基于Inception模块和残差学习策略的U-net野值压制

**摘要：**野值干扰是实际地震资料的常见非相干噪声，实际数据处理中多采用基于阈值设置的相关去噪方法，其对有效信息的损伤较大且需人工干预。本文将Inception模块和残差学习策略同时应用于U-net神经网络中，用并行多尺度卷积层取代常规卷积层，结合残差学习和批标准化提升网络性能，该方法在不显著增加神经网络参数量以及计算量的前提下拓展了U-net横向宽度和纵向深度，进而实现了地震数据中野值干扰的高精度剔除。基于模型数据和实际数据的实验均表明，基于Inception模型的U-net神经网络训练速度快且泛化性强，噪音剔除精度高。相比于常规阈值压制类技术，其处理效率更高且噪音剔除效果更好，显示出该技术在实际数据处理中具有良好的应用前景。

**关键词：**野值干扰；噪音剔除；U-net；Inception模型；残差学习

## 0 引言

野值干扰是实际地震资料的常见非相干噪声，在炮集域多表现为弧形干扰和线性干扰（潘军等，2016），其振幅相较于有效信号呈高值异常的特点降低了地震资料的信噪比，因此，地震记录中异常振幅的准确剔除对于后续高精度数据处理具有重要意义。

针对实际数据中的压制干扰，常用的压制方法有滤波法（Dabov et al., 2007;Buadesetal.,2005;李小红，2004；刘洋等，2011;王伟等，2012;袁勇和张固澜，2013；Chen，2009）、块匹配（Zhang and Mirko， 2018；孙成禹等，2019）、分频噪声识别与压制（蔡希玲，1999；蒋立等，2007；Bekara and Mirko, 2010; Niu and Wu, 2013；）等。以上方法基本是基于对噪声的识别—作用域变换—迭代衰减—振幅补偿这一流程，部分方法需要人工设置阈值或多次迭代计算，影响数据处理效率和精度，不利于后续的高精度处理。

AlexNet（Krizhevsky et al., 2012）网络的问世、GPU显存容量和计算能力的持续性提升为不同神经网络架构（Goodfellow et al., 2014；Simonyan and Zisserman, 2014；Shelhamer et al., 2014；Ronneberger et al., 2015；Christian et al., 2015；He et al., 2016；）的提出和应用创造了可能，也有力地推动了以卷积神经网络为代表的深度学习技术在地震数据处理中的应用（Biswas et al., 2018; Ma et al., 2018；Alwon, 2018; Mandelli et al., 2018; Xie et al., 2018；Picetti et al., 2018; Halpert et al., 2018；Das et al., 2018; Phan and Sen, 2018；Zhang et al., 2016）。

Ma（2018）、Si等（2019）、Li等（2019）分别利用不同深度的CNN对地震数据进行去噪，王钰清等（2019）利用Unet和数据增广策略对共偏移距道集数据进行噪音剔除。但以上工作所处理的噪声均具有干扰值幅值相对较小、覆盖面广的特点，即取任意的、与感受野大小相同的数据块时，其信噪比水平相当，而野值一类的异常振幅并不具有该特性，局部信噪比的差异较大且剖面上沿时间方向的连续性强。另外，U-net架构和传统CNN（DNN）相邻层之间的卷积核固定，使得高维特征无法得到充分利用，冗余的特征层和参数降低了网络优化的效率。所以仅利用残差学习结合U-net这一思路去除地震剖面中异常振幅时会损伤同相轴连续性并且去噪效果也不尽人意。

针对这一问题，本文基于并行多尺度卷积（NIN中采用多个卷积核并进行特征融合）的Inception模型（Szegedy et al, 2015; Szegedy et al., 2016a; Szegedy et al., 2016b）和残差学习策略（Zhang et.al，2017），构建了适应于野值噪声剔除的U-net网络（Inception and Residual Learning U-net，简称IRLUnet），该网络不但能够在上采样过程中融合下采样特征，同时能够充分利用每个特征图中的不同尺度特征并将相似特征在层内融合，同时设计了适合于野值去噪的卷积核。基于模拟地震记录和实际地震记录数据集的网络训练均表明，所提出的IRLUnet方法相较于常规去噪方法的智能化程度更高，且比DnCNN和DnUnet等架构具有更好的泛化和去噪能力。

## 1 方法原理

### 1.1 残差学习

基于残差学习策略的神经网络去噪—优化问题可以表示为：



其中，为神经网络待优化的参数，表示由含噪数据到纯净噪声数据的残差映射过程，表示含有异常振幅噪声的地震数据，******表示纯净（不含噪声）地震记录，表示每个批次内样本（对）的数量。

常规的Unet神经网络去噪-优化问题可以表示为：



其中，为神经网络待优化的参数，表示由含噪数据到纯净地震记录的去噪映射过程，表示含有异常振幅噪声的地震数据，***x***表示纯净（不含噪声）地震记录，*N*表示每个批次内样本（对）的数量。

若神经网络的标签为纯净地震记录，基于朴素神经网络的异常振幅压制的优化过程（式1.2），可视作求解当网络预测的纯净地震记录与用作标签的纯净地震记录在最小二乘意义下最小时网络中参数的非线性问题。将残差学习策略这一线性运算添加到朴素神经网络末端并不会引起网络参数量的变化和计算量复杂度的增加，而是将网络输出修改为，进而求解纯净噪声与预测噪声在最小二乘意义下最小时的网络参数。经简单推导可知，标签为纯净地震数据的残差网络与标签为噪声的朴素神经网络，二者在loss函数计算上保持一致（需做实验对比效率）。

