

# SAR-UNet: 一种基于空间注意力机制的大气河识别网络模型

罗月梅, 胡金蓉, 李桂钢, 帅梓涵, 郎子鑫

成都信息工程大学计算机学院, 四川 成都

收稿日期: 2023年3月18日; 录用日期: 2023年4月17日; 发布日期: 2023年4月24日

## 摘要

准确、高效地识别大气河可以有效地预防洪涝和干旱等自然灾害, 对地区的经济、社会和生态发展具有非常重大的意义。现有的大气河识别算法主要是基于多个物理量的阈值来实现的, 它们能有效地识别大气河, 但需要手动设置阈值和几何度量标准。由于现有方法在复杂度和泛化能力上的局限性, 本文基于深度学习的方法, 提出了名为SAR-UNet的网络模型, 该模型基于UNet网络结构, 在特征提取时创新性地加入了空间注意力机制, 从而能够更好地获取大气河的局部特征信息。通过深度学习的方法对大气河进行自动识别, 有效地解决了现有方法需要确定阈值和度量标准的不足。实验结果表明, 基于SAR-UNet的大气河识别方法在ERA-Interim气候再分析数据集上具有最佳的识别效果, 大气河识别的精度上达到了97.27%, MIoU值高达85.22%。

## 关键词

大气河识别, 语义分割, 深度学习, 空间注意力机制

# SAR-UNet: A Model of Atmospheric River Recognition Network Based on Spatial Attention Mechanism

Yuemei Luo, Jinrong Hu, Guigang Li, Zihan Shuai, Zixin Lang

School of Computer Science, Chengdu University of Information Technology, Chengdu Sichuan

Received: Mar. 18<sup>th</sup>, 2023; accepted: Apr. 17<sup>th</sup>, 2023; published: Apr. 24<sup>th</sup>, 2023

## Abstract

Accurate and efficient identification of atmospheric rivers can effectively prevent natural disas-

文章引用: 罗月梅, 胡金蓉, 李桂钢, 帅梓涵, 郎子鑫. SAR-UNet: 一种基于空间注意力机制的大气河识别网络模型[J]. 计算机科学与应用, 2023, 13(4): 819-832. DOI: 10.12677/csa.2023.134081

ters such as floods and droughts, and is of great significance to the economic, social and ecological development of the region. Existing atmospheric river identification algorithms are mainly based on the thresholds of multiple physical quantities, which can effectively identify atmospheric rivers, but need to manually set thresholds and geometric metrics. Due to the limitations of existing methods in complexity and generalization ability, this paper proposes a network model named SAR-UNet based on deep learning, which is based on UNet network structure and innovatively adds spatial attention mechanism during feature extraction, so as to better obtain local feature information of atmospheric rivers. The automatic identification of atmospheric rivers by deep learning effectively solves the shortcomings of existing methods that need to determine thresholds and metrics. The experimental results show that the atmospheric river identification method based on SAR-UNet has the best identification effect on the ERA-Interim climate reanalysis dataset, the accuracy of atmospheric river identification reaches 97.27%, and the MIOU value is as high as 85.22%.

## Keywords

Atmospheric River Recognition, Semantic Segmentation, Deep Learning, Spatial Attention Mechanism

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

大气河(Atmospheric River, AR)是位于对流层下层的一条狭长的水汽传输路径[1], 通常长度会有几千公里, 而宽度却只有几百公里, 形似一条长河而得名。它经常出现在低纬度至中纬度地区的海洋上方[2], 同时也会伴随着海洋性气旋、低空急流和锋面活动。一条大气河所携带的水量甚至可以超过地球上最大的河流——亚马逊河[3], 因此其在全球水汽分布的影响中起着非常关键的作用, 故而经常作为地区降水的一个重要考量因素。由于大气河经常会引起极端降水, 而造成洪涝灾害[4], 因此研究大气河的识别具有重要的意义。描述大气河的物理量通常有两种: 1) 水汽输送(Integrated Water Vapor Transport, IVT), 它能够显示大气河中水汽的运输情况; 2) 水汽含量(Integrated Water Vapor, IWV), 它能够显示大气河中的水汽总量。由于 IVT 拥有更好的方向性, 能在全球范围内应用, 同时也是产生强降雨和洪水的关键因素[5], 因而 IVT 是更加常用的物理量, 包括本文的数据集也是基于 IVT 的。

随着人们对大气河的了解, 越来越多的人开始关注并研究大气河的识别方法[6]。现有的大气河识别方法主要是基于 IVT 和 IWV 两个物理量来判定[7], 此外, 还要根据大气河的形状, 如长、宽, 方向等来辅助判定。现有方法可以大致分成三个大类: 1) 基于单个观测点或者模型网络单元的低层风和 IWV 值, 无需考虑大气河的几何形状特征。这种方法适合研究小的、局部地区的大气河登录[8]。2) 先确定研究区域的预选面积, 跟踪流经这个截面水汽的空间特征, 再根据大气河的长度、方向来实现对大气河的识别, 由于需要预先选择特定的区域, 因此这种方法适用于指定区域内大气河登陆的研究, 而不适用于全球范围内的大气河识别[9]。3) 充分地考虑大气河的地理形态, 包括给定区域内任何位置大气河的长宽特征, 这类方法适用于大气河在大区域空间内的分布研究[10]。这些大气河识别算法都可以有效地识别出符合条件的大气河, 但还是存在着一些不足: 1) 泛化能力较差。不同的算法对大气河确定的阈值条件和几何标准也不尽相同, 这会导致不同的算法对大气河的识别结果也不完全相同, 使得算法的泛化

能力不强。2) 大气河识别过程较为复杂。它们通常需要计算 IVT/IWV 的具体数值, 然后将 IVT/IWV 阈值化, 还要计算 IVT 的方向, 甚至大气河持续的时间, 满足这些条件之后, 还要计算它们的几何特征是否符合要求等。

鉴于现存算法的局限性, 我们想要简化大气河识别的过程, 同时增强算法的泛化能力, 因此将大气河识别看成一个语义分割的问题, 创新性地提出了结合大气河特征的 SAR-UNet 网络模型, 使用深度学习的方法来识别大气河。本文的主要贡献如下: 1) 根据大气河的特性提出了一种新颖的基于深度学习的大气河识别方法; 2) 简化了大气河识别的过程并增强了大气河识别算法的泛化能力。

