

# 基于IASPP-ResNet分割算法的手势识别

雷 玉<sup>1,2</sup>, 崔振超<sup>1,2</sup>, 陈丽萍<sup>1</sup>, 陈向阳<sup>1</sup>, 王煜骁<sup>1</sup>

(1. 河北大学网络空间安全与计算机学院, 河北保定 071002; 2. 河北省机器视觉工程研究中心, 河北保定 071002)

**摘要:** 手势识别是计算机视觉领域中研究的重要领域, 是人机交互领域的重要组成部分. 由于其识别结果受到复杂背景的影响, 手势识别面临着巨大挑战. 为了解决复杂背景影响的问题, 本文利用了密集分割+手势分类的组合型模型, 提出了一种新的手势识别算法. 在密集分割部分, 本文提出了改进型空洞空间金字塔池化(Improved Atrous Spatial Pyramid Pooling, IASPP). IASPP通过密集的连接不同空洞率的空洞卷积获取了在不同视野上的手势多尺度信息, 从而提高了特征表述的精确性. 另外, 为了融合不同层级上的细节和空间位置信息, 提升整体网络的分割性能, 本文将IASPP嵌入编码器-解码器结构中, 提出了IASPP-ResNet手势分割算法. 在手势识别部分, 我们利用了深度卷积神经网络模型, 获得了较高的识别率. 实验结果表明, 在目前常用的公共数据集上, 与传统的机器学习方法以及基于深度学习的方法相比, IASPP-ResNet分割算法的准确率更高, 并且本文提出的密集分割+手势分类的组合型模型在NUS-II数据集上的手势识别率可达98.63%, 优于现有的手势识别算法.

**关键词:** 手势分割; 复杂背景; 编码器-解码器; 空洞卷积; 手势识别

中图分类号: TP391.4

文献标识码: A

文章编号: 0372-2112(XXXX)XX-0001-12

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20210049

## Hand Gesture Recognition Based on IASPP-ResNet Segmentation Algorithm

LEI Yu<sup>1,2</sup>, CUI Zhen-chao<sup>1,2</sup>, CHEN Li-ping<sup>1</sup>, CHEN Xiang-yang<sup>1</sup>, WANG Yu-xiao<sup>1</sup>

(1. School of Cyber Security and Computer, Hebei University, Baoding, Hebei 071002, China;

2. Hebei Machine Vision Engineering Research Center, Baoding, Hebei 071002, China)

**Abstract:** Gesture recognition is an essential research area in the field of computer vision, and it is also a significant component of the human-computer interaction. Due to its recognition results can be influenced by complex backgrounds, gesture recognition faces huge challenges. To solve the problem that is affected by the complex background, this paper proposes a new gesture recognition algorithm based on the combined model of dense segmentation and gesture classification. In the dense segmentation part, this paper shows the Improved Atrous Spatial Pyramid Pooling (IASPP). IASPP is a pooling layer in a convolution neural network, which can obtain the refine features by connecting cascade model and parallel model in atrous spatial pyramid pooling. Otherwise, in order to improve the segmentation performance by integrating details and spatial location information at different levels, the IASPP was embedded in a ResNet with encoder-decoder, and we name the method the Improved Atrous Spatial Pyramid Pooling-ResNet (IASPP-ResNet) for gesture segmentation. In the part of gesture recognition, we use the deep convolutional neural network model to obtain a higher recognition rate. The experimental results show that the IASPP-ResNet segmentation algorithm has a higher accuracy rate on the commonly used public data sets, compared with the traditional machine learning methods as well as the deep learning-based methods, and the gesture recognition rate of the combined model of dense segmentation and gesture classification proposed in this paper can reach 98.63% on NUS-II dataset, which is superior to the existing gesture recognition algorithm.

**Key words:** hand gesture segmentation; complex background; encoder-decoder; atrous convolution; gesture recognition

## 1 引言

基于手势识别的手势交互是人机交互领域中基本的交互方式之一,是机器视觉、计算机应用领域研究的重点方向之一<sup>[1]</sup>. 手势识别在交互领域有着广泛的应用,并且在多种环境中有着较强的优势,例如非接触环境、声音嘈杂或安静的环境等.

目前,手势交互方法主要包括手部检测及手势识别两个主要的过程. 手部检测方法主要包括基于传统的机器学习检测算法、基于深度图像的手部检测算法及基于深度学习的手部检测算法. 基于传统机器学习的手部检测算法包括基于肤色模型的检测算法、基于手部局部特征的检测算法等. 常见的肤色模型包括椭圆肤色模型<sup>[2]</sup>、阈值判定法<sup>[3]</sup>、高斯模型<sup>[4]</sup>以及肤色概率分布模型<sup>[5]</sup>等. 目前基于肤色的手势分割技术都是基于皮肤颜色可以正确建模的假设. 虽然上述方法利用皮肤颜色特征可以快速分割手势,但是,由于图像中光源的位置、角度会造成手部肤色不均匀,最终导致分割精度不高. 基于手部局部特征的检测算法目前主要以边缘检测算子<sup>[6]</sup>以及模板匹配<sup>[7]</sup>等方法为主. 虽然边缘检测算子检测速度快,但却无法平衡抗噪性能与分割精确率之间的矛盾. 模板匹配的本质是依次对比像素点与预先设置的模板的匹配程度,至到找到最优匹配点. 在文献<sup>[7]</sup>中, Tofighi 等人提出使用皮肤自适应直方图模板找到肤色区域,接着基于手势的轮廓和边缘进行分割. 对比其他方法可知模板匹配的最大优点是在特定场景下抗噪能力强,但由于手在大小,形状上极具灵活性,因此很大概率会造成提取的轮廓信息不精准. 基于深度图像的手部检测算法<sup>[8]</sup>所用的图像具有深度信息,通过该信息判断手部的位臵,并进行检测,该方法对图像要求较高,推广较为困难.

由于深度学习方法在检测方面的优势,该类方法很快成为计算机应用领域研究的重点. Long<sup>[9]</sup>等人首先提出了全卷积网络(Fully Convolutional Networks, FCN),文中提到的融合了三个尺度特征的FCN-8s要比融合了一个尺度特征的FCN-32s效果好很多,但边缘细节的处理不够精细. 之后,在文献<sup>[10]</sup>中, Ronneberger 等人提出的UNet网络采用了编解码结构,其添加的跳跃连接更加有利于恢复物体边界轮廓,但是不同数据集所对应网络的最优深度不同,最终会增加训练时间. 接着出现的PSPNet<sup>[11]</sup>使用金字塔池化模块将四种不同尺度的特征连接起来,以此提高获取全局信息的能力. ASPP(Atrous Spatial Pyramid Pooling, ASPP)首先在DeepLabv2<sup>[12]</sup>中提出,在DeepLabv3<sup>[13]</sup>和DeepLabv3+<sup>[14]</sup>中也有应用. 值得注意的是,ASPP生成的特征表示并不丰富,因此处理复杂场景任务的能力较差. 对于基于深度学习的手部检测算法,在文献<sup>[15]</sup>中, Dadashzadeh 等

人利用残差网络和空洞空间金字塔池化提高了手势分割的鲁棒性. 在文献<sup>[16]</sup>中, Wei 等人则将单阶段多盒检测器(Single Shot multi-box Detector, SSD)融合在手势分割中,但其中对肤色概率图进行阈值分割时,造成了手部细节信息的丢失.

在手势识别方面,基于卷积神经网络的识别方法成为该领域的主要方法. 在文献<sup>[17]</sup>中, Mohanty 等人在复杂的背景和变化的光照条件下识别手势,他们提出的模型已经在三个基准数据集上进行了测试,均获得良好识别结果. Gongfa 和 Xing 等人的方法超过了手势识别的经典方法,因为它们避免了获取复杂的手工特征的需求<sup>[18,19]</sup>. Ameen 等人利用两个平行 CNN 提取手势特征,在 ASL 基准数据集上达到了 80.34% 的识别精度<sup>[20]</sup>.

