

在本文中，我们采用三维卷积神经网络来分割体积医学图像。虽然已经证明深度神经网络在许多 2D 视觉任务上非常有效，但由于有限的注释 3D 数据和有限的计算资源，将它们应用于 3D 任务仍然具有挑战性。我们提出了一种新颖的基于 3D 的粗到细框架，以有效地解决这些挑战。所提出的基于 3D 的框架优于 2D 对应物，因为它可以利用沿所有三个轴的丰富空间信息。我们分别对两个数据集进行实验，包括健康和病理性胰腺，并根据 Dice-Sørensen 系数 (DSC) 实现当前的状态。在 NIH 胰腺分割数据集中，我们的平均优于平均超过 2%，最差的情况提高了 7%，达到近 70%，这表明我们的框架在临床应用中的可靠性。

一、

在对计算机辅助诊断系统的巨大需求的推动下，医学图像（例如计算机断层扫描 (CT) 和磁共振成像 (MRI)）的自动器官分割已成为医学图像处理和计算机视觉社区中的一个活跃的研究课题。这是许多临床应用的先决条件，例如糖尿病检查，有机癌症诊断和手术计划。因此，探索自动分割系统以加速医学图像分析中的计算机辅助诊断是非常值得的。在本文中，我们关注 CT 扫描的胰腺分割，这是最具挑战性的器官分割问题之一。如图 1 所示，主要困难来自三个方面：1) 整个腹部 CT 体积中胰腺的尺寸比较小；2) 胰腺的质地，位置，形状和大小的变化巨大；3) 胰腺囊肿等异常可以改变胰腺的外观。

图 1. NIH 数据集[25]上的正常胰腺和第一和第二行中分别显示的 JHMI 数据集[33]上的异常囊性胰腺的图示。正常胰腺区域被标记为红色，异常胰腺区域被标记为蓝色。胰腺通常在整个 CT 扫描中占据很小的区域。最好看的颜色。

随着深度神经网络的快速发展[15] [29]及其在许多计算机视觉任务中的成功，如语义分割[18] [4]，边缘检测[28] [30] [27]和 3D 形状检索[35] [7]，已经提出了许多基于深度学习的方法用于胰腺分割并取得了相当大的进展[34] [25] [26]。然而，这些方法基于 2D 完全卷积网络 (FCN) [18]，其逐片地执行分割，而 CT 体积确实是 3D 数据。虽然这些 2D 方法使用策略来融合来自不同 2D 视图的输出以获得 3D 分割结果，但是它们不可避免地丢失一些 3D 上下文，这对于捕获胰腺相对于背景区域的辨别特征是重要的。

使用 3D 深度网络进行器官分割是最近的趋势，但尚未应用于胰腺。一个障碍是训练 3D 深度网络遭受“内存不足”问题。2D FCN 可以接受整个 2D 切片作为输入，但是由于 GPU 内存大小有限，3D FCN 不能被馈送到整个 3D 体积。一个常见的解决方案是从小的子体积训练 3D FCN 并以滑动窗口方式测试它们[20] [1] [5] [3] [31]，即在密集均匀采样的子体积上一个接一个地执行 3D 分割。通常，这些相邻的采样子体积彼此重叠，以提高最终 3D 结果的稳健性。值得注意的是，重叠尺寸是分割精度和时间成本之间的折衷。设置更大/更小的重叠尺寸通常会导致更好/更差的分割精度，但在测试期间需要更多/更少的时间。

为了解决这些问题，我们提出了一种基于 3D 深度网络的简洁有效的胰腺分割框架，可以同时实现高分割精度和低时间成本。我们的框架是以从粗到细的方式制定的。在训练阶段，我们首先从整个 CT 体积中采集的子体积训练 3D FCN。我们称之为 **ResDSN Coarse 模型**，其目的是通过充分利用整体 3D 背景从整个 CT 体积获得目标胰腺的粗略位置。然后，我们从仅从目标胰腺的标记的真实边界框采样的子体积中训练另一个 3D FCN。我们称之为 **ResDSN Fine 模型**，它可以根据粗略结果细化分割。在测试阶段，我们首先将滑动窗口方式的粗模型应用于整个 CT 体积，以提取胰腺最可能的位置。由于在此步骤中我们只需要目标胰腺的粗略位置，因此重叠大小设置为较小的值。然后，我们以滑动窗口方式将精细模型应用于粗胰腺区域，但是通过设置更大的重叠尺寸。因此，我们可以有效地获得精细的分割结果，我们称之为 **ResDSN C2F 的粗到细框架**。