本文的其他部分组织如下: 第二节概括了相关工作; 第三节描述了数据集的相关情况; 第四节描述了使用的模型结构; 第五节是实验部分; 第六节对文章内容进行总结。

## 2. 相关工作

现存的大气河识别算法多种多样, 它们有的是针对全球区域内的大气河, 有的是针对特定的区域, 但大部分都是通过计算一些物理量是否满足条件来判定的, 其中对 IVT/IWV 值有相对量和绝对量之分。如 Ralph 等人[11]提出的算法要求是每小时的 IWV 值大于等于 2 cm, 并且至少持续 12 小时, 每个 IWV 的通量大于等于 20 cm m/s (相当于 IVT 值为 250 kg/m/s), 这是针对局部地区的方法。Wick 等人[12]也提出了一种基于 IWV 的算法, 不同的是, 还需要考虑大气河的形状。该算法需要满足的三个基本条件是: 1) IWV 含量 > 20 mm; 2) 长度 > 2000 km; 3) 宽 < 3000 km。Rutz 等人[13]在 2014 年提出了基于 IVT 的算法, 他们对大气河的标准定义为长度  $\geq 2000$  km, IVT  $\geq 250$  kg/m/s, 不考虑大气河的曲率, 同时也不依赖于持续时间。在 2015 年, Guan 等人[14]提出了一种针对全球区域的基于 IVT 的大气河识别方法, 这个算法的标准是 IVT 的强度要高于每个季节和网络单元计算机的 85 百分位 IVT, 约为 166~254 kg/m/s, 最小的限制为 100 kg/m/s, 同时平均 IVT 方向要在  $45^\circ$  以内, 除此之外, 大气河的长度要大于 2000 km, 长宽比大于 2。

相比于地球上其他地区(如北美和欧洲), 东亚地区的大气河研究还处于初步阶段, 甚至还难以找到一个可适应东亚地区独特而又复杂的气候条件的大气河识别追踪算法。在 2020 年, Pan 等人[15]研发了适用于东亚地区的大气河识别追踪算法, 这个算法包括 6 个步骤: 1) 基于高斯平滑算法和两步阈值法的潜在大气河区域提取; 2) 平滑大气河轨迹的生成; 3) 大气河多维度特征的归纳与计算; 4) 非大气河引起的强水汽传输现象(如热带气旋)的识别与过滤; 5) 不同天气系统所控制的大气河片段的分割和识别; 6) 大气河全生命周期(life cycle)的识别与追踪。使用不同的阈值标准得到的大气河识别结果不同, 从而给理解和量化大气河带来不确定性。对此, Ohanuba 等人[16]提出了新的方法, 他们将拓扑数据和机器学习结合起来, 利用数据拓扑分析为大气河提供形状特征, 并且作为大气河分类任务的特征描述符, 联合查找算法跟踪 IWV 场中拓扑结构的变化, 再把生成的特征描述符利用支持向量机分类器进行训练, 最后使用极端气候事件分析的并行工具包(TECA)识别的结果作为真实值评估分类器的性能。这种方法为研究大气河的识别带来了新的思路, 分类的精度最高也可以达到 91%, 但是分类器的性能容易受到训练数据质量的影响, 我们需要高质量的真实数据[17]。

## 3. 数据集

本文的数据集来自欧洲中期天气预报中心(ECMWF)的 ERA-Interim 气候再分析数据的第一版, 网格分辨率是  $1.5 \times 1.5$ 。它涵盖了 1979 年 1 月 1 日至 2019 年 8 月 31 日的数据, 每日有四次数据, 间隔时间为 6 小时, 分别是 00UTC、06UTC、12UTC 和 18UTC。我们根据原始 ERA-Interim 数据库中的整层积分水汽输送量值(IVT)生成了 IVT 的分布图, 一共生成了 59,412 张图片, 图片的大小为  $1024 \times 1024$ , 如图

1 所示。我们通过 Guan、Wailser [18]提出的大气河识别算法对图像进行标记，并邀请了研究大气河识别的专家对标记的结果进行核验。有了准确的标签，也可以使得我们训练出更好的模型。用这种方式，我们得到了 59,412 张尺寸为 1024\*1024 的大气河标记图，如图 2 所示。

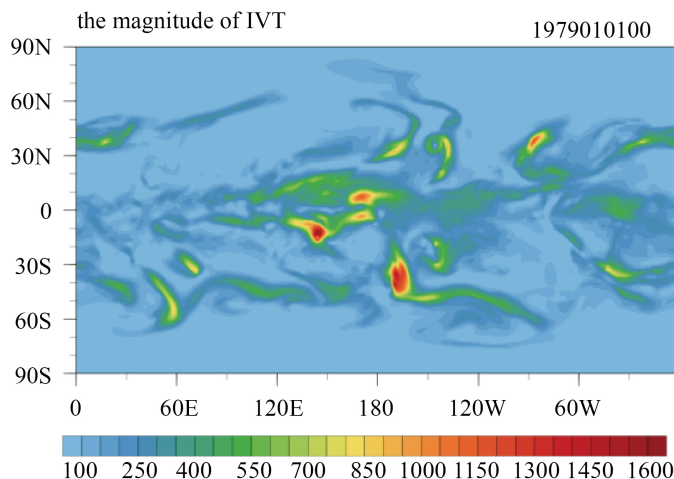


Figure 1. Original image

图 1. 原始图像

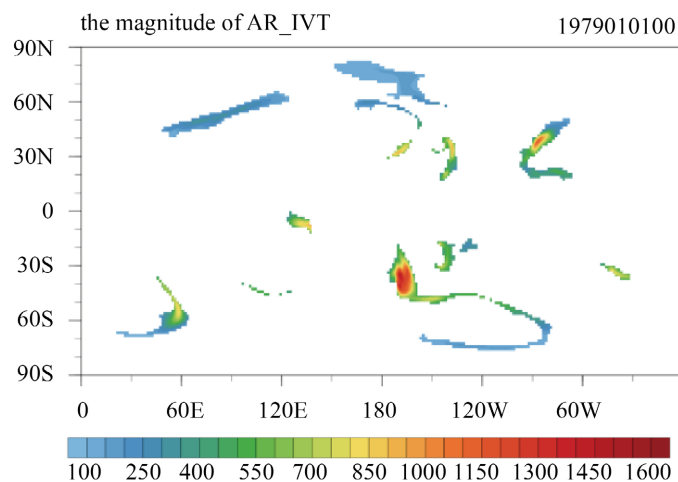


Figure 2. Marked atmospheric river image

图 2. 标记的大气河图像

通过裁剪和 `resize` 函数将原始图片和标记图片的大小均缩放为  $256 \times 256$ ，通过缩放后的原始图片如图 3 所示，标记图片如图 4 所示。为了全面准确地评价模型的性能，同样将数据集划分了训练集、验证集和测试集三个部分。把 2018 年到 2019 年的 2432 张图片作为测试集，剩余的从 1979 年到 2017 年的 56,980 张图片按 17:3 随机划分为训练集和验证集。具体划分结果如表 1 所示。

