# 摘要

本文主要是针对基于深度学习对鲕粒灰岩的地质显微图像的分割问题进行了较为深入的研究，在地质学显微图像研究领域，对显微图像的识别和记录是一个繁琐单调的过程，本文用深度学习的方法对地质显微图像中的鲕粒灰岩显微图像进行了处理，实现了对鲕粒灰岩显微图像的全自动分割，能够准确将传统方法无法识别出的鲕粒灰岩进行精确识别，提高了识别的准确率和识别精度，且在世界范围内暂时没有深度学习在此领域方向上应用的先例，因此本文将有助于丰富深度学习研究的应用领域，在地质显微图像分割方面有着重要的应用价值。

深度学习各个领域都有着广泛的应用前景，并且近些年来在如自动驾驶、手机的人脸识别、高铁站的安检系统等都有了较为显著的成果诞生，而当前在地质显微图像分割领域，虽然有人提出过类似的构想但却暂时没有得到实现，本文对此进行了较为深入的研究。通过大量的实验分析，最终选择了由google提供的Tensorflow框架进行开发，使用PS等工具对LeiCa电子显微镜下拍摄的鲕粒灰岩薄片显微图像进行了数据的加工处理，人工标注出了供训练使用的大量数据集，并且用垂直、水平、水平垂直翻转等对数据集进行了数据增强，之后利用BiSeNet卷积神经网络对数据集进行卷积、池化、全连接以及反向传播等过程，通过数百轮的训练得出了一个准确率在95%以上的鲕粒灰岩分割模型。

通过BiSeNet卷积神经网络训练出的网络模型能对测试数据进行较为准确的预测，可以有效分割出预测数据集中的鲕粒灰岩，并且针对预测图像可以与原图做轮廓对比，同时还能够计算一些关于鲕粒的参数，可以完成对鲕粒灰岩的进一步研究分析并判断其应用价值。

**关键词：**地质显微图像分割；深度学习；鲕粒灰岩；卷积神经网络

# Abstract

In this paper, the segmentation of Oolitic Limestone micro-image based on depth learning is studied in depth. In the field of geological micro-image research, the recognition of micro-image is a tedious and monotonous process. In this paper, the Oolitic Limestone micro-image in geological micro-image is processed by depth learning method, which reduces the recognition process of oolitic limestone. This paper will help enrich the research and application fields of in-depth learning, and has important application value in geological micro-image segmentation.

Deep learning has a broad application prospect in various fields, and in recent years, it has made remarkable achievements in autonomous driving, face recognition of mobile phones, security check system of high-speed railway stations and so on. At present, in the field of geological microscopic image segmentation, although some people have put forward similar ideas but have not been realized. This article has carried on the more thorough research to this problem. Through a lot of experimental analysis, and ultimately chose Tensorflow framework for development provided by Google, use tools such as photoshop for LeiCa oolitic limestone chip under the electron microscope microscopic image processing of the data, marking out for training a large number of data sets, and with a vertical, horizontal and vertical flip horizontal enhanced data set for the data. After that, I used BiSeNet convolutional neural network to perform the processes of convolution, pooling, full connection and back propagation on the data set. Through hundreds of rounds of training, an oolitic limestone segmentation model with an accuracy rate of more than 95% is obtained.

The network model trained by BiSeNet neural network can predict the test data more accurately, and can effectively identify the Oolitic Limestone in the forecasting data set. The forecasting image can be compared with the original contour, and some parameters about Oolitic Limestone can be calculated. It can complete the further research and analysis of Oolitic Limestone and judge its application value.

**Key words:** geological microscopic image segmentation; Deep learning; Oolitic limestone; Convolutional neural network

# 目录

[1. 绪论 1](#_Toc11028223)

[1.1 课题的研究背景 1](#_Toc11028224)

[1.2 课题研究的意义以及问题的提出 2](#_Toc11028225)

[1.2.1 课题研究的目的与意义 2](#_Toc11028226)

[1.2.2 问题的提出 3](#_Toc11028227)

[1.3 论文研究的主要内容 4](#_Toc11028228)

[2. 图像分割算法概述 5](#_Toc11028229)

[2.1 图像分割的定义 5](#_Toc11028230)

[2.2 传统图像分割方法 5](#_Toc11028231)

[2.2.1 边缘分割算法 5](#_Toc11028232)

[2.2.2 基于图的分割算法 6](#_Toc11028233)

[2.3 基于深度学习的语义分割算法 7](#_Toc11028234)

[2.4 本章小结 7](#_Toc11028235)

[3. 实验软件工具及介绍 9](#_Toc11028236)

[3.1 实验环境概述 9](#_Toc11028237)

[3.2 PS工具概述 9](#_Toc11028238)

[3.3 Tensorflow工具概述 10](#_Toc11028239)

[3.4 Tensorflow框架的安装 10](#_Toc11028240)

[3.5 本章小结 14](#_Toc11028241)

[4. 鲕粒灰岩显微图像数据集的制作 15](#_Toc11028242)

[4.1 鲕粒灰岩显微图像的获取 15](#_Toc11028243)

[4.2 鲕粒灰岩数据的采集 17](#_Toc11028244)

[4.3 鲕粒灰岩数据的标注 18](#_Toc11028245)

[4.4 鲕粒灰岩数据集的数据增强 21](#_Toc11028246)

[4.5 本章小结 23](#_Toc11028247)

[5. 深度神经网络模型的构建 25](#_Toc11028248)

[5.1 卷积神经网络基本组成结构 25](#_Toc11028249)

[5.1.1 卷积层 25](#_Toc11028250)

[5.1.2 池化层 26](#_Toc11028251)

[5.1.3 激励层 28](#_Toc11028252)

[5.2 卷积神经网络中的前向与反向传播算法 29](#_Toc11028253)

[5.2.1 前向传播算法 29](#_Toc11028254)

[5.2.2 反向传播算法 30](#_Toc11028255)

[5.3 端到端卷积神经网络的构建 30](#_Toc11028256)

[5.3.1 ResNet深度神经网络概述 30](#_Toc11028257)

[5.3.2 深度神经网络的基本原理简介 32](#_Toc11028258)

[5.4 BiSeNet双向分割神经网络的构建 35](#_Toc11028259)

[5.4.1 BiSeNet网络架构 35](#_Toc11028260)

[5.4.2 空间路径（Spatial Path） 37](#_Toc11028261)

[5.4.3 环境路径（Context Path） 37](#_Toc11028262)

[5.4.4 注意力优化模块（ARM） 38](#_Toc11028263)

[5.4.5 特征融合模块（FFM） 38](#_Toc11028264)

[5.4.6 损失函数 39](#_Toc11028265)

[5.5 本章小结 39](#_Toc11028266)

[6. 鲕粒灰岩显微图像分割结果及实验分析 41](#_Toc11028267)

[6.1 实验分割结果 41](#_Toc11028268)

[6.2 识别准确率分析 41](#_Toc11028269)

[6.2.1 检测的准确率（test\_accuracy） 41](#_Toc11028270)

[6.2.2 平均交并比（mean iou） 42](#_Toc11028271)

[6.2.3 损失率（loss） 43](#_Toc11028272)

[6.2.4 鲕粒预测效果展示 43](#_Toc11028273)

[6.3 相关参数计算结 44](#_Toc11028274)

[6.4 本章小结 45](#_Toc11028275)

[7. 总结与展望 47](#_Toc11028276)

[7.1 论文工作总结 47](#_Toc11028277)

[7.2 进一步研究方向 48](#_Toc11028278)

[致谢 49](#_Toc11028279)

[参考文献 51](#_Toc11028280)

# 绪论

## 课题的研究背景

在地质学领域，学者们常常要对岩石进行仔细鉴别，并测定其中的一些数据以确认一个地区岩石的种类、品性和分布情况。而传统的工作方法要求专业人士对其进行测定鉴别，有一部分较小的岩石就需要制成显微薄片在拍摄显微图像后计算数据并进行观察分析，之后将其中要研究岩石从众多杂志中分离出来，达到对其中某种类别的岩石进行鉴别的效果。而本文要研究的鲕粒灰岩便是需要在显微图像下进行鉴别的。

鲕粒灰岩主要组成成分为碳酸钙，同时还会含有铁矿、白云石、赤铁矿、磷酸盐以及燧石，鲕粒灰岩坚固的特性主要是因为其中含有白云岩还有燧石，其形状多为近似的同心圆或椭圆状，鲕粒灰岩由于粒径较小，在0.2~1.22毫米左右，所以经常被用来制作家庭水族箱，同时也因为其粒径较小，常常需要在显微薄片下进行观察[1] 。地质显微图像中常规的鉴别鲕粒的方法主要是通过人工在电子显微镜下观察，然后对显微薄片中的鲕粒颗粒进行逐一鉴别，并计算相关参数，这是需要耗费相关专业人士大量的时间和精力的，因此本文提出了一种基于深度学习的分割方法，将显微图像下的鲕粒灰岩进行全自动识别，达到节省人力成本，提高效率的工作方法。

近些年来，深度学习技术迅速发展，应用极为广泛，计算机视觉相关领域迎来了一波发展热潮。而在深度学习的实际应用中，当前也有很多相对成熟的一些案例，如高铁站的人脸检测过安检、智能驾驶技术的应用以及一些目标检测技术的应用等，这些方方面面的深度学习的应用都为我们的工作、学习、生活带来了数不胜数的便利，更是极大地提高了生产效率和生产力，促进了人类社会的快速发展。

虽然深度学习在某些领域已经得到了普及，效果也颇令人满意，但在一些专业的研究领域，深度学习还没能得到发展，甚至没有被加以应用。深度学习在一些可能会发挥重要作用的专业领域内尚未得到关注，在地质显微图像中鲕粒灰岩显微图像的识别上就有这种情况。传统的地质学方法，在对鲕粒灰岩的鉴别中需要耗费专业人士大量的精力去观察确定显微薄片中的物质是否确认是鲕粒，而选用深度学习的方法则可以免去这种耗费精力的工作，将这种工作交给机器去做，这将会大大提高工作效率甚至准确率，提高社会生产力和生产效率。本文的主要研究方向就是针对深度学习在地质显微图像方面的应用。

## 课题研究的意义以及问题的提出

* + 1. 课题研究的目的与意义

岩石和矿物的识别和分类是地质研究的重要内容。许多学者用不同的方法研究它们。它们大致可以概括为三类方法[2] ：

（1）物理实验方法，例如在2014年通过物理检测，根据实验室对Zaini et al的高光谱图像的分析，得出了如下结论：波的定义方法和波的有效性，可用于评估碳酸盐矿物表面的化学物质。

（2）数学统计分析方法，通过传统的数理统计和计算分析，识别和提取岩石分类特征。 这种方法对于岩石特征及其岩性分析有着不错的结果，但对研究人员的理论水平和实验设备的要求都偏高，往往会受到限制。

（3）智能学习分析方法，通过机器学习深度学习等方法对岩石基本特征进行处理，减轻对专业人员和高额设备的依赖，从计算机视觉领域出发，利用图像识别技术达到分析岩石岩性的目的。

传统的这三类方法中，智能学习分析方法还处于不完善的阶段，之前一直受制于计算机的计算能力以及没有合适的网络模型，但近些年来随着深度学习的高速发展以及计算机的计算能力每年都在成倍的增长，这使得对一些结构复杂的岩石显微图像进行分类分析有了极大可能。

鲕粒为的鲕粒灰岩的主要组成成分，而鲕粒灰岩多存在于石灰岩、白英岩、铝土质岩、碳酸岩等岩石之中，它具有良好的储油性。它的成因同时具备了了物理和化学作用，鲕粒灰岩主要形成于湖、海等有波浪活动以及潮汐活动的热带到亚热带环境，位于南北回归线之间。鲕粒在最初只是一些在沿海附近的生物碎屑、沙子、气泡等微小物质，在碳酸钙过饱和的水中被一次次的冲刷、翻滚，同时结晶，在其表面产生一层层的碳酸钙外壳，经过数百万年时间的冲刷，才形成了鲕粒的放射状同心圆结构。从力学性质上来讲，鲕粒的硬度大，耐压性强，在古代常被用作庙宇楼阁等的建筑材料，而现今鲕粒灰岩可被用于建筑地基、水族箱制作等，它是一种质地坚硬、结构紧密的良好的建筑材料[1] 。传统的鲕粒识别分析方法是通过在电子显微镜下对岩石显微薄片进行观察，耗费大量人力物力，且受到研究人员专业水平的限制，容易将一些近似的岩石误判为鲕粒灰岩，而现在通过深度学习的方法，能够达到更高的识别准确率，且工作效率更胜人工鉴别，在研究、生产方面有着重要的应用价值。同时这项研究也拓宽了深度学习的研究领域，填补了深度学习在地质显微图像领域上研究的空缺，为在地质上进一步应用深度学习的方法进行理论研究打下了坚实的基础。

* + 1. 问题的提出

本文我们主要解决三个问题：

（1）鲕粒灰岩显微图像数据集的制作

鲕粒灰岩的显微图像数据集是深度学习中进行训练过程的基础，没有数据集作为支撑，自然无法通过训练得到一个良好的模型，因此鲕粒灰岩数据集的制作就显得尤为重要。数据集的制作虽然有很多种方法，但我们希望得到质量更高的数据集，所以要选用高精度的设备对鲕粒灰岩的显微图像进行拍摄，确保得到的图片质量能够支撑后期的训练过程。之后对显微图像进行人工识别标注数据集的过程也较为繁琐，工作量巨大，极其容易在高强度的标注工作中产生少量错误的标注，因此需要选用合适的软件简化标注过程，增加标注的准确率，同时还要对标注过的数据进行反复的检查核对。而深度学习需要大量的训练数据，完全通过人工标注来产生鲕粒灰岩数据集显然不现实，因此可以利用图像增强方面的知识来帮助得到数据，还可以通过前期图像处理增强得到的显微图像质量等，总而言之，我们要得到大量的、标注精确的、效果令人满意的用于在网络中训练的数据集，以上是我们在本文研究中要解决的第一个问题。