请注意，我们框架中“粗到细”的含义是双重的。首先，它意味着 ResDSN Coarse 模型和 ResDSN Fine 模型的输入感兴趣区域 (RoIs) 是不同的，即，前者的整个 CT 体积和后者的

目标胰腺的粗略边界框。我们将此称为粗到细 RoI，旨在实现更好的分割性能。粗略步骤移除了大量不相关的背景区域，然后使用相对较小的区域作为输入进行采样，精细步骤可以更容易地学习将胰腺与局部背景区分开的线索，即利用局部环境使得 更容易获得更准确的分割结果。其次，这意味着在推理期间用于 ResDSN Coarse 模型和 ResDSN Fine 模型的重叠大小是不同的，即，它们分别具有小的和大的重叠大小。 我们将其称为粗到细重叠尺寸，其设计用于有效的 3D 推理。

据我们所知，我们是首批使用利用丰富空间信息的 3D 网络分割具有挑战性的正常和异常胰腺的研究之一。所提出的 3D 粗细框架的有效性和效率在两个胰腺分割数据集上得到证实，其中我们以相对较低的

时间成本实现了最先进的技术。 值得一提的是，虽然我们的重点是胰腺分割，但我们的框架是通用的，可以直接应用于分割其他医疗器官。

二、

医学图像分析界正面临着深度网络快速发展带来的革命[15] [29]。 基于深度卷积神经网络（CNNs）的方法在过去几年中主导了体积医学图像分割的研究领域。 一般而言，基于 CNN 的体积医学图像分割方法可以分为两大类：基于 2D CNN 和基于 3D CNN。

2.1 用于体积分割的 2D CNN

基于 2D CNN 的方法[25] [26] [11] [21] [23]从不同视图逐片地执行体积分割，然后融合 2D 分割结果以获得 3D 体积分割结果。在早期阶段，基于 2D 分割的模型是从图像补丁中训练出来的，并通过一个补丁接一个补丁的方法测试[25]，这非常耗时。自从引入完全卷积网络（FCN）[18]以来，几乎所有的 2D 分割方法都建立在 2D FCN 上，以便在训练和测试期间执行整体切片分割。Havaei 等[11]提出了一种双路径 FCN 架构，它通过两条路径同时利用局部特征和更多全局上下文特征。Roth 等[26]通过整体学习方法进行胰腺分割，该方法首先通过整体嵌套网络对胰腺区域进行分割[30]，然后通过使用随机森林进行稳健空间聚合获得的边界图对其进行细化。U-Net [23]是用于医学图像分割的最流行的 FCN 架构之一，其是编码器 - 解码器网络，但在编码器和解码器路径之间具有额外的短连接。基于胰腺仅占整个扫描的一小部分这一事实，Zhou 等人[34]建议找到粗糙的胰腺区域，然后学习基于 FCN 的固定点模型，迭代地改进胰腺区域。 他们的方法也是基于粗到细的框架，但它只考虑粗到细的 RoI。除了粗到细的 RoI 之外，我们的粗到细方法还考虑了粗到细的重叠尺寸，这是专为高效的 3D 推理而设计的。但它只考虑粗到细的 RoI。除了粗到细的 RoI 之外，我们的粗到细方法还考虑了粗到细的重叠尺寸，这是专为高效的 3D 推理而设计的。

2.2 用于体积分割的 3D CNN

尽管基于 2D CNN 的方法取得了相当大的进步，但它们对于医学图像分割并不是最佳的，因为它们不能充分利用体积数据中编码的 3D 上下文。已经提出了几种基于 3D CNN 的分割方法。 3D U-Net [5]通过用 3D 对应物替换所有 2D 操作，扩展了之前的 2D U-Net 架构 [23]。 基于 3D U-Net 的架构，V-Net [20]将残差结构[12]（短期跳过连接）引入到网络的每个阶段。Chen 等[3]提出了一种用于三维脑分割的深度体素残差网络。I2I-3D [19]和 3D-DSN [6]都通过侧输出将辅助监督纳入其 3D 深度网络。 尽管 3D CNN 作为分割目标器官的技术（如前列腺[20]和肾[5]）取得了成功，但很少有技术可用于利用有关胰腺分割的三维空间信息。 吉布森等人。 [8]提出了 DenseVNet，然而由于计算要求密集的连接，它被限制为具有浅编码器。罗斯等人[24]将 3D U-Net 扩展到细分胰腺，具有以下缺点：1) 其网络输入固定为 $120 \times 120 \times 120$ ，由于体积较大，计算量很大，2) 粗糙的胰腺边界框在其网络输入时被重新采样到固定大小，这会丢失信息和灵活性，并且无法处理形状和大小上胰腺本身存在的大变化。因此，我们提出了适用于正常和异常的 3D 粗到细框架，以确保