#### 4. 模型介绍

我们把大气河的识别看成是图像分割的问题，因此我们可以使用深度学习的方法来解决[19]。UNet [20]

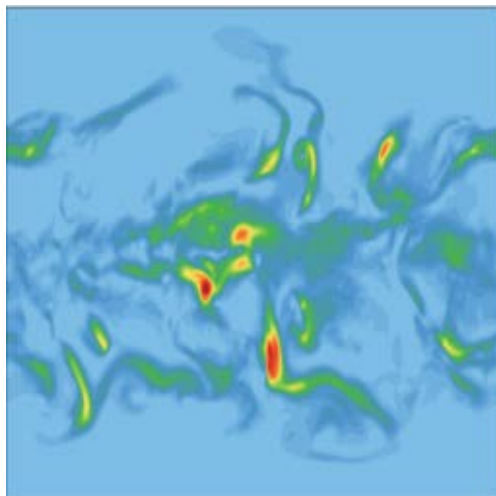


Figure 3. Input image

图 3. 输入图像

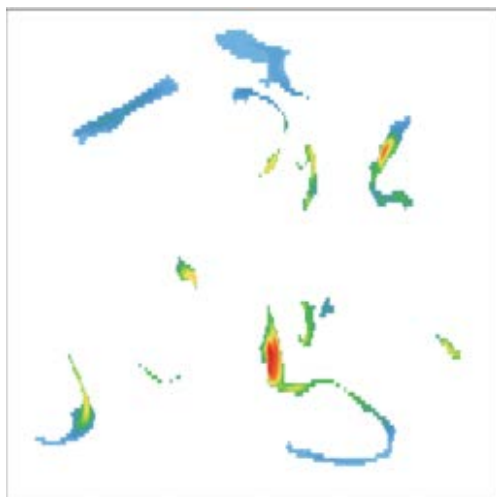


Figure 4. Label image

图 4. 标签图像

Table 1. Dataset partitioning

表 1. 数据集划分

Train set	Valid set	Test set
48433	8547	2432

是图像语义分割的一个经典模型，在各种应用场景都有着优秀的性能表现，比如生物医学界、工业界以及航空航天等领域。它是编码器解码器的完全对称结构，通过拼接的方式融合深层和浅层的信息，使用跳跃连接的架构，在每个阶段都允许解码器学习在编码器中丢失的相关特征，最后在上采样时采用转置卷积。UNet 模型结构简单且分割效果良好，经常作为语义分割任务的 **baseline** 与其他模型对比[21]。

本文提出的 SAR-UNet 模型把 UNet 模型作为基本的框架，UNet 是一个 U 型对称结构，左边的部分可以看成是一个编码器，右边的部分对应解码器。UNet 整体结构可如图 5 所示。

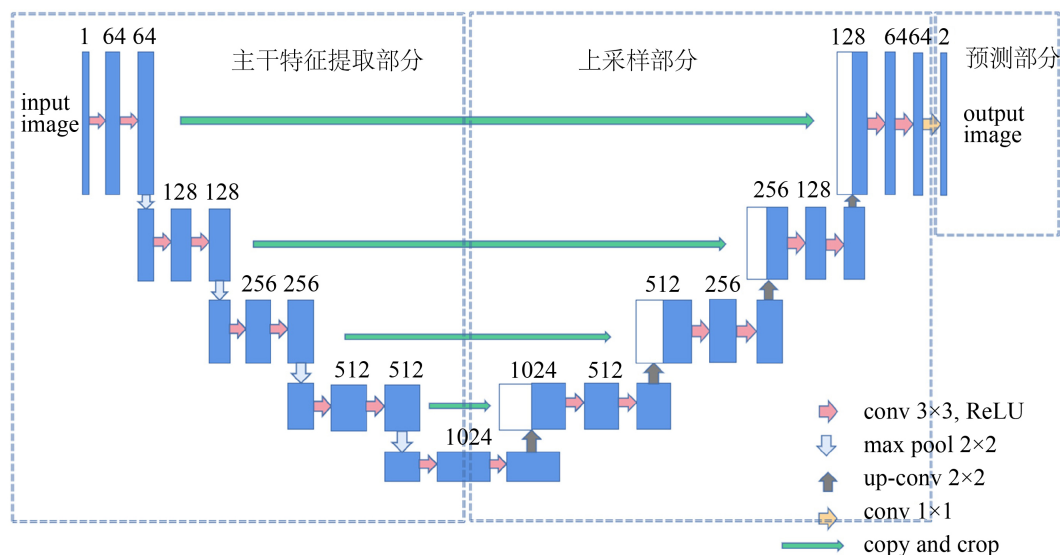


Figure 5. Network structure of UNet  
图 5. UNet 网络结构

为了使得模型更好地提取大气河的特征,我们把 UNet 的编码器部分换成了更适用于提取大气河特征的网络结构。我们利用 VGG16 网络[22]来提取大气河的位置信息,并通过空间注意力机制[23]更好地提取到大气河的局部特征;防止网络模型过度的关注局部区域特征而忽略整体特征,使用残差网络结构[24]来提取深层的整体特征,然后再把局部特征和整体特征进行融合,得到 5 个不同尺寸的特征层,最后将这 5 个特征层进行上采样和跳跃连接。基于 SAR-UNet 的大气河识别方法的整体概况可如图 6 所示。

