使用双路径 U-Net 对血管内超声图像中的动脉壁进行稳健分割

关键词：双路径 UNet DPU-Net 实时增强器血管内超声 IVUS 分割

ABSTRACT

提出了一种基于全卷积网络 (FCN) 的深度架构，称为双路径 U-Net (DPU-Net)，用于自动分割血管内超声 (IVUS) 帧中的管腔和中外膜，这对许多心血管疾病的诊断至关重要以及促进人体动脉的 3D 重建。医学图像分析中最普遍的问题之一是缺乏训练数据。为了克服这个限制，我们提出了双重解决方案。首先，我们引入了一种深度架构，该架构能够使用少量训练图像进行学习，并且仍然具有高度的泛化能力。其次，我们通过实时增强器控制图像增强过程来增强所提出的 DPU-Net。我们的实时增强器包含专门设计的操作，可模拟三种类型的 IVUS 伪影并将它们集成到训练图像中。我们详尽地评估了我们对 Balocco 标准公开可用的 IVUS 20 MHz 和 40 MHz B 模式数据集的双重贡献，这些数据集分别包含 109 个训练图像、326 个测试图像和 19 个训练图像、59 个测试图像。使用提供的训练图像从头开始训练模型，并使用 IVUS 分割文献中的两个常用指标进行评估，即 Jaccard Measure (JM) 和 Hausdorff Distance (HD)。实验结果表明，DPU-Net 在 40 MHz 数据集上分别实现了 0.87 JM、0.82 mm HD 和 0.86 JM、1.07 mm HD，用于分割流明和媒体。此外，DPU-Net 在 20 MHz 图像上分别实现了 0.90 JM、0.25 mm HD 和 0.92 JM、0.30 mm HD，用于分割流明和媒体。此外，DPU-Net 在 HD 距离方面比现有方法高出 8-15%。 DPU-Net 还显示出强大的泛化特性，可用于预测测试集中的图像，这些图像包含大量主要伪像，例如训练集中不常见的分叉、阴影和侧分支。此外，DPU-Net 在 0.03 秒内运行以使用单个现代 GPU（Nvidia GTX 1080）分割每一帧。拟议的工作利用基于现代深度学习的方法对 20 MHz 和 40 MHz IVUS B 模式图像中的管腔和介质血管壁进行分割，并在没有任何人工干预的情况下实现了最先进的结果。该代码可在 <https://github.com/Kulbear/IVUS-Ultrasonic> 在线获得。

1. Introduction

心血管疾病由于其高发病率、高死亡率和不可逆的后遗症，需要准确/快速的检测和治疗。 血管内超声 (IVUS) 是一种成像方式，常用于协助医务人员并帮助他们诊断心血管疾病。 IVUS 为医学专家提供了一个由内而外的冠状动脉视图。 为了获取 IVUS 帧，将携带超声发射器的导管插入冠状动脉，以提供动脉的横截面断层扫描视图。 尽管 IVUS 帧允许评估血管形态 [1]，但获取的冠状动脉图像对于人眼来说并不容易解释。毫无疑问，从 IVUS 框架自动准确地标记（分段）血管壁对许多应用都是有益的。 例如，医务人员可以受益于圆形层边界的自动标记来诊断心血管疾病。 因此，获得的 IVUS 图像的分割具有重要的临床意义，尽管它一直是一项具有挑战性的任务，因为 IVUS 图像通常包含显着的成像伪影。 特别是，IVUS 图像中内部（腔）和外部（介质）血管壁的准确分离在创建精确的动脉 3D 重建和诊断心血管疾病方面起着至关重要的作用，其中冠状动脉的这种定量测量 边界会影响临床决策。

从传统的角度来看，IVUS 图像的分割是一个经过充分研究的问题，其中使用了许多计算机视觉和图像处理的想法和方法 [2-8]。较旧的方法使用各种策略，例如形状和强度先验[6]、非参数能量函数中的梯度矢量流 [5]、具有概率成本函数的参数可变形模型 [4,2]，甚至射频信号 [3] ] 来分割管腔和介质。近年来，大量的研究得到了成功的实验结果的支持。在 [7,8] 中，作者利用一种称为极值水平极值区域 (EREL) [9,10] 的区域检测器来聚类感兴趣的区域，即管腔和介质。为了在 IVUS 框架中分割动脉壁，Yang 等人。 [11] 和 Kim 等人。 [12]，提出了众所周知的 U-Net [13] 的两个衍生物，它们基于深度卷积神经网络的概念。自 2012 年 AlexNet 在 ImageNet 竞赛 [14] 中取得成功之后，卷积神经网络 (CNN) 在视觉图像识别中发挥着重要作用。语义分割是计算机视觉和图像处理中最活跃的研究领域之一，已成为大量研究的主题，例如 [15-19]。 [16] 中提出了一种非常流行的架构，它试图通过以完全卷积的方式构建网络架构，将从图像分类任务中学到的知识转移到语义分割中。从那时起，全卷积网络（FCN）被广泛用于解决像素级密集预测的问题。

深度学习 [14] 出现后几年过去了，尽管许多基于深度学习的方法几乎主导了与计算机视觉和医学图像分析相关的每个研究领域，但基于深度学习的方法用于血管内超声分割的并不多帧。在这个领域只进行了少量研究的主要原因之一是缺乏足够的训练数据。许多深层架构需要非常大量的训练示例才能实现高质量的泛化能力。然而，可用的 IVUS 数据集包含非常少量的训练图像。例如，Balocco 的数据集 [1] A 仅包含 19 个训练图像。拥有少量训练图像会显着降低深度模型的性能，尤其是在 IVUS 分割中，并且由于大多数最先进的预训练模型都是在自然照片图像上训练的 [14,20]，迁移学习不能作为 IVUS 分割的有效选项，因此需要从头开始训练模型。

为了克服上述问题，我们在本文中提出了一种深度架构，它不仅能够比当前的深度模型更好地泛化（在少量图像上训练后），而且不需要初始化一个预训练的模型。 DPU-Net，我们提出的全卷积网络是基于 UNet 架构构建的，是 IVUSNet [11] 的改进扩展，可以自动描绘管腔和媒体血管壁的边界。我们通过将所提出的 DPU-Net 与在相同数量的帧上训练相同时间的其他两种最先进的架构进行比较来评估其泛化能力，即 SegNet [21] 和 UNet [ 13]。我们的结果表明，DPU-Net 在由 Jaccard Measure (JM) 表示的分割精度方面优于它们两者，或者称为联合交集 (IoU) 和 Hausdorff 距离 (HD)。 DPU-Net 可以使用少量图像进行训练并实现更好的泛化，这一事实使其成为各种医学图像分析问题的合适选择。

除了提议的 DPU-Net 之外，我们还引入了几种增强操作，这些操作专门用于减轻三种常见 IVUS 伪影的影响，即阴影、侧血管和分叉。 这些操作包含在我们的实时增强框架中，该框架能够根据要求尽快生成用于训练的增强图像。

我们在公开可用的 IVUS B 模式基准数据集的两个测试集上评估了拟议的工作 [1]。 这些集不仅包含少量的训练图像，而且还具有基于与训练集不同的工件分布生成的测试集。 评估结果表明，实时增强器增强的拟议 DPU-Net 优于所有现有的最先进方法。 此外，由于我们基于 IVUS 的增强操作，被伪影污染的图像的准确性结果远远优于其他自动方法。 我们的工作是一种端到端的方法，不需要人工干预。

论文的其余部分组织如下：第 3 节详细解释了我们提出的架构。 2.1 节说明了我们如何构建增强管道。 我们还展示了多个实验，这些实验加强了我们在第 4 节中的贡献。最后，我们在第 6 节中总结了我们的研究。

1. Datasets

我们利用公开可用的 IVUS 数据集 [1] 来验证 DPU-Net。该数据集旨在用于可能需要单帧或多帧数据集的不同方法 [1]。因此，有两个可用的数据集是用不同的两个超声频率（包括 20 MHz 和 40 MHz）获得的。请注意，这两个数据集以不同的分辨率获得并存储，20 MHz 数据集为 384×384，40 MHz 数据集为 512×512。 20 MHz 数据集包含两组（训练和测试）IVUS 门控帧，使用来自 10 名患者在心脏舒张末期完全回拉。数据集框架由四位临床专家手动注释。具体而言，其中两人在第一次标记后一周重复该任务 [1]。训练和测试集中分别有 109 和 326 个 IVUS 帧。此外，测试集包含大量 IVUS 伪影，包括分叉（44 帧）、侧血管（93 帧）和阴影（96 帧）伪影。剩余的 143 帧不包含除斑块之外的任何伪影。

