

中图分类号: TP391

单位代号: 10280

密 级: 公开

学 号: 20721629

# 上海大学



## 专业学位硕士学位论文

SHANGHAI UNIVERSITY  
MASTER'S DISSERTATION

题 目	科学文献中的曲线信息处理方法 研究
--------	----------------------

作 者 王迎港

学科专业 软件工程

导 师 陈侨川

完成日期 二〇二三年六月

姓 名：王迎港

学号：20721629

论文题目：科学文献中的曲线信息处理方法研究

## 上海大学

本论文经答辩委员会全体委员审查，确认  
符合上海大学硕士学位论文质量要求。

答辩委员会签名：

主席：

委员：

导 师：

答辩日期：

姓 名：王迎港

学号：20721629

论文题目：科学文献中的曲线信息处理方法研究

## 原创性声明

本人声明：所呈交的论文是本人在导师指导下进行的研究工作。除了文中特别加以标注和致谢的地方外，论文中不包含其他人已发表或撰写过的研究成果。参与同一工作的其他同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

签名：\_\_\_\_\_ 日期：\_\_\_\_\_

## 本论文使用授权说明

本人完全了解上海大学有关保留、使用学位论文的规定。即：学校有权保留论文及送交论文复印件，允许论文被查阅和借阅；学校可以公布论文的全部或部分内容。

(保密的论文在解密后应遵守此规定)

签名：\_\_\_\_\_ 导师签名：\_\_\_\_\_ 日期：\_\_\_\_\_

# 上海大学工程硕士学位论文

## 科学文献中的曲线信息处理方法研究

作者: 王迎港

导师: 陈侨川

学科专业: 软件工程

计算机工程与科学学院

上海大学

2023 年 6 月



A Dissertation Submitted to Shanghai University for the  
Degree of Master in Engineering

# **Research on Curve Information Processing Method in Scientific Literature**

Candidate: Yinggang Wang

Supervisor: Qiaochuan Chen

Major: Software Engineer

**School of Computer Engineering and Science**

**Shanghai University**

**March, 2023**

## 摘 要

随着数据可视化技术的不断发展，曲线坐标数据图已经成为了一种常见的数据可视化方式。曲线坐标数据图通常用于展示多个变量之间的关系，例如在生物学、医学、金融、工程、科研等领域中，曲线坐标数据图被广泛应用于研究和分析数据中，呈现数据和趋势的关系。其中曲线数据提取有利于不同层面的发展，包括更容易的数据分析、创建各种可视化的能力、改进的决策、提高的准确性和时间效率。提取的数据可用于计算统计值、创建不同类型的图表、做出更好的业务决策、提高数据准确性并节省时间。通过利用曲线数据提取，科研人员可以对其数据获得有价值的见解，做出明智的决策，并更有效地分配时间等。

曲线坐标数据图包含复杂的数值信息，但从中提取数值却是一项繁琐和耗时的任务，需要人工进行提取和分析。因此，曲线坐标数据图的应用范围和效率受到限制。目前，主流的方法包括传统的基于图像处理的方法和基于学习机制的神经网络方法。

传统的曲线坐标图数值提取方法包括基于特征提取和模式匹配的方法，以及基于模型拟合和优化的方法。这些方法虽然在某些情况下具有较高的准确率，但也存在着一系列的局限性。例如，基于特征提取和模式匹配的方法需要手动设计特征，难以适应不同的数据特征。而基于模型拟合和优化的方法则需要选择合适的模型和优化方法，不同的数据可能需要不同的模型和优化方法。因此，这些方法的应用范围受到了一定的限制。相比之下，基于学习机制的神经网络方法可以自动从数据中提取特征和模式，并具有更强的适应性和泛化能力。它们通过学习大量数据，自动学习如何提取关键特征并进行数值提取和分析。因此，神经网络方法在曲线坐标图数值提取领域具有较深的挖掘价值。为了提高曲线提取任务的处理能力，本文针对曲线的特点，结合多种深度学习技术，提出了两种曲线数据提取的检测方法。

首先，针对当前科学文献中普遍存在的曲线坐标图曲线数据提取困难与耗时费力问题，本文探索从边缘检测角度实现曲线信息自动提取。曲线图像大多数绘制种类多样、密度较大、连续性较强，导致不同的方法提取的曲线信息不准确。本文尝试在密集网络模型基础上提出端到端的曲线提取模型以用来解决曲线检测方法产生的

曲线线条杂乱且模糊等问题，提高曲线信息提取准确率。通过增加自适应空洞卷积模块扩大感受野，添加逐层反向细化路径模块，每一层的中间输出引入后续细化模块，精心设置损失函数参数，进而优化了网络性能。此外还自主构建了曲线检测的数据集，对改进后的模型进行训练，进一步提高了网络对曲线图像中曲线边缘信息提取的能力。定性评价结果进一步体现了本文方法相比于其他方法的优越性。

其次，针对上述方法堆叠过多的密集卷积模块而造成通道特征信息损失，模型可训练参数较大、训练推理耗时较长等问题，本文提出了基于双重高效通道注意力机制的曲线提取网络结构。该方法以 Vgg 作为主干特征提取网络，采用双重高效通道注意力机制结构表征不同通道的特征的权重，更好地学习图像中通道之间的关系，提高特征表示的能力；然后嵌入阶段特征融合模块，减少特征丢失并增加特征表达能力，通过将不同阶段的低分辨率和高分辨率特征进行融合，使得模型可以更好地理解图像中的语义信息，同等情况下模型参数大幅降低。在曲线数据集上进行训练与测试，结果表明，该方法提取的曲线结构轮廓清晰、层次分明、定位准确，用更少的参数解决了曲线周边模糊等问题，提升了曲线提取精度。

最后，曲线数据提取算法的复杂性和实现难度往往限制了其应用范围。因此，开发易用的数据提取软件可以使曲线数据提取更加普及和易于使用，进一步推动其在实践中的应用。本文着眼于曲线数据提取的实用价值，开发桌面端数据提取软件，推动算法落地，可以让更多的人受益于曲线数据提取的实用价值，促进其在各个领域的广泛应用。

本文提出的两种深度学习方法可以实现对曲线坐标图像的端到端训练，并在曲线数据集上进行了详细的对比实验和消融实验。利用这两种方法，本文能够准确提取多种曲线结构，为曲线提取方法提供了新的思路，为相关领域的结构分析提供了重要的见解。

**关键词：**曲线提取，深度学习，数据挖掘，密集网络，注意力机制，边缘检测

## ABSTRACT

With the continuous development of data visualization technology, curve coordinate data graphs have become a common way of visualizing data. Curve coordinate data graphs are usually used to display the relationship between multiple variables, and are widely used in fields such as biology, medicine, finance, engineering, and scientific research to study and analyze data relationships and trends. Extracting curve data is beneficial for different levels of development including easier data analysis, ability to create various visualizations, improved decision-making accuracy and time efficiency. The extracted data can be used to calculate statistical values, create different types of charts, make better business decisions, improve data accuracy and save time.

However extracting numerical information from complex curve coordinate graphs is a tedious task that requires manual extraction and analysis. Therefore the application scope and efficiency of curve coordinate graph is limited. Currently mainstream methods include traditional image processing-based methods and neural network-based learning mechanisms.

Traditional numerical extraction methods for curve coordinate graphs include feature extraction- based pattern matching method as well as model fitting- based optimization method. Although these methods have high accuracy in some cases they also have limitations such as requiring manual design features which makes it difficult to adapt to different types of features or selecting appropriate models or optimization methods which may vary depending on the type of dataset being analyzed. In contrast neural network-based learning mechanism can automatically extract features from datasets with stronger adaptability and generalization capabilities by learning how to extract key features through large amounts of training datasets thus making them more valuable in extracting numerical information from curves.

To improve the processing capability for extracting curves this paper proposes two detection approaches combining multiple deep learning techniques targeting specific characteristics unique to curves:

Firstly this paper explores automatic extraction using edge detection approach due difficulties associated with current conventional line detection approaches resulting in inaccurate results when applied on dense multi-type lines found within most curve images leading researchers spending significant amount time manually correcting errors caused by these inaccuracies . This paper attempts an end-to-end solution using dense networks to improve accuracy by increasing the receptive field through adaptive dilated convolution modules, adding reverse refinement path modules and introducing intermediate output to subsequent refining module while carefully setting loss function parameters. Additionally a curve detection dataset was created to train the improved model further improving its ability to extract edge information from curves as demonstrated in qualitative evaluation results.

Secondly this paper proposes a double efficient channel attention mechanism-based curve extraction network structure targeting issues associated with too many stacked dense convolutional layers leading to loss of channel feature information, large number of trainable parameters and long training inference times. This method uses Vgg as the main feature extraction network and adopts a dual efficient channel attention mechanism structure that characterizes the weight of different channels' features better learning relationships between channels thus improving feature representation capabilities. Then an embedded phase fusion module is used reducing feature losses while increasing expression capability by fusing low-resolution and high-resolution features from different stages allowing for better understanding of semantic information within images while significantly reducing model parameters. Training on curve datasets showed that this method extracted clear, well-defined contours with accurate positioning using fewer parameters solving problems such as blurring around curves thereby enhancing precision.

Finally, complexity and implementation difficulties often limit the application scope of numerical data extraction algorithms therefore developing easy-to-use data extraction software can make it more accessible promoting wider use across various fields. This paper focuses on practical value for extracting curve data developing desktop-based data extraction software promoting algorithm deployment making it easier for people to benefit from practical value offered by extracting numerical information from curves facilitating widespread

use across various fields.

The two deep learning methods proposed in this paper enable end-to-end training of curve coordinate images, and detailed comparison experiments and ablation experiments are conducted on curve datasets. Using these two methods, this paper is able to accurately extract a variety of curve structures, which provides new ideas for curve extraction methods and important insights for structure analysis in related fields.

**Keywords:** curve extraction; deep learning; data mining; dense networks; attention mechanisms; edge detection

## 目 录

第一章 绪论 .....	1
1.1 课题来源 .....	1
1.2 课题研究的背景和意义 .....	1
1.3 研究目的和意义 .....	2
1.4 国内外研究现状 .....	4
1.4.1 基于传统图像处理的边缘检测 .....	4
1.4.2 基于机器学习方法的边缘检测 .....	5
1.4.3 基于深度学习方法的边缘检测 .....	6
1.5 本文研究内容与主要工作 .....	8
第二章 相关理论及技术概述 .....	10
2.1 引言 .....	10
2.2 神经网络技术 .....	10
2.2.1 卷积层 .....	10
2.2.2 池化层 .....	11
2.2.3 全连接层 .....	12
2.2.4 激活函数 .....	12
2.3 深度边缘检测网络 .....	13
2.3.1 VGG 卷积网络 .....	13
2.3.2 残差网络 .....	15
2.3.3 密集型网络 .....	16
2.4 上采样方法 .....	17
2.4.1 转置卷积 .....	18
2.4.2 亚像素卷积 .....	19
2.5 注意力机制 .....	21
2.6 特征融合 .....	22
2.7 曲线提取评价指标 .....	23

2.8 本章小结 .....	24
<b>第三章 基于密集结构的曲线提取网络 .....</b>	<b>25</b>
3.1 方法概述 .....	25
3.2 基于密集网络的曲线信息提取模型 .....	26
3.2.1 密集网络结构 .....	27
3.2.2 反向细化结构 .....	27
3.2.3 尺度增强模块 .....	29
3.3 实验与讨论 .....	31
3.3.1 实验环境和评价标准 .....	31
3.3.2 数据集介绍 .....	31
3.3.3 损失函数 .....	32
3.3.4 现有方法和 DCEN 对比 .....	32
3.3.5 与 DexiNed 的对比 .....	38
3.3.6 消融实验 .....	39
3.4 本章小结 .....	44
<b>第四章 基于双重高效通道注意力机制的曲线信息提取网络 .....</b>	<b>45</b>
4.1 方法概述 .....	45
4.1.1 双重高效通道注意力机制 .....	46
4.1.2 阶段特征融合 .....	49
4.1.3 改进的自适应权重损失 .....	51
4.2 实验与讨论 .....	52
4.2.1 实验环境和评价标准 .....	52
4.2.2 与现有方法检测结果对比 .....	52
4.2.3 消融实验 .....	57
4.2.4 和 DCEN 对比实验 .....	59
4.3 曲线数据提取应用设计 .....	60
4.3.1 开发环境和相关工具介绍 .....	61
4.3.2 软件设计流程 .....	61
4.3.3 应用软件介绍与效果展示 .....	65



4.4 本章小节 .....	67
第五章 总结与展望 .....	69
5.1 总结 .....	69
5.2 展望 .....	70
参考文献 .....	72
作者在攻读硕士学位期间发表的论文与研究成果 .....	78
作者在攻读硕士学位期间所作的项目 .....	79
致 谢 .....	80

# 第一章 绪论

## 1.1 课题来源

本课题得到国家重点研发计划（编号 2018YFB0704400, 2018YFB0704402, 2020YFB0704503），国家自然科学基金（面上，编号 52273228），上海市自然科学基金项目（编号:20ZR1419000），之江实验室科研攻关项目资助（编号: 2021PE0AC02），上海市“科技创新行动计划”启明星项目（扬帆专项）（编号：23YF1412900）。

## 1.2 课题研究的背景和意义

随着数据可视化技术的不断发展，曲线坐标数据图已经成为了一种常见的数据可视化方式。曲线坐标数据图通常用于展示多个变量之间的关系，例如在生物学、医学、金融等领域中，曲线坐标数据图被广泛应用于研究和分析数据。然而，曲线坐标数据图中包含的数值信息往往比较复杂，需要通过人工的方式进行提取和分析，这限制了曲线坐标数据图的应用范围和效率。

提取坐标曲线图的数据可以帮助我们更深入地分析和理解图表所代表的关系。通过提取数据，我们可以更精确地计算数据点的值，并对它们进行统计分析、建模或预测。具体而言，提取坐标曲线图的数据可以有诸多用途，例如：

（1）数据分析和统计：通过提取坐标曲线图的数据，我们可以对数据点进行统计分析，例如计算均值、标准差、方差、相关系数等等。这些统计指标可以帮助我们更深入地理解数据之间的关系和趋势。

（2）数据建模和预测：提取坐标曲线图的数据可以用于构建数学模型，以便预测未来数据点的值。例如，我们可以使用回归分析等方法，根据已有的数据点，建立一个预测模型，以预测未来数据点的值。

（3）数据可视化：提取坐标曲线图的数据可以帮助我们重新绘制或修改图表，以更好地呈现数据。例如，我们可以使用其他数据可视化工具，如散点图、柱状图、饼图等，来更好地呈现数据的特征和趋势。

因此，提取坐标曲线图的数据具有很重要的意义和价值，可以帮助我们更好地了解数据的分布和趋势，发现数据中的异常值和潜在问题，从而更好地理解 and 利用

数据。在实际应用中，提取坐标曲线图的数据是数据分析过程中不可或缺的一步，可以帮助我们更好地掌握数据的全局情况，做出更加明智的决策。

### 1.3 研究目的和意义

坐标曲线图是一种数据可视化方式，通常用于呈现数学函数或其他类型的数据关系。它是由两个坐标轴（通常是  $x$  轴和  $y$  轴）组成的平面，其中每个坐标都代表一个数据点，数据点通过在坐标轴上绘制的线条连接起来，形成一条曲线。

在坐标曲线图中， $x$  轴和  $y$  轴通常代表不同的数据变量而曲线则代表这些变量之间的关系。例如，一个简单的坐标曲线图可能显示温度随时间的变化。在这种情况下，时间可能沿着  $x$  轴，而温度可能沿着  $y$  轴。将这些数据点绘制在坐标轴上，并将它们连接起来，就可以形成一条显示温度随时间的变化趋势的曲线。坐标曲线图可视化广泛应用于许多不同类型的数据，包括科学实验数据、金融数据、社会经济数据等等。曲线坐标图数据提取是指从曲线坐标图中提取数据的过程，曲线数据图也是基本的数字图像，可以采取图像处理的方法根据计算坐标轴的刻度值和数据点的位置来提取数据并导出为相应格式。通过将数字图像的曲线图中曲线数据提取出来，转换为计算机可读的数字形式有利于进行后续的分析 and 处理，例如，可以利用这些数据点进行曲线拟合、数据统计、模型建立等操作。

目前，曲线提取方法主要有提取软件手工提取和基于传统图像处理方法的自动提取。基于手工提取的软件有 GraphDigitizer, WebPlotDigitizer, Matlab 等工具，其中工作流程是手工标注出曲线坐标  $x$  轴与  $y$  轴并输入相应的下标值，同时绘制出曲线图样，导出所需要的数据。这种检测方法，虽然能够详细的检测出曲线，但是需要投入大量的人力，效率较低。一般的，作者根据实验数据的不同绘制多样的曲线坐标图，这进一步加剧了曲线的复杂度与提取的繁琐程度。与此同时基于传统图像处理的曲线提取方法会采取曲线坐标图的不同特征进行提取，如颜色、连续性、边缘等信息，但往往由于曲线坐标图种类的繁杂以及形式的多样，不同类别的曲线图可能就需要设计一种特需的提取算法，进而给快速提取带来了极大的困难。

曲线通常由有语义性质的边缘组成，所以曲线的边缘检测是曲线信息提取的基础。在坐标曲线图中，边缘检测方法可以用来提取曲线的轮廓信息，并加以分析转换，即可得到具体曲线数据。在面临曲线边缘提取时，也有许多不同的方法，如基于

数学和统计的传统图像处理方法与基于学习机制的深度学习方法，然而传统图像处理面临曲线数据提取时存在着以下不足。

(1) 非常依赖于前期预处理：在传统图像处理方法中，曲线数据提取通常需要先对图像进行一些预处理，例如二值化、去噪等，这些预处理过程可能会影响曲线数据的准确性，需要经过多次试验和调整才能得到满意的结果。

(2) 对曲线形状限制较大：传统图像处理方法中通常使用的曲线提取算法，如边缘检测、霍夫变换等，对曲线的形状有一定的限制，只适用于某些特定形状的曲线，例如直线、圆等，对于其他非常规形状的曲线提取效果不佳。

(3) 难以处理复杂背景下的曲线：在复杂背景下，曲线的边缘可能会与其他物体的边缘相互交错，传统图像处理方法通常只关注图像本身，而忽略了深度图像的上下文信息，没有考虑每一层的视觉信息，在应用中需要不断调整参数，使检测结果不尽如人意。

因此，对于复杂的曲线数据提取任务，传统的图像处理方法存在很多限制和不足，需要采用更为高级的方法和技术。深度学习方法能够强有力的实现相关目标的边缘检测，特别是卷积神经网络（CNN）。通过训练 CNN，可以学习到边缘的特征，并且可以识别出不同类型的曲线边缘，例如强边缘和弱边缘。相较于传统的边缘检测方法，深度学习模型具备以下特点：首先，深度学习模型具有更强的表征能力和学习能力，因此其边缘检测准确率更高；其次，深度学习模型通过充分的训练和数据增强等手段，具有更好的泛化能力，可以在不同的图像中准确地检测出边缘；此外，深度学习模型能够自动学习特征，不需要手动设置参数和阈值，从而减少了人工干预的工作量，自动化程度更高；最后，深度学习模型可以并行计算，可以处理大规模的图像数据，从而提高了处理效率。

本文基于深度学习方法提出了端到端的曲线提取模型，通过将曲线坐标图输入到深度学习模型中，自动识别曲线的特征并提取出曲线的像素坐标点。其优势在于可以自动学习数据的特征和模式，而无需手动设计特征提取器，这种自动化的特征学习能力可以适应各种不同类型和复杂度的曲线图像数据，并取得较好的性能表现。

## 1.4 国内外研究现状

边缘检测是计算机视觉领域中非常重要的一项任务，其目的是在图像中识别出物体和场景的边缘。边缘是指在图像中明显的亮度、颜色或纹理变化处的位置，例如物体的轮廓、地平线和阴影。边缘检测是计算机视觉任务的基础，它在图像处理、目标跟踪、图像分割、计算机视觉和机器人领域中都有广泛的应用。近几十年来，研究人员提出了许多用于检测边缘的方法。这些方法可以广泛地分为三类：基于图像处理的传统方法、机器学习方法和深度学习方法。

### 1.4.1 基于传统图像处理的边缘检测

现有的基于传统图像处理的边缘检测方法包括基于差分 [1-2]、统计学 [3-8]。

**基于差分的方法。**在图像边缘轮廓检测任务中，基于差分的边缘检测方法是数字图像处理中非常常用的一种方法，通过计算图像中像素值的差分或导数来确定图像中边缘的位置和方向。目前，基于差分的边缘检测方法主要包括一阶导数算子、二阶导数算子以及其它高阶导数算子。

早期的单尺度方法（例如，Sobel 方法 [9]）采用简单的核或运算符来获取边缘强度图。这种方法虽然高效，但这些早期方法对噪声敏感 [2]，因此一些单尺度方法采用具有可变大小的核。在文献 [10] 中，拉普拉斯高斯核被用于产生边缘响应。通过设置不同的内核大小，可以将拉普拉斯高斯内核应用于细线或宽线上。Canny 检测器 [11] 是一种更先进的方法，并且已经成为边缘检测中最受欢迎的方法，该方法使用各向同性高斯内核的一阶导数来检测被加性白噪声污染的孤立步进边缘，并被认为是理论上最佳操作符，而在 Canny 方法中选择比例仍然是一个问题。为了克服这个问题，许多边缘检测器 [12-15] 已经开发出来以改善 Canny 方法。

基于差分的检测方法通常根据图像灰度值的一阶或二阶导数计算得到不同的差分算子，之后对灰度图像应用差分算子，得到边缘响应图像，最后通过阈值处理得到二值化边缘图。基于差分方法的传统图像边缘检测具有计算速度快、实现简单等优点，但也存在一些缺陷，如对图像噪声敏感、边缘检测结果不够精细等。

**基于统计的方法。**传统的基于差分的边缘检测方法难以抑制不需要的纹理，同时也无法准确识别由纹理定义的边缘结构。此外，在使用基于差分的方法进行边缘检测时，关键参数需要经过精心的选择。为了克服这些限制，近年来提出了一种统

计边缘检测方法,它能够分析每个像素周围的局部图像模式。**Rakesh** 等人 [4] 使用统计原则来执行边缘阈值处理过程。该方法通过对梯度向量的统计变异性进行本地标准化,来确定每个像素的阈值。然后,使用基于梯度向量的标准化统计数据来确定每个像素是否为边缘像素。这种方法具有抑制非目标区域的能力,在图像包含复杂的纹理背景时能够提高边缘检测的性能。

基于统计方法的图像处理技术,它通过统计图像中像素值的分布来检测边缘。常用的统计方法有基于概率密度函数的方法、基于灰度共生矩阵的方法、基于小波变换的方法等。基于统计方法具有较好的抗噪性能、边缘检测结果更加精细等优点,但也存在一些缺陷,如计算量较大、实现复杂等。

#### 1.4.2 基于机器学习方法的边缘检测

基于机器学习方法的边缘检测是指使用机器学习算法来自动检测图像中的边缘。传统的边缘检测方法往往依赖于手动设置阈值或运用图像处理算法,而基于机器学习的方法则可以自动从大量的图像样本中学习出边缘特征,并对新的图像进行边缘检测。常见的基于机器学习的边缘检测方法包括基于随机森林的方法、基于支持向量机的方法、基于决策树的方法等。这些方法通常需要使用已标注好的图像数据进行训练,训练完成后就可以用于新的图像中的边缘检测。

随着纹理特征描述技术的发展 [16], **Martin** 等人 [17] 利用图像局部线索(即亮度梯度、纹理梯度和颜色梯度信息)提出了一种基于概率边界 (**Pb**) 的边缘检测方法。在文献 [17] 中,利用逻辑回归结合图像局部线索(即亮度梯度、纹理梯度和颜色梯度信息)的  $v$  距离,并学习一个判别模型来从图像中检测边缘轮廓。**Ren**[18] 扩展了 **Pb** 方法 [17] 并提出了一种多尺度基于概率边界 (**MsPb**) 的边缘检测方法,该方法利用本地边界线索,即对比、定位和相对对比,并结合大尺寸和小尺寸的优势以实现噪声鲁棒性和保留细节。**Arbelaez** 等人 [19] 扩展了 **Pb** 方法 [17] 并提出了一种全局基于概率边界 (**g-Pb**) 方法,该方法使用多尺度和光谱聚类获取本地提示并从图像中检测边缘轮廓。**Payet** 等人 [20] 提出了基于纹理的 **Pb(tPb)**,它使用随机放置的窗口对纹理变化的平均值进行估计,以获得判别性边缘响应。在文献 [21] 中指出,上述基于手动设计的特征开发的边缘检测方法将边缘检测技术提升到更高的抽象层次,这些方法在表示用于语义重要边缘检测的高级信息方面存在较大的局限性,缺乏对象级

知识，仍然低于人类感知水平。

### 1.4.3 基于深度学习方法的边缘检测

深度学习技术在计算机视觉和图像识别领域取得了巨大的进展，尤其是在图像边缘检测方面。研究人员采用不同的网络结构，使用卷积层和全连接层生成局部边缘特征图，然后采用非极大值抑制等后处理技术来生成最终的边缘图像。在这些方法中，VGG 模型是应用最广泛的。根据数据集是否具有直观的标签特征，现有的基于深度学习的边缘检测方法可以分为监督学习（即带标签特征）图像边缘检测方法 [22-23] 和无监督学习（即不带标签特征）图像边缘检测方法 [24-25]。

**基于监督学习的图像边缘检测。**当前的图像处理任务主要通过监督学习来完成。在这种情况下，每个样本都有相应的标签，并且使用深度神经网络（例如 ResNet[26]）来学习与每个标签对应的特征，以执行不同的视觉任务。基于监督学习的图像边缘检测示意图如图1.1所示。

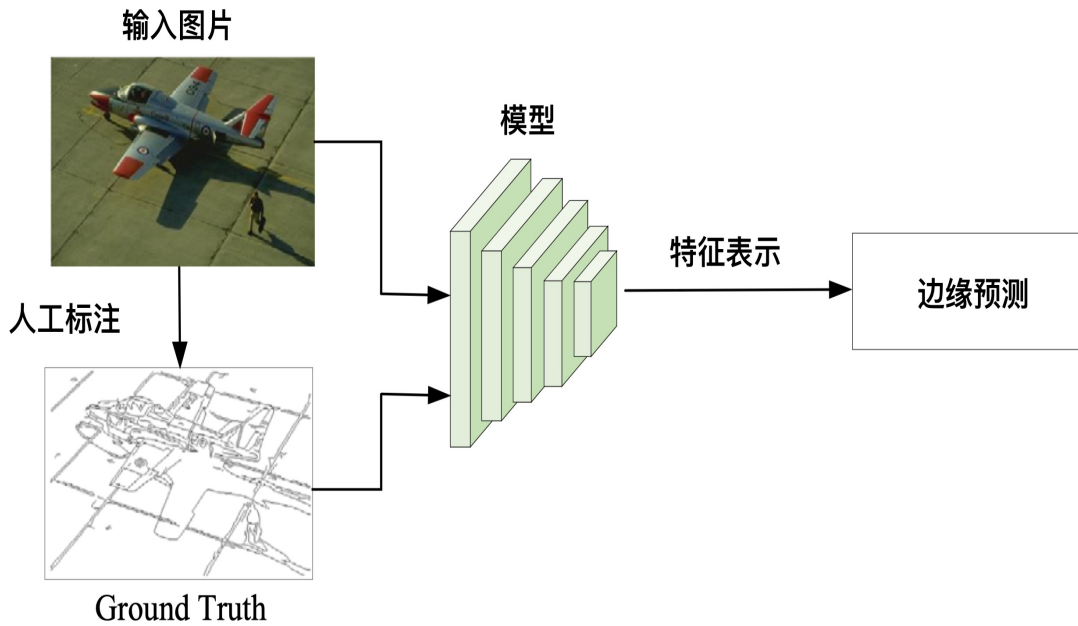


图 1.1 基于监督学习的图像边缘检测示意图

文献 [27] 中提出了一种扩展版本的概率增强分类方法，该方法利用大量具有多个尺度的特征来学习从图像中检测边缘和物体边界的判别模型。基于蚁群算法 [28]，Rahebi 等人 [29] 使用带有监督学习和动量的人工神经网络来改进医学图像边缘检测。Lim 等人 [30] 将人类生成轮廓的小块聚类成 sketch token 类别 [31]，然后使用随