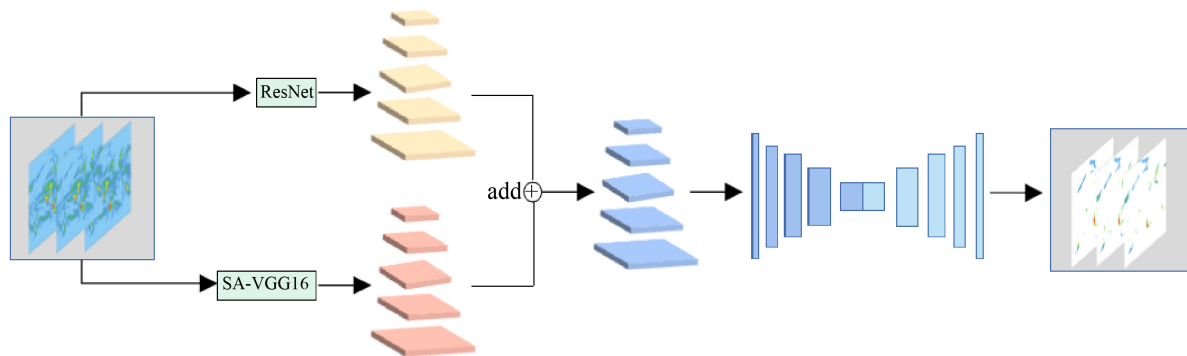


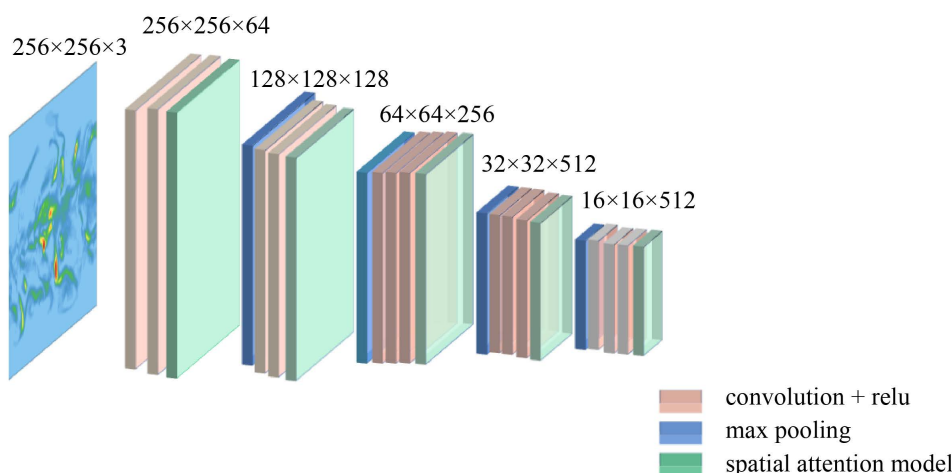
Figure 6. Overview of the model  
图 6. 模型概况

本文模型可以大致分为四部分:第一部分是提取局部特征,第二部分是提取整体特征,第三部分是特征融合,第四部分是上采样和跳跃连接。接下来我们分别对每个部分进行详细的介绍:

1) 提取局部特征

使用 VGG16 作为提取大气河区域位置的网路,这个网路主要使用卷积操作和最大池化,同时引入空间注意力机制对每个位置分配注意力权重,使得网路更加的关注目标区域[25],从而更好地提取大气河的局部信息,把这个结构称为 SA-VGG16,具体如图 7 所示。





**Figure 7.** Network structure of SA-VGG16  
**图 7.** SA-VGG16 网络结构

给定一张待识别的大气河图片  $I \in \mathbb{R}^{H \times W \times C_{in}}$  (其中  $H$ 、 $W$  和  $C_{in}$  分别为图片的高度、宽度以及通道数), 使用两次卷积  $H_{SF}(\cdot)$  来提取特征  $F_0 \in \mathbb{R}^{H \times W \times 64}$  :

$$F_0 = H_{SF}(I)$$

其中, 第一次特征提取的通道数为 64。该卷积能够简单快速地将图片映射到一个更高的维度, 从而方便添加空间注意力机制模块(Spatial Attention Module), 经过一次注意力模块  $H_{SA}(\cdot)$  过后, 便得到了浅层特征  $F_{SF} \in \mathbb{R}^{H \times W \times 64}$  :

$$F_{SF} = H_{SA}(F_0)$$

接着从浅层特征  $F_{SF}$  中提取更高维的深层特征  $F_{DF} \in \mathbb{R}^{H_r \times W_r \times 512}$  :

$$F_{DF} = H_{DF}(F_{SF})$$

其中,  $H_{DF}(\cdot)$  包含四次升维过程, 该过程由最大池化、卷积和空间注意力模块组成(其中第一次卷积层数为 2, 剩余三次卷积层数为 3)。每次升维其特征通道数就会翻倍, 最终通道数会变成 512, 而特征图大小会缩小为原来的一半, 最终特征图大小会变成初始图片大小的 1/32。

本文使用的空间注意力机制模块(Spatial Attention)的实现过程如图 8 所示。首先对输入进来的特征层, 在每一个特征点的通道上取的最大值和平均值, 分别得到 Feature 1 和 Feature 2, 再把这两个特征层进行堆叠。堆叠之后的结果利用一次通道数为 1 的卷积调整通道数, 然后取一个 sigmoid 获得输入特征层每一个特征点的权值(在 0 到 1 之间), 这些权值代表了网络模型对这块区域的关注程度, 权值越大, 关注度越高, 因此在加入了空间注意力机制后, 经常出现大气河的中高纬度地区会更多地被网络模型关注, 从而提高识别大气河的精度。

## 2) 提取整体特征

使用残差网络提取大气河的深层整体特征, 经过 ResNet 同样获得五个不同大小的特征层。残差网络是由一系列残差块组成的, 残差块分成直接映射部分和残差部分, 如图 9 所示。

残差结构的处理分成两个部分, 左边的  $F(X)$ 以及右边的  $X$ , 最后的结果为两者相加。右边部分不会对  $X$  做任何处理, 没有可学习的参数; 左边部分  $F(X)$ 是学习特征的部分, 这个结构学习的其实是输出结果与输入的差值。完整的残差网络由多个残差快组成, 使用残差网络除了可以提取到图像的深层整体特

