

# 基于 Generalized Region Loss 的代价函数及在图像分割中的应用

张凯, 余义斌

(五邑大学 智能制造学部, 广东 江门 529020)

**摘要:** 针对图像分割中的困难样本, 提出了一种对像素区域细分计算的 Generalized Region Loss 的新的代价函数; 首先通过引入一项参数, 改变了以往代价函数主要通过设置权重或 Focal 等关注困难样本的方法, 其次通过对标签图像和预测图像进行区域划分, 并且对划分四区域的困难样本分类关注, 最后分别计算其四区域绝对损失, 进而进行加权组合; 为验证算法性能, 使用 CamVid 数据集作为实验数据, 该代价函数在 FCN 和 U-Net 两种图像分割网络上得到验证, 同当前图像分割领域常用的 12 种代价函数相比, IoU 指标分别提高 1.93% 和 2.99%, 由此证明此代价函数优于大多数图像分割代价函数; 最终实验结果表明, 提出的基于像素区域细分计算的代价函数能够有效提高图像分割精度, 为图像分割的研究提供借鉴。

**关键词:** 代价函数; 图像分割; 网络收敛; 分割精度; 困难样本

## Cost Function Based on Generalized Region Loss and its Application in Image Segmentation

ZHANG Kai, YU Yibin

(Department of Intelligent Manufacturing, Wuyi University, Jiangmen 529020, China)

**Abstract:** Aiming at difficult samples in image segmentation, a new cost function of generalized region loss is proposed; Firstly, a parameter is introduced to change the previous method of focusing on difficult samples by setting weights or Focal in the cost function. Secondly, label image and prediction image are divided into several regions, and difficult samples divided into four regions are classified and focused. Finally, the absolute losses of four regions are calculated respectively, and then the weighted combination is carried out; In order to verify the performance of the algorithm, the CamVid data set is used as experimental data. The cost function is verified on the FCN and U-Net image segmentation networks. Compared with 12 cost functions commonly used in the current image segmentation field, the IOU index is increased by 1.93% and 2.99% respectively, which proves that this cost function is superior to most image segmentation cost functions; The final experimental results show that the proposed cost function based on pixel segmentation can effectively improve the accuracy of image segmentation and provide a reference for the research of image segmentation.

**Keywords:** cost function; image segmentation; network convergence; segmentation accuracy; difficult sample

## 0 引言

深度学习已经彻底改变了从软件到制造业的各个行业, FCN<sup>[1]</sup>奠定了语义分割深度模型的先驱, 医学界也从深度学习中受益。在疾病分割方面有多种创新, 使用 U-Net<sup>[2]</sup> 网络模型进行肿瘤分割, SegNet<sup>[3]</sup> 网络模型进行癌症检测等。图像分割是深度学习社区对医学领域的重要贡献之一, 除了说明某些疾病的存在外, 它还显示了疾病确切存在的位置, 它极大地帮助创建了在各种类型的医学扫描中检测肿瘤、病变等的算法。在设计基于复杂图像分割的深度学习体系结构时, 有许多可变的参数, 包括代价函数、优化器、学习率等, 都会影响模型最终的结果。而代价函数对深度学习中神经网络的训练有非常重要的作用, 对于不同的分

割任务使用合适的代价函数可以提高最终的分割精度, 最终得到很好的分割效果。

本文从代价函数的角度出发, 针对解决分割中困难样本的问题, 国内外学者提出了各种各样的代价函数。然而, 在处理图像分割任务中困难样本的问题时, 没有一种损失能够始终达到最佳性能。本文通过研究各类型代价函数的特性, 总结背后的一般规律, 最终构建一种通用关注困难样本的代价函数, 例如脑 CT、胰腺、肾脏等图像语义分割问题。

## 1 图像分割代价函数的分类

### 1.1 图像分割代价函数按目的分类

近几年, 随着深度学习技术的快速发展, 图像语义分

收稿日期: 2022-08-02; 修回日期: 2022-08-31。

作者简介: 张凯(1997-), 男, 山西运城人, 硕士, 主要从事先进信息处理与应用技术方向的研究。

余义斌(1966-), 男, 湖北北京山人, 博士, 教授, 主要从事现代信号处理与模式识别方向的研究。

引用格式: 张凯, 余义斌. 基于 Generalized Region Loss 的代价函数及在图像分割中的应用[J]. 计算机测量与控制, 2023, 31(3): 215-222.