机森林分类器有效地检测边缘轮廓。Dollar 等人 [32] 将图像边缘检测问题表述为基于随机决策森林 [33] 并预测图像局部分割掩码, 从而实现了图像边缘的检测。该方法具有较高的准确性和鲁棒性, 并且可以在较短的时间内处理大量的图像数据。

Ganin 等人 [34] 设计了神经网络最近邻场以提取边缘, 这是第一个基于深度卷积神经网络 (CNNs) 的边缘检测方法, 其中利用最近邻技术匹配从图像 patch 中提取的特征以便从图像中检测边界轮廓。Xie 等人 [35] 提出了全方位嵌入边缘检测 (Holistically-nested edge detection, HED) 方法, 充分利用卷积神经网络和深监督网络进行图像到图像预测。此外, 在深度监督网络 [36] 的指导下, HED 方法还使用了多尺度和多层次的策略并在卷积层的一侧插入一个侧输出层, 通过真实值的映射对其进行深度监督, 将输出与不同层连接起来获得最终预测结果, 有效地解决边缘和物体边界检测中具有挑战性的模糊问题。He 等人 [37] 在 VGG16[36] 基础上用 ImageNet[38] 数据集预训练骨干网络, 并设计了一个双向级联模块 (BDCN) 来提取边缘。BDCN 中专门设计了一个尺度增强模块, 它具有利用扩张卷积生成多尺度特征并在极不同的尺度上检测边缘轮廓的能力。Liu 等人 [39] 为了有效地利用分层 CNN 特征对像素进行分类到边缘或非边缘集合, 提出 RCF 网络 (Rich Convolutional Features for Edges)。在图像到图像方式下使用所有卷积层的 CNN 特征执行逐像素预测以从图像中提取边缘轮廓。他认为不同卷积层之间的信息可以互补, 该方法使用了每一层之间的信息, 但只使用了池化层之前最后一个卷积层的信息作为输出。因此, RCF 方法利用每个卷积层的信息进行边缘检测。Wang 等人 [40] 提出了 CED (Deep Crisp Boundaries) 网络, 一种新的细化体系结构, 借鉴于在密集图像标记方面的进展, 通过增加额外的非线性细化路径, 进一步降低了相邻像素内边缘响应之间的相关性。

基于监督学习的图像边缘检测算法具有准确性高、泛化能力强、可解释性好等优点, 但是也存在数据标注成本高、对噪声敏感、需要大量计算资源等缺点。在实际应用中, 需要根据具体的应用场景和需求来选择合适的算法, 并结合其他技术手段来提高边缘检测的准确性和鲁棒性。

**基于无监督学习的图像边缘检测。** Wang 等人 [41] 指出标记高质量数据集的成本随目标数量和识别难度呈指数增长。基于无监督学习的图像边缘检测不需要手动标记边缘特征进行训练, 它可以从图像本身的基础知识估计边缘轮廓, 如图1.2所示。



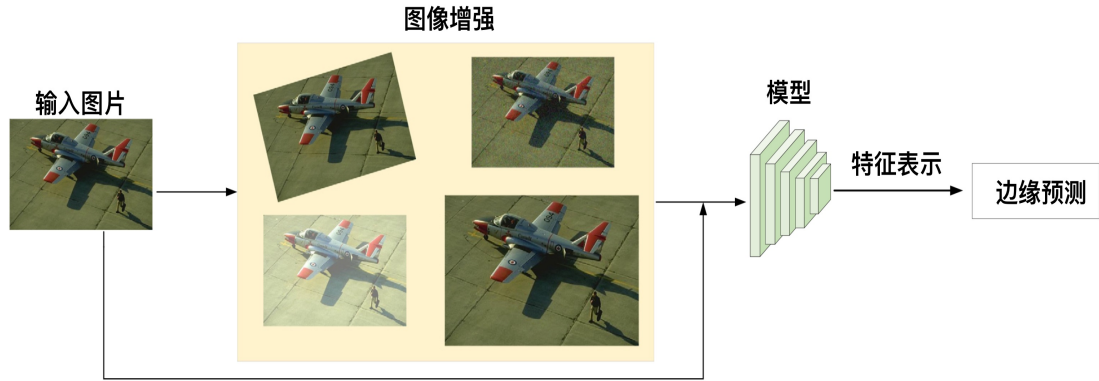


图 1.2 基于非监督学习的图像边缘检测示意图

Ren 等人 [42] 认为, 通过获取稀疏编码梯度 (SCG) [43] 可以显著提高边缘轮廓检测的准确性, 并且基于多尺度池化和幂变换对稀疏编码梯度进行分类具有有效学习如何测量局部对比度和找到边缘轮廓的能力, 然后使用线性支持向量机 (SVM) [44] 进行图像边缘分类。Isola 等人 [45] 指出, 基于同属于一个目标的像素之间显示出比不同目标更高的统计相关性这一基本原理, 采用点间互信息结构 [46] 从图像中提取边缘轮廓。Yang 等人 [21] 开发了一种完全卷积编码器-解码器网络来自图像检测物体轮廓, 该模型是基于全卷积网络和反池化的。该模型训练集为 PASCAL-VOC 数据集, 其具有很好地泛化到相同超类别下未见过对象类别的能力。Xia 等人 [47] 提出了一种用于无监督语义分割的编码器-解码器架构, 用于检测边缘轮廓, 在该架构中, 编码器将输入图像编码为多类别分割图, 然后对其进行解码以生成原始输入。通过软切损失约束分割图, 并通过条件随机场和区域层次合并进行后处理。

基于无监督学习的图像边缘检测算法具有多项优点, 例如不需要标注数据、适用范围广、可以学习到更加抽象和高级的特征, 准确性和鲁棒性较高等。不过, 这种方法也存在一些缺点, 例如模型训练难度大、准确性相对较低、解释性较差等。在实际应用中, 需要权衡这些优缺点, 选择最适合特定场景的边缘检测算法。

## 1.5 本文研究内容与主要工作

本文以作者攻读硕士学位期间承担的课题工作为基础, 针对曲线坐标图像数据提取特点, 研究提出了两种曲线提取算法, 收集并制作了曲线提取数据集, 通过该数据集上进行对比实验和消融实验以验证本研究方法的有效性, 最后为方便用户使用和理解提取结果开发桌面端提取软件。本文的组织安排结构如下:

第二章简要介绍本研究涉及的相关理论知识，首先介绍了边缘检测基本知识，然后介绍本研究使用技术的研究现状，主要是特征融合、注意力机制和多尺度学习模块，最后介绍曲线提取相关评价指标及其原理。

第三章紧紧围绕曲线坐标图像特征，在曲线图像边缘特征复杂、形状轮廓多变情况下，提出一种基于密集网络的曲线边缘提取算法，着重介绍了细化模块，尺度增强模块对于曲线提取的积极作用。之后通过对比实验和消融实验验证方法的有效性。

第四章针对第三章方法存在的缺陷，提出一种基于双重注意力机制的曲线提取算法。首先阐述网络框架设计原理，然后详细介绍了注意力机制的设计与实现，阶段特征融合模块的细节。最后进行了对比实验和消融实验，并对结果进行了说明和分析。最后完成了曲线提取桌面端软件系统设计与实现，将曲线提取算法开发封装，实现曲线提取真实落地可用。

第五章对全文回顾与展望，回顾了本文的主要研究成果、创新点以及不足之处，并提出对未来研究方向的展望。

## 第二章 相关理论及技术概述

### 2.1 引言

深度学习技术的快速发展为计算机视觉领域带来了重大进展，同时也为边缘检测和曲线图像处理等领域提供了新的发展机遇。针对曲线图像的特殊性质，本研究采用成熟的深度学习技术，设计了更为有效的方法，用于解决这些领域存在的困难问题。因此，本章节将回顾与本研究相关的技术理论基础和研究现状，为后续章节提供理论支持。

### 2.2 神经网络技术

卷积神经网络 (Convolutional Neural Network, CNN) 是一种深度学习模型，主要应用于图像处理和计算机视觉领域。CNN 通过学习输入数据的特征，将其转换成适合于分类、检测、分割等任务的形式。相较于传统的机器学习算法，CNN 可以自动提取并学习更加复杂的图像特征，使得模型的表现更加优秀。

CNN 主要由卷积层 (Convolutional Layer)、池化层 (Pooling Layer)、全连接层 (Fully Connected Layer) 等组成，其中卷积层和池化层是 CNN 最重要的组成部分。卷积层通过卷积操作来提取图像中的特征，池化层则通过降采样操作来减少特征图的大小，全连接层则用于将特征图转换为分类输出。

#### 2.2.1 卷积层

卷积层是 CNN 中最为重要的组成部分，它可以自动提取输入数据中的特征。在卷积层中，会定义一组卷积核 (Convolution Kernel) 或者滤波器 (Filter)，这些卷积核会与输入数据进行卷积操作，得到一组新的特征图 (Feature Map)。每一个卷积核都是一个小的矩阵，它会在输入数据的不同位置进行滑动，将局部信息进行提取。卷积操作如图2.1所示。通过将多个卷积核应用到输入数据上，可以得到多组不同的特征图，每一组都对输入数据的不同方面进行了提取。卷积操作的数学计算公式如2.1所示。

$$S(i, j) = (I * K)(i, j) = \sum_m \sum_n I(m, n) K(i - m, j - n) \quad (2.1)$$

其中， $I$  表示输入数据， $K$  表示卷积核， $S$  表示特征图。 $m$  和  $n$  分别表示卷积核中心点和输入数据中心点的相对位置。

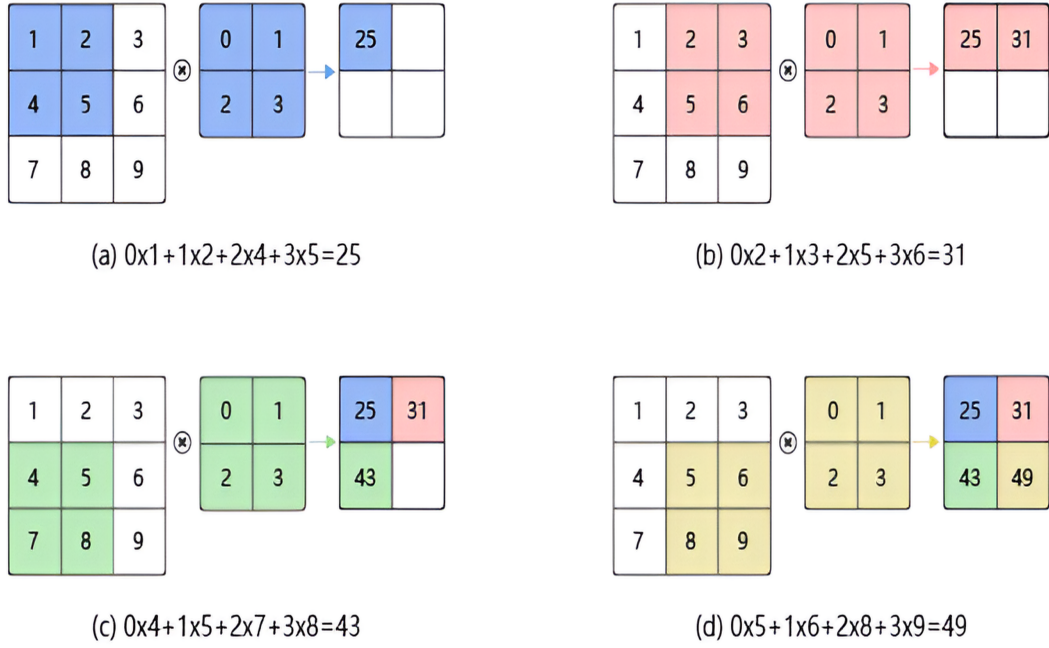


图 2.1 卷积操作示意图

### 2.2.2 池化层

池化层通常紧随卷积层之后，用于对特征图进行下采样。它的作用是减少特征图的尺寸，同时保留重要的特征信息，从而降低了网络的复杂度。池化层的运算方式与卷积层基本相同，其卷积核只取对应位置的最大值或平均值等（最大池化和平均池化），具体操作如图2.2。

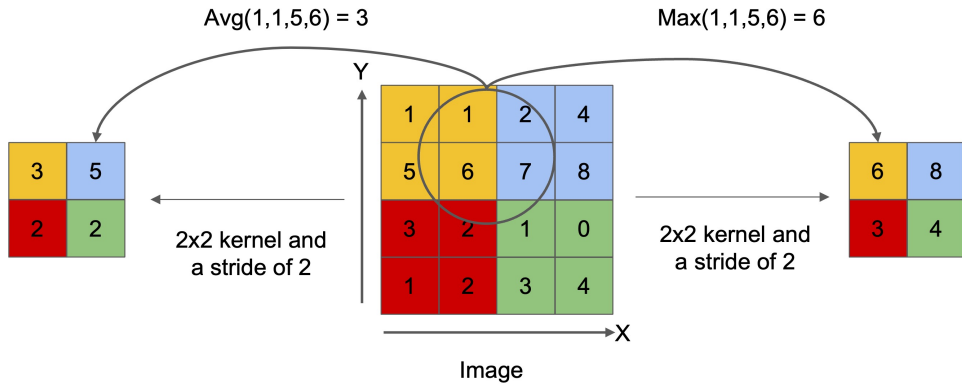


图 2.2 最池化原理示意图（最大值池化与平均值池化）。

### 2.2.3 全连接层

全连接层（Fully Connected Layer），也称为稠密层（Dense Layer），是神经网络中最常用的一种层，也是最基本的一种层。全连接层是指每一个神经元都和前一层的所有神经元相连，形成一个完全连接的网络结构。

全连接层通常用于深度学习的前馈神经网络（Feedforward Neural Network）中，它将输入数据通过矩阵乘法和加法运算进行线性变换，然后通过激活函数进行非线性变换，最终得到输出结果。在这个过程中，全连接层的参数包括权重和偏置，需要通过训练数据进行学习，以使神经网络能够对输入数据进行准确的预测。

### 2.2.4 激活函数

在神经网络中，每一层的输出都是由输入的线性变换和激活函数的非线性变换组成的，具体输出如公式2.2所示。

$$\mathbf{y} = \text{activation}(\mathbf{W}\mathbf{x} + \mathbf{b}) \quad (2.2)$$

其中， $\mathbf{x}$  是输入数据， $\mathbf{W}$  是权重矩阵， $\mathbf{b}$  是偏置向量， $\text{activation}$  是激活函数， $\mathbf{y}$  是输出向量。

在 AlexNet[48] 之前，神经网络通常使用 Sigmoid 作为激活函数。然而，Sigmoid 函数会饱和，导致梯度消失的问题，这严重限制了神经网络的深度，为了缓解这个问题，直到非饱和的激活函数 ReLU 被提出，在一定程度上解决了梯度消失的问题，使

得神经网络能够更深。常见的激活函数包括 Sigmoid 函数、ReLU 函数、LeakyReLU 函数等，如图2.3所示。

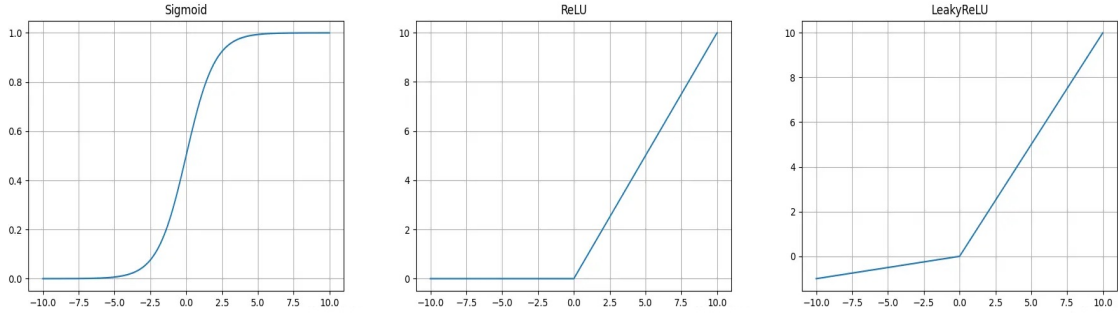


图 2.3 常见激活函数可视化图，从左至右分别是 Sigmoid，ReLU，LeakyReLU

## 2.3 深度边缘检测网络

### 2.3.1 VGG 卷积网络

VGG 网络 [49] 是一种深度卷积神经网络，旨在解决 ImageNet 数据集上的图像分类任务。其主要贡献是将深度卷积神经网络的深度推到了 16 或 19 层，VGG16 网络结构如图2.4所示，通过堆叠多个卷积层和池化层，以有效增加网络深度。为提高模型性能 VGG 网络采用小尺寸的卷积核 ( $3 \times 3$ ) 和“same”填充，并在每个卷积块之间加入了 Dropout 层，以减少过拟合。VGG 网络包含 5 个卷积块，每个卷积块包含 2-3 个卷积层和一个池化层，池化层采用  $2 \times 2$  的最大池化，其中每层细化结构如图2.5所示。在卷积层之后 VGG 网络使用了三个全连接层，前两个全连接层都包含 4096 个神经元而最后一个全连接层包含 1000 个神经元，对应于 ImageNet 数据集的 1000 个类别，最后使用了 softmax 激活函数，将网络的输出转换为概率分布。

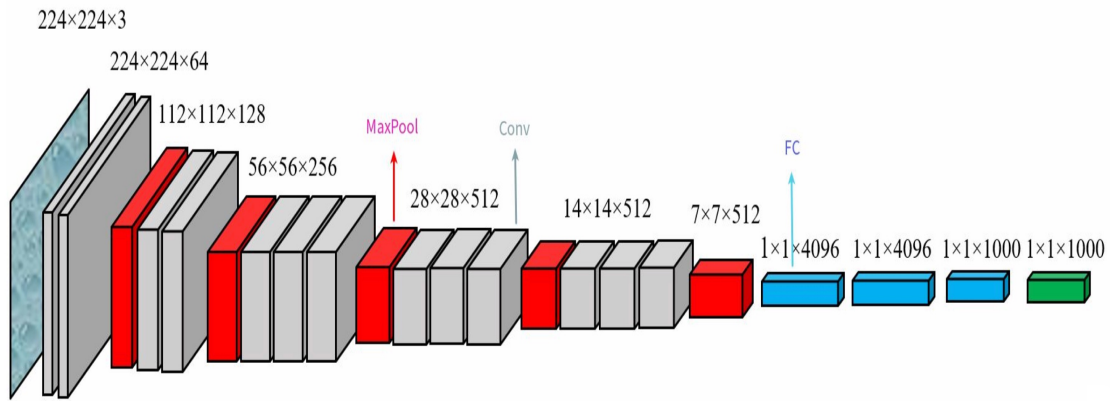


图 2.4 Vgg16 网络结构示意图

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 × 224 RGB image)					
conv3-64	conv3-64 <b>LRN</b>	conv3-64 <b>conv3-64</b>	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 <b>conv3-128</b>	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 <b>conv1-256</b>	conv3-256 conv3-256 <b>conv3-256</b>	conv3-256 conv3-256 conv3-256 <b>conv3-256</b>
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 <b>conv1-512</b>	conv3-512 conv3-512 <b>conv3-512</b>	conv3-512 conv3-512 conv3-512 <b>conv3-512</b>
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 <b>conv1-512</b>	conv3-512 conv3-512 <b>conv3-512</b>	conv3-512 conv3-512 conv3-512 <b>conv3-512</b>
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

图 2.5 VGG16 每层结构具体示意

VGG 网络的简单架构证明了深度卷积神经网络的卓越性能，正是由于其简单的结构，VGG 网络可以轻松地进行修改和扩展以适应不同的任务和数据集。例如，可以

通过在全连接层之前添加平均池化层来将其应用于图像边缘检测任务。此外，VGG网络的卷积块可以轻松地复制和堆叠，以构建更深的网络。

### 2.3.2 残差网络

残差网络（Residual Network, ResNet）[26] 是一种深度学习网络结构，由何凯明等人于 2015 年提出。它通过引入残差块（Residual Block）来解决深度神经网络中的梯度消失和梯度爆炸问题，从而实现更深的网络结构和更好的性能。在传统的深度神经网络中，每个层都是通过非线性变换来学习输入和输出之间的映射关系。然而，当网络变得非常深时，这种非线性变换可能会导致梯度消失或梯度爆炸，从而使得网络难以训练和优化。为了解决这个问题，ResNet 提出了残差块的概念，即在每个残差块中引入一个跨层连接（Shortcut Connection），使得网络可以直接学习输入和输出之间的残差（Residual），从而避免了梯度消失和梯度爆炸的问题。残差块的结构如 2.8 所示。

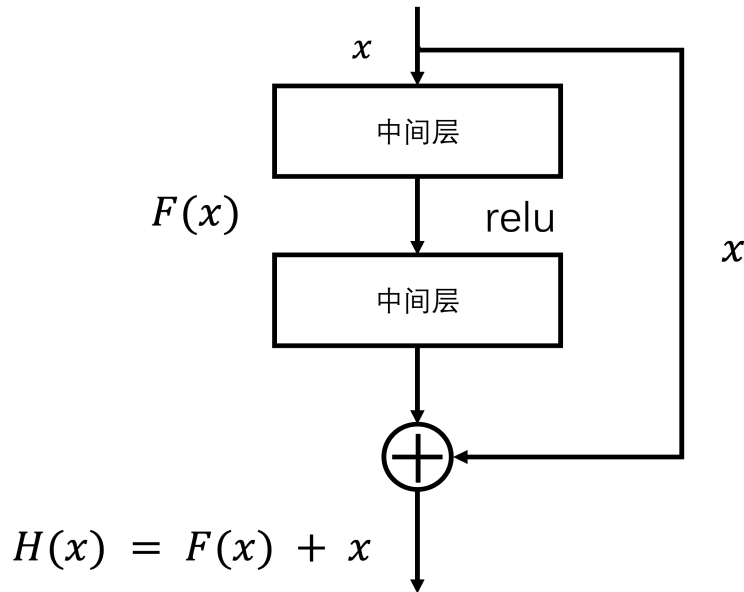


图 2.6 残差结构示意图

其中， $x$  表示输入， $F$  表示残差块中的非线性变换， $H$  表示跨层连接。残差块的输出为  $H(x) = F(x) + x$ ，即输入和残差之和。

在 ResNet 中，可以分为多个阶段（Stage），每个阶段包含多个残差块。多个堆叠在一起形成更深的网络结构。在每个阶段中，网络的分辨率会逐渐减小，通道数会逐渐增加，从而实现更好的特征提取和分类性能。



### 2.3.3 密集型网络

DenseNet 是由 Huang 等人 [50] 于 2017 年提出的。在此之前，卷积神经网络的设计都是基于 ResNet 或者 HighwayNet 等思想，这些连接方式都是将前面层的输出与后面层的输入相加，普遍采用跳层（shortcut）连接结构，以便更好地训练深度网络。然后作者实验通过随机弃用神经网络中的某些层，改变网络深度继续训练，观察到随着网络的加深，虽然图像的特征能够提取的更加丰富，但是深层次的网络中某些层对提高任务性能指标的贡献甚微，说明过于深度的神经网络中也存在着许多的冗余结构。例如在残差连接中，后面层的输入可能只是前面层的输出的一个微小变化，因此不能充分利用前面层的信息。为了解决这个问题，Huang 等人受到 GoogLeNet 和 FractalNet 的启发提出了 DenseNet，认为网络不必过于追求深层次的网络，在合理的层次下增加层内通道个数就能提升网络性能。DenseNet 通过将前面所有层的特征图连接起来，使得前面层的信息可以充分利用。如图 2.7，是该网络的一个基本模块。

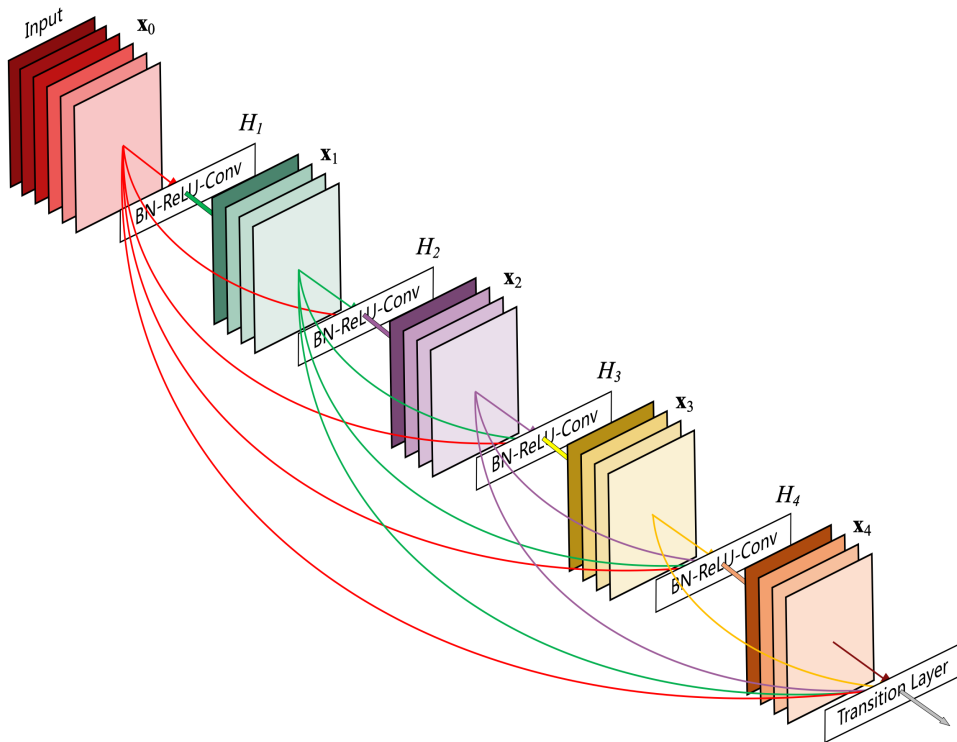


图 2.7 稠密网络基本模块

从图中可以看出，每个稠密网络的基本模块包含多个密集连接层。与普通卷积块相比，密集块（Dense Block）在每一层都与前面所有层的输出进行连接，而不是

只与前一层的输出连接。这种连接方式被称为密集连接 (Dense Connection)。每个密集连接层包含以下三个步骤：

(1) Batch Normalization：对输入特征图进行归一化。(2) ReLU 激活函数：使用 ReLU 激活函数来增加非线性性。(3) 卷积操作：使用卷积操作来提取特征。

