

第1章 问题重述

1.1 问题背景

在我国，煤炭作为主要能源，支撑着国民经济的持续发展。随着国家对新质生产力的高度重视，煤矿安全、高效、智能化开采成为行业发展的核心目标。井工煤矿作为一种典型的地下工程系统，其结构复杂，运行环境特殊。其中，巷道不仅是煤炭生产的咽喉要道，更承担着井下通风、物资运输等关键功能。巷道的安全稳定直接关系到煤矿整体生产的顺利进行。

然而，巷道围岩内部存在大量肉眼难以识别的裂隙网络，这些裂隙若探测不清或控制不及时，极易引发冒顶、突水、瓦斯突出等重大安全事故，严重威胁井下作业人员的生命财产安全。因此，如何定量分析围岩裂隙的几何特征，成为保障煤矿安全生产的关键技术难题。

传统的围岩探测方法主要依赖“地质切片式”的岩芯取样。这种方法虽然能够获取一定的地质信息，但存在耗时长、成本高、劳动强度大等显著缺点，难以满足现代煤矿对高效、精准探测的需求。钻孔成像技术的出现，为煤矿巷道围岩结构的可视化探测提供了新的解决方案。该技术利用锚杆支护过程中形成的钻孔，通过高清摄像头对孔壁岩层进行 360° 全景扫描，实现岩层内部结构的高精度数据采集。钻孔成像系统由探头、主机、深度计数器、推杆和电缆线等核心部件组成，能够实时记录探头位置并传输高清图像，为后续分析提供丰富的数据基础。

钻孔成像展开图能够直观呈现岩体内部的二维裂隙图像。尽管实际地质环境中的裂隙多为复杂的三维曲面结构，但大部分裂隙曲率较低，可近似视为平面。当裂隙平面与钻孔轴线斜交时，其与孔壁的交线呈椭圆形，在展开图中则表现为周期性变化的正弦曲线，即“正弦状”裂隙。这种特征在煤系地层钻孔成像中尤为常见，成为识别裂隙的重要依据。

然而，钻孔成像图像中的“正弦状”曲线并不全部代表裂隙，有些可能是岩层界面。区分裂隙与岩层界面需结合“张开距离”和“填充材质”等特征进行判断：若曲线有明显张开且内部填充有水、气体或其他异质物质，则可判定为裂隙；反之，若曲线两侧紧密贴合且岩石过渡自然，则为岩层界面。

尽管钻孔成像技术极大提升了围岩结构的探测效率和精度，但在实际应用中仍面临三大挑战：一是地质“杂音”干扰严重，岩石天然纹理、孔壁残留泥浆及钻头痕迹常在图像中形成干扰，影响判读准确性；二是人工判读耗时长，资深工程师需依赖经验对单孔图像进行长时间分析，且不同专家间判读结果差异较大；三是地质模型重构难，现有技术难以将多钻孔的二维图像拼接为连续的三维地质模型，制约了地质信息的综合利用和智能化分析。

综上所述，如何高效、准确地识别钻孔成像图像中的裂隙，并实现三维地质模型的自动重构，是煤矿安全生产亟需解决的重要技术问题。

1.2 问题要求

本赛题希望通过研究典型煤系地层的随钻探测图像数据，将附件中的原始影像数据转化为工程可用的结构化裂隙信息，在此基础上，进一步分析裂隙的多层次特征，具体任务如下：

问题 1：针对附件 1 中的钻孔成像展开图数据，建立能够自动识别裂隙像素的数学模型。模型需充分考虑岩石天然纹理、钻头钻进痕迹、孔壁泥浆污染等多种干扰因素对裂隙识别的影响，以提升识别的准确性和鲁棒性。随后，利

用所建立的模型对附件 1 中的全部数据进行裂隙识别与分类，并生成与原图像像素尺寸一致的二值化图像，裂隙像素用黑色表示，其他像素用白色表示。

问题 2：在钻孔成像展开图的二维平面中，建立以横轴 x 表示钻孔周向展开距离、纵轴 y 表示钻孔轴向深度的坐标系。基于该坐标系，利用已知的“正弦状”裂隙的数学模型对附件 2 中的数据进行分析，实现对钻孔成像展开图中裂隙的自动聚类与表征；并应用该模型对附件 2 中的数据进行聚类分析，给出裂隙的聚类表征结果及相关分析。

问题 3：针对附件 3 中的复杂裂隙数据，建立数学模型对裂隙进行表征，并对每一条已识别裂隙提取其轮廓线离散坐标点，计算粗糙度轮廓指数（JRC）值。需结合等间距采样等方法，分析离散坐标点选取方式对 JRC 计算结果的影响，探讨是否存在更适合裂隙轮廓线特征的采样方法，比较不同采样方式的优缺点，并讨论裂隙面积在粗糙度评价中的作用。

问题 4：针对附件 4 中的数据，建立数学模型对多钻孔裂隙的连通性进行定量评估。需结合钻孔成像展开图，分析相邻钻孔中裂隙的空间表征和 JRC 值，推断裂隙的连通概率，并以三维空间结构图的方式呈现分析结果。考虑到实际工程中钻孔数量有限，连通性分析存在不确定性，需基于所建模型识别当前钻孔布局下高不确定性区域，并结合工程资源限制，按优先级排序提出 3 个最优补充钻孔位置建议。

第2章 模型假设

1. 假设题目给的数据真实可靠；
2. 假设所有图片的清晰度足够；
3. 假设钻孔成像展开图中的裂隙可以通过像素级分类模型实现自动识别；
4. 假设岩石天然纹理、泥浆污染和钻头痕迹是影响裂隙识别准确性的主要干扰因素；
5. 假设给定数据集中的裂隙和干扰因素的区分性很强。
6. 假设复杂裂隙的三维粗糙度（JRC）可以通过将展开图中的轮廓投影并结合局部曲率/高频成分估算，从而实现近似的 JRC 计算；
7. 假设不同的离散采样策略对 JRC 计算结果有系统性影响，但对裂隙分类的一致性仍然较好。

第3章 符号说明

第4章 问题 1 模型建立与求解

4.1 问题分析

在进行钻孔成像展开图中裂隙识别的任务时，面临的最大挑战是如何处理图像中的干扰因素。裂隙的边界常常受到岩石纹理、钻进痕迹以及泥浆污染等因素的影响，这些干扰不仅增加了识别的难度，还可能导致裂隙和其他不相关结构之间的混淆。岩石的自然纹理往往带有高频噪声，这种噪声在图像中与裂隙的边界形成类似的结构，而钻进痕迹和泥浆污染则会在成像中产生伪影或阴影，进而影响图像的准确性。因此，如何在噪声干扰下有效地识别裂隙，避免误识别成为了一个核心问题。

为了解决这一问题，本文提出了两条技术路径。第一条是基于传统图像处理的边缘检测方法。这种方法依赖于图像中的梯度信息，通过检测亮度的剧烈变化来提取图像中的边缘。虽然边缘检测方法在处理简单、对比度明显的图像

时效果较好，但当图像受到干扰时，传统的边缘检测算法往往会产生大量的伪边缘，导致裂隙的边界提取不准确，特别是在岩石纹理复杂或泥浆污染严重的情况下，效果十分有限。

第二条路径是基于数据驱动的深度学习方法。通过大量标注数据的训练，深度学习模型能够自动从图像中学习到复杂的特征，不仅能够识别清晰的裂隙边界，还能有效应对图像中的噪声和干扰。深度学习方法的优势在于其较强的鲁棒性和适应性，能够在复杂环境下对裂隙进行精确识别，即使在泥浆污染和钻进痕迹的影响下，仍能保持较高的识别精度。然而，这种方法也存在一定的缺点，如对计算资源的要求较高，训练过程需要大量的标注数据，同时模型的决策过程缺乏透明度，难以解释其具体的识别机制。

4.2 模型建立

4.2.1 基于传统图像处理的裂隙检测方法

一、白色泥浆伪影去除

首先对读取到的图像进行白色泥浆伪影去除，由于白色竖直条纹源于钻孔泥浆在镜头表面的强反射，其灰度值与裂隙边缘的局部高反差极为相似；若直接执行 Canny 或 Sobel 边缘检测，这些高亮伪影极易被误判为“强边缘”，导致裂隙像素的召回率虚高而精度骤降。因此预先剔除白色伪影，可从源头消除误检隐患。同时传统的边缘检测算子默认“边缘两侧灰度差异由物体结构引起”。白色泥浆区域破坏了这一空域一致性假设——其内部灰度恒定但绝对值极高，使全局高低阈值失去统计意义。移除白色伪影后，自适应阈值或非极大值抑制才能重新满足“边缘-背景”灰度分布的连续性前提，从而保证检测结果的可靠性。

可采用基于快速行进（FMM）的图像修复算法，它从待修补区域的边界出发，按像素到已知区域的距离确定修复次序，并在每一步利用邻域内已知像素的加权平均来估计缺失点的颜色与纹理；权重由梯度方向和几何距离共同决定，使得信息优先沿等照度线方向传播，从而在填补白色泥浆条纹的同时保持岩石纹理的连续性与裂隙边缘的清晰度。通过“检测-掩膜-修复”这一空域预处理链路，模型能够在传统边缘检测框架下，显著抑制泥浆伪影干扰，为裂隙像素的自动、精准识别奠定坚实基础。

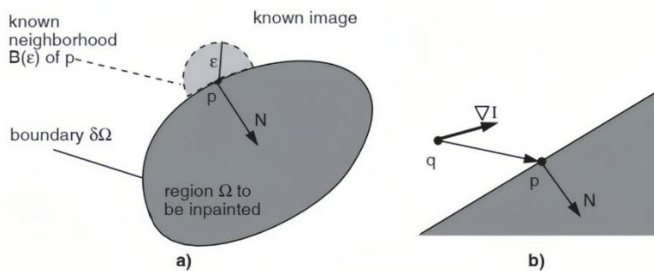


图 4-2 FMM 算法示意图

先在掩膜边界上建立到未知区域内部每个像素的距离函数 $T(x)$ ，然后按 T 从小到大逐像素填充。距离函数由 Eikonal 方程在待填充区域 Ω 中并在边界 $\partial\Omega$ 处满足 $T=0$ 来定义，数值上通过 FMM 求解得到每个像素 x 的到达时间或到边界的最短距离 $T(x)$ ，从而保证信息由边界向内部逐层传播，减少远处信息对近处细节的不当影响。在具体的像素值估计上，待填像素 p 的强度（或颜色向量） $I(p)$ 由其邻域内已知像素 $q \in N(p) \cap \Omega_{\text{known}}$ 的加权平均给出，数学表达为：