低计算成本和高胰腺分割准确性。

三、

在本节中，我们详细阐述了我们提出的 3D 粗粒度框架，其中包括粗糙阶段和之后的精细阶段。我们首先制定一个可以推广到粗阶段和精细阶段的分割模型。后来在 Sec. 3.1 和 Sec. 3.2，我们将分别为这两个阶段定制分割模型。

我们用 X 表示 3D CT 扫描体积。这与人类标记的每体素注释 Y 相关联，标签分割掩模 Y 具有二进制值 y_i , $i = 1, \dots, WHD$ ，在每个空间位置 i ，其中 $y_i = 1$ 表示 x_i 是胰腺体素。 $M: P = f(X; \Theta)$ 表示分割模型，其中 Θ 表示模型参数， P 表示二元预测体积。特别是在具有 L 层和参数 $\Theta = \{W, B\}$ 的神经网络中， W 是一组权重， B 是一组偏差，其中 $W = \{W^1, W^2, \dots, W^L\}$, $B = \{B^1, B^2, \dots, B^L\}$ 。假设 $p(y_i|x_i; \Theta)$ 表示体素 x_i 的预测概率是输出的最后一层的标记类，则负对数似然损失可以表示为：

$$\mathcal{L} = \mathcal{L}(X; \Theta) = - \sum_{x_i \in X} \log(p(y_i|x_i; \Theta)). \quad (1)$$

它也被称为我们的二元分割设置中的交叉熵损失。通过阈值化 $p(y_i|x_i; \Theta)$ ，我们可以得到二元分割掩模 P 。我们还为这样一个神经网络添加了一些辅助层（在本文的其余部分将称为主流网络），它在深度监督下产生侧输出。这些辅助层形成分支网络，便于在主干网络的较低层进行特征学习。每个分支网络共享来自主流网络的前 d 层的权重，其由 $\Theta_d = \{W_d, B_d\}$ 表示。除了共享权重之外，它还拥有各自的权重 Θ_d 来输出每个体素的预测。同样，辅助网络的丢失可以表述为：

$$\mathcal{L}_d(X; \Theta_d, \hat{\Theta}_d) = \sum_{x_i \in X} -\log(p(y_i|x_i; \Theta_d, \hat{\Theta}_d)), \quad (2)$$

缩写为 L_d 。最后，应用随机梯度下降来最小化负对数似然，这通过以下正则化目标函数给出：

$$\mathcal{L}_{overall} = \mathcal{L} + \sum_{d \in \mathcal{D}} \xi_d \mathcal{L}_d + \lambda (\|\Theta\|^2 + \sum_{d \in \mathcal{D}} \|\hat{\Theta}_d\|^2), \quad (3)$$

其中 \mathcal{D} 是辅助监督的一组分支网络， ξ_d 平衡每个辅助网络的重要性，并且将 L_2 正则化添加到目标以防止网络过度拟合。为了以下部分中的简明性考虑，我们保持从由等式 3 中描述的整体函数获得的分割模型，记作 $M: P = f(X; \Theta)$ ，其中 Θ 包括主流网络和辅助网络的参数。