### 1.2 MLP、NIN前馈网络表达能力分析

对于非线性问题的解决可以采用分段线性（Piecewise Linear，PWL）近似的方法，即用分段线性函数的组合逼近非线性函数。Chua和Kang (1977)首先提出了CR（Canonical representation, 经典表示）方法用于单值单变量、双变量、多值以及n变量分段线性函数的表示，当PWL函数具有一致变差性时，其可以用CR方法表示。Lin和Unbehauen (1995)进一步拓展了CPWL函数的表示范围并构建起了其与感知机网络之间的联系，为前馈神经网络的非线性拟合能力分析、卷积神经网络中NIN结构剖解、激活函数应用分析等提供了强有力的工具，在此我们首先引入该理论并在后文中利用其中的思想分析部分问题。

**定义1(**Lin & Unbehauen , 1995**)** **阶CPWL函数**：

假设定义在空间的紧子集上的PWL函数，被个维超平面切割为个区域，同时我们将此维超平面称为一阶切片，当D中的维线性流形能够被两个及以上阶切片所表示时，称其为阶切片。





其中为的闭包，。

当由式(1.3)、(1.4)定义的PWL函数具有一致变差性时，可将其表示为



其中。

我们称该式为一阶CPWL（Canonical Piecewise Linear，经典分段线性）函数。当所有为零时，称其为一阶CPWC（Canonical Piecewise Constant，经典分段常数）函数，当所有为零时，称其为一阶CCPWL（Continuous Canonical Piecewise Linear，连续经典分段线性）函数。

进一步地，可将阶CPWL函数可以表示为



其中，，三者均为不超过阶CPWL函数，且至少存在定义于或上的一个使得或为阶CPWL函数。

在具有个隐藏层（即层网络）的MLP多层感知机模型中，有M个节点的层第个节点的输出是由其连接的上一层所有节点的线性组合而成，若该节点连接一个线性输出单元，则其输出可表示为



其中和分别表示由隐藏层第个神经元到隐藏层第个神经元（即第个突触结构）的权重和层与层之间的偏置，表示第隐藏层神经元个数，表示激活函数。

显然，阶CPWL函数和MLP模型都具有链式迭代的形式：第阶（层）的表达式是由输入层经过逐阶（层）迭代运算得到。此外，根据MLP输出的结构还可以建立起其与高阶CPWL函数的数学联系，当取不同形式时，MLP网络的高阶CPWL具体表现形式也有所差别，例如当所有层激活函数均为单位阶跃函数或阈值函数中的一种时，由于这两种阶跃函数都可以转化为符号函数的形式，因此二者可以分别用阶CPWC函数或CCPWL函数将其表示。

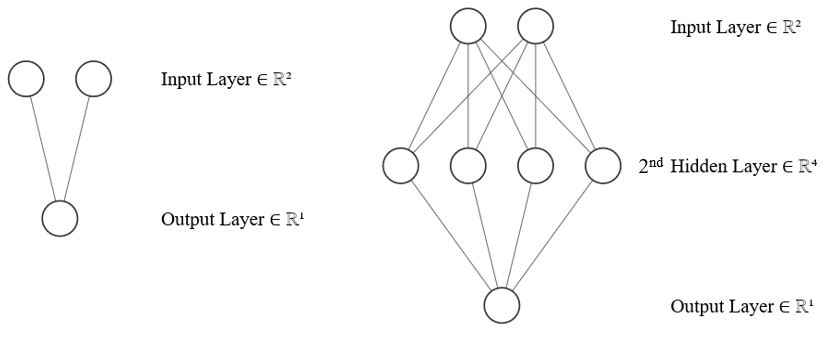
在分析MLP与CPWL关系时使用的激活函数均为挤压型函数（例如sigmoid、Heaviside、threshold、tanh等），即当输入通过该激活函数时，所有的值被压缩到一个较小的范围内，但当前计算机视觉应用中使用的主流激活函数（Maxout、ReLU及其变种）并不具有这一属性，为此我们首先讨论此类激活函数在前馈架构中的近似能力。

图1.1 MLP(左)与Maxout(右)模型

前馈架构Maxout模型（Goodfellow et al., 2013）在原多层感知器模型的两相邻层间添加k个节点并将所有节点的最大值作为总输出（式1.8），对卷积神经网络中的特征图进行Maxout则体现为沿特征图通道的池化。Maxout模型（激活函数）架构上相当于引入二级隐藏层，即隐藏层之间的隐藏层，对二级隐藏层参数的训练使其能够同时学习到一级隐藏层之间和二级隐藏层内部各神经元的相对关系，其对复杂凸函数具有很强的表达能力。同样是为了提升神经网络的非线性表达能力，Maxout之于ReLU、Sigmoid、tanh等激活函数的优势在于其表现形式并不是固定的，而是可以根据不同的学习任务自动优化调整参数。



其中，分别表示从一级隐藏层到二级隐藏层（图右从Input Layer到2nd Hidden Layer，k=4）的权重和偏置，k为二级隐藏层的神经元个数，由二级隐藏层到一级输出层（图右从2nd Hidden Layer到Output Layer）的运算即为求最大值的过程。

若定义，可将Maxout中激活函数拆解为嵌套的绝对值函数形式，嵌套阶数与隐藏层神经元个数相等，我们称式(1.9)为函数的链式迭代公式。



显然为二维空间中的连续PWL凸函数，由于可以写为 即的形式，且，所以根据两个PWL函数的复合仍为PWL函数可知仍为连续PWL凸函数，进一步地，由链式迭代公式可推得为维空间中的连续PWL凸函数。