割算法的性能有明显的提升。在图像分割方面,国内外学者提出了许多新的技术致力于提升其分割精度,主要包括对数据集的处理、网络模型的修改、代价函数的改进以及优化器的调整等方面。针对代价函数,国内外学者提出各种各样不同的代价函数用于提升分割效果,从最初的 CE Loss, 到后来的 WCE Loss<sup>[2]</sup>、Sensitivity Specificity (SS) Loss<sup>[5]</sup>、Topk Loss<sup>[6]</sup>、Dice Loss<sup>[7]</sup>、IoU Loss<sup>[8]</sup>、Tversky Loss<sup>[9]</sup>、Lovasz Loss<sup>[10]</sup>、Focal Loss<sup>[11]</sup>、Boundary (BD) Loss<sup>[12]</sup>、Hausdorff Distance (HD) Loss<sup>[13]</sup>、Focal Tveraky Loss<sup>[15]</sup>、Asymmetric Similarity Loss<sup>[16]</sup>、DPCE Loss<sup>[17]</sup>、Generalized Dice Loss<sup>[18]</sup>、Penalty Loss<sup>[19]</sup>等,都是为了解决图像语义分割中某些问题而提出,主要包括类不平衡问题、关注分割中困难样本、关注分割边界等,以及一些复合代价函数 Combo Loss<sup>[20]</sup>、Exponential Logarithmic Loss (ELL)<sup>[21]</sup>等都是为了更好地解决这些问题。

对于分割中类不平衡的问题来说旨在通过提升不频繁标签的重要性来缓解损失偏差,主要有两种常见的方法,其一是通过对标签频繁下采样来重新平衡类别的先验分布,然而这种方法限制了训练图像的信息;其二是通过类别与标签频率成反比的方法来设置权重,例如 WCE Loss<sup>[2]</sup>、Generalized Dice Loss<sup>[18]</sup>等,尽管该方法对某些分割中的不平衡问题有效,但对于不常见的那些像素区域使用交叉熵计算有时会加入一些噪声,并且使用大权重对其加权会放大该错误从而导致训练不稳定。

对于分割中困难样本来说旨在提升损失对困难样本的关注度,同样主要有 3 种常见的方法,其一是对单个像素进行加权,通过数字一减去预测像素概率值的方法加权损失,例如 Focal Loss<sup>[11]</sup>等;其二是将计算好的损失整体加权,例如 Focal Tveraky Loss<sup>[15]</sup>等;其三是使用权重控制 FP (False Positive) 和 FN (False Negative) 之间的平衡,例如 Tveraky Loss<sup>[9]</sup>、Asymmetric Similarity Loss<sup>[16]</sup>、Penalty Loss<sup>[19]</sup>等。

对于分割边界来说旨在提升损失对图像边界的关注度,例如 DPCE Loss<sup>[17]</sup>、Boundary (BD) Loss<sup>[12]</sup>、Hausdorff Distance (HD) Loss<sup>[13]</sup>等,但本文认为图像标签多为人工标注,标注的标签边界区域或多或少都存在一些误差,如果对标签真值边界的关注度过分大于对标签真值内部的关注度,对于计算分辨率比较低的图像是不适合的。