3.1

在粗略阶段，从由 X^C 表示的整个 CT 扫描体积中采样“ResDSN Coarse”的输入，是粗分割模型 $M^C: P^C = f^C(X^C; \Theta^C)$ 在这上训练。所有的 C 上标都描绘了粗略的阶段。该阶段的目标是从复杂背景中有效地产生粗略的二元分割 P^C ，其可以去掉被分割为非胰腺的区域，具有高置信度以获得近似的胰腺体积。基于这个近似的胰腺体积，我们可以从原始输入 X^C 中裁剪出来自 P^C 的矩形立方体，以获得更小的 3D 图像空间 X^F ，与 X^C 相比，它被简单且变化较小的上下文所包围。 X^F 的数学定义表示为：

$$X^F = \text{Crop}[X^C \otimes P^C; P^C, m], \quad (4)$$

其中 \otimes 表示元素内积 每元素乘积，公式 $\text{Crop}[X; P, m]$ 表示通过一个矩形立方体裁剪 X ，该立方体覆盖了一个二进制体积 P 的所有 1 的体素，并沿 3 个轴添加了一个填充边 m 。给定 P ，对 X 施加的功能约束是它们在 3D 空间中具有完全相同的维度。填充参数 m 在实验中凭经验确定，其中它用于在精细阶段期间更好地分割胰腺的边界体素。剪裁操作可以减少维数，促进精细分割，这对于减少分割的消耗时间至关重要。值得注意的是，记录矩形立方体的 3D 位置，其指定从 X^C 裁剪 X^F 的位置，以将精细分割结果映射回其在全 CT 扫描中

的位置。

3.2

在精细阶段，ConvNet 的输入从裁剪体积 X^F 中采样，我们在其上训练精细分割模型 M^F : $P^F = f^F(X^F; \Theta^F)$ ，其中 F 上标表示精细阶段。此阶段的目标是细化前一阶段的粗分割结果。在实践中， P^F 具有与 X^F 相同的体积尺寸，其小于 X^C 的原始尺寸。

3.3

我们的分割任务是对 X^C 的每个体素进行体积预测，因此我们需要将 P^F 映射回完全相同的 X^C 大小：

$$P^{C2F} = \text{DeCrop}[P^F \odot P^C; X^F, X^C], \quad (5)$$

其中 P^{C2F} 表示最终的体积分割， \odot 表示逐个元素的替换，并且在 P^F , P^C , X^F 和 X^C 上定义的 DeCrop 操作是用 P^F 替换 P^C 内的预定义的矩形立方体，其中替换的位置由从等式 $4X^C$ 中裁剪 X^F 的定义给出。

总而言之，我们在测试过程中完整的基于 3D 的粗到细分割框架如图 2 所示。

3.4

如图 3 所示，我们提供了卷积网络架构的图示。受 V-Net [20]，3D U-Net [5] 和 VoxResNet [3] 的启发，我们有一个编码器路径，后跟一个解码器路径，每个路径有四个分辨率步骤。网络的左侧部分充当特征提取器以学习更高和更高级别的表示，而网络的右侧部分将紧凑特征解压缩为更精细和更精细的分辨率以预测每体素分段。每层的填充和步幅 (Conv, Pooling, DeConv) 都经过精心设计，以确保密集预测的输出与输入的大小相同。

左侧的编码器子网分为不同的步骤，可以处理不同的分辨率。每个步骤由一到两个卷积组成，其中每个卷积由 $3 \times 3 \times 3$ 卷积，然后是批量归一化 (BN [13]) 和整流线性单位 (修正线性单元) (ReLU [22]) 以达到更好的收敛，然后最大池化层，内核大小为 $2 \times 2 \times 2$ ，步长大小为 2，以降低分辨率并学习更紧凑的特征。通过 max-pooling 实现的下采样操作可以减小中间特征映射的大小，同时增加感受野的大小。在给定有限的计算资源的情况下，具有较少的激活函数使得在特征聚合期间可以使信道的数量加倍。

右侧的解码器子网络由几个步骤组成，这些步骤也可以在不同的分辨率上运行。每个步骤有两个卷积，每个卷积后跟一个 BatchNorm 和一个 ReLU，然后连接一个卷积核大小为 $4 \times 4 \times 4$ 和步长为 2 的反卷积，来扩大特征图，在最后一层最终预测分割掩码。通过反卷积执行的上采样操作扩大了每个步骤之间的分辨率，这中间激活函数增加了尺寸，因此由于 GPU 卡的有限存储器，我们需要将信道数量减半。

除了左子网和右子网之外，我们还建立了一个残差连接 [12] 来连接低层和高层之间的特征。在前向阶段期间，由网络提取的低级提示被直接添加到高级别提示，这可以帮助精细化精细缩放的分割，例如，靠近边界的小部分，由于高级层的感受野大，因此在特征聚合期间可能会被忽略。对于反向传播阶段，高级别层的监督线索可以通过残差连接通过快捷方式反向传播。这种机制可以防止网络梯度消失和爆炸 [9]，这阻碍了训练期间的网络融合。