在具体的像素值估计上，待填像素 p 的强度（或颜色向量） $I(p)$ 由其邻域内已知像素 $q \in N(p) \cap \Omega_{\text{known}}$ 的加权平均给出，数学表达为：

$$I(p) = \frac{\sum_{q \in N(p) \cap \Omega_{\text{known}}} w(p, q) I(q)}{\sum_{q \in N(p) \cap \Omega_{\text{known}}} w(p, q)}.$$

式中 $N(p)$ 表示以 p 为中心的邻域窗口， $I(q)$ 为像素 q 的已知值，权重函数 $w(p, q)$ 决定了不同已知像素对 p 的贡献大小。按照此顺序逐像素填充，填完一个像素后即将其加入已知集合继续处理下一个像素，直至区域填充完成。

权重 $w(p, q)$ 的设计旨在使信息优先沿等照度线或边缘方向传播，同时兼顾几何距离和局部纹理稳定性。可将权重分解为几部分的乘积形式

$w(p, q) = w_d(p, q) \cdot w_{\text{dir}}(p, q) \cdot w_{\text{lev}}(q)$ ，其中几何距离项 $w_d(p, q)$ 使近邻像素贡献更大；方向项 $w_{\text{dir}}(p, q)$ 借助边界法向量 $\frac{\nabla M(p)}{\|\nabla M(p)\|}$ （ M 为掩膜，已知处取 1，待填充处取 0）衡

量 $(p-q)$ 在法向方向上的投影，常取 $\frac{|(p-q) \cdot n(p)|}{(\|p-q\| + \delta)}$ ，以强化沿等照度线传播的优先性；强度/纹理抑制项 $w_{\text{lev}}(q)$ 用于降低高梯度点（可能为边缘或噪声）对填充值的影响。将这些因子组合后可以得到类似：

$$w(p, q) \approx \frac{|(p-q) \cdot n(p)|}{(\|p-q\| + \delta)(\|p-q\|^2 + \varepsilon)(1 + \|\nabla I(q)\|)}$$

将进行了白色泥浆伪影剔除后的图像转换成单通道灰度图像，因为裂隙在亮度轴上以“暗谷”形式呈现，梯度幅值可直接、唯一地反映岩石—裂缝间的真实对比度，避免三通道梯度合并带来的方向歧义和颜色噪声。

二、双边滤波

接下来对去除白色伪影后的灰度图像进行双边滤波处理。双边滤波既能有效抑制局部纹理噪声与小尺度伪影，又能尽量保持裂隙等显著边缘这一重要特性。双边滤波在后续的边缘检测中发挥两重作用：一方面降低了由岩石天然纹理、细粒噪声或修补伪影产生的高频干扰，从而减少了伪边缘的响应；另一方面由于其“强度相似性”的加权机制，真实的强边缘不会被平滑掉，从而保证了边缘算子能识别出连贯且清晰的裂隙轮廓。对钻孔成像这类既有粗糙纹理又有窄弱裂隙的图像，双边滤波能够在不破坏结构信息的前提下提高信噪比，这对于后续基于阈值的边缘检测尤为重要。

从原理上讲，双边滤波对像素 p 的输出值是其邻域内像素 q 的加权平均，权重由空间距离和像素强度差两个因子共同决定。用数学形式可表示为：

$$I_{\text{filtered}}(p) = (1/W_p) * \sum_{q \in N(p)} G_s(\|p-q\|) \cdot G_r(|I(p)-I(q)|) \cdot I(q)$$

其中 $W_p = \sum_{q \in N(p)} G_s(\|p-q\|) \cdot G_r(|I(p)-I(q)|)$ 为归一化项， G_s 和 G_r 分别为空间高斯核和强度高斯核。空间项 $G_s(\|p-q\|)$ 随像素间几何距离衰减，保证近邻像素贡献更大；强度项 $G_r(|I(p)-I(q)|)$ 则根据灰度差判断相似性，灰度差越大，该像素对 p 的贡献越小，从而实现“跨边缘抑制平滑”的效果。

在本问题中优先选择双边滤波，主要因为它具备“保边缘去噪”的特性：一方面通过空间权重抑制邻域噪声，另一方面通过强度相似性权重避免跨越明显亮暗跳变

（即裂隙边缘）进行平滑，从而在降低岩石纹理与泥浆带来的高频干扰时尽量保持窄而清晰的裂隙边界。相比之下，简单的线性平滑虽能去噪但会模糊细窄裂隙；中值滤波对椒盐噪声有效但可能破坏细长结构；更复杂的方法在效果上可能更好，但实现复杂、参数敏感且计算开销更大。双边滤波在效果、可解释性与计算成本之间取得了较

好的折衷，并且在工程实现中易于通过窗口大小与色差参数进行直观调优，因此作为传统预处理流程的首选更为合适。

三、Canny 边缘检测

再对图像进行 Canny 边缘检测。Canny 边缘检测算法是计算机视觉领域中常用的边缘检测算法之一。它的主要特点是能够准确地检测到图像中的边缘，同时对于噪声有较好的抵抗能力。该算法主要由以下几个部分构成：

1. 灰度化处理：将输入的彩色图像转换为灰度图像，灰度化公式如下：

$$I_{\text{gray}} = 0.299 \times I_R + 0.587 \times I_G + 0.114 \times I_B$$

其中， I_{gray} 表示灰度化后的像素值， I_R 、 I_G 、 I_B 分别为原图像对应像素的红、绿、蓝通道上的像素值。

2. 高斯滤波：使用高斯滤波器对灰度图像进行平滑处理，以减少噪声的影响；

$$G(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right)$$

其中， $G(x, y)$ 表示高斯函数， σ 表示滤波器的标准差。

3. 计算图像梯度：使用 Sobel 算子计算灰度图像在水平和垂直方向的梯度值，从而得到每个像素点的梯度强度和方向。Sobel 算子公式和梯度幅值和方向公式如下：

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} * I_{\text{gray}}, G_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * I_{\text{gray}}$$

G_x 和 G_y 分别表示水平和垂直方向上的梯度值， $*$ 表示卷积运算符。

$$M(x, y) = \sqrt{G_x^2(x, y) + G_y^2(x, y)}, \theta(x, y) = \tan^{-1}\left(\frac{G_y(x, y)}{G_x(x, y)}\right)$$

其中， $M(x, y)$ 表示该像素点的梯度强度， $\theta(x, y)$ 表示该像素点的梯度方向。

4. 非极大值抑制：对于每个像素点，比较其梯度方向上的邻近两个像素点之梯度值的大小，只保留梯度方向上梯度值极大的像素点，有利于消除边缘上的模糊效果。非极大值抑制公式如下：

$$M'(x, y) = \begin{cases} M(x, y) & \text{if } M(x, y) \geq M_L \text{ and } M(x, y) \geq M(x + \delta_x, y + \delta_y) \\ 0 & \text{otherwise} \end{cases}$$

其中，为梯度强度的低阈值， δ_x 和 δ_y 为梯度方向的偏移值。

5. 双阈值处理：将像素点分为三类：强边缘、弱边缘和非边缘。当像素点的梯度值高于较高的阈值时，将其分类为强边缘。当像素点的梯度值低于较低的阈值时，将其分类为非边缘。当像素点的梯度值介于较高和较低的阈值之间时，将其分类为弱边缘。双阈值处理公式如下：

$$\text{Canny}(x, y) = \begin{cases} \text{Strong} & \text{if } M'(x, y) \geq M_H \\ \text{Weak} & \text{if } M'(x, y) < M_H \text{ and } M'(x, y) \geq M_L \\ \text{Non-edge} & \text{otherwise} \end{cases}$$

其中，为梯度强度的高阈值，Strong、Weak 和 Non-edge 分别表示强边缘、弱边缘和非边缘。

6. 边缘连接：将弱边缘和其周围的强边缘连接起来形成完整的边缘。

在本问题中选择 Canny 边缘检测的优点在于它能以严格且可控的步骤同时兼顾边缘定位精度与抗噪能力：通过先平滑再求梯度计算出可靠的幅值与方向，再经非极大值抑制将边缘细化为近单像素宽度，从而获得高精度的裂隙中心线位置；其双阈值与滞后连接机制能够在抑制孤立噪声的同时保留与强边缘连通的弱边缘，使得对窄弱裂隙的召回率得到保障。Canny 的参数易于结合验证集或分组噪声场景进行调优，且算法成熟、计算代价适中、实现与调试方便。

在前面的数据处理过程中仅对白色泥浆区域进行伪影剔除是因为白色条纹是镜头表面泥浆的强反射，灰度值饱和且梯度极性恰好与裂隙“暗谷”相反，一旦进入 Canny 链路就会被当成正向强边缘而直接淹没真实裂隙，因此必须在边缘检测前用修复手段彻底抹除；黑色竖纹则源于孔壁泥浆流痕或阴影，其灰度虽低却仍处于裂隙“暗谷”区间，梯度极性与裂隙一致，后续中值滤波即可在抑制随机噪点的同时保留该低频暗区，Canny 高低阈值进一步把低于裂隙对比度的弱阴影抑制为零，而不损伤显著裂隙边缘，因而无需额外去除即可实现可靠提取。

四、轮廓提取

最后对检测到的裂隙进行轮廓提取。首先对整幅二值边缘图进行拓扑扫描，仅提取最外层曲线，将内部封闭或孤立的边缘全部剔除，从而一次性滤除岩石颗粒、刀痕划痕等产生的细碎噪点与纹理碎片；随后引入弧长门槛，对保留下来的曲线进行第二次筛选——长度低于设定阈值的短枝被视为无关毛刺而舍弃，剩余的长轮廓则具备跨度大、走向连贯的特点，其几何特征与钻孔展开图中呈正弦状延展的真实裂隙高度吻合，由此形成高置信度的裂隙候选集，为后续的形状分析、曲率计算或分类识别提供了干净且可靠的数据基础。