40 MHz 数据集还带有两个预定义集（训练和测试），它们都包含带注释的地面实况。 训练集和测试集分别由 19 个和 59 个 IVUS 帧组成。 训练图像数量较少会显着增加训练泛化深度模型以产生合理水平的结果的难度。 我们解决了使用 DPU-Net 以及主要数据增强的有限训练示例的问题，如第 2.1 节所述。

* 1. Augmentation

数据增强的想法是通过增强观察到的数据来努力填充缺失和看不见的值，并且在机器学习中已经研究了几十年 [22]。然而，在大型深度模型出现后，“对数据的渴望”变得更加强烈，尤其是在计算机视觉和图像分析中，问题的输入域的大维度会在潜在空间中产生显着的可变性。因此，一旦决定采用深度模型，扩充训练集就成为必然，尤其是当我们没有庞大而多样的训练集时，这在医学图像分析应用中总是如此。由于获取医学图像的一系列困难，大多数医学数据集包含的图像数量相对较少。例如，Balocco 数据集 A 和 B [1] 分别仅包含 19 和 109 个训练图像，这突出了人为增加训练集的必要性。为了最好地利用训练集中的可用信息，我们设计了一个实时增强类，能够同时转换和扭曲训练图像，以及一些专门为 IVUS 图像设计的新操作。

* + 1. Augmentation operations

文献中已经尝试了各种类型的增强（操作）来创建新的训练图像：弹性失真用于视觉文档分析 [23]，RGB 色移、平移、缩放、水平剪切、水平翻转和旋转变换是示例[24-26,14,20,27,17,13] 中已经考虑过这些。另一方面，在医学应用中，并非所有类型的增强都可能有用。事实上，医学图像增强（转换）的类型在很大程度上取决于应用。因此，执行冗余增强会增加训练时间，而不会实现显着的准确度增益 [28]。例如，平移（移动）IVUS 帧听起来并不合理，因为导管始终位于图像的中心。因此，模型永远不会看到导管位于其他地方而不是框架中心的 IVUS 框架。相比之下，旋转和翻转似乎是合理的，因为导管很有可能在血管内旋转或移动。因此，在本节中，我们将探讨各种类型的现有转换对训练模型准确性的影响。我们还介绍了三个专为增强 IVUS 帧而设计的新过滤器。我们表明，在训练帧上应用现有转换和我们设计的掩码的组合可以赋予模型更大的泛化能力，从而提高最终分割的准确性。

* + 1. Proposed augmentation operations

在各种传统研究中已经多次实践了设计特定策略以减轻超声伪影的影响。 在 [29] 中，拖动圆形光标以覆盖导管伪影。 在[30]中，该方法添加了一个预处理步骤以清除运动伪影。 许多其他研究也努力包括伪影检测步骤 [6,31-33]。 在拟议的研究中，我们不仅使用上述常见的增强操作，如旋转和重新缩放，而且我们还设计了三种不同类型的增强来模拟常见的 IVUS 伪影，即分叉、侧血管和阴影。 这些操作包含在我们的增强管道中，以提高 DPU-Net 架构在处理多种类型的 IVUS 工件时的稳健性。

为了模拟分叉、侧血管和阴影伪影，我们设计了三个不同的掩码，其尺寸等于训练图像的大小。 这些掩码中的每一个都旋转了 360 次。2 图 1 说明了所有掩码。 然后可以在增强时将设计的掩码与训练图像相乘。 在增强期间使用掩码的一个优点是其计算成本低，因为应用掩码只需要乘法运算，不需要变形或卷积。 使用这种类型的增强的另一个好处是我们可以同时对一个图像应用多个操作。 考虑到模拟阴影、侧血管和分叉的掩膜被称为阴影掩膜、侧血管掩膜和分叉掩膜，并用 SH、SV、BF 表示，一个新的增强图像 (I\*) 可以得到如下： 

其中 I 表示原始训练图像，\* 表示元素乘法。 需要注意的是，也可以定义基于每个特定掩膜的其他操作，例如  仅应用阴影掩膜，  仅应用侧脉掩膜，  仅应用分叉掩膜， 应用侧脉 和分叉掩码一起等等。 图 2 显示了随机选择的三个掩码以及将它们应用于 Balocco 数据集 [1] 的 20 MHz 训练帧之一后的结果图像。

* + 1. Real-time augmenter

从实用的角度来看，大多数类型的增强都是通过对图像应用特定的变换（例如相似性或仿射）来完成的。这表明图像增强本质上是一项计算成本很高的任务，因为在对训练图像应用变换之后需要执行昂贵的变形过程（包括插值）。增加新图像的繁琐过程显着增加了训练时间。因此，为了防止训练过程被延迟，新图像以离线或在线形式增强。基本上，如果增强训练集由少量图像组成并且有足够的可用磁盘空间，则可以在训练过程开始之前增强训练集并将其保存在磁盘上。这种创建新图像的方式称为 oﬄine 增强。然而，由于高磁盘空间要求，离线增强在许多倾向于在大型数据集（如 ImageNet [34] 或 Microsoft coco [35]）上训练大型模型的应用中并不实用。一个简单的补救措施是让 CPU 增强图像，同时在 GPU 上执行训练过程，这称为在线增强。该技术已在许多最先进的深度学习架构的训练过程中采用 [14,26]。但是 GPU 硬件和软件的最新进展使得图像的前馈和后向传播速度非常快，因此即使在 CPU 上进行少量计算也会延迟训练过程。因此，在我们的研究中，我们提出了一个实时图像增强器模块，该过程在后台运行并控制许多旨在增强所需图像的并行进程。训练过程（在 GPU 上运行）可以立即从始终充满增强图像的缓冲区中读取图像。一旦从缓冲区中读取图像，在每个 CPU 内核的后台运行的进程就会扩充另一个图像并将其放置到一个空位置。图 3 说明了我们提出的实时增强管道是如何工作的。

1. Proposed method

自 [14] 以来，深度学习 (DL) 方法已被证明可以在许多视觉识别任务上实现卓越的性能。 特别是，卷积神经网络 (CNN) 及其衍生物在几乎所有可以被视为监督学习任务的视觉识别任务中都优于传统方法 [20,18,21,27]。 在本节中，我们首先介绍用于训练和验证建议工作的流行 IVUS B 模式数据集。 然后，我们提出了一个双路径 UNet (DPU-Net)，它为管腔或介质区域生成二进制预测掩码以描绘血管壁，并在其直觉和设计方面进行了详细解释。

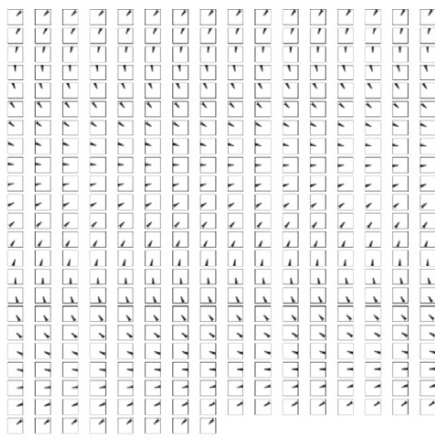
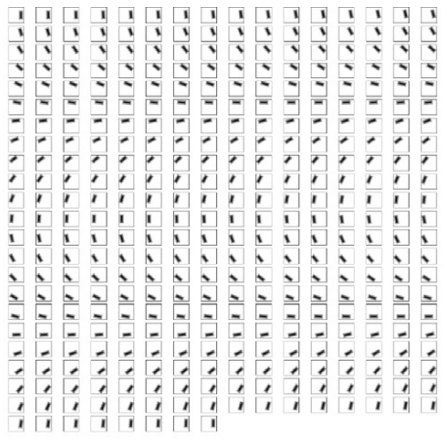
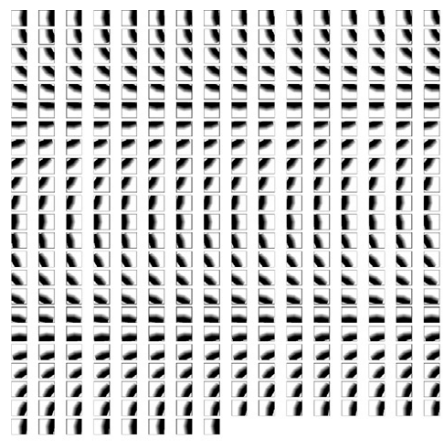
* 1. Dual Path UNet

Dual Path UNet (DPU-Net) 的设计结合了人类感知的直觉和从现有流行的完全卷积网络 (FCN) [16,21,36] 中吸取的经验教训以及 FCN 的分段改进。 根据整体架构设计，我们采用 U-Net [13] 作为我们提出的工作的基础架构。 与流行的 1-stage 架构一样，DPU-Net 有两个主要组件：

1. 一个编码器网络，可以对输入进行下采样和处理以生成低分辨率的深度特征图。

2. 一个解码器网络，可以将编码器网络输出的深度特征图的分辨率恢复到原始大小。

由于计算能力和内存有限，我们首先对输入图像进行下采样，然后再将它们提供给深度模型，并在模型结束时对预测结果进行上采样。 本节稍后将解释此过程。 上采样的输出特征图被发送到另一个卷积层，然后是 sigmoid 激活以产生最终结果。



1. (b) (c)

Fig. 1. 360 designed masks for emulating the common artifacts of IVUS frames.