我们有一个主流损耗层从“Res / Conv1b”连接到另外两个辅助损耗层，分别从“Conv2b”和“Conv3b”连接到 ground truth 标签。对于“Res / Conv1b”中的主流丢失作为与输入之一具有相同数据流大小的最后一层，遵循 $1 \times 1 \times 1$ 卷积以将通道数减少到标签类的数量，在我们的案例中是 2。对于两个辅助损耗层，反卷积层连接到上采样特征图以与输入相同。辅助损耗所施加的深度监督为超参数选择提供了鲁棒性，因为低级层与直接分段损失有

关，导致更快的收敛速度。在整个工作中，我们有两个辅助分支，其中默认参数在方程式 3 中为 $\xi_1 = 0.2$ 和 $\xi_2 = 0.4$ ，以控制深度监督与主流损失相比的重要性，以控制所有分割网络。

如表 1 所示，我们从四个方面给出了网络配置的详细比较：长剩余连接，短剩余连接，深度监督和损失函数。我们的骨干网络架构，名为“ResDSN”，与 VoxResNet [3]，3D HED [19]，3D DSN [6]和 MixedResNet [32]相比，在长剩余连接和短剩余连接的组合方面提出了不同的策略。]。在该表中，我们还描绘了“FResDSN”和“SResDSN”，其中“FResDSN”和“SResDSN”分别类似于 MixedResNet [32]和 VoxResNet [3]。我们的定量实验证实了这一点。5.1，而不是向网络添加短剩余连接，例如“FResDSN”和“SResDSN”，我们只选择长剩余每元素相加，这可以在计算上更有效，同时甚至比“FResDSN”架构表现更好。配有长短路残余连接。此外，ResDSN 在 V-Net [20]和 3D U-Net [5]方面存在明显差异。一方面，与将低级局部特征连接到更高级别全局特征的 3D U-Net 和 V-Net 相比，我们采用这些特征之间的元素之和，输出更少数量的通道以进行有效计算。另一方面，我们通过辅助损耗将深度监督引入网络，以产生更好的收敛。

4.、

在本节中，我们首先详细描述我们如何分别在粗略和精细阶段进行训练和测试。然后我们将比较我们提出的方法与两个胰腺数据集上的先前技术水平：NIH 胰腺数据集[25]和 JHMI 病理胰腺数据集[33]。

4.1 网络训练和测试

我们所有的实验都在配备 NVIDIA TITAN X (Pascal) GPU 的桌面上运行，深度神经网络基于 CAFFE [14]平台实现，该平台定制为支持所有必要层的 3D 操作，例如“卷积”，“反卷积”对于数据预处理，我们简单地将原始强度值截断为 $[-100, 240]$ ，然后将每个原始 CT 情况标准化为均值为零和方差为 1，以减少由医学图像的物理过程引起的数据差异。至于训练阶段的数据增强，与其他人使用的复杂处理不同，例如，弹性变形，我们对所有训练切片采用简单但有效的增强，即旋转 (90° , 180° 和 270°) 并在所有三个轴（轴向，矢状和冠

状）上翻转，以增加 3D 训练样本的数量，这可以通过昂贵的人类注释来减轻 CT 扫描的稀缺性。请注意，不同的 CT 情况具有不同的物理分辨率，但我们保持其分辨率不变。我们所有网络的输入大小用 $W \times H \times D$ 表示，其中 $W = H = D = 64$ 。

对于粗阶段，我们在训练阶段从整个 CT 扫描中随机采样 $64 \times 64 \times 64$ 子体积。在这种情况下，子体积可以覆盖一部分胰腺体素，或者可以从具有非胰腺体素的区域裁剪，作为难分样本挖掘减少加正向（false positive）。

False Positive：把合法的判断成非法的，译为“误报”。（相当于医学上的假阳性）

False Negative：把非法的判断成合法，译为“漏报”。（相当于医学上的假阴性）

在计算机领域中，False Positive 通常指误报，从字面上来看就是说，一个东西是被查出来了，但这是错误的(false)。False Negative 通常指漏报，也就是说，一个东西是没有被查出来的，但这是错误的(false)。