（2）选用什么网络模型能达到更高的识别准确率

在深度学习中不同的网络模型对着最后的训练结果都有着很大的影响，而深度学习主要应用各种卷积神经网络，它们的不同点在于网络的深度，卷积核的选择、步长的选择、卷积的方式、池化的方式、激活函数的选择等方面。卷积核方面主要是关注局部特征以便于有效提高网络的表达能力，它是对频域信息的一种选择方式。而池化层的存在是为了压缩输入进去的特征图以达到进一步提取主要特征的目的，更能够让网络计算由繁化简，由难到易。还有一些其他参数同时也要考虑到。鲕粒灰岩显微图像中的背景杂质较多也有一些在颜色形状上与鲕粒相近因此分割难度较大，要注意其中的空间信息易缺失和感受野缩小的问题，要尽可能使用深度较大的网络。而计算机视觉中对图片分割的处理是基于每个像素点实现的，所以在对卷积网络修改的过程中卷积层重点要尽量加快计算速度，而全连接层重点要进行参数优化和权值剪切。网络模型的选择以及对其优化就是本文要解决的第二个问题。

（3）对鲕粒灰岩图像中参数的计算

在实际的研究中，单单是识别出鲕粒并不能满足研究的需求，还需要对鲕粒的一些实际数据进行测定，根据这些数据对鲕粒灰岩的岩性等方面进行判断，确定某一区域的鲕粒灰岩的性状，这就需要涉及到对最后模型预测出的显微图像进行二次处理，进一步利用显微图像测定出鲕粒的诸如周长、面积等数据。数据的测定就是本文要解决的第三个问题。

## 论文研究的主要内容

在论文的第一章中，主要介绍了鲕粒灰岩研究的背景，研究了本文研究的目的和意义，并且列举了待研究的问题，以及本文研究的主要内容。

在论文的第二章中，主要对一些图像分割的传统算法进行了介绍，并指出了其中存在的一些不足指出，为后文选择利用深度神经网络进行图像分割建立理论依据。

在论文的第三章中，主要介绍了实验环境以及使用到的仪器设备，并详细讲述了在环境配置中容易出现的问题。

在论文的第四章中，主要讲述了鲕粒灰岩显微图像数据集的制作流程，并对其中遇到的一些问题提出了合理可行的解决方案，完成了鲕粒灰岩显微图像数据集的构建。

在论文的第五章中，首先对卷积神经网络的主要模块进行了介绍说明，并解释了卷积过程中的一些算法，之后对本文应用到的BiSeNet双向分割网络以及其前缀网络ResNet深度神经网络做了详细的解释，说明了它的特点，完成了深度神经网络的架构并对网络的参数进行了优化。

在论文的第六章中，主要从检测的准确率（test\_accuracy）、平均交并比（mean iou）、损失率（loss）三个方面对模型的训练结果进行了详细的分析，之后展示了鲕粒预测的效果图并与人工标注的图像行了对比，最后对鲕粒灰岩显微图像中鲕粒进行了边缘分割，统计出了鲕粒的数量，并对鲕粒相关的一些参数做了计算。

在论文的第七章中，主要介绍了进一步需要研究的内容，在接下来的研究中，应当能对鲕粒的种类进行详细的划分，依据鲕粒的形状以及间隙物的种类来对鲕粒灰岩进行深入的分析研究，依据鲕粒的种类以及其他数值方面的参数对鲕粒所处的地理环境进行分析研究，进一步确定其应用价值。

# 图像分割算法概述

## 图像分割的定义

图像分割是将数字图像细分成若干图像子区域(“超级像素”)的技术，这是将图像分割为具有特定的单独属性的区域，并指出感兴趣的对象的技术和过程。图像处理是图像解析的重要领域[3] 。图像分割的意义是对图像的显示格式修饰或者简化达到更容易地分析图像的效果。图像分割通常用于指出在图像中的对象的方位和边缘（例如，线条、曲线等），进一步来说，图像分割的过程就是对图像中的所有像素点做标记的过程。这可以为拥有同样的标签的像素点提供一些共有的视觉特点、性质，有助于将各种具有特殊意义的区域从图像中分离出来。这些领域都是相对独立的，每个区域都满足该区域的同一性[4] 。总之，图像分割是对图像中的各个像素的类别或对象的预测。

已知某个图像所有像素表示为个区域全集的一种分割：

(1) ，表示一幅图像由所有的像素子集组成。

(2) 对每一个和j，，有，即每一个子区域在分割结果中是不相互重叠的。

(3) ，即拥有某些相似性的像素点存在于同一区域。

(4) ，即相邻的单一区域不可以进行合并。

(5) 对是连通的区域，即一些连通的像素点形成了一个子区域。  在区域上的所有像素点与其模型的相似程度由一致性谓词定义。

## 传统图像分割方法

基于阈值的分割、边缘分割（边缘检测）、聚类分析、区域分裂合并、区域生长、小波变换等等都是一些现存的图像分割方法[12] 。本文对基于边缘分割算法以及基于图的分割算法进行简述。

* + 1. 边缘分割算法

根据人类视觉的直观分析，图像边缘像素点往往伴随着颜色、纹理、明暗等属性的变换，产生明显的不连续性，这也是边缘分割算法的核心思想。边缘检测算子有很多，应用比较广泛的边缘差分算子包括交叉差分Roberts算子、Sobel算子、Prewitt算子、Canny算子等，下面对这些算子进行简单介绍。Roberts算子是利用局部差分算子来检测图像边缘[4] ，如公式(2-1)所述，该算法对于水平方向和垂直方向上的边缘检测的效果更好。



如下方程式(2-2)中表示出了Sobel边缘检测运算符的梯度幅度的计算方法，Sobel运算符使用微分运算来获得图像的梯度，选择适当的分割边缘阈值t，根据所选择的阈值t来确定图像的边缘点，其中梯度值大于阈值t的像素是边缘点。



其中，式(2-2)中

与Roberts算子相比而言，Sobel算子具有一定的噪声耐性，可以提高边缘检测效果。但由Sobel算子检测出的边缘粗糙，定位精度不够准确。Prewitt算子和Sobel算子类似，是先对噪声进行平滑处理后再求像素梯度。

Canny边缘检测运算符是常用的效果较好的方法，在通过高斯模糊去除噪声之后，使用一次偏微分的差分来计算梯度值，通过抑制非最大值来保存梯度方向的最大值。最后用双线性阈值来除去伪边缘点。它能够在抗噪性和边缘检测效果二者中取得平衡。Canny算子也是在之后对鲕粒灰岩显微图像进行对比验证时用到的较好的方法。

* + 1. 基于图的分割算法

基于图论的分割算法也有多种，下面只介绍目前基于图论的最好分割算法是GrabCut图像分割算法，该算法结合人工智能，是一种无监督的机器学习分割算法。通过将图中的每个像素当做图的节点，从而建立网络图，相邻节点的相似程度定义为权重也就是相邻节点的边，定义一个源点和一个汇聚点，通过计算能量方程，获取最小割，根据最小割重新计算权重，从而进行迭代不断更新，完成图像自动分割出前景目标。GrabCut图像分割算法能够方便的分割出前景目标，分割效果较好，因此在图像研究领域得到了广泛应用。

## 基于深度学习的语义分割算法

语义分割（应用于静态2D图像、视频甚至3D数据、体数据）是计算机视觉的关键问题之一，语义分割旨在像素级别理解图像，如我们期望对图像中的所有像素标注以目标类别。

在宏观级别中，语义分割是一个高层任务，用于理解场景。作为计算机视觉的中心问题，场景理解越来越重要，因为实际上越来越多的应用场景需要从图像中演绎推理有关的知识或语义（即具体到抽象的过程）。这些应用包括自动驾驶，人机交互，计算摄影学，图像搜索引擎，增强现实等。应用各种传统的计算机视觉和机器学习技术，这些问题已经得到了解决。虽然这些方法很流行，但深度学习革命让相关领域发生了翻天覆地的变化，因此，包括语义分割在内的许多计算机视觉问题都开始使用深度架构来解决，通常是卷积神经网络CNN，而CNN在准确率甚至效率上都远远超过了传统方法[17] 。然而，相比于固有的计算机视觉及机器学习分支，深度学习还远不成熟。也因此，还没有一个统一的工作及对于目前最优方法的综述。该领域的飞速发展使得对初学者的启蒙教育比较困难，而且，由于大量的工作相继被提出，要跟上发展的步伐也非常耗时。于是，追随语义分割相关工作、合理地解释它们的论点、过滤掉低水平的工作以及验证相关实验结果等是非常困难的。

## 本章小结

本章首先介绍了图像分割的定义，随后大致介绍了传统的图像分割方法，着重介绍了效果比较好的边缘分割算法以及基于图的分割算法，然后指出了他们的不足之处，引出了现阶段应用比较广泛的基于深度学习的语义分割算法。

# 实验软件工具及介绍

## 实验环境概述

前期工作主要是在LeiCa电子显微镜进行下图像的采集。接下来就是重要的标记工作，用PS对收集的岩石图像进行标记。

实验是在anaconda配置TensorFlow环境然后进行的，训练机器的配置是内存为12G、显卡为1060T。采用的是BiSeNet网络，batch\_size大小为7，learning\_rate为0.0001，epoch设置为600轮。把数据集分成train、test、validation，按照8:1:1的比例来划分数据集。它们分别对应相应train\_label、test\_label、validation\_label的标签。

## PS工具概述

Adobe Photoshop是功能丰富且比较容易掌握的图像处理软件，它是由Adobe Systems公司开发并发布的图像处理软件[5] 。由于一开始需要对大量的图片数据集进行标注工作，需要较快的工作效率，因而就需要一款较为便捷的标注工具。经过反复的筛选比较，在网上提供的大量的标注方法中，我选择了PS这款软件。它能为我们提供一个强大的快速选择工具，如图3-1是最初进行标注工作时的工作场景。如图3-2所示，而这恰好能够满足对大量数据集标注工作的需求，且PS本身含有批处理工具，方便对数据集进行批量的格式转化以及命名。



图3-1 PS工具展示



图3-2 快速选择工具图

## Tensorflow工具概述

随着深度学习技术的逐步兴起，多种多样支持深度学习的框架都慢慢涌现出来，现在业界内普遍使用的框架十分之多，如Tensorflow、Caffee、Theano等等。本文选用Tensorflow是因为Tensorflow作为Google重要的开源项目，其未来的社区热度是可以保证的，而一个火热的社区热度对推动一个开源项目的发展有很大帮助，我们可以比较容易在上面提问并获得其他人的帮助。

正如其名，Tensorflow是用于计算数据流程图的开源软件库。他有很多很棒的特色。首先，它支持多环境和集群。它在PC的CPU、GPU、甚至Android上都能发挥出色的性能。甚至能在它们之间并行计算，如有多个GPU则可以由应用程序控制在哪个GPU上进行任务的运行。

Tensorflow还可以领用同步以及异步两种方式训练模型，他可以将共享参数在参数服务器上进行更新[6] 。Tensorflow还内置了一个功能强大的组件TensorBoard，他可以让训练的过程变得可见，通过仪表篇可以观察到网络的现状以及数据流动的方向[7] 。综合了训练模型的时间、网络的构建难度、接口的支持程度以及一些算法包的多少等条件充分分析之后，结合本文的具体研究内容决定选择了Tensorflow作为本文开发鲕粒灰岩显微图像分割的框架。

## Tensorflow框架的安装

首先需要进行一些前期的准备工作。

（1）安装Anaconda3.X版本

如下图3-3所示，Anaconda是一个简单实用的包管理工具，它可以在机器上配置多个Python的运行环境，并对每个环境中的工具包进行管理，有了Anaconda就可以更加方便的解决在安装Tensorflow中由于一些版本不匹配引起的问题。

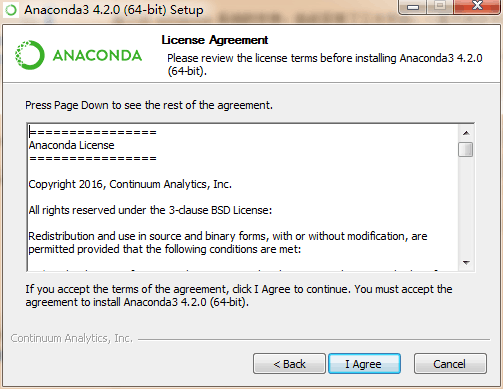


图3-3 Anaconda的安装

（2）安装CUDA与cuDNN

CUDA简单来说就是NVIDA推出的且只能在NVIDA显卡的GPU上运行的并行计算框架。而cuDNN则是一款应用深度学习时效果优秀的加速库，它可以完成深度神经网络训练模型时对GPU的加速任务。

通常CUDA与cuDNN需要有严格的版本匹配关系，他们之间与Tensorflow的版本也需要严格匹配才行，而这些关系通常可以在NVIDA的官网上找到得到，具体匹配关系图如下面的3-4图。

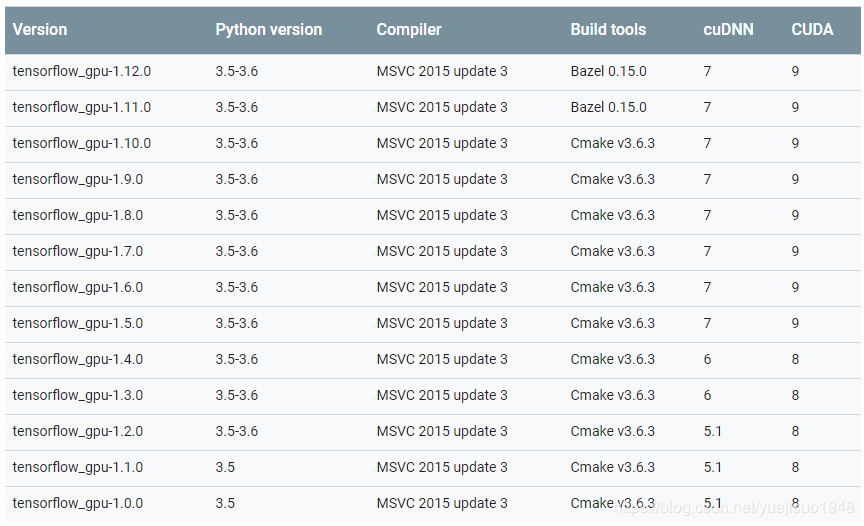


图3-4 版本匹配关系图

安装后，请在正常使用前设定系统环境变量。它可以在高级系统中的控制面板>系统和安全系统>的设置中进行添加。如下图3-5所示。

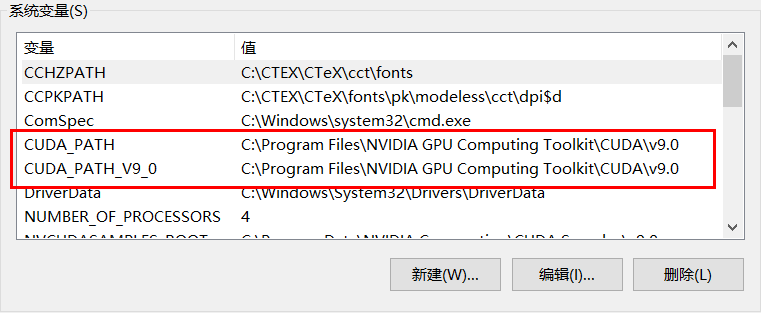


图3-5 添加CUDA与cuDNN环境变量

（3）安装3.X版本的python

本文中使用了3.6版本的python，python是一门集成了多种优点的编程语言，它接近自然语言且相较其他语言有更多可以使用的开源库，同时python在升级到3.X版本后有了巨大的变动。按照上述3-4图中的匹配关系，本文选用了3.6版本的python以及10.0版本的CUDA以及与其匹配的cuDNN。