由于手势的多变性,手部检测的结果易产生丰富的背景,从而干扰手势识别,降低交互性. 为了解决这一问题,本文提出了密集分割+手势识别的策略. 手势分割可极大化地去除背景带来的冗余信息,减少对手势识别算法的干扰,从而提高手势识别的精确度. 为了提高手势分割的精确度,本文提出一种改进的空洞空间金字塔池化方法(IASPP),该方法结合了级联模式和并行模式进行特征提取,获得了更加丰富的手部特征信息. 并且本文将 IASPP 嵌入到了一个编码器-解码器模型中,我们称之为 IASPP-ResNet. 为了获得更高的手势识别率,本文在深度卷积模型上对 IASPP-ResNet 的分割结果进行识别.

文章的结构安排如下:在第二节中详细介绍本文所提出的改进型空洞空间金字塔池化以及整体网络架构. 在第三节中,为了验证所提方法的有效性对比了现有的手势分割算法和识别算法,并分析了实验结果. 在第四节中,总结了本文的工作.

## 2 基于 IASPP-ResNet 的手势交互算法

为了降低手部背景对手势识别的干扰,本文提出了手部分割+手势识别策略. 在分割任务中,为了获得精确的分割结果,提高手势识别的准确率,本文提出了一种新的卷积池化方法:IASPP,并将 IASPP 应用于 ResNet 主干网络上,形成了 IASPP-ResNet 手势分割算法;在识别任务中,本文在深度卷积模型上对 IASPP-ResNet 的结果进行分类,并与不同的手势识别算法进行了对比.

### 2.1 空洞卷积池化

在语义分割任务中,空洞卷积<sup>[21]</sup>的提出是为了解决特征图获取过程中,图像分辨率与感受野获取之间的矛盾问题. 假设在二维空间中,对于输入特征图进行空洞卷积可以用式(1)表示:

$$y_{i,j} = \sum_{r=1}^{k_r} \sum_{c=1}^{k_c} x[i+d*(r-1), j+d*(c-1)] * w[r, c] \quad (1)$$

其中,  $i, j$  表示特征图的第  $i$  行和第  $j$  列的值;  $k_r$  表示卷积核的行数;  $k_c$  表示卷积核的列数;  $d$  表示空洞率, 通过改变  $d$  的值, 可以修改卷积核的感受野.

丰富的多尺度特征对图像中不同大小物体的分割与检测十分重要. Deeplabv2 中首先提出的 ASPP 利用了具有不同空洞率的空洞卷积对输入进行采样, 获得了具有不同感受野信息的上下文编码. 其 ASPP 可用式 (2) 表示:

$$y = H_{3,6}(x) \oplus H_{3,12}(x) \oplus H_{3,18}(x) \oplus H_{3,24}(x) \quad (2)$$

其中, 我们为了简化公式, 用  $H_{k,d}(x)$  来表示滤波器大小为  $k$ , 空洞率为  $d$  的空洞卷积, 用  $\oplus$  表示不同尺度的特征在通道上进行拼接.

传统的 ASPP 虽然使用了不同空洞率的空洞卷积并行的对图片进行特征提取, 使得特征表达包含了多尺度信息, 提高了分割性能. 但是 4 个分支之间并不会共享信息, 对于形状百变的手势来说产生的特征尺度和特征表达并不丰富, 最终会影响分割性能. 并且在空洞卷积中并不是所有的像素点都会参与计算, 尤其当空洞率增长到一定程度时, 虽然感受野增大但却只选取了少量的像素点进行计算, 在卷积过程中大量的信息会被舍弃, 结果是不利于目标分割任务. 为了解决以上问题, 我们提出了改进的空洞空间金字塔池化, 如图 1 所示.

## 2.2 改进的空洞空间金字塔池化

在 ASPP 的启发下, IASPP 的设计构架结合了并行和级联两种模式. 在并行模式下, 使用不同的空洞率对输入的特征图进行特征编码, 以获取手势的多尺度信息. 在级联模式中, 除第一、二层之外的每一层都将并行模式的输出串联上一层的输出, 以更加密集的方式提取手势信息, 产生更好的特征表达. 由于图像分割对像素的空间位置信息极为敏感, 为了在恢复图像大小的同时融合更多细节信息, 本算法设计了带有不同空洞率的反卷积, 自下而上的与并行模式的输出相连接, 用以恢复局部特征, 使得图像边缘更加光滑. 在并行模式中, 本文使用空洞率为  $\{2^0, 2^1, 2^2, \dots, 2^n\}$  的空洞卷积, 其中一共包括  $n+1$  个空洞卷积对特征图进行多尺度特征提取, 以产生更加丰富的特征表达. 本文取  $n=4$ , 并行模式的输出定义为式 (3):

$$o_i = \begin{cases} H_{1,d[i]}(x), & i=0 \\ H_{3,d[i]}(x), & 1 \leq i \leq 4 \end{cases} \quad (3)$$

其中用  $d$  表示空洞率为  $\{2^0, 2^1, 2^2, 2^3, 2^4\}$  的数组,  $o_i$  表示  $n+1$  个并行模式的输出.

级联模式是将并行模式中第  $i$  层的输出与第  $i-1$  层

的输出串联起来. 并对串联之后的特征图采用与本层相同的空洞率再次进行空洞卷积. IASPP 中级联模式的输出定义为式 (4). 其中  $p_i$  表示级联模式的输出.

$$p_i = \begin{cases} H_{3,d[i]}(o_1), & 0 < i \leq 1 \\ H_{3,d[i]}(o_i \oplus p_{i-1}), & 1 < i \leq 3 \end{cases} \quad (4)$$

接着本文将并行模式中的局部特征通过跳跃连接与不同空洞率的反卷积信息相融合, 使得模型更易于恢复图像的空间、位置信息. 根据式 (2)~式 (4), 可用式 (5)、式 (6) 表示为

$$q_i = \begin{cases} DH_{3,d[3]}(o_4 \oplus p_3), & 0 < i \leq 1 \\ DH_{3,d[4-i]}(o_{4-i+1} \oplus q_{i-1}), & 2 \leq i \leq 3 \\ DH_{3,d[1]}(o_1 \oplus q_3), & 3 < i \leq 4 \end{cases} \quad (5)$$

$$y = o_0 \oplus q_4 \quad (6)$$

其中  $q_i$  表示反卷积之后的输出,  $y$  表示 IASPP 最后的输出,  $DH_{3,d[i]}$  表示卷积核为 3, 空洞率为  $d$  的反卷积.

