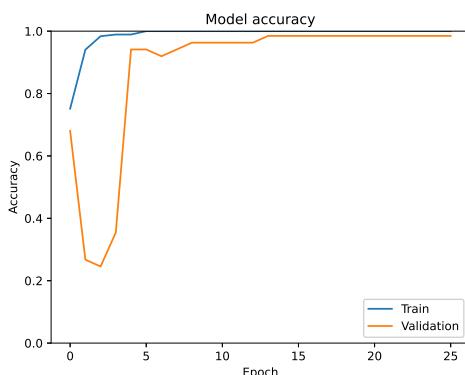
在这里我们讨论模型的进一步提高的空间。

将高分辨率图片缩放为 InceptionV3 模型默认的 299x299 大小的确可能导致信息和细节的丢失，特别是对于原始分辨率远高于此标准的图像。例如，从 VHX7000 设备采集的 2880x2160 分辨率图像就含有大量的细节，直接缩放可能不利于模型捕捉到所有的细微差别，尤其是在医学影像或其他细节丰富的领域。

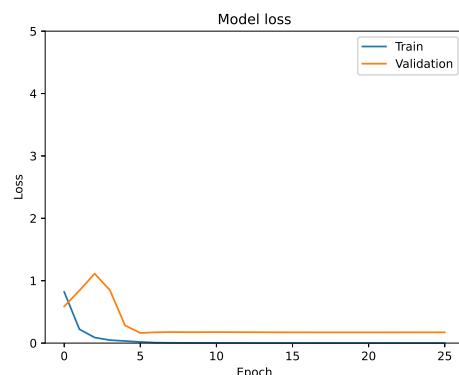
改变模型的输入层接受更大的图片尺寸是一个潜在的解决方案。这样做的优势是它允许模型处理更高分辨率的图像，保留更多的原始信息和细节，可能导致更好的性能和更高的准确度。InceptionV3 的架构设计有助于处理更大图片，因为其含有多个大小不一的卷积核，这使得它能够捕捉不同尺度的特征。

受制于实验室机器性能(显存为 16G)，在这里将图像缩放到分别为原图像的 0.4 倍，即 1152\*864，进行再一次训练。

新的模型为 model4，训练效果如下所示：



**Figure 5.1 Model-4 accuracy**

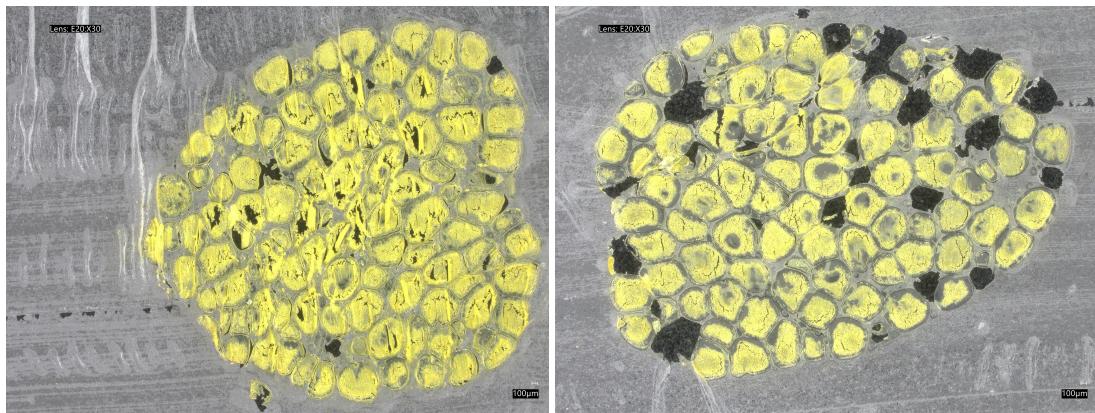


**Figure 5.2 Model-4 loss**

观察训练准确度和损失随步长的变化可以发现，模型的性能有惊人的明显提升。训练和验证准确度都接近 1，同时验证损失降至 15% 左右，这通常表明模型具有很强的泛化能力。这种情况下，模型不仅在训练数据上表现出色，而且能够很好地泛化到新的、未见过的数据上。