在每个密集块中，每个密集连接层的输入都是前面所有层的输出的连接，假设第  $i$  层的输入是  $x_i$ ，输出是  $H(x_i)$ ，则第  $i + 1$  层的输入就是  $[x_0, x_1, \dots, x_i, H(x_i)]$  的连接，因此输出的特征图大小不变，但是通道数会增加。

这种连接方式可以使得前面层的信息在后面层中得到充分的利用，同时可以缓解梯度消失问题，因为前面层的梯度可以直接传递到后面层。同时后面层也可以通过前面层的特征来学习更加丰富的特征，这也是 DenseNet 相对于其他卷积神经网络的一个优势。相比之下，普通卷积块中的每一层只与前一层的输出连接，前面层的特征信息可能会被丢失，也会限制网络的深度。

完整的密集型网络结构则是由多个稠密基本块组成，如图2.8。但是不同子模块之间的大小不完全一致，所以该网络还采用了过渡块 (Transition Block) 来解决这个问题，减少特征图的大小和通道数，从而减少参数量和计算量。过渡块中通常包括一个  $1 \times 1$  的卷积层和一个平均池化层。这样可以避免特征图通道数过大，同时也可以控制特征图的大小，提高模型的计算效率。

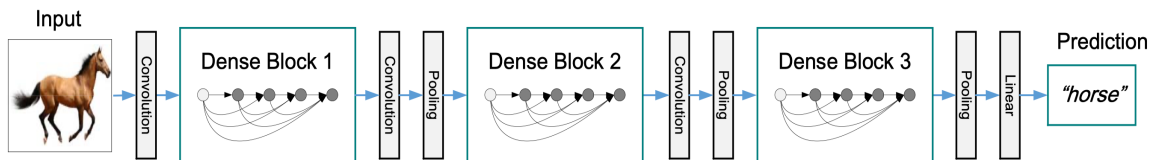


图 2.8 密集型网络整体结构

## 2.4 上采样方法

通常情况下，对图像进行卷积运算时，经过多层的卷积运算后输出图像的尺寸会变得很小，即图像被约减。对于某些特定的任务，比如在图像分割中，目标是预测一个分割掩模，其中输入图像中的每个像素被分类为几个类之一。为了生成具有与输入图像相同空间分辨率的分割掩模，神经网络的最后一层必须输出一个与输入图像大小相同的特征映射。然而，用于此任务的大多数卷积神经网络架构会随着它们通过网络进行进展而降低特征映射的空间分辨率。因此，必须在生成分割掩模之前

恢复特征映射的空间分辨率。恢复图像尺寸，实现图像由小分辨率到大分辨率映射的操作，叫做上采样（Upsample），如图2.9所示。

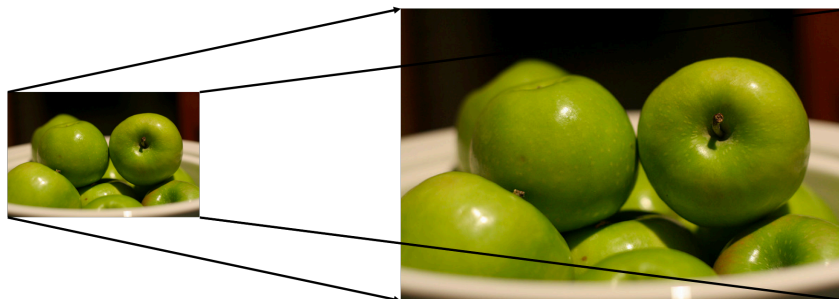


图 2.9 上采样示例

上采样有多种方式，常见的包括：转置卷积、亚像素卷积和双线性插值等。

#### 2.4.1 转置卷积

转置卷积 [51] 是一种卷积神经网络中常用的操作，主要目的是将输入信号的大小扩展到更大的空间维度，以便进行更深层次的特征提取和处理。

转置卷积的操作过程与正常的卷积操作有所不同。在正常的卷积操作中，卷积核在输入信号上滑动并计算出一组输出特征。而在转置卷积操作中，卷积核将被“翻转”并放置在输出信号的位置，然后在该位置上进行计算。这种操作可以将输出信号的大小扩大，因为在计算每个输出值时，可以将多个输入值映射到一个输出值。转置卷积的工作原理可以简单概括为：通过在输入特征图之间插入零值来进行升采样，然后使用与常规卷积相反的卷积核来进行卷积操作。更具体地说，转置卷积可以分为两个主要步骤：扩充和卷积。详细示意如图2.10所示。

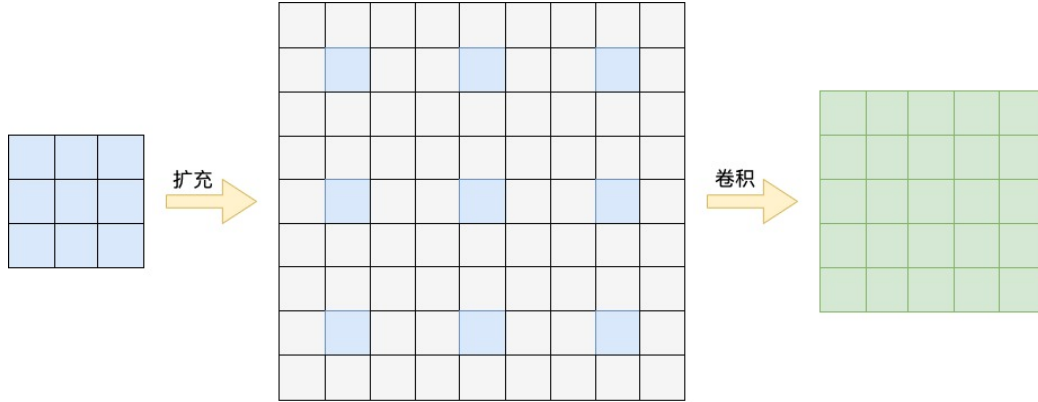


图 2.10 转置卷积示例

在扩充步骤中，转置卷积会在输入特征图之间插入零值，从而扩大特征图的空间尺寸。例如，如果要将一个  $2 \times 2$  的特征图升采样到  $4 \times 4$  的特征图，那么可以在每个原始像素之间插入一个零像素，得到一个  $3 \times 3$  的特征图，然后对该特征图进行卷积操作，以生成一个  $4 \times 4$  的特征图。在卷积步骤中，转置卷积会使用与常规卷积相反的卷积核来进行卷积操作。与常规卷积不同的是，转置卷积的卷积核会将输入特征图中的每个像素与输出特征图中多个像素相连。例如，在一个  $2 \times 2$  的转置卷积操作中，卷积核中的每个像素都会与输出特征图中 4 个不同的像素相连。

通过这样的升采样和卷积操作，转置卷积可以将低分辨率的特征图恢复到原始高分辨率，并且可以用于生成更高分辨率的图像或特征图。

## 2.4.2 亚像素卷积

亚像素卷积 (sub-pixel convolution) [52] 是一种深度学习中常用的图像处理技术，其主要应用是在图像超分辨率重建中。在传统的图像超分辨率算法中，通常采用双三次插值等方法将低分辨率图像放大到目标尺寸，但这种方法存在明显的缺点，如图像模糊、失真等。而亚像素卷积则是一种更为有效的方法，能够在保持图像细节的同时提高分辨率。在相机成像系统中，通过一种离散化处理方法对相机获取的图像数据进行了处理。由于光传感器的限制，图像受限于原始像素分辨率，换句话说，图像上的每个像素代表现实世界中一小块颜色区域。在数字图像中，像素与像素是连在一起的，而在微观世界中，两个物理像素之间存在着许多微小的像素。那些微小的像素称为亚像素。

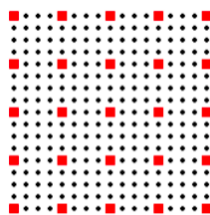


图 2.11 亚像素卷积示意图

如图2.11所示，每个被四个红色小方块包围的正方形区域是相机成像平面中的像素，黑色的点就是亚像素。可以根据相邻像素之间的插值来调整子像素的精度。这样就可以通过亚像素插值实现小方块到大方块的映射。

亚像素卷积的核心思想是将低分辨率图像转化为高分辨率图像的过程，看作是对高分辨率图像的逐像素预测，从而避免了图像放大过程中的信息丢失和失真。具体地，亚像素卷积操作会将输入的低分辨率图像进行卷积运算，得到的输出结果是一个更高分辨率的图像，通常采用反卷积（Deconvolution）或转置卷积（Transpose Convolution）实现。

在卷积运算过程中，亚像素卷积将输入的低分辨率图像进行多通道卷积，生成一个带有多个通道的高分辨率图像，然后将每个通道进行特定的重排列操作，以得到最终的高分辨率输出。这种操作类似于对输入图像进行像素重排列，通过亚像素插值的方法，从而使得输出图像的每个像素都由多个原始像素组成，实现图像的超分辨率重建。亚像素卷积具体过程如图2.12所示，将不同通道的像素分成子像素并重新排列再通过插值填充到对应的像素点上，每个特征上的像素点就相当于新的特征上的亚像素。

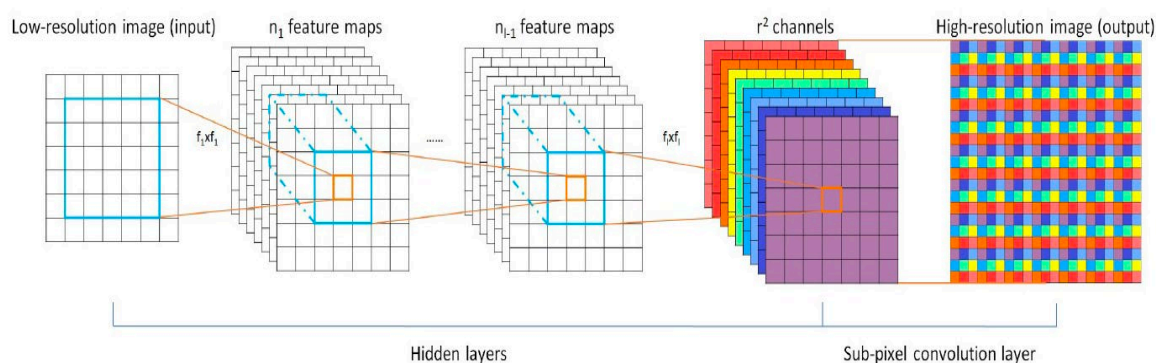


图 2.12 亚像素卷积过程

更简单的来说就是一幅图像经过多个卷积核形成中间特征图，这些特征图的像



素被认为是亚像素，经过这些特征图被融合形成特定尺寸的新的图，如图2.13就是一幅简易的亚像素卷积的过程图。

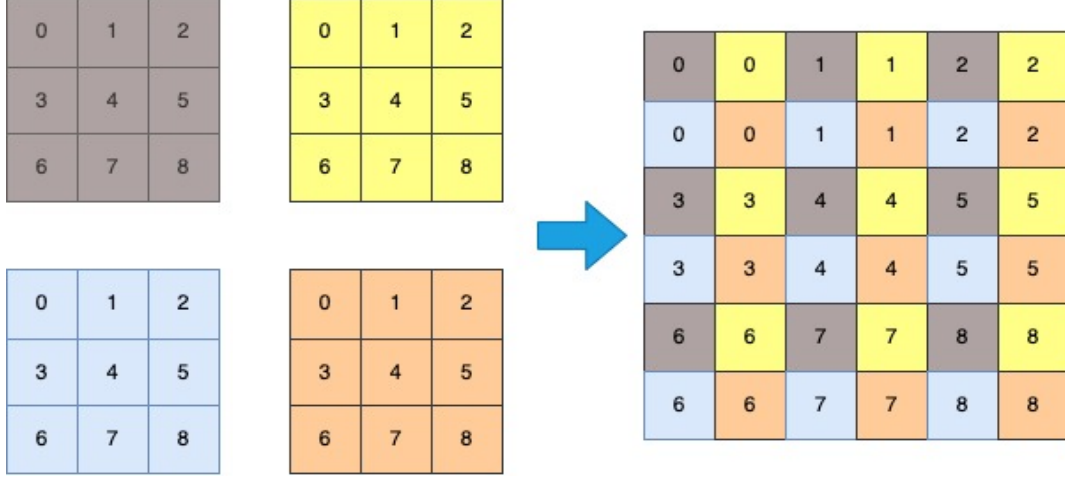


图 2.13 亚像素卷积简易图

## 2.5 注意力机制

注意力机制起源于认知学领域的研究，在处理信息时人类由于信息处理能力有限，往往会有选择性地关注一部分信息，而忽略其他可见信息，这极大的提升了信息处理效率。这种现象被称为注意力机制 [53]，其核心思想是在处理信息时，根据信息的重要性来分配有限的资源，在深度学习中注意力机制的动机也是模仿人类的这一特性。注意力机制本质是对特征权重的再分配，根据分配权重的所在维度，计算机视觉中的注意力机制可以分为空间注意力、通道注意力、时间注意力和分支注意力等 [54]，其中最常见的是空间注意力和通道注意力，在目标跟踪等具有时间维度的任务里则可以应用时间注意力。

2018 年，Hu 等人 [55] 提出的 SE-Net 将原特征图全局平均池化到  $1 \times 1 \times C$  的尺寸，将每个通道的特征抽象为一个点，之后使用多层感知机（Multi-Layer Perceptron, MLP）和 Sigmoid 函数计算各通道的注意力，是通道注意力的代表作。

同年，Wang 等人 [56] 把自然语言处理领域中的自注意力（Self-Attention）[57] 引入计算机视觉领域，提出了 Non-Local 模块。该模块用某空间位置的元素与其他位置元素的相似度作为该位置的注意力权重，其输入  $X = \{x_1, x_2, \dots, x_h\}$  与输出  $Y = \{y_1, y_2, \dots, y_h\}$  的尺寸相同，计算公式如2.3所示：

$$y_i = \frac{1}{C(x)} \sum_{j=1 \dots h} f(x_i, x_j) g(x_j) \quad (2.3)$$

其中  $i$  和  $j$  代表空间位置,  $x_i$  和  $y_i$  代表该位置所有通道的元素组成的向量,  $f$  为相似度计算函数,  $g$  为  $1 \times 1$  的卷积操作, 用于映射原特征,  $C(x)$  为归一化系数, 通常设定为特征图的元素数。

总的来说, 各种注意力的实现方式多种多样, 但核心思想不变, 即为重要的特征分配更高的权重。

## 2.6 特征融合

在深度学习中, 特征融合是一种重要的技术手段, 可以将不同来源的特征信息融合在一起, 以提高模型的性能和泛化能力。特征融合的作用是将不同层次、不同尺度和不同来源的特征信息组合成一个更加丰富和强大的特征表示, 从而提高模型对输入数据的理解和表达能力。

特征融合的原理可以通过深度学习模型中的多路输出来解释。多路输出是指将深度学习模型的不同层次或不同部分的特征输出进行融合。在这个过程中, 每个特征的输出都会通过一定的变换 (例如卷积、池化等操作) 进行下采样, 使得输出的特征具有不同的尺度和抽象程度。将这些不同尺度和不同抽象程度的特征进行融合, 可以使得模型更好地理解 and 表达输入数据。特征融合常用包括以下方式, 如图2.14所示:

(1) 横向特征融合 (Concatenation): 将不同来源的特征按照通道维度拼接在一起, 形成一个更加丰富的特征表示。这种方法适用于特征来源较少的情况, 可以增加特征的维度和多样性, 提高模型的表达能力。

(2) 纵向特征融合 (Summation): 将不同来源的特征按照通道维度相加, 形成一个更加强大的特征表示。这种方法适用于特征来源较多的情况, 可以增强特征之间的相关性, 提高模型的鲁棒性和泛化能力。

(3) 逐点相乘 (Element-wise Multiplication): 将不同来源的特征在每个空间位置逐点相乘, 从而使得特征的局部信息得到更好的保留。这种方法适用于特征之间存在较强的互补性的情况, 可以提高模型的表达能力和鲁棒性。

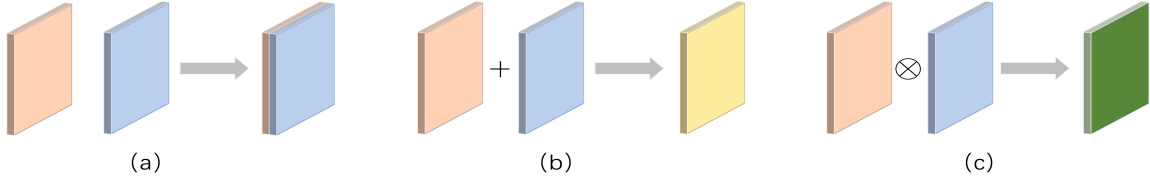


图 2.14 特征融合示意图。(a) 横向特征融合, (b) 纵向特征融合, (c) 逐点相乘。

特征融合是深度学习中重要的技术手段, 可以将不同来源的特征信息融合在一起, 从而提高模型的性能和泛化能力。特征融合的原理可以通过深度学习模型中的多路输出来解释, 通过将不同层次、不同尺度和不同来源的特征信息组合成一个更加丰富和强大的特征表示, 来提高模型的表达能力。

## 2.7 曲线提取评价指标

对于曲线提取的评价标准, 这里采用了 SSIM (Structural Similarity) [58]、FSIM [59] 指标以及边缘检测任务基础指标 ODS [60] (最佳数据集尺度), OIS [60] (最佳图像尺度) 和 AP (平均精度) 五个指标来定量评价曲线图像检测效果。

SSIM 是一种广泛使用的衡量图像之间相似程度的指标, 它是一种感知模型更加符合曲线提取的直观感受, 其具体计算公式如公式 2.4 所示:

$$SSIM_{x,y} = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (2.4)$$

其中,  $x$  表示原图;  $y$  表示曲线提取图像;  $\mu_x$  和  $\mu_y$  分别代表对应图像的均值;  $\sigma_{xy}$  表示  $x$  和  $y$  的协方差;  $\sigma_x$  和  $\sigma_y$  分别代表  $x$ ,  $y$  的标准差。SSIM 取值范围为  $[0, 1]$ , 其值越大表示输出图像和无失真图像的差距越小, 即图像质量越好。

FSIM (Feature Similarity Index Measure) 是一种基于特征相似性进行图像质量评价的方法。人类视觉系统通过感知图像的低层次特征来进行视觉判断, 而图像的局部结构可以很好地被相位一致性特征 (phase congruency, PC) 所刻画。由于 PC 具有相对不变性, 因此可以提取图像中稳定的特征。但有时图像的变化会影响观感, 此时需要使用梯度特征 (Gradient Magnitude, GM) 来弥补。FSIM 使用 PC 和 GM 两个特征互相补充, 以获得更准确的结果。FSIM 的计算公式较为繁琐复杂, 但 FSIM 值越大表示曲线边缘检测效果越好。



ODS/OIS 是指设置阈值  $\eta$  的不同方法，其中阈值  $\eta$  是指曲线边缘检测点概率的置信度，如若图像检测的概率结果大于阈值  $\eta$  则认为是正例，反之则认为是反例。ODS (Optimal Dataset Scale) 表示数据集的整体最优值，即选取一个固定的阈值  $\eta$  应用于所有图片，使得整个数据集上的 F-score 最大；OIS (Optimal Image Scale) 表示单幅图像的最佳阈值，即在每一张图片上均选取不同使得该图片 F-score 最大的阈值  $\eta$ ；AP (平均精度) 表示平均置信度。这三个指标的计算依赖于混淆矩阵的真实值判断。

混淆矩阵的四个参数如下：TP (True Positive)，FP (False Negative)，TN (True Negative)，FN (False Negative)。精确率  $P$ ，召回率  $R$  以及  $F_1$  定义分别如公式2.5所示。

$$\begin{cases} P = \frac{TP}{TP + FN} \\ R = \frac{TP}{TP + FP} \\ F_1 = \frac{2 * P * R}{P + R} \end{cases} \quad (2.5)$$

## 2.8 本章小结

本章首先介绍了边缘检测的相关理论，紧接着简要回顾了与本工作相关的技术及其原理。本章还介绍了几个经典的深度学习模块或者方法在边缘检测任务中的应用，例如上采样方法、注意力机制和多尺度的特征融合等。最后介绍了曲线提取中常见的评价指标。

## 第三章 基于密集结构的曲线提取网络

由于曲线坐标图像通常有复杂形态、断裂不连续、多曲线重叠和轮廓不一致等特点。本文提出了一种基于密集结构的曲线提取网络。具体而言，针对曲线图像的复杂特性，该方法的主干网络利用密集的卷积结构堆叠而成，以实现曲线复杂特征提取。通过尺度增强模块，在不另外增加卷积结构的情况下增强浅层网络的特征获取能力，降低训练难度。针对目前主流边缘检测方法所提取的结果较为粗糙与厚重的问题，引入细化模块，逐层细化并精炼主干网络多尺度侧方输出，使得最终检测结果的视觉效果更加优秀。

### 3.1 方法概述

曲线坐标图像数据提取主要面临三个难点。第一，曲线坐标图像通常具有多个分支和交叉点，断裂或不连续的部分，以及多条曲线叠加的特点，增加了曲线提取难度；第二，目前深度学习方法相较于传统方法有了显著提高，但曲线边缘的“正确性”和“清晰度”是曲线提取的两个最为重要的方面仍不能得到很好的平衡。目前的方法主要注重边缘的正确性，而忽略了清晰度，导致检测到的边缘比较粗糙。另一种方法虽然考虑了边缘的清晰度，但过分关注清晰度会导致边缘正确性的不理想；第三，曲线图像轮廓的不一致性也是一个严重的问题，例如主体轮廓和蜿蜒边缘细节的比例差距很大。

为了解决以上问题，本文提出了一种名为 DCEN (Dense Curves Extraction Network) 的密集型曲线提取网络，旨在解决曲线边缘检测中准确性和清晰度之间的平衡问题，网络结构如图3.1所示。该网络由五个密集卷积网络组成的前向特征提取网络和细化模块组成。其核心思想是通过主干密集网络提取多尺度、多维度的曲线特征信息，并设置了一个自上而下的细化路径，该路径使用有效的亚像素卷积逐步提高特征图的分辨率。逐层细化路径的增加为网络提供了额外的非线性特征，进一步降低了相邻像素内边缘响应之间的相关性，以便于提取出纤细的曲线边缘。同时增加自适应的空洞卷积模块，扩大神经元感受范围，缓解了整体参数的增加。因此，DCEN 在曲线坐标轴中曲线提取任务中提供了一种新的解决思路。具体而言，DCEN 网络

的前向特征提取网络能够提取输入图像的多层特征信息，并将其侧方输出同步输入到下方细化模块中进行反向细化操作。逐层的细化过程有助于更好地融合高层和底层信息，捕获曲线边缘的纹理特征，从而提高曲线边缘检测的准确性和清晰度。

在实验中，本文对 DCEN 进行了广泛的测试，并与现有的曲线边缘检测方法进行了比较。实验结果表明，DCEN 在各种曲线坐标轴中的曲线提取任务中表现出了很好的性能和鲁棒性。与此同时，本文还对 DCEN 的各个组成部分进行了详细的分析和讨论，下面将一一介绍。

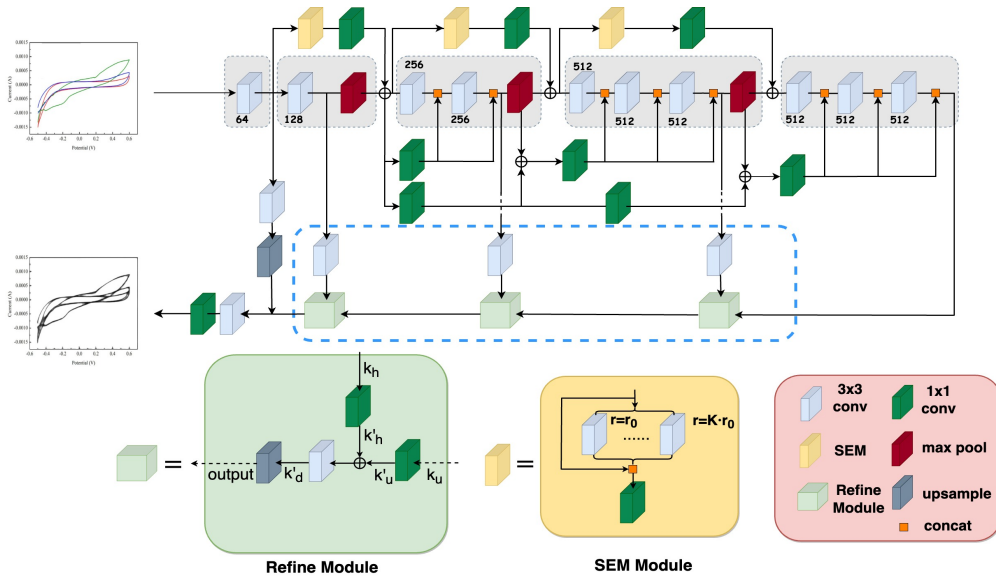


图 3.1 DCEN 网络结构

### 3.2 基于密集网络的曲线信息提取模型

DexiNed[61] 是一种基于 HED 和 Xception 网络的极密集型网络。密集型的网络结构能够获取丰富的卷积信息，与独特的上采样方法结合能够很好的保留高、低级别特征，所以 DexiNed 具有相对较好的整体检测效果，但其中仍存在一些问题。DexiNed 方法使用多个侧面输出，这使得无法完全融合中间特征图之间的差异，在浅层网络中获取大量纹理特征可能会对最终图像产生重大影响，致使最终曲线边缘检测清晰度降低。

DCEN 能够端到端的实现曲线提取任务，在于其采用和 DexiNed 类似的密集型主干特征提取网络，不需要像大多数深度学习模型一样从预训练模型进行权重初始化。通过实验发现，在深层次上，浅层计算出的边缘特征容易丢失，为此，DCEN 受

到 DexiNed 的启发采用两个平行跳跃连接来涵盖不同层次计算出的边缘信息。DCEN 主要由密集网络结构和细化模块组成，前者接收曲线图像作为输入并将特征映射馈送到后者进行处理。所有从细化模块得出的边缘特征图都被连接起来以馈送网络末端，并融合产生最终的曲线边缘。所有五个细化模块并不共享权重。

### 3.2.1 密集网络结构

密集网络结构包含五个编码模块，每个块由一组卷积层的小型子块组成。这些子块通过跳跃连接耦合在一起，形成一个密集的网络（如图3.1中浅灰色和浅蓝色所示）。每个块生成的特征映射被传输到单独的细化模块中，创建中间边缘映射。这些边缘映射组成一个学习滤波器，最终被融合成单个边缘映射。从左到右的浅蓝色矩形块包含 2 两个卷积层，其卷积核大小为  $3 \times 3$ 。每个卷积层紧随批量归一化和 ReLU 作为激活函数，并最终跟随具有  $3 \times 3$  内核和步长 2 的最大池化层，这部分称为前向路径特征提取。

当网络的深度逐渐增加的时候，执行大量的卷积操作会使得曲线边缘信息丢失的更快，所以仅有一个主连接是不够的。为了解决这个问题，本文采 DexiNed 中的跳跃连接结构，从图3.1中第三个浅灰色结构块开始，引入跳跃连接（图中绿色矩形块），每个子块的输出都与另一个跳跃连接进行平均操作。在最大值池化之后，这些跳跃连接的子块将已连接子块的输出平均到与下一层的跳跃连接中。同时，最大池化层的输出直接馈送给后续子块。