与传统的 ASPP 相比, 本文所提出的 IASPP 通过堆叠不同大小的空洞卷积获得更大感受野的同时, 也获得了多尺度特征金字塔, 丰富了特征表达. 对于空洞率等于  $d$  的空洞卷积的等效感受野由式 (7) 定义<sup>[22]</sup>:

$$R = k + (k-1)(d-1) \quad (7)$$

其中  $k$  为原始卷积核大小,  $R$  为新的感受野, 将不同空洞率的卷积层堆叠起来可以获得更大的感受野, 假设卷积核大小分别为  $k_1$  和  $k_2$ , 则新的感受野大小由式 (8) 定义<sup>[22]</sup>:

$$k = k_1 + k_2 - 1 \quad (8)$$

根据式 (7) 可以计算出特征金字塔的最大感受野  $R_{\max}$ ,  $R_{k,d}$  表示卷积层的核大小为  $k$ , 空洞率大小为  $d$ . 因此, 由式 (7)、式 (9) 可以得出空洞率为  $\{2^1, 2^2, 2^3, 2^4\}$  的 ASPP 的最大感受野为  $R_{3,16}$  其值为 33.

$$R_{\max} = \max(R_{3,2}, R_{3,4}, R_{3,8}, R_{3,16}) \quad (9)$$

对于 IASPP, 由式 (8)、式 (10) 计算得最大感受野为 89.

$$R_{\max} = R_{3,2} + R_{3,2} + R_{3,4} + R_{3,4} + R_{3,8} + R_{3,8} + R_{3,16} - 6 \quad (10)$$

当神经元的感受野值越大时, 它所对应的原图的映射范围就越大, 也就是包含了更多全局和高级语义特征. 当感受野值越小时, 它所对应的原图的映射范围就越小, 也就是包含了更多的局部和细节特征. 对于复杂背景中的手势存在尺度多样的问题, 本文在 IASPP 中的并行和级联模式中设计了不同的空洞率, 并将不同空洞率的空洞卷积堆叠在一起, 获得了不同感受野上的手势多尺度信息, 丰富了特征表达. 由此可知, IASPP 将全局、高级的语义特征与局部、细节的语义特征联合在一起过滤了背景中的冗余信息, 提高了分割准确率.

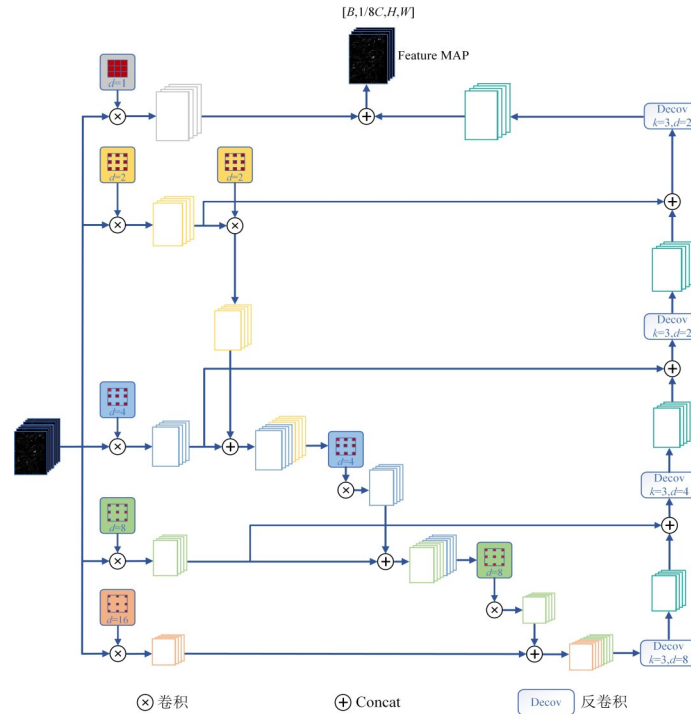


图1 IASPP框架图:图中的 $d$ 表示空洞率大小, $k$ 表示卷积核大小,Concat表示特征图在通道上进行拼接。

图2表示ASPP与IASPP的特征金字塔,其中矩形右侧的 $k$ 表示不同空洞率组合的感受野大小,由图可

知,当空洞率为 $\{2^1, 2^2, 2^3, 2^4\}$ 时IASPP中空洞卷积的堆叠产生了更大的感受野和尺度多样性。

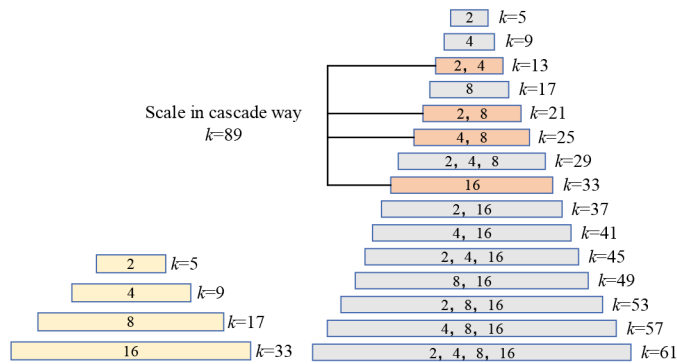


图2 模型的特征金字塔.(a) ASPP模型;(b) IASPP模型.其中,每个条带中的数字表示不同空洞率的组合,每个条带的长度表示每个组合的等效核大小,条带右边的 $k$ 表示相应的感受野大小。

图3表示不同空洞率的空洞卷积堆叠的像素参与量.其中图3(a)代表一维空间下, $d=8$ 的空洞卷积的感受野为17.但是从图中观察到只有3个像素参与计算,也就是参与计算的像素并不连续,有很多像素点会丢失,结果是在更高维的空间中,会损失更多的信息.图3(b)将空洞率为4,8的空洞卷积堆叠在一起,可以看出比原始 $d=8$ 的空洞卷积多出的4个像素将参与计算.图3(c)将空洞率2,4,8的空洞卷积堆叠在一起,可以看出比原始 $d=8$ 的空洞卷积多出的12个像素将参与计算。

### 2.3 手势分割算法

为了过滤背景中的冗余信息,实现手部的精准分

割,本文将IASPP嵌入到一个带有编码-解码器的模型中,其主要由三部分组成,依次为:深度卷积神经网络(Deep Convolutional Neural Network, DCNN)、IASPP和解码器,网络结构如图4所示。

其网络的输入是RGB图像,编码部分由DCNN和IASPP组成.其中DCNN是由1个 $7 \times 7$ 和1个 $3 \times 3$ 的卷积以及4个残差组构成的特征提取的骨干网络,值得注意的是为了使解码结构在减少计算量的同时融合更多的局部细节信息,在 $7 \times 7$ 的卷积以及第1,4个残差组的输出特征后依次添加了 $1 \times 1$ 的卷积,网络参数由表1所示.其中4个残差组的残差单元数为 $\{3, 4, 6, 3\}$ ,并且

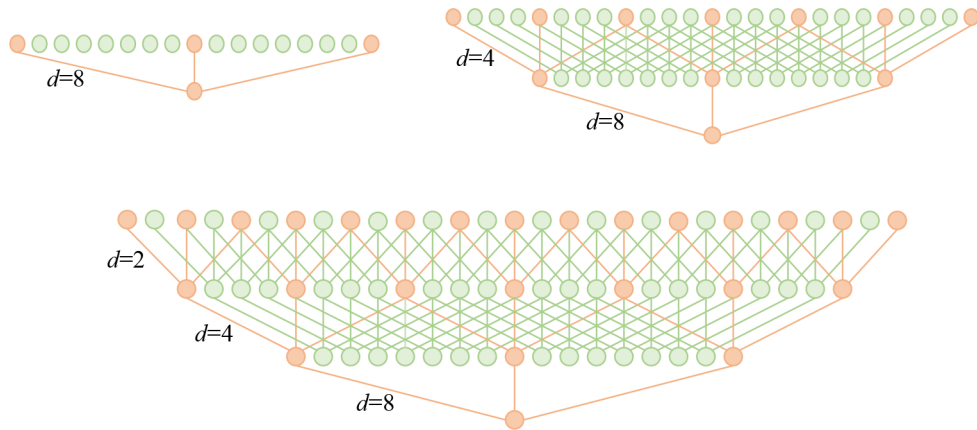


图3 空洞卷积堆叠图:(a) $d=8$ 的空洞卷积;(b)将 $d=4$ 的空洞卷积堆叠在 $d=8$ 的空洞卷积之上;(c)将 $d=2,4$ 的空洞卷积堆叠在 $d=8$ 的空洞卷积之上. 其中橙色表示信息的来源.

