Title

Abstract

Table of Contents

1 Introduction and background

1.1 Introduction

1.1.1 Project Overview

This project addresses the critical task of optimizing the cutting parameters of a biological tissue slicer, an essential instrument in biomedical research and clinical diagnostics. The aim is to enhance the precision and efficiency of tissue sample preparation by identifying the optimal slicing conditions. Through the collection of tissue samples under various cutting parameters and subsequent artificial image classification, this study employs deep learning techniques to analyze and predict the most effective slicing parameters. This endeavor not only promises to improve the quality of tissue samples for microscopic examination but also to streamline the workflow in laboratories, thereby contributing to the advancement of biological and medical sciences.

1.1.2 Objectives

1. Collect a comprehensive dataset of tissue samples sliced under different parameters.
2. Employ artificial image classification to categorize the quality and characteristics of these samples.
3. Develop and train a deep learning model capable of assessing tissue sample quality.
4. Use the model's insights to determine the optimal cutting parameters for the tissue slicer.
5. Validate the model's predictions through empirical testing and refinement.

1.1.3 Structure of the Report

This project is organized into the following chapters, each designed to systematically explore the research background, methodologies, experimental work, results presentation, discussions and conclusions, as well as considerations for project management, sustainability, and health and safety:

Introduction and Background - This chapter outlines the project's objectives, goals, and structural arrangement. It provides a brief introduction to the motivation and necessity for the research, along with the technical protocols and specifications adopted.

Literature Review - An in-depth discussion on the use of biological tissue slicers, image classification, and deep learning in the preparation of biological samples. This section positions the current study within the context of existing research.

Methodology and Theory - Detailed descriptions of the experimental methods, theoretical frameworks, and the specific plans for data collection and processing are presented here.

Experimental Work/Analytical Investigation/Design - Describes the detailed steps of experimental design, implementation, and analytical investigation. It elaborates on the strategies and methods adopted to achieve the project's objectives.

Presentation of Experimental or Analytical Results/Descriptions of Final Constructed Product - This chapter showcases the experimental data, analysis results, or the final design product, providing detailed accounts of the experimental or design outcomes.

Discussion and Conclusions - The results are analyzed, and their scientific significance and practical value are discussed. This chapter also offers the research conclusions and suggests potential directions for future studies.

Project Management, Consideration of Sustainability and Health and Safety - Discusses strategies for project management, sustainability issues, and health and safety measures to ensure the research work is conducted efficiently and safely.

References - Lists all the bibliographic materials cited, supporting the research and providing the basis for the study.

1.1.4 Assumptions and Technical Specifications

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

1. The consistency in tissue sample properties across different batches.
2. The reliability and precision of the biological tissue slicer and imaging equipment.
3. The adequacy of the deep learning model in interpreting complex biological image data.

Technical specifications regarding the tissue slicer settings, image classification criteria, and deep learning architecture are detailed in **Methodology and Theory**.

1.2 Background

1.2.1 Importance of Tissue Sample Quality

High-quality tissue samples are pivotal for accurate diagnosis and research. The quality of a tissue sample can significantly affect the results of histological analysis, making the optimization of slicing parameters a crucial endeavor.

1.2.2 Advancements in Image Classification and Deep Learning

Recent advancements in image classification and deep learning have opened new avenues for automating and enhancing the analysis of biological samples. By leveraging these technologies, it is possible to achieve greater accuracy and efficiency in identifying optimal tissue slicing parameters.

1.2.3 Gap in Current Research

While there have been significant strides in both biological sample preparation and computational analysis, a gap remains in integrating these approaches to optimize tissue slicing parameters. This project aims to bridge this gap by developing a predictive model that can guide the adjustment of slicing conditions for optimal outcomes.

2 Literature review

This literature review examines the convergence of technologies in biological tissue slicing, with a particular focus on the application of image classification and deep learning to optimize slicing parameters. It aims to highlight significant advancements, identify gaps in current methodologies, and set the groundwork for the proposed project.

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/o
），

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/S010956412300
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/

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

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/s1 017-1685-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 实验操作和数据采集??? 是否需要保留还是放到下节

在实验中，使用预先从生物实验室制备好的石蜡包埋好的组织切片（预先处理好的鱼的卵巢，肺泡等），将其放在 HM355s 自动切片机上依据切片机的使用手册，不同参数执行切片操作（详细步骤见下一节）。

之后，将切好的不同类型的组织切片放在载玻片上，待其晾干后转移至 VHX7000 显微镜下，通过显微镜对没份样品进行拍照，获取到每份样品的电子图像数据。

3.2 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。

3.2.1 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/docun

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。

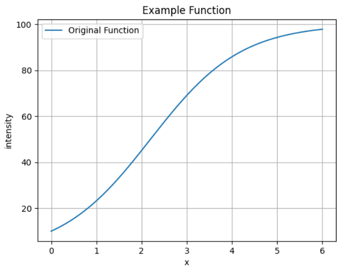


Figure 3.1　f(x)

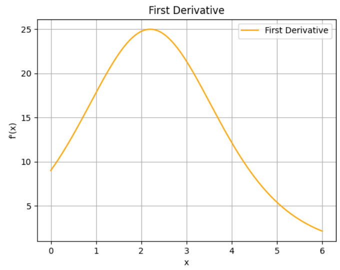


Figure 3.2　f'(x)

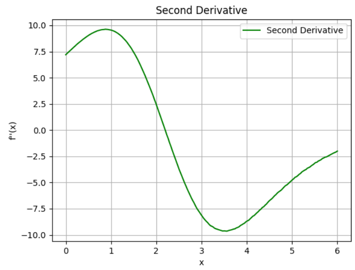


Figure 3.3　f''(x)

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：

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \tag{3.1}$$

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。

3.2.2　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（样品在制备是已染色）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。

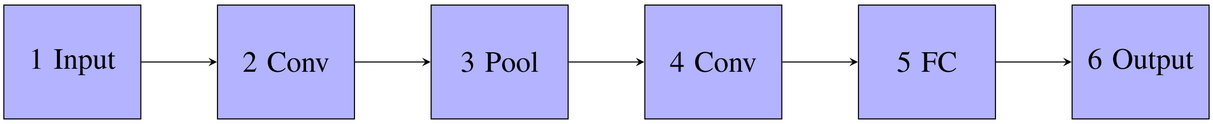
此外，关于阈值分割还有更多的方法，比如下面就是一个基于 Otsu 方法的指纹提取算法。将其用在此处能够显著提高生物组织的分割效果。Yue Yaru 和 Zhu Jialin 在《Algorithm of fingerprint extraction and implementation based on OpenCV》一文中提出了一种基于 OpenCV 的指纹提取算法。该算法对 Otsu 方法进行了改进，特别是在光照不均匀、图像模糊的情况下能够实现准确、简单、运行时间短的指纹提取。https://ieeexplore.ieee.org/document/7984539

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

3.3 深度学习

3.3.1 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。

3.3.2 迁移学习

显然，对于复杂的图像来说

4 Experimental work/analytical investigation/ design

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

5.1 Subsection 3.1

5.2 Subsection 3.2

6 Discussion and conclusions

6.1 Subsection 4.1

6.2 Subsection 4.2

7 Project management, consideration of sustainability and health and safety

7.1 Subsection 5.1

7.2 Subsection 5.2