<https://blog.csdn.net/u010532666/article/details/79161093>

在测试阶段，对整个 CT 体积执行滑动窗口，其具有在每个相邻子体积内具有小重叠的粗糙步长。具体而言，对于尺寸为 $W \times H \times D$ 的测试体积，我们具有总数

$(\lfloor \frac{W}{W_I} \rfloor + n) \times (\lfloor \frac{H}{H_I} \rfloor + n) \times (\lfloor \frac{D}{D_I} \rfloor + n)$ 的子体积块被送入网络然后合并以获得最终预测，其

中 n 是控制滑动重叠的参数，较大的 n 导致较大的重叠，反之亦然。在低时间成本考虑的粗略阶段，我们设置 $n = 6$ 以从整个 CT 扫描 X^C 有效地定位在等式 4 中定义的胰腺 X^F 的粗糙区域。

对于精细阶段，我们随机裁剪 $64 \times 64 \times 64$ 子体积，限制在训练期间由 ground truth 标签定义的胰腺区域。在这种情况下，确保训练子体积覆盖胰腺素，其专门设计为能够进行细分分割。在测试阶段，我们仅在 X^F 上应用尺寸为 $W_F \times H_F \times D_F$ 的滑动窗口。要测试的子体积总数是 $(\lfloor \frac{W_F}{W_I} \rfloor + n) \times (\lfloor \frac{H_F}{H_I} \rfloor + n) \times (\lfloor \frac{D_F}{D_I} \rfloor + n)$ 。在高精度性能关注的精细阶段，我们设置

$n = 12$ 以从粗分割体积 X^F 精确估计胰腺掩模 P^F 。最后，我们将 P^F 映射回 PC ，以获得最终胰腺分割的 P^{C2F} ，如公式 5 所示，其中映射位置由 X^C 的 X^F 裁剪位置给出。

在得到最终的二元分割掩模后，我们将 P 和 Y 分别表示为预测和地面实况中的胰腺像素集，即 $P = \{i \mid p_i = 1\}$ 和 $Y = \{i \mid y_i = 1\}$ 。评估度量由 Dice-Sørensen 系数 (DSC) 定义为 DSC。该评估测量范围为 $[0,1]$ ，其中表示完美预测。

4.2

我们在 NIH 胰腺分割数据集[25]上进行实验，其中包含由经验丰富的放射科医师提供的 82 个对比增强的腹部 CT 体积。CT 体积的大小为 $512 \times 512 \times D$ ，其中 $D \in [181,466]$ ，其空间分辨率为 $w \times h \times d$ ，其中 $d = 1.0\text{mm}$ ， $w = h$ ，范围为 0.5mm 至 1.0mm 。第 4.1 节描述了数据预处理和数据增加。请注意，我们没有将空间分辨率规范化为同一个，因为我们想让网络学习处理不同体积情况之间的差异。根据训练方案[25]，我们在 82 名患者的随机分组中进行 4 次交叉验证，用于训练和测试褶皱，其中每个测试褶皱分别有 21,21,20 和 20 个病例。我们通过 SGD 优化器训练了图 3 所示的网络，其中 16 个小批量，0.9 个动量，基本学习率为 0.01，通过多项式衰减（功率为 0.9），共计 80,000 次迭代，并且权重衰减为 0.0005。粗略和精细阶段的训练网络共享相同的训练参数设置，除了它们采用从第 4.1 节中描述的不同基础分布采样的 $64 \times 64 \times 64$ 输入，其中还包括测试设置的细节。我们平均来自滑动窗口的重叠区域的得分图，并丢弃小部分小于总预测的 0.2 的小孤立预测，这可以消除小的误报。对于 DSC 评估，我们报告了所有 82 个测试案例的标准偏差，最大值和最小值统计量，如表 2 所示。

首先，我们的整体粗到细框架在平均 DSC 方面优于先前的最新技术水平 (Cai et al. [2] 和 Zhou et al. [34])，这是一个很大的改进。DSC 的较低标准偏差表明我们的方法在所有不同的 CT 病例中最稳定和最稳健。虽然由于饱和度，我们框架的最大 DSC 增强很小，但最小 DSC 的改进 (第二等) (Dou 等[6]) 从 62.53% 提高到 69.62%，这是超过 7% 的进步。最坏的情况几乎达到 70%，这是一个合理且可接受的分割结果。在从粗到细之后，在基于 3D 的粗略结果的基于 3D 的细化之后，最坏情况的分割结果改善了超过 11%。总体平均 DSC 也提高了 1.41%，这证明了我们框架的有效性。