（4）Tensorflow-gpu版本的安装

训练鲕粒灰岩数据集是一项较为庞大的工程，因此选用GPU进行训练模型会更加高效，因而选用了tensorflow-gpu版本进行了安装，具体过程如下：

输入如下命令，创建一个tensorflow-gpu版本的python环境，如下图3-6所示：

conda create -n tensorflow\_gpu python=3.6

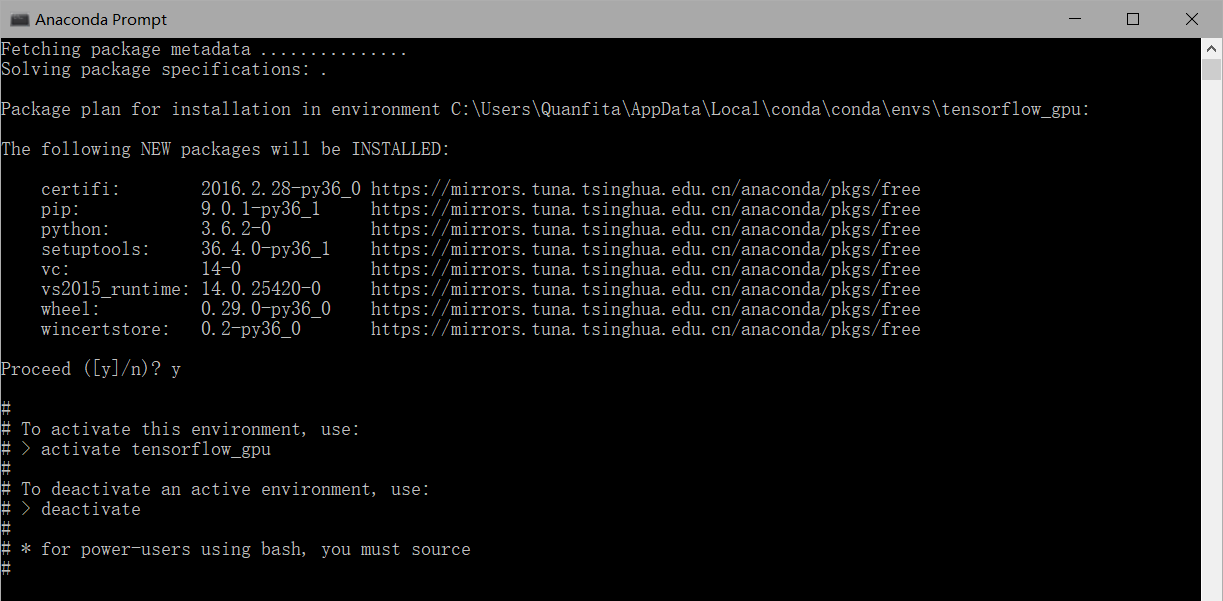


图3-6 创建tensorflow环境

激活刚创建好的环境并进入该环境，中间会让我们确认一下，输入个y回车就好了。安装好后会给我们提示用activate，和deactivate进行环境的切换。

activate tensorflow\_gpu

基本环境已经配置好了，此时需要安装一些重要的Python科学运算库，Anaconda已经为我们准备好的一系列常用的Python库，例如numpy，pandas，matplotlib等等，所以我们只需要安装一次anaconda库就可以把这些库全部安装好。

​conda install anaconda

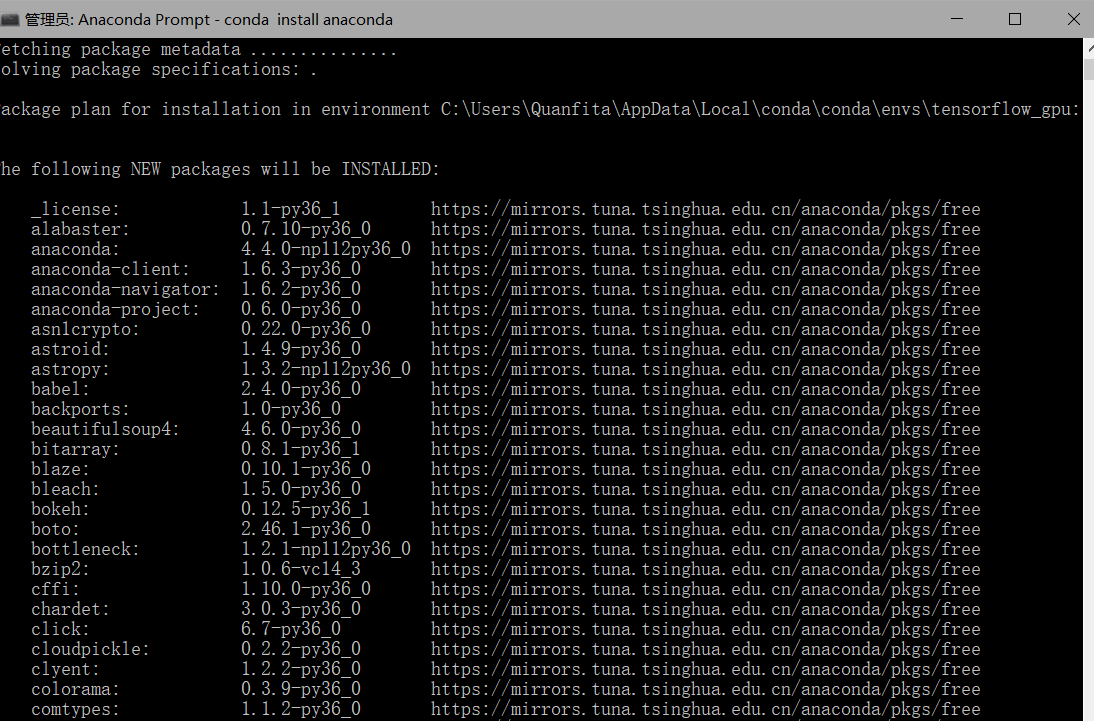


图3-7 python科学运算库的安装

最后准备工作都就绪了，进行tensorflow-gpu版本的安装：

conda install tensorflow-gpu

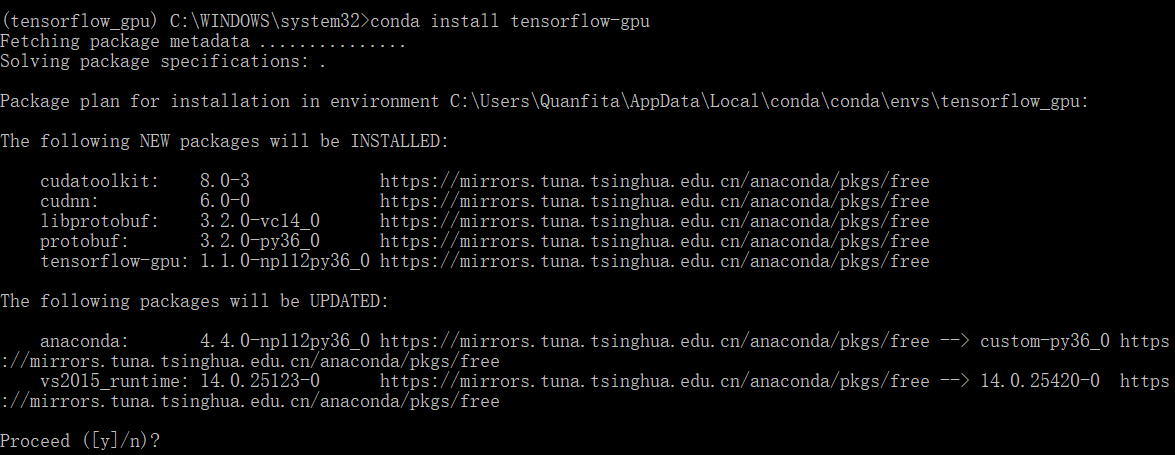


图3-8 tensorflow-gpu版本安装

## 本章小结

本章对实验所使用的器械、运行环境和框架进行了简要介绍，并详细讲解了在配置环境变量中容易遇到的问题。

# 鲕粒灰岩显微图像数据集的制作

由于前人并未对鲕粒灰岩进行过相关数据集的制作，因而本文采用了自己进行数据集制作的方法获取训练使用的数据，经过与有关专业人士进行沟通了解，在他们悉心的帮助下，使用相关领域专家提供的显微薄片样本，经过了长时间的摸索与实践，形成了一套完备的岩石显微图像数据集的制作体系[8] 。接下来的内容将详细介绍鲕粒灰岩显微图像获取数据集以及标注及相关数据增强的主要过程。

## 鲕粒灰岩显微图像的获取

由于进行深度学习需要大量的显微图像样本来制作初始的数据集，因此我们找到了资环专业领域的相关教授和学生，经过多次沟通，商讨并确认了方案的可行性，在得到了相关的专业知识的指导后，学习到了鲕粒灰岩中鲕粒的辨识技巧并更加具体的了解到了鲕粒的成因用途等方面的知识，并且成功与他们达成合作关系，由资环专业的教授提供他们仔细打磨好的鲕粒灰岩薄片，如下图4-1所示，并选定在Leica DM 2700 P偏光显微镜下进行显微图像的拍摄，如下图4-2所示。



图4-1 岩石显微薄片样本



图4-2 Leica DM 2700 P偏光显微镜

拍摄每周进行两到三次每次持续四个小时左右，前前后后差不多用了三周左右的时间，每次的拍摄任务都在资环学院学生的帮助下完成的，需要两到三个人协作进行拍摄，分别观察下显微镜更换岩石薄片、调整薄片位置，在电脑上筛并存储录入拍摄到的显微图片。拍摄是在目镜10倍焦距、物镜5倍焦距总计对显微薄片放大50倍的条件下下完成的，这个大小刚好可以观察到大小适中且范围较广的鲕粒情况，下面4-3图是拍摄过程中的情况。



图4-3 显微图像拍摄过程

## 鲕粒灰岩数据的采集

在显微图像获取完成后，又对拍摄完成后得到的显微图像进行了筛选工作，去除了其中因为拍摄时人为因素导致的效果较差的图像，得到了一个大约160张图片的初始数据集，如下4-4图。

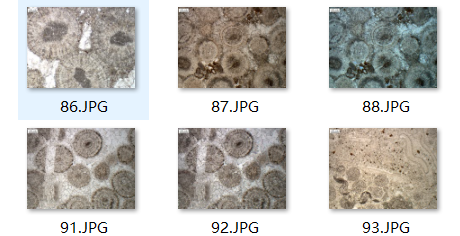


图4-4 鲕粒灰岩显微图像样本

进行到了这里本应继续进行接下来数据标注的然后数据增强之后训练模型，当然我也这样做过了，但这样做并没有得到一个特别理想的训练效果。之后通过查阅资料了解到国际上比赛中的数据集需要识别的目标多数占到整张图的左右甚至更多，所以借助网上大多数数据集提供的参考，这里提出了一个相当不错的解决方案，我们知道训练的目的是让网络提取出的模型更好的识别鲕粒同心圆放射状的特征，基于此，鲕粒的数据集得到了重制，我对已经得到的图像进行了二次采集，截取了其中比较完整特征明显的鲕粒，并将其用PS修改成了大小为600\*600像素、水平与垂直分辨率都为96dpi、通道数为3、位深度为24的新的图像，制作成了一个全新的鲕粒灰岩数据集。如图4-5所示，其中鲕粒灰岩占整张图像的比例也得到了调整。其中参数如图4-6所示。

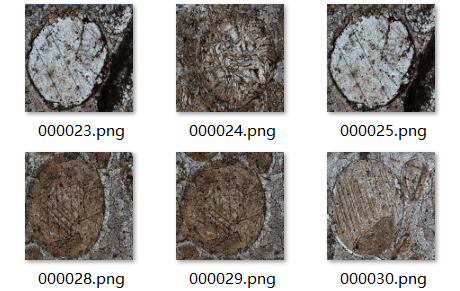


图4-5 重制的鲕粒灰岩显微图像



图4-6 修改后的图片参数

## 鲕粒灰岩数据的标注

为了用更小的数据集更快的训练速度得到较好的模型训练效果，本文采用了BiSeNet双向分割网络来进行训练过程，要用到的的数据集不仅仅只有显微镜下得到的显微图像，还需要一个人为进行标注的图像数据集一同输入进去，借用ResNet残差网络，令其对每一轮的残差进行学习，以达到更高效完成训练的目的。因此在对鲕粒灰岩显微图像进行标注时也要分外仔细，要尽量切合鲕粒本身的轮廓，减少由于人为因素导致的误差，令网络训练出更精确的模型。

PS工具是标注过程中一个很好的助手，经过了各方面性能的比较筛选，PS工具得以应用到标注工作中。在前文中提到过，快速选取工具是一个标注鲕粒的及其有效且迅速的手段，如图4-7所示，它结合套索攻击可以极快的完成大部分的标注工作。

除了快速选区工具外，对以那些在光线较暗情况下拍出的图像以及鲕粒与背景色及其接近的一些样本图像，钢笔工具更是一个不错的选择。先利用钢笔工具建立工作路径，然后利用PS自带的批处理命令将其转化为选区，之后填充颜色。下图4-8所示是钢笔工具标注效果图。