### 3.2.2 反向细化结构

受 CED 方法启发，本文提出自上而下反向曲线细化结构，从高层级特征逐步往低层级特征进行曲线边缘细化。通过自适应地选择预测分割标签的点，并使用亚像素卷积来消除上采样带来的模糊边缘映射，从而提高输出分辨率和边缘清晰度。通过细化结构的特征图可以很好地保留浅层中间特征并清晰地检测曲线边缘，有利于更好地融合高低层次的信息，捕捉更丰富的边缘纹理。具体结构如图3.1蓝色虚线矩形框所示。

在细化结构中，每个中间层的输出都被添加到后续的细化模块中，并且网络的反向路径进入相应的细化模块，重复多次使用细化模块以逐步增加特征映射的分辨率。其中细化的关键思想是使用中间特征映射聚合整个路径上边缘特征，该模块详

细结构如图3.2所示。每个模块将来自后向路径的自顶向下特征映射与当前层前向路径中的特征映射融合，并通过一个较小的因子（2x）进一步上采样，然后沿着路径传递下去。该模块有两个核心组件，即融合和上采样。

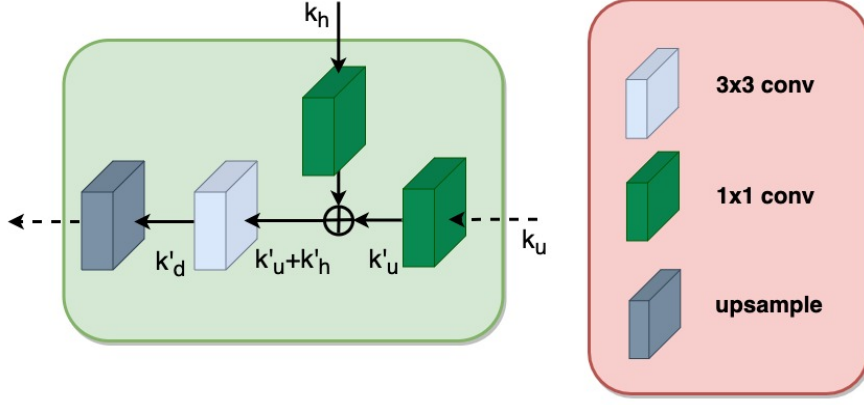


图 3.2 细化模块结构

**融合：**一种简单的策略是直接连接两个特征图。然而每个特征图都有大量通道，直接连接这些特征会极大地增加模型中的参数数量。出于这个原因，本文通过额外卷积减少了两个特征图的维度，并将两个低维度、相等通道数的特征图连接起来。

**上采样：**在融合之后，细化模块还需要恢复特征图像素分辨率。所以对融合后的特征图进行亚像素卷积上采样处理。上采样的方法有很多，其中亚像素卷积不同于双线性插值，在进行标准卷积之后还要进行额外的特征值重排操作，称为相移操作。该方法有助于消除图像超分辨率任务中出现的块状伪影，降低计算成本。

本文同时考虑三种上采样策略实施结果：双线性插值、亚像素卷积和转置卷积。上采样是生成纤细边缘所必须的结构，在例如增强边缘等任务可视化中非常重要。本文在 DCEN 设计中考虑了这一点，并在本章消融实验小节中对其进行了详细评估。在细化模块中，具体计算如公式3.1所示。

$$\begin{cases} k'_h = W_p^1(k_h) \\ k'_u = W_p^1(k_u) \\ output = upsample \left( W_p^3 \left[ \sigma(k'_h) \oplus \sigma(k'_u) \right] \right) \end{cases} \quad (3.1)$$

其中  $W_p^1$  和  $W_p^3$  分别表示  $1 \times 1$  和  $3 \times 3$  卷积， $k_h$  是细化模块中输入特征图的通道数，卷积和池化后的通道数为  $k'_h$ ，远低于输入值。中间层前向输出到细化模块的

特征通道数为  $k_u$ ，经过卷积和池化后道数为  $k'_u$ 。之后，将两者拼接得到  $k'_h + k'_u$ ，并在此基础上使用  $3 \times 3$  卷积得到最终通道。

在反向路径中，每一层特征图的分辨率会逐层减半，因此在每个细化模块中，需要使用亚像素卷积对输入特征图进行两次上采样，并处理边缘周围的像素。通过逐层细化操作，可以更好地融合高层和低层信息，从而增加检测结果的准确性，并捕捉到更丰富的边缘纹理特征。

### 3.2.3 尺度增强模块

为了丰富 DCEN 浅层学习到的多尺度表示，本文添加了多个尺度增强模块 (Scale Enhancement Module, SEM)。由于不同的卷积层携带了不同尺度的信息，神经网络的规模与深度制约了它可以建模特征的尺度范围。浅层网络在特征有限的情况下可能无法获取全部尺度边缘信息，但通过加深网络则会引入较多的参数，训练就变得异常困难。为了实现浅层网络在不同尺度上进行边缘检测，本文引入 SEM 来增强浅层卷积层可学习的多尺度表示。

在卷积神经网络中，池化通常用于减小图像的尺寸，并进行多次卷积以增加感受野。然而，边缘检测任务是像素级别的，需要在预测之前通过上采样恢复原始图像大小。在这个过程中，池化操作是不可逆的，并且上采样可能会导致图像信息的丢失，使得无法重构小目标物体。为了解决这个缺陷，空洞卷积引入了一个可变的采样率 (dilation)，在卷积核中间插入一些空洞 (dilation)，扩大了感受野，保留了更多的上下文信息，增强了特征的表达能力，从而提高了模型的性能，如图3.3所示。

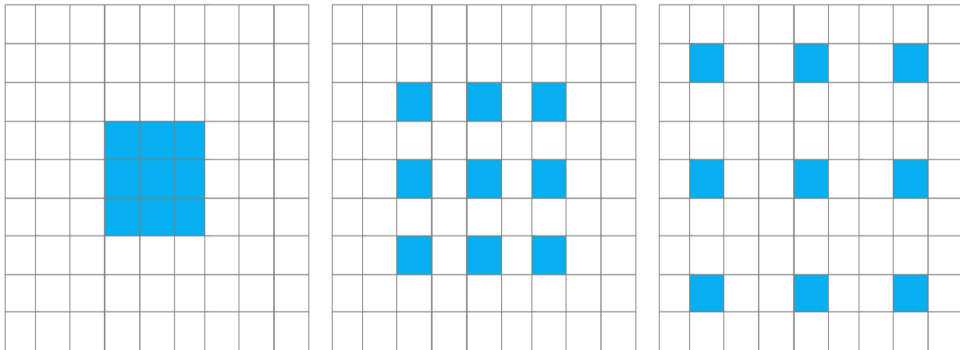


图 3.3 不同空洞率下的空洞卷积示意图

简单来说，空洞卷积是一种提高了感受野的卷积操作，在不增加参数量的情况

下，使卷积核的有效大小变大。空洞卷积通过在卷积核之间添加 0 来实现，这些 0 被称为空洞率，可以控制网络的感受野，同时还能保留图像的细节信息。常用的空洞卷积有  $3 \times 3$  和  $5 \times 5$  的卷积核，其中的空洞率可以取 1、2、3 等。

SEM 利用空洞卷积生成多尺度特征，而不是使用更多的 CNN 结构。在 SEM 模块中，空洞率参数  $r_0$  表示输入特征图的采样步长，标准卷积可以被视为  $r_0 = 1$  的特殊情况。对于每个 SEM 模块，应用具有不同空洞率的  $k$  个空洞卷积。例如，对于每个空洞卷积，本文将空洞率设置为  $r_0$ 。为了取得更好的结果，在本文中进行了许多实验，并且最优值  $k = 3$ ， $r_0 = 4$ 。

对于一个输入的二维特征图  $\mathbf{x} \in R^{H \times W}$  通过卷积核  $\mathbf{w} \in R^{h \times w}$  进行卷积运算，空洞卷积在位置  $(i, j)$  处计算得到的输出  $\mathbf{y} \in R^{H' \times W'}$  如公式 3.2 所示。

$$\mathbf{y}_{i,j} = \sum_{m,n}^{h,w} \mathbf{x}_{[i+r \cdot m, j+r \cdot n]} \cdot \mathbf{w}_{[m,n]} \quad (3.2)$$

其中， $r$  是空洞率，表示采样输入特征图的步幅。标准卷积可以视为空洞率  $r = 1$  的特殊情况。空洞卷积避免了使用池化，将原本的数据内部结构保存下来。但空洞卷积中的卷积核是不连续的，仅对图像中的部分像素点进行卷积计算，对信息连续性造成了损害，影响像素级的曲线提取的精度。因此对于每个尺度增强模块，如图 3.4 所示，对于每个 SEM 使用具有不同空洞率的  $k$  个空洞卷积。对于第  $k$  个空洞卷积，本文将空洞率设置为  $r_k = \max(1, r_0 \times k)$ ，其中涉及 SEM 中的两个参数：空洞率因子  $r_0$  和卷积个数  $k$ 。对于两个超参数的选择在实验部分详细介绍。

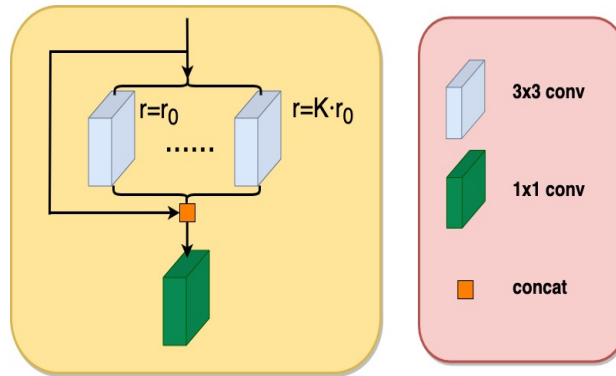


图 3.4 SEM 尺度增强模块

### 3.3 实验与讨论

#### 3.3.1 实验环境和评价标准

本文实验环境使用 Xeon E5 处理器，64GB 内存，RTX 3060 显卡，显存 12GB，Ubuntu20.04 操作系统，使用 PyTorch 搭建整体模型网络。初始训练集 batch size 为 4，学习率为 0.0001，学习率衰减因子为 0.1，权重衰减因子为 0.0002，迭代次数最大为 2000 次。

#### 3.3.2 数据集介绍

本文研究的曲线数据集 CURVES-DATA 截取自开源学术期刊和会议论文中的曲线坐标图像，包含 200 个训练集图片、80 个验证集图片和 150 个测试集图片。每张图片均由人工进行精细标注，并绘制曲线边缘，保证了数据集的高质量和可靠性。为了进一步提高数据集的丰富性和泛化能力，本文对训练集和验证集进行了图像扩充，使用了多种数据增强方法，将数据集的图片数量扩充到 35000 张。

其中，缩放是一种常见的数据增强方法，可以通过按比例缩小或放大图像来增加数据的多样性。本文对数据集进行了随机缩放处理，缩放比例范围为 0.8 到 1.2，以增加数据的多样性。此外，本文还使用了分割、随机裁剪、旋转、翻转等变换方式对数据集进行了扩充，以增加数据的变化性和泛化能力。在分割和随机裁剪方面，通过对图像进行不同的分割和裁剪方式，如随机分割、中心分割、多尺度分割等方式，增加数据的变化性和难度。同时，在旋转和翻转方面，本文使用了随机旋转、水平翻转和垂直翻转等方式，以增加数据的多样性和鲁棒性。

此外，本文还应用了伽马校正技术对图像进行处理，以增加数据集对不同光照条件的适应能力。伽马校正可以调整图像的对比度和亮度，使得图像在不同光照条件下仍能保持较好的边缘信息。通过这些数据增强方法的应用，本文成功地扩充了数据集的规模和多样性，提高了数据集的质量和可用性，为曲线边缘检测算法的研究提供了有力的支持。具体数据集示例如图3.5所示。本章主要使用 SSIM, FSIM, ODS, OIS 和 AP 作为评价指标。



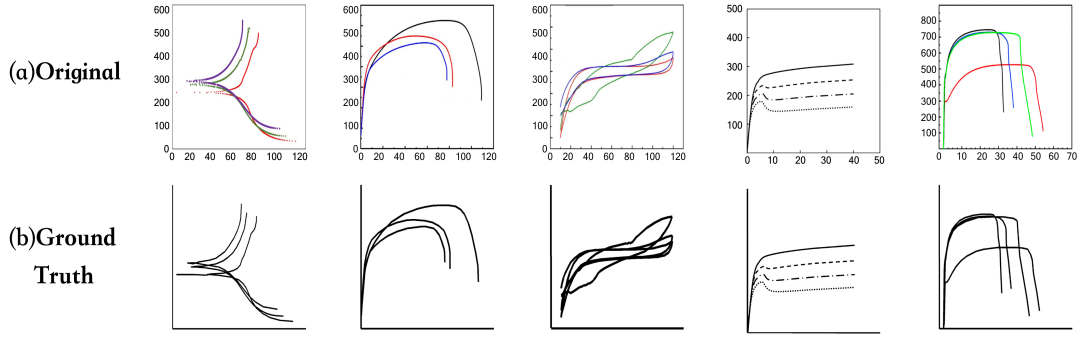


图 3.5 曲线数据集展示图

### 3.3.3 损失函数

DCEA 的损失函数主要源自 HED[35] 网络，由于曲线图像中边缘和非边缘像素之间的比例差距较大，采用标准交叉熵函数时网络难以收敛。为了平衡正负样本之间的差距，在每个像素基础上引入一个类别平衡权重  $\beta$ ，类别平衡权重  $\beta$  来自于动态平衡下正负样本之间的损失。具体的损失函数如公式3.3所示。

$$L_w = -\beta \sum_{j \in Y_+} \log Pr(y_j = 1 | X; W, w^{(m)}) - (1 - \beta) \sum_{j \in Y_-} \log Pr(y_j = 0; W, w^{(m)}) \quad (3.3)$$

其中  $\beta = |Y_-| / |Y|$ ,  $1 - \beta = |Y_+| / |Y|$ ,  $w = (w^{(1)}, \dots, w^{(M)})$  表示网络的所有参数值,  $w^{(M)}$  表示第  $M$  个分支的参数值;  $X$  表示原始输入图像;  $Y$  表示图像像素数,  $Y_-$  和  $Y_+$  分别表示曲线边缘和非边缘像素数量。  $Pr(y_j = 1 | X; W, w^{(m)}) = \sigma(a_j) \in [0, 1]$  是表示像素点  $j$  经激活后是边缘点的概率

### 3.3.4 现有方法和 DCEN 对比

为了展示所提出方法的性能，本文将其与一些著名的方法进行比较，包括传统方法 Canny[11]，和一些深度学习方法 HED[35]、RCF[39]、BDCN[37]、CED[40]、Dex-iNed[61]、TIN[62]、PiDiNet[63] 和 LDC[64] 在 CURVESDATA 数据集上的表现。

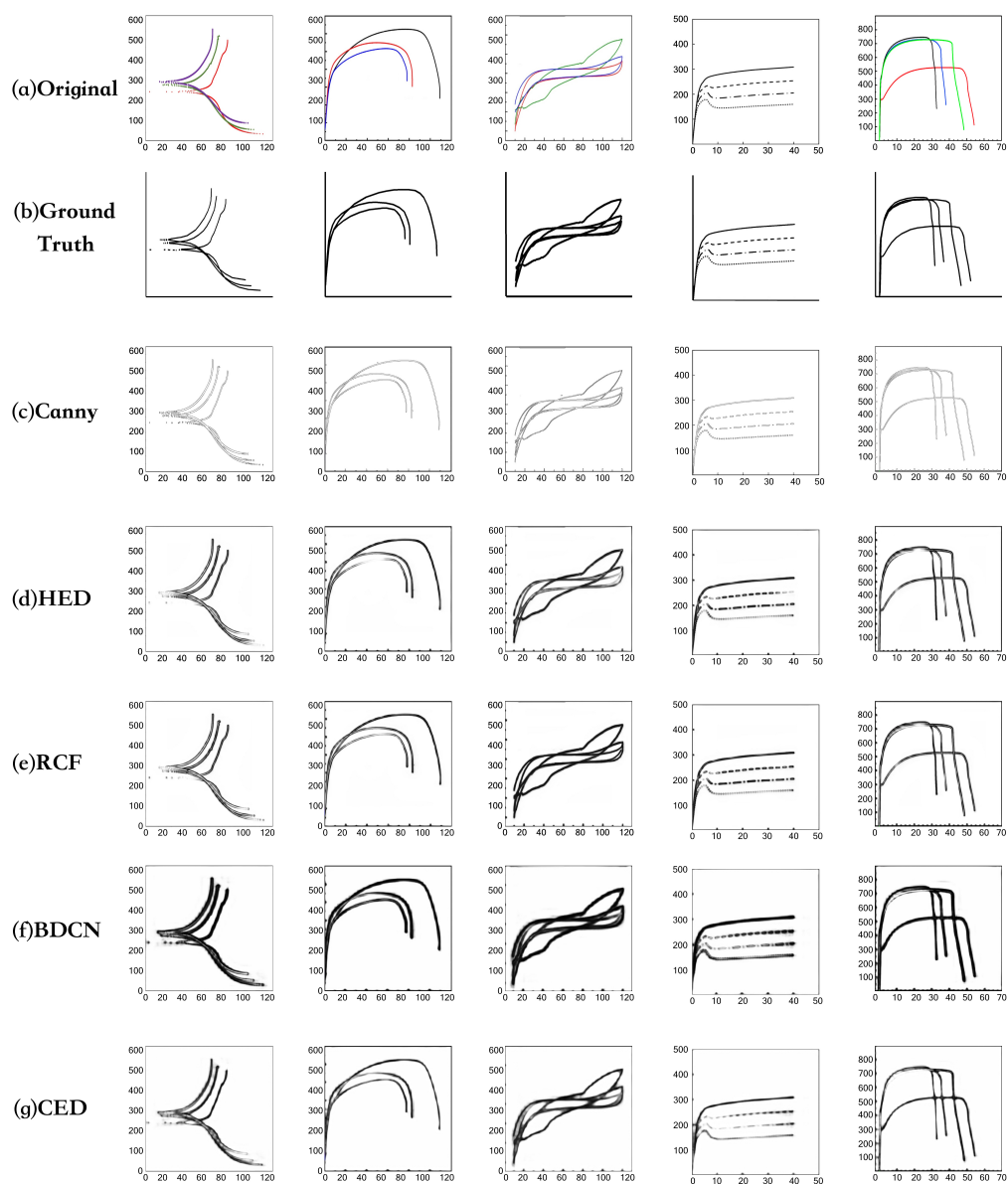


图 3.6 现有方法对比结果可视化结果图。图中分别是原始输入图像、输入图像的标注图，传统方法中 Canny[11] 的提取结果，还有一些深度学习方法 HED[35]、RCF[39]、BDCN[37]、CED[40] 的提取结果

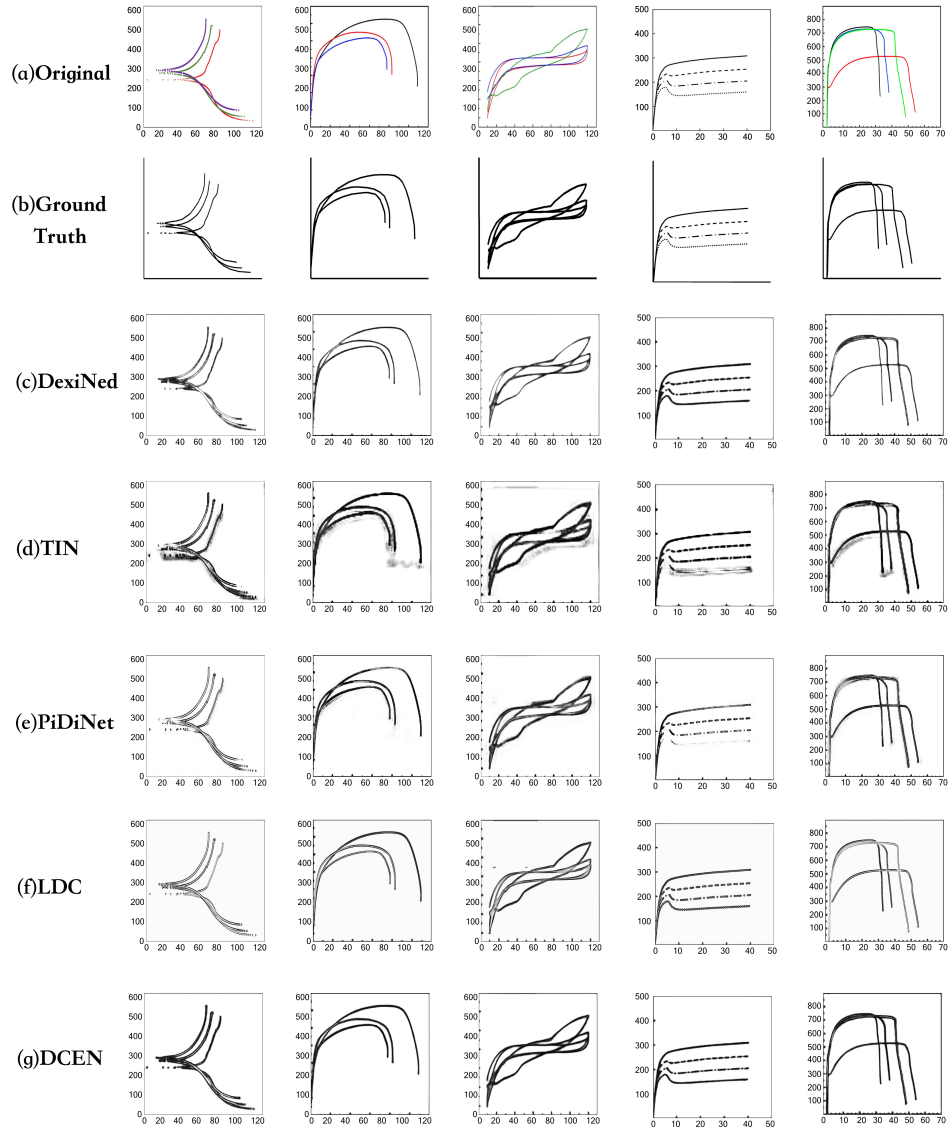


图 3.7 现有方法对比结果可视化结果图。图中分别是原始输入图像、输入图像的标注图，Dex-iNed[61]、TIN[62]、PiDiNet[63] 和 LDC[64] 和 DCEN 的提取结果

表 3.1 评估指标对比

方法	SSIM	FSIM	ODS	OIS	AP
Canny[11]	0.5873	0.6234	0.6001	0.6405	0.6812
HED[35]	0.7201	0.7612	0.7885	0.8085	0.8264
RCF[39]	0.7432	0.7573	0.7673	0.8137	0.7744
BDCN[37]	0.7313	0.7821	<b>0.8283</b>	0.8337	0.8034
CED[40]	0.7424	0.7565	0.7957	0.8151	0.8307
DexiNed[61]	0.7986	0.8593	0.8159	0.8446	0.8767
TIN	0.7169	0.6938	0.7273	0.7437	0.6984
PiDiNet[62]	0.7936	0.8149	0.7926	0.7851	0.8713
LDC[64]	0.7874	0.8462	0.7964	0.8115	0.8425
DCEN	<b>0.8263</b>	<b>0.8764</b>	0.8122	<b>0.8471</b>	<b>0.8923</b>

首先,如图3.6所示,对本文中曲线提取网络的提取结果进行详细分析,并与其他边缘检测方法进行比较。仅分析坐标图像中的曲线内容,并忽略坐标上的数据。其中由 Canny 代表的传统方法,如图3.6(c)所示, Canny 的检测结果并不理想,对于曲线图像中的边缘过于敏感,导致绝大多数曲线细节特征都能够检测出来,但同时造成了一种一根曲线不能完全分割的现象。进一步说明的是,一根曲线的上下两端能够检测出,但是中间的曲线主体部份未能够进行识别,如图3.6(c)中的 Canny 方法第 1 列和第 3 列所示。因此这种被过度检测的结果不适合曲线提取。

HED 方法是一种用于提取曲线图像的方法,然而其性能表现存在较大的不稳定性。尤其是对于均匀且连续的曲线结构, HED 方法的检测结果往往难以令人满意,如图3.6(d)中 HED 方法的第一张和第四张结果图所示。在连续明显的曲线结构检测中, HED 方法存在着检测断裂和错误检测的问题。由于曲线结构的特殊性,该方法在曲线数据集上无法充分学习到曲线的特征,导致检测效果较差。相比之下, RCF 方法在实验结果方面表现稍微出色。例如,在第五张图中, HED 方法出现了断裂,但 RCF 方法成功弥补了 HED 的缺陷,提取出了完整的曲线区域结果。因此, RCF 方法在处理曲线图像方面表现出更高的稳定性和准确性,能够更好地满足实际应用的需求。RCF 的实验结果整体上与 HED 方法基本相似,二者方法对比第五张图, HED 出现了断裂但是 RCF 很好的弥补 HED 的缺陷,在区域内提取出完整的曲线结果。第三张图片对比结果中, RCF 方法对密集的重叠曲线提取效果更为出色,能够整体上识别到曲线,并获得一个很好的检测结果。但是当面临断续的点的曲线类型的时候,

RCF 的结果不尽人意, 如图3.6(e) 中的第四张图部分所示, 可视化结果说明 RCF 的感受野较小, 对于上下文信息的结合能力较差, 导致没有学习到全局的特征信息。

BDCN 方法在具有复杂曲线的图像中容易受到干扰, 如图3.6(f) 中所示, 导致许多细节和纹理丢失, 边界检测结果较为厚实与粗糙, 各个边界被过的分类, 得到曲线提取结果质量较低。虽然该方法能够很好的覆盖到所需检测的曲线, 但是其未能精确定位到曲线真实边缘的中心, 产生了较多的假边缘。造成这种现象的主要原因是由于其内部众多尺度增强模块的存在而使得空洞卷积模块也随之增加, 导致网络过拟合, 降低泛化能力。结果说明 BDCN 对于曲线特征的学习和训练与真实标注结果偏差较大, 没有在学习到正确的曲线特征表示。CED 方法的检测结果从整体上来看, 曲线提取的更加纤细, 在单一曲线上的视觉效果较为出色, 如图3.6(g) 的第一张图所示部分。但 CED 方法的分割结果层次不齐, 虽然该方法是以细化边缘检测结果图为目的可是在一些断续的点上出现了漏判, 如图3.6(g) 的第四列可以看出, 在一些关键点上处理能力较差, 可见对于断续的边缘点上, CED 难以实现较好的检测效果。

DexiNed 是一种基于密集 CNN 的重量级网络, 它能够获取曲线的整体结构, 生成的曲线边缘较厚, 但对曲线边缘细节的感知能力相对较弱。在图3.7(c) 中, 可以明显看出 DexiNed 方法很好地提取了全部的轮廓信息, 但是当需要考虑曲线中部边缘这样的细节时, DexiNed 会出现部分中空的现象。相反, DCEN 方法采用尺度增强模块和细化路径来提高边缘检测的精度。其中, 尺度增强模块可以在多个尺度下提取边缘特征, 并通过跨层连接融合到中间层的输出中。细化路径能够对边缘预测进行后处理, 去除噪声和不连续性, 从而增强边缘的清晰度和完整性。这些技术的应用可以更好地处理各种情况, 并实现更为精准的曲线检测结果。