如图 4 所示，我们在同一切片上通过“ResDSN Coarse”和“ResDSN C2F”报告分割结果以进行比较。请注意，黄色区域是正确预测的胰腺。对于 NIH 案例 # 33，这是“ResDSN Coarse”和“ResDSN C2F”报告的最小 DSC 情况，“ResDSN C2F”成功预测了底部更正确的胰腺区域，“ResDSN Coarse”明显忽略了这一点。如果粗分割是坏的，例如情况 # 33 和 # 63，我们的 3D 粗到细可以在 DSC 中显着改善分割结果多达 10%。然而，如果粗分割已经非常好，例如，情况 # 74，我们提出的方法不能改进太多。我们得出结论，我们提出的“ResDSN C2F”通过聚合丰富的空间信息显示其在 2D 方法上的进步，并且在具有挑战性的胰腺分割任务上比其他 3D 方法更强大。

4.3 JHMI 病理性胰腺数据集

我们验证了我们提出的关于腹部 CT 扫描的 JHMI 病理囊肿数据集[33]的想法。与 NIH 健

康胰腺数据集不同，该数据集包括病理性囊肿，其中一些可能或可能变成癌症。胰腺癌阶段在很大程度上影响胰腺的形态[16]，这使得该数据集在考虑大变异时极具挑战性。他的数据集总共有 131 个对比增强的腹部 CT 体积与人类标记的胰腺注释。CT 体积的大小为 $512 \times 512 \times D$ ，其中 $D \in [358, 1121]$ 的厚度范围比 NIH 数据集的厚度多。根据培训方案，我们对该数据集进行了 4 次交叉验证，其中每个测试折叠分别有 33, 33, 32 和 33 个案例。我们使用与 NIH 相同的训练设置在粗略和精细阶段训练图 3 所示的网络，除了我们在该病理数据集上训练了总共 300,000 次迭代，因为具有囊肿的胰腺比正常的更难分割。在测试阶段，我们对来自滑动窗口的重叠区域的预测图进行投票，并忽略其小部分小于总预测的 0.05 的小孤立胰腺预测。如表 3 所示，我们将该框架与该数据集上仅有一个可用的已发布结果进行比较。“ResDSN C2F”实现了平均 80.56% 的 DSC，其始终优于基于 2D 的粗 to fine 方法，这证实了沿三个轴利用丰富空间信息的优势。就平均 DSC 而言，“ResDSN Coarse”为 2.60%，这是一个显著的改进，证明了所提出的 3D 粗 - 细框架的有效性。[33]和我们的方法都有多个失败案例，其测试 DSC 为 0，这表明病理器官的分割是一项更艰巨的任务。由于这些失败案例，我们观察到该病理性胰腺数据集与 NIH 健康胰腺数据集上的结果相比存在较大偏差。

5、

在本节中，我们进行关于残余连接，时间效率和深度监督的消融研究，以进一步研究我们提出的胰腺分割框架的有效性和效率。

5.1

我们讨论了残差连接的不同组合如何对 NIH 数据集上的胰腺分割任务做出贡献。所有残差连接都以元素方式求和实现，并且它们共享完全相同的深度监督连接，交叉验证分裂，数据输入，训练和测试设置，除了残差结构彼此不同。如表 4 所示，我们仅在粗略阶段比较基于 3D 网络的残差连接的四种配置。表 1 中描述了我们的骨干网络“ResDSN”与“FResDSN”，“SResDSN”和“DSN”之间的主要差异。“ResDSN”在平均 DSC 和小标准偏差方面优于其他网络架构，即使通过网络也没有“FResDSN”那么复杂，这也是我们采用“ResDSN”来解决粗略阶段效率问题的原因。