征，还可以在一定的程度上解决深层神经网络中由于网络层数太深而产生梯度消失的问题。

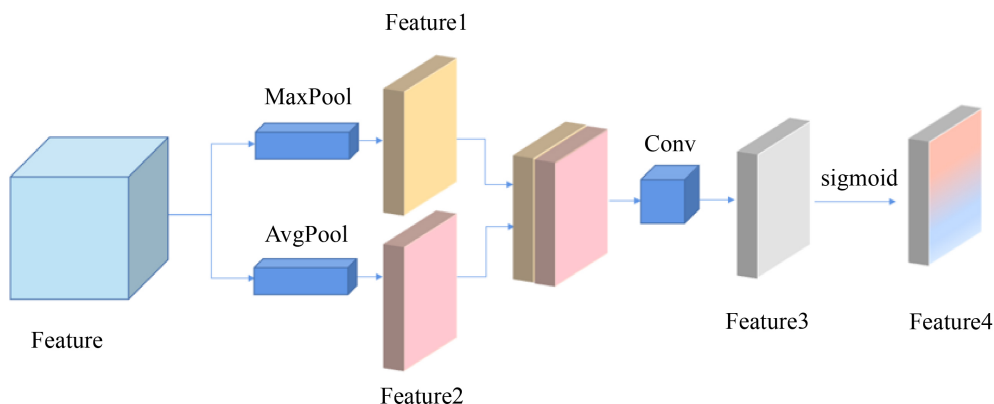


Figure 8. Spatial Attention module  
图 8. Spatial Attention 模块

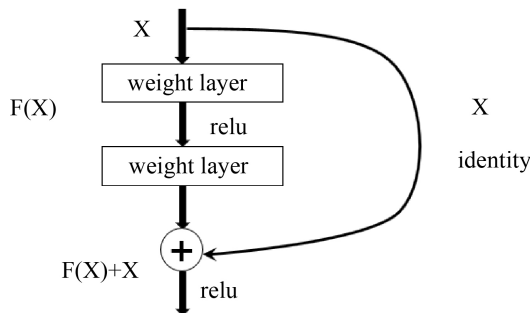


Figure 9. Residual block  
图 9. 残差块

### 3) 特征融合

由于深层特征分辨率较低，图片比较模糊，可以得到目标区域大致的位置，往往容易忽略一些细节信息；而低层特征的分辨率较高，能体现更多的细节，具有更多的语义信息。因此，我们把整体特征和局部特征进行了一个融合，这样可以更好地利用不同特征的特性，融合特征之间的优点，从而达到提升模型性能的目的。通常特征融合的方式有 `concat` 和 `add` 这两种方式，`concat` 的方式是系列特征融合，直接将两个特征进行连接。两个输入特征  $x$  和  $y$  的维数若为  $p$  和  $q$ ，输出特征  $z$  的维数为  $p + q$ ，也可以看作是通道数的增加。而 `add` 使用的是并行策略，将这两个特征向量组合成复向量，对于输入特征  $x$  和  $y$ ， $z = x + iy$ ，其中  $i$  是虚数单位，即特征图相加，特征的信息量增多了，但是通道数不变。在这里使用的特征融合方式是 `add`，这样可以更好地融合大气河的特征。

### 4) 上采样和跳跃连接

在这部分我们利用主干特征提取网络特征融合后得到的五个特征层进行上采样，并且 `concat` 的方式进行特征融合，最后得到一个最终的特征层。为了让网络更好地进行构建以及使它有更好的通用性，我们在上采样时直接进行两倍的上采样，再进行特征融合，使最终获得的特征层和输入图像的高宽相同。

### 5) 预测部分

我们利用最终获得的有效特征层再经过一个  $1 \times 1$  的卷积对每一个特征点进行分类，也就是对图像进



行像素级的分类，使用的激活函数是 softmax。

## 5. 实验部分

### 5.1. 评价指标

在本文中，我们使用像素准确率(PA)、平均像素准确率(MPA)、平均交并比(MIOU)和频权交并比(FWIOU)作为我们的评价指标。假设有  $k + 1$  类(包括  $k$  个目标类和 1 个背景类)， $P_{ij}$  表示本属于  $i$  类却预测为  $j$  类的像素点总数，称为假阳性(FP)。同样地， $P_{ii}$  表示真阳性(TP)， $P_{ji}$  表示假阴性(FN)， $P_{jj}$  表示真阴性(TN)。

像素准确率(PA)表示分类正确的像素点数和所有像素点数的比例，可以表示为真阳性(TP)和真阴性(TN)之和除以真阳性(TP)、真阴性(TN)、假阳性(FP)和假阴性(FN)之和，即  $PA = (TP + TN)/(TP + TN + FP + FN)$ ，完整的公式如下：

$$PA = \frac{\sum_{i=0}^k p_{ii}}{\sum_{i=0}^k \sum_{j=0}^k p_{ij}}$$

平均像素准确率(MPA)是计算每一类分类正确的像素点数和该类的所有像素点数的比例，然后再求平均，具体如下：

$$MPA = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij}}$$

平均交并比(MIOU)是语义分割中常用的评价指标，交并比(IOU)是真实值和预测值两个集合的交集和并集之比，它可以表示成真阳性(TP)和真阳性(TP)、假阳性(FP)和假阴性(FN)之和的比值，即  $IOU = TP/(TP + FP + FN)$ 。平均交并比(MIOU)是对每一类求 IOU 后再求平均，具体计算公式如下所示：

$$MIOU = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij} + \sum_{j=0}^k p_{ji} - p_{ii}}$$

频权交并比(FWIOU)是根据每一类出现的频率设置权重，权重乘以每一类的 IOU 并进行求和，计算公式如下所示：

$$FWIOU = \frac{1}{\sum_{i=0}^k \sum_{j=0}^k p_{ij}} \sum_{i=0}^k \frac{p_{ii} \sum_{j=0}^k p_{ij}}{\sum_{j=0}^k p_{ij} + \sum_{j=0}^k p_{ji} - p_{ii}}$$

### 5.2. 实验参数设置

为了更好地评价模型检测的结果，对比了 PSPnet [26]、SegNet [27]、FCN [28]和 DeepLabv3+ [29]三个模型。这些模型都在语义分割中有着重要的地位，也在不同的场景中展示出了较高的性能。PSPNet 通过全局均值池化操作和特征融合实现引入更多的上下文信息，对不同区域的语境进行聚合，让模型拥有理解全局语境信息的能力。本文使用 resnet50 为 PSPNet 的主干网络进行实验。SegNet 的主要结构和自编码器类似，通过编码解码复原图片上的每一个点所属的类别，也是对图片进行像素级的分割。SegNet 在编码部分的最大池化操作时记录了最大值所在位置(索引)，节省了内存空间，然后在解码时通过对应的索引实现非线性上采样，这样在上采样的时候就不需要学习。FCN 是语义分割的开山之作，是非常经典的一个分割模型。它将 CNN 网络最后面的全连接层换成了卷积层，因此该网络的输入可以是任意尺寸。DeepLabv3+为了融合多尺度信息，DeepLabv3+引入了语义分割常用的编码-解码结构，在这个结构中引入可任意控制编