1. Designed shadow masks (SH).
2. Designed side vessel masks (SV).
3. Designed bifurcation masks (BF)

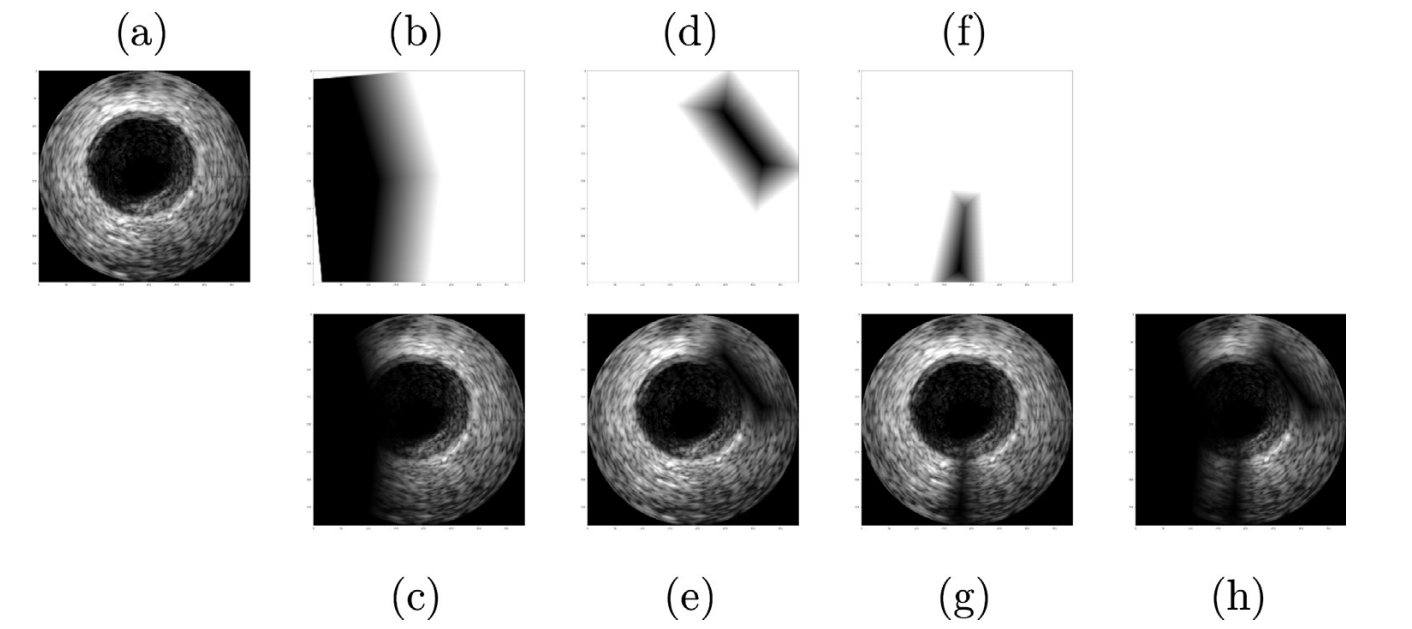


Fig. 2. The proposed artifact simulation operations to mimic the three common IVUS artifacts.

1. An original training Image .
2. Designed SHadow mask .
3. The image obtained after multiplying the shadow mask (b) by the image .
4. Designed Side Vessel mask .
5. The result of multiplying the image by the side vessel mask .
6. Designed BiFurcation mask.
7. The result of multiplying the image by the bifurcation mask.
8. The image obtained after applying all of the masks designed to the original image at the same time .

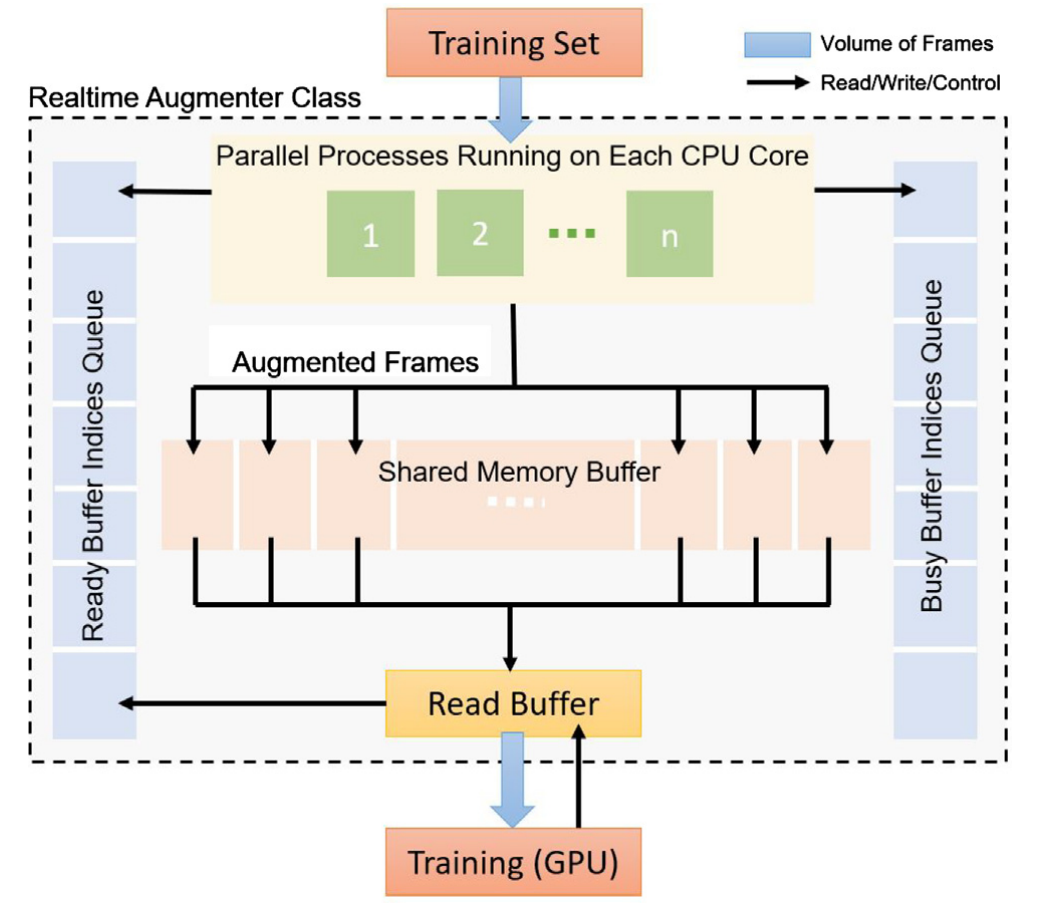


Fig. 3. The proposed real-time augmente

编码器网络包含 6 个编码块，而解码器网络包含 5 个解码块。 从网络中的第二个块开始，每个块从其前一个块接收特征图； 特别是对于每个解码器块，有一个额外的跳过连接可以帮助从编码器网络转发信息。 跳过从编码器网络到解码器网络的连接提供额外的信息，以帮助将特征图恢复到原始大小。 特别是，像素之间的空间关系被保留，因为跳跃连接实际上连接了编码器和解码器之间的相应块。 请注意，这些跳跃连接还可以帮助加强深度模型中的梯度流，从而避免常见的梯度消失问题并加快训练过程 [37]。 因此，整个架构是对称的，如图 4 所示。架构中的块之间存在细微差别。