图4-7 套索工具建立选区效果

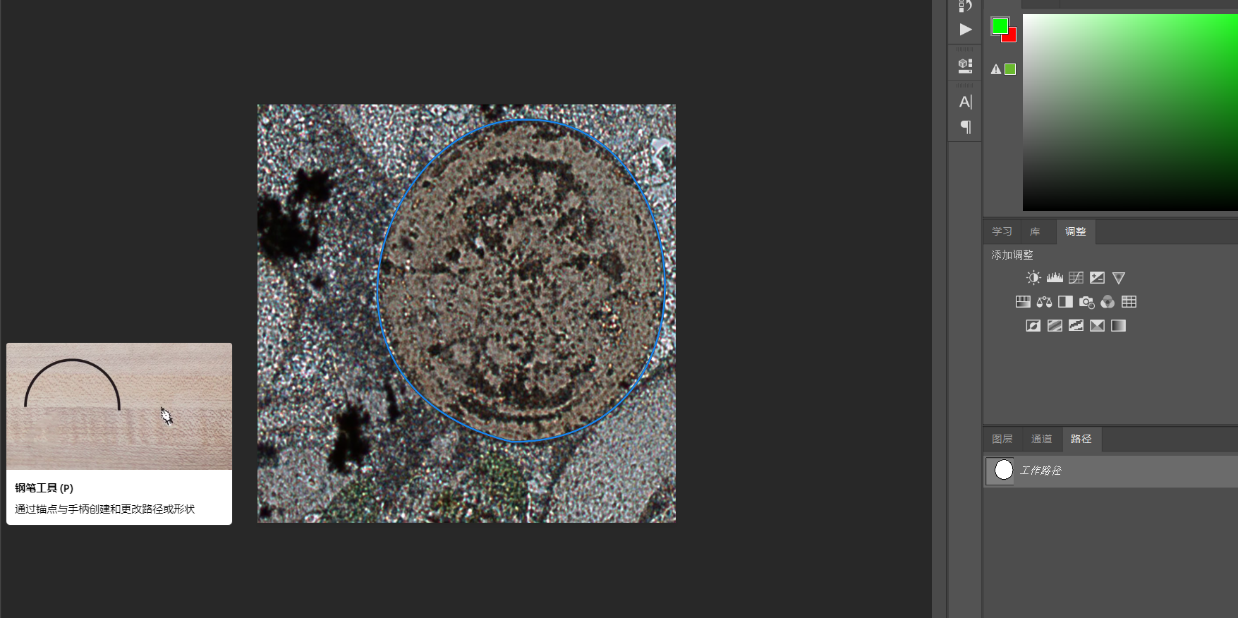
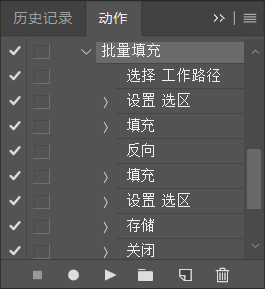


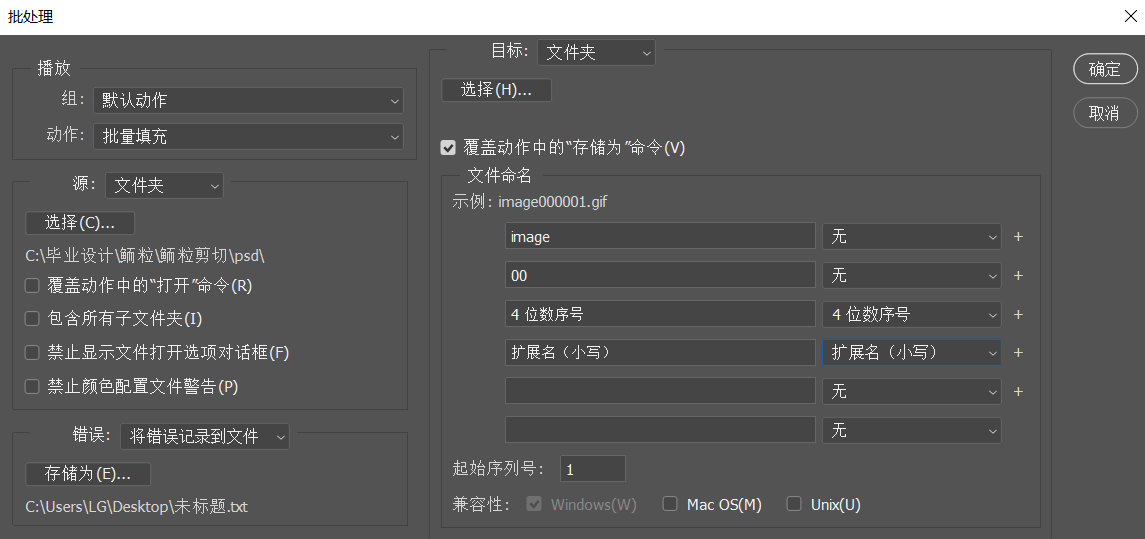
图4-8 钢笔标注工具

标注工作需要将鲕粒与背景区分开，然后其中鲕粒填充成绿色，背景填充成红色，最后另存为一个PNG格式的鲕粒灰岩标注数据集，要注意与其相应的原图片名称一一对应。

在这里利用批处理命令会更快的帮助完成上色任务，批处理命令是PS中对图片进行批量操作的命令，它可以提前设置好一个动作，之后选择执行批处理命令的文件夹中的图片都会进行该动作中的操作，达到对图片进行批量建立选区并填充前景色和背景色的目的。如图4-9所示，用一张需要操作的图片提前设置好批处理的动作，然后选定导入和导出的文件夹，选择覆盖存储并在执行完动作之后关闭文件，之后我们只需要静静等待电脑将我们标注好的PSD文件中的选区上好色，然后另存为PNG格式的标注数据集，完成数据集的初步制作。



（a）批处理动作



（b）进行批处理操作

图4-9 设置批处理动作并进行批处理

标注要完成的任务听起来并不是十分困难，但通常在实际完成时一张图需要五分钟左右，长期标注会导致注意力不是很集中，容易出现人为误差，且标注过的图需要逐一反复检查是否有错标甚至漏标现象，确保其中没有“漏网之鱼”，因此标注也是除了构建网络之外最繁琐的工作，大约花费了三个星期左右的时间才得以完成，下图4-10所示是标注的效果图。同时在批处理操作的过程中也要注意将覆盖动作中的“存储为”命令勾选上，这样才能保证上色过后的文件存储到之前选定的目标文件夹，否则将可能直接让标注过后的psd文件被上色，导致原始标注的数据丢失，不利于之后对上色图片进行审核工作，出现错误时甚至要重新进行标注，浪费人力物力，因此做好备份也很重要。

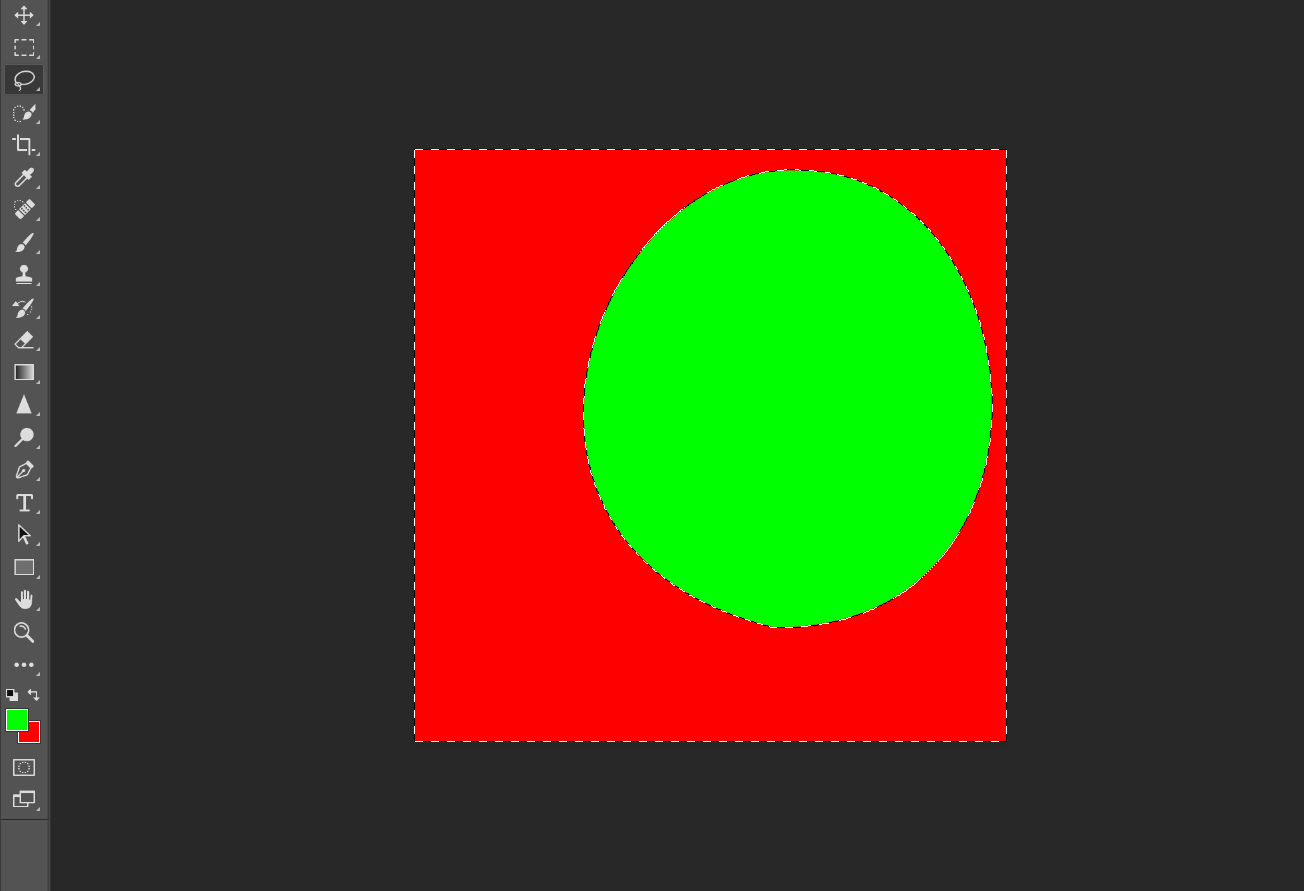


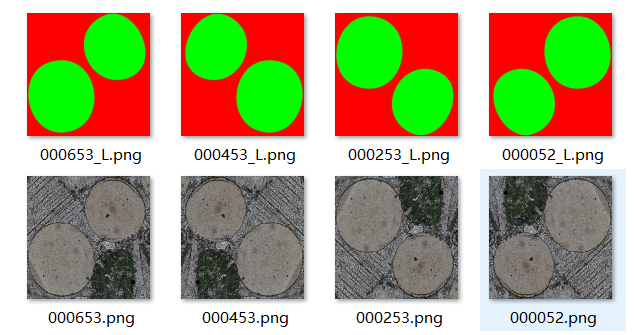
图4-10 鲕粒灰岩标注数据集上色

## 鲕粒灰岩数据集的数据增强

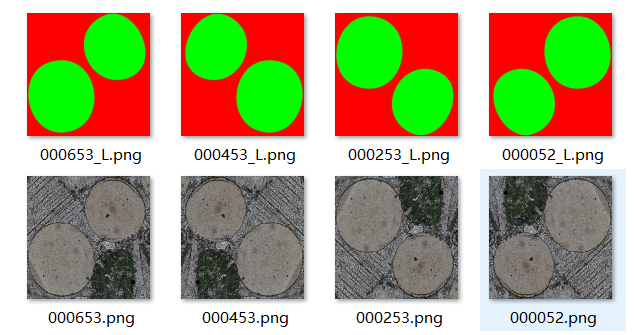
由于采集到的鲕粒灰岩显微图像样本珍贵，且数量有限，难以满足深度学习时对大量数据集的需求，因此本文对得到的两百张鲕粒灰岩标注图像进行了数据增强操作以扩大训练用的数据集，增强训练模型的效果。

数据增强就是通过对图像进行水平、竖直方向的反转以及一定角度的旋转得到新的图像的过程。由于网络训练模型时输入的图像会得到像素级别的处理，因此对图像进行翻转等操作虽然在我们看来还是同一张图没错，但因为像素点的改变，网络会将其当做不同的图像来处理。数据增强目前有离线增强和在线增强两种方式，本文将同时进行这两个方面的增强[9] 。离线增强正适合数据集数量不多的情况。

如图4-11所示，本文同时对数据集进行了垂直、水平、水平垂直翻转三种方式的数据增强。图（a）为原图数据增强后一张图片对应的四张翻转过后的图片，图（b）为其对应的鲕粒灰岩显微图像标注的数据集。可以明显的看到在翻转过后图片每个位置的像素点对应都得到了改变，这样在模型训练的过程中，就由原来的对一张图片训练变成了对四张图片进行训练，扩大了训练的数据集，最终显微图像的数据集大小扩充到了800张图像。



（a）鲕粒灰岩翻转后的图像



（b）反转后对应的填充图像

图4-11 鲕粒灰岩显微图像数据增强

离线增强虽然将鲕粒灰岩数据集扩大到了800张，但仍需要尽可能多的数据才能使训练出的模型达到更好效果，因此在将图像输入到网络中进行训练时，又进行而二次的在线数据增强。考虑到图像大小要适应卷积层的需要，图像输入进网络时就需要对输入图像的像素大小进行随机裁剪，因此对随机裁剪后的图像做二次的在线增强时仍可以利用水平，垂直翻转等操作来增强数据，因为对随机裁剪后的图像而言，进行翻转后又会生成一组各个像素点与原图不同的图像，同时还可以对图片的角度进行旋转或是进行图片亮度的调整以适应数据增强的需要，在在线的数据增强后数据集又选择性的扩充了多倍，此时的数据集大小便能够满足深度学习中训练的要求。

在线增强代码如下：

水平翻转：



垂直翻转：





水平垂直翻转：



随机角度旋转：



## 本章小结

本章对实验所使用的器械、运行环境和框架及其安装配置进行了简要介绍，并描述了数据采集工作的进行过程和采集到的数据情况，接下来又因为采集到的数据不满足训练的要求，又对采集到的数据进行了数据集的重制，然后对重制过的数据集进行了准确的标注，最后对标注过后的数据进行了数据增强操作。

# 深度神经网络模型的构建

在深度学习中卷积神经网络是不可或缺的重要的一部分。LeCun自1989年首次提出真正的卷积神经网络以来，已经过去了29年。自2012年AlexNet网络问世以来，卷积神经网络在过去6年间迅速发展，并在许多问题上取得了很好的结果。尤其是在2015年将卷积式神经网络模型应用在图像分类领域，它打开了图像分类的新时代。卷积神经网络是最广泛使用的深度学习技术之一[10] 。它的构建主要由卷积层，活性化层，游泳层等一层一层堆叠起来的。