码器提取特征的分辨率，通过空洞卷积平衡精度和耗时。和之前的语义分割模型相比，其最大的特点就是引入了空洞卷积[30]，在不损失信息的情况下，加大了感受野，让每个卷积输出都包含较大范围的信息。

我们为这些模型设置了相同的参数，并且都使用了预训练的权重，具体如下：batch size 设置为 8、优化器使用的是 Adam，损失函数则由 Cross Entropy Loss 和 Dice Loss 组成，其中 Cross Entropy Loss 在 Softmax 对像素点进行分类时使用。Dice Loss 定义为  $1 - \text{Dice}$ ，其中 Dice 系数是一种集合相似度的函数，用于计算两个集合之间的相似度，取值范围是 0 到 1 之间。假设有 X 和 Y 两个集合，Dice 系数和 Dice Loss 的计算公式如下：

$$\text{Dice} = \frac{2|X \cap Y|}{|X| + |Y|} \quad \text{Dice Loss} = 1 - \text{Dice}$$

迁移学习就是把已经训练好的模型(预训练模型)参数迁移到新的模型来帮助新模型的训练[31]。因为神经网络主干特征提取部分所提取到的特征是通用的，所以我们可以将在良好数据集上已经学到的模型参数通过某种方法分享给新模型，从而加快模型训练的速度以及优化模型[32]。实现迁移学习有三种方式，分别是 Transfer Learning、Extract Feature Vector 和 Fine-tuning [33]，在本文的实验中我们都通过 Fine-tuning 来实现迁移学习，并且我们的预训练权重都是在 Pascal VOC 数据集上训练得到的。

当冻结靠近输入这部分的卷积层时，学习率设置为  $1e-4$ ，epoch 为 30，解冻之后，学习率设置为  $1e-5$ ，epoch 为 20，总共训练 50 个 epoch。

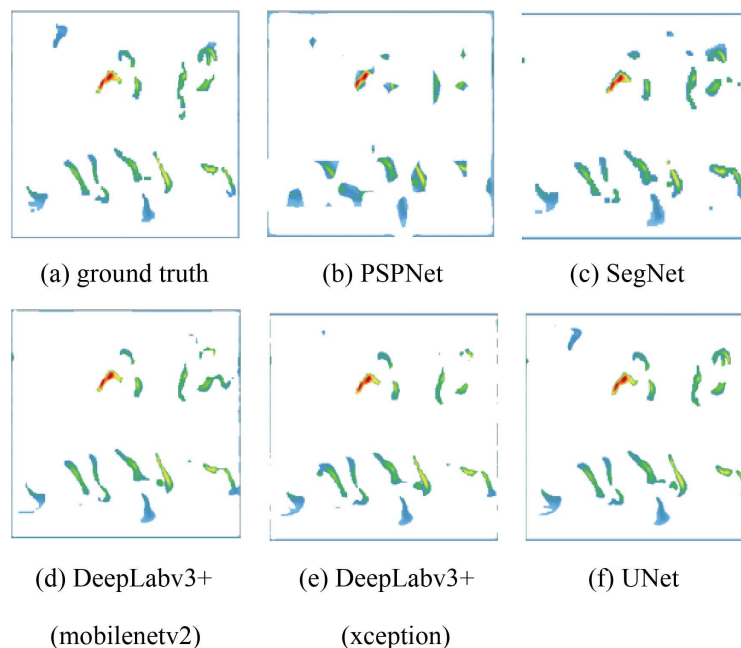
### 5.3. 实验结果

表 2 展示了这几种模型的定量对比结果，可以看出，UNet 模型在我们的数据集上表现的最好，MIoU 值达到 83.00%，在各项评价指标上都要优于其他几个模型。

**Table 2.** Quantitative comparison results of basic models  
**表 2.** 基础模型定量对比结果

Methods	Backbone	Epochs	Transfer	PA	MPA	MIoU	FWIoU
UNet	vgg16	50	✓	<b>96.73%</b>	<b>92.43%</b>	<b>83.00%</b>	<b>94.17%</b>
PSPNet	resnet50	50	✓	93.64%	82.99%	70.90%	89.42%
SegNet	convent	50	✓	93.68%	89.47%	73.21%	89.80%
FCN	mobilenet	50	✓	96.47%	89.15%	80.86%	92.91%
DeepLabv3+	xception	50	✓	95.72%	85.20%	77.28%	92.35%

为了更直观地展示实验对比结果，我们还把分割的结果可视化出来和真实值作对比，如图 10 所示。图 10(f)是 UNet 模型识别大气河的结果，识别效果非常不错，几乎检测出了所有的大气河，形状也最接近真实值，在测试集的平均准确率也高达 92.43%，MIoU 值则达到 83.00%。图 10(b)、图 10(c)分别是 PSPNet 和 SegNet 的检测结果，它们在测试集的各项指标相较 UNet 都有所下降，尤其是 PSPNet 的平均准确率下降了将近 10 个百分点，MIoU 值只有 70.90%。这两个模型的分割结果相比其他模型都要粗糙一些，没有其他模型分割结果那么光滑。图 10(d)和图 10(e)则是 FCN 和 DeepLabv3+得到的实验结果，它们的各项指标也都比较高，但相比 UNet 的各项指标还是下降了几个百分点，这可能是因为它对大气河特征的提取能力不如 UNet。



**Figure 10.** Visual comparison results of basic models  
**图 10.** 基础模型可视化对比结果

通过这些模型的实验对比结果,最终选择了 UNet 作为基本网络结构。为了更好地结合大气河的特性,我们加入了一些特定结构后得到了 SAR-UNet,同时使用相同的实验参数也做了一个对比实验。同样地,我们给出了定量的对比结果,如下表 3 所示。从我们的结果可以看出,SAR-UNet 的 MIoU 值提高了 2 个百分点,达到了 85.22%,同时其他评价指标也有所提升。