4.2.2 基于深度学习的裂隙检测方法

一、性能评估指标

模型的性能评估指标是对模型预测性能的一种衡量方式，它们可以帮助我们了解模型在各种情况下的表现。以下是本文所涉及到的相关指标的简要介绍和对应的计算公式：

(1) 准确率 (Accuracy)：准确率是模型预测正确的样本数占总样本数的比例。它是一个直观的性能指标，但有时并不能完全反映模型的性能，特别是在样本分布不平衡的情况下。计算公式为：

$$\text{Accuracy} = \frac{(\text{TP} + \text{TN})}{(\text{P} + \text{N})}$$

(2) 召回率 (Recall)：召回率又称为真正例率 (TPR)，它表示的是模型正确识别出来的正样本占有所有实际正样本的比例。对于预测正样本的模型，召回率是一个重要的指标。计算公式为：

$$\text{Recall} = \frac{\text{TP}}{\text{P}}$$

(3) 精确率 (Precision)：精确率表示的是模型预测为正样本的样本中，实际为正样本的比例。它反映了模型预测为正样本的准确性。计算公式为：

$$\text{Precision} = \frac{TP}{(TP + FP)}$$

(4) F1 Score: F1 Score 是精确率和召回率的调和平均数，它旨在平衡这两个指标，以便更全面地评估模型的性能。在图像分割任务中，也可以称之为 Dice 系数，反映预测区域和真实区域的相似性。计算公式为：

$$\text{F1 Score} = \frac{2 \times (\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})}$$

(5) 交并比 (Intersection over Union, IoU) : IoU 表示模型预测区域与真实区域的交集与并集之间的比率，取值范围为[0, 1]。假设目标的标签包围区域为 A，检测结果输出的包围区域为 B，则 IoU 定义为：

$$\text{IoU} = \frac{|A \cap B|}{|A \cup B|}$$

其中，IoU 数值越大表示模型预测的区域与真实区域重叠程度越高，算法性能越好。

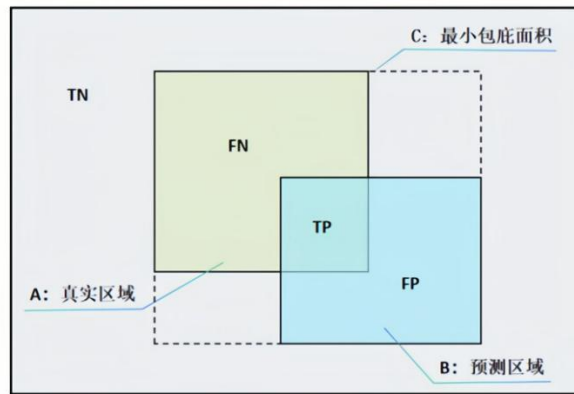


图 1 IoU 体系计算方式示意图

二、数据预处理

1. 数据集标注

在模型训练数据准备阶段，首先需要进行数据集的标注，将附件 1、附件 2 和附件 3 中初始的图片进行打标签操作，以提高模型在语义分割任务中的性能。

本文选择了开源标注工具 Labelme，来实现对图像的像素级标注。标注的过程中，聚焦于绘制裂隙的边缘，确保每个裂隙的区域都被准确标注。这一步骤的重点在于为每个像素赋予相应的语义标签，从而为模型提供清晰的训练信号。

在对原始图像进行标签标注后，将得到相应的 JSON 文件，每个 JSON 文件记录了详细的标注信息，包括对象的类别、位置和形状等。

```
"label": "structure",
"points": [
  [
    5.606557377049199,
    60.24590163934426
  ],
  [
    16.754098360655746,
    61.22950819672131
  ],
  [
    42.327868852459034,
    67.78688524590163
  ],
],
```

图 1-10 JSON 文件内容示意图

标注完成后，将标签结果转化为像素级标签，即生成与每个图像大小相匹配的掩码。在掩码中，每个像素的数值代表其所属的类别，而裂隙的区域则被赋予特定的像素值。得到的标签不仅可以有效地指导模型对坑洼进行语义分割，还使得模型在训练过程中能够更好地理解图像中的局部结构。

再根据 JSON 标注文件中的几何形状信息，在对应的图像上生成掩码图像。

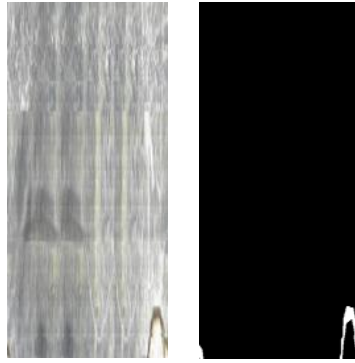


图 1-11 掩码图示例

2. 图像增强

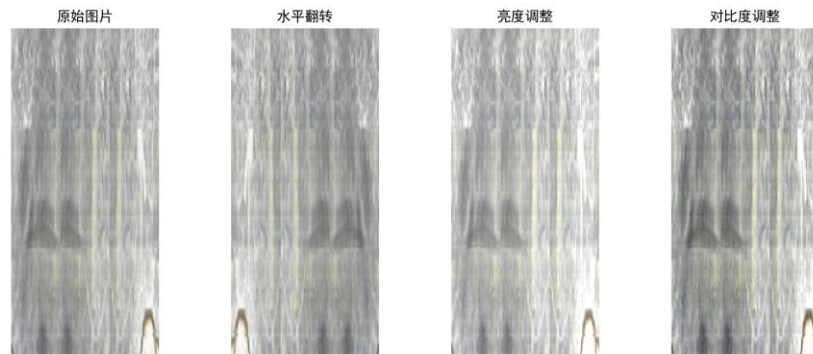
由于初始图像数据量有限，可能难以满足深度学习模型的训练需求，因此可以考虑对图像进行适当的增强处理，以提升模型的泛化能力和训练效果。

本问中对初始图像和掩码图像进行了滑动窗口裁剪操作。初始图像的大小为 244×1350 ，因此裁剪图像大小为 122×270 ，横向和纵向裁剪步距分别为 61 和 50，共获取 630 张像素尺寸为 122×270 的图像。



图【1】 裁剪后的图像示例

在裁剪图像的基础上，进一步采用水平翻转、随机调整亮度和对比度等图像增强手段，以增加数据的多样性和数量，提升模型的泛化能力和鲁棒性。



图【2】 图像增强示例

3. 数据集划分

完成标注后，将得到的全部图片按照 8：2 划分为训练集和测试集，为模型训练和评估提供相应的数据集。

只使用裂隙图片的打标签策略专注于坑洼的语义信息，使得模型在裂隙区域的分割、效果得到显著提升。通过这样的训练方式，使模型在裂隙检测中更准确地检测和钻孔成像展开图中的裂隙，提升裂隙检测的效果。

三、UNet

UNet 是一种广泛应用于医学图像分割任务的深度学习模型，首次由 Olaf Ronneberger 等人在 2015 年提出。UNet 的核心思想是通过编码器和解码器架构进行逐步提取和恢复图像的空间信息，特别是在像素级别的图像分割任务中表现突出。与传统的卷积神经网络不同，UNet 在设计上特别关注了图像的细粒度特征恢复，采用了跳跃连接来有效结合编码器和解码器的特征图，从而弥补了下采样过程中丢失的空间信息。

UNet 网络结构的核心包含一个对称的编码器和解码器部分。编码器主要通过卷积层进行特征提取，并通过池化操作逐步降低特征图的空间分辨率。在这个过程中，图像的语义信息逐渐得到丰富，但细节部分逐步丢失。解码器则采用反卷积或上采样操作逐步恢复图像的分辨率，最终输出与输入图像大小相同的分割结果。

在具体的网络架构中，UNet 使用了跳跃连接的策略。跳跃连接将编码器中的特征图直接连接到解码器的相应层，从而避免了解码器上采样过程中的空间信息丢失。跳跃连接在每个卷积块后都会将该层的特征图与解码器上采样后得

到的特征图进行拼接，这使得解码器可以利用编码器中的低级特征来改善上采样后的细节恢复。

UNet 的计算过程可以通过以下几个关键的卷积操作和反卷积操作来表示。首先，输入图像 X 通过一系列的卷积操作 Conv 和池化操作 Pool 进行下采样。对于输入的图像 X ，编码器部分的卷积层输出特征图 F_{enc} 可以表示为：

$$F_{\text{enc}} = \text{Conv}(\text{Pool}(X))$$

这里， Conv 表示卷积操作， Pool 表示池化操作， F_{enc} 为编码器部分生成的特征图。

在解码器部分，反卷积操作（即上采样）用于逐步恢复空间分辨率。设解码器输入为 F'_{enc} （编码器特征图），解码器的输出 F_{dec} 为：

$$F_{\text{dec}} = \text{Conv}(\text{UpSample}(F'_{\text{enc}}))$$

这里， UpSample 表示上采样操作，通常使用反卷积或者转置卷积操作。

跳跃连接的关键是将编码器的特征图 F_{enc} 与解码器相应层的特征图进行拼接。设解码器第 i 层的上采样输出为 $F^{(i)}_{\text{dec}}$ ，跳跃连接后的特征图 $F^{(i)}_{\text{skip}}$ 表示为：

$$F^{(i)}_{\text{skip}} = \text{Concatenate}(F^{(i)}_{\text{enc}}, F^{(i)}_{\text{dec}})$$

跳跃连接通过拼接编码器和解码器的特征图，增强了解码器的细节恢复能力。这一策略显著提升了网络在高分辨率图像上的分割效果。

UNet 通常使用交叉熵损失函数 \mathcal{L}_{CE} 来优化分割模型，公式为：

$$\mathcal{L}_{\text{CE}} = - \sum_{i=1}^N y_i \log(\hat{y}_i)$$

其中， y_i 为图像中第 i 个像素的真实标签（0 或 1）， \hat{y}_i 为该像素的预测概率， N 为图像中像素的总数。

本任务中也可结合 Dice 系数损失，该损失函数可以通过以下公式表示：

$$\mathcal{L}_{\text{Dice}} = 1 - \frac{2 \sum_{i=1}^N y_i \hat{y}_i}{\sum_{i=1}^N y_i + \sum_{i=1}^N \hat{y}_i}$$

这里， y_i 和 \hat{y}_i 分别为真实标签和预测概率，Dice 系数衡量了预测与真实标签的相似度，值越大代表分割效果越好。Dice 损失函数适用于数据集中的类别不平衡问题，能够鼓励模型关注难以识别的小区域。