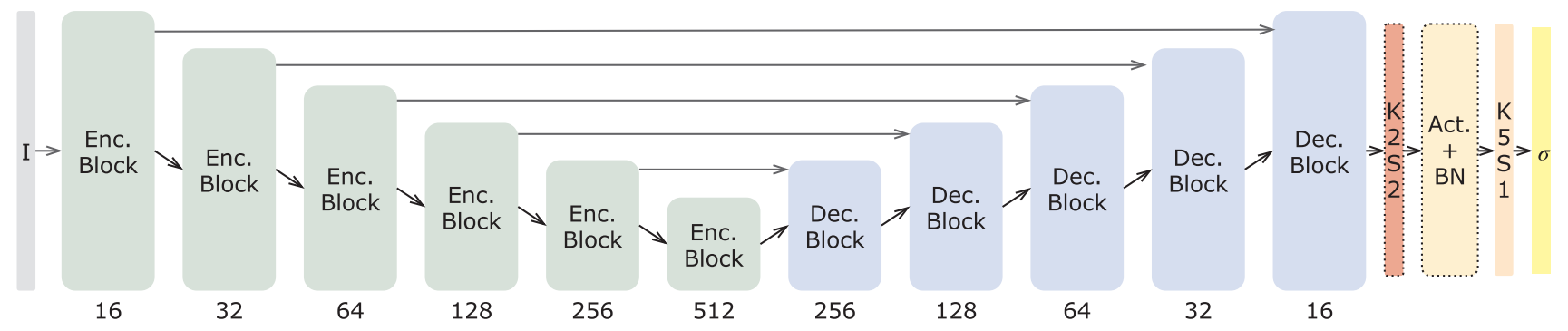


图 4. DPU-Net 架构。 同一块中的每个卷积层都具有与块底部标记的相同的输出深度。 最后，红色层表示步长为 2 的 2×2 转置卷积（反卷积），其中 sigmoid 输出（标有黄色）之前的层是 5-by 5 卷积层。 图中使用的缩写含义如下：“K2S2”表示“kernel size 2 and stride size 2”，“BN”表示“batch normalization”。 （为了解释这个图例中对颜色的引用，读者可以参考本文的网络版本。）

池化层有两种常见的选择，即平均池化和最大池化。 最大池化广泛用于对从前一层获得的输入特征图进行下采样。 我们选择最大池化而不是平均池化，因为我们认为 IVUS 图像在低分辨率下相对模糊。 最大池化迫使网络在内核大小的子区域中捕获最活跃的神经元中的信息，但丢弃其他不重要的信息。

除了第一个编码块，每个编码块都包含一个下采样分支，对接收到的特征图进行下采样，然后是一个两分支的卷积路径，如图 5（a）所示。 为了避免使用最大池化导致的信息丢失并降低输入的空间分辨率，我们通过使用 2×2 平均池化层和 3×3 卷积层来构建和扩展下采样分支 步幅为 2。最后，我们在深度维度将两个输出连接在一起。 这种聚合思想类似于[27,36]。

下采样后，下采样分支输出的聚合特征图被传递到两个后续分支，即精炼分支和主分支。 首先，我们按照 [21,13] 中的设计包含一个具有连续卷积层的分支，然后是激活和批量归一化，这里我们称之为主分支。

最近的一个趋势是使用小内核大小进行特征图细化 [17,19]，网络 [27] 中的网络概念也在文献中广泛使用。因此，我们引入了一个精炼分支，它具有一个 3×3 内核大小的卷积层，后跟一个内核大小为 1×1 的卷积层，以生成相似但经过改进的特征图。 1×1 卷积能够细化或修剪特征图，因为它仅覆盖单个像素而不受其邻居的影响。但在所有深度上，这个想法类似于全局平均池化 [38]，具有更多的学习能力。此外，由于通常需要在不同尺度上捕获特征的能力，我们在主分支中设置了内核大小为 5 的卷积层，相比之下，在精炼分支中的内核大小为 3 和 1。主分支和精炼分支的输出相加并传递到下一个块及其相应的解码块。一个关键问题是，由于梯度消失问题，深度网络难以训练。多分支和本地网络中网络架构不仅提供了良好的本地拓扑，而且还加强了梯度流以加速训练。

解码块需要稍微不同的配置，如图5（b）所示。 每个解码块都从其前一个块和相应的编码块接收特征图。 只有从前一个块接收到的特征图通过 2×2 转置卷积进行上采样，然后与来自其相应编码块的特征图连接。 请注意，此连接的特征图只会传递到主分支，其中精炼分支仅处理上采样的特征图。

DPU-Net 中使用的激活是参数整流线性单元 (PReLU) [39]。

与普通的 ReLU 激活相比，PReLU [39] 允许一部分梯度在未激活时流过神经元，而 ReLU 仅在神经元处于活动状态时通过梯度。 正如 [39,40] 中所建议的，PReLU 在许多基准测试中都优于 ReLU，并且性能也更稳定。

为了减少训练过程中涉及的计算成本，我们最初将输入图像下采样了 2 倍（即，20 MHz 数据集中的 384 x 384 图像将被调整为 192 x 192 图像）。 但是，ground truth 掩码保持原始大小。 我们不对地面实况掩码进行下采样的原因是为了使模型预测更平滑的地图，因为如果我们对二元地面实况掩码进行下采样，错误（尤其是边界周围的像素）将添加到地面实况中。 因此，通过不改变地面实况维度，我们实现了更平滑的预测，尤其是在边界附近。 所有可用架构中的特征图在宽度和高度维度上都应该是下采样的尺寸，同时我们添加了一个额外的调整大小分支，有助于将特征图恢复到其原始尺寸，这与地面实况相同 掩码的宽度和高度。

最后，最后一个解码块的输出特征图首先由步长为 2 的 2×2 转置卷积层上采样，然后由 5×5 卷积层细化，实验证明这有助于提高性能 [14]。 由于额外的 2×2 转置卷积层，获得的最终输出现在具有与原始数据集图像完全相同的大小。 由于我们希望 DPU-Net 生成二进制掩码，因此在最后一个卷积层之后使用的激活是一个 sigmoid 函数。 此外，值得一提的是，相应的编码和解码块之间的跳过连接为解码器网络添加了上下文信息，并为当前架构提供了额外的梯度流。