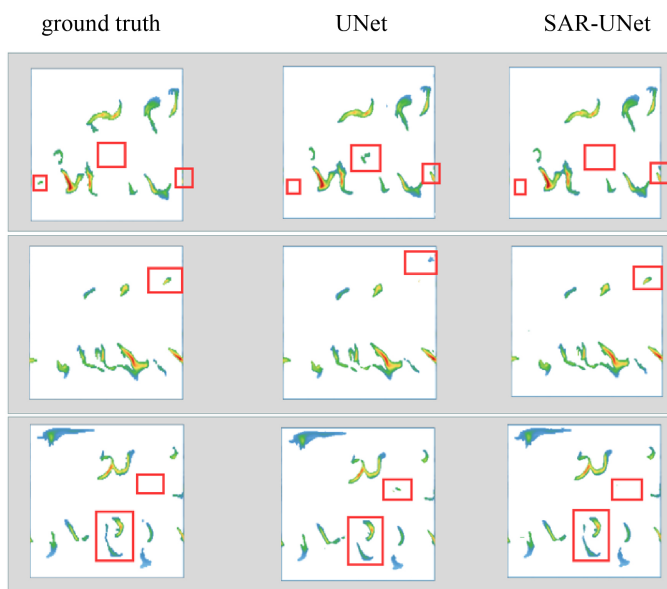
**Table 3.** Quantitative comparison results  
**表 3.** 定量对比结果

Methods	Epochs	Transfer	PA	MPA	MIoU	FWIoU
UNet	50	✓	96.73%	92.43%	83.00%	94.17%
<b>SAR-UNet</b>	50	✓	<b>97.27%</b>	<b>93.05%</b>	<b>85.22%</b>	<b>95.03%</b>

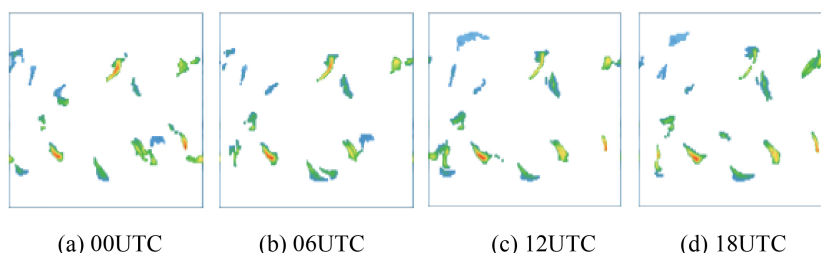
为了更加直观地展示他们之间的差异,随机选择了三个时刻的大气河图像来进行更细致的对比,图 11 是我们给出的可视化对比结果。我们可以清楚地看到,SAR-UNet 对大气河的细节处理得更好,相比 UNet 的识别结果,SAR-UNet 少了一些冗余信息。这说明,SAR-UNet 在能正确识别大气河的同时,比 UNet 更加地注重大气河的局部信息,因此达到了更精确的识别效果。

### 5.4. 未来工作

根据我们的数据集和我们在测试集上的检测结果,我们发现,大气河具有时间连续性,大气河的发生、发展和消亡具有一定的周期性。同一天的大气河,它的位置都差不多,只有较小程度的演变,如图 12 展示了 2018 年 1 月 1 日的大气河的变化。因此,我们想在未来加入时间序列的模块,希望通过引入时间序列进一步提高大气河检测的精度。



**Figure 11.** Comparison of UNet and SAR-UNet visualization results  
**图 11.** UNet 和 SAR-UNet 可视化对比结果



**Figure 12.** Atmospheric river on January 1, 2018  
**图 12.** 2018 年 1 月 1 日的大气河

## 6. 总结

本文提出了一种新的深度学习的方法对大气河进行识别,我们利用语义分割中的 UNet 模型作为基础网络,提出了 SAR-UNet 模型,从而对大气河进行像素级的识别。我们通过用几个比较经典的语义分割模型和我们的模型进行对比,验证了我们的模型更适合用来识别大气河,我们为大气河的识别提供了一个全新的思路。在未来,我们会考虑针对大气河的周期性加入一些时间序列结构,从而实现对大气河更精细的分割。

## 基金项目

四川省科技计划项目(编号: 2023YFQ0072, 22QYCX0082)。

## 参考文献

- [1] Zhu, Y. and Newell, R.E. (1994) Atmospheric Rivers and Bombs. *Geophysical Research Letters*, **21**, 1999-2002. <https://doi.org/10.1029/94GL01710>
- [2] Ralph, F.M., Wilson, A.M., Shulgina, T., *et al.* (2019) ARTMIP-Early Start Comparison of Atmospheric River Detection Tools: How Many Atmospheric Rivers Hit Northern California’s Russian River Watershed? *Climate Dynamics*, **52**,