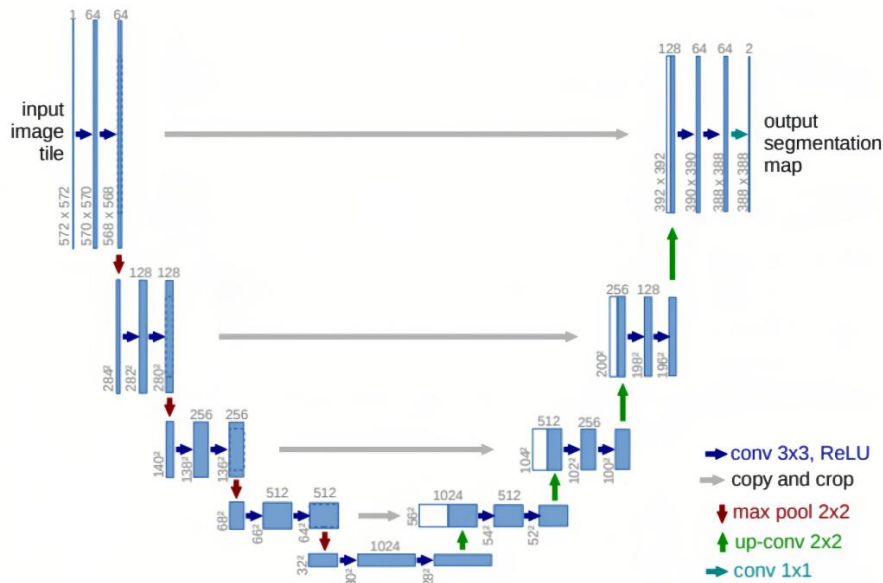


图 4.2 UNet 结构示意图

4.3 模型求解

4.3.1 基于传统图像处理的裂隙检测方法

基于传统图像处理的裂隙检测模型求解流程图如图 4.3 所示：

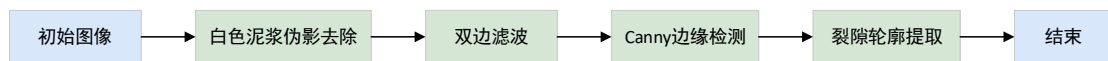


图 4.3 基于传统图像处理的裂隙检测模型求解流程图

一、白色泥浆伪影去除

进行白色伪影去除后的图像如图 4.4 所示：



图 4.4 去除白色伪影后图像示例

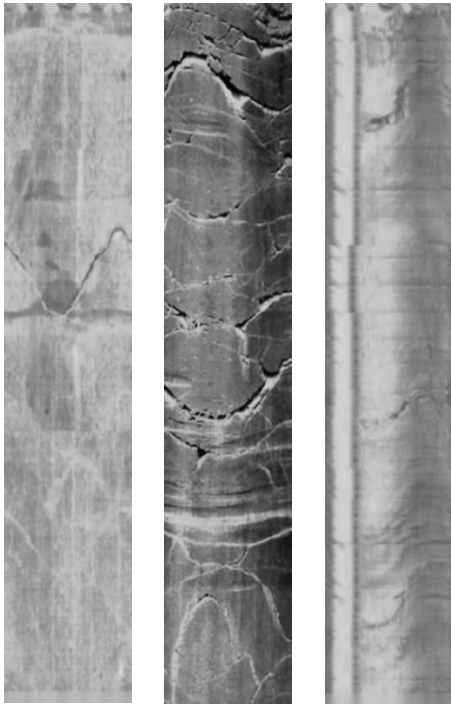
从图 4.4 可以看出，在本次图像处理中，通过白色伪影去除有效地消除了因钻孔泥浆粘附在成像镜头表面而形成的白色竖直区域。这些区域由于泥浆的高反射特性，在光源照射下呈现出过度的高亮度，从而在图像中形成了明显的

白色伪影，影响了图像的质量和后续分析。通过应用图像修复算法，针对这些白色伪影区域，本文通过填补像素值并使其与周围区域的视觉信息相融合，成功地去除了这些不真实的亮区。

这一处理过程有效地恢复了图像的真实性，使得背景和物体的边缘更加清晰，减少了由于伪影干扰可能导致的错误分析。去除伪影后，图像中原本受影响区域的细节得以恢复，有助于后续的图像处理和分析，提高了结果的准确性。

二、双边滤波

进行双边滤波处理后的图像如图【】所示：



图【】 双边滤波处理后图像示例

在对白色伪影去除后的图像进行双边滤波处理时，图像的细节和边缘得到了有效保留，同时噪声和细微的灰度变化被平滑处理。双边滤波是一种非线性滤波技术，它能够在去除噪声的同时，保留图像中的边缘信息。与传统的平滑方法不同，双边滤波考虑了像素的空间距离和灰度值的差异，因此能够精确地处理图像中的局部细节，尤其是在边缘和纹理区域。

从图【】可以看出，在处理过程中，双边滤波有效地抑制了图像中的细小噪声，并平滑了背景区域，使得图像的整体质量得到改善。特别是在白色伪影去除后的图像中，滤波后不仅减少了可能存在的细小杂点，而且保留了钻孔周围的关键结构和细节。这为后续的边缘检测和轮廓分析提供了更加干净、清晰的图像基础，从而提高了分析结果的精确性。

三、Canny 边缘检测

进行 Canny 边缘检测后的图像如图【】所示：

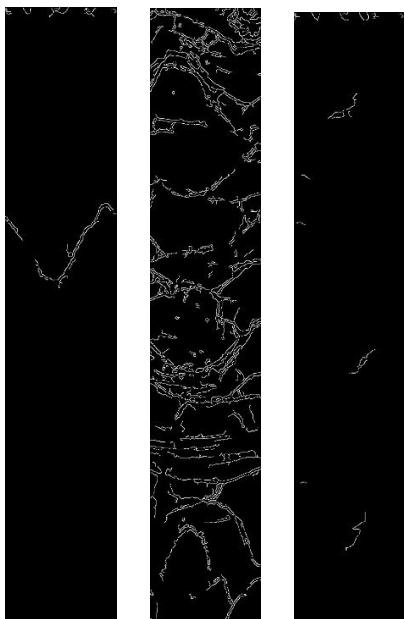


图 10 Canny 边缘检测后图像示例

从图 10 可以看出，经过 Canny 边缘检测后的图像，边缘信息得到了显著提取，图像中的主要结构和轮廓变得更加清晰。Canny 边缘检测是一种经典的边缘检测算法，通过多阶段处理来精确识别图像中的裂隙边缘。其优势在于能有效区分边缘和非边缘区域，尤其在处理具有明显对比度变化的图像时，能够准确地捕捉到物体的轮廓。

在本次图像处理中，Canny 算法成功地提取了图像中的显著边缘，尤其是钻孔周围的结构边界和细节。通过边缘检测，可以清晰地识别出原图中细微的轮廓变化，同时滤除掉一些不必要的噪声，使得图像的轮廓更加简洁明了。Canny 边缘检测后的图像提供了一个干净的基础，便于后续的轮廓分析和特征提取，有助于提高后续处理步骤的精度。

四、轮廓提取

进行轮廓提取后的图像如图 11 所示：

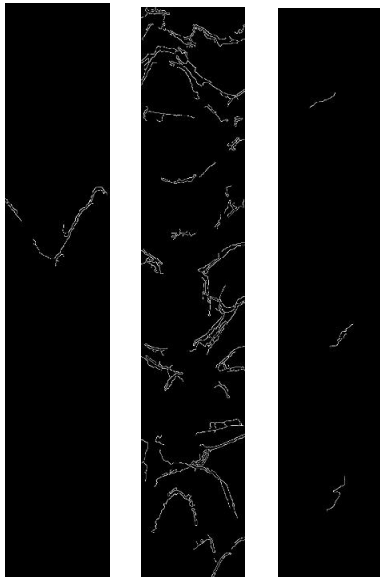


图 11 轮廓提取后图像示例

从图【】可以看出经过轮廓提取后的图像，物体的形状和结构得到了有效的突出，主要轮廓被准确地识别出来，图像中的重要特征变得更加明显。轮廓提取是图像分析中常用的一种技术，能够从复杂的背景中分离出感兴趣的区域，尤其是在图像中存在明显边缘时，轮廓提取能够清晰地呈现出物体的外部形态。

在本次处理过程中，轮廓提取成功地识别了图像中的裂隙结构，特别是钻孔周围的复杂形状和曲线部分，能够精确地捕捉到其边界。通过过滤掉较小或不相关的轮廓，系统有效地减少了噪声的干扰，使得图像中的重要信息更加突出，确保了后续分析能够集中在真正有意义的区域。

最后对提取到的图像进行转色处理以满足题目的要求。转色结果如图【】所示：



图【】 传统图像处理方法裂隙提取结果

4. 3. 2 基于深度学习的裂隙检测方法

基于深度学习的裂隙检测模型求解流程图如图【】所示：



图【】 基于深度学习的裂隙检测模型求解流程图

一、模型参数

参数符号	大小设置	意义
batch_size	8	每次训练和验证时加载的图像样本数量

epochs	50	模型训练的总轮数
lr	1e-4	优化器学习率，控制每次更新参数的步长
in_channels	3	输入图像的通道数
classes	1	输出类别数
encoder_name	resnet50	编码器网络架构，选择 ResNet50 做基础网络
encoder_weights	imagenet	用 ImageNet 上预训练的权重初始化编码器

二、模型训练效果

1. loss 曲线

使用处理后的数据集对 UNet 模型进行训练，训练过程中的损失曲线如图 1 所示：

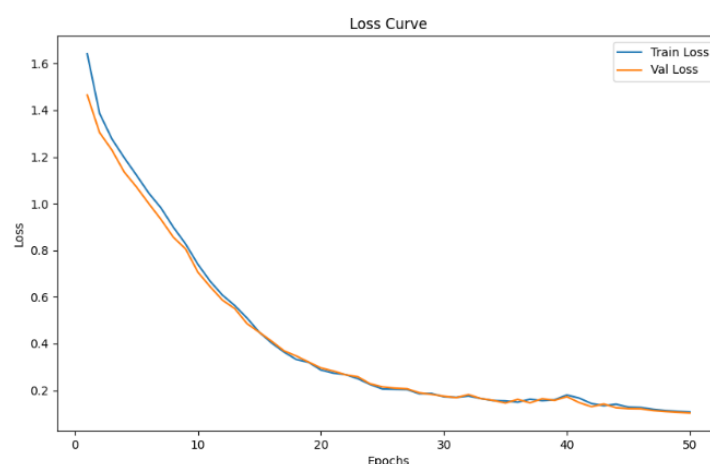


图 1 模型训练 loss 曲线

从图 1 可以看出在，在模型训练过程中，训练损失从第一轮的 1.6414 降至第 50 轮的 0.1076，验证损失从 1.4651 下降 0.1033，这表明模型在目标裂隙图像的学习任务中逐渐取得了显著的进展。最初阶段，较高的损失值表明模型未能有效地提取裂隙特征，可能存在较多的误差和不精确的预测。然而，随着训练的深入，模型通过优化参数逐步降低了误差，使得训练损失和验证损失都呈现出持续下降的趋势。