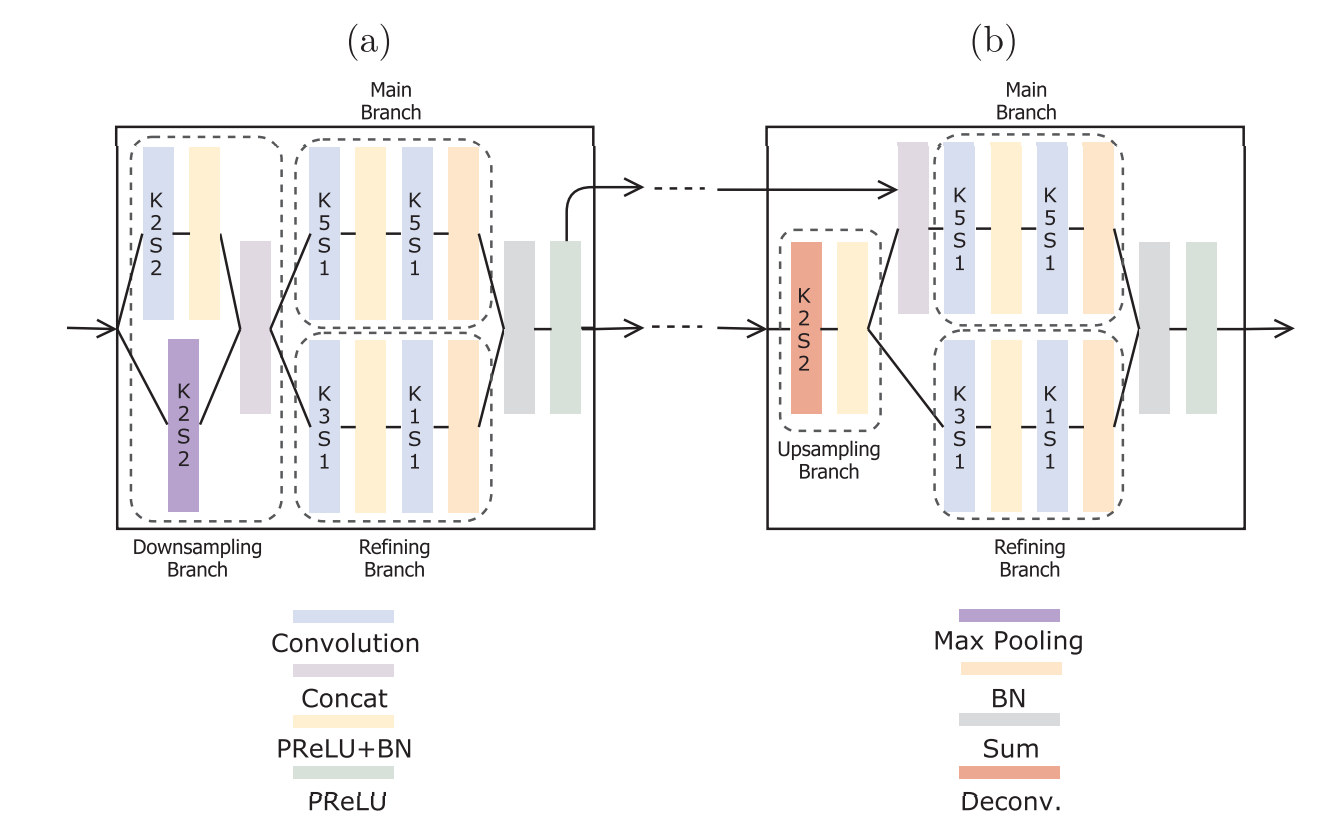


图 5. 编码块和解码块的详细说明。 注意第一个编码块没有下采样分支； 因此主分支和精炼分支将直接接受原始图像作为输入。 (a) 具有下采样分支的编码块，其后是主分支和精炼分支。 (b) 一个典型的解码块，它接受来自前一个块和跳过连接的特征图。 图中使用的缩写表示如下：“K2S2”表示“内核大小2和步幅大小2”，“BN”表示“批量归一化”。

1. Experiments

在本节中，我们首先介绍我们如何设置和训练 DPU-Net。 然后，我们详细研究了用于训练 IVUS 分割模型的不同增强技术的有效性。 此外，我们提供了一组比较实验，表明我们的架构超越了一些以前最好的工作，即 U-Net [13] 和 SegNet [21] 的性能。 最后，我们报告了 DPU-Net 分割结果的准确性，并将其与现有的 IVUS 分割文献进行了比较。

* 1. Experiments setup

训练和评估基于两个公开可用的 IVUS B 模式数据集 [1]。这两个数据集都已广泛用于 IVUS 分割文献 [7,41,29,8]。这两个数据集是用不同的超声频率获得的，即 20 MHz 和 40 MHz。这两组IVUS图像的图案和纹理相似。由于 [1] 中没有提供官方验证集，我们在训练期间从增强图像中随机选择一个小的子集作为我们的验证集进行 5 折交叉验证。作为 5 折交叉验证的结果，我们需要基于 5 个不同的训练/验证拆分来训练 5 个模型。然后我们使用 5 个模型对其相应的验证集进行推理。这样做之后，我们应该最终得到完整的折叠推理训练集。我们评估了这个 out-of-fold 训练集以选择架构的最佳训练配置，即批量大小、学习率和架构中每个块的深度。所有模型仅在两个给定数据集的训练集上进行端到端训练，不涉及任何其他外部资源，例如额外的训练图像和预训练的模型权重。

用于评估分割结果的指标已由数据集中提供的函数计算得出，该函数需要预测段的轮廓和地面实况作为输入参数。 Jaccard Measure (JM) 和 Hausdorff Distance (HD) 是使用上述函数计算的两个流行的分割性能指标。 Jaccard Measure，有时也称为联合交集 (IoU)，它是根据深度模型的自动分割 (ypred) 与专家描绘的手动分割 (ytrue) 的比较计算得出的。

自动 (Cpred) 和手动 (Ctrue) 曲线之间的 Hausdorff 距离是属于 Cpred 的所有点到 Ctrue 中最近点 [7] 的最大距离，定义如下：

* 1. Training the model

所有模型都在配备 Core i7-8700 K 处理器、16 GB RAM 和 GTX 1080 8 GB 显卡的计算机上进行训练和评估。 从头开始训练模型通常需要不到 40 分钟即可完成。 为了使训练更快并享受相对较大的批量大小，如前所述，我们将数据集的每一帧缩小了 2 倍。

我们使用 TensorFlow [42] 实现了 DPU-Net。 模型中的权重均使用 He 初始化 [39] 进行初始化，使用 TensorFlow 中的默认设置。 然后，我们用 Adam 优化训练模型 [43]。 学习率设置为 0.0001，没有衰减方案。 通过将每个模型训练 96 个 epoch 来使用实时增强器，每个 epoch 的批次大小为 6 次，总共 144 次迭代。 请注意，由于输出激活是一个 sigmoid 函数，我们需要两组独立的模型来预测流明面积和介质面积：

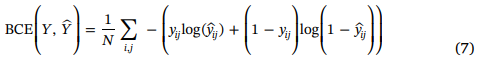
4.2.1. Loss function

We propose a heuristically designed loss function, that helps the model beneﬁt from both Binary Cross Entropy (BCE) and the Soft Dice Loss (SD). We minimize this objective function which is in fact a weighted average of the aforementioned loss functions.

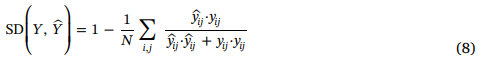


我们提出了一种启发式设计的损失函数，它有助于模型从二元交叉熵 (BCE) 和软骰子损失 (SD) 中受益。 我们最小化这个目标函数，它实际上是上述损失函数的加权平均值。

where λ = 0.8 denotes the weight of each term, the pixel-wise binary cross entropy is



and the soft dice loss which is a diﬀerentiable form of intersection over union is deﬁned as:



where N represents the total number of pixels in each image, yij is the ground truth value of ith row and jth column in the image annotation (either 1 or 0) and ŷ is the predicted probability of being a foreground pixel (either the lumen or the media).

For training each model, we monitor the average Jaccard Measure without extracting contours.

为了训练每个模型，我们在不提取轮廓的情况下监控平均 Jaccard Measure。 单个模型给出的预测是输入图像大小的概率图。 我们遵循 [15] 中相同的简单平均集成实践来产生最终结果。 二值化是通过使用搜索阈值来执行的，其中阈值是在训练集的折叠版本上搜索的。 一旦生成、集成和二值化预测图，我们使用 [44] 中解释的“填充孔”算法填充二值区域内的任何孔，使用 [45] 中描述的算法提取和跟踪边界，平滑轮廓坐标 使用“rloess”[46] 方法并将其报告为最终的分割输出。

4.3. Results

在本节中，我们报告了我们对各种增强集的彻底实验的结果。 特别是，我们展示并比较了 DPU-Net 在 20 MHz 和 40 MHz 帧的 10 个增强集上的分割输出，以及在同一组上训练的 UNet 和 SegNet 的分割结果。 此外，我们将在最佳增强集上训练的 DPU-Net 的分割输出与现有最先进的 IVUS 分割方法进行比较。

4.3.1. Augmentation results

4.3.1. 增强结果

为了牢牢掌握各种增强操作的影响，我们已经在基于提议操作的各种组合生成的增强图像上训练了我们提出的 DPU-Net。 结果在表 1 中报告并与 Balocco 数据集 [1] 的原始训练集进行了比较，该数据集仅包含 19 和 109 个图像，分别用于 40 MHz 和 20 MHz IVUS 帧（没有任何增强）。 DPU-Net 的性能可以在表 2 中看到，其中所有评估结果都是在官方测试集上获得的。