从 TIN 方法的可视化结果来看, 如图3.7(d), 检测结果相对较差, 出现了不同程度的误检与阴影噪声, 曲线的整体结构杂糅在一起不能明显区分。具体到曲线的边缘细节, TIN 网络出现了大量的检测错误, 这说明, 在处理深层次的曲线特征时, TIN 网络由于其轻量级的结构, 主干网络的多次卷积操作, 会损失重要的边界细节。与此相反的是, 低分辨率的高级特征图则占据了主导地位, 导致损失的边界信息无法被充分利用。同时, 表3.1中也反映了该网络检测结果的差异。PiDiNet 采用了像素差卷积, 将传统的边缘检测算子融入现代 CNN 中流行的卷积运算中, 以提高性能。从图3.7(e) 和表3.1中可以看出 PiDiNet 网络的可视化结果和性能指标很好。然而, 该方

法在细节处理方面不够精细，如图3.7(e)中，PiDiNet的第三张图中间部分和第五张图上方曲线部分可以看出，该方法对曲线边缘细节部分的识别不够稳定，对曲线与曲线之间的粘连边界划分不太敏感。LDC方法是一种轻量化、低复杂度的边缘检测实现方式，其基于DexiNed模型进行改进而来。通过减少卷积块的使用，LDC方法在较小的计算复杂度下达到了接近DexiNed的检测效果。然而，相较于多尺度、大规模的模型，LDC方法提取对象的高级特征能力较弱。虽然LDC方法具有计算复杂度上的显著优势，但其模型结构也因此更加简单。因此，在某些场景下，其精度可能不如一些更复杂的模型。通过图3.7(f)中LDC方法与DexiNed模型的检测结果对比可以看到，LDC方法的可视化结果中出现了很多曲线主干空洞的现象，表征曲线主体结构的能力相对较弱。

从表3.1中可以看出，DCEN在曲线数据集上各项指标分别是0.8263, 0.8764, 0.8122, 0.8471和0.8923，大多指标超过现有方法。BDCN的ODS指标在曲线数据集上表现突出，但是其他指标上DCEN的指标表现更优异。从柱状图3.8中可以看出，本章方法所表示的曲线提取结果在四个指标上稳定的处于领先地位。

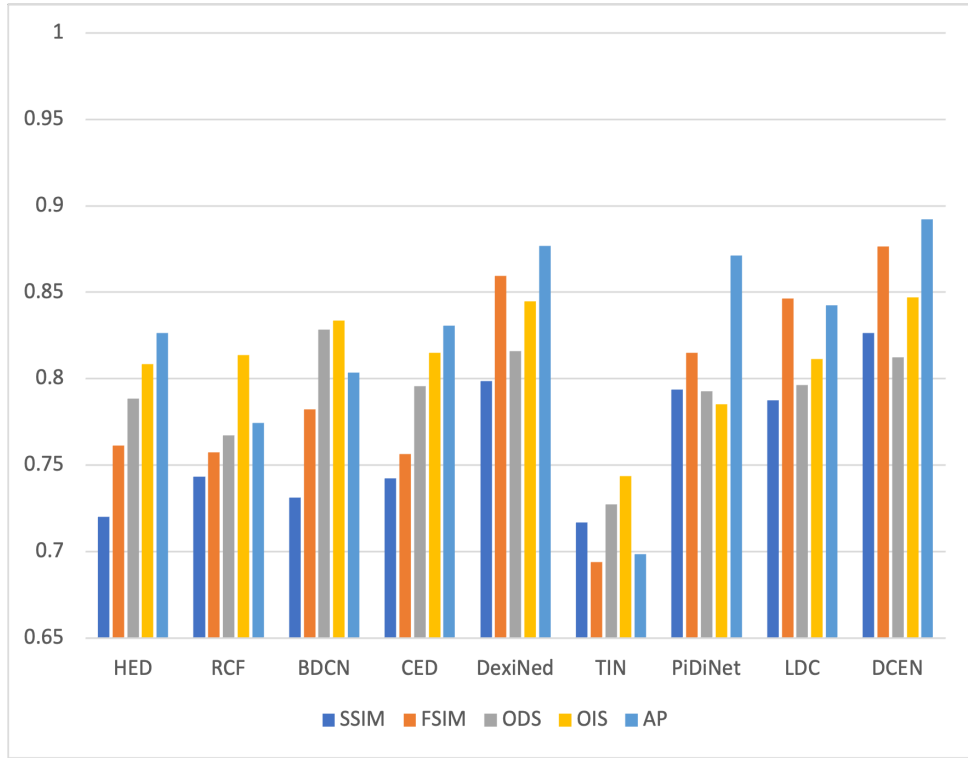


图 3.8 深度学习对比实验结果柱状图。深度学习对比实验方法包括 HED[35]、RCF[39]、BDCN[37]、CED[40]、DexiNed[61]、TIN[62]、PiDiNet[63]、LDC[64] 和 DCEN 的提取结果

本文提出的 DCEN 网络模型对曲线数据集的提取效果具有稳定的性能，不仅可以正确定位曲线边缘，而且在检测时还可以对其进行精细化处理。首先，该方法有效地提取了曲线主干结构，从而准确地识别和检测了曲线蜿蜒区域；其次，该方法能够更加精细地检测不同曲线之间的狭窄边界和细节部分；最后，在存在断续点的情况下，该方法仍然能够准确、清晰地预测出曲线的边界点与非边界点。本章方法较好的权衡了曲线提取的准确度和清晰度之间的关系。

### 3.3.5 与 DexiNed 的对比

受 DexiNed 网络的启发，DCEN 采用密集边缘提取网络作为骨干网络，保留上下文信息的同时有效提取曲线图像边缘信息，对于边缘细节部分信息也能够很好的获取，有利于快速准确地提取粗粒度边缘特征。原网络 DexiNed 利用密集网络结构融合局部和整体特征来优化网络结构，采取多步上采样步骤提高边缘的准确性，但是，这种多输出网络在各阶段融合时候方法较为僵硬，仅依靠一层卷积通道进行融合，中间特征图不能充分融合。此外，较浅层获取的大量纹理特征会对最终图像产生较大影响。

为此，DCEN 使用自顶向下反向细化模块来更好地保留浅层中间特征并精细化曲线边缘，弥补 DexiNed 方法各阶段输出融合僵硬的缺陷，在曲线提取过程中同时考虑检测结果的准确性和清晰度，使得曲线提取结果更加平衡。如图3.7所示的结果表明，与 DexiNed 方法相比，所提出的方法具有更好的边缘锐度和准确性。另外，相较于 DexiNed 方法 DCEN 的网络参数量也有所削减，这不仅使得网络更易于训练，并且获得了更好的检测性能。如图3.9展示了不同方法在曲线数据集上的训练 loss 收敛图，可以发现相较于 DexiNed 方法，本文所提出的模型具有更快的收敛速度，并且在 15 个 epoch 之后表现更为稳定。

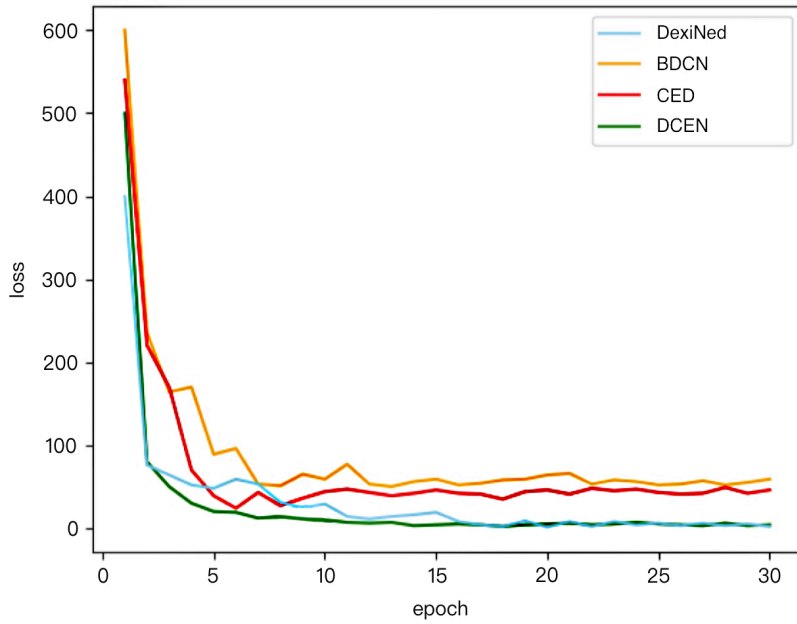


图 3.9 训练 loss 收敛图。

出于对 DexiNed 网络结构与参数量的思考, 经实验发现 DexiNed 仍旧有进一步的优化空间。本文未直接使用 DexiNed 网络是因为 DexiNed 具有良好性能但牺牲了模型容量, 包含 35.2M 可训练参数, 而 BDCN 和 RCF 分别只包含 16.3M 和 14.8M 的可训练参数。原始 Dexined 网络由 6 个主要块组成, 并且其预测结果是 6 个主要块输出结果融合或平均值。在本文数据集上进行训练和测试时, Dexined 的第五、六个主要块输出结果非常相似 (见表 3.2)。因此第六个主要块是否存在对最后融合结果没有决定性影响, 为简化网络结构并降低网络容量考虑, 只保留五个主要块并将参数降至 30M, 相对原始容量大大降低。

表 3.2 DexiNed 最后一个主结构模块对实验结果的影响

卷积块结构对结果影响	SSIM	FSIM	ODS	OIS	AP
减去第六个卷积块结构	0.798	0.860	0.813	0.832	0.872
保留第六个卷积块结构	0.798	0.858	0.816	0.833	0.871

### 3.3.6 消融实验

为了探究上采样方法、SEM 模块和细化模块对最终数据集评估效果的影响, 本文进行了消融研究以探索这些影响。值得注意的是, 消融实验的设计思想并不是判断三组变量的所有条件, 而是分别探索条件并选择最优设计方法来构建模型。



### 3.3.6.1 不同的上采样方法

为了研究不同上采样方法对结果的影响，本文对常用的上采样方法进行了消融实验，包括双线性插值、转置卷积和亚像素卷积。本文在3.3中展示了曲线数据集的实验结果，虽然三种采样方法的评估指标存在一些差异，但总体来说并不显著。如图3.10所示，可以看出亚像素卷积可以获取更详细的边缘信息。双线性插值仍然可以获得更好的检测结果，但在边缘处理方面失去了很多细节。此外，转置卷积可以在曲线边缘区域获得强烈反应但是对细节不敏感，因此本文选择使用亚像素卷积进行上采样。

表 3.3 不同上采样方法对比结果

不同上采样方法	SSIM	FSIM	ODS	OIS	AP
Sub-pixel convolution	<b>0.868</b>	0.859	0.852	0.846	0.861
Bilinear interpolation	0.796	0.823	<b>0.859</b>	<b>0.885</b>	0.875
Transposition convolution	0.826	<b>0.876</b>	0.832	0.867	<b>0.894</b>

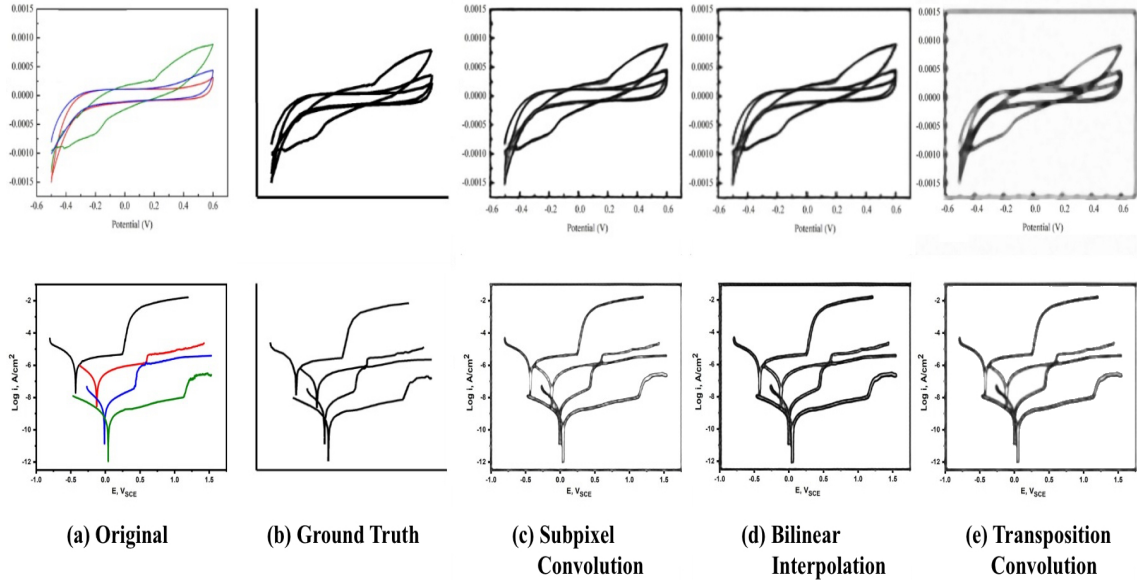


图 3.10 不同上采样方法结果图。

### 3.3.6.2 SEM 模块

SEM (Scale Enhancement Module) 被嵌入到 DCEN 中以丰富多尺度表示，为了验证 SEM 模块的作用及其参数对模型性能的影响，本文在不同的参数配置下重新训

练模型，并进行评估。

首先，本文测试了 SEM 中不同参数对曲线提取效果的影响，即空洞卷积数量  $k$  和空洞率  $r_0$ 。本文固定  $r_0$  并逐渐增加空洞卷积数量，对不同空洞卷积数量产生的曲线边缘图进行评估，实验结果如表3.4所示。在空洞率  $r_0 = 4$  时，空洞卷积数  $k$  对曲线提取效果的影响。 $k = 0$  表示输入直接用于输出而没有使用空洞卷积。结果表明，空洞卷积有效地改善了网络边缘检测性能，但是过大的  $k$ ，即太多空洞卷积，并不能持续改善性能。由于太多空洞卷积会产生高维输出，从这样高维的数据中提取曲线边缘数据信息变得很困难。

表3.5表示了当  $k = 3$  时，不同空洞率  $r_0$  对曲线提取效果的影响，当  $r_0 = 0$  时，表示是使用标准卷积进行特征图的卷积运算。从表中可以看出，更大的空洞率可以更好地改善网络性能，但是当  $r_0$  太大时，如  $r_0 = 8$ ，网络性能开始降低。在固定空洞卷积数量情况下，过多的空洞填充在卷积核中，导致卷积核无法完全覆盖特征图像中的所有像素信息，造成大量像素信息未被有效利用，网络的曲线提取能力开始下降。因此，在空洞率较大时可能需要更多具有不同空洞速率的卷积核来完成多尺度增强任务，相应的根据实验数据表3.4和表3.5可得到最优参数  $k = 3$  和  $r_0 = 4$ 。

表 3.4 SEM 模块固定  $r_0 = 4$  参数，变量  $k$  对结果影响

$k$	SSIM	FSIM	ODS	OIS	AP
0	0.726	0.777	0.816	0.819	0.821
1	0.727	0.778	0.816	0.824	0.824
2	0.727	0.778	0.814	0.825	0.837
3	<b>0.731</b>	<b>0.782</b>	<b>0.818</b>	<b>0.827</b>	<b>0.875</b>
4	0.730	0.780	0.814	0.827	0.837

表 3.5 SEM 模块固定  $k = 3$  参数，变量  $r_0$  对结果影响

$r_0$	SSIM	FSIM	ODS	OIS	AP
0	0.721	0.780	0.815	0.825	0.871
1	0.728	0.781	0.815	0.825	0.871
2	0.728	0.781	0.813	0.825	0.872
4	<b>0.731</b>	<b>0.782</b>	<b>0.818</b>	<b>0.827</b>	<b>0.875</b>
8	0.727	0.779	0.813	0.826	0.872

### 3.3.6.3 细化模块

对于细化模块的添加是否能对网络起到贡献作用，本文设置两组对比实验。第一个是未修改侧向细化模块路径的基本原始网络，第二个是经过侧向细化模块修改的网络。检测效果比较如图3.11所示，具体值分析如表3.6所示。结果表明，本文的细化模块对原始基础网络带来了更好的检测效果。DCEN 可以在检测过程中更好地融合底层和高级信息的中间特征，并保留更多详细信息以确保检测准确性，使检测结果更加精细并提高检测结果清晰度。

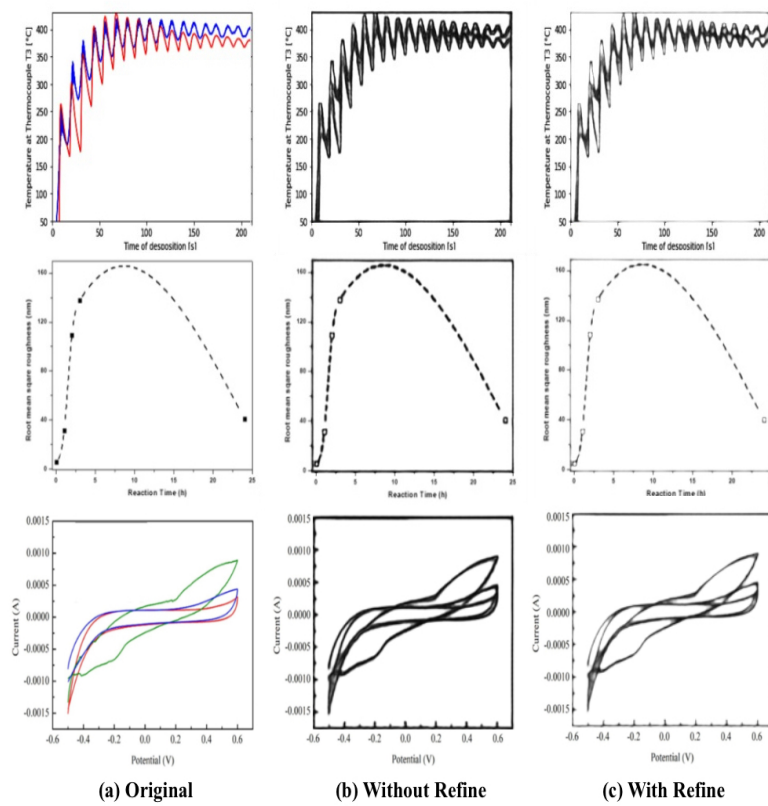


图 3.11 细化模块消融对比图。

表 3.6 网络中细化模块对比

是否添加细化模块	SSIM	FSIM	ODS	OIS	AP
否	0.788	0.859	<b>0.826</b>	0.834	0.861
是	<b>0.824</b>	<b>0.873</b>	0.817	<b>0.875</b>	<b>0.897</b>

### 3.3.6.4 非极大值抑制

在边缘检测中，卷积和梯度运算会在图像中产生梯度方向和大小的信息，这些信息需要被进一步处理以确定真实的边缘位置。如果不进行非极大值抑制，将会产生一些错误的边缘，因为在梯度方向上相邻像素的梯度值可能都很高，而没有局部最大值点的存在。此外，噪声也可能会使得梯度值变得非常小或非常不稳定，从而产生假边缘，非极大值抑制的使用可以很好地滤除这些噪声和错误的边缘。

另一个采用非极大值抑制的原因是，它能够提高边缘检测的准确性和可靠性，由于它可以过滤掉梯度值较小的边缘，同时保留梯度值较大的边缘，因此可以更好地定位和描述图像中的真实边缘。这对于许多计算机视觉应用来说都非常重要，例如目标检测、图像分割、物体跟踪等等。

所以接下来的实验需要验证非最大抑制（NMS）对检测结果的影响，在所有条件相同情况下，本文探究了非最大抑制对实验结果的影响，具体评价指标如表3.7所示，可视化结果如图3.12所示。从图中可以清楚地看到，使用非极大值抑制之后有效地减少了噪声和假边缘对于曲线检测的干扰，并且能够保留曲线中真实边缘的形状和位置。另外从表3.7也可以发现，非极大值抑制对结果的影响较为明显。

表 3.7 非极大值前后性能指标对比

是否使用 NMS	SSIM	FSIM	ODS	OIS	AP
否	0.648	0.597	0.726	0.734	0.711
是	<b>0.826</b>	<b>0.876</b>	0.812	<b>0.861</b>	<b>0.892</b>

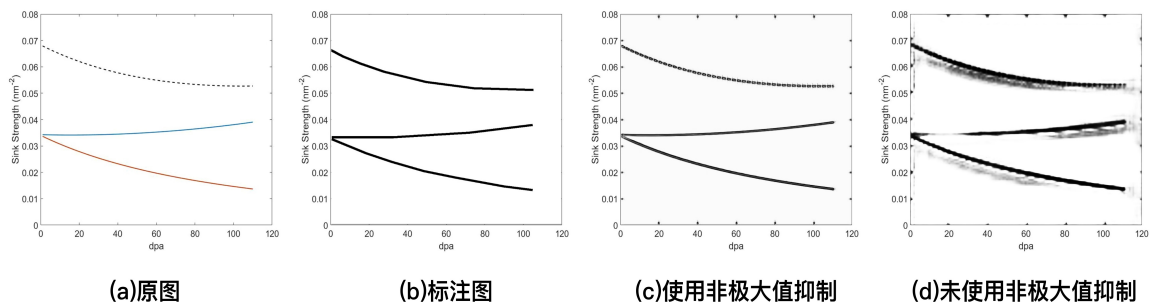


图 3.12 非极大值抑制效果对比。

### 3.4 本章小结

在曲线图像复杂、多样的前提下，缺乏成熟的深度学习技术来较好的解决曲线提取问题。为此，本章利用密集网络的灵活扩展性的优势，结合细化模块提出了精细化曲线提取网络。

本章详细主要工作内容如下：首先，提出了一种全新的曲线检测方法，并详细介绍了改网络架构的实现与各个子模块之间的构成。该网络采用自顶向下的单网络结构，摒弃了 DexiNed 网络的多分枝部分以及最后的复杂模块，既能充分融合低层和高层信息的中间特征，又能简化网络结构。该网络不仅可以正确检测曲线，而且可以提高曲线检测的清晰度，细化曲线边缘。紧接着，介绍了本文手工标注的新型精巧的曲线数据集，实验结果表明，所提出的方法在曲线数据集上取得了较好的结果。与原方法相比，参数量减少的同时评价指标均有所提高，收敛速度也有显著提升。

但是，本章所提的方法也存在一些不足。第一，网络中密集模块的特性，叠加过深的网络虽然能够获取丰富特征，但是会带来参数量的增加影响整体性能。第二，网络中叠加多个卷积块为基础的结构，忽略了曲线图像中不同通道之间的关系，无法避免像素间通道信息的损失。本研究后续会为了解决以上问题而进一步探索。

## 第四章 基于双重高效通道注意力机制的曲线信息提取网络

在上一章中，本文提出了基于密集网络的曲线信息提取网络模型，可以有效的探索曲线图像特征细节并予以检测提取，很好的平衡了曲线提取的正确性和清晰性，然而过多的卷积操作不可避免的丢失曲线图像中的结构化信息和上下文联系，而且过多的密集模块还增加了可训练参数量和训练推理时长。上一章网络是由密集的卷积结构组成，丰富的卷积结构带来了许多视觉特征，但是未能充分利用全局上下文关联的信息，也未限制相邻像素之间的标签以加强深度监督。同时还面临着训练时间长以及大量的存储空间占用等问题。本文在第四章提出基于双重注意力机制的网络，减少密集卷积对于特征提取的影响，从通道特征入手，以多维阶段特征融合方式实现更好的提取效果，进而降低参数与训练时间。

### 4.1 方法概述

本章提出了一个用于曲线图像提取的双重通道注意力曲线提取网络 (Dual Channel Attention Curve Extraction Network, DCACE)，如图4.1所示。该模型利用了 VGG16 网络中的所有卷积层，使用图像到图像的预测来获得多尺度物体或物体局部区域的精确表示。该模型克服了传统卷积神经网络 (CNN) 的局限性，即计算效率低，难以优化，以及在遥远的位置之间传递信息的能力差，通过纳入双重高效通道注意模块 DECA (Dual Efficient Channel Attention)，捕捉全局和通道的背景。此外，该模型还融入了空洞卷积结构，以改善卷积的感受野，在曲线数据集上的实验结果表明，DCACE 模型的性能优于传统的 CNN 和其他最先进的方法。

由于深度卷积神经网络对于网络层次越深的情况下，对于图像的感受野有一定要求。为此，模型的前三个阶段保持不变，在后两个阶段中引入空洞卷积，其中空洞率为 2，以用来在保证参数量不变的基础上增加模型的卷积感受野。为了融合不同阶段的中间层特征，设计阶段特征融合结构 (Stage Fusion Module, SFM) 旨在另不同层级特征得到有效融合，将高尺度的特征反向传递，低层级也能关注到全局尺度信息，有效解决了低阶段生成的特征图中细节杂乱模糊等问题。

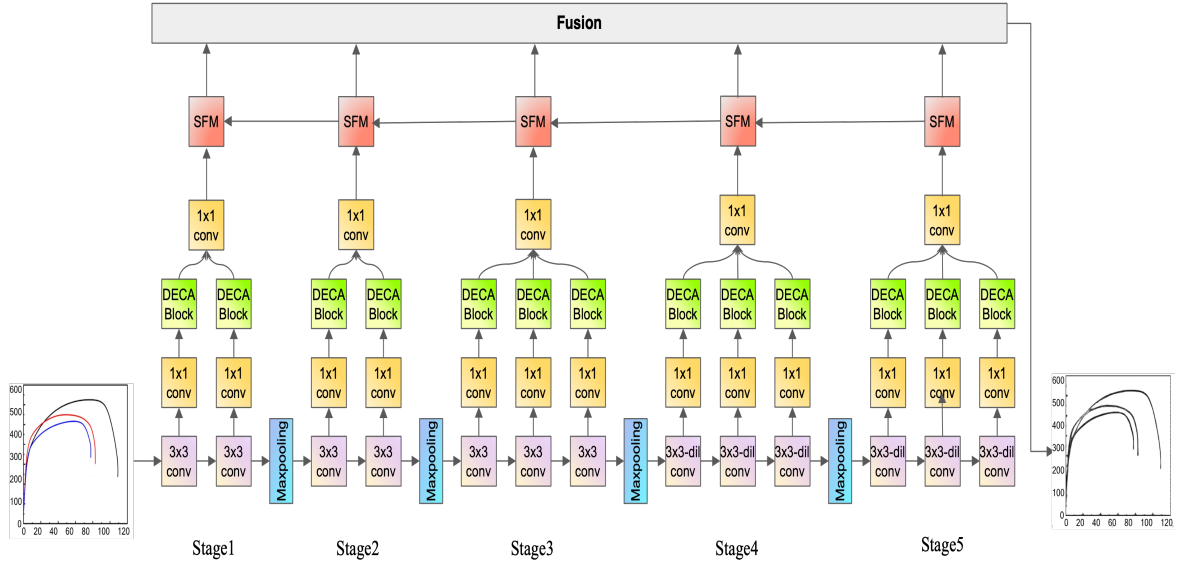


图 4.1 基于双重通道注意力机制的曲线提取模型

#### 4.1.1 双重高效通道注意力机制

人的视觉系统可以轻易地过滤掉无关部分，专注于自己关心的内容，图像领域的研究者希望神经网络能够模仿并学习到这一特点。计算机视觉中的注意力机制（Attention Mechanism）的核心思想就是基于原有的数据找到其间的关联性，然后突出某些重要特征，从而能够忽略无关噪声信息，关注重点信息。注意力机制可对特征进行校正，校正后的特征可保留有价值的特征，剔除没价值的特征。