5.2

我们讨论了所提出的粗到细框架的时间效率，在粗略阶段中较小的重叠用于低消耗时间关注，而在精细阶段中较大的重叠用于高预测精度问题。重叠大小取决于我们在第 4.1 节中定义的大小。我们在粗略阶段选择 $n = 6$ ，而在精细阶段期间 $n = 12$ 。实验结果如表 5 所示。“ResDSN Coarse”是最高效的，而准确度是三种方法中最差的，这是有道理的，我们更关心获得粗略胰腺分割的效率。“ResDSN Fine”是在整个 CT 扫描上使用大的重叠来进行最耗时的分割。在我们的粗到精框架中，我们将两个优势结合起来提出“ResDSN C2F”，它可以实现最佳分割结果，而每个案例的平均测试时间成本从 382s 减少到 245s 与“ResDSN”相比减少了 36% 精细”。相比之下，一个经验丰富的腹部放射科医生需要 20 分钟一个案例，这证明了我们框架的临床应用。

5.3

我们讨论辅助损失的有效性，以证明深度监督对我们的 3D 粗细框架的影响。基本上，我们训练我们的主流网络，没有粗糙和精细阶段的任何辅助损失，表示为“Res C2F”，同时保持所有其他设置相同，例如交叉验证分割，数据预处理和后处理。如表 6 所示，“ResDSN C2F”在最小 DSC 上在很大程度上优于“Res C2F”17.79%，在平均 DSC 上优于 0.53%，尽管在最大 DSC 上稍微差一点。我们得出结论，具有深度监督的 3D 粗到精在胰腺分割上表现更好，尤其更稳定。

6

在这项工作中，我们提出了一个名为“ResDSN”的新型 3D 网络，它集成了粗到精的框架，同时实现了高分割精度和低时间成本。骨干网络“ResDSN”经过精心设计，只有很长的剩余连接才能进行有效的推理。据我们所知，我们是首批使用 3D 网络对具有挑战性的胰腺进行分割的作品之一，该网络利用丰富的空间信息来实现最先进的技术。在广泛使用的数据集中，最糟糕的分割案例通过我们的粗到精框架进行了实验性的改进。更重要的是，我们的 coarse-to-fine 框架可以在正常和异常胰腺上工作，以实现良好的分割准确性。

图 1. NIH 数据集[25]上的正常胰腺和第一和第二行中分别显示的 JHMI 数据集[33]上的异常囊性胰腺的图示。正常胰腺区域被掩盖为红色，异常胰腺区域被标记为蓝色。胰腺通常在整个 CT 扫描中占据很小的区域。最佳浏览颜色

图 2.在测试阶段提出的 3D 粗到细分割系统的流程图。我们首先应用“ResDSN Coarse”和一个小的重叠滑动窗口来获得粗糙的胰腺区域，然后使用“ResDSN Fine”模型用一个重叠的滑动窗口来细化结果。

图 3.用于体积分割的 3D 卷积神经网络的示意图。编码器路径由“Conv1a”到“Conv4b”组成，而解码器路径从“DeConv3a”到“Res / Conv1b”。每个卷积或反卷积层由一个卷积，后跟 BatchNorm 和 ReLU 组成。为了阐明，“Conv1a, 32,3×3×3”表示具有 32 个通道的卷积操作和 3×3×3 的内核大小。“池 1, 最大, 2”表示最大池操作，内核大小为 2×2×2，步长为 2。蓝色混凝土线表示长的残余连接。具有相同颜色的块意味着相同的操作

表 1.医学图像分析中不同 3D 分割网络的配置比较。对于所有缩写短语，“Long Res”表示长剩余连接，“Short Res”表示短残留连接，“Deep Super”表示由辅助损耗层实施的深度监督，“Concat”表示连接，“DSC”表示 Dice-Sørensen 系数和“CE”表示交叉熵。对于剩余连接，它有两种类型：连接（“Concat”）或元素和（“Sum”）。

表 2.对 NIH 数据集的不同方法的评估。与以前的艺术水平相比，我们提出的框架可以大大提高现有技术水平。

表 3.对 JHMI 病理性胰腺的评估。

图 4.分别来自 NIH 案例 # 33, # 63 和 # 74 的轴向视图中的“ResDSN Coarse”和“ResDSN C2F”报告的分割结果的示例。“粗略”或“C2F”之后的数字表示测试 DSC。红色，绿色和黄色分别表示地面实况，预测和重叠区域。浏览彩色图片最佳。

表 4. NIH 上不同残余连接的评估。

表 5.测试阶段的平均时间成本，其中 n 控制推理期间滑动窗口的重叠大小。

表 6.对 NIH 的深度监督的讨论。