## 卷积神经网络基本组成结构

* + 1. 卷积层

卷积其实是一种数学上的运算，它的效果在某种意义上要远远由于全连接网络，尤其是在图像识别领域。我们称(f\*g)(n)为f、g的卷积，其连续的定义为：



其离散的定义为：



也就是说，和满足一定的条件的情况下，对和两个函数进行卷积运算(f\*g)(n)，对这样的乘积进行求和或者积分。如下图5-1所示，在卷积神经网络中，两个矩阵之间的运算被叫做卷积，左边的是输入矩阵，中间的是卷积核（也称作滤波器），在上以一定的步长移动，同时完成点积运算，得到右边的输出矩阵。这个过程就是卷积神经网络中最重要的卷积部分[13] 。

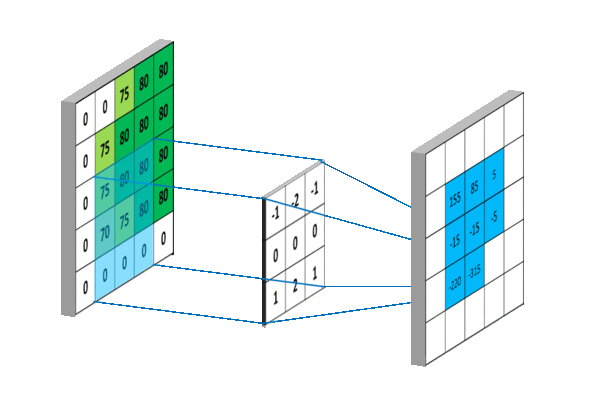


图5-1 卷积运算示意图

通过将偏移量bisa置于点积运算中的与滤波器的不同深度对应的矩阵，将其应用到相应深度输入矩阵的不同位置可以获得输出矩阵的各个元素，如下图中5-2所示。对卷积层的改进可以通过卷积核小型化，卷积，Network In Network，Inception机制，卷积分解（Factorization），反卷积运算等来进行。在这些方法中，卷积核的小型化也广泛地被接受，这也是最优化方法，本篇论文将使用这种方法。

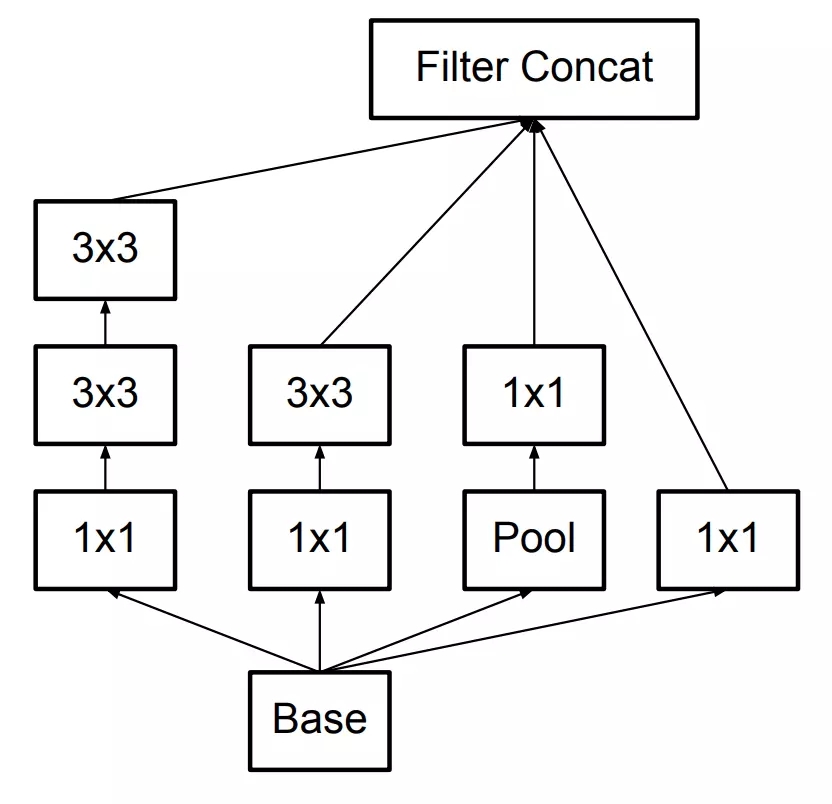


图5-2 几种卷积核的设计方式

* + 1. 池化层

池化操作也被称为下采样。主要功能是维数的削减、数据和参数的压缩、过度匹配的削减、以及模型的容错性的提高。图像的下采样可以减少数据处理，并保持有用的信息。这都通过使用图像的局部相关原理实现。使用上述数据，池化后的图像的长度和宽度可以按以下方式获得：





池化操作主要就是降低张量的维度，对图片进行压缩处理，它大致的作用可以概括为以下三点[14] ：

（1）特征不变性，即在图像处理中经常被提及的的比例不变性的特征，池化操作是能让我们像认识处理前的图像一样认识压缩了两倍的图像。以一张狗的图像为例，它显示着这幅画保持着狗最重要的特征，我们可以看到画中的画是狗而不是别的什么无关紧要的东西。这些留下的信息最能表现图像的特征。

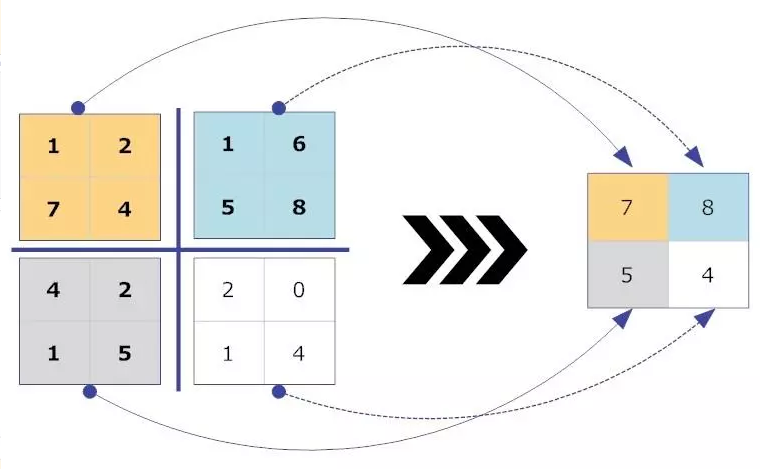
（2）特征降维，即特征的维数减少。一副图像可以包括许多信息和特点，然而，其中一些信息对处理图像没有那么重要。此时，可以删除这些冗余的信息，并提取最重要的特征。

（3）在某种程度上，池化可以防过拟合，让优化更加的方便。此外，即使在特征图中的像素偏移了少量的情况下，池化层的存在也可以保证输出层的输出不改变，从而提高网络的鲁棒性。

如下图5-3中为两种池化方式，图（a）所示为区域平均值池化和下（b）图所示为区域最大值池化，这两种池化方式为常用的池化方式，本文采用的为区域最大值池化，池化的简要过程在下面进行了展示。



（a） 区域平均值池化



（b） 区域最大值池化

图5-3 两种池化方式

* + 1. 激励层

激励层可以使神经网络近似为任何非线性函数，它的功能是提供规模的非线性化能力，它模拟激励的神经元的状态变化。不使用激活函数将导致每层的输出都与上层输入线性相关，那所有的输出层将都是输入层的线性组合，这将极大的影响训练结果。当前主流的三种激励函数为，Sigmoid激励函数，Thah激励函数，ReLU激励函数[15] ：

（1）Sigmoid激励函数：用于隐层神经元输出，取值范围为(0,1)，可以用来做二分类。



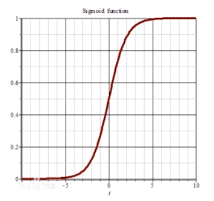


图5-4 Sigmoid函数

（2）Thah激励函数：Tanh是由Sigmoid变形得到的，它在 sigmoid 上做的改进是，tanh 的均值是0。但是梯度消失和幂运算的问题仍然没有得到解决。在实际应用中，tanh 比 sigmoid 应用的更广泛。



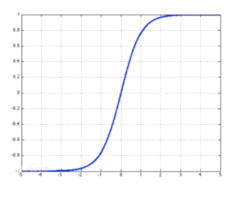


图5-5 Thah函数

（3）ReLU激励函数：输入信号 时，输出全部为0，输入信号时，输出全部等于输入。ReLU 函数会比sigmoid或是tanh收敛速度更快。相较之下，ReLU 得到激活值的过程中只需要一个阈值，在此之上它还同时解决了梯度消失这一问题，它是目前应用效果最好的激活函数。



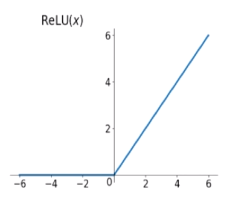


图5-6 ReLU函数

## 卷积神经网络中的前向与反向传播算法

在整个神经网络传播的过程中，分为前向传播算法和反向传播算法，下面分别介绍一下这两种算法的传播过程以及其主要的作用。

* + 1. 前向传播算法

整个神经网络中，我们最终要训练的其实是每个像素点对应的权重值，而这些权重值对应的分类结果是从输入层开始层层递进最后在输出层得到的，现在我们令第l-1层的输出作为第l层的输入并设为其输出值，l= 1,2,3,…L：对应第l层三种不同情况可的下面三个推导公式：

第l层为卷积层：





第l层为池化层：



第L层为输出层：



其中为第l层的权值为第l层的偏置fl-1是第l层的激活函数，P为通过区域最大值池化对输入的图片进行池化的过程，C为分类器，本文中是Softmax。

* + 1. 反向传播算法

在进行一个魔性训练的过程中，如果只有前向的传播显然是不够的，前项传播尽管可以得到一个分类结果，但却无法在接下来的训练过程中不断地优化我们训练中使用到的权值，因此反向传播过程就显得尤为重要，反向传播的工作速度远远快于早期的学习方法，因此可以使用神经网络解决以前不可解决的问题。今天，反向传播算法是神经网络学习的主力。反向传播的过程就是在不断的计算梯度值并更新模型直到损失函数收敛的过程，它通过后向传播残差来逐步修正卷积神经网络完成了极大地简化多层复合函数中变量求解的过程，下图4-7所示是一张神经网络最后两层的计算过程图：

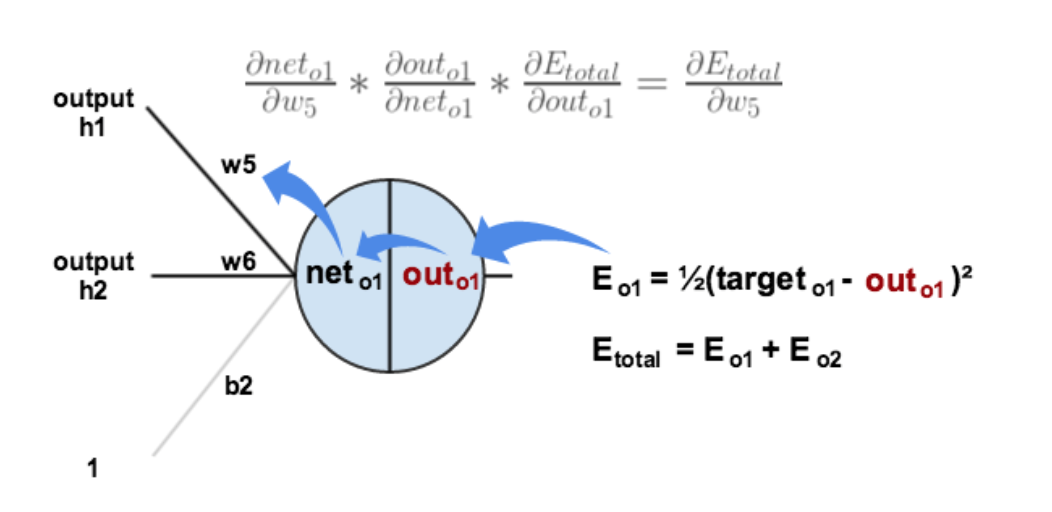


图5-7 神经网络参数传递示意图

## 端到端卷积神经网络的构建

* + 1. ResNet深度神经网络概述

ResNet是一种深层的残差神经网络，它是由来自微软研究院的Kaiming He等四名华人提出来的。它们利用Residual Unit成功训练出了深度最高高达152层的深层次的神经网络并以此获得了在ILSVRC 2015年度大赛的冠军头衔。ResNet网络的结构不仅仅在准确率上提升十分显著同时也极快的加速了神经网络的训练，他们夺冠时得到的错误率仅仅有3.57%但参数量却比VGGNet低得多，训练成效十分显著。

如下图5-8所示，ResNet最开始的提出是因为大家在加深网络训练深度的时候会因为深度过深而出现Degradation的问题，这会导致准确率在上升到一定程度后逐渐达到饱和，而过深的训练会导致准确率不如之前。

经过研究发现，这并不是因为出现了过拟合的问题，而是因为不只是测试集数据上因为人为原因存在一定的误差同时训练集本身也因为深层次的训练会加大误差。

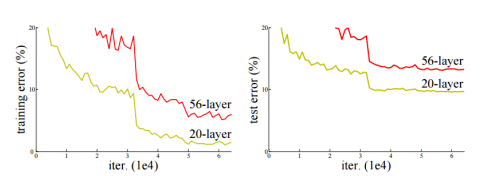


图5-8 不同深度网络训练误差比较

现在令一个浅层的神经网络达到饱和，那么如果在之后加上一些的全等映射理论上不会导致误差的增加，那么便可以推论出更深层的网络实际上不并不是导致训练集产生误差的罪魁祸首。而这里提到的全等映射，也就是ResNet产生的根源了。

假设一段神经网络的输入为x、期望得到的输出为Y(x)，那么基于此如果我们把x直接传到输出作为初始的输出结果，此时应该的学习任务应该是F(x) = Y(x) – x，如下图5-9便是 RseNet的一个残差学习单元的结构示意图，在这个结构单元中可以很直观的看到上述的计算过程，其中直接将输入结果传到输出层然后与输出结果进行比较计算出差异就是残差，让计算机学习这个残差而不是直接对输出结果进行学习。

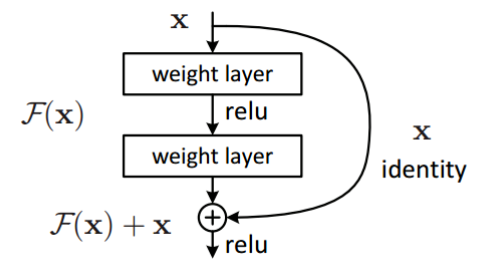


图5-9 残差学习单元示意图

* + 1. 深度神经网络的基本原理简介

ResNet变相改变了学习目标，它只是输出和输入的差别Y(x)-x，即残差，不再学习一个完整的输出Y(x)。通过训练集重制将一个问题转化为多个尺度的直接映射问题同时能够起到很好的优化训练效果。这个剩余的模块通过残差实现，通过这个残差将这个模块的输入和输出进行一个参数上的叠加，这个简单的操作并不会给网络增加额外的参数和计算量，同时，它可以大大提高模特训练速度并改善训练效果。此外，当模型的层数变深时，该简单结构能够有效地解决退化问题。图5-10是ResNet对应的各层网络的输入输出图像大小。