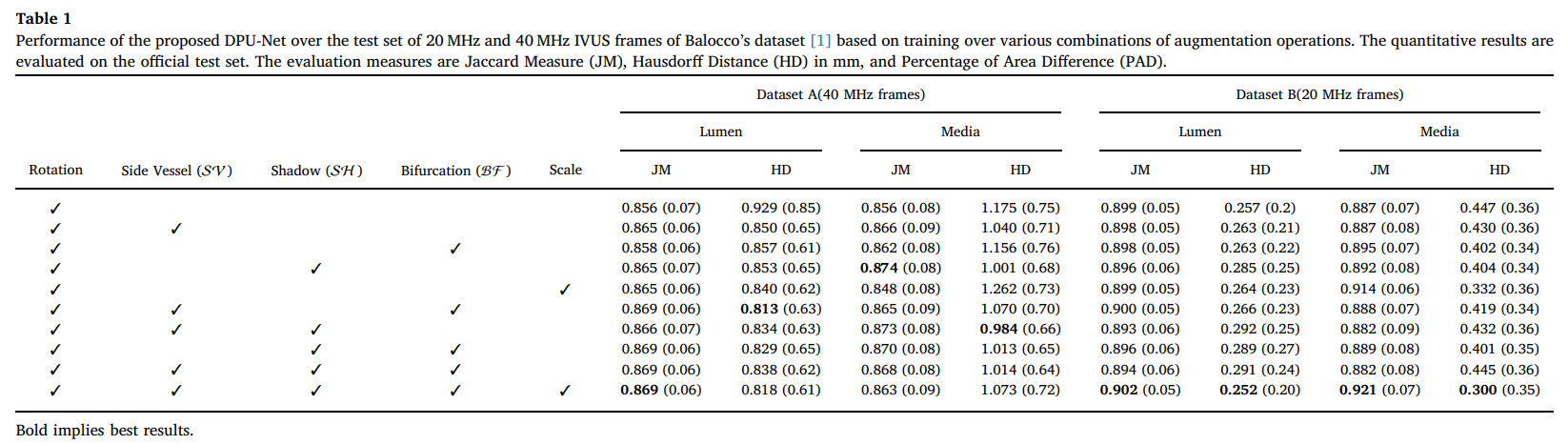


表 1 基于对各种增强操作组合的训练，建议的 DPU-Net 在 Balocco 数据集 [1] 的 20 MHz 和 40 MHz IVUS 帧的测试集上的性能。 定量结果在官方测试集上进行评估。 评估量度是 Jaccard 量度 (JM)、以毫米为单位的豪斯多夫距离 (HD) 和面积差异百分比 (PAD)。

4.3.2. DPU-Net vs SegNet and U-Net

在图像分割文献中，有两种著名的架构，即用于街道场景分割的 SegNet [21] 和用于电子显微镜堆栈中神经元结构分割的 U-Net [13]。 由于两种架构都可以输出原始输入大小的预测图（对于 U-Net，我们需要将所有卷积层的有效填充更改为相同的填充）。 为了保持比较公平，我们将最后一个上采样转置卷积层添加到两种架构的末尾。 出于说明目的，让训练图像的原始大小为 W × H，正如我们在第 3 节中提到的，我们将训练图像下采样到  的大小。 我们使用重新缩放的低分辨率图像和原始大小的掩码训练这两个网络和我们的 DPU-Net，而没有应用增强。 比较结果如表2所示。

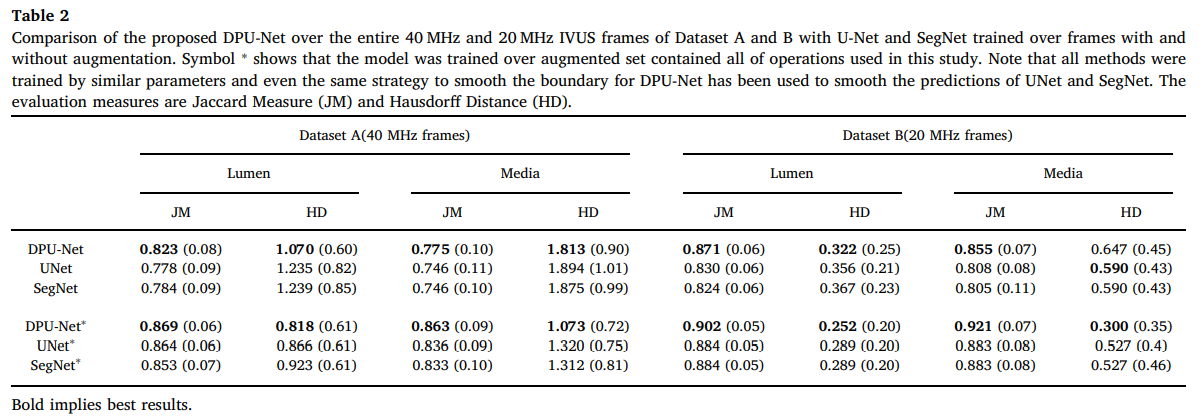


表 2 建议的 DPU-Net 在数据集 A 和 B 的整个 40 MHz 和 20 MHz IVUS 帧上与在有和没有增强的帧上训练的 U-Net 和 SegNet 的比较。 符号 \* 表明模型是在包含本研究中使用的所有操作的增强集上训练的。 请注意，所有方法都使用相似的参数进行训练，甚至用于平滑 DPU-Net 边界的相同策略也已用于平滑 UNet 和 SegNet 的预测。 评估措施是 Jaccard 措施 (JM) 和 Hausdorff 距离 (HD)。

4.3.3. Comparison with existing methods

4.3.3. 与现有方法的比较

在本节中，我们展示了 20 MHz 和 40 MHz IVUS B 模式数据集测试集的实验结果 [1]。 我们根据每个数据集的 5 倍交叉验证获得了 5 个模型，使用第 4.2 节中提到的配置，并通过简单的平均投票来集成预测图。 定量结果见表3。

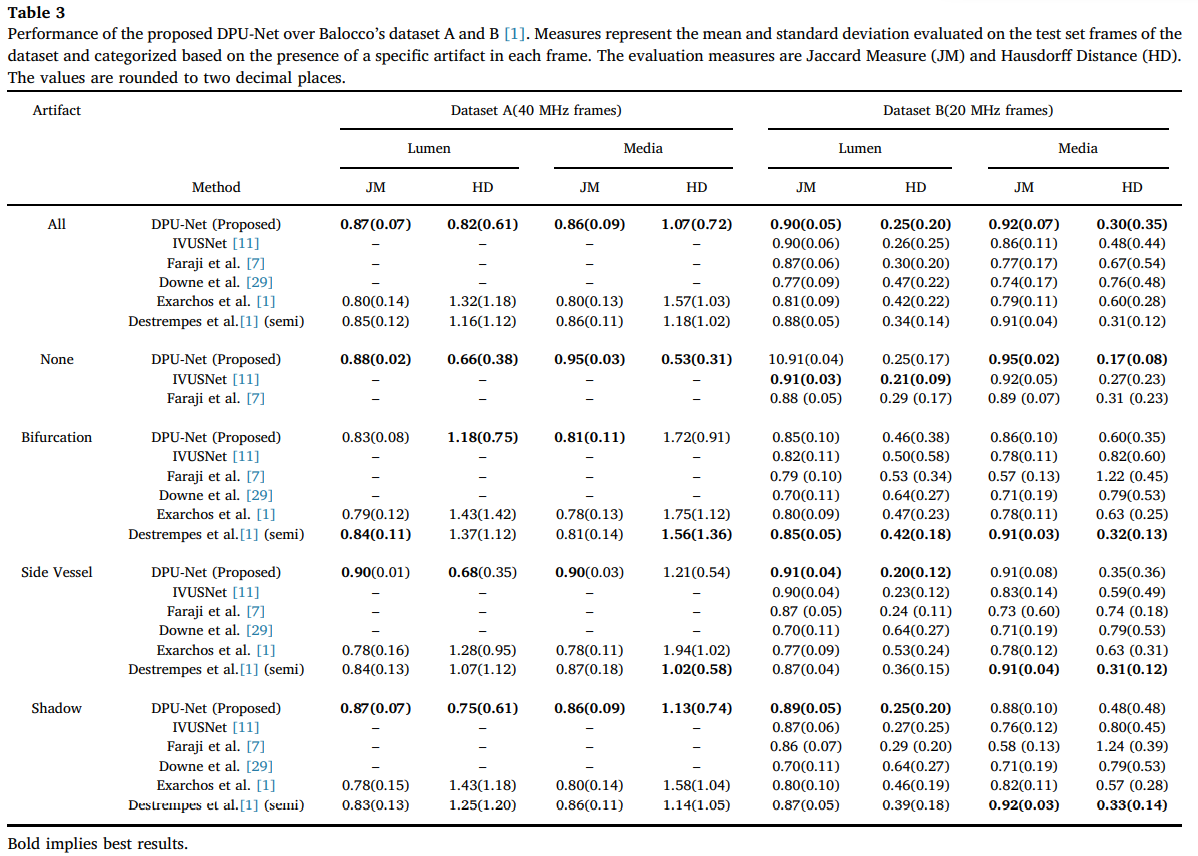


表 3 提议的 DPU-Net 在 Balocco 的数据集 A 和 B [1] 上的性能。 度量表示在数据集的测试集帧上评估的平均值和标准偏差，并根据每个帧中特定工件的存在进行分类。 评估措施是 Jaccard 措施 (JM) 和 Hausdorff 距离 (HD)。 这些值四舍五入到两位小数。