为了增大网络的输出大小以适应图像分割任务,本文在第三个和第四个残差组中使用不同空洞率的空洞卷积替换普通卷积,以此来调节卷积核的大小更好的捕获目标的高层语义特征<sup>[23]</sup>,我们用 $d$ 表示每一层卷积中的空洞率, $s$ 则用来表示卷积步长. 最后RGB图像在经过DCNN的特征提取之后,输出的特征图变为原始图像的 $1/8$ .

接着在本文方法中,我们将ResGroup4的输出特征图作为IASPP的输入,其利用堆叠的不同空洞率的空洞卷积对DCNN输出的2048维特征进行特征编码,在丰富特征表达的同时挖掘多尺度的上下文信息.

为了在解码过程中恢复更多的细节特征,本文选择了Conv1、ResGroup1、ResGroup4的三个尺度特征. 并且使用了三个上采样操作来调整特征图的大小,在每次上采样后将其与来自编码部分特征图连接起来. 最后,使用 $3\times 3$ 和 $1\times 1$ 卷积核来细化手势分割. 由此可知,本文利用获取高级语义信息的编码器和利用编码阶段的信息对图进行放大以恢复图像的细节信息的解码器<sup>[24]</sup>,得到了更加准确的手势分割结果.

## 2.4 基于IASPP-ResNet的手势识别方法

在实际应用中,手势识别的鲁棒性并不强. 因此,我们提出了一种密集分割+手势分类的组合型模型,

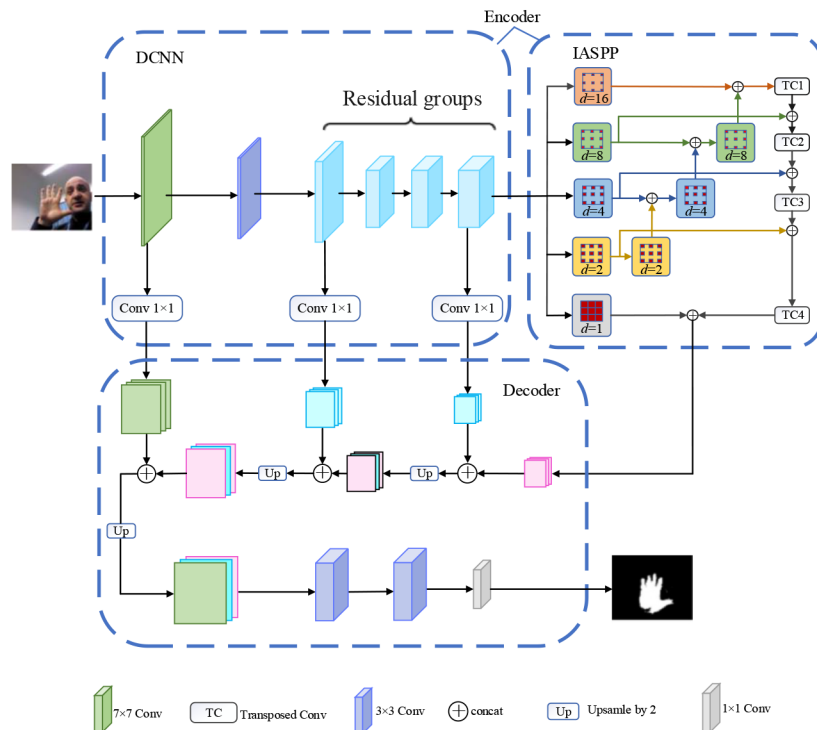


图4 IASPP-ResNet框架

表1 全卷积深度残差网络参数设置

Layer name	Output shape	layers
input	512×512×3	
Convolution1	256×256×64	
Convolution2	128×128×64	
ResGroup1	128×128×256	$\begin{bmatrix} 1 \times 1 & 64 \\ 3 \times 3 & 64 \\ 1 \times 1 & 256 \end{bmatrix} \times 3, d=1, s=2$
ResGroup2	64×64×512	$\begin{bmatrix} 1 \times 1 & 128 \\ 3 \times 3 & 128 \\ 1 \times 1 & 512 \end{bmatrix} \times 4, d=1, s=1$
ResGroup3	64×64×1024	$\begin{bmatrix} 1 \times 1 & 256 \\ 3 \times 3 & 256 \\ 1 \times 1 & 1024 \end{bmatrix} \times 6, d=2, s=1$
ResGroup4	64×64×2048	$\begin{bmatrix} 1 \times 1 & 512 \\ 3 \times 3 & 512 \\ 1 \times 1 & 2048 \end{bmatrix} \times 3, d=4, s=1$

如图5所示. 在密集分割中, 使用提出的手势分割算法来减少背景中的冗余信息, 从而为手势识别阶段做准备. 在识别任务中, 将来自IASPP-ResNet分割算法信息输入到手势识别算法中, 继续进行分类. 在此项工作中, 我们在文献<sup>[25]</sup>的网络框架上改进识别算法, 该模型由三个卷积层以及用于特征提取的ReLU和MaxPooling层, 一个全连接层和一个Softmax层构成手势分类层. 在分类算法中, 我们首先将分割算法的输出随机分为训练集和测试集, 然后再作为输入输入到手势分类层. 第一层中有64个卷积核, 其大小为19×19, 第二层中有128个卷积核, 其大小为17×17, 第三层中有128个卷积核, 卷积核的大小分别为15×15, 步长都为2.

### 3 实验结果与分析

为了进一步证明本文提出的密集分割+手势分类组合型模型的有效性, 在此节中, 我们在OUTHANDS和HGR公共数据集上进行手势分割实验, 在NUS-II数据集上与其他基于深度学习的识别算法进行了对比.

#### 3.1 实验平台及评价指标

实验平台参数为: Windows 10操作系统, CPU型号: Inter Xeon Silver 4215, 内存: 64GB, GPU型号为Nvidia GTX 2080. 实验利用Python进行实现, 所用框架为Pytorch.

本文采用平衡F分数(F-score)<sup>[26]</sup>、平均交并比(Mean Intersection over Union, MIoU)<sup>[27]</sup>、准确率(Precision, Pr)<sup>[28]</sup>、召回率(Recall, Re)<sup>[28]</sup>、像素准确率(Pixel Accuracy, PA)<sup>[29]</sup>作为评价指标来评估IASPP-ResNet分割算法的有效性.

MIoU作为主要评估指标来计算所有类别交并集的平均值, 由式(11)定义.

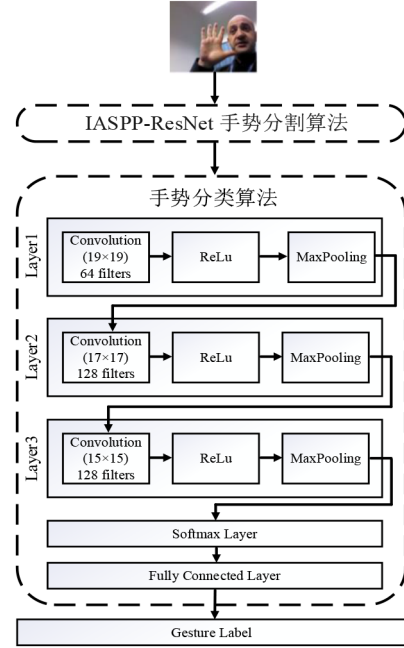


图5 整体框架

$$MIoU = \frac{1}{n+1} \frac{\sum_{i=0}^n p_{ii}}{\sum_{j=0}^n p_{ij} + \sum_{j=0}^n p_{ji} - p_{ii}} \quad (11)$$

式(11)中  $n+1$  表示数据集中所要分割的类别数量.  $p_{ii}$  表示正确分类的像素.  $p_{ij}$  表示原本为第  $i$  类被预测为第  $j$  类的像素,  $p_{ji}$  表示原本为第  $j$  类被预测为第  $i$  类的像素.