Wang（2004）将连续PWL函数表示为lattice形式并证明了连续PWL函数可以表示为两个凸PWL函数之差的形式，因此理论上可以用两个Maxout单元之差来逼近任意PWL函数，进而用PWL函数近似任意的连续函数。

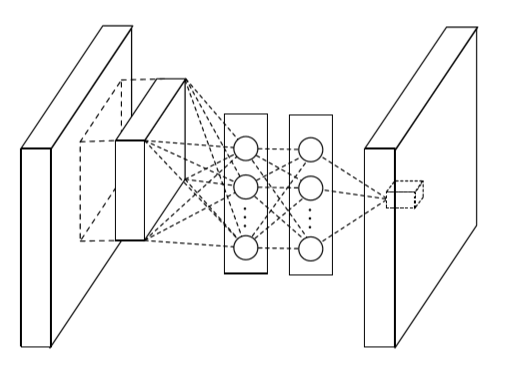


图1.2 Mlpconv层（据Lin等，2013）

Maxout在神经网络中的实际作用相当于一个激活函数，但从架构上讲其更像一个特殊的内置网络结构。与Maxout的“网络内嵌”思想相类似，NIN (Network in Network)（Lin et al., 2013）中的Mlpconv层由多个全连接层+非线性激活函数的MLP (Multilayer perceptron, 多层感知器)组合而成，其与常规卷积层相连之后，形成了由局部数据块到特征向量的非线性映射，进而滑动MLP（与卷积层操作类似）对输入数据进行扫描，使得特征图中的所有局部感受野能够共享多层感知器模型，从实现上讲Mlpconv相当于一个卷积核大小为1的卷积层。这种由多级仿射变换、ReLU激活函数以及一个线性输出结构构成的卷积神经网络本质上也是一个PWL函数，同时Xu & Wang（2018）证明了深度最多为层（个隐藏层）的ReLU DNN即可表示任意的PWL函数。

经过以上分析我们可以得出以下结论：以ReLU函数为激活的深度前馈神经网络对非线性函数有足够强的逼近能力，但网络的具体表达能力还受到卷积层组织形式、参数设置、训练集统计分布规律以及优化算法等多方面的影响，仍旧需要具体问题具体分析。

### 1.3 Inception模型

神经网络强大的特征提取能力归功于连续的仿射变换和非线性激活，但在应用上述提到激活单元和网络结构时，不难发现其层间特征图的传递是串行连续且感受野尺寸唯一的，而且Maxout单元的应用实质上是对特征图子空间的最大池化，即选取最大特征图作为子空间的代表，导致非最大特征无法传递到下一层，影响了这一激活单元这一微网络的池化能力[赵慧珍等，2017]，所以要尽量使用ReLU等常用激活函数，避免使用全连接或自适应优化类型的激活函数（层）。

诚然，在训练集规模足够大的前提下，简单而暴力的网络加宽（提高每个卷积层级的特征图数量或增加MLP每级神经元个数）、加深（使用更多连续的层）是提升DNN数据拟合能力最直接和最有效的措施。但是，神经网络规模的扩大意味着训练参数量、计算复杂度以及训练时长的非线性增长，网络参数的过度稀疏（非零权重较多）也会导致计算资源的浪费，而且当训练集样本容量有限网络的泛化能力较弱。

在DNN中，卷积、池化以及激活函数等运算均作用于输入图像的局部区域，由于核的尺寸远小于输入图像，因此深层特征图中的像素单元能够间接地连接全部或绝大部分输入，所以卷积网络相较于全连接类型的MLP等网络架构而言是稀疏的，能够有效提升统计效率。

为理解Inception模型与常规神经网络的区别，我们首先定义神经网络中作用在二维平面上的卷积计算（以下均假设卷积核窗口滑动步长为1）为



其中表示第个卷积核与的卷积，表示第个隐藏层的第个特征图，表示由到、高度和宽度分别为和的卷积核，。个卷积核将前一层尺寸为特征图映射为。若前一层输入尺寸为，想要得到含有个通道的特征图，则需要设置个的卷积核（即卷积层中的滤波器数量和尺寸），其中每个卷积核与输入进行卷积表现为对应通道进行二维卷积并沿通道方向求和。同时，我们在二维平面上定义ReLU激活函数为