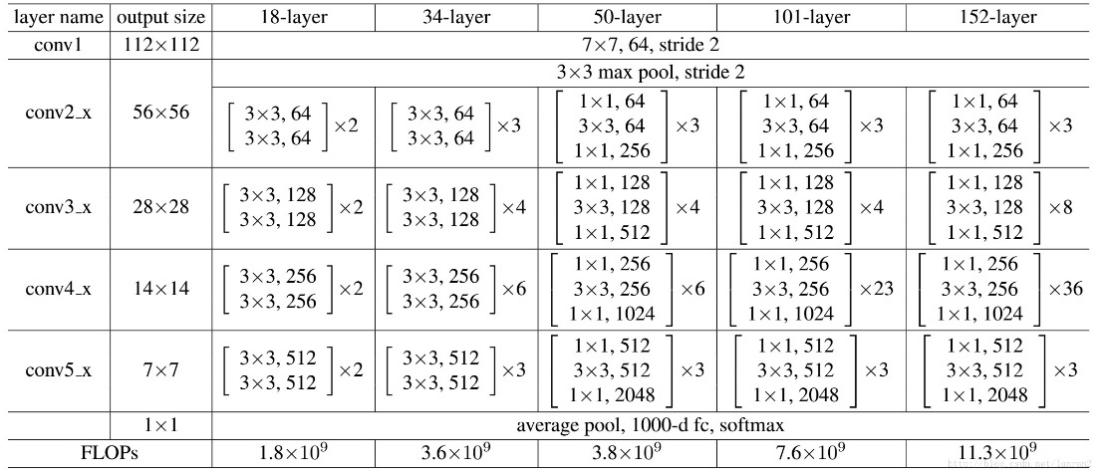


图5-10 ResNet网络各层网络输入输出结构

传统的神经网络在卷积层或者全连接层传递信息时，大都存在这训练集信息丢失损耗的问题但RseNet另辟蹊径解决了这个问题，它使得输入信息直接传到输出让网络学习输入、输出的差别部分，不仅保护了信息的完整性还简化了学习目标。下图5-11是比较VGNet 19和通常的34层的卷积神经网络的示意图，比起ResNet神经网络层，它们最大的区别在于ResNet网络有很多支路能够将输入直接连接到后面的输出令后面的层可以直接学习残差。

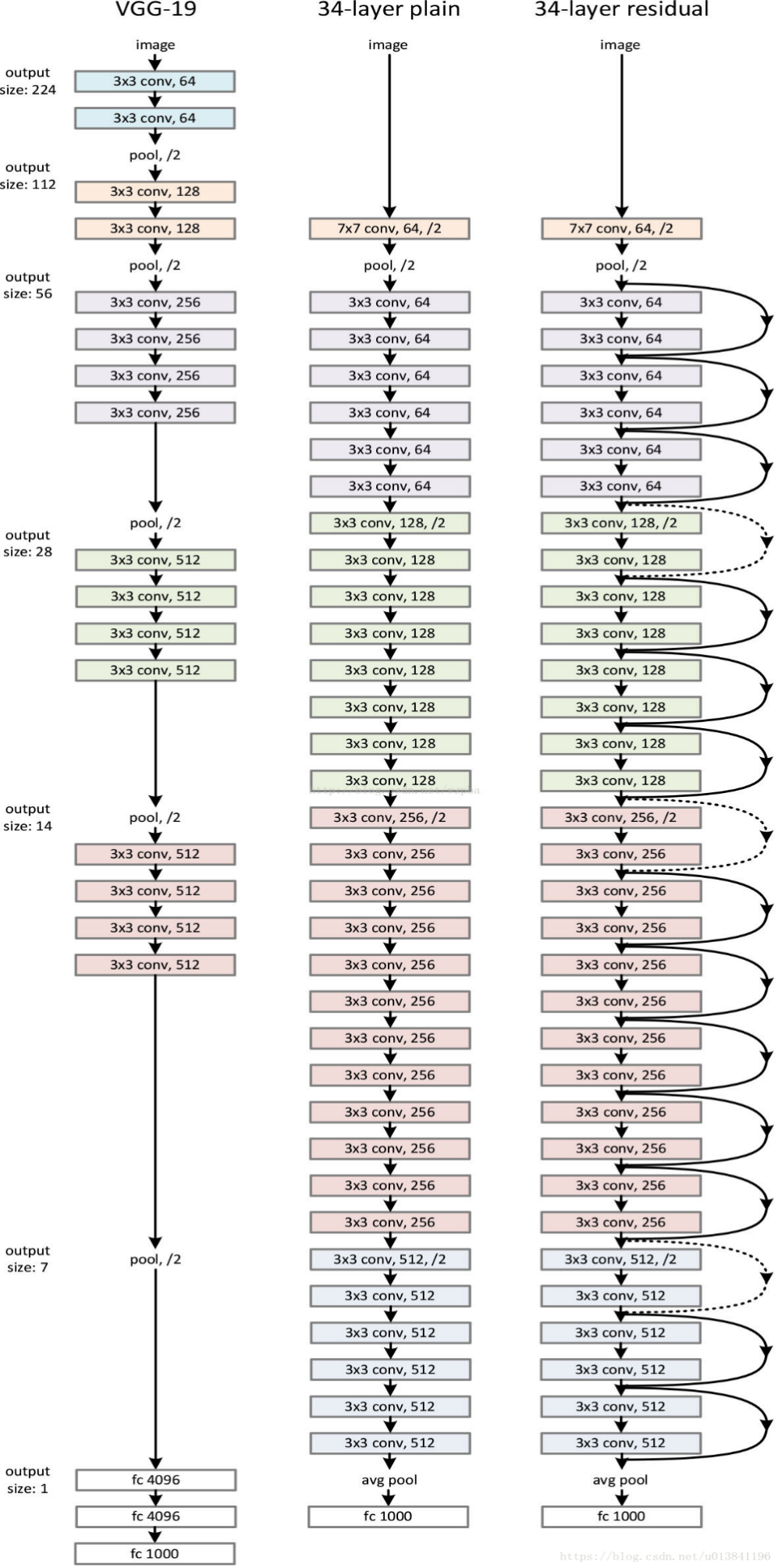


图5-11 三种神经网络比较图

下面提出了三种常用的映射方法：

（1）都用恒等映射，输入输入模块大小相同，此时如果残差模块的输出维度与输入维度不同，那就对其进行0填充。

（2）在模块输入维度与输出维度相的情况也使用恒等映射，在输入输出维度不一致时使用线性投影方式以确保维度变得一致。

（3）将线性投影应用到所有模块上，这样可以保证每一层的效果都达到最佳的效果但相应的模型复杂度会提高很多。

在本文的实验中，由于前缀网络选用的是ResNet101深度神经网络，网络深度很深，为了确保实验中模型的复杂度比较低，所以选用了全部进行恒等映射并进行0填充，而不是选择效果更好但模型复杂度更高的线性投影方式。

如下图5-12所示，在两层残差学习单元的基础上进一步的提出了三层残差学习单元的结构，其中H(x)-x残差它的输入输出要相同才能运算，也就是目标中的输入输出维度要保持相同，在两层残差单元中用到了两个的卷积核，因此要在三层残差单元中达到同样的效果就要改变卷积核大小，这里先运用了一个1\*1的卷积核，之后又用了一个3\*3的卷积核，最后又是一个1\*1的卷积核，这样卷积出来的结果达到了先降维再升维的效果，这样可以在后面的x层对维度利用线性变换进行调整。

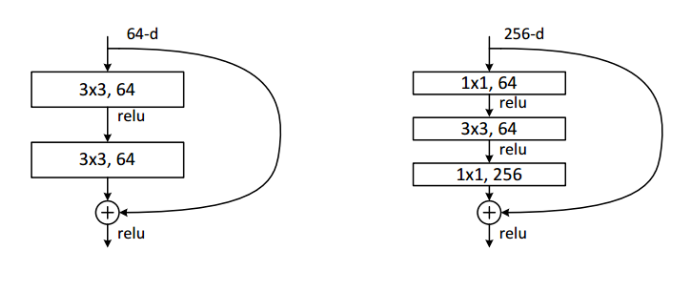


图5-12 两层与三层残差学习单元

三层结构拥有相同数量的层又减少了参数量，因此可以拓展成更深的模型。于是作者提出了50、101、152层的ResNet，而且不仅没有出现退化问题，错误率也大大降低，同时计算复杂度也保持在很低的程度，因为鲕粒灰岩的显微图像训练集中干扰因素较多，所以选择了深度高达101层的ResNet深度神经网络进行训练，以求之后能够达到更精准的识别效果。

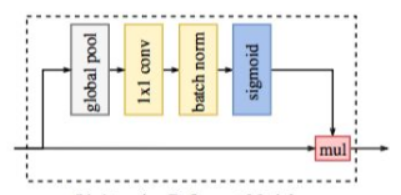
## BiSeNet双向分割神经网络的构建

在计算机视觉领域，语义分割是其中的一项基本任务，他可以对图像中的每一个像素点进行一次标记。在深度学习应用越来越广泛的今天，一种新的双边分割网络(Bilateral Segmentation Network/BiSeNet)被创造出来，抛开了传统神经网络的设计方法，它主要包含两个部分，Spatial Path (SP) 和 Context Path (CP)。这两个比分设计理念十分清晰，其中SP用来防止空间信息缺失，CP则是用来保证感受野不会缩小[11] 。

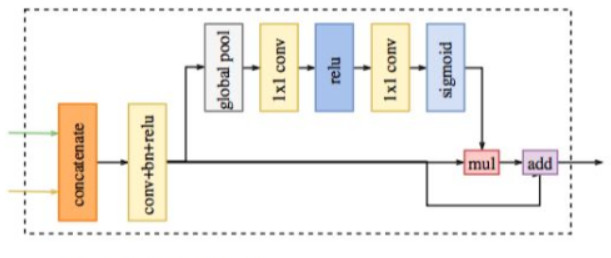
* + 1. BiSeNet网络架构

BiSeNet深度神经网络运用了前缀网络，把提前训练好的Xception作为环境路径的backbon，在之后的空间路径中运用了三个卷积层，重新定义了步长，最后通过特征融合模块将这两个输出结果进行特征提取并融合，完成最后的模型结果。BiSeNet是在空间路径（Spatial Path）和 环境路径（Context Path）的支持下完成的实时的语义分割，他可以同时实现高性能以及解决感受野缺失的问题，在本文中前缀网络用的是ResNet101深度神经网络。

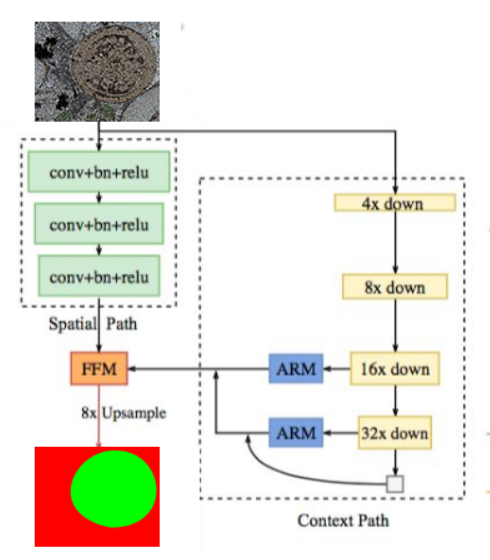
首先，就实际计算而言，Spatial Path只有三个卷积层，尽管它拥有很大的感受野，因此，它的计算并不密集。而Context Path则可以使用轻量级模型进行快速池化操作。此外，这两个组件可以并行计算，大大提高了效率。其次，BiSeNet还关心到了网络精度问题。Spatial Path编码了丰富的空间信息，Context Path提供了较大的感受野，两者互补结合，实现了更高的性能。下图5-13为BiSeNet网络的模型构建过程。



（a） 注意力优化模块（ARM）



（b） 特征融合模块（FFM）



（c） BiSeNet网络模型构建图

图5-13 BiSeNet网络架构图

下面为构建网络的部分代码：



* + 1. 空间路径（Spatial Path）

在语义划分任务中，难以实现空间分辨率和感受野同时满足，特别是实时语义划分时，传统方法往往基于较小的输入图像或轻量级主模型来加速。然而，较小的图像比原始图像缺少更多的空间信息。轻量级模型由于切割通道而丢失了空间信息。在此基础上，BesiNet提出了Spatial Path用以保存来自原始输入图像的空间尺度，并对丰富的空间信息进行编码。Spatial Path包含3层，每层包含一个卷积操作，其步长为2，然后进行批处理归一化，最后经过ReLU激活函数激活。因此，该网络可以提取与原始图像1/8大小对应的输出特征图，而大规模地使用特征图就可以编码更丰富的空间信息[11] 。

下面为构建空间路径的部分代码：



* + 1. 环境路径（Context Path）

感受野的情况在一个语义分割的实现中线的尤为重要，传统的一些方法，为了增大感受野，采用了空洞池化（ASPP）或称为“large kernel”，它的模型呈现为一个金字塔形，而这种结构显然是十分耗费计算和内存的，这通常会导致计算速度很慢。而BesiNet网络提出的Context Path同时考虑到了计算效率和感受野两方面的要求，它充分利用轻量级模型与全局平均池化以提供大感受野，而本文经过反复试验比较，选用了前文中提到的ResNet深度残差神经网络来作为网络中的Context Path来提取图片的上下文语义特征信息。

在BesiNet网络中如 Xception的轻量级模型可以快速下采样特征图以获得大感受野并编码高层语义语境信息。接着，BesiNet网络在轻量级模型末端添加一个全局平均池化，通过全局语境信息提供一个最大感受野。

下面为构建环境路径的部分代码：





* + 1. 注意力优化模块（ARM）

BesiNet 网络在Contect Path中提出了一个单独的注意力优化模块（ARM）来优化各个阶段的特征。ARM通过全局平均池化获取全局上下文语境，计算注意向量并指导特征学习。该设计可以优化Contect Path中各阶段的输出特性，它的计算成本几乎可以忽略不计，因为它无需进行上采样操作就可以轻松集成全局语言环境信息。