训练过程中的损失下降速度的逐步减缓表明，模型逐渐接近于最优解，并且在进一步优化时需要更精细的调整。在后期，损失曲线趋于平稳，表明模型在训练和验证集上都达到了较好的性能平衡，进一步的损失下降变得困难，通常是因为模型已经捕捉到了大部分有用特征，且不易通过常规训练过程进一步优化。

从损失值的变化可以推测，模型在目标裂隙图像的分割任务中逐渐学会了如何区分裂隙与背景，在图像的裂隙区域提取上取得了较好的效果。这意味着模型能够有效识别裂隙区域的形状和纹理，并且减少了假阳性和假阴性的产生。训练集和验证集上的损失下降趋势一致，表明模型不仅在已知数据上有良好的表现，也具备较强的泛化能力，能够对未见数据进行较为准确的预测。

2. 评价指标

使用测试集对模型的性能进行评估，各项指标如表【】所示：

Accuracy	Recall	Precision	F1 Score	IoU
89.56%	86.58%	86.88%	86.71%	86.10%

表【】 Unet 评价指标

通过对表【】的观察，可以看出模型准确率达到了 89.56%，这一数值表明模型整体的分类正确率较高，能够有效地将图像中的裂隙区域与背景进行区分。模型的召回率为 86.58%，这一指标表明模型在裂隙区域的检测上表现较为出色。召回率反映了模型能够识别出所有实际裂隙区域的能力，较高的召回率意味着模型能够较好地捕捉到裂隙的存在，减少漏检的情况。这对于裂隙图像分割任务尤为重要，因为漏检裂隙可能会导致严重的后果，特别是在一些需要高精度检测的实际应用场景中。模型精确率为 86.88%，这个结果表明模型对于检测出的裂隙区域的预测质量较高，即在所有被模型判定为裂隙的区域中，约 86.88%是正确的。这一数值的较高表现表明，模型在分割裂隙时能够有效避免过多的假阳性，减少了将背景区域误判为裂隙的情况，这对于实际应用中的可靠性和准确性具有重要意义。模型的 F1 Score 为 86.71%，这一指标是精确率和召回率的调和平均，综合反映了模型在裂隙检测任务中的平衡性能。较高的 F1 分数表明，模型在精确率和召回率之间取得了良好的平衡，既能保证较高的裂隙检测率，又能有效控制假阳性，避免过多的错误分割。模型的交并比为 86.10%，这一指标是衡量预测分割区域与真实区域重叠程度的标准之一。较高的 IoU 值意味着模型在裂隙区域的分割上与真实标签之间有较高的重叠度，能够较为准确地界定出裂隙的边界。

综合来看，模型在裂隙图像分割任务中的表现非常优秀。准确率、召回率、精确率、F1 分数和 IoU 都处于较高水平，说明该模型能够较为准确地分割裂隙区域，且在不同评估指标上都能取得较好的平衡。尽管在实际应用中，依然可能存在个别复杂裂隙形态导致的分割不完美，但整体的性能已经表明该模型在目标任务中的有效性和实用性。这些性能指标为模型的后续优化提供了有力的依据，同时也为实际部署提供了可靠的保障。

3. 预测结果

使用训练得到的模型对附件 1 中的图片进行预测，得到模型的的预测结果示例如图【】所示：

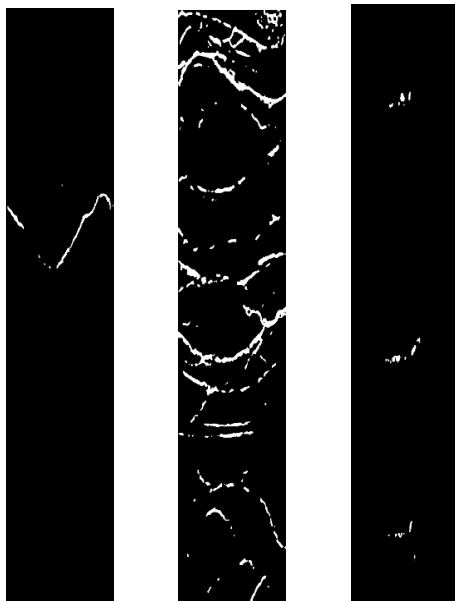


图 4-10 UNet 模型预测结果

最后对提取到的裂隙图像进行转色处理以满足题目的要求。转色结果如图 4-11 所示：

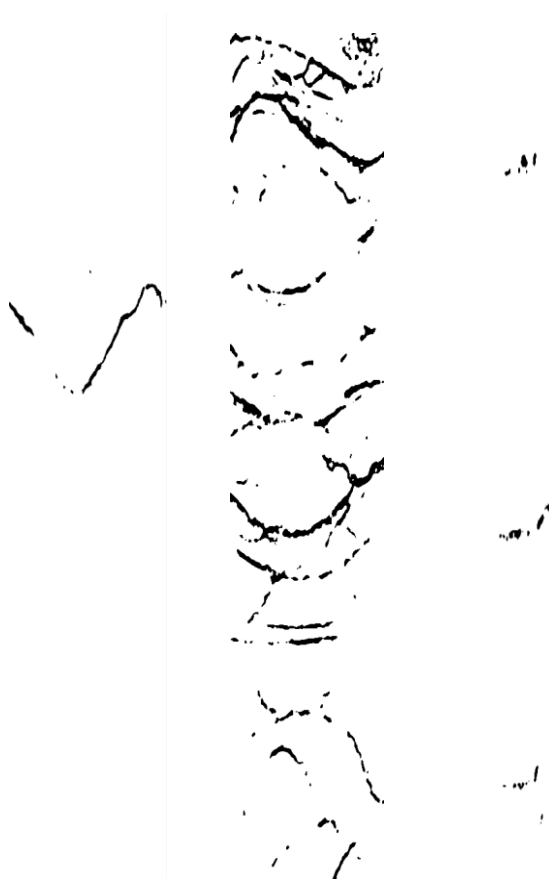


图 4-11 深度学习方法裂隙提取结果

4.4 章节小结

在钻孔成像展开图中的裂隙识别任务中，传统的图像处理方法和深度学习方法都能实现题目要求。尽管传统方法在一些简单任务中可能表现出较好的性

能,但对于复杂且多变的裂隙图像,尤其是在裂隙形态复杂、背景干扰强烈的情况下,UNet 方法显然更具优势。

从模型的可解释性角度来看,传统的图像处理方法通常依赖于手工设计的特征提取步骤,比如边缘检测、阈值分割、形态学操作等。这些方法尽管在特定条件下能有效处理图像中的简单结构,但在裂隙图像的识别中往往无法处理多样化和复杂的形态。传统方法的可解释性相对较强,因为每个步骤的逻辑都是明确的,且通常依赖于人工设计的规则。然而,随着图像背景、光照、噪声等因素的变化,这些方法的效果往往大打折扣。相比之下,UNet 作为深度学习模型,其核心优势在于通过大量的训练数据自动学习图像中的特征。虽然 UNet 的“黑箱”性质使得它在理论上缺乏直接的可解释性,但通过可视化其卷积层的特征图和反向传播的梯度信息,依然可以获得一定程度的模型可解释性,理解其关注的关键特征区域。此外,UNet 的网络架构能够在多尺度特征上进行高效学习和融合,从而有效捕捉到裂隙的细节和全局信息,这是传统方法所无法实现的。

从实际工作场景中的可靠性角度来看,传统图像处理方法在固定且较为简单的场景中可能能够提供足够的可靠性,但在实际的钻孔成像任务中,裂隙的形态多变且背景环境复杂,传统方法往往会受到光照变化、噪声干扰等因素的影响,导致识别准确率下降。传统方法依赖于预设的阈值和规则,在面对复杂背景时容易出现误检或漏检,尤其是裂隙区域的形态不规则或在图像中的位置较为微小时,传统方法常常难以精确识别。而 UNet 模型通过大规模的数据训练,能够在不同的工作环境下自适应调整,显著提高对复杂裂隙的识别能力。由于 UNet 能够处理高维的特征信息,且通过端到端的训练过程自动优化网络参数,它在处理复杂的裂隙形态、背景变化、噪声干扰等方面表现出更强的鲁棒性和适应性。

此外,UNet 在处理高复杂度任务时的效率也远远超过传统图像处理方法。传统方法需要通过一系列手工设计的特征提取、预处理、分割步骤,这些步骤需要精确的参数调节,并且对不同场景和图像的适应能力差,通常需要根据不同的数据进行多次调整和优化。而 UNet 通过深度学习的端到端训练,能够从数据中自动提取特征,省去了大量人工设计的工作,并且其训练出的模型能够迅速适应不同的钻孔成像图像,减少了对人工干预的依赖。

在进行综合比较后,本文最终选择了深度学习作为裂隙识别的主要方案。虽然传统图像处理方法在计算速度和部署上具有优势,但由于其在复杂图像环境下的识别精度较低,难以应对图像中的各种干扰,因此无法满足实际应用的需求。深度学习方法尽管在训练时需要更多的计算资源和数据,但其在识别精度、鲁棒性以及复杂环境的适应性方面展现出了明显的优势,特别是在面对岩石纹理、钻进痕迹和泥浆污染等干扰因素时,能够有效地分辨出裂隙区域。因此,深度学习被选定为最终的解决方案。

第5章 问题 2 模型建立与求解

5.1 问题分析

解决钻孔成像展开图中裂隙的自动聚类与表征问题时,需要从多个技术角度出发,综合运用不同的数学与计算方法来实现对裂隙形态的准确提取和归类。该问题涉及到的核心任务是对“正弦状”裂隙进行拟合、聚类和分析,以便得到裂隙的特征参数并进一步分析其变化规律。为了实现这一目标,首先需要对数据进

行适当的预处理，然后通过合适的数学模型进行建模与拟合，最后通过评估结果的精度和拟合效果来完成问题的解决。

原始的钻孔成像数据中可能包含噪声或不完整的信息，这会影响后续分析的准确性。为了提高数据的质量，可以通过应用高斯滤波技术来对噪声进行平滑处理。高斯滤波通过对数据点进行加权平均，能够有效地减少噪声的影响，使得裂隙的轮廓更加清晰。