PA 代表着正确分类的像素点在所有像素点上所占的比例, 由式(12)定义.

$$PA = \frac{\sum_{i=0}^n p_{ii}}{\sum_{i=0}^n \sum_{j=0}^n p_{ij}} \quad (12)$$

Pr, Re 是分割任务中简单的评价指标, 由式(13)定义, 其中  $N_c$  表示被判定为正类的像素中其标签也为正类的像素,  $N_r$  表示判定为正类的像素,  $N_g$  表示真值为正类的像素.

$$Pr = \frac{N_c}{N_r}, Re = \frac{N_c}{N_g} \quad (13)$$

F-score 则用来权衡 Pr, Re 之间的权重占比, 一般用式(14)定义, 其中使用  $\beta$  来调节 Pr 与 Re 之间的权重.

$$F_\beta = (1 + \beta^2) \times \frac{Pr \times Re}{(\beta^2 \times Pr) + Re} \quad (14)$$

#### 3.2 实验结果与分析

表2及图6显示了IASPP-ResNet算法与几种经典的机器学习方法在HGR数据集上的测试结果. 这几种经典的机器学习方法包括肤色检测方法<sup>[30]</sup>(Discrimina-

tive Skin-Presence Features, DSPF), 基于判别式纹理特征分割算法<sup>[31]</sup> (Discriminative Textural Feature, DTF), 基于皮肤像素检测器的贝叶斯算法<sup>[32]</sup> (Bayesian skin model, Bayes), 基于全局对比度的显著性分割方法<sup>[33]</sup> (the RC saliency detection with Saliency Cut method, RC), 以及基于最大化协议算法的图像分割算法<sup>[5]</sup>. 从表2及图6中, 我们可以看出基于IASPP-ResNet的分割算法在手势分割方面有着较大优势, 其中Pr、Re、F-score及ROC-area等指标分别达到了0.994 8、0.992 9、0.993 9、0.998 2. 这些评价指标均高于对比算法, 这说明了本文所提方法在各方面优于基于传统机器学习类算法.

表2 本文算法与机器学习方法在HGR数据集下的对比结果

Method	Pr	Re	F-score	ROC-area
DSPF <sup>[30]</sup>	0.8297	0.9531	0.8552	0.8546
DTF <sup>[31]</sup>	0.7816	0.9596	0.8166	0.8823
Bayes <sup>[32]</sup>	0.6998	0.9297	0.7422	0.7341
RC <sup>[33]</sup>	0.8085	0.8097	0.8087	0.5798
Conaire <sup>[5]</sup>	0.7858	0.8228	0.7940	0.7034
Ours	<b>0.9948</b>	<b>0.9929</b>	<b>0.9939</b>	<b>0.9982</b>

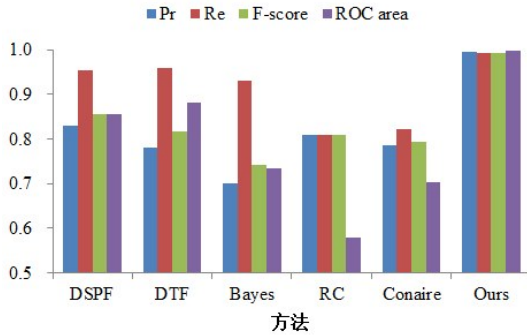


图6 本文算法与机器学习方法的测试结果柱状图

表3记录了本文方法与其他基于深度学习的分割算法, 在OUTHANDS数据集上的测试结果. 从表中可以观察到对于复杂背景中的手势分割任务, IASPP-ResNet算法在MIoU、PA、Re及F-score四种评价指标中, 均优于其他对比算法. 并且IASPP-ResNet算法在四种评价指标上, 分别达到了0.8960、0.9775、0.9820和0.9863, 比次优算法分别提高了0.29% (SPNet)、0.12% (SPNet)、0.07% (SPNet)及0.02% (SPNet). 表中的第七列记录了每种算法的运行速度, 其中算法<sup>[34]</sup>具有良好的运行速度, 但MIoU、F-score值较低, 对比其他算法可知本文的算法在时间和精度上表现都较为良好.

表4显示了各算法在HGR数据集上的测试结果. 从表中我们可以看出, 在MIoU、PA及F-score评价指标上, IASPP-ResNet方法均优于其他算法结果, 分别获得

表3 在OUTHANDS数据集上的测试结果

Method	MIoU	PA	Pr	Re	F-score	Time
FCN-8s	0.783 0	0.939 3	0.993 9	0.936 7	0.964 4	73.05
PSPNet	0.813 4	0.950 6	0.994 2	0.949 2	0.971 2	92.26
ICNET[34]	0.772 8	0.935 1	0.988 4	0.937 5	0.962 3	<b>39.05</b>
DeepLabv3	0.797 0	0.944 2	<b>0.994 3</b>	0.941 9	0.967 4	113.74
DeepLabv3+	0.887 2	0.973 6	0.990 4	0.979 5	0.984 9	114.66
DUNet[35]	0.892 9	0.974 1	0.990 6	0.979 9	0.985 2	82.59
SPNet[36]	0.893 1	0.976 3	0.991 0	0.981 3	0.986 1	96.23
IASPP	<b>0.896 0</b>	<b>0.977 5</b>	0.990 7	<b>0.982 0</b>	<b>0.986 3</b>	121.01

表4 在HGR数据集上的测试结果

Method	MIoU	PA	Pr	Re	F-score
FCN-8s	0.816 8	0.937 6	0.926 0	<b>0.997 6</b>	0.960 5
PSPNet	0.892 4	0.960 1	0.996 8	0.950 2	0.972 9
ICNET[34]	0.885 4	0.957 3	0.994 8	0.948 7	0.971 2
DeepLabv3	0.882 6	0.956 0	<b>0.997 4</b>	0.944 5	0.970 2
DeepLabv3+	0.959 4	0.985 5	0.993 0	0.987 4	0.990 2
DUNet[35]	0.961 1	0.986 3	0.995 0	0.986 6	0.990 8
SPNet[36]	0.971 2	0.990 1	0.993 4	0.990 1	0.991 7
IASPP	<b>0.975 2</b>	<b>0.990 9</b>	0.994 8	0.992 9	<b>0.993 9</b>

了0.975 2、0.990 9及0.993 9, 比次优算法结果高出了0.4% (SPNet)、0.08% (SPNet)及0.22% (SPNet). 并且在Pr及Re两项指标中, IASPP-ResNet分别获得了0.994 8及0.992 9的结果, 均高于SPNet的结果. 总体上, 在HGR数据集的测试中, 本文方法优于对比方法.

图7显示了OUTHANDS数据集的测试结果. 我们可以看出数据集中的典型图片具有较强的干扰信息, 例如复杂的背景以及与手部有相似颜色、纹理的人脸信息. 从结果图片中, 我们可以看出输入图片在手部与脸部发生重合 (test\_01和test\_07), 复杂背景 (test\_02、test\_03及test\_05)及光照变化较大 (test\_04和test\_06)等复杂性情况下, 其他对比算法结果均出现分割结果与手动分割标准差别大、边缘细节处理不够精准以及检测出非手部区域等情况. 而本文提出的IASPP-ResNet算法的结果, 如图7第三列所示, 最接近数据集的真值图 (Ground Truth, GT). 这说明本文算法对光照、复杂背景及相近肤色干扰等情况并不敏感, 在这些情况下可以获得较为精确的分割结果.