- 4973-4994. <https://doi.org/10.1007/s00382-018-4427-5>
- [3] Shields, C.A., Rutz, J.J., Leung, L.Y., *et al.* (2018) Atmospheric River Tracking Method Intercomparison Project (ARTMIP): Project Goals and Experimental Design. *Geoscientific Model Development*, **11**, 2455-2474. <https://doi.org/10.5194/gmd-11-2455-2018>
- [4] Leung, L.R. and Qian, Y. (2009) Atmospheric Rivers Induced Heavy Precipitation and Flooding in the Western US Simulated by the WRF Regional Climate Model. *Geophysical Research Letters*, **36**, L03820. <https://doi.org/10.1029/2008GL036445>
- [5] Tang, Y., Huang, A.N., Wu, P.L., *et al.* (2021) Drivers of Summer Extreme Precipitation Events over East China. *Geophysical Research Letters*, **48**, e2021GL093670. <https://doi.org/10.1029/2021GL093670>
- [6] Chapman, W.E., Subramanian, A.C., Delle Monache, L., Xie, S.P. and Ralph, F.M. (2019) Improving Atmospheric River Forecasts with Machine Learning. *Geophysical Research Letters*, **46**, 10627-10635. <https://doi.org/10.1029/2019GL083662>
- [7] Xu, G.Z., Ma, X.H., Chang, P. and Wang, L. (2020) Image-Processing-Based Atmospheric River Tracking Method Version 1 (IPART-1). *Geoscientific Model Development*, **13**, 4639-4662. <https://doi.org/10.5194/gmd-2020-135>
- [8] Neiman, P.J., White, A.B., Ralph, F.M., *et al.* (2013) A Water Vapour Flux Tool for Precipitation Forecasting. *Water and Energy International*, **70**, 72-73.
- [9] Lavers, D.A., *et al.* (2012) A Water Vapour Flux Tool for Precipitation Forecasting. *Proceedings of the Institution of Civil Engineers Water Management*, **162**, 83-94. <https://doi.org/10.1680/wama.2009.162.2.83>
- [10] 苏亮宇, 余晓禹. 基于骨架模型的大气河流识别算法[J]. 移动信息, 2021(4).
- [11] Ralph, F.M., Coleman, T., Neiman, P.J., *et al.* (2013) Observed Impacts of Duration and Seasonality of Atmospheric-River Landfalls on Soil Moisture and Runoff in Coastal Northern California. *Journal of Hydrometeorology*, **14**, 443-459. <https://doi.org/10.1175/JHM-D-12-076.1>
- [12] Wick, G.A., Neiman, P.J. and Ralph, F.M. (2012) Description and Validation of an Automated Objective Technique for Identification and Characterization of the Integrated Water Vapor Signature of Atmospheric Rivers. *IEEE Transactions on Geoscience and Remote Sensing*, **51**, 2166-2176. <https://doi.org/10.1109/TGRS.2012.2211024>
- [13] Rutz, J.J., Steenburgh, W.J. and Ralph, F.M. (2014) Climatological Characteristics of Atmospheric Rivers and Their Inland Penetration over the Western United States. *Monthly Weather Review*, **142**, 905-921. <https://doi.org/10.1175/MWR-D-13-00168.1>
- [14] Guan, B. and Waliser, D.E. (2015) Detection of Atmospheric Rivers: Evaluation and Application of an Algorithm for Global Studies. *Journal of Geophysical Research: Atmospheres*, **120**, 12514-12535. <https://doi.org/10.1002/2015JD024257>
- [15] Pan, M. and Lu, M. (2020) East Asia Atmospheric River Catalog: Annual Cycle, Transition Mechanism, and Precipitation. *Geophysical Research Letters*, **47**, e2020GL089477. <https://doi.org/10.1029/2020GL089477>
- [16] Ohanuba, F.O., Ismail, M.T., Ali, M.K.M., *et al.* (2021) On the Cluster Validity Test(s) in Unsupervised Machine Learning TDA Approach for Atmospheric River Patterns on Flood Detection in Nigeria. <https://doi.org/10.21203/rs.3.rs-459258/v1>
- [17] Muszynski, G., Kashinath, K., Kurlin, V., Wehner, M. and Prabhat (2019) Topological Data Analysis and Machine Learning for Recognizing Atmospheric River Patterns in Large Climate Datasets. *Geoscientific Model Development*, **12**, 613-628. <https://doi.org/10.5194/gmd-12-613-2019>
- [18] Guan, B. and Waliser, D.E. (2019) Tracking Atmospheric Rivers Globally: Spatial Distributions and Temporal Evolution of Life Cycle Characteristics. *Journal of Geophysical Research: Atmospheres*, **124**, 12523-12552. <https://doi.org/10.1029/2019JD031205>
- [19] Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., *et al.* (2017) A Review on Deep Learning Techniques Applied to Semantic Segmentation. *Applied Soft Computing*, **70**, 41-65. <https://doi.org/10.1016/j.asoc.2018.05.018>
- [20] Ronneberger, O., Fischer, P. and Brox, T. (2015) U-net: Convolutional Networks for Biomedical Image Segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, Cham, 234-241. [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)
- [21] Sultana, F., Sufian, A. and Dutta, P. (2020) Evolution of Image Segmentation Using Deep Convolutional Neural Network: A Survey. *Knowledge-Based Systems*, **201**, Article ID: 106062. <https://doi.org/10.1016/j.knosys.2020.106062>
- [22] Simonyan, K. and Zisserman, A. (2014) Very Deep Convolutional Networks for Large-Scale Image Recognition. ArXiv: 1409.1556.
- [23] Hu, J., Shen, L. and Sun, G. (2018) Squeeze-and-Excitation Networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 7132-7141. <https://doi.org/10.1109/CVPR.2018.00745>

- 
- [24] He, K., Zhang, X., Ren, S., *et al.* (2016) Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, 27-30 June 2016, 770-778. <https://doi.org/10.1109/CVPR.2016.90>
- [25] Vaswani, A., Shazeer, N., Parmar, N., *et al.* (2017) Attention Is All You Need. *31st Conference on Neural Information Processing Systems (NIPS 2017)*, Long Beach, CA, 4-9 December 2017, 1-15.
- [26] Zhao, H., Shi, J., Qi, X., *et al.* (2017) Pyramid Scene Parsing Network. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, 21-26 July 2017, 2881-2890. <https://doi.org/10.1109/CVPR.2017.660>
- [27] Badrinarayanan, V., Kendall, A. and Cipolla, R. (2017) Segnet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **39**, 2481-2495. <https://doi.org/10.1109/TPAMI.2016.2644615>
- [28] Chen, L.C., Zhu, Y., Papandreou, G., *et al.* (2018) Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. *Proceedings of the European Conference on Computer Vision (ECCV)*, Munich, 8-14 September 2018, 801-818. [https://doi.org/10.1007/978-3-030-01234-2\\_49](https://doi.org/10.1007/978-3-030-01234-2_49)
- [29] Long, J., Shelhamer, E. and Darrell, T. (2015) Fully Convolutional Networks for Semantic Segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, 7-12 June 2015, 3431-3440. <https://doi.org/10.1109/CVPR.2015.7298965>
- [30] Chen, L.C., Papandreou, G., Schroff, F., *et al.* (2017) Rethinking Atrous Convolution for Semantic Image Segmentation. ArXiv: 1706.05587.
- [31] Weiss, K., Khoshgoftaar, T.M. and Wang, D.D. (2016) A Survey of Transfer Learning. *Journal of Big Data*, **3**, Article No. 9. <https://doi.org/10.1186/s40537-016-0043-6>
- [32] Pan, S.J., Tsang, I.W., Kwok, J.T., *et al.* (2010) Domain Adaptation via Transfer Component Analysis. *IEEE Transactions on Neural Networks*, **22**, 199-210. <https://doi.org/10.1109/TNN.2010.2091281>
- [33] Zhuang, F., Qi, Z., Duan, K., *et al.* (2020) A Comprehensive Survey on Transfer Learning. *Proceedings of the IEEE*, **109**, 43-76. <https://doi.org/10.1109/JPROC.2020.3004555>