为了从处理后的数据中提取出裂隙的轮廓，骨架化技术可以得到裂隙的骨架形态，使得裂隙的几何特征更加明确。骨架化是将裂隙的形态简化为一条中心线，方便后续的聚类和分析。此外，连通分析可以进一步帮助我们识别裂隙的连通性与分布特征，去除孤立或不连贯的部分，确保所提取的裂隙数据在空间上具有一定的连贯性。

在完成数据预处理后，可使用 DBSCAN（密度聚类）方法对裂隙进行聚类。DBSCAN 是一种基于密度的聚类算法，它能够自动识别数据中的密集区域，并将相似的裂隙归为同一类。相比于传统的 K 均值聚类，DBSCAN 不需要事先指定聚类数目，且对于噪声和异常点具有较强的鲁棒性，这使得其在处理钻孔成像数据时表现出较高的可靠性。

一旦完成裂隙的聚类，我们便可以通过正弦函数对每个聚类结果进行拟合。通过非线性最小二乘拟合方法，能够精确地拟合每个裂隙的形态，并求解出上述模型中的各个参数。拟合过程中的优化目标是 minimized 模型预测值与实际数据之间的误差，这通常通过最小二乘法来实现。

为评估拟合效果，可以使用均方根误差（RMSE）和决定系数（ R^2 ）等指标对拟合结果的精度进行量化。均方根误差可以衡量拟合曲线与实际数据点之间的偏差，决定系数则能反映模型拟合的总体精度和拟合效果。理想的情况下，RMSE 越小， R^2 越大，表示模型的拟合效果越好。

5.2 模型建立

一、连通组件标记算法

连通组件标记算法（Connected Components Labeling, CCL）是图像处理中的一个基本且核心的算法，旨在从二值图像中识别并标记出所有的连通区域。在图像分析中，连通区域通常是指图像中由具有相同像素值且相邻的像素组成的区域。该算法在许多计算机视觉任务中具有广泛应用，尤其是在物体识别、形态学分析以及特征提取等领域。其原理依赖于遍历图像并根据像素间的相似性和邻接性对图像进行区域划分。

在二值图像中，图像的像素通常只有两种取值：0 表示背景，1 表示前景。连通组件标记算法通过检测前景像素的连通性，来区分图像中不同的连通区域。根据定义，连通区域是由一组相邻的前景像素组成的集合，这些像素之间通过边界相连。连通性通常分为两类：4 连通和 8 连通，其中 4 连通要求相邻的像素仅通过上下或左右方向连接，而 8 连通则包括对角线方向的相邻像素。

算法的实现流程主要分为两个部分：标签分配和标签合并。在开始时，图像中每个像素都未被标记，算法通过逐像素扫描图像并检查其邻域像素是否已经标记，从而确定每个像素所属的连通区域。若当前像素与其邻域中的某个像素相连通，且它们的值相同，则它们应当属于同一连通区域，算法会将它们标记为相同的标签。如果当前像素没有相邻的前景像素，那么该像素被视为一个

新的连通区域的起始点，算法会为其分配一个新的标签。为了确保每个连通区域的像素在扫描过程中具有统一的标签，标签分配通过不断检查邻域来实现。通过这种方式，算法逐渐识别出所有的连通区域并为其分配不同的标签。

然而，在标签分配过程中，某些连通区域可能会被分配多个标签。为了解决这个问题，算法需要进行标签合并操作。标签合并的目的是将那些属于同一个连通区域的像素的标签合并为一个统一的标签。标签合并通常采用并查集结构来实现，它能够高效地合并标签并保持图像中连通区域的正确性。在算法的执行过程中，合并操作会不断检查已标记区域之间的标签是否冲突，若有冲突，则通过合并操作使得所有属于同一区域的像素最终共享同一个标签。

在具体实现中，算法首先需要对图像中的每个像素进行逐行逐列扫描，对于每个前景像素，它会检查其左侧和上侧的像素是否已经标记。如果左侧或上侧的像素已经标记且其标签一致，则当前像素被赋予相同的标签；否则，它会为当前像素分配一个新的标签。随着扫描的进行，图像中的前景像素逐渐被标记并归类到不同的连通区域中。在标签合并的阶段，算法会根据标签间的相邻关系合并不同的标签，确保同一连通区域内的像素最终具有一致的标签。

连通组件标记算法的核心计算过程可以用以下公式来表示：设图像的尺寸为 $h \times w$ ，其中 h 为图像的高度， w 为图像的宽度。对于图像中的每个像素 $I(y, x)$ ，如果该像素是前景像素（即 $I(y, x)=1$ ），则其标签的赋值规则为：

$$\text{labels}(y, x) = \begin{cases} \min(\text{labels}(y', x'), \forall (y', x') \in N(y, x)), & \text{if } N(y, x) \neq \emptyset \\ C + 1, & \text{otherwise} \end{cases}$$

其中， $N(y, x)$ 表示当前像素 $I(y, x)$ 的邻域像素集合，根据 4 连通或 8 连通的定义，邻域集合 $N(y, x)$ 可能包含上下、左右以及对角线方向的像素。若当前像素与邻域中的某个像素具有相同的标签，则当前像素的标签为邻域像素标签的最小值；如果没有邻域像素已被标记，则当前像素为新连通区域的起始点，赋予一个新的标签 $C+1$ ，其中 C 是当前已有标签的数量。

在标签合并阶段，若发现两个像素属于同一个连通区域，但由于扫描顺序的原因它们分别被赋予了不同的标签，则需要将这两个标签进行合并。标签合并的操作可以表示为：

$$\text{labels}(y, x) \leftarrow \text{labels}(y, x) \text{ for } \text{labels}(y, x) = \min(L_1, L_2)$$

这里， L_1 和 L_2 分别表示需要合并的两个标签。标签合并操作确保同一连通区域的所有像素都具有一致的标签，从而为后续的图像处理和区域分析提供准确的信息。



图【1】 连通组件示例（左为形态学分割后图像，右为连通域随机上色后图像）

二、形态学骨架化算法

形态学骨架化算法（Morphological Skeletonization）是图像处理中的一种基本操作，旨在将二值图像中的对象或形状转化为细线状的骨架，同时保留原始图像的拓扑结构和形状特征。该算法通常用于提取图像中物体的中心结构，广泛应用于形态学分析、目标识别、图像分割等领域。骨架化的目标是通过一系列形态学变换，逐渐将图像的前景区域精细化为单像素宽度的骨架，而不丢失物体的形态信息。

形态学骨架化的核心思想是通过迭代地对图像进行腐蚀操作，在不断减少图像前景的面积的过程中，逐步保持图像的形态特征。骨架化的过程可以视为对图像进行多次腐蚀（erosion）操作，并在每次腐蚀后进行图像更新，直到图像的前景区域只剩下一个由单像素宽度的线条组成的骨架。这个操作遵循一个重要的原则：在进行腐蚀时，尽量保留物体的拓扑结构，避免物体的主要特征部分被过度删除。

在数学上，形态学操作通常基于集合理论，利用结构元素对图像进行形态学变换。对于二值图像，腐蚀操作是通过结构元素在图像上滑动，检查每个位置的邻域是否满足特定条件，从而决定图像中该位置的像素值。在骨架化过程中，腐蚀操作不断剔除图像中的边缘部分，直到只剩下物体的骨架。

骨架化操作的实现通常采用迭代腐蚀和差分的方法。首先选择一个结构元素，然后将该结构元素与图像进行腐蚀操作，剔除图像的边界部分。再计算原图像与腐蚀后的图像之间的差异，得到一个“差图像”，这个差图像包含了骨架部分，即图像中的前景部分被保留为线条。然后，将这个差图像与原图像结合，更新图像，重复该过程，直到图像收敛为一个单像素宽度的骨架。

在具体实现时，骨架化的核心计算公式可以表示为：对于一个给定的二值图像 I ，其形态学骨架化过程通过迭代腐蚀与差分操作实现。假设 S 为结构元素， I 为原始二值图像， $I^{(k)}$ 表示经过第 k 次腐蚀后的图像，则骨架化的核心操作可以用以下公式描述：

$$I^{(k)} = I \ominus S$$

其中， \ominus 表示腐蚀操作， S 是定义结构元素的集合，腐蚀操作的结果 $I^{(k)}$ 为图像 I 与结构元素 S 的交集，即只有当结构元素完全位于图像的前景区域时，腐蚀才会保留对应位置的像素。腐蚀操作的公式为：

$$(I \ominus S)(y, x) = \min\{I(y + s_y, x + s_x) \mid (s_y, s_x) \in S\}$$

其中， (y, x) 是图像 I 中某一像素的位置， (s_y, s_x) 是结构元素 S 中的某一元素，腐蚀操作的结果仅在结构元素的所有像素都在图像的前景区域内时才会保留原图像的像素。



图【1】 腐蚀操作效果示意图

接着差分操作用于提取骨架部分。假设 $I^{(k)}$ 是第 k 次腐蚀后的图像，骨架部分 $B^{(k)}$ 可以通过计算原图像和腐蚀后的图像之间的差异得到，公式为：

$$B^{(k)} = I - I^{(k)}$$

通过这一步骤，差分图像 $B^{(k)}$ 就包含了每次腐蚀过程中剔除掉的部分，即前景区域的边缘。然后，将这些差图像逐步合并，更新原图像 I ：

$$I^{(k+1)} = I^{(k)} - B^{(k)}$$

此过程不断进行，直到图像中的前景区域已经缩减为一个单像素宽度的骨架。

骨架化方法的关键点在于保持拓扑结构，即使图像中的物体不断缩小，骨架化的过程仍然能保持物体的连接性和形状特征，而不会导致物体的断裂或不连续。通过多次迭代腐蚀操作和差分操作，算法能够高效地提取出图像中物体的骨架。

三、高斯滤波

高斯滤波（Gaussian Filtering）是一种广泛应用于图像处理中的线性平滑滤波技术，主要用于图像的去噪、平滑、模糊以及特征提取等任务。该算法基于高斯函数（Gaussian function）的数学性质，将图像的每个像素值与其邻域像素的加权平均值结合，从而实现去除高频噪声并保留低频信息的效果。高斯滤波的核心优势在于其平滑效果相较于其他滤波方法更加自然，并且能够有效保留图像的边缘信息。