5. Discussion

5. 讨论

在这项研究中，我们进行了几次实验来评估我们提出的 DPU-Net，同时展示了通过采用新提出的增强操作所实现的改进。我们在原始 IVUS 帧和增强帧上对 DPU-Net、SegNet 和 UNet 进行了比较。从定量结果我们可以看出，DPU-Net 的性能明显优于其他两种现有架构。我们讨论了有助于实现这一结果的三个经验原因。首先，DPU-Net 具有更多的卷积层，但每个卷积层的深度较低。这类似于具有相对少量神经元的多层，而不是将许多神经元放在单个层中。多层架构具有更强的学习表示能力 [20,14]，因为它提供了更多可能的交叉特征方式，并且可以通过多个非线性激活引入更多非线性。由于神经网络中更深的层通常具有更大的感受野，因此它可以捕获多个尺度的特征。但是，始终存在固定大小的内核不能是通用解决方案的限制。在两个 IVUS 数据集中，尤其是 40 MHz 的数据集中，管腔或介质区域的形状和大小在图像之间存在显着差异。在 U-Net 和 SegNet 中，都没有特殊的设计来处理或改进多尺度分割。我们的主分支和精炼分支通过具有不同内核大小的卷积层自然地处理这个问题。 DPU-Net 对原始 SegNet 和 U-Net 的其他改进包括同时采用两种下采样方法，即池化和跨步卷积。对这两个基础架构的这种修改有助于确保我们能够以多种方式利用信息来增加特征的多样性。

展示了 DPU-Net 相对于 UNet 和 SegNet 的优越性，所有这些都在非增强的 IVUS 框架上进行了训练，我们比较了表 1 中我们提出的增强操作的各种组合，以找出操作的最佳组合。尽管几乎很明显包含所有操作的训练集将获得最佳结果，但我们对不同的操作组合进行了繁琐的训练过程，以支持和证明我们的假设，如表 1 所示。查看表 2 中报告的结果，我们看到在这些操作上训练 UNet 和 SegNet 显着改善了它们的分割结果，尽管 DPU-Net 仍然优于它们。对使用我们提出的增强操作创建的集合进行训练可以显着提高分割的准确性。更具体地说，它可以将 40 MHz 图像的流明 JM 从 0.823 增加到 0.869，将 20 MHz 帧的流明 JM 从 0.871 增加到 0.902。 40 MHz 图像的媒体分割精度也从 0.775 增加到 0.863，20 MHz 图像从 0.855 增加到 0.921。

在 20 MHz 数据集上，DPU-Net 显着改善了分割媒体区域的结果，从 JM 得分 0.79 [1] 到 0.92，HD 从 0.60 到 0.30。 DPU-Net 没有超过 [1] 中每个类别中的所有方法的原因可以从两个角度解决。 首先，训练集太小，无法捕获现实世界中的所有常见工件，甚至是测试集。 然而，该架构仍然相当有效，因为训练集仅包含 1 个带有侧血管伪影的图像，而测试集包含 93 个具有侧脉伪影的帧。 其次，阴影伪影通常与媒体区域的部分重叠，这使得分割变得更具挑战性，因为媒体区域泄漏到背景中。 一些预测如图 6 所示。

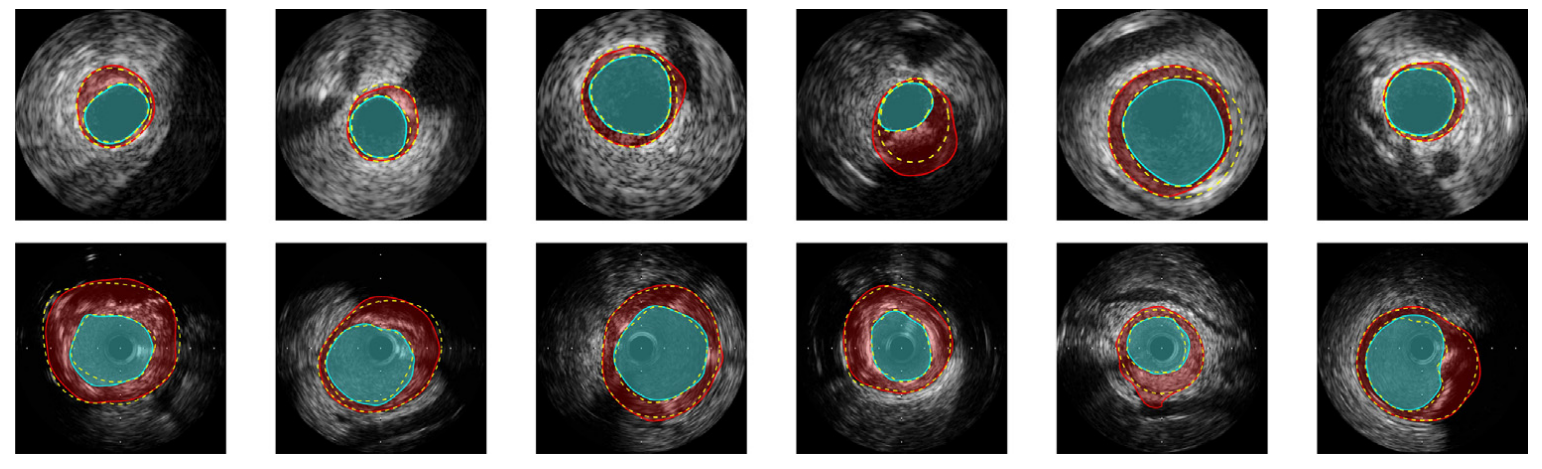


图 6. 来自数据集 B（第一行）和数据集 A（第二行）的图像的一些流明和媒体分割结果。 分段的流明和媒体分别用青色和红色突出显示。 黄色虚线表示由四位临床专家描绘的金标准 [1]。 （为了解释这个图例中对颜色的引用，读者可以参考本文的网络版本。）

6. Conclusion

在本文中，我们提出了 DPU-Net，一种完全卷积的深度网络，即使在 IVUS 图像中用于分割动脉壁的训练图像数量很少，它也能够进行泛化。 我们通过将我们提出的 DPU-Net 与两个现有的通用分割架构（即 SegNet 和 UNet）进行比较来评估我们提出的 DPU-Net 的泛化能力，这些架构在相同数量的图像上训练相同的时间而不进行任何增强。 结果表明 DPU-Net 比 SegNet 和 UNet 有显着改进。 具体来说，DPU-Net 在 40 MHz 和 20 MHz 数据集的 JM 方面的准确度分别提高了 4% 和 5% 以上。 这些实证结果表现出比 SegNet 和 UNet 更高的泛化能力。 本文的贡献可以总结如下：

• 我们为图像增强引入了一个特定领域的设计，它可以：

– 实时生成各种类型的增强图像。

– 将三种常见 IVUS 伪影的各种组合添加到训练图像中。 我们凭经验证明我们可以产生大量有效的增强图像。 这可以算作一个有效的增强管道，并且可以推广到不同的深层架构和任务。

• 我们提出的 DPU-Net 在公开可用的 IVUS 基准数据集 [1] 上优于现有方法，该数据集包含具有大量伪影的 IVUS 图像。 我们还将其与深度学习文献中现有的几个有影响力的架构进行了比较，即 SegNet [21] 和 U-Net [13]。 这表明所提出的工作也有可能用于解决其他分割问题。

为了进一步提高分割性能，我们设计了自己的增强框架，称为实时增强器。 我们的实时增强器不仅以不中断 GPU 训练过程的方式生成增强图像，而且还包含我们提出的基于 IVUS 伪影的增强操作，将三种常见的 IVUS 伪影包含在训练数据中以模拟 带有伪像的帧。 我们彻底研究了各种增强操作如何影响模型的最终精度。 因此，实验结果表明，使用这些增强数据训练的 DPU-Net 模型在很大程度上优于所有最先进的自动和半自动 IVUS 分割方法。