注意力机制的目的是突出局部特征所捕获的重要特征，并消除全局特征所传达的不相关信息，这一机制旨在改善长距离全局相关性的建模，同时确保计算和统计效率。注意力模块使用每个位置的反应的加权平均数来计算上下文感知的表示，然后将其嵌入特征空间，有助于建立本地特征和其相关的全球背景之间的关系，使网络能够集中于最重要的信息。所以，注意力机制为处理大量数据和提高深度学习模型的性能提供了一个更有效的方法。

经典卷积神经网络 CNN 的设计使其能够学习到图像中的局部特征，但其结构感受野的局限性会导致一些限制，例如：

(1) 局部感受野限制。每个神经元只关注输入图像的局部区域，这种设计使得神经元只能感知到该区域内的特征信息，而无法获取更广泛的上下文信息。这种局限性会导致神经元无法在全局范围内识别图像的复杂模式，例如识别整个对象或物体。

(2) 信息损失。在网络的较深层中，由于感受野的大小限制，神经元只能接收到



前一层的局部特征信息。这会导致在信息传递过程中丢失了一些重要的上下文信息，这可能会影响网络在复杂任务上的性能。

在全局上下文信息缺乏的情况下，局部特征存在误导模型判断的可能，不同特征的像素被错误的划分为同一类别，进而造成类内特征的不一致，影响检测效果。与此同时，通道是图像特定类型的响应，不同语义间的响应高度相关，因此，可以从提高语义特定特征表示的角度出发，增强通道之间的相关性，缓解深度卷积网络的缺陷。为了根本性解决这一问题，本文从采用注意机制建立特征之间的相关性出发，整合上下文信息，提高通道间的相关性，以增强语义特定的特征表示。将卷积网络提取的多尺度特征输送到全局通道注意力模块中，为不同通道赋予权重不同系数，强化重要特征抑制无用特征，逐步将全局上下文信息嵌入至局部特征中。全局 DECA 具体结构如图4.2所示。

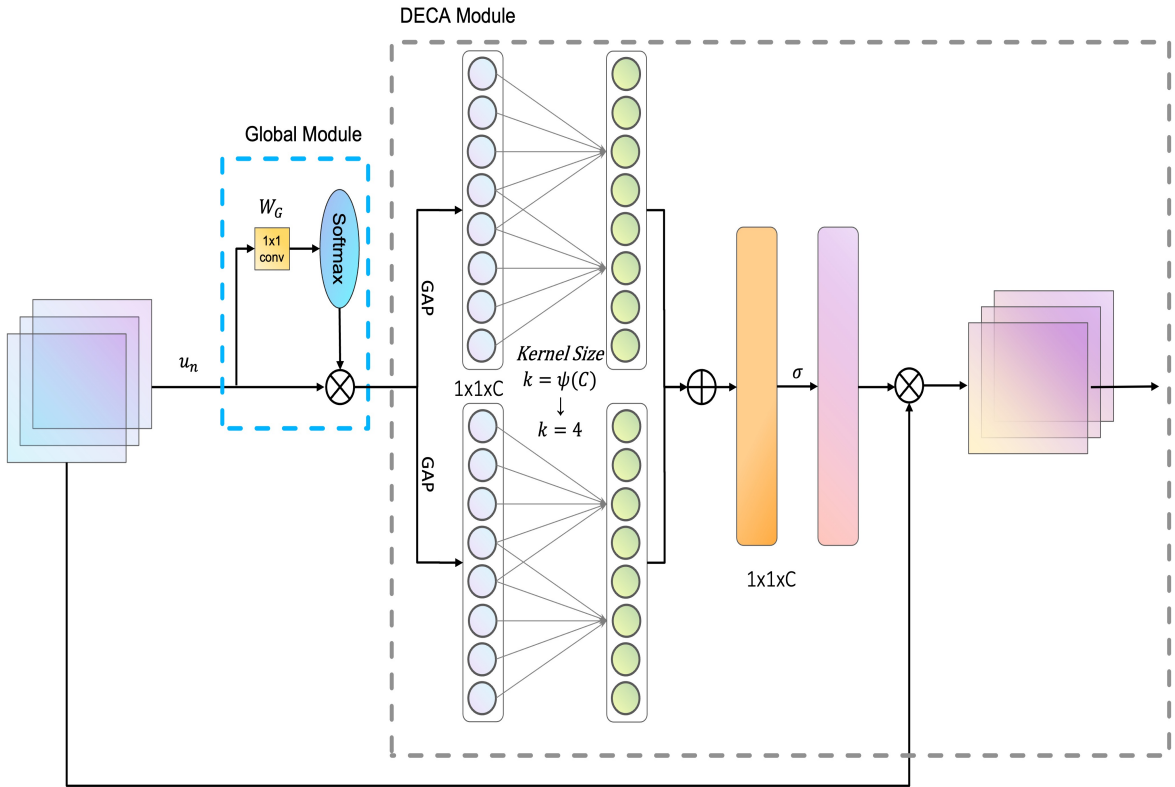


图 4.2 DECA 双重高效通道注意力模块

全局模块如图4.2中浅蓝色矩形所示，通过使用  $W_G$  卷积和 softmax 函数聚合并池化全局特征，它为每个像素生成了特定于当前位置的全局上下文注意力图  $S$ ，通过使用全局注意力网络能够汇总各个位置的特征信息，并将其加权分配至每个位



置，以形成特定于每个位置的全局上下文特征。这样做可以增强每个位置的特征，使网络能够获取全局长程上下文信息，并从其他位置汇集信息以实现聚合。本文用  $U = u_n, n = 1, \dots, N$  代表全局模块输入的特征图，如公式4.1所示：

$$S = \sum_{n=1}^N \frac{f(u_n)}{C(U)} u_n \quad (4.1)$$

其中  $N$  表示输入特征图中所有的像素个数， $n$  表示了输入图中所有不同像素的位置， $f(u_n) = e^{W_G u_n}$  是广泛使用的嵌入式高斯核函数，用于计算尺度空间的相似程度， $C(U) = \sum_{m=1}^N e^{W_G u_m}$  是一个归一化因子。

紧接着将全局模块输出传入双重高效通道注意力机制（DECA）中，如图4.2中浅灰色矩阵所示，增强低维特征中有效的部份。为了在卷积神经网络中学习到高性能和高效率的通道注意力，有效的方法是避免降维，并采用适当的跨通道交互策略。DECA 考虑到这一点，使用自适应一维卷积核和不降维的跨通道交互策略，更有效地增强了低维特征与外部关联关系之间的信息表达。此外，该策略还可以避免低维外部特征受到噪音的干扰，从而保持了网络模型的效率和有效性。使用全连接层在保持大幅降低参数量的同时提供相当的性能，对于曲线特征对应的通道的学习更加有利，帮助有效的提取曲线特征。最终的通道注意力权重计算公式如公式4.2所示：

$$\omega_i = \sigma \left( \sum_{j=1}^k (w_i^j y_i^j)_1 + \sum_{j=1}^k (w_i^j y_i^j)_2 \right); y_i^j \in \Omega_i^k \quad (4.2)$$

其中  $\sigma$  代表激活函数， $y_i$  表示不同的通道， $w_i$  为相应通道  $y_i$  的权重， $\Omega$  是根  $y_i$  相邻的  $k$  个不同通道。卷积核  $k$  的计算如公式4.3所示：

$$k = \psi(C) = \left\lfloor \frac{\log_2 C}{2} + \frac{1}{2} \right\rfloor_{\text{odd}} \quad (4.3)$$

其中  $C$  代表通道维度， $|t|_{\text{odd}}$  代表变量  $t$  最近的奇数。采用全局平均池化（GAP）去处理每个特征图以生成具有  $n$  个元素的列向量，其中  $n$  表示要分类的类别。GAP 提高了计算效率和对输入空间信息变化的鲁棒性，与全局最大池化（GMP）仅考虑特征图中的最大值而忽略其余部分不同，GAP 考虑所有元素以获得整体特征信息。为

了加强对有效特征值的提取，防止重要特征被遗漏。

DECA 网络模型改进了 ECA 网络结构模块，通过并行的 2 次全局平均池化和一维卷积运算进行特征融合，强化特征提取能力。利用内核大小为  $k$  的自适应一维卷积来促进通道与相邻  $k - 1$  个通道之间的信息交互，如公式 4.4 所示。

$$w = \sigma((Conv_1 D_k(y))_1 + (Conv_1 D_k(y))_2) \quad (4.4)$$

其中  $Conv_1 D_k$  代表卷积核大小为  $k$  的一维卷积， $(Conv_1 D_k)_1$  和  $(Conv_1 D_k)_2$  为 2 次独立的一维卷积， $y$  代表通道。

作为一个轻量级模块，与其他通道注意力方法 [55, 65] 相比，ECA-Net 最大的优势就是简单高效，是采用该模块的主要原因，从上述结构图中可以看到，整个计算过程中需要训练的参数只有一个一维卷积核。

#### 4.1.2 阶段特征融合

曲线提取任务中，复杂曲线坐标图像包含丰富的边缘信息，深度挖掘这些特征是实现高精度边缘提取任务的关键。融合高层和低层的特征是提升边缘检测提取性能的一个重要方法，低层特征具有高分辨率，主要包含位置和细节特征，但包含较少的语义信息并且有噪声信息。相反，高层特征有较强的语义表征能力，但对细节较为不敏感，因此，将高层特征与低层特征充分融合能够使得网络更深层次挖掘复杂的边缘特征，使得决策信息更加充分，从而提高边缘检测效果。

DCACE 模型采用了特征金字塔技术，将多尺度的特征在每个层次上无缝融合。这种方法允许在低层就能关注到高层和全局特征，有效解决了特征图在低层次阶段生成中细节模糊不清的问题。

阶段融合模块 (Stage Fusion Module, SFM) 采用了一个多步骤的过程来无缝融合高层次和低层次的特征。首先，对高层次的特征进行上采样并与低层特征进行 concat，最后再通过卷积来压缩合并后的特征，这使得低层能够有效地融合高层的信息。为了防止模型遗漏低层特征中的重要细节，SFM 通过残差网络将上一层的特征直接添加到下一层的特征中，以保持信息的连续性和完整性。最终输出的特征图包含了来自不同层次的特征信息，同时也保留了每个层次的局部细节和全局上下文信

息。SFM 结构如图4.3所示。

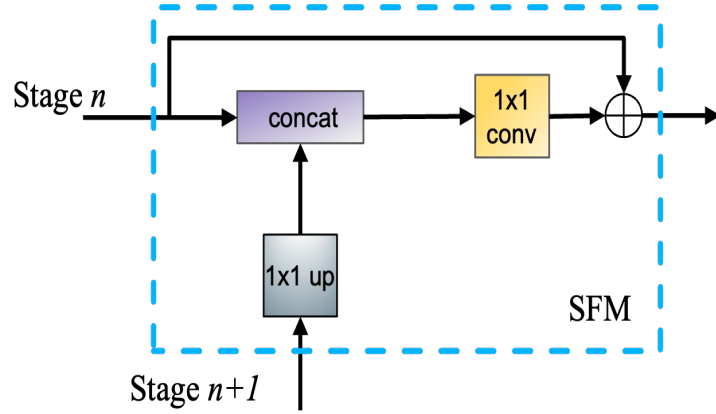


图 4.3 阶段特征融合模块

最后对每个阶段的输出进行上采样，使得每个阶段都能产生一个边缘图，模型在此边缘图基础上进行监督学习。具体融合部份结构细节如图4.4所示。采用了全卷积网络（FCN）中常用的反卷积层（Deconvolutional Layer），以恢复图像边缘的空间位置信息。反卷积层能够将低维的特征还原为高维的图像原始分辨率，这样就可以更精确地定位边缘位置，紧接着采用单一的  $1 \times 1$  卷积来合并所有层的多尺度结果，产生 DCACE 模型的最终输出边缘图像，之后是进行监督学习。

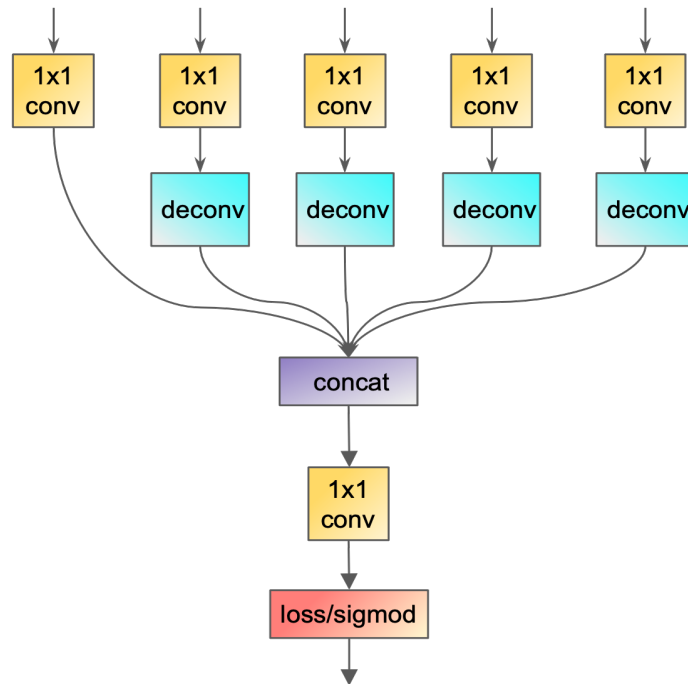


图 4.4 多尺度特征融合模块

### 4.1.3 改进的自适应权重损失

对于像曲线提取这样的像素级别的分类问题，是把曲线坐标图上曲线上的每个像素看作是一个二分类问题，即分为曲线上的点和非曲线上的点，故本文采用交叉熵作为损失函数。为了提升模型的鲁棒性，对标注图中的每个像素值进行归一化处理，将标签图变成概率图，并将概率为 0 的点作为非曲线点，概率大于  $\eta$  的像素点作为曲线点。本模型对每个像素点的损失函数表示如公式4.5，4.6所示。

$$l(X_i^k; W) = \begin{cases} \alpha \cdot \log(1 - P(X_i; W)) & y_i = 0 \\ 0 & 0 < y_i \leq \eta \\ \beta \cdot \log(P(X_i; W)) & y_i > \eta \end{cases} \quad (4.5)$$

$$\begin{cases} \alpha = \lambda \cdot \frac{|Y^+|}{|Y^+| + |Y^-|} \\ \beta = \frac{|Y^-|}{|Y^+| + |Y^-|} \end{cases} \quad (4.6)$$

其中  $|Y^+|$  和  $|Y^-|$  分别代表正负样本的数量， $\alpha$ ， $\beta$  和  $\lambda$  用来平衡正负样本数量差， $X_i$  代表网络中的激活值， $y_i$  表示标注图中像素点  $i$  是曲线上的点的概率， $P(X)$  是标准的 sigmoid 函数， $W$  表示网络中可学习的参数。

在多尺度结构中，每个阶段侧方输出检测边缘的纹理结果差异相当大，每个层的输出可能需要不同的损失来调整，值得注意的是融合阶段的损失是至关重要的，因为它对最终的预测结果起着主导作用。在具体的试验中，发现训练 25 轮之后，最初 stage1 和 stage2 的特征图显示缺乏复杂的纹理特征。这些对最终预测结果产生不利影响的原因可能是当低层特征与高层特征融合时，出现特征混淆，导致模型对不相关的特征进行关注，从而影响结果。为了抑制这种影响，本文为不同的阶段设置不同的损失权重，同时重点提高最终融合时的权重系数，以平衡不同损失之间的关系，总损失如下公式4.7所示。

$$L(W) = \sum_{i=1}^{|I|} \left( \sum_{k=1}^K \alpha^k \cdot l(X_i^{(k)}; W) + \beta^{fuse} \cdot l(X_i^{fuse}; W) \right) \quad (4.7)$$

其中  $X_i^{(k)}$  是第  $k$  个阶段输出中的第  $i$  个像素点的激活值； $X_i^{fuse}$  是融合输出中

第  $i$  个像素点的激活值； $\alpha^k$  表示第  $k$  个阶段的损失权重； $\beta^{fuse}$  表示最终融合部份的损失权重； $|I|$  代表图片中所有像素点； $K$  是模型中所有 stage 的数量。实验 loss 如图4.5所示。

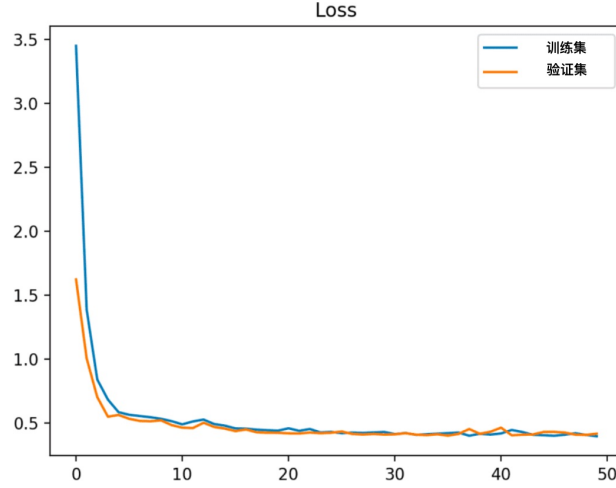


图 4.5 训练 loss 示意图

## 4.2 实验与讨论

### 4.2.1 实验环境和评价标准

本文实验环境使用 Xeon E5 处理器，64GB 内存，RTX 3060 显卡，显存 12GB，ubuntu20.04 操作系统，使用 PyTorch 搭建整体模型网络。

模型初始 batch size 为 4，基础学习率为 0.00001，学习率衰减因子为 0.1，权重衰减设置为 0.0003。不使用预训练模型，用 Gaussian 分布初始化网络参数。数据集部分采用第三章所提曲线数据集，本章主要使用 SSIM，FSIM，ODS，OIS 和 AP 作为评价指标。

### 4.2.2 与现有方法检测结果对比

为了进一步证明本文的网络结构的有效性，在相同的曲线数据集上与当前先进的方法进行了比较。包括传统方法 Canny[11]，和一些深度学习方法 HED[35]、RCF[39]、BDCN[37]、CED[40]、DexiNed[61]、TIN[62]、PiDiNet[63] 和 LDC[64]。对于一些基于深度学习的方法，重新在曲线数据集上训练，以显示不同方法之间的结果差异。

首先，详细分析了本文提出的曲线提取网络的提取结果，并与其他边缘检测方

法进行了对比,如图4.6, 4.7和表4.1所示。对于本文的结果分析,本文的数据集的评价标准是针对曲线坐标图中的曲线部分,其坐标部分未参与评价。

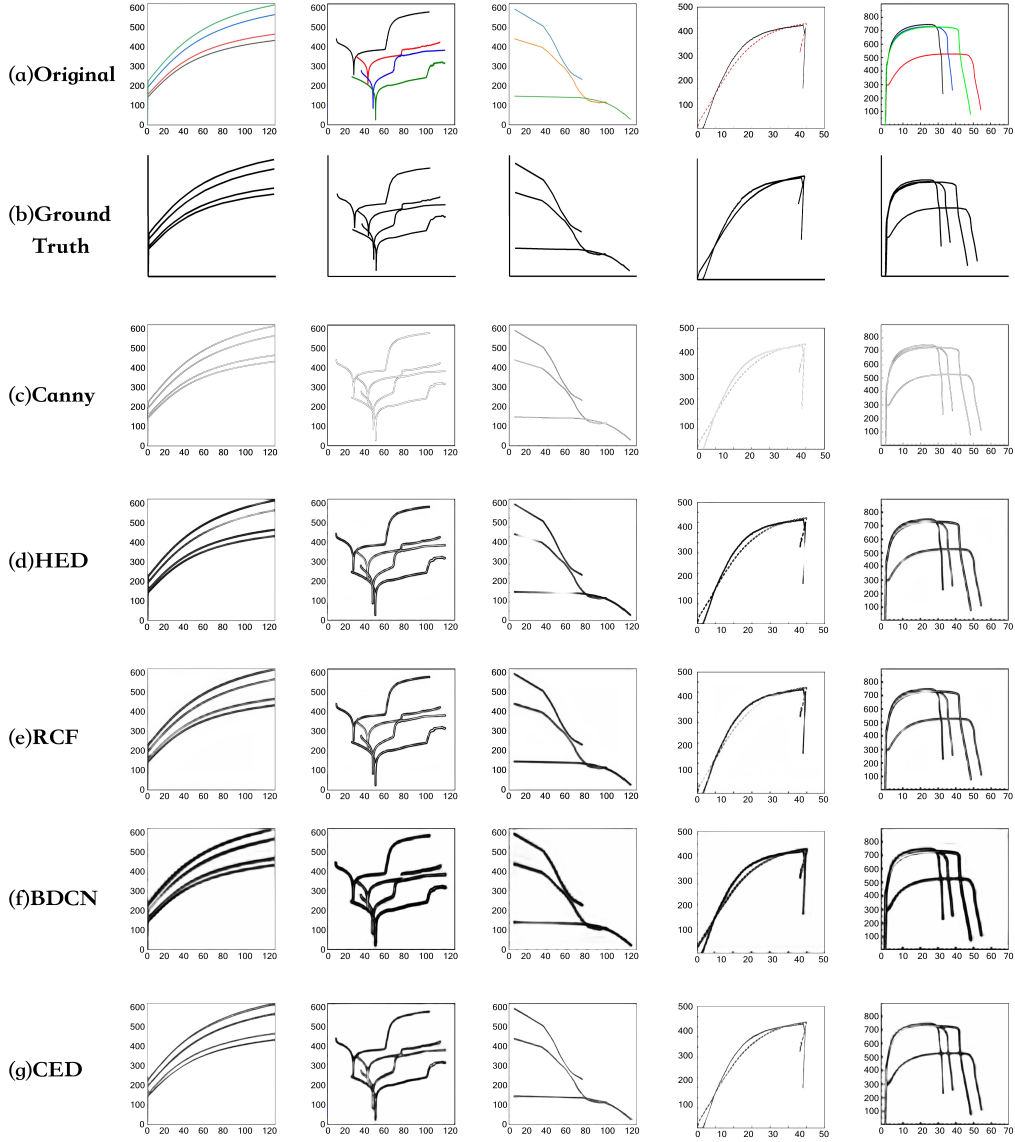


图 4.6 现有方法对比结果可视化结果图。图中分别是原始输入图像、输入图像的标注图,传统方法中 Canny[11] 的提取结果,还有一些深度学习方法 HED[35]、RCF[39]、BDCN[37]、CED[40] 的提取结果

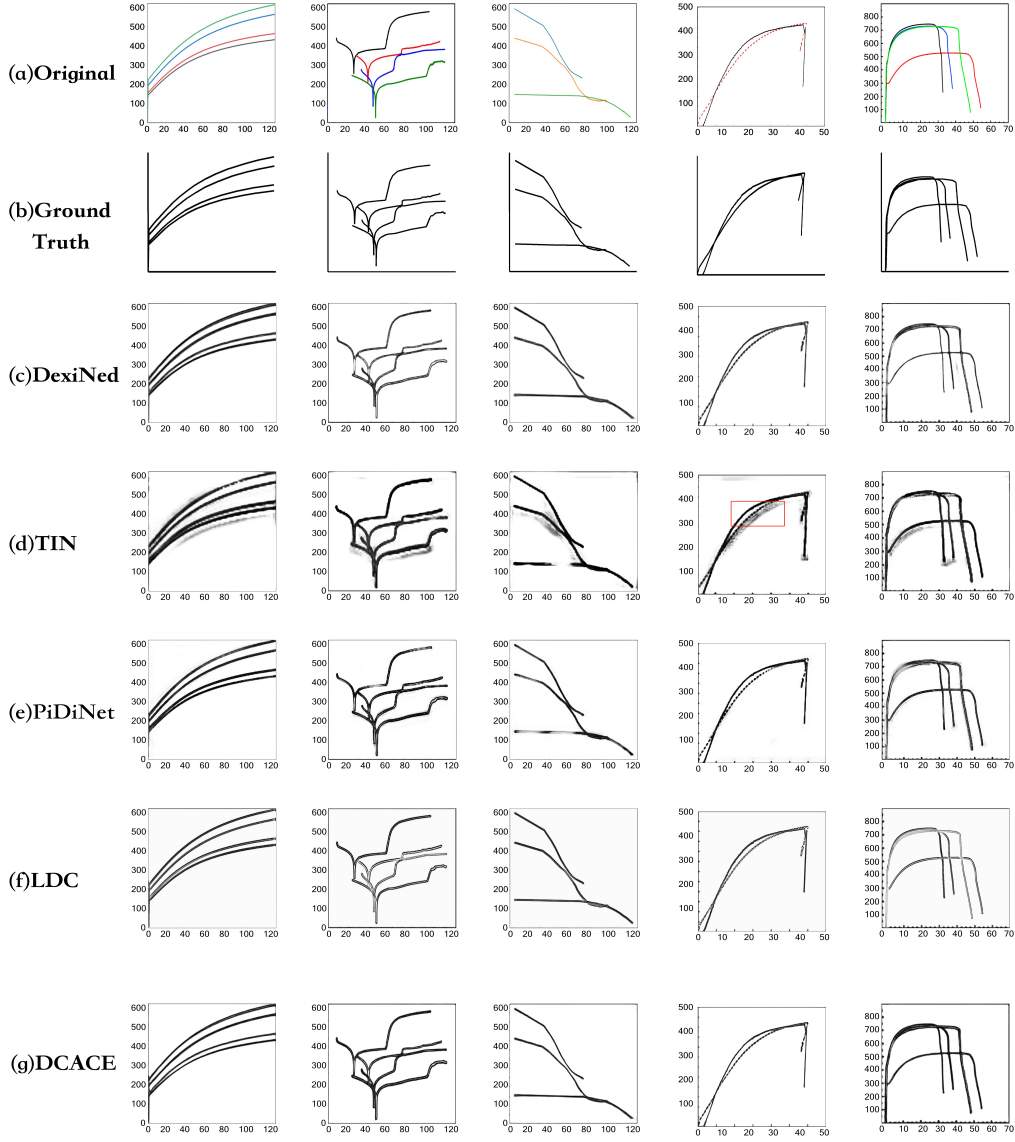


图 4.7 现有方法对比结果可视化结果图。图中分别是原始输入图像、输入图像的标注图，DexiNed[61]、TIN[62]、PiDiNet[63]、LDC[64] 和 DCACE 的提取结果

在表4.1列出了不同方法得出的定量结果。基本上,与最先进的方法 (DexiNed) 相比,曲线数据集上的指标均有所提高。本文认为,DCACE 的出色效果是因为在网络设计过程中,充分考虑到了高效通道信息在曲线提取中的作用。因此,在检测过程中,所提方法可以更好的保留细节信息。

表 4.1 评估指标对比。DCACE 与现有主流或先进的方法对比，主要为传统和深度学习方法。

方法	SSIM	FSIM	ODS	OIS	AP
Canny[11]	0.5873	0.6234	0.6001	0.6405	0.6812
HED[35]	0.7201	0.7612	0.7885	0.8085	0.8264
RCF[39]	0.7432	0.7573	0.7673	0.8137	0.7744
BDCN[37]	0.7313	0.7821	<b>0.8283</b>	0.8446	0.8367
CED[40]	0.7424	0.7565	0.7957	0.8151	0.8307
DexiNed[61]	0.7986	0.859	0.8159	0.8337	0.8713
TIN	0.7169	0.6938	0.7273	0.7437	0.6984
PiDiNet[62]	0.7936	0.8149	0.7926	0.7851	0.8034
LDC[64]	0.7874	0.8462	0.7964	0.8115	0.8425
DCACE	<b>0.8659</b>	<b>0.8776</b>	0.8048	<b>0.8735</b>	<b>0.9038</b>