高斯滤波的数学原理基于高斯分布的卷积操作。假设输入图像为 $I(x, y)$ ，高斯滤波过程通过将一个二维高斯核与图像进行卷积，从而得到平滑后的图像 $I'(x, y)$ 。二维高斯核是通过高斯函数生成的，该函数定义为：

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right)$$

其中， σ 是高斯函数的标准差，控制着滤波的平滑程度。 (x, y) 是二维空间中的任意位置， $G(x, y, \sigma)$ 表示高斯函数在该位置的值。标准差 σ 越大，高斯核的宽度越大，滤波效果越强，能够对图像进行更为平滑的模糊处理；相反， σ 越小，滤波效果越弱，图像的细节和边缘信息能保留得更多。

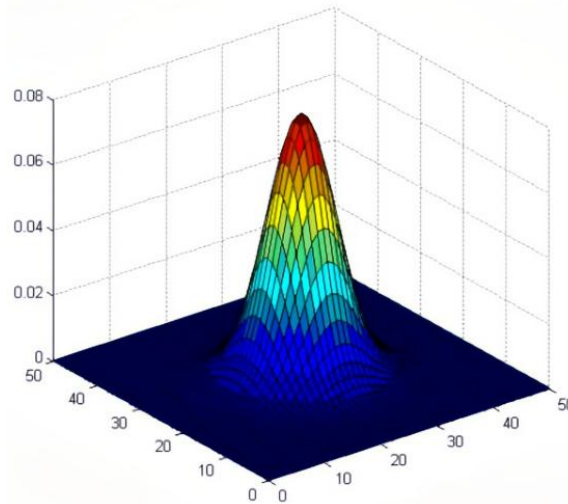


图 1 高斯函数的图像分布

高斯滤波的核心计算过程是将图像与高斯核进行卷积，卷积操作本质上是计算图像中的每个像素与其邻域像素的加权和。在二维空间中，卷积操作可以表示为：

$$I'(x,y) = \sum_{i=-k}^k \sum_{j=-k}^k I(x+i,y+j) \cdot G(i,j,\sigma)$$

其中， $I'(x,y)$ 是滤波后的图像像素值， $I(x+i,y+j)$ 是原图像中邻域像素的值， $G(i,j,\sigma)$ 是高斯核在位置 (i,j) 上的权重， k 是核的半径，通常是与标准差 σ 相关的一个常数。卷积操作对图像的每个像素都进行一次，并通过高斯核加权邻域像素值，从而得到平滑效果。

高斯核 $G(i,j,\sigma)$ 通过高斯函数计算得到，每个像素与周围像素的加权平均值由高斯核决定。由于高斯函数具有平滑性，它会赋予图像中较中心的像素较高的权重，而离中心较远的像素权重较低。因此，高斯滤波在图像的平滑处理上较为自然，能够有效去除噪声而不引入明显的模糊边界。

四、DBSCAN

DBSCAN (Density-Based Spatial Clustering of Applications with Noise) 是一种基于密度的聚类算法，它的核心思想是通过识别数据点在空间中的局部密度差异来进行聚类，能够有效处理噪声数据，并能发现任意形状的聚类。与传统的聚类算法如 K-means 不同，DBSCAN 不要求事先设定聚类的个数，而是通过数据点之间的密度关系自动识别聚类的数量和形状。这种密度驱动的聚类方法在面对复杂数据结构时展现出了很强的适应性，尤其对于形状不规则的聚类和含有噪声的情况，DBSCAN 能够提供较为精确的聚类结果。

DBSCAN 的工作原理依赖于两个关键的参数： ϵ 和 minPts 。其中， ϵ 表示一个数据点的邻域范围，即在该范围内，数据点与其他点的最大允许距离； minPts 则是指在 ϵ 邻域内，数据点必须包含至少 minPts 个点才能成为一个“核心点”。核心点是指周围密度足够高的点，而边界点是指位于核心点的 ϵ 邻域内但不满足成为核心点的条件的点，噪声点则是指既不属于任何核心点的邻域，也没有足够多的邻居点的点。基于这些定义，DBSCAN 通过检查每个点的

邻域来判断它是否可以归属于某个聚类，并根据点之间的密度关系进行聚类的扩展。

在实际应用中，DBSCAN 首先选择一个未访问的数据点，然后计算该点的 ϵ 邻域。如果邻域内的点数大于或等于 minPts ，则该点被标记为核心点，且该核心点的邻域内的所有点会被归为同一个聚类，并继续检查这些邻域点的密度可达性，递归地将密度相连的点归入该聚类中。若某个点的邻域内的点数少于 minPts ，则该点被标记为噪声。通过这种方式，DBSCAN 能够识别出数据中密度较高的区域，并将它们归为同一聚类，同时将噪声点与稀疏区域隔离开。

在数学上，DBSCAN 的核心计算依赖于 ϵ -邻域和密度可达性的概念。给定数据点 x 和 y ，其距离度量通常采用欧几里得距离，公式为：

$$\text{dist}(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

其中， $x = (x_1, x_2, \dots, x_n)$ 和 $y = (y_1, y_2, \dots, y_n)$ 分别为点 x 和 y 在 n 维空间中的坐标。根据该欧几里得距离，DBSCAN 定义了 ϵ -邻域，即数据点 x 的 ϵ -邻域是所有与 x 的距离小于或等于 ϵ 的点的集合，形式化表达为：

$$N_\epsilon(x) = \{y \mid \text{dist}(x, y) \leq \epsilon\}$$

此外，DBSCAN 还通过密度可达性来判断两个点是否可以被归为同一聚类。点 x 被称为是点 y 的密度可达点，若存在一个从 x 到 y 的路径，路径上的所有点的 ϵ -邻域内至少包含 minPts 个点。这一概念确保了即使两个点之间没有直接连接，如果它们分别与其他密度较高的点相连接，也可以归为同一聚类。DBSCAN 的聚类过程基于两个重要的操作：邻域计算和密度扩展。在遍历数据集时，算法首先检查每个点的 ϵ -邻域，确定其是否是核心点；若是核心点，则算法会尝试将其邻域内的密度可达点加入当前聚类，并递归扩展。整个过程确保了密度较高的点被聚集到一起，而那些稀疏的点则被标记为噪声。

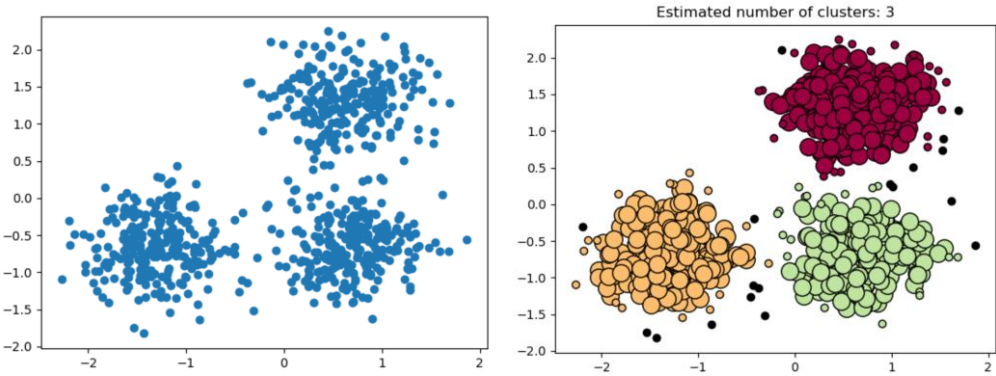


图 1 DBSCAN 聚类示例

五、非线性最小二乘拟合

非线性最小二乘拟合是一种通过最小化数据与模型之间的误差来拟合模型参数的数学方法。其核心思想是通过调整模型参数，使得模型输出与实际观测数据之间的误差达到最小。与线性最小二乘拟合不同，非线性最小二乘拟合的模型中包含非线性参数，因此无法像线性最小二乘那样通过直接的解析解来求解参数。为了应对这一挑战，非线性最小二乘拟合通常采用迭代优化的方法来逐步逼近最优解。

设定一组数据点 (x_i, y_i) ，目标是找到一个非线性模型 $f(x, \theta)$ ，使得该模型的输出尽可能接近真实观测值 y_i 。其中，模型的输出是由参数 $\theta = (\theta_1, \theta_2, \dots, \theta_p)$ 控制的。非线性最小二乘拟合的目标是最小化所有数据点的残差平方和，该残差为每个数据点的观测值与模型预测值之差。目标函数 $S(\theta)$ 被定义为：

$$S(\theta) = \sum_{i=1}^n [y_i - f(x_i, \theta)]^2$$

其中， y_i 是第 i 个数据点的观测值， $f(x_i, \theta)$ 是模型在该点的预测值， θ 是模型的参数， n 是数据点的数量。通过最小化这个目标函数，非线性最小二乘拟合能够找到一组参数 θ ，使得模型 $f(x, \theta)$ 对于所有数据点的拟合误差最小。

在实际计算中，非线性最小二乘拟合通常通过数值优化方法来求解，最常用的方法是基于迭代更新的技术，如高斯-牛顿法或勒文贝格-马夸特算法。高斯-牛顿法通过泰勒展开近似目标函数的二阶导数，并利用雅可比矩阵来进行参数更新。假设已经获得了当前参数 $\theta^{(k)}$ ，接下来计算当前残差向量：

$$r_i(\theta) = y_i - f(x_i, \theta)$$

然后计算雅可比矩阵 J ，它包含了残差对模型参数的偏导数。雅可比矩阵的每一元素 J_{ij} 是第 i 个数据点的残差 r_i 对第 j 个参数 θ_j 的偏导数：

$$J_{ij} = \frac{\partial f(x_i, \theta)}{\partial \theta_j}$$

通过计算雅可比矩阵 J 和当前残差 $r(\theta)$ ，可以通过以下的更新公式来得到新的参数 $\theta^{(k+1)}$ ：

$$\theta^{(k+1)} = \theta^{(k)} - [J^T J]^{-1} J^T r(\theta^{(k)})$$

其中， J^T 是雅可比矩阵的转置， $[J^T J]^{-1}$ 是 $J^T J$ 的逆矩阵， $r(\theta^{(k)})$ 是当前参数下的残差向量。通过这一更新过程，参数 θ 被逐步调整，直到目标函数 $S(\theta)$ 收敛到最小值。