图8显示了HGR数据集中不同颜色背景下, 各分割算法对不同手势图像分割的结果. 其中当手势的手指趋于聚拢时 (test\_02和test\_05), 本文的测试结果更加接近GT, SPNet、DUNet以及DeepLabv3+虽然也能完整的将手势分割出来, 但却产生了噪声, 而其余四种方法无法保留手指轮廓的细节信息. 由此可见, 在不同颜色背景下, 本文提出的基于IASPP-ResNet的手势分割算法比其他对比算法具有更高的精确度, 结果更加接

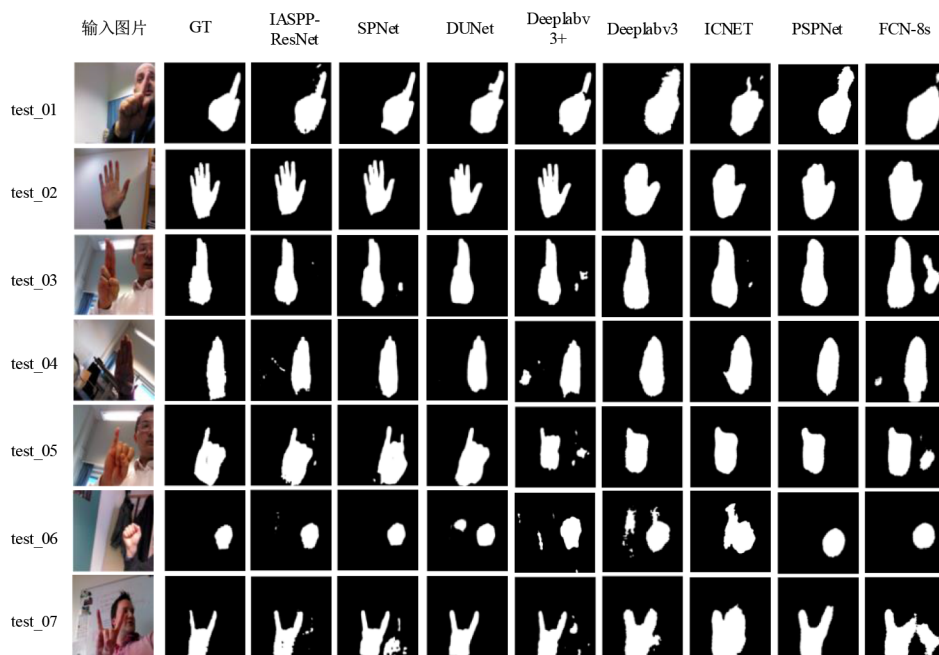


图7 OUTHANDS数据集上各算法测试结果

近GT.

手势分割是手势识别的一个预处理环节,其分割结果的好坏将直接影响手势识别的准确率. 鉴于OUTHANDS数据集中的手势图像不仅有真值图,其每个模板还有相应的分类标签,因此本文在识别阶段中利用分割以后的图像进行分类,对不同算法的性能进行了评估. 如表5所示,本文提出的密集分割+手势分类的识别精度可达98.61%,比手势识别算法<sup>[25]</sup>提高了3.99%,并且在优于其他对比算法同时其运

行时间并没有大幅度增加. 由此可见,本文提出的分割算法可极大化的过滤背景中的干扰信息,提高手势识别的准确率.

为了进一步证明本文提出的密集分割+手势识别算法可以提高手势识别率,我们还在NUS-II数据集上与其他基于深度学习的算法进行了对比. 结果如表6所示,本文算法的手势识别率可达98.63%,比次优算法提高了0.33%. 由此可见,本文算法可使得手势与背景的分割更加准确,可进一步提高手势识别率.

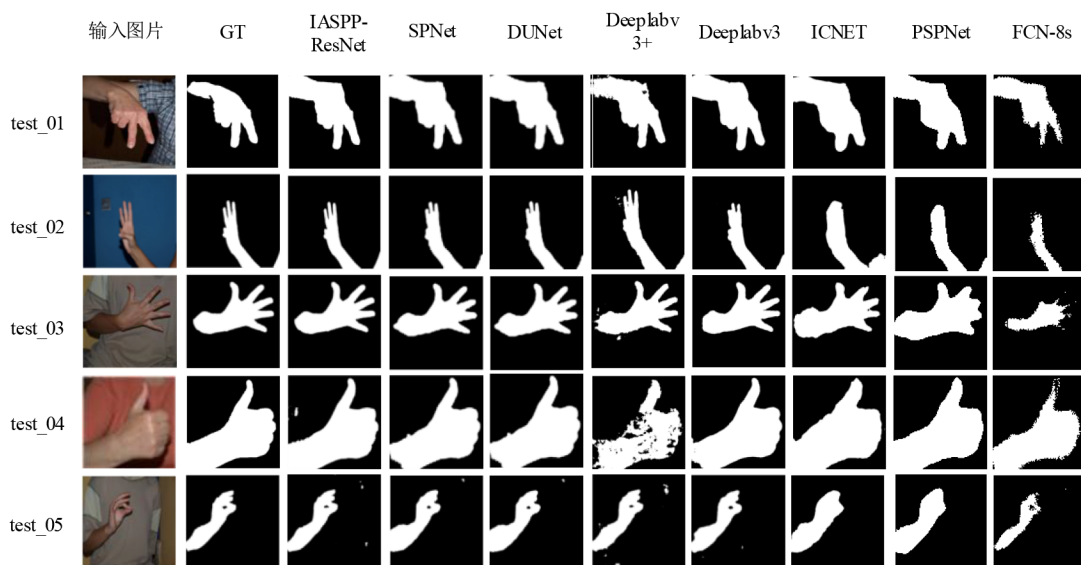


图8 HGR数据集上各算法测试结果

表 5 在 OUTHANDS 数据集上的识别率

Method	Time In OUTHANDS	Accuracy
文献[25]	/	0.946 2
Deeplabv3+文献[25]	123.65	0.965 3
DUNet+文献[25]	92.60	0.979 4
SPNet+文献[25]	109.26	0.980 9
本文	130.42	<b>0.986 1</b>

表 6 在 NUS-II 数据集上的识别率

Method	Accuracy
文献[26]	0.965 3
文献[37]	0.983 0
DAG-CNN[38]	0.910 5
CNN-SPP[39]	0.959 5
文献[25]	0.966 2
本文	<b>0.986 3</b>

### 3.3 空洞率的选择对实验结果的影响

空洞率的值决定了空洞卷积中各个元素之间的间隔,因此不同的组合会影响卷积层的特征提取效果和分割的精度.为了更好的处理数据集中的大小物体的关系,本文采用分层特征融合(Hierarchical Feature Fusion, HFF)<sup>[40]</sup>策略的思想,将 IASPP 中的空洞率设置为与其相同的大小.通过多组对比实验,选取了  $\{2^0, 2^1, 2^2, 2^3, 2^4\}$ 、 $\{2^0, 2^1, 2^2, 2^1+2^3, 2^2+2^4\}$ 、 $\{2^0, 2^2, 2^3, 2^2+2^3, 2^2+2^4\}$ 、 $\{2^0, 2^1, 2^1+2^2, 2^1+2^3, 2^4\}$  四组实验结果较好的空洞率组合,其实验结果如表 7、表 8 所示.