则某一层所有特征的集合可以表示：



其中表示沿通道（）方向将所有二维特征图物理连接，形成厚度为的特征簇。

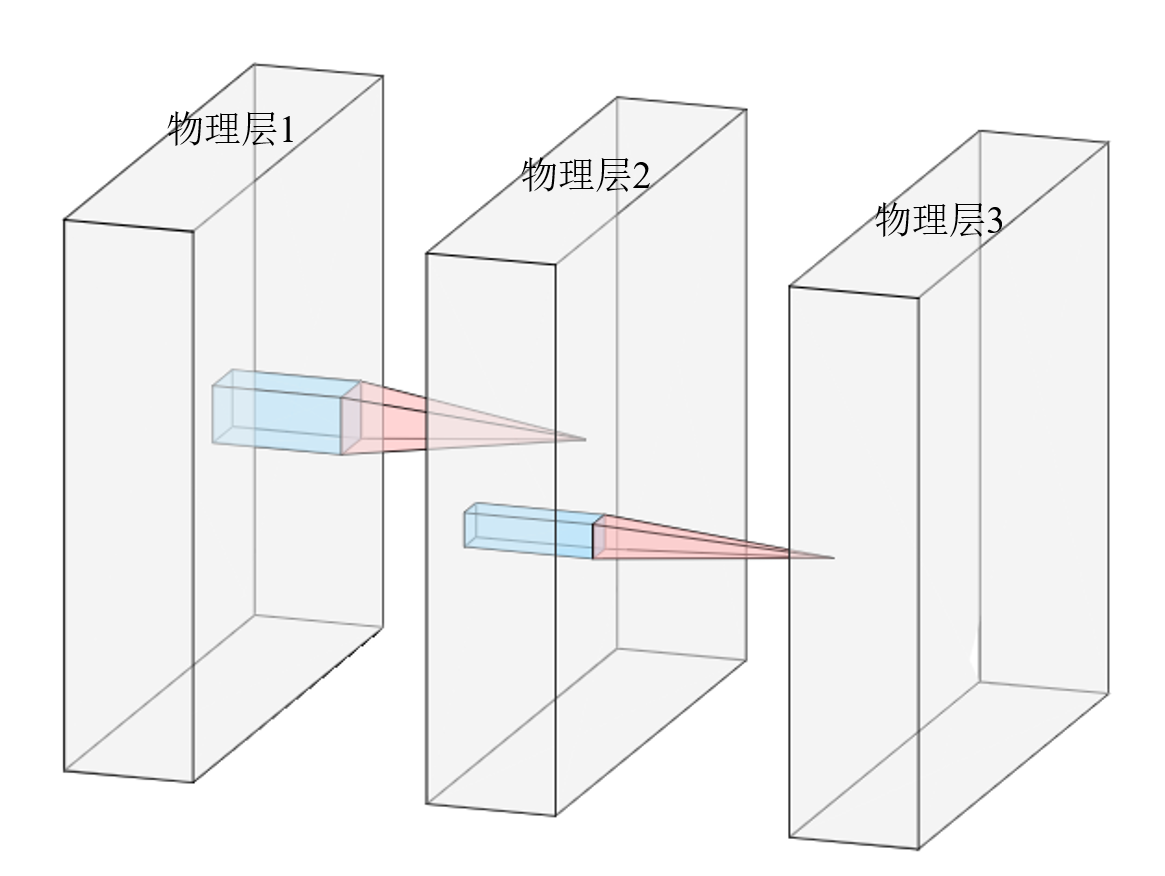


图1.3 原始卷积物理层

在常规的神经网络中（如图1.3），相邻物理层之间个卷积核权重有所区别而尺寸完全一致，因此中每个像素点对应的前一层中局部感受野大小也是固定的。这里我们将原始卷积物理层之间的部分称为层间卷积， Inception (Szegedy et al., 2015)的层间卷积（如图1.4）是由多个不同尺寸的卷积核构成，因此在相邻两个物理层内该模块对特征的提取是并行多尺度的，另外，不同感受野的所有的特征图沿厚度方向连接在一起。若不考虑Inception-v1-a（图1.4左）步长为1的Maxpooling层，可将其表示为



其中，表示不同尺寸的卷积核，在Inception-v1中。

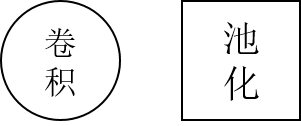
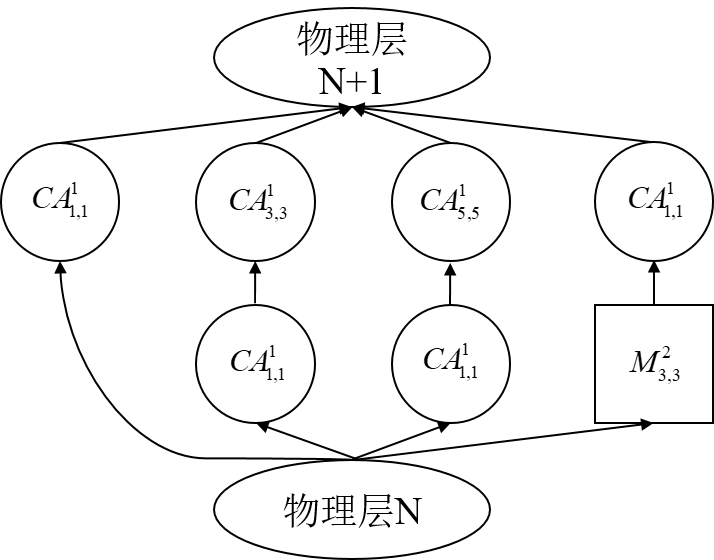
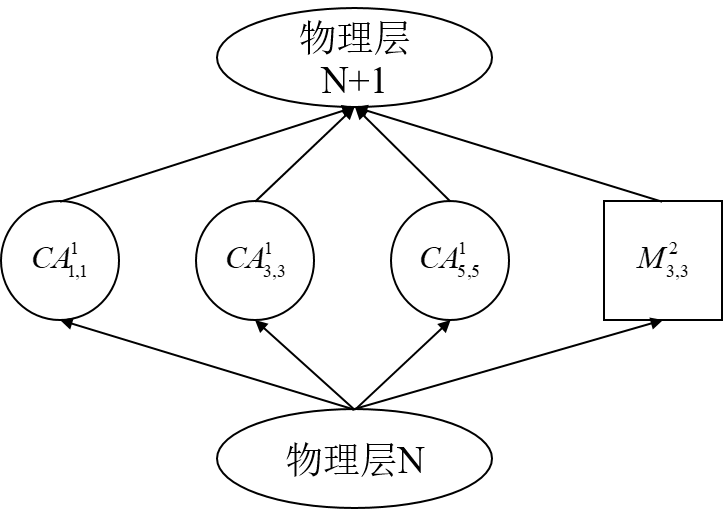


图1.4 Inception-v1-a(左) 、Inceptionv1-b(右)