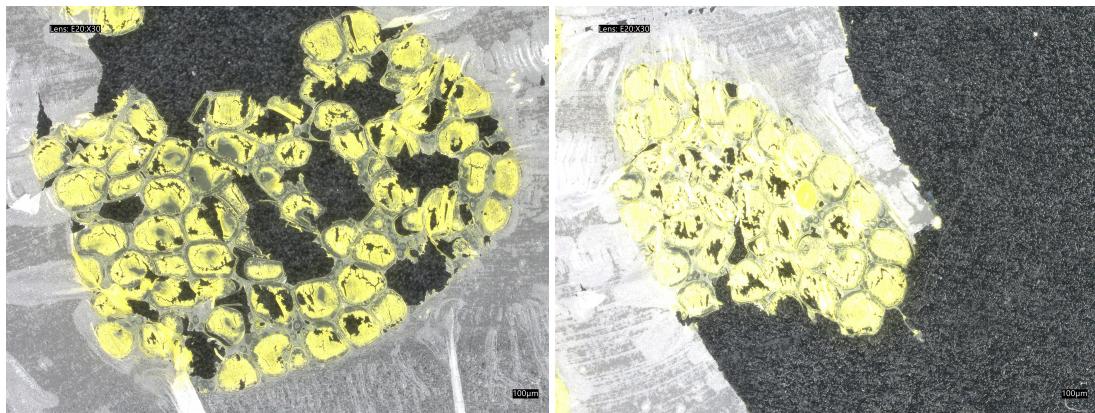
将其再一次带入测试集进行准确度评估，结果如Table 5.3：

对比修改分辨率前后的模型准确度，可以发现模型的准确度虽有提升但并不显著的，可能是由于准确度已经很接近于 1，提升的空间较小导致的。



**Figure 5.3 good fish lung**

**Figure 5.4 normal fish lung**



**Figure 5.5 bad fish lung**

**Figure 5.6 other fish lung**

## 5.4 模型通用性

在上文的实验中我们采用的是鱼的卵巢为组织切片，而在实际应用中，我们可能会遇到其他组织切片，如其他器官或其他动物标本等。因此，我们需要考虑模型的通用性。

在这里有另一组已经采集好的数据集，是鱼的肺切片，一共分成 4 类，分别是 good, normal, bad,other 四类（样本如Figure 5.3到Figure 5.6）。现在更改输入数据集，保持原有模型 4 架构不变，采用 1152\*864（0.4 倍长宽）分辨率图片作为输入数据进行训练。

训练的准确度和损失如Figure 5.7和Figure 5.8所示。

通过观察图片可以得出：模型 5 的训练和验证准确度迅速上升并保持在高位，表明模型在这两个数据集上均有良好表现，损失图显示训练损失快速降低并趋于零，而验证损失在初始阶段出现尖峰后迅速降低并稳定，整体来看，这些迹象表明模型具有较好的拟合能力和泛化性能。