在实际应用中，若高斯-牛顿法收敛较慢或不稳定，可以采用勒文贝格-马夸特算法，它结合了高斯-牛顿法和梯度下降法的优点，在每次迭代中通过引入一个调节因子 λ 来调整步长，使得算法在接近最优解时保持较好的收敛性。该算法的更新公式为：

$$\theta^{(k+1)} = \theta^{(k)} - [J^T J + \lambda I]^{-1} J^T r(\theta^{(k)})$$

其中， λ 是调节因子， I 是单位矩阵。通过适当地选择 λ ，该算法可以有效地避免步长过大而导致的不收敛问题。

六、均方根误差

均方根误差 (Root Mean Squared Error, RMSE) 是一种广泛应用于回归分析、机器学习和信号处理中的误差评估指标。其核心目的是衡量模型预测结果与实际观测数据之间的差异或误差大小。RMSE 通过计算每个预测值与真实值之间的差异，然后对这些差异进行平方、求和、求平均，并最终对结果取平方根，从而得到一个对误差幅度具有直观意义的量度。

均方根误差的计算公式为：

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

其中， y_i 是第 i 个真实数据点的观测值， \hat{y}_i 是第 i 个数据点的预测值， n 是数据点的总数。RMSE 的含义可以通过公式逐步解释：首先对于每个数据点 i ，计算真实值 y_i 和预测值 \hat{y}_i 之间的误差，即 $y_i - \hat{y}_i$ 。这个误差通常称为“残差”，它表示模型在该数据点的预测误差。接着对所有数据点的误差进行平方操作，目的是消除正负误差的相互抵消，并将较大的误差赋予更大的权重。通过平方操作后，再将所有平方误差进行求和，得到所有数据点的总误差平方和。最后，对平方和取平均，得到均方误差（Mean Squared Error, MSE）。为了恢复误差的原始尺度，通常对均方误差进行平方根操作，得到均方根误差。

RMSE 的数值表示了模型预测误差的大小，其单位与数据本身的单位相同。较小的 RMSE 值表示模型的预测结果与真实数据之间的差异较小，模型拟合较好；而较大的 RMSE 值则意味着模型预测存在较大偏差，模型拟合较差。由于 RMSE 是通过平方残差求得，因此它对较大误差较为敏感，使得在面对数据中的异常值时，RMSE 的值会被这些异常值放大。因此，RMSE 常用于模型比较和模型选择，尤其是在不同模型之间进行误差量化和评估时具有直观意义。

七、决定系数

决定系数（ R^2 ）是回归分析中用于评估模型拟合质量的一个重要指标。它表示了自变量对因变量变异的解释程度，衡量的是模型所能解释的因变量总变异的比例。决定系数的取值范围为 0 到 1 之间，值越接近 1，表示模型对数据的拟合效果越好；值越接近 0，则表明模型无法有效解释数据的变化。

决定系数的计算是基于回归模型的预测值和实际观测值之间的差异。设有一组观测数据点 (x_i, y_i) ，其中 x_i 是自变量的取值， y_i 是因变量的观测值， \hat{y}_i 是由回归模型预测得到的因变量值。决定系数的核心计算公式为：

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

其中， y_i 是第 i 个数据点的真实值， \hat{y}_i 是回归模型对第 i 个数据点的预测值， \bar{y} 是因变量 y_i 的均值， n 是数据点的数量。分子部分 $\sum_{i=1}^n (y_i - \hat{y}_i)^2$ 表示模型未能解释的误差平方和，通常称为“残差平方和”，即模型预测与实际观测值之间的差异的平方和。分母部分 $\sum_{i=1}^n (y_i - \bar{y})^2$ 是因变量的总变异平方和，称为“总平方和”，表示数据点相对于其均值 \bar{y} 的总变动程度。

决定系数的计算通过将残差平方和与总平方和进行比值计算，反映了回归模型能在多大程度上解释因变量的变化。如果模型能够很好地拟合数据，残差

平方和将较小，从而 R^2 值接近 1；相反，若模型无法有效解释数据变动，则残差平方和接近总平方和，导致 R^2 值接近 0。

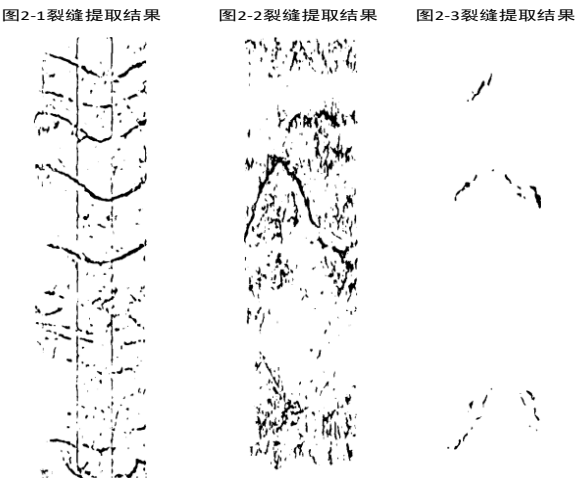
5.3 模型求解

本文建立的裂隙自动聚类与表征模型的求解流程图如图【】所示：



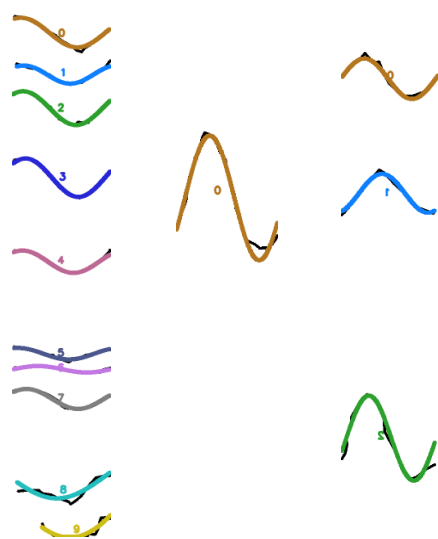
图【】 裂隙自动聚类与表征模型求解流程图

基于问题 1 的裂缝提取算法，本文对附件 2 中的数据进行了识别，如图【】所示：



图【】 基于问题 1 算法对图 2-1 至 2-3 的裂缝提取结果

再对提取出来的结果依次进行连通分析、骨架提取、高斯滤波去除噪声数据、聚类、曲线拟合，并计算曲线拟合的 RMSE 和 R^2 ，将拟合后的曲线进行绘制结果如图【】所示：



对所有图像数据给出聚类表征结果如表【】所示：

图像编号	裂隙编号	振幅 R(mm)	周期 P (mm)	相位 β (rad)	中心线位置 C (mm)
2-1	1	-27.185	107.347	-2.057	925.874
2-1	2	-16.635	88.272	-2.363	848.664

2-1	3	30.986	100.741	0.921	787.070
2-1	4	35.312	100.736	0.782	659.354
2-1	5	20.494	97.221	0.941	505.060
2-1	6	10.560	97.103	1.228	336.213
2-1	7	5.984	92.857	-0.085	307.625
2-1	8	18.271	98.226	0.700	253.141
2-1	9	39.393	179.773	-3.056	109.998
2-1	10	-1000.000	848.793	7.449	998.831
2-2	1	110.694	91.396	-0.557	598.785
2-3	1	37.097	89.561	-3.479	841.025
2-3	2	35.587	86.630	3.832	630.744
2-3	3	78.021	84.766	2.825	181.201
2-4	1	113.428	91.189	-0.794	469.670
2-5	1	-1000.000	253.090	9.971	-445.889
2-5	2	1000.000	180.319	1.725	1289.978
2-5	3	22.350	24.782	-17.817	49.847
2-6	1	20.996	26.627	-4.806	786.451
2-6	2	-15.590	39.791	-3.915	367.135
2-6	3	44.639	23.275	-15.488	319.941
2-6	4	67.529	52.112	-3.168	271.600
2-6	5	1000.000	365.532	9.429	1159.520
2-7	1	305.246	180.693	-2.641	1083.405
2-7	2	338.074	125.270	-1.359	95.435
2-7	3	1000.000	320.197	3.773	1033.089
2-8	1	1000.000	164.937	2.953	1629.764
2-8	2	160.077	89.630	0.562	235.988
2-9	1	98.249	139.899	3.344	1047.736
2-9	2	74.114	100.457	2.847	839.686
2-9	3	65.324	105.214	-3.333	700.257
2-9	4	56.355	114.250	-3.222	632.219
2-9	5	102.341	105.362	-2.935	305.660
2-9	6	1000.000	394.205	-1.967	1156.593
2-9	7	87.441	103.115	3.409	86.452
2-10	1	57.366	132.335	3.778	816.712
2-10	2	31.472	84.580	1.888	580.821
2-10	3	61.679	123.641	-2.319	493.348
2-10	4	60.851	75.833	1.394	263.144

2-10	5	-51.413	116.308	6.485	229.747
------	---	---------	---------	-------	---------

表【1】 附件 2 聚类表征结果

通过进一步分析曲线拟合评估数据 RMSE 和 R^2 可以观察到，模型的拟合效果在不同的情况中表现出较为明显的差异。在整体数据中，RMSE 的值大多处于一个较低的范围，这意味着模型预测与实际观测数据之间的误差相对较小，拟合效果较为精确。同时 R^2 值普遍处于一个较高的范围，这表示模型能够解释数据中的大部分变异性。 R^2 值接近于 1 的情况下，说明模型成功地捕获了因变量与自变量之间的关系，大部分的变异性都被模型所解释，而仅有少部分的波动无法被模型解释。这进一步表明，在这些数据集上，模型的拟合能力较强，能够准确地预测大部分的观测值。

然而，也有少数样本表现出相对较大的 RMSE 值和较低的 R^2 值。这表明在这些特定情况下，模型在拟合过程中存在一定的偏差或不完美的拟合，可能是数据的特殊性所致。尽管如此，总体上来看，RMSE 和 R^2 的结果表明，模型在大部分情况下能够提供较为准确的拟合。

5.4 本章小结

本章主要讨论了钻孔成像展开图中裂隙的自动聚类与表征问题，并从多个技术角度综合运用不同的数学与计算方法来实现对裂隙形态的准确提取和归类。本文提出的裂隙拟合与聚类方法通过连通分析、骨架化、高斯滤波、DBSCAN 聚类和非线性最小二乘拟合等技术手段，有效地解决了钻孔成像展开图中裂隙的自动提取和表征问题。实验结果表明，该方法不仅能够精确提取裂隙的几何特征，还能较为准确地拟合裂隙的形态，且拟合效果得到了较好的验证。因此，本文提出的模型较好地满足了题目给出的拟合曲线的要求，为裂隙的自动化分析提供了有效的技术支持。