图中定义为以平面尺寸为的卷积核对卷积，上标表示与卷积层相连的激活层是（1）否（0）起作用，同时将最大池化定义为，其中表示步长，表示池化窗口大小。另外，v1-b中与和相连的（即卷积）用于降低输入特征图的厚度并增加非线性程度。

此时Inception结构可以简化为式1.14（v1-a）和式1.15（v1-b）。





另外，我们可以将一个转化为（Szegedy et.al, 2015）以减少层间卷积的参数量并提升层内单通道的非线性程度。

### 1.4 IRLUnet架构

网络深层的感受野较大但对细节描述能力较弱，因此，为了在去除野值噪声的同时尽可能不损伤其它有效信号，本文采用的网络架构以深度为4（即四个下采样和四个上采样）的Unet为基础，其通过在Decoder-Encoder模型中添加跨层连接的方式使得网络在利用深层高维度特征还原原始信号的过程中能够结合下采样的高分辨率细节信息。另外，我们将（部分）原始物理卷积层替换为Inception模块并采用残差学习策略，形成IRLUnet，使其能够更好地去除野值噪声并提升网络泛化能力。

IRLUnet的架构如下所示：

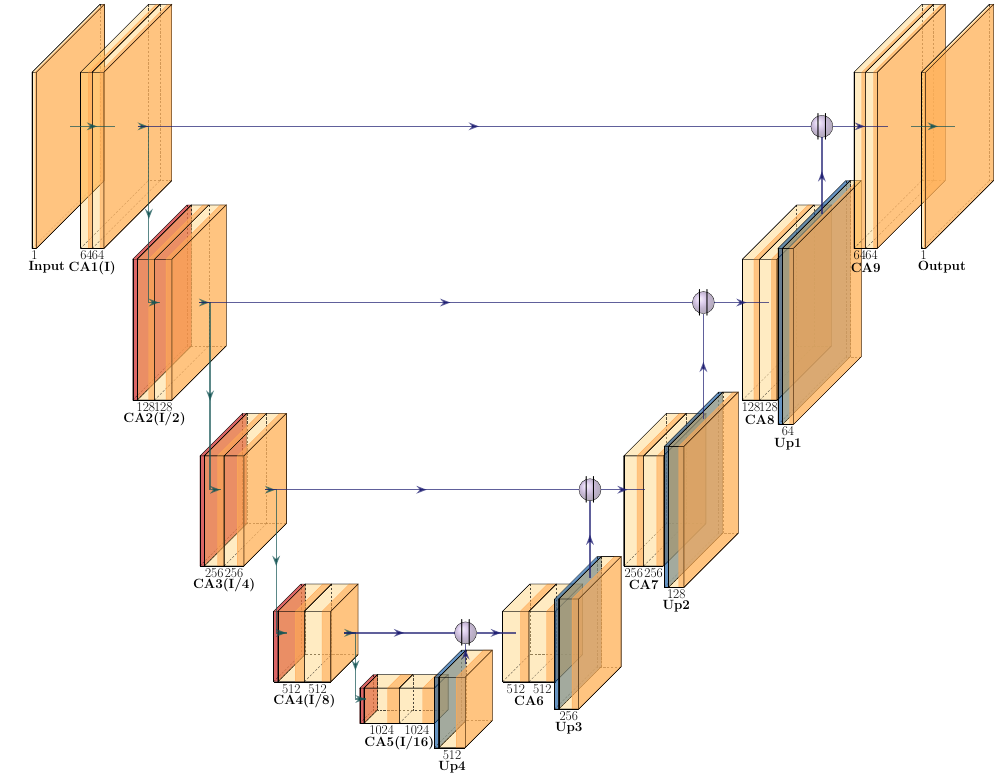


图1.5 IRLUnet架构（圆圈表示厚度方向的物理连接）

在IRLUnet中，每个物理层中都有两个NIN结构，其中物理层、、、前端各包含一个池化层（图中红色层）来对前一层的特征簇降采样，、、、对特征簇上采样（插值或反卷积，图中蓝色层）和降维（减少特征簇厚度）后分别与来自、、和的下采样输出进行物理融合（即沿特征通道方向首尾连接），然后再分别输入到、、和中对图像逐级还原，最终通过到 层的卷积降维并输出去噪后的数据。

## 2实验分析

### 2.1 数据集设计

首先，通过其野值剔除方法从数据中提取纯净的野值噪声，利用有效手段进一步将噪声中的有效信号剔除，利用变换道、复制平移以及翻转等方法增广噪声数据集并加入到正演模拟出的地震记录中，进一步利用上述手段增广加入噪声后的训练数据集，为网络训练提供充足的样本。

### 2.2 IRLUnet模型参数设置

DnCNN、DnUnet等去噪网络主要针对全局信噪比均衡的数据，所以采用方形卷积核较为合适，而野值在地震剖面中具有沿时间方向呈条带状分布的特点，所以理论上讲纵向的条状卷积核更为妥当（如图1.6）。