* + 1. 特征融合模块（FFM）

在特征表示方面上，两种网络的特征是不同的。因此，这些特征能仅仅进行加权。Spatial Path捕获的空间信息编码了大量丰富的细节信息。Context Path的输出特性主要是对上下文语境信息进行编码。也就是说，Spatial Path的输出特征在低层，而Context Path的输出特征在高层[11] 。在此之上，BiSeNet提出了一种独特的特征融合模块来融合链各个路径得到的特征。

在给定不同层次特征的情况下，BiSeNet首先将Spatial Path和Context Path的输出特征连接起来，然后通过批量归一化均衡特征的尺度来实现，利用连通特征池作变为一个特征向量，之后计算该权重向量[16] 。该向量可以对在特征选择和结合中起到重新加权特征的作用。

下面为特征融合模块的执行过程：







* + 1. 损失函数

BiSeNet通过辅助的损失函数监控模型的训练过程，损失函数均选用了Softmax函数，通过主损失函数监控整个BiSeNet的输出。同时还可以通过添加两个特殊的辅助损失函数像多层监督一样来监督Context Path 输出结果。最后,借助了参数α对主损失函数与辅助损失函数的权重进行了平衡。

## 本章小结

本章一开始主要介绍了卷积神经网络的基本组成单元，包括卷积层、池化层、激励层等，之后简单介绍了前向传播与反向传播的过程，进一步引出了ResNet残差神经网络的结构，最后介绍了如何构建BiSeNet双向分割神经网络。为本文的算法和实践提供了充分的理论依据。

# 鲕粒灰岩显微图像分割结果及实验分析

## 实验分割结果

本试验主要从检测的准确率（test\_accuracy）、检测精度（precision）、召回率（recall）、平均交并比（mean iou）、前景识别准确率（Inner）、背景识别准确率（Outer）等方面分析本次试验的预测结果[18] 。下面主要介绍了识别准确率（test\_accuracy）、平均交并比（mean iou）、损失率（loss）三个方面的具体结果并对其进行简要分析。

具体结果如下表6-1所示：

表6-1 检测结果统计表

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  | | --- | | 测试  图像 | | |  | | --- | | 测试  准确率 | | |  | | --- | | 检测  精度 | | |  | | --- | | 召回率 | | |  | | --- | | 测试分数 | | |  | | --- | | 平均  交并比 | | |  | | --- | | 背景  识别率 | | |  | | --- | | 鲕粒  识别率 | |
| 41 | 0.987 | 0.988 | 0.987 | 0.987 | 0.941 | 0.944 | 0.993 |
| 42 | 0.987 | 0.943 | 0.934 | 0.934 | 0.836 | 0.882 | 0.989 |
| 43 | 0.932 | 0.980 | 0.980 | 0.979 | 0.983 | 0.982 | 0.979 |
| 44 | 0.980 | 0.991 | 0.991 | 0.990 | 0.968 | 0.993 | 0.990 |
| 45 | 0.991 | 0.991 | 0.991 | 0.991 | 0.973 | 0.978 | 0.995 |

## 识别准确率分析

* + 1. 检测的准确率（test\_accuracy）

下图6-1所示为训练轮数（epochs）与准确率（accuracy）的折线关系图，准确率简单来解释就是鲕粒灰岩通过模型预测得到的图像与对应人工标记的图像的相似程度，通过这个参数可以准确的反映出一个模型的优劣，而由下图6-1可见鲕粒灰岩显微图像的识别准确率随着训练轮数的提升总体呈上升趋势，在训练了100轮以后准确率保持在95%上下波动，最高时的准确率接近99%，可见模型训练的效果十分出色。

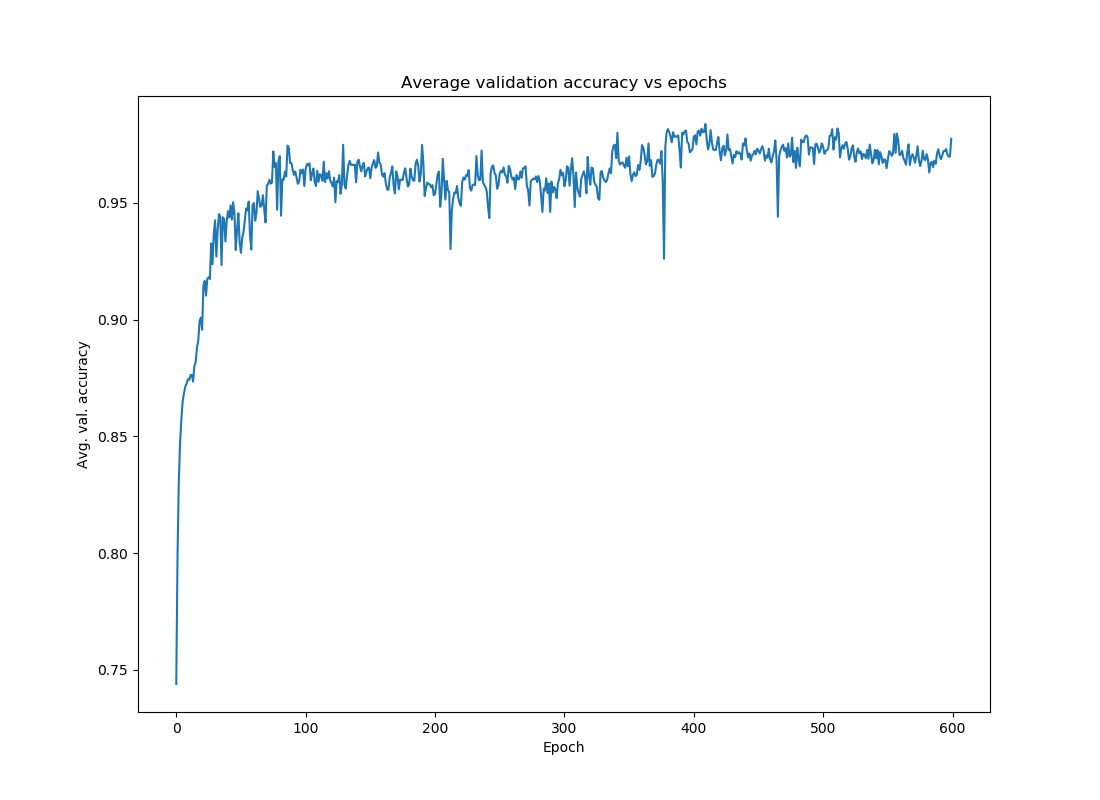


图6-1 训练轮数与准确率关系图

* + 1. 平均交并比（mean iou）

平均交并比表示预测出的候选框与原标记框的交叠程度又或是重叠率，即平均交集与并集的比值关系，相关性越高，平均交并比就越高，在理想情况下，比值为1，即预测图与标号图完全重合。如下图6-2所示，随着训练轮数的提升，平均交并比维持在95%左右波动，达到了很好的训练效果。

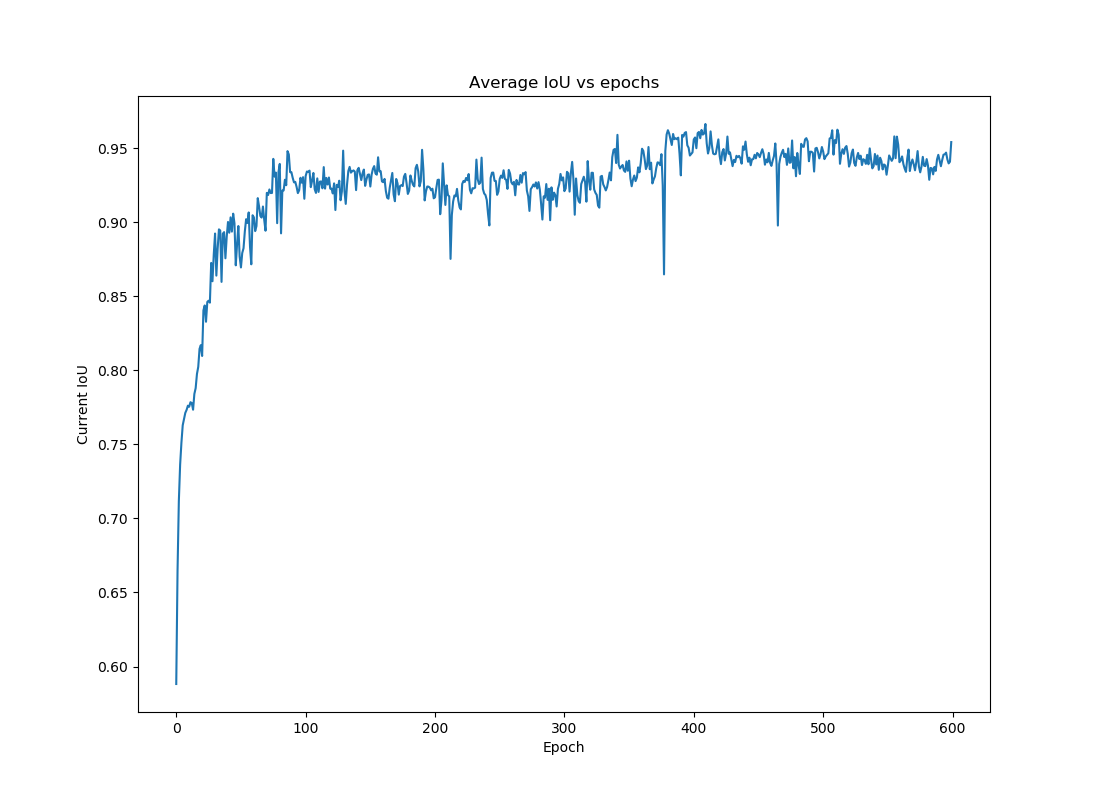


图6-2 平均交并比与训练轮数关系图

* + 1. 损失率（loss）

本文运用到的损失函数计算方法是Tensorflow中内置的tf.reduce\_mean函数，这个函数主要用于计算张量tensor沿着指定的某一维度上的平均值，它主要用作降维或者计算tensor的平均值。

具体代码如下：



如下图6-3所示，随着训练轮数的增加，预测图像的平均损失值在有效降低，并在200轮开始达到稳定状态，平均损失在%0.04左右趋于稳定，训练效果明显，预测出的鲕粒灰岩显微图像与真实值十分接近。

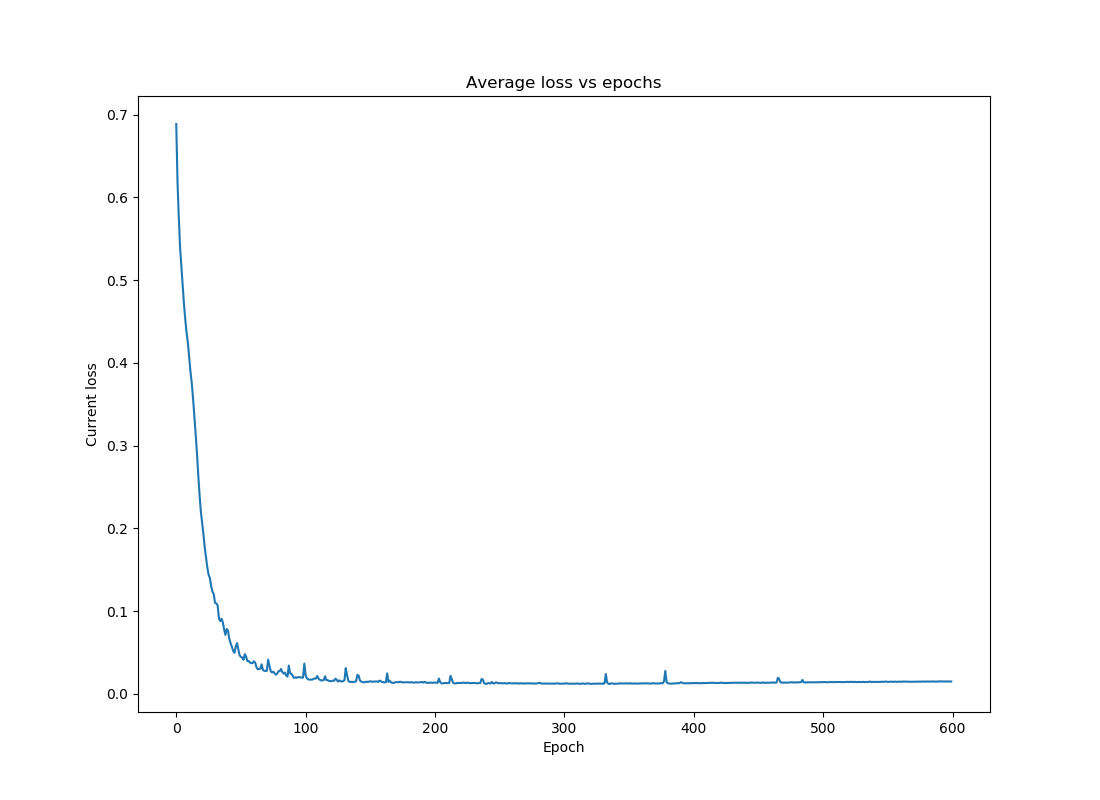
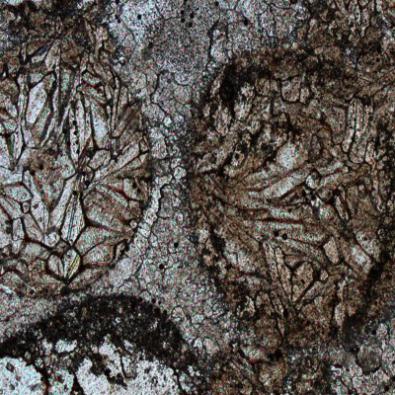
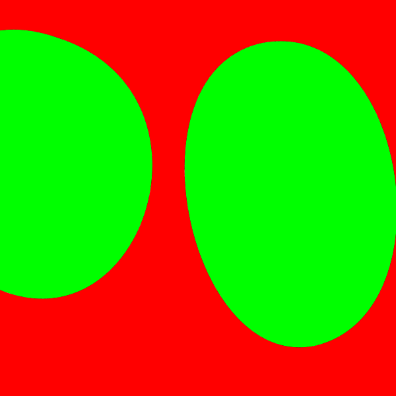
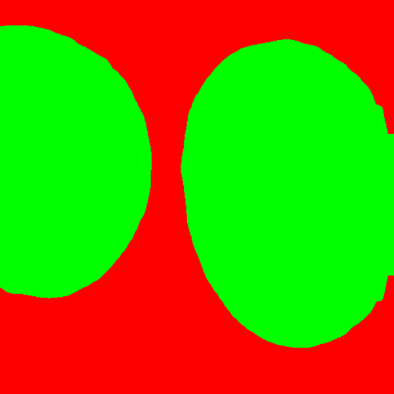