表 7 在 HGR 数据集上的测试结果

Rate	MIoU	F-score
$\{2^0, 2^1, 2^2, 2^3, 2^4\}$	0.975 2	<b>0.993 9</b>
$\{2^0, 2^1, 2^2, 2^1+2^3, 2^2+2^4\}$	<b>0.975 3</b>	<b>0.993 9</b>
$\{2^0, 2^2, 2^3, 2^2+2^3, 2^2+2^4\}$	0.973 4	0.993 5
$\{2^0, 2^1, 2^1+2^2, 2^1+2^3, 2^4\}$	0.974 1	0.993 7

表 8 在 OUTHANDS 数据集上的测试结果

Rate	MIoU	F-score
$\{2^0, 2^1, 2^2, 2^3, 2^4\}$	0.896 0	0.986 3
$\{2^0, 2^1, 2^2, 2^1+2^3, 2^2+2^4\}$	<b>0.912 5</b>	<b>0.989 1</b>
$\{2^0, 2^2, 2^3, 2^2+2^3, 2^2+2^4\}$	0.899 6	0.987 7
$\{2^0, 2^1, 2^1+2^2, 2^1+2^3, 2^4\}$	0.900 2	0.987 2

由表 7 的实验结果可知,四组不同空洞率的组合对 HGR 数据集的影响并不大,MIoU 和 F-score 的差距都极小.从表 8 中可以看出,当空洞率组合为  $\{2^0, 2^1, 2^2, 2^1+2^3, 2^2+2^4\}$  时,此时模型在 OUTHANDS 数据集上的 MIoU 以及 F-score 表现最佳,相比于空洞率组合  $\{2^0, 2^1, 2^2, 2^3, 2^4\}$  分别提升了 1.65%、0.28%.

### 3.4 消融实验

为了更好的证明分割任务中各个模块的作用,本次实验评估了 IASPP 中的级联模式、反卷积以及不同解码器对 IASPP-ResNet 手势分割算法的影响.空洞率组合采取固定值  $\{2^0, 2^1, 2^2, 2^3, 2^4\}$ ,实验分四组进行.第一组去掉 IASPP 中的反卷积,使得 IASPP 中经过级联模式的输出特征图直接上采样到与 DCNN 的输出特征图大小一样,接着继续输入到原有的解码器中.第二组,我们去掉 IASPP 中的级联模式,将并行模式的输出输入到带有不同空洞率反卷积中,接着继续输入到原有的解码器中.第三组,我们改变了原本的解码结构,在解码阶段使得 IASPP 的输出特征与 DCNN 中 Convolution1、ResGroup1、ResGroup2、ResGroup4 的特征在通道上进行拼接.第四组为本文所提模型.表 9、表 10 记录了不同网络架构的 MIoU、PA 和 F-score 的数值.

表 9 不同网络架构对 OUTHANDS 数据集的影响

Rate	MIoU	PA	F-score
DCNN+并行、级联模式+解码器	0.894 8	0.975 9	0.985 3
DCNN+并行模型、反卷积+解码器	0.892 0	0.973 2	0.983 3
DCNN+ IASPP+新的解码器	<b>0.902 9</b>	0.976 4	<b>0.986 6</b>
DCNN+IASPP+解码器(本文模型)	0.896 0	<b>0.977 5</b>	0.986 3

表 10 不同网络架构对 HGR 数据集的影响

Rate	MIoU	PA	F-score
DCNN+并行、级联模式+解码器	0.973 7	0.987 8	<b>0.993 9</b>
DCNN+并行模型、反卷积+解码器	0.971 8	0.984 9	0.993 4
DCNN+ IASPP+新的解码器	0.975 0	<b>0.990 9</b>	0.993 8
DCNN+IASPP+解码器(本文模型)	<b>0.975 2</b>	<b>0.990 9</b>	<b>0.993 9</b>

从表 9、10 中我们可以观察到,无论是去掉 IASPP 中的反卷积还是级联模式,模型的分割性能明显下降.在 OUTHANDS 数据集上,本文模型与第二组模型相比在 MIoU、PA、F-score 评价指标上分别提升了 0.4%、0.43%、0.3%,在 HGR 数据集上分别提升了 0.34%、0.6%、0.05%.对比表 9、10 中的三、四行可知,多融合了 ResGroup2 特征尺度的解码器的网络性能并没有明显的提升.在综合考虑网络的复杂性之后,可知我们设计的解码器更为简单有效.

## 4 结论

为了最大化过滤掉复杂背景中的冗余信息,提高手势识别的准确率,本文提出了密集分割+手势分类的组合型模型.在分割任务中,我们在ASPP的基础上进行改进提出了IASPP方法.该方法在ASPP的并行模式中融合了级联模式,形成一种新的具有两种模式特性的混合模式.在混合模式下,本文用五种不同空洞率的空洞卷积并行的对全卷积残差网络所提取的特征进行编码,捕获不同尺度的信息,并且低层次空洞卷积的结果会传入高层次空洞卷积中,真正做到了局部特征和全局特征的融合,丰富了特征表达.算法最后设计的自下而上的反卷积用以恢复空间位置信息,细化图像的边缘和细节.最后将IASPP嵌入编解码结构中,提出IASPP-ResNet手势分割算法.在识别任务中,将分割阶段学到的特征表示送入深度卷积神经网络中.实验证明了本文方法在类肤色、手脸遮挡、以及非均匀的光照条件等情况下,都具有较强的鲁棒性并能够获得更高的识别率,且分割算法优于现有的基于传统的机器学习方法以及SPNet等基于深度学习框架的分割算法.因此可以证明,本文所提算法能够极大化的过滤背景的冗余信息,提高手势识别的准确率.

## 参考文献

- [1] 王勇,王沙沙,田增山,等.基于FMCW雷达的双流融合神经网络手势识别方法[J].电子学报,2019,47(7):1408-1415.  
WANG Y, WANG S S, TIAN Z S, et al. Two-stream fusion neural network approach for hand gesture recognition based on FMCW radar[J]. Acta Electronica Sinica, 2019, 47(7): 1408-1415. (in Chinese)
- [2] SAYED U, MOFADDEL M A, BAKHEET S, et al. An elliptical boundary skin model for hand detection based on HSV color space[J]. Information Sciences Letters, 2018, 7(1):13-17.
- [3] LIU C, WANG J, ZHANG T, et al. Adaptive threshold gesture segmentation algorithm based on skin color[C]//Proceedings of 2016 2nd International Conference on Advances in Mechanical Engineering and Industrial Informatics (AMEII 2016). Hangzhou: Computer Science and Electronic Technology International Society, 2016: 1602-1605.
- [4] ZHENG Y, ZHENG P. Hand segmentation based on improved gaussian mixture model[C]//2015 International Conference on Computer Science and Applications (CSA). Wuhan: IEEE, 2015:168-171.
- [5] CONAIRE C O, O'CONNOR N E, SMEATON A F. Detector adaptation by maximising agreement between independent data sources[C]//Conference on Computer Vision and Pattern Recognition. Minneapolis, MN, USA: IEEE, 2007: 1-6.
- [6] WANG X, FANG Y, LI C, et al. Static gesture segmentation technique based on improved sobel operator[J]. The Journal of Engineering, 2019, 2019(22).
- [7] TOFIGHI G, MONADJEMI S A, GHASEM-AGHAEI N. Rapid hand posture recognition using adaptive histogram template of skin and hand edge contour[C]//Iranian Conference on Machine Vision and Image Processing. Isfahan: IEEE, 2010: 1-5.
- [8] CHEN D, LI G, SUN Y, et al. Fusion hand gesture segmentation and extraction based on CMOS sensor and 3D sensor [J]. International Journal of Wireless and Mobile Computing, 2017, 12(3): 305-312.
- [9] LONG J, SHEHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston:IEEE, 2015: 3431-3440.
- [10] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation [C]//International Conference on Medical Image Computing and Computer-assisted Intervention. Cham: Springer International Publishing, 2015: 234-241.
- [11] ZHAO H, SHI J, QI X, et al. Pyramid scene parsing network[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017: 2881-2890.
- [12] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 40(4): 834-848.
- [13] CHEN L C, PAPANDREOU G, SCHROFF F, et al. Rethinking atrous convolution for semantic image segmentation[EB/OL]. (2017-12-05)[2020-12-31]. <https://arxiv.org/abs/1706.05587>.
- [14] CHEN L C, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]//Proceedings of the European Conference on Computer Vision. Munich: Springer, 2018: 801-818.
- [15] DADASHZADEH A, TARGHI A T, TAHMASBI M, et al. HGR-Net: A fusion network for hand gesture segmentation and recognition[J]. IET Computer Vision, 2019, 13(8): 700-707.