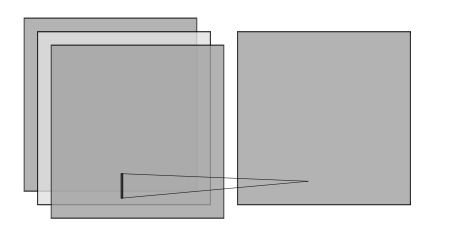
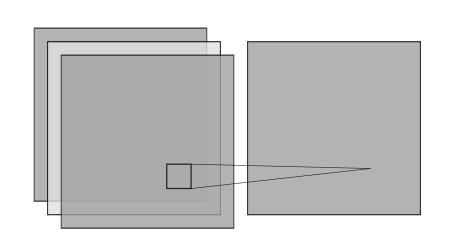
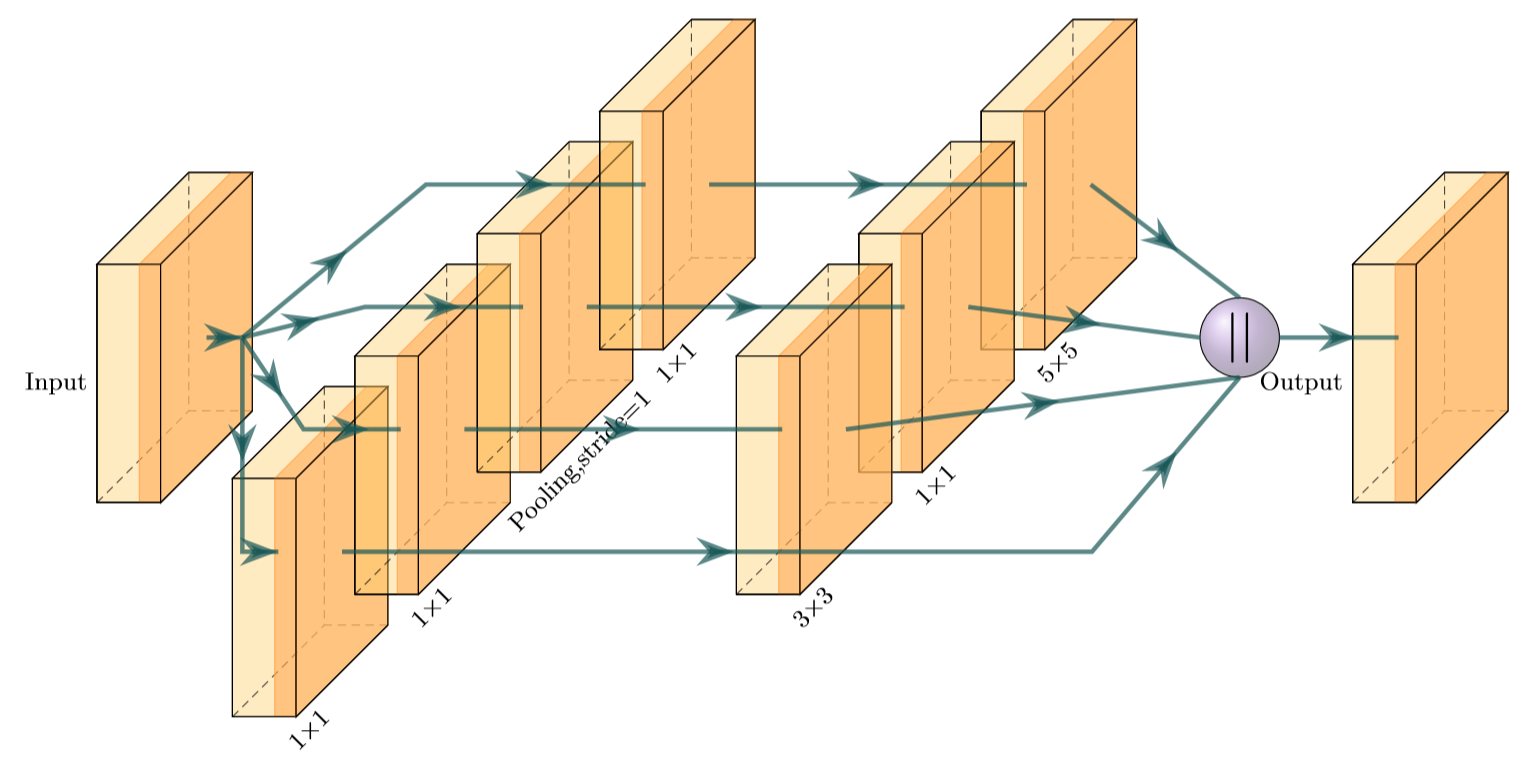
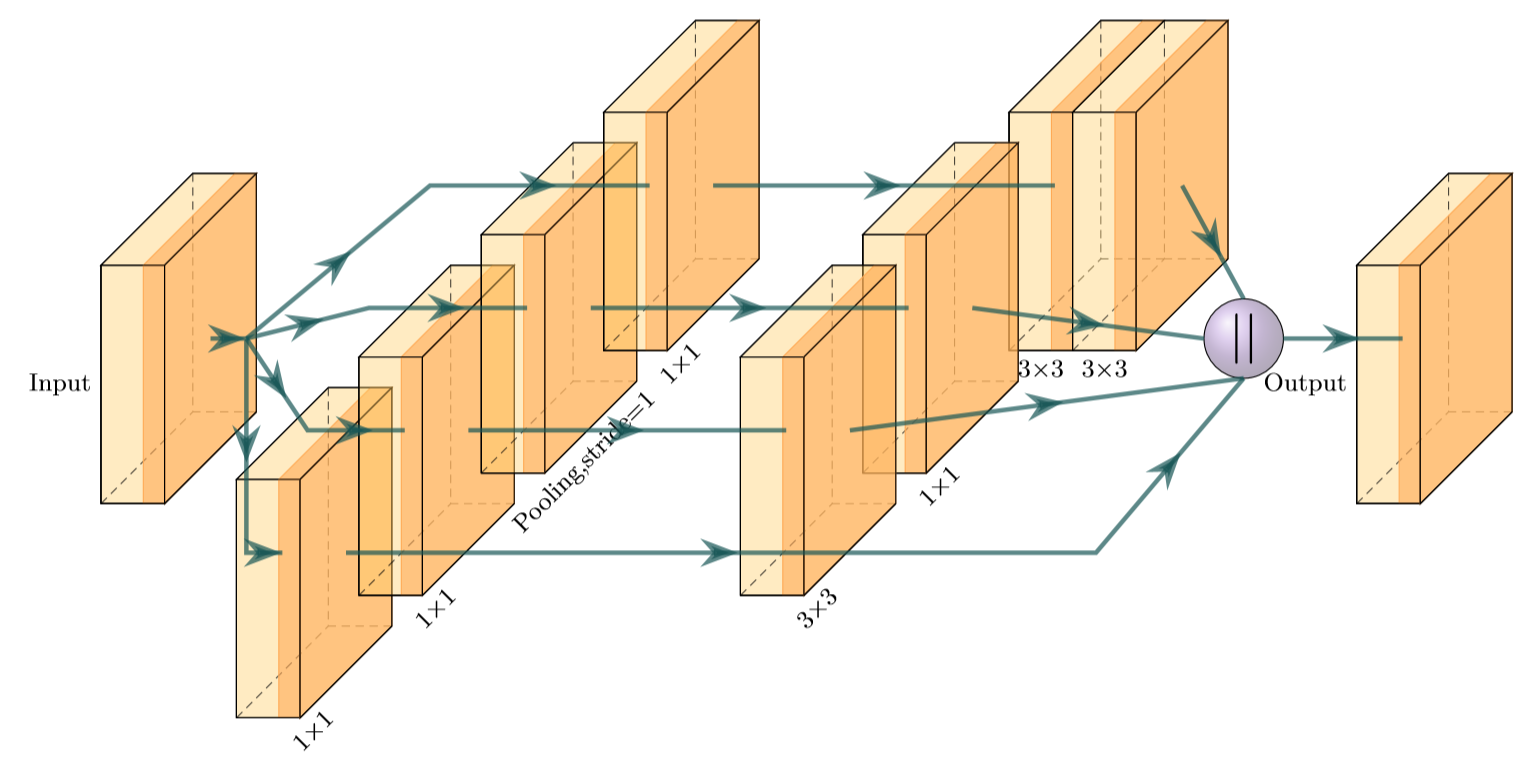