图6-3 损失率与训练轮数关系图

* + 1. 鲕粒预测效果展示

在分析过上边的参数之后，大概对模型的训练成果有了一定的了解，从识别准确率（test\_accuracy）、平均交叠率（mean iou）、损失率（loss）等方面随着训练轮数的变化信息，我们了解到了模型识别的准确率稳定在95%以上，达到了十分不错的训练效果。如下图6-4所示，图（a）为鲕粒灰岩显微图像原图，图（b）为人工进行标注的鲕粒灰岩显微图像，而图（c）则是我们通过网络模型训练后得到的机器自动识别预测出的鲕粒的图像，肉眼基本上难以分辨出机器自动分割的与人工标注的有什么不同，识别效果十分显著。

（a）鲕粒原图 （b）鲕粒人工标注图 （c）鲕粒自动预测图

图6-4 鲕粒预测效果展示图

## 相关参数计算结

在完成鲕粒灰岩显微图像中鲕粒的识别的同时，出于在地质学上对地质显微图像的研究分析需要，以及能够更加方便的对原显微图像进行观察，我们利用了高频滤波对x，y 轴梯度进行了图像的边缘提取[19] ，也就先将预测图转化为灰度图像然后对其中的灰度值产生明显变化的位置进行边缘提取，再将提取到的边缘拟合到原鲕粒灰岩显微图像中去，得到一个带有鲕粒轮廓线的鲕粒灰岩显微图像。

其具体操作步骤如下：

（1）图像灰度化，就是将原预测图中0-255的颜色值都除以255得到一个每个像素点颜色值都在0到1之间的灰度图像。

（2）用高斯滤波器平滑图像，去除图像中的少量噪声点，排除干扰。

（3）用一阶偏导有限差分计算梯度幅值和方向。

（4）进行非极大值抑制，也就是沿着梯度方向寻找像素点局部最值。

（5）双阈值的选取，通常来讲，边缘检测算法通常指需要用到一个阈值，它的作用是除去较小梯度值的干扰，但本文选用的Canny算子是取一个最高阈值和一个最低阈值，并将其分为三种点[20] ：

边缘像素梯度值高阈值的点，称为强边缘点。

低阈值边缘梯度值高阈值的点，称为弱边缘点。

边缘梯度小于低阈值的点，称为抑制边缘点。

（6）滞后边界跟踪，在若边缘点中，有一类是因为颜色或是噪声变化引起的，它们通常与强边缘点不连通，而其他的若边缘点则与强边缘点连通，通过这个准则又可以排除一部分的噪声干扰，使边缘线划分的更精确[21] 。

预测边缘的效果如下图6-5所示。

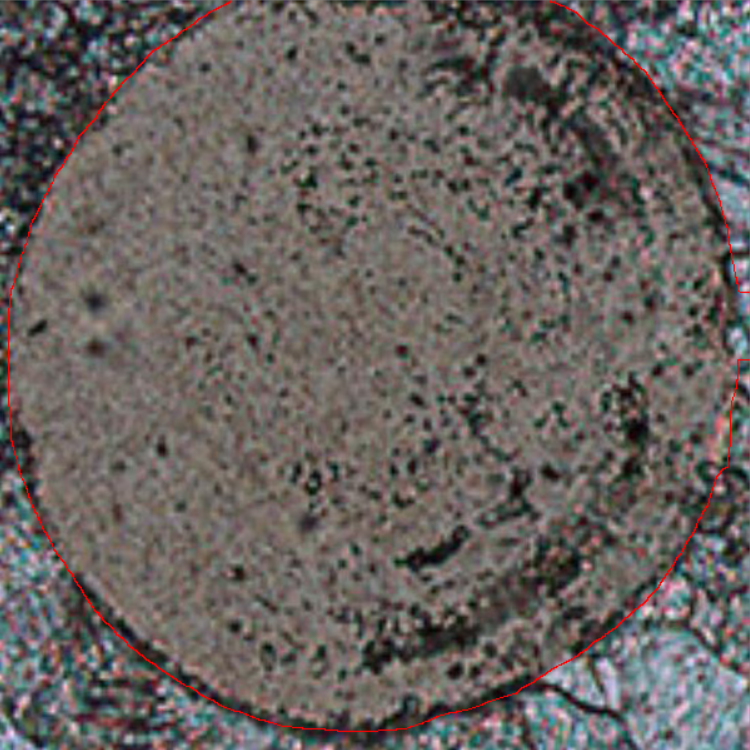


图6-5 鲕粒边缘检测图

针对检测出的鲕粒部分，图中包含的鲕粒数量以及它们的面积和周长都得到了计算，可以用作进行鲕粒的数据分析，上图6-5计算出的结果如下，我对鲕粒按照面积大小进行了排序，取了其中最大鲕粒部分的计算结果加以显示：

鲕粒总数量为：1

最大部分面积为：277064.0

最大部分周长为：2003.6996958255768

## 本章小结

本章主要介绍了实验的分割结果，以表格的形式呈现出了各项分析数据的测试结果，同时将训练过程中的识别准确率（test\_accuracy）、平均交并比（mean iou）和损失率（loss）做成了随轮数增长的折线图并进行了详细的分析，之后又向大家展示了鲕粒的具体自动识别的情况，最后对鲕粒的边缘进行了检测并拟合到原图像上，并进行了如周长面积等的对鲕粒的一些参数计算分析。

# 总结与展望

## 论文工作总结

鲕粒灰岩显微图像的自动分割，在地质显微图像分割领域的研究中较为重要。卷积神经网路拥有人类所不能比拟的强大学习训练能力，它在各个领域已经取得了初步的成功，并走向成熟。随着计算速度的不断提高，卷积神经网络在图像处理方面应用的必定更加广泛，利用BiSeNet双向分割神经网络将分割过程分为了空间路径与环境路径两条通路，并在最后进行了特征融合，弥补了传统的卷积神经网络在细节分割上的不足，完成了更好的分割效果。主要工作如下所示：

（1）对鲕粒灰岩显微图像数据的采集、数据集的制作以及数据的标注和数据增强的工作，由于关于地质显微图像方面的应用技术在当今世上暂时还存于理论阶段，没有前人制作好的可供训练使用的数据集，因此要做相关领域的研究就要先从数据集的制作开始，经过六周左右时间的摸索，制作鲕粒灰岩显微图像数据集的方法终于逐渐变得成熟。一方面训练用到的数据集成功得到了制作，另一方面，也为之后进行相关领域的研究奠定了一定的实践基础。

（2）介绍了一些图像分割方面的传统的算法，指出了传统地质显微图像分割方法的一些缺点，本文引入了BiSeNet双向分割网络，分为空间路径（Spatial Path）和环境路径（Context Path），它们分别用来理解图像的特征和语义，同时考虑了空间信息缺失与感受野缩小的问题，最后通过特征融合模块（Feature Fusion Module）将两个路径提取的特征融合在一起，最后借助全局平均池化捕获全局并计算注意力以指导特征学习。形成了感受野更广，鲁棒性更强的网络模型。达到了更细致的对图像进行分割的目的。

（3）对鲕粒灰岩显微图像的训练效果做了各方面的分析，主要从检测的准确率（test\_accuracy）、平均交并比（mean iou）、损失率（loss）等方面对网络模型的训练效果做出了具体分析，得出了它们随着轮数增长的变化折线图，展示了鲕粒的预测效果图。同时作了进一步的研究，从鲕粒预测图中提取出鲕粒边缘，拟合到原鲕粒灰岩显微图像上去，并计算出了图中鲕粒数量，对鲕粒的周长、面积等进行了简单的计算分析。

## 进一步研究方向

鲕粒灰岩显微图像的全自动分割，在理论研究以及实际应用中都有着重要的意义，本文提出的选择使用BiSeNet双向分割网络进行鲕粒的分割在应用中取得了较好的分割结果，但该分割结果还存在着一些细节上的问题，一些破碎的鲕粒会对正确的识别结果有一定的干扰效果，同时在制作数据集时由于受到专业知识的限制，没能将鲕粒按照更细的中种类划分进行标注，在后续的工作中仍然还需要继续进行以下方面的研究：

（1）对鲕粒的种类进行进一步区别

鲕粒灰岩显微图像中鲕粒的类别是多种多样的，而不同种类的鲕粒也有着一些不同之处，这都是由于它们形成的环境不同引起的，在接下来的工作中，应当不仅仅能够识别出鲕粒灰岩中的鲕粒，更应该能分别出其种类，这样才能更深入的研究某一具体地区鲕粒灰岩的形成环境、成因、岩性等的一些性质特点，同时也能根据其鲕粒类别的不同确定其不同的应用价值。

（2）基于鲕粒灰岩的填充物进行二次分类

鲕粒灰岩中不仅仅鲕粒是其重要组成部分，鲕粒灰岩可按照填充物分为亮晶鲕粒灰岩及微晶蛹粒灰岩，它们由于在不同的地质环境下形成，性质上也有所区别，地质学上可以根据鲕粒灰岩的填充物对其形成环境进行细致的分析解读，因此对鲕粒间的填充物进行进一步分类也显得尤为重要，在理论研究上也有着重要的价值。

# 致谢

本文是在芦碧波导师以及杨文艺学长的精心指导下完成的。在兴趣的指引下，我选择了这个关于鲕粒灰岩图像处理方向上的研究。芦碧波导师不仅知识渊博，而且平易近人，乐于与学生们进行沟通交流、答疑解惑。当我有什么不懂需要帮助的时候，他都会耐心的为我解答，为我解释清楚，并对我提出的一系列问题提出建设性的意见，教会了我学习新知识的方法，非常感谢您！杨文艺学长在课题的研究中也为我提供了很大的帮助，他在毕业设计中传授给我的经验让我受益匪浅，同时我也跟他们在图像分割方面学到了很多新的知识。同时还要感谢资环学院的老师为我提供的帮助，感谢您提供的鲕粒显微薄片样本以及观察设备，也感谢您专门为我讲解了相关地质方面的知识，扩宽了我的知识面，让我能更好了了解到鲕粒灰岩的背景知识。芦碧波导师在我后期的论文指导中也给了我很多帮助，我会定期报告自己的进展进度，对于格式的要求也给了我们许多的指导，谢谢你们！

在毕业设计期间，我不会的问题有时候也会和大家一起探讨，也会得到满意的结果，在这里要感谢我们专业的同学们，感谢你们对我的帮助。也要感谢我的女朋友，在我毕业设计期间帮我查阅外籍资料，在我写代码、改论文烦躁的时候安慰我，督促我提早完成了毕业设计，真的很感谢你！

在本科学习期间，我还得到了学院、各个老师以及同学们的帮助，感谢你们一路的陪伴和指导。如果没有你们的一直陪伴和教导，我就不能很好的独立完成自己的学业。在此，向你们表示最真诚的谢意，谢谢你们！

最后，我要感谢我的父母，感谢他们对我的养育，感谢他们开明的教育，他们为我提供了能让我专注学习的条件，成就了现在的我，谢谢你们！

# 参考文献

王宁, 何幼斌, 李向东, 苑伯超, 罗进雄. 深水异地鲕粒灰岩的特征及成因探讨——以宁夏中宁地区香山群徐家圈组为例[J]. 海相油气地质, 2012, 17(03):72-78.

李航, 统计学习方法, 清华大学出版社, 2012

赵诗宇. 图像分割算法的分类与研究[J]. 科技风, 2019(05):58.

张淑姚, 孙婷婷, 张路园. 浅谈图像处理技术[J]. 科技风, 2019(02):62.

刘武辉. PS可选颜色工具的色彩学原理及使用方法[J]. 广东印刷, 2019(01):23-24.

高扬, 卫峥. 白话深度学习与TensorFlow[M]. 万娟. 插画设计. 北京：机械工业出社, 2017-08.

加日拉·买买提热衣木, 常富蓉, 刘晨, 要秀宏. 主流深度学习框架对比[J]. 电子技术与软件工程, 2018(07):74.

芦碧波, 刘利群, 郑艳梅, 等. 一种线束端子显微图像全自动分割方法[J]. 光电工程, 2016, 43(10): 49–55.

高友文, 周本君, 胡晓飞. 基于数据增强的卷积神经网络图像识别研究[J]. 计算机技术与发展, 2018, 28(08):62-65.

Changqian Yu, Jingbo Wang, Chao Peng, Changxin Gao, Gang Yu, and Nong Sang. Bisenet: Bilateral segmentation network for real-time semantic segmentation. In Proceedings of the European Conference on Computer Vision (ECCV), pages 325–341, 2018.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun. 2015. Deep residual learning for image recognition. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770-778, 2016.

Vaillant, R. Monrocq, C. & LeCun, Y. Original approach for the localisation of objects in images. In Proc. Vision, Image, and Signal Processing 141, 245–250 (1994).

Kavukcuoglu, K. et al. Learning convolutional feature hierarchies for visual recognition. In Proc. Advances in Neural Information Processing Systems 23 1090–1098 (2010).

G.E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R.R. Salakhutdinov. Improving neural networks by preventing co-adaptation of feature detectors. arXiv preprint arXiv:1207.0580, 2012.

K. He, X. Zhang, S. Ren, and J. Sun. Spatial pyramid poolingin deep convolutional networks for visual recognition. In ECCV. 2014. 1, 2

高琳, 陈念年, 范勇. 融合多尺度上下文卷积特征的车辆目标检测[J]. 光电工程, 2019, 46(04):31-38.

朱明君. 基于卷积神经网络的图像语义分割技术[J]. 电子技术与软件工程, 2019(01):69-70.

O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, “ImageNet Large Scale Visual Recognition Challenge, ” in International Journal of Computer Vision (IJCV), 2015.

张永梅, 王世伟, 许静. Canny算子与小波变换相结合的道路边缘特征提取算法[J]. 计算机时代, 2015(08):44-46.

朱凤芹, 杨慕升, 马晨晖, 闫怀仁. 改进Canny算法在再制造视觉检测中的应用[J]. 中国农机化学报, 2015, 36(03):270-273.

拓小明, 李云红, 刘旭, 曹浏, 霍可, 田冀达, 陈航. 基于Canny算子与阈值分割的边缘检测算法[J]. 西安工程大学学报, 2014, 28(06):745-749.