- [16] 卫保国, 徐勇, 刘金玮, 等. 融合SSD目标检测的自适应手势分割方法[J]. 信号处理, 2020, 36(07): 1038-1047.  
WEI B G, XU Y, LIU J W, et al. Adaptive gesture segmentation based on SSD object detection[J]. Journal of Signal Processing, 2020, 36(7): 1038-1047. (in Chinese)
- [17] MOHANTY A, RAMBHATLA S S, SAHAY R R. Deep gesture: Static hand gesture recognition using CNN[C]// Proceedings of International Conference on Computer Vision and Image Processing. Singapore: Springer, 2017: 449-461.
- [18] YANG H L, XUAN S B, MO Y B. Hand gesture recognition based on convolution neural network[J]. Cluster Computing, 2019, 22(2): 2719-2729.
- [19] XING K, DING Z, JIANG S, et al. Hand gesture recognition based on deep learning method[C]//2018 IEEE Third International Conference on Data Science in Cyberspace (DSC). Guangzhou: IEEE, 2018: 542-546.
- [20] AMEEN S, VADERA S. A convolutional neural network to classify American Sign Language fingerspelling from depth and colour images[J]. Expert Systems, 2017, 34(3): e12197.
- [21] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions[EB/OL]. (2016-04-30)[2020-12-31]. <https://arxiv.org/abs/1511.07122>.
- [22] YANG M K, YU K, ZHANG C, et al. DenseASPP for semantic segmentation in street scenes[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. 3684-3692.
- [23] 李宝奇, 贺昱曜, 何灵蛟, 等. 基于全卷积神经网络的非对称并行语义分割模型[J]. 电子学报, 2019, 47(5): 1058-1064.  
LI B Q, HE Y Y, HE L J, et al. Asymmetric parallel semantic segmentation model based on full convolutional neural network[J]. Acta Electronica Sinica, 2019, 47(5): 1058-1064. (in Chinese)
- [24] 孟磊, 徐磊, 郭嘉阳. 一种基于改进的MobileNetV2网络语义分割算法[J]. 电子学报, 2020, 48(9): 1769-1776.  
MENG L, XU L, GUO J Y. Semantic segmentation algorithm based on improved MobileNetV2[J]. Acta Electronica Sinica, 2020, 48(9): 1769-1776. (in Chinese)
- [25] ADITHYA V, RAJESH R. A deep convolutional neural network approach for static hand gesture recognition[J]. Procedia Computer Science, 2020, 171:2353-2361.
- [26] ZHANG Q, YANG M, KPALMA K, et al. Segmentation of hand posture against complex backgrounds based on saliency and skin colour detection[J]. IAENG International Journal of Computer Science, 2018, 45(3): 435-444.
- [27] 罗会兰, 张云. 基于深度网络的图像语义分割综述[J]. 电子学报, 2019, 47(10): 2211-2220.  
LUO H L, ZHANG Y. A survey of image semantic segmentation based on deep network[J]. Acta Electronica Sinica, 2019, 47(10): 2211-2220.(in Chinese)
- [28] 张庆锐. 复杂场景下的手势分割算法研究[D]. 山东: 山东大学, 2018.  
ZHANG Q R. Research on Hand Gesture Segmentation Algorithm with Complex Background[D]. Shan Dong: Shandong University, 2018. (in Chinese)
- [29] GARCIA-GARCIA A, ORTOS-ESCOLANO S, OPREA S, et al. A review on deep learning techniques applied to semantic segmentation[EB/OL]. (2017-04-22)[2020-12-31]. <https://arxiv.org/abs/1704.06857>.
- [30] KAWULOK M, KAWULOK J, NALEPA J. Spatial-based skin detection using discriminative skin-presence features[J]. Pattern Recognition Letters, 2014, 41: 3-13.
- [31] KAWULOK M, KAWULOK J, SMOLKA B. Discriminative textural features for image and video colorization [J]. IEICE Transactions on Information and Systems, 2012, 95(7): 1722-1730.
- [32] JONES M J, REHG J M. Statistical color models with application to skin detection[J]. International Journal of Computer Vision, 2002, 46(1): 81-96.
- [33] CHENG M M, MITRA N J, HUANG X, et al. Global contrast based salient region detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 37(3): 569-582.
- [34] ZHAO H, QI X, SHEN X, et al. Icnnet for real-time semantic segmentation on high-resolution images[C]//Proceedings of the European Conference on Computer Vision. Munich: Springer, 2018: 405-420.
- [35] TIAN Z, HE T, SHEN C H, et al. Decoders matter for semantic segmentation: Data-dependent decoding enables flexible feature aggregation[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach: IEEE, 2019: 3126-3135.
- [36] HOU Q B, ZHANG L, CHENG M M, et al. Strip pooling: Rethinking spatial pooling for scene parsing[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE, 2020: 4002-4011
- [37] SUN J H, JI T T, ZHANG S B, et al. Research on the hand gesture recognition based on deep learning[C]//2018 12th International Symposium on Antennas, Propagation

- and EM Theory (ISAPE). Hangzhou: IEEE, 2018: 1-4.
- [38] ARENAS J O P, MORENO R J, BELEÑO R D H. Convolutional neural network with a DAG architecture for control of a robotic arm by means of hand gestures[J]. Contemporary Engineering Sciences, 2018, 11(12): 547-557.
- [39] TAN Y S, LIM K M, TEE C, et al. Convolutional neural network with spatial pyramid pooling for hand gesture recognition[J]. Neural Computing and Applications, 2020: 1-13.
- [40] MEHTA S, RASTEGARI M, CASPI A, et al. Espnet: Efficient spatial pyramid of dilated convolutions for semantic segmentation[C]//Proceedings of the European Conference on Computer Vision (ECCV). Munich: Springer, 2018: 552-568.

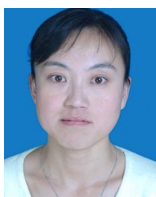
#### 作者简介



**雷 玉** 女,1995年生,山西晋中人.河北大学网络空间安全与计算机学院硕士研究生,研究方向为深度学习、图像分割技术及其应用.



**崔振超(通讯作者)** 男,1983年生,河北邯郸人,讲师.中国计算机学会会员,2007年于燕山大学获得学士学位,2010年于燕山大学获得硕士学位,2015年于哈尔滨工业大学获得博士学位.现为河北大学网络空间安全与计算机学院教师,主要从事人工智能、机器视觉方向研究.  
E-mail: cuizhenchao@gmail.com



**陈丽萍** 女,1974年生,河北保定人,讲师.1997年于河北农业大学获得学士学位,2000年获得硕士学位.现为河北大学网络空间安全与计算机学院教师,主要从事机器视觉方面的研究.



**陈向阳** 女,1977年生,河南三门峡人,讲师.2000年毕业于燕山大学获得学士学位,2007年毕业于河北大学获得硕士学位.现为河北大学网络空间安全与计算机学院教师,研究方向为深度学习.



**王煜骁** 男,1997年生,河北廊坊人.河北大学网络空间安全与计算机学院硕士研究生,研究方向为深度学习、图像分类.