Appendix A. Supplementary material

与本文相关的补充数据可在在线版本中找到，网址为 <https://doi.org/10.1016/j.ultras.2019.03.014>。

References

1. S. Balocco, C. Gatta, F. Ciompi, A. Wahle, P. Radeva, S. Carlier, G. Unal, E. Sanidas, J. Mauri, X. Carillo, et al., Standardized evaluation methodology and reference database for evaluating IVUS image segmentation, Comput. Med. Imag. Graph. 38 (2) (2014) 70–90.
2. E.G. Mendizabal-Ruiz, M. Rivera, I.A. Kakadiaris, Segmentation of the luminal border in intravascular ultrasound b-mode images using a probabilistic approach, Med. Image Anal. 17 (6) (2013) 649–670.
3. G. Mendizabal-Ruiz, I.A. Kakadiaris, A physics-based intravascular ultrasound image reconstruction method for lumen segmentation, Comput. Biol. Med. 75 (2016) 19–29.
4. A. Taki, Z. Najaﬁ, A. Roodaki, S.K. Setarehdan, R.A. Zorooﬁ, A. Konig, N. Navab, Automatic segmentation of calciﬁed plaques and vessel borders in IVUS images, Int. J. Comput. Assis. Radiol. Surg. 3 (3–4) (2008) 347–354.
5. X. Zhu, P. Zhang, J. Shao, Y. Cheng, Y. Zhang, J. Bai, A snake-based method for segmentation of intravascular ultrasound images and its in vivo validation, Ultrasonics 51 (2) (2011) 181–189.
6. G. Unal, S. Bucher, S. Carlier, G. Slabaugh, T. Fang, K. Tanaka, Shape-driven seg- mentation of the arterial wall in intravascular ultrasound images, IEEE Trans. Inform. Technol. Biomed. 12 (3) (2008) 335–347.
7. M. Faraji, I. Cheng, I. Naudin, A. Basu, Segmentation of arterial walls in in- travascular ultrasound cross-sectional images using extremal region selection, Ultrasonics 84 (2018) 356–365.
8. Y. Li, M. Faraji, Erel Selection Using Morphological Relation. Available from: arXiv preprint arXiv:1806.03580.
9. M. Faraji, J. Shanbehzadeh, K. Nasrollahi, T. Moeslund, Extremal regions detection guided by maxima of gradient magnitude, IEEE Trans. Image Process.
10. M. Faraji, J. Shanbehzadeh, K. Nasrollahi, T.B. Moeslund, Erel: Extremal regions of extremum levels, 2015 IEEE International Conference on Image Processing (ICIP), IEEE, 2015, pp. 681–685.
11. J. Yang, L. Tong, M. Faraji, A. Basu, Ivus-net: An Intravascular Ultrasound Segmentation Network. Available from: arXiv preprint arXiv:1806.03583.
12. S. Kim, Y. Jang, B. Jeon, Y. Hong, H. Shim, H. Chang, Fully automatic segmentation of coronary arteries based on deep neural network in intravascular ultrasound images, Intravascular Imaging and Computer Assisted Stenting and Large-Scale Annotation of Biomedical Data and Expert Label Synthesis, Springer, 2018, pp. 161–168.
13. O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2015, pp. 234–241.
14. A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classiﬁcation with deep con- volutional neural networks, Advances in Neural Information Processing Systems, 2012, pp. 1097–1105.
15. D. Ciresan, A. Giusti, L.M. Gambardella, J. Schmidhuber, Deep neural networks segment neuronal membranes in electron microscopy images, Advances in Neural Information Processing Systems, 2012, pp. 2843–2851.
16. J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic seg- mentation, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 3431–3440.
17. L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A.L. Yuille, Deeplab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFS, 2016. Available from: arXiv preprint arXiv:1606.00915.
18. P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, C. Langlotz, K. Shpanskaya, et al., Chexnet: Radiologist-Level Pneumonia Detection on Chest X-rays with Deep Learning. Available from: arXiv preprint arXiv:1711.05225.
19. C. Peng, X. Zhang, G. Yu, G. Luo, J. Sun, Large Kernel Matters–Improve Semantic Segmentation by Global Convolutional Network. Available from: arXiv preprint arXiv:1703.02719.
20. K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.
21. V. Badrinarayanan, A. Kendall, R. Cipolla, Segnet: a deep convolutional encoder- decoder architecture for image segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 39 (12) (2017) 2481–2495.
22. M.A. Tanner, W.H. Wong, The calculation of posterior distributions by data aug- mentation, J. Am. Stat. Assoc. 82 (398) (1987) 528–540.
23. P.Y. Simard, D. Steinkraus, J.C. Platt, Best practices for convolutional neural net- works applied to visual document analysis, Null, IEEE, 2003, p. 958.
24. D. Cireşan, U. Meier, J. Schmidhuber, Multi-column deep neural networks for image classiﬁcation. Available from: arXiv preprint arXiv:1202.2745.
25. D.C. Cireşan, U. Meier, J. Masci, L.M. Gambardella, J. Schmidhuber, High-perfor- mance neural networks for visual object classiﬁcation. Available from: arXiv pre- print arXiv:1102.0183.
26. K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition. Available from: arXiv preprint arXiv:1409.1556.
27. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 1–9.
28. M. Paulin, J. Revaud, Z. Harchaoui, F. Perronnin, C. Schmid, Transformation pur- suit for image classiﬁcation, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014, pp. 3646–3653.
29. R. Downe, A. Wahle, T. Kovarnik, H. Skalicka, J. Lopez, J. Horak, M. Sonka, Segmentation of intravascular ultrasound images using graph search and a novel cost function, in: Proc. 2nd MICCAI Workshop on Computer Vision for Intravascular and Intracardiac Imaging, Citeseer, 2008, pp. 71–79.
30. M.-H. Cardinal, J. Meunier, G. Soulez, R.L. Maurice, É. Therasse, G. Cloutier, Intravascular ultrasound image segmentation: a three-dimensional fast-marching method based on gray level distributions, IEEE Trans. Med. Imag. 25 (5) (2006) 590–601.
31. J. Dijkstra, G. Koning, J. Tuinenburg, P. Oemrawsingh, J. Reiber, Automatic border detection in intravascular iltrasound images for quantitative measurements of the vessel, lumen and stent parameters, Computers in Cardiology 2001, IEEE, 2001, pp. 25–28.
32. M.E. Plissiti, D.I. Fotiadis, L.K. Michalis, G.E. Bozios, An automated method for lumen and media-adventitia border detection in a sequence of IVUS frames, IEEE Trans. Inform. Technol. Biomed. 8 (2) (2004) 131–141.
33. J.D. Klingensmith, R. Shekhar, D.G. Vince, Evaluation of three-dimensional seg- mentation algorithms for the identiﬁcation of luminal and medial-adventitial bor- ders in intravascular ultrasound images, IEEE Trans. Med. Imag. 19 (10) (2000) 996–1011.
34. J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei, Imagenet: a large-scale hierarchical image database, IEEE Conference on Computer Vision and Pattern Recognition, 2009. CVPR 2009, IEEE, 2009, pp. 248–255.
35. T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, C.L. Zitnick, Microsoft coco: common objects in context, European Conference on Computer Vision, Springer, 2014, pp. 740–755.
36. S. Xie, R. Girshick, P. Dollár, Z. Tu, K. He, Aggregated residual transformations for deep neural networks, 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2017, pp. 5987–5995.
37. M. Drozdzal, E. Vorontsov, G. Chartrand, S. Kadoury, C. Pal, The importance of skip connections in biomedical image segmentation, Deep Learning and Data Labeling for Medical Applications, Springer, 2016, pp. 179–187.
38. B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, A. Torralba, Learning deep features for discriminative localization, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 2921–2929.
39. K. He, X. Zhang, S. Ren, J. Sun, Delving deep into rectiﬁers: surpassing human-level performance on imagenet classiﬁcation, Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 1026–1034.
40. B. Xu, N. Wang, T. Chen, M. Li, Empirical evaluation of rectiﬁed activations in convolutional network. Available from: arXiv preprint arXiv:1505.00853.
41. L. Lo Vercio, J.I. Orlando, M. del Fresno, I. Larrabide, Assessment of image features for vessel wall segmentation in intravascular ultrasound images, Int. J. Comput. Assis. Radiol. Surg. 11 (8) (2016) 1397–1407.
42. M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G.S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, X. Zheng, TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems, Software Available from tensorﬂow.org, 2015. < https://www.tensorﬂow.org/ >.
43. D.P. Kingma, J. Ba, Adam: A Method for Stochastic Optimization. Available from: arXiv preprint arXiv:1412.6980.
44. P. Soille, Morphological Image Analysis: Principles and Applications, Springer Science & Business Media, 2013.
45. R.C. Gonzalez, R.E. Woods, S. Eddins, Digital Image Processing using Matlab: Pearson Prentice Hall, Upper Saddle River, New Jersey.
46. S.J. Orfanidis, Introduction to Signal Processing, Prentice-Hall Inc., 1995.