（补充额外的表格-关于遇到肺泡验证准确度的表格）

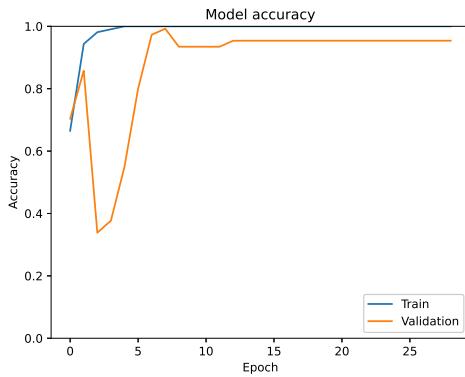


Figure 5.7 Model-5 accuracy

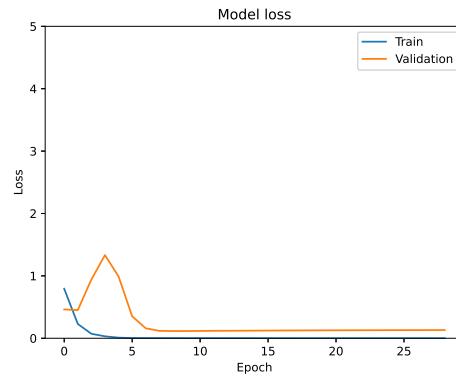


Figure 5.8 Model-5 loss

## 6 Discussion and conclusions

### 6.1 讨论

正如上文中所详述，这里将详细讨论研究过程中的发现和结果。

在第 4 节中，我们试图建立模型，尝试了从简单的 cnn 模型开始，之后进行图像的预处理，发现效果都不理想的情况下尝试了迁移学习的方法，得出采用 InceptionV3 模型的迁移学习效果最好。

在这其中值得注意的是伴随着不同模型的尝试，当模型参数调整或模型架构变得更加复杂时（例如 InceptionV3），模型性能显着提高，即验证集的准确度越来越高，损失越来越低。

此外，对比模型系列 1 和 2，我们发现在在图像分类的任务中使用预处理图像的手段去辅助机器提取特征的行为不是非常有效。采用图像处理有可能会导致重要细节及信息的丢失，进而影响机器对特征的提取，从而影响模型的准确度和性能。

在第 5 节中，我们对模型进行了应用测试，首先选取了额外的测试集测试模型的准确度，发现模型在所有测试集上的准确度均大于 85%。之后带入使用模型对不同切削角进行评估，发现若要在保证切削质量为百分之 80 的情况下，切削角度应在 9 度到 10.5 度之间。最后，我们使用了另外的鱼的肺泡切片的图像数据集对模型进行二次验证，发现模型能够很好地应用于其他数据集。（贴数据）

### 6.2 Future work

#### 分类方法的改进

本研究提供了多个改进和深化的路径。尤其是，所采用的分类方法已经展现出理想的结果；然而，不足的是，五个分类仍旧是不充足的。通过提升分类的多样性，可以进一步得到切削角度和样本完整度的关系，进而提供更精确的分析。

此外，在拥有足够多的分类点的情况下，可以考虑从二元分类框架向线性分析方法过渡。在分类种类足够多的情况下，可以近似认为这是线性关系下的离散点-即可以通过线性拟合来得到最佳切削角度和样本完整度的线性关系。

在模型显示出显著能力和鲁棒性的情况下，采用线性分析——特别是在确定最优切割角度方面——可能呈现出一种更精细的方法，用于关联组织质量与切割参数。这样的方法有可能简化切割参数的预测准确性，并有助于更精确地控制组织切割过程。

当然，将分类问题转变为线性回归问题的难度是非常大的。其中一个关键的问题是，线性回归模型的训练和验证需要巨量的数据（甚至高出几个数量级），而这些数据采集和收集将会是一个非常漫长且困难的过程。不仅如此，线性回归模型的训练和验证也需要更多的计算资源，目前在采用 tensorflow 框架下选择 InceptionV3 模型，输入分辨率为 1152\*864 的情况下，显卡的显存已经达到了极限，因此在这一方面的改进需要更多的硬件和更加强大的计算资源。所以，这一方面的改进是一个长期的目标，需要更多的资源和时间。（如果可以增加相关论文证明观点：二分类和线性回归问题的对比和性能开销）

### **6.3 性能提升与优化**

随着本研究目标向大规模应用前进，性能提升成为了必须面临的重要挑战。性能优化不仅仅局限于算法效率的提升，还包括模型框架的扩展性、稳定性和部署能力的增强以及底层的语言和代码优化。

显然，针对计算资源的利用和调度策略，我们应当寻求更高效的计算框架和并行处理算法。例如，利用分布式计算资源，可以显著缩短模型训练的时间，并提高处理大型数据集时的效率。同时，考虑到能耗和计算成本的约束，优化模型的计算结构和参数设置，以期达到在有限资源下最大化计算输出的目的。（论文：tf 和 torch 的对比，libtf libtf，opencv c++ 和 py 的对比）

### **6.4 总结**

## **7 Project management, consideration of sustainability and health and safety**

### **7.1 Subsection 5.1**

### **7.2 Subsection 5.2**