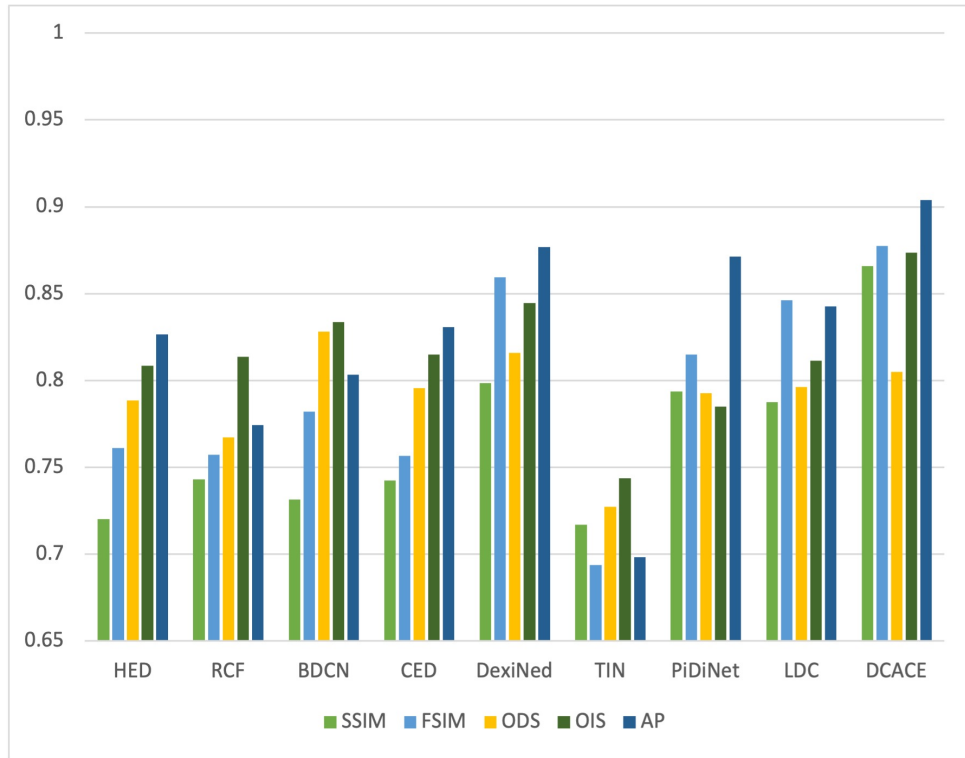


图 4.8 深度学习方法对比实验结果柱状图。深度学习方法包括 HED[35]、RCF[39]、BDCN[37]、CED[40]、DexiNed[61]、TIN[62]、PiDiNet[63]、LDC[64] 和 DCACE 的提取结果

表4.1展示了本章所提方法与诸多先进方法的对比结果。图4.6和4.7展示了本章方法与其他方法的可视化结果。柱状图4.8主要体现了不同深度学习方法之间在测试集上的评价指标结果。

Canny 是传统图像处理方法，对于对于图像灰度变化十分敏感而不具有对象感



知性质，容易出现过度检测，如图4.6(c)中，Canny 的检测结果中绝大多数的梯度变化较大部分都被检测到，各个曲线结构被过度分割，断续的点也无法很好的呈现。HED 方法在曲线图像提取中性能不稳定，对于均匀且连续的曲线上检测结果没有达到很好的效果，如图4.6(d)中 HED 方法的第三张和第四张结果图。HED 方法对于连续明确的曲线结构的出现了检测断裂甚至是错误检测的结果，该网络在曲线数据集上，由于曲线结构的特殊性无法充分学习到曲线特征，导致效果检测较差。RCF 的实验结果整体上比 HED 出色，二者方法对比第三张图，HED 出现了断裂但是 RCF 很好的弥补 HED 的缺陷，在区域内提取出完整的曲线结果。但是当面临断续的点的曲线类型的时候，RCF 的结果不尽人意，如图4.6(e)中的第四张图部分所示。可视化结果说明 RCF 的感受野较小，对于上下文信息的结合能力较差，导致没有学习到全局的特征信息。BDCN 的实验结果在视觉效果上是突出了检测的准确性，如4.6(f)所示 BDCN 方法所示，由于其内部网络结构中存在过多的空洞卷积模块容易导致曲线整体更加粗厚，曲线边缘更为粗糙，同时会出现噪声。结果说明 BDCN 对于曲线特征的学习和训练与真实标注结果偏差较大，没有在足够的曲线数据集上学习到正确的特征表示。CED 方法的检测结果从整体上来看，曲线提取的更加纤细，在单一曲线上的视觉效果较为出色如4.6(g)的第一张图所示部分。但是当面临复杂曲线杂糅的情况时，如图4.6(g)的第四列，CED 方法对于检测正确性方面有所欠佳。

同时，DexiNed 是在 HED 和 Xception 的基础上构建的极度密集网络，它真正能够获取到曲线的整体结构，但它感知曲线边缘细节的能力相对较弱，因此它只能生成厚实型曲线边缘。如4.7(c)DexiNed 方法对应结果明显可见，该算法虽然很好地提取了总轮廓信息，但是对纤细的曲线检测出的结果更加宽厚且曲线中部出现的空洞现象，反观 DCEN 方法能够处理得更好并实现了更好的结果。TIN 方法从可视化结果上展示结果来看检测结果相对较差，出现了不同程度的误检与阴影噪声。曲线整体结构不能明显区分，杂糅在一起，在曲线的边缘出现大量的检测错误，说明 TIN 网络在处理高分辨率的曲线特征时，由于主干网络的多次卷积操作，损失重要的边界细节。相反，低分辨率的高级特征图则占据了主导地位，导致损失的边界信息无法被充分利用。同时在表4.1中也体现出了改网络检测结果的差异。

PiDiNet 提出了一种简单、轻量的像素差分网络结构，用于有效的边缘检测。PiDiNet 采用了新颖的像素差卷积，将传统的边缘检测算子集成到现代 CNN 中流

行的卷积运算中，以增强任务性能。从4.7(e)和表4.1中可以发现，PiDiNet的实验结果很好，但是细节处理不够精细，如4.7(e)中PiDiNet的第三张图中间部分和第五张图上方曲线部分。该方法在部分图像上表现结果较为优异，然而对曲线之间粘连的边界划分不敏感。从表4.1中也可以看到PiDiNet方法的AP指标相对DECN较低。LDC方法是DexiNed模型的一种轻量化、低复杂度实现，该方法保留部分DexiNed的主干结构并减少部分卷积块的使用，仅用674K个可训练参数达到接近DexiNed的效果。LDC在卷积之间引入密集连接来增强特征利用，但是对于提取对象高级特征的能力明显弱于多尺度、大规模的模型，如图4.7(f)中和DexiNed的检测结果对比中可以看到，LDC的可视化结果中出现了很多曲线主干空洞的现象，对于曲线主体结构的表征能力较弱。

本文的网络模型对曲线提取效果呈现稳定的性能。曲线的整体结构能够很好的检测，对于大面积的曲线之间的非联通域检测准确，边缘的大部分噪点都能很好地过滤。多重曲线之间交叉部份的检测结果优于绝大多数方法，在曲线略微重叠的情况下依然很好地平衡了检测准确率和边缘定位精度。

由于存在部分曲线高度重叠的情况，RCF和DexiNed方法对一些无关的细节处理能力较弱，每个不同阶段中都包含了一些模糊的线条，难以准确比较曲线。为了解决这个问题，引入了DECA注意力机制，DECA利用权重分配来有效地提取重要特征，防止无效的特征被重复使用。此外，多尺度融合被纳入底层以帮助融合多个尺度的特征，减少了较大曲线中空洞的出现。视觉和定量评估的结果表明，与其他方法相比，DECA能很好地适应不同的重叠曲线场景，减少边缘混淆改善边缘混乱的问题，没有过多的复杂纹理，提高特征提取的完整性。

### 4.2.3 消融实验

本文提出了一个新颖双重高效通道注意力模块，能够精确的将坐标图像中的曲线边缘提取出来，并且在未使用非极大值抑制的情况下，DCACE模型生成的曲线边缘也非常清晰，达到前列水平。为了了解不同注意力机制以及空洞卷积对网络模型的影响，本文做了详细的消融实验。第一组是基础的VGG16网络结构不添加任何空洞卷积（Dilated Convolution, DC）和注意力模块；第二组的网络模型中stage3, stage4, stage5结构中替换为空洞卷积模块；第三组是添加了普通的通道注意力模块(CA)；第

四组则是更换为高效的通道注意力模块 (ECA); 第五组是本文网络模型采用空洞卷积和双重高效通道注意力机制模块 (DECA), 并使用阶段特征融合结构 (SFM)。

表4.2展示了消融实验中不同指标的对比结果, 从指标上可以看出本文提出的双重通道对于模型性能有着明显提升。图4.9展示了消融实验部分对比的可视化结果, 对于重点关注位置由红色矩形框标出。在基础网络上添加空洞注意力能稍微的改善曲线提取结果, 从表4.2中可以看出, ODS 和 OIS 指标分别提升了 0.6% 和 3% 左右。当单纯添加一层通道注意力时, 虽然相对于基础网络性能略有提高, 但仍不及 DECA 结构, 如图4.9(d) 中所示, 在第二列和第五列图中出现了漏提取等情况, 图中红色框标出。通过第四组实验, 将普通的通道注意力模块 (CA) 替换为高效通道注意力模块 (ECA), 能够带来计算效率的提升, 同时在可视化结果中, 如图4.9(e) 所示, 能够很好的改善 CA 不足, 提升检测结果。

表 4.2 消融实验指标结果

Approachs	SSIM	FSIM	ODS	OIS	AP
-	0.7432	0.7573	0.7673	0.8137	0.7744
DC	0.7336	0.7453	0.7725	0.8384	0.7791
DC+CA	0.8147	0.8679	0.7901	0.8489	0.8641
DC+ECA	0.8244	0.8729	0.7984	0.8369	0.8825
DC+DECA+SFM(Ours)	<b>0.8659</b>	<b>0.8776</b>	0.8048	<b>0.8735</b>	<b>0.9038</b>

针对不同消融实验的配置组别, 可视化结果相近, 人眼较难发现组别之间区别。然而, 在检测的边界中, 结果发现在每一组实验的第四张图中重叠部分, 本文提出的模块对于识别重合曲线结构有着很好提升作用, 最后一组实验提取的曲线结构相比于其他组更加连续和明显。由此可看出, 基于双重通道注意力模块提升了对于曲线微小细节的提取效果。另外, 针对融合模块的部份, 由于多尺度特征融合结构的加入使得 SSIM 和 FSIM 相对于原来的模型有所提升, 这是因为多支路上的融合模块可以充分利用不同尺度特征的互补性, 将它们有机地结合起来, 从而提高了模型对目标的表示能力和区分度。普通的层叠结构只能逐层提取特征, 容易导致信息丢失和冗余, 从而影响检测质量。

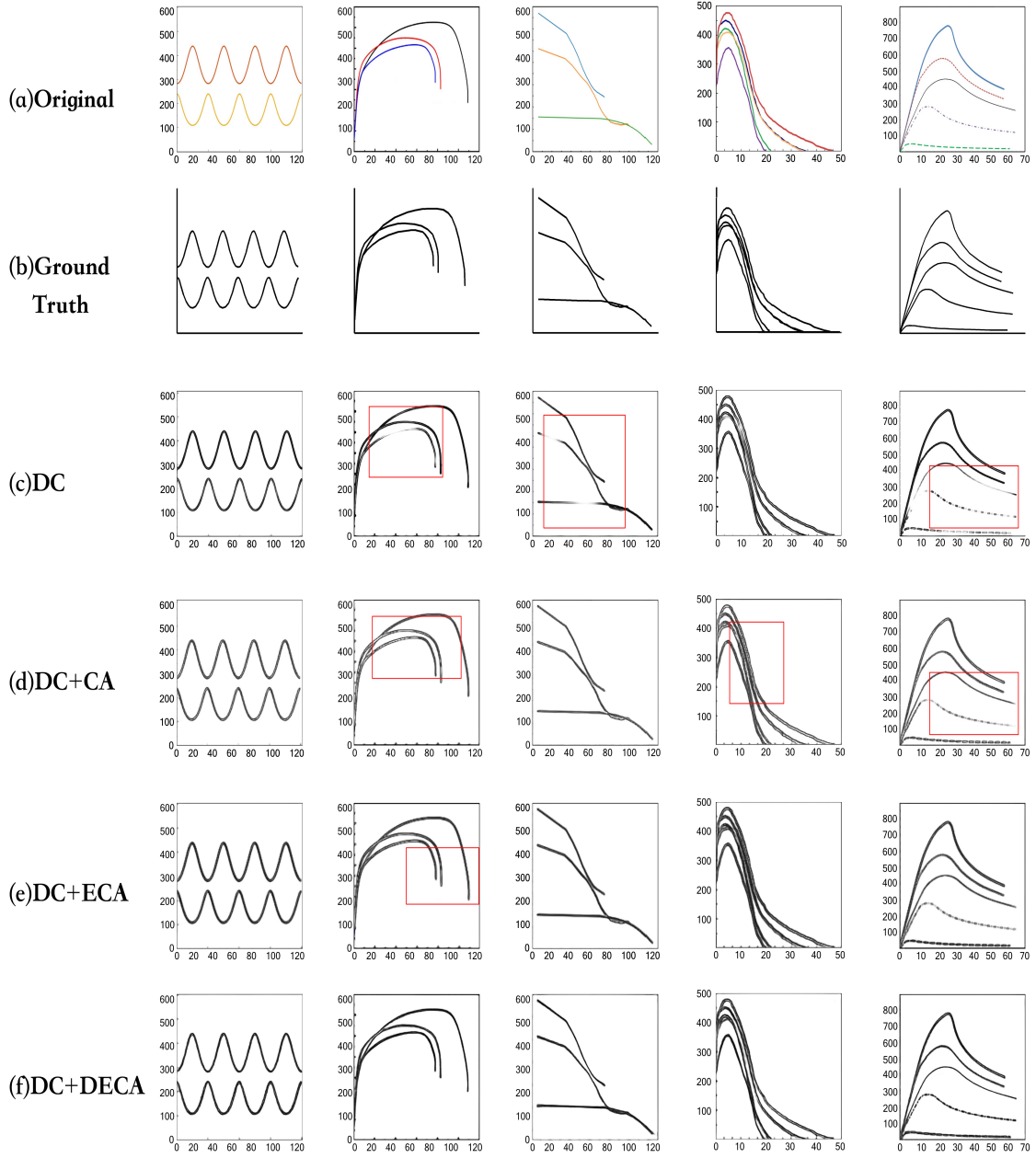


图 4.9 消融实验对比结果展示。

#### 4.2.4 和 DCEN 对比实验

为了进一步探究本章方法和上一章方法 DCEN 的各自优缺点, 本文在 CURVES-DATA 数据集上对比了两种方法的曲线提取结果。

表 4.3 DCACE 和 DCEN 方法对比实验结果

Approachs	SSIM	FSIM	ODS	OIS	AP	参数量	检测速度 (FPS)
DCEN	0.8263	0.8764	<b>0.8122</b>	0.8471	0.8923	30M	0.5
DCACE	<b>0.8659</b>	<b>0.8776</b>	0.8048	<b>0.8735</b>	<b>0.9038</b>	<b>18.3M</b>	6.7

DCACE 网络以较少的可训练参数数量, 在 AP 等性能指标方面已经超越了上一章提出的 DCEN 模型, 而且在检测速度上 DCACE 更有优势, 如图4.10所示, 这表明 DCACE 的网络设计对于曲线提取任务更为有效。总而言之, DCEN 提高了网络在曲线信息提取任务上的性能和稳定性, 但是过多的卷积模块和密集网络结构在曲线数据上容易出现丢失曲线图像中结构化信息和上下文联系的问题, DCACE 关注通道信息, 避免添加过多的卷积操作, 充分利用全局上下文关联的信息, 限制相邻像素之间的标签以加强深度监督, 以双重注意力的方式关注重要特征提升了分割精度, 在解决曲线信息提取任务上, DCACE 的优势更加明显。

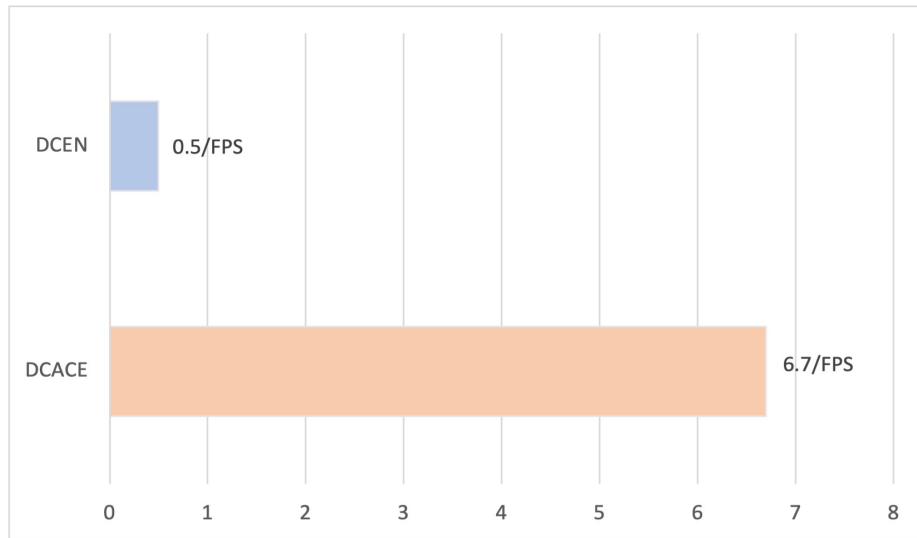


图 4.10 DCEN 和 DCACE 方法的检测速度对比。

### 4.3 曲线数据提取应用设计

本节主要介绍将所论述的曲线提取算法进行封装和开发成桌面端提取软件, 并展示其具体应用。在这一节中, 本文将会讨论如何将曲线提取算法转化为可执行软件, 并通过用户友好的界面呈现提取结果。

### 4.3.1 开发环境和相关工具介绍

本软件使用基于 Python 的桌面软件开发框架 PyQt 开发，为 Windows 操作系统创建图形用户界面 (GUI)。PyQt 是一个跨平台的应用框架，为开发者提供了建立用户界面的必要功能，它专门为应用程序和用户界面开发而设计，可以在各种桌面、嵌入式和移动操作系统上部署，具有良好的封装机制，高度模块化和可重用性，为开发者提供了极大的便利。

PyQt Designer 是 PyQt 库的可视化工具，用于创建 GUI 应用程序的用户界面。它提供易于使用的界面，支持多种窗口小部件、布局 and 对话框，可以方便地导出为 Python 代码并与其他 Python 模块集成。此外，PyQt Designer 还提供属性编辑器，大大提高了 GUI 应用程序的开发效率。

本软件的开发将 PyQt 的各种功能和优势与曲线提取需求相结合，完成后进行了相关测试，整体上达到了预期的效果。

### 4.3.2 软件设计流程

曲线数据提取应用设计是为了实现对曲线坐标图像的数字化信息提取，对科研人员后续进一步分析数据提供辅助作用。结合本文所提算法模型，使用 python 编程语言和 PyQt GUI 开发了桌面端曲线提取软件，具体实施流程图如图4.11所示。

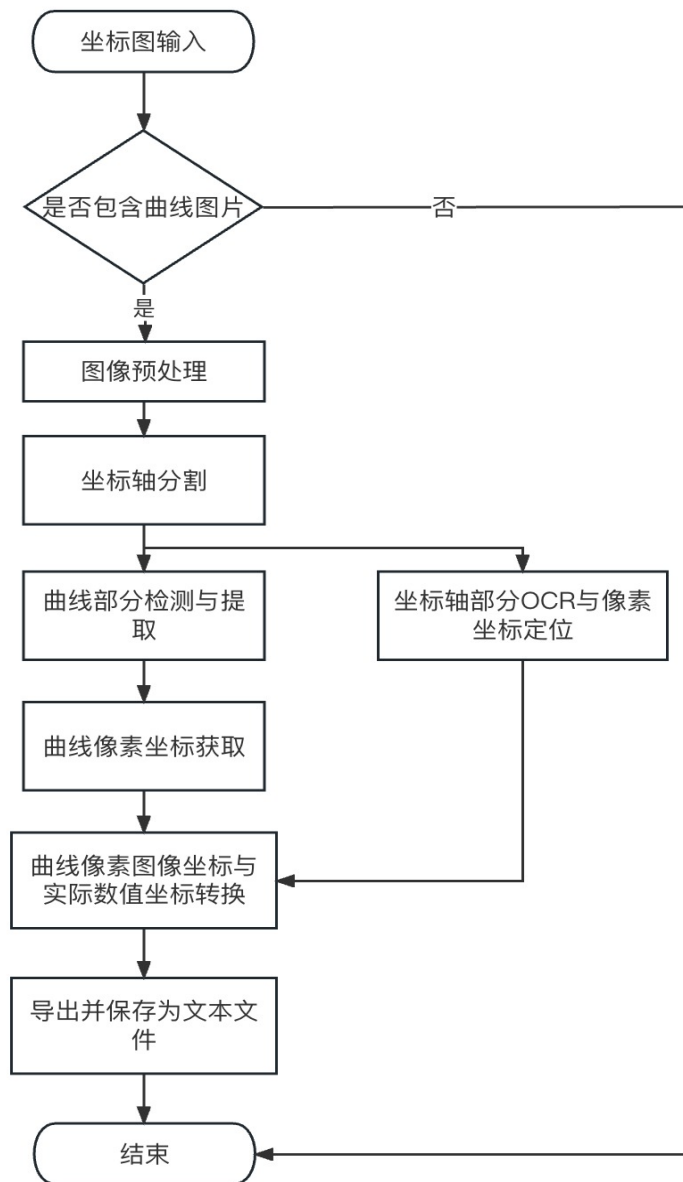


图 4.11 曲线提取软件主流程图

一般的，对于每次读取的图片都会经历以下步骤，首先会通过算法判断图片中是否包含有坐标图片，然后通过分割算法将曲线坐标图分割为具体的 X，Y 坐标轴部分和曲线部分，如图4.12所示。

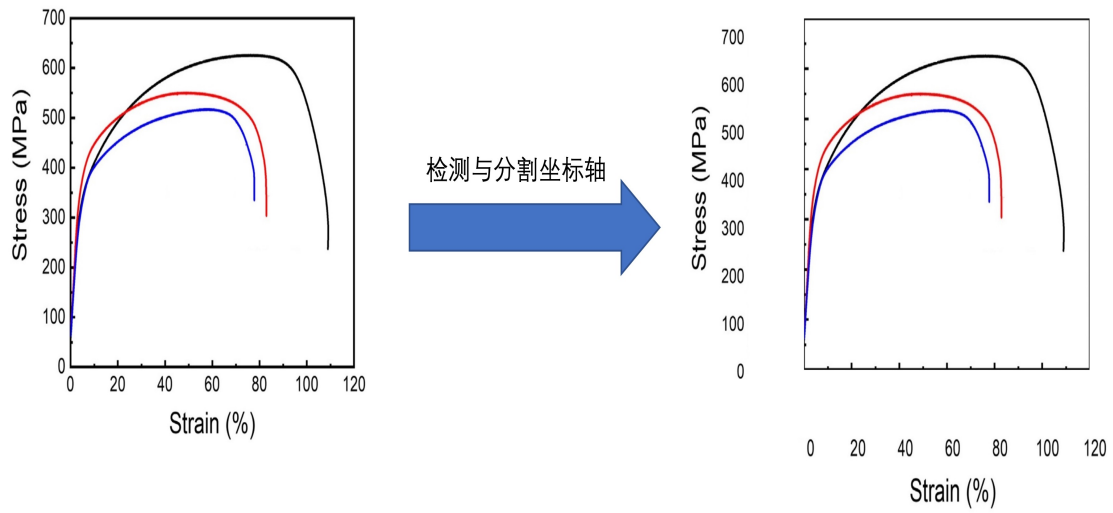


图 4.12 曲线图片分割 x, y 轴结果

紧接着对 X, Y 轴坐标进行 OCR 识别, 获取坐标轴中置信度高的真实数值以及四周红色预选框中的像素坐标位置并排除非数值的预选框, 如图4.13, 4.14所示。

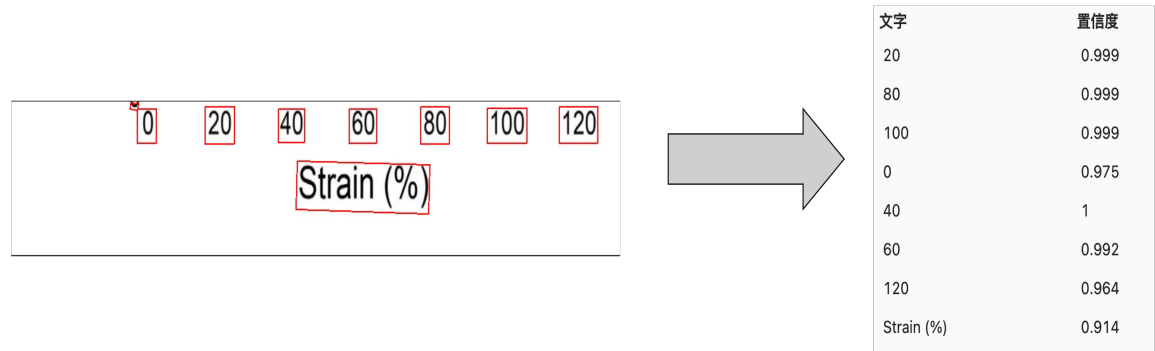


图 4.13 X 轴识别结果



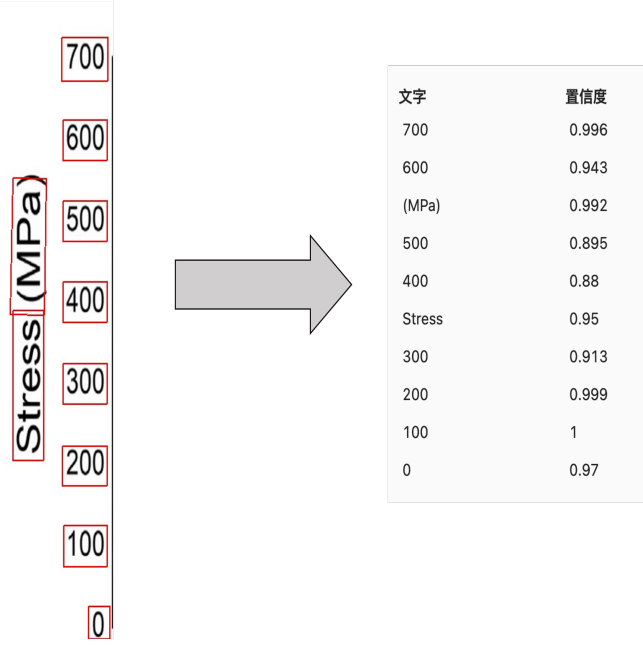


图 4.14 Y 轴识别结果

之后应用 DCACE 算法，获取曲线像素坐标数值并通过公式4.8和4.9进行转换，将曲线的像素坐标转换为真实值坐标。

$$X = X_{min} + pX \frac{X_j - X_i}{pX_j - pX_i} \quad (4.8)$$

$$Y = Y_{min} + pY \frac{Y_j - Y_i}{pY_j - pY_i} \quad (4.9)$$

其中  $X$ ,  $Y$  代表最终转换过来的所需要的真实曲线点数据坐标值； $pX$ ,  $pY$  为在曲线上每个点的像素层次的坐标； $X_{min}$ ,  $Y_{min}$  分别为  $X$ ,  $Y$  轴上从原点开始的第一个检测到的数字坐标数据的数据值。在对坐标轴上的字符数值进行获取的时候需要进行字符检测与识别，同时字符检测会对每个数字字符绘出矩形检测框， $pX_i$ ,  $pY_j$  代表前后两个坐标轴数值检测过程中矩形框的左上角顶点像素的横坐标， $X_i$ ,  $Y_j$  为坐标检测的具体数值。对于公式4.8的  $\frac{X_j - X_i}{pX_j - pX_i}$  部分，表示单位像素上对应的真实数据数值。公式  $pX \frac{X_j - X_i}{pX_j - pX_i}$  部分代表，由像素坐标转换成真实数据坐标之后的具体数值。以此计算出每个曲线上像素点的相对坐标，再加上起始点的坐标  $X_{min}$ ，即可获得每个数据点转换后真实的坐标。对于公式4.9的  $Y$  轴像素坐标转换计算同理。