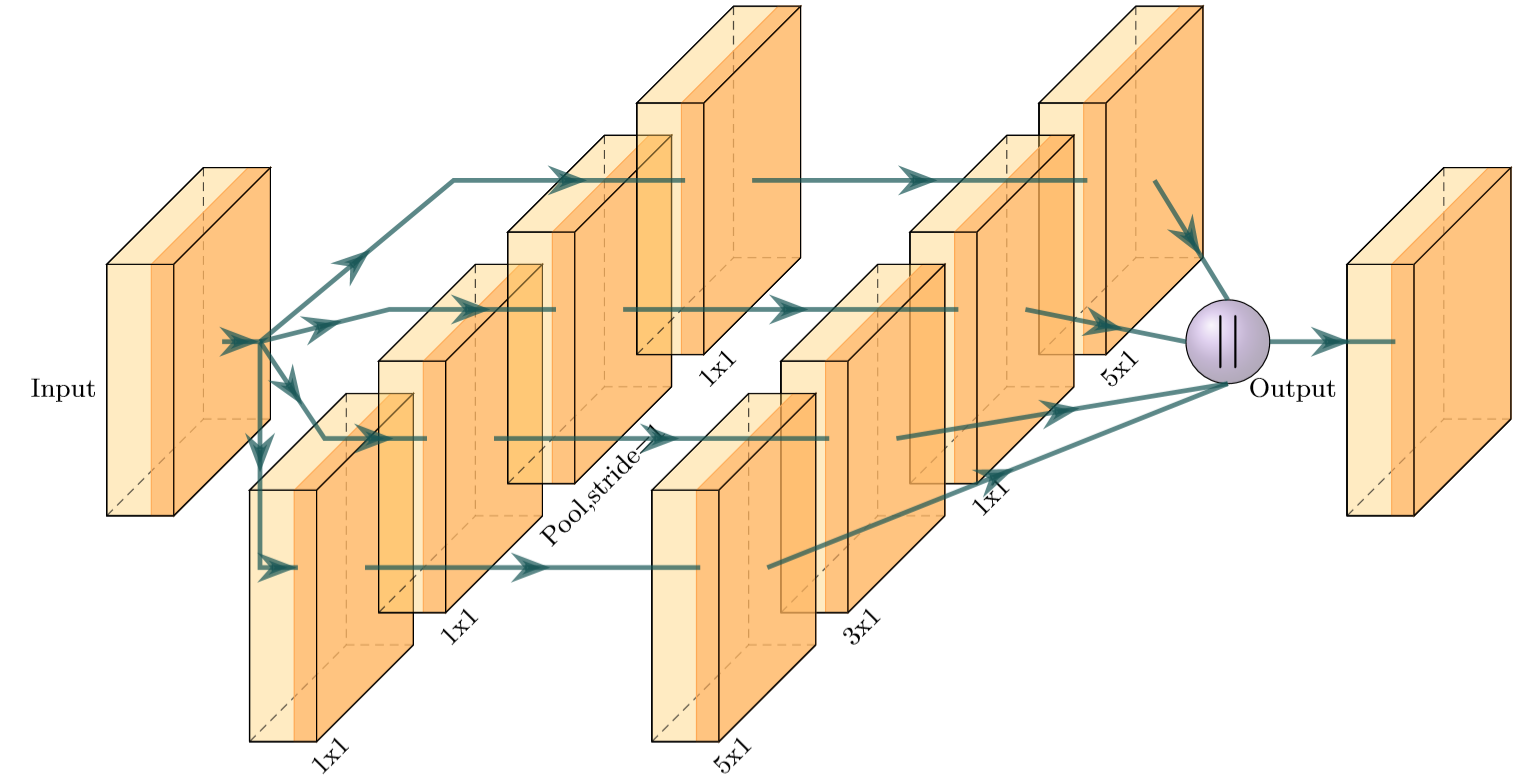
图1.6 方形卷积和条形卷积



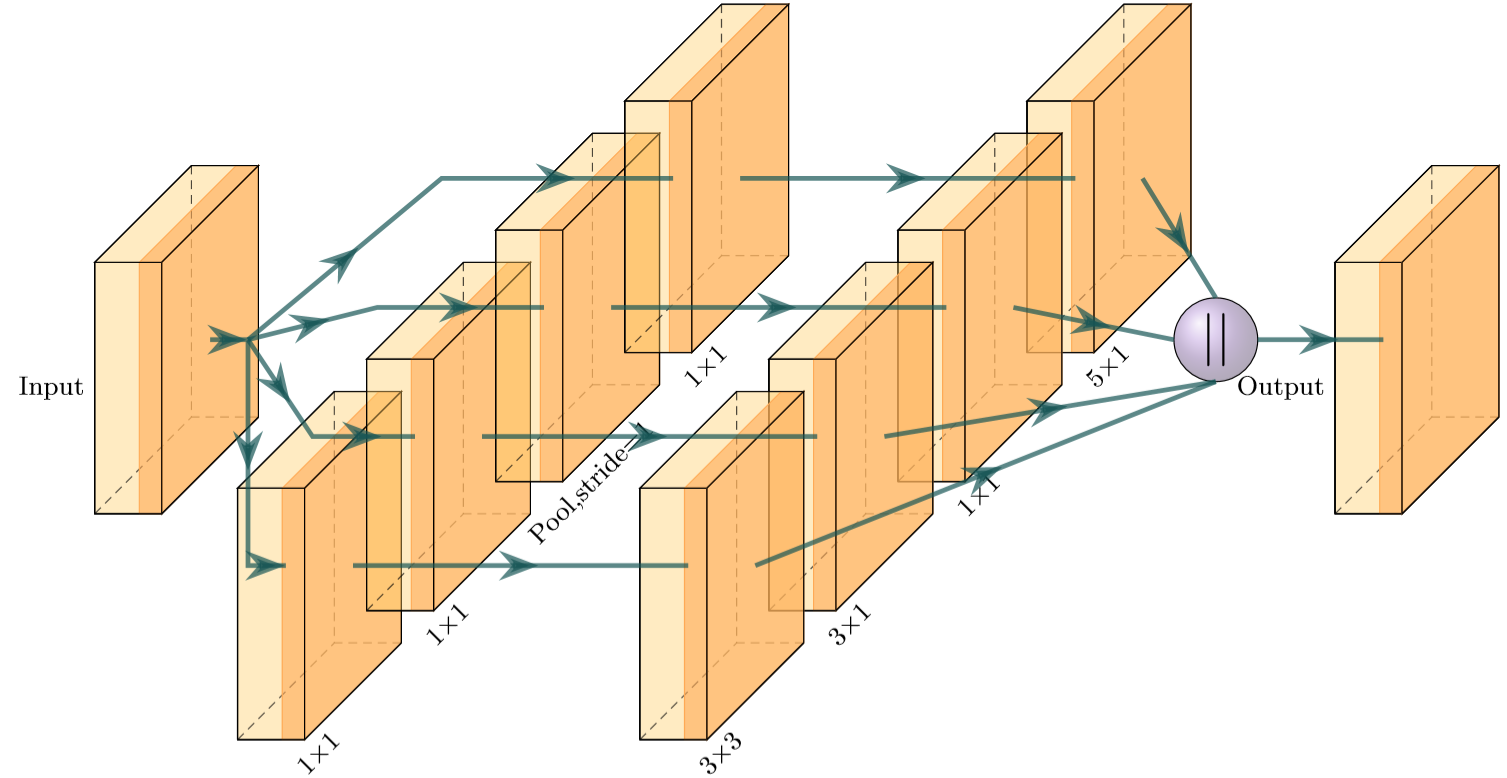
1. Inception\_v1b



1. Inception\_v2a



1. Incepiton\_IRLU\_v1



1. Inception\_IRLU\_v2

图1.7 Inception模型对比简图

另外，在Inception-v2a中，使用了两个连续的代替Inception-v1b中的,以减少训练参数量并降低优化难度。本文另外设计了Inception\_IRLU\_v1和Inception\_IRLU\_v2用于对比实验。

### 2.3 实验结果分析

#### 2.3.1 模拟数据测试

与传统Unet、Re-Unet纵向对比，与四种Inception架构横向对比。

#### 2.3.2 实际数据测试

训练集、测试集分类对比。深层特征可视化分析等

## 3 讨论与结论

本文提出的基于Inception模型和残差学习策略的IRLUnet网络用于合成数据和实际数据的实验表明，其能够在几乎完全保留有效信号的基础上有效剔除野值一类的非相干噪声，Inception模块的应用使得层间网络能够结合不同感受野范围内的有效信息，有利于不同尺度野值特征的提取、融合，以及高分辨率有效信号向深层的传递，Unet跨层传递则为去噪之后纯净地震记录的还原提供了下采样的高分辨率信息。另外，残差学习策略的引入加速了网络优化进程并且有效提升了去噪能力。