### 4.3.3 应用软件介绍与效果展示

本文用 PyQt 框架封装前文提出的曲线检测模型，并实施坐标数值转换流程。开发完成的软件主界面如图 4.15 所示，共四个区域。中间空白部分为曲线图片读取展示区域，显示所需要提取的曲线，可以通过右上角选择文件输入。右侧清除按钮可清除当前选择的曲线图片并重新选取图片，点击导出按钮即可完成曲线数据提取。图 4.16 是图片读取后展示效果，图 4.17 展示了中间曲线分离结果，图 4.18 是最终的提取结果。



图 4.15 曲线提取程序主界面

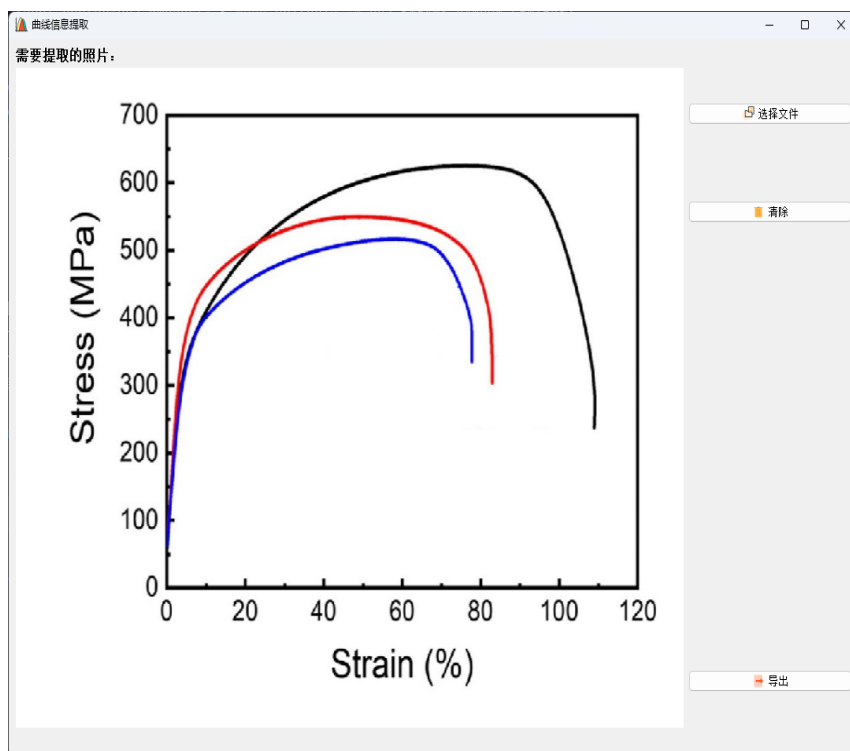


图 4.16 曲线提取程序读取图片界面

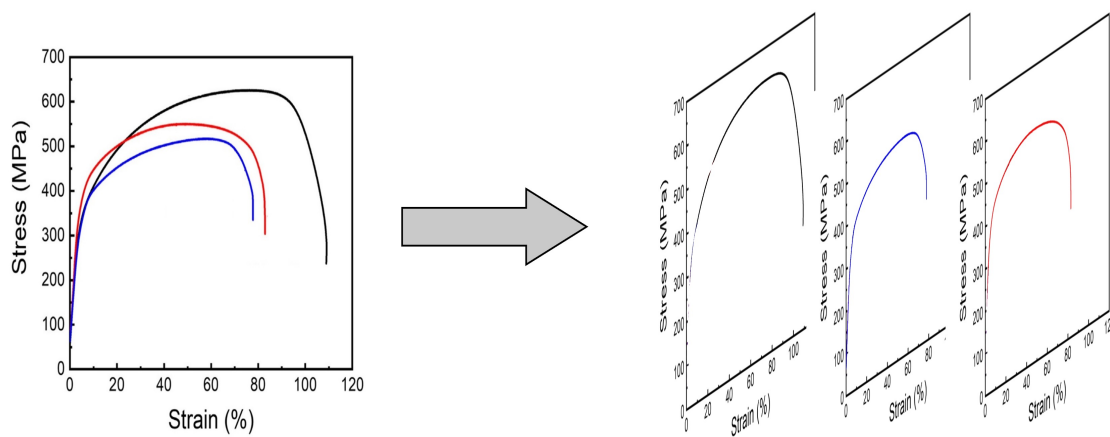


图 4.17 曲线中间分离结果

黑色曲线		蓝色曲线		红色曲线	
X	Y	X	Y	X	Y
3.183435804701620	298.12927756654000	0.11855334538879000	83.47832699619780	1.367209162145870	213.1802281368820
3.6091139240506300	310.7414448669200	-0.08577215189873040	67.56258555133090	1.7191030741410500	237.7536121673000
4.034792043399640	323.1908745247150	0.45909584086799500	107.56640412819100	1.4807233273056000	226.0365019011410
4.545605786618450	335.39619771863100	0.20368896925858600	95.68365019011410	1.9064014466546100	251.78973384030400
5.141555153707050	348.2117870722430	0.7712597950572620	130.91634980988600	2.0766726943942100	262.5304182509510
5.652368896925860	358.342205323194	0.5442314647377910	119.60608365019000	2.360458107293550	276.1596958174910
6.546292947558770	372.98859315589400	0.8847739602169970	144.74904942965800	2.473972272453290	289.422813688213
7.4827848101265800	385.0718631178710	1.1061265822784800	157.49140684410600	2.8428933092224300	303.6623574144490
8.461844484629300	395.9346007604560	1.4320743993800100	177.77316675719700	3.159111340738830	321.3077675176540
9.653743218806510	407.0937533948940	1.1401808318264000	165.49809885931600	3.467221217600970	335.64030418251000
10.760506329113900	418.31102661597000	1.7726168948592100	202.0443237371000	3.8645207956600300	350.5307984790880
11.867269439421300	428.8541010320480	1.6509945750452100	188.93231939163500	4.375334538878840	364.77034220532300
13.05916817359860	439.6645301466600	2.1192405063291100	220.30000000000000	4.8861482820976500	377.4638783269960
14.336202531645600	450.6144486692020	2.417215189873420	240.8456273764260	5.482097649186260	390.13707224334600
15.69837251356240	461.5382129277570	2.7577576853526200	257.2821292775670	6.102371480237670	401.79315589353600
17.145678119349000	472.47553865652700	3.183435804701620	275.34600760456300	6.759132007233270	414.2914068441070
18.678119349005400	483.16197718631200	3.438842676311030	290.1144486692020	7.610488245931290	425.4226615969580
20.380831826401400	493.976633252679	3.7793851717902300	301.0992395437260	8.632115732368900	436.10755024443200
21.913273056057900	501.87680608365000	4.20506329113924	313.91482889733800	9.909150090415910	446.58669201520900
24.637613019891500	519.5412374697550	4.715877034358040	326.60836501901100	11.441591320072300	457.4006083650190
26.51059674502710	527.7520912547530	5.251015241539660	339.5111352525800	13.229439421338200	468.1901140684410
...	...	...	...	...	...
...	...	...	...	...	...
102.6218444846290	472.7467680608370	77.52386256781190	380.7023574144490	82.61497287522600	335.8844106463880
103.15698269181100	461.076154263987	77.50683544303800	368.5946768060840	82.61497287522600	325.1437262357420

图 4.18 曲线提取结果图

## 4.4 本章小节

本章针对第三章模型中堆叠过多的卷积模块而出现曲线图像中结构化信息和上下文联系丢失的问题，提出了一种基于双重通道注意力机制的曲线提取方法，结合通道权重分配机制和阶段特征融合方法来充分利用全局上下文关联信息。具体而言，双重通道注意力模块可以在通道维度上对特征进行加权，从而强化有用特征的贡献，而多尺度级联特征融合则利用不同尺度的特征来获取更全面的信息，增强模型对不同尺度目标的识别能力。

此外，本章还采用了局部特征和全局长程上下文的融合策略，从而提高特征的有效性和选择性，增强模型的精度和鲁棒性。实验结果表明，本章方法 DCACE 相比于第三章的模型在评价指标和视觉效果上均有所提升，同时该方法的参数量也大幅

减少，提高了训练和推理速度，这进一步验证了该方法的高效性和实用性。最后，将 DCACE 曲线提取算法使用 PyQt 框架进行封装，并进行二次开发将曲线提取算法模块与用户界面进行集成，创建一个功能完善的曲线提取软件，实现用户友好的操作和实时曲线提取功能。

## 第五章 总结与展望

### 5.1 总结

图像是一种简单且高效的信息表达方式，在科学文献中存在着大量的数据支撑的曲线图像，将实验数据、观测数据、预测数据等可视化的呈现出来。曲线图是一种常见的图像类型，用于表示数据之间的关系，例如时间序列数据、变量之间的关系等。因此将文献中的曲线图形准确且完整的提取出来非常重要，有利于下游任务去分析相应的数据为各种后续的文献信息挖掘，内容提取分析提供基础。

曲线提取是科学文献中实验分析的基础任务，它可以将曲线图像从文献中精确地提取出来，并转化为数字形式，以便进行后续的数据分析和信息挖掘。曲线提取需要应对图像质量不一、曲线形状复杂、噪声干扰等问题，因此需要使用先进的图像处理和机器学习算法来实现。

准确的曲线提取有助于科研人员快速获取文献中的实验结果、观测结果和预测结果，并进行数据分析和模型构建。同时，曲线提取也有助于自动化文献信息抽取和文本摘要生成，为科研人员提供更快速、更精准的科学数据分析和研究工具。曲线提取是科学文献中的基础任务，它可以提高数据分析和信息挖掘的效率和准确性，有助于推动科学研究的发展。本文针对曲线图像的特点，结合现有的科学理论知识分别提出两种曲线提取的方法并实现相关软件开发。本文的研究成果主要如下：

(1) 针对采用边缘检测方法提取曲线时，曲线获取结果十分粗糙、精细化程度较低等问题，第三章采用了 DexiNed 的密集网络结构和 CED 的后向细化路径模块以及尺度增强模块。这些技术不仅可以准确地检测曲线，还可以提供更加清晰的曲线边缘和细节。本文使用了专门为曲线检测设计的数据集进行测试，结果表明所提出的相关检测指标与视觉效果均有显著提高。换言之，本文的方法能够提高曲线检测的精度和效果，从而更好地满足实际应用需求。

(2) 针对曲线图像中全局上下文与通道信息忽略的问题，设计双重通道注意力模块来加权特征，通过局部特征和全局长程上下文的融合，提高了特征的有效提取和筛选，减少像素级信息和通道信息的损失，进而提高了模型的精度和鲁棒性。同时，采用多尺度级联特征融合来获得更全面的信息，增强模型对不同尺度曲线的识

别能力。

(3) 本文将所提模型针对性的开发了相关软件,以便轻松提取相关曲线数据。该软件具有高精准度和高效性,能够对高清晰度的曲线图像进行处理,并可以自动化地进行曲线数据提取处理,方便快捷地获取所需的数据信息。

本文对两种方法进行了对比和消融实验,结果表明两种方法在曲线提取任务上均表现出较高的精度和鲁棒性,具有较好的泛化性和通用性,对未来的曲线提取和相关领域的发展提供了一定的贡献和帮助。

## 5.2 展望

本文主要探讨了基于深度学习方法在曲线提取任务中的应用,提出了两种网络模型,这些模型可应用于不同类型的数据,并能够高效地解决曲线提取问题。实验结果充分证明了基于深度学习技术的曲线提取方法是实现自动化曲线提取的优秀工具。为了更好地促进深度学习技术在曲线提取任务中的应用,作者认为有以下三个方面值得深入探究:

(1) 首先,本文第三章在曲线图像上采样方面已经取得了不错的效果,但是仍然存在一些可以改进的问题。特别是,当前的上采样方法可能会导致一些图像细节丢失,从而影响图像的还原质量。为了更精确地还原图像并保留更多的细节,可以探索一些更合适的上采样技术,例如基于深度学习的超分辨率方法,这些方法可以从低分辨率图像中重建出高分辨率图像,并在这个过程中保留更多的细节信息。未来的研究方向还可以包括在不同类型的数据集上优化这种方法,并将其扩展到曲线语义分割等更加复杂的场景中。

(2) 其次,尽管深层卷积神经网络能够更好地适应复杂数据分布,但是模型的计算复杂度和收敛速度仍然存在较大的改进空间。在本文所研究的曲线提取任务中,边缘检测是一种基于二分类的检测算法。然而,考虑到实际应用需求,使用多分类算法可能更加合适。进一步地,结合深度学习中的边缘检测知识,可以探索更加高效的模型结构和训练算法,以提高模型的计算效率和收敛速度,并实现更准确和鲁棒的曲线边缘检测。此外,未来的研究可以将多分类算法应用于更加复杂的曲线提取任务中,以实现更高级别的图像分析和理解。

(3) 最后,未来还可以引入多模态信息进行曲线坐标图数值提取。除了曲线坐标

图本身的信息，曲线的颜色、线条宽度、文字标注等多种信息，都可以为曲线提取提供依据，更有利于提高数值提取的准确率和鲁棒性。

以上内容的深入研究能够进一步完善深度学习技术在现实生活中的应用，更能提高曲线提取效果，为数据分析提供坚实基础，提升科研工作者的分析和研究的效率。



## 参考文献

- [1] SUN J. Image edge detection based on relative degree of grey incidence and sobel operator[C]// Artificial Intelligence and Computational Intelligence: 4th International Conference, AICI 2012, Chengdu, China, October 26-28, 2012. Proceedings 4. Springer, 2012: 762-768.
- [2] ZHANG W, ZHAO Y, BRECKON T P, et al. Noise robust image edge detection based upon the automatic anisotropic gaussian kernels[J]. Pattern Recognition, 2017, 63: 193-205.
- [3] KONISHI S, YUILLE A L, COUGHLAN J M, et al. Statistical edge detection: Learning and evaluating edge cues[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(1): 57-74.
- [4] RAKESH R R, CHAUDHURI P, MURTHY C. Thresholding in edge detection: a statistical approach[J]. IEEE Transactions on image processing, 2004, 13(7): 927-936.
- [5] ANDO S. Image field categorization and edge/corner detection from gradient covariance[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(2): 179-190.
- [6] PIÓRKOWSKI A. A statistical dominance algorithm for edge detection and segmentation of medical images[C]//Information Technologies in Medicine: 5th International Conference, ITIB 2016 Kamień Śląski, Poland, June 20-22, 2016 Proceedings, Volume 1. Springer, 2016: 3-14.
- [7] LIU Y, XIE Z, LIU H. An adaptive and robust edge detection method based on edge proportion statistics[J]. IEEE Transactions on Image Processing, 2020, 29: 5206-5215.
- [8] WILLIAMS I, BOWRING N, SVOBODA D. A performance evaluation of statistical tests for edge detection in textured images[J]. Computer Vision and Image Understanding, 2014, 122: 115-130.
- [9] SOBEL I. Camera models and machine perception aim-21[M]. Stanford Artificial Intelligence Lab, Palo Alto, CA, 1970.
- [10] MARR D, HILDRETH E. Theory of edge detection[J]. Proceedings of the Royal Society of London. Series B. Biological Sciences, 1980, 207(1167): 187-217.
- [11] CANNY J. A computational approach to edge detection[J]. IEEE Transactions on pattern analysis and machine intelligence, 1986(6): 679-698.

- [12] DING L, GOSHTASBY A. On the canny edge detector[J]. Pattern recognition, 2001, 34(3): 721-725.
- [13] JACOB M, UNSER M. Design of steerable filters for feature detection using canny-like criteria[J]. IEEE transactions on pattern analysis and machine intelligence, 2004, 26(8): 1007-1019.
- [14] XU Q, VARADARAJAN S, CHAKRABARTI C, et al. A distributed canny edge detector: algorithm and fpga implementation[J]. IEEE Transactions on Image Processing, 2014, 23(7): 2944-2960.
- [15] JING J, LIU S, LIU C, et al. A novel decision mechanism for image edge detection[C]//Intelligent Computing Theories and Application: 17th International Conference, ICIC 2021, Shenzhen, China, August 12–15, 2021, Proceedings, Part I 17. Springer, 2021: 274-287.
- [16] MALIK J, BELONGIE S, LEUNG T, et al. Contour and texture analysis for image segmentation [J]. International journal of computer vision, 2001, 43: 7-27.
- [17] MARTIN D R, FOWLKES C C, MALIK J. Learning to detect natural image boundaries using local brightness, color, and texture cues[J]. IEEE transactions on pattern analysis and machine intelligence, 2004, 26(5): 530-549.
- [18] REN X. Multi-scale improves boundary detection in natural images[C]//Computer Vision–ECCV 2008: 10th European Conference on Computer Vision, Marseille, France, October 12-18, 2008, Proceedings, Part III 10. Springer, 2008: 533-545.
- [19] ARBELAEZ P, MAIRE M, FOWLKES C, et al. Contour detection and hierarchical image segmentation[J]. IEEE transactions on pattern analysis and machine intelligence, 2010, 33(5): 898-916.
- [20] PAYET N, TODOROVIC S. Sledge: Sequential labeling of image edges for boundary detection [J]. International journal of computer vision, 2013, 104: 15-37.
- [21] YANG J, PRICE B, COHEN S, et al. Object contour detection with a fully convolutional encoder-decoder network[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 193-202.

- [22] ABDULRAHMAN H, MAGNIER B, MONTESINOS P. A new objective supervised edge detection assessment using hysteresis thresholds[C]//New Trends in Image Analysis and Processing—ICIAP 2017: ICIAP International Workshops, WBICV, SSPandBE, 3AS, RGBD, NIVAR, IWBAAS, and MADiMa 2017, Catania, Italy, September 11-15, 2017, Revised Selected Papers 19. Springer, 2017: 3-14.
- [23] FLORES-VIDAL P A, OLASO P, GÓMEZ D, et al. A new edge detection method based on global evaluation using fuzzy clustering[J]. *Soft Computing*, 2019, 23: 1809-1821.
- [24] RAY K. Unsupervised edge detection and noise detection from a single image[J]. *Pattern Recognition*, 2013, 46(8): 2067-2077.
- [25] LIU C, WANG C, DAI R. Text detection in images based on unsupervised classification of edge-based features[C]//Eighth International Conference on Document Analysis and Recognition (ICDAR'05). IEEE, 2005: 610-614.
- [26] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.
- [27] DOLLAR P, TU Z, BELONGIE S. Supervised learning of edges and object boundaries[C]//2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06): volume 2. IEEE, 2006: 1964-1971.
- [28] DORIGO M, MANIEZZO V, COLORNI A. Ant system: optimization by a colony of cooperating agents[J]. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 1996, 26(1): 29-41.
- [29] ETEMAD K, CHELAPPA R. A neural network based edge detector[C]//IEEE International Conference on Neural Networks. IEEE, 1993: 132-137.
- [30] LIM J J, ZITNICK C L, DOLLÁR P. Sketch tokens: A learned mid-level representation for contour and object detection[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2013: 3158-3165.
- [31] EITZ M, HAYS J, ALEXA M. How do humans sketch objects?[J]. *ACM Transactions on graphics (TOG)*, 2012, 31(4): 1-10.

- [32] DOLLÁR P, ZITNICK C L. Fast edge detection using structured forests[J]. IEEE transactions on pattern analysis and machine intelligence, 2014, 37(8): 1558-1570.
- [33] GEURTS P, ERNST D, WEHENKEL L. Extremely randomized trees[J]. Machine learning, 2006, 63: 3-42.
- [34] GANIN Y, LEMPITSKY V. -fields: neural network nearest neighbor fields for image transforms [C]//Computer Vision—ACCV 2014: 12th Asian Conference on Computer Vision, Singapore, Singapore, November 1-5, 2014, Revised Selected Papers, Part II. Springer, 2015: 536-551.
- [35] XIE S, TU Z. Holistically-nested edge detection[C]//Proceedings of the IEEE international conference on computer vision. 2015: 1395-1403.
- [36] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[A]. 2014.
- [37] HE J, ZHANG S, YANG M, et al. Bdcn: Bi-directional cascade network for perceptual edge detection[J]. IEEE transactions on pattern analysis and machine intelligence, 2020, 44(1): 100-113.
- [38] DENG J, DONG W, SOCHER R, et al. Imagenet: A large-scale hierarchical image database[C]//2009 IEEE conference on computer vision and pattern recognition. Ieee, 2009: 248-255.
- [39] LIU Y, CHENG M M, HU X, et al. Richer convolutional features for edge detection[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 3000-3009.
- [40] WANG Y, ZHAO X, HUANG K. Deep crisp boundaries[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 3892-3900.
- [41] WANG N, SONG Y, MA C, et al. Unsupervised deep tracking[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019: 1308-1317.
- [42] REN X, BO L. Discriminatively trained sparse code gradients for contour detection[C]//Proceedings of the 25th International Conference on Neural Information Processing Systems—Volume 1. 2012: 584-592.
- [43] CHEN X, LIN Q, KIM S, et al. Smoothing proximal gradient method for general structured sparse learning[A]. 2012.
- [44] HEARST M A, DUMAIS S T, OSUNA E, et al. Support vector machines[J]. IEEE Intelligent Systems and their applications, 1998, 13(4): 18-28.

- [45] ISOLA P, ZORAN D, KRISHNAN D, et al. Crisp boundary detection using pointwise mutual information[C]//Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part III 13. Springer, 2014: 799-814.
- [46] FANO R M. Transmission of information: A statistical theory of communications[J]. American Journal of Physics, 1961, 29(11): 793-794.
- [47] XIA X, KULIS B. W-net: A deep model for fully unsupervised image segmentation[A]. 2017.
- [48] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [49] LIU S, DENG W. Very deep convolutional neural network based image classification using small training sample size[C]//2015 3rd IAPR Asian conference on pattern recognition (ACPR). IEEE, 2015: 730-734.
- [50] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 4700-4708.
- [51] SHI W, CABALLERO J, THEIS L, et al. Is the deconvolution layer the same as a convolutional layer?[A]. 2016.
- [52] SHI W, CABALLERO J, HUSZÁR F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 1874-1883.
- [53] ASHISH V. Attention is all you need[J]. Advances in neural information processing systems, 2017, 30: I.
- [54] GUO M H, XU T X, LIU J J, et al. Attention mechanisms in computer vision: A survey[J]. Computational Visual Media, 2022, 8(3): 331-368.
- [55] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 7132-7141.
- [56] WANG X, GIRSHICK R, GUPTA A, et al. Non-local neural networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 7794-7803.
- [57] VASWANI A, SHAZEER N, PARMAR N, et al. Advances in neural information processing systems. 2017[J]. Attention is all you need, 2017: 5998-6008.

- [58] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity[J]. IEEE transactions on image processing, 2004, 13(4): 600-612.
- [59] ZHANG L, ZHANG L, MOU X, et al. Fsim: A feature similarity index for image quality assessment [J]. IEEE transactions on Image Processing, 2011, 20(8): 2378-2386.
- [60] ZIOU D, TABBONE S, et al. Edge detection techniques-an overview[J]. Pattern Recognition and Image Analysis C/C of Raspoznavaniye Obrazov I Analiz Izobrazhenii, 1998, 8: 537-559.
- [61] POMA X S, RIBA E, SAPPA A. Dense extreme inception network: Towards a robust cnn model for edge detection[C]//Proceedings of the IEEE/CVF winter conference on applications of computer vision. 2020: 1923-1932.
- [62] WIBISONO J K, HANG H M. Traditional method inspired deep neural network for edge detection [C]//2020 IEEE International Conference on Image Processing (ICIP). IEEE, 2020: 678-682.
- [63] SU Z, LIU W, YU Z, et al. Pixel difference networks for efficient edge detection[C]//Proceedings of the IEEE/CVF international conference on computer vision. 2021: 5117-5127.
- [64] SORIA X, POMBOZA-JUNEZ G, SAPPA A D. Ldc: Lightweight dense cnn for edge detection [J]. IEEE Access, 2022, 10: 68281-68290.
- [65] ZHANG H, ZU K, LU J, et al. Epsanet: An efficient pyramid squeeze attention block on convolutional neural network[C]//Proceedings of the Asian Conference on Computer Vision. 2022: 1161-1177.

## 作者在攻读硕士学位期间发表的论文与研究成果

### 发表的学术论文

1. Lightweight and Crisp Curves Extract Network. 2023 4nd International Conference on Computer Information and Big Data Applications Engineering (CIBDA 2023) (EI 检索会议)
2. 发明名称：一种基于图像处理的文献曲线分离与坐标信息提取方法；发明人：韩越兴、陈侨川、王迎港、张瑞、潘登、夏锦桦、张家旺；申请号或专利号：202111490372.3
3. 软件名称：曲线坐标图像信息提取软件 V1.0；开发人：陈侨川，王迎港，韩越兴；登记号：2022SR1365016

## 作者在攻读硕士学位期间所作的项目

1. 项目来源：国家重点研发计划项目

项目名称：材料基因组工程专用数据库平台建设与示范应用

项目编号：2018YFB0704400

执行期限：2018.07-2022.06

2. 项目来源：之江实验室科研攻关项目

项目名称：智能计算材料平台建设与示范应用

项目编号：2021PE0AC02

执行期限：2021.11-2024.11



## 致 谢

在完成我研究生学业的岁月里，有许多人和事让我受益匪浅。在此，我向这些关心、支持和帮助过我的人们致以最深切的谢意。

首先，我要感谢我的导师陈老师和韩老师。陈老师不仅传授给我专业知识，而且还教会了我做学问的方法和思路，他耐心地指导我，在学习和生活上都给予了我很大的帮助。韩老师也是我的良师益友，每次与他交流都可以收获颇丰。因为她的指导，我的研究才能够更加深入。

此外，我还要感谢同期的同学们。一路走来，我们相互鼓励、相互帮助，共同度过了繁忙但充实的求学岁月。感谢他们尽可能分担我的压力，给我提供了一个愉快而难忘的学习环境，让我度过了最好的一段时光。

最后，我要特别感谢我的家人和朋友。在我研究生求学期间，他们一直是我的坚实后盾，给予了我无微不至的照顾和支持。家人们默默地为我提供物质和精神上的支持，让我全身心投入到研究中，朋友们也鼓励我坚持自己的信仰，永不言败。

在这里，我要向所有给予我帮助和关心的人表达我最深情的感谢。正是你们，让我走进了学问的殿堂，让我成为一个更加自信成熟的人。衷心的感谢每一位给予过我帮助的人，愿大家一直生活在幸福和健康之中。