## 1.2 图像分割代价函数按类型分类

图像分割的代价函数<sup>[4]</sup>伴随着深度学习的发展,逐渐演化为 4 种类型,即基于分布、基于区域、基于边界和基于复合的代价函数。

基于分布的代价函数旨在最小化两个分布之间的差异,包括 CE Loss、Topk Loss<sup>[6]</sup>、Focal Loss<sup>[11]</sup>、DPCE Loss<sup>[17]</sup>等,该类型代价函数是以交叉熵为基准进行变化,只关注标签图真值内部的损失,而对于真值之外的损失计算没有涉及。交叉熵的计算公式如式 (1) 所示:

$$L_{CE} = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C g_i^c \log s_i^c \quad (1)$$

基于区域的代价函数旨在最小化不匹配或最大化真实值和预测分割之间的重叠区域,包括 Sensitivity Specificity Loss<sup>[5]</sup>、Dice Loss<sup>[7]</sup>、IoU Loss<sup>[8]</sup>、Tversky Loss<sup>[9]</sup>、Lovasz Loss<sup>[10]</sup>、Focal Tveraky Loss<sup>[11]</sup>、Asymmetric Similarity Loss<sup>[16]</sup>、Generalized Dice Loss<sup>[18]</sup>、Penalty Loss<sup>[19]</sup>等,该类型代价函数都是以 Dice Loss 为基准进行变化,虽然对标签图真值内外的损失都进行了计算,但计算过于笼统,针对标签真值内部以及外部关于分类正确或错误的像素区域都没有进行精确区分。Dice Loss 的计算公式如式 (2) 所示:

$$L_{Dice} = 1 - \frac{2 \sum_{i=1}^N \sum_{c=1}^C g_i^c s_i^c}{\sum_{i=1}^N \sum_{c=1}^C s_i^c + \sum_{i=1}^N \sum_{c=1}^C g_i^c} \quad (2)$$

基于边界的代价函数<sup>[14]</sup>是一种相对较新的代价函数,旨在最小化真实值与预测分割之间的距离,包括 Boundary (BD) Loss<sup>[12]</sup>、Hausdorff Distance (HD) Loss<sup>[13]</sup>等,该类型代价函数主要是通过距离图加权计算损失值,BD Loss 使用边界匹配度去监督网络的损失,而 HD Loss 使用真实值和分割的距离变换图加权损失,两者都容易导致损失值偏向距离边界近的区域,而对于标签真值内部的区域给予很少的关注。

基于复合的代价函数是基于分布、区域和边界的代价函数之间的加权组合,包括 Combo Loss<sup>[20]</sup>、Exponential Logarithmic Loss (ELL)<sup>[21]</sup>等,该类型代价函数主要是为了解决单个代价函数不能很好的处理分割任务的问题,通常需要至少两种及两种以上的代价函数进行某种变换加权得到,但该方法在实际应用中变得非常困难,需要用户必须对各类型代价函数都有精准的估计,然后从中选择适合处理某种任务的代价函数。

为了解决以上的这些问题,本文提出了一种新的代价函数 Generalized Region Loss,该代价函数也是一种基于区域计算损失的代价函数,同以往基于区域的代价函数不同的是 Generalized Region Loss 致力于将图像损失区域进行精确地分类计算,使计算的损失结果更加符合该区域的实际损失值,更加有利于激发算法的学习过程,使计算结果更容易收敛。

## 2 Generalized Region Loss 代价函数算法研究

### 2.1 Generalized Region Loss (GR) 函数

本文提出了一种基于像素区域细分计算的 Generalized Region Loss 代价函数用于图像分割,其表达式如式 (3) 所示:

$$GR \text{ Loss} = \beta \cdot (l_1 + l_3) + (1 - \beta) \cdot (l_2 + l_4) \quad (3)$$

该代价函数主要是计算以下两部分区域的损失值:真值内部区域和真值外部区域,具体如图 1 所示。真值内部区域主要包括预测和真值重合区域  $l_1$ 、预测和真值非重合区

域  $l_3$ ; 真值外部区域主要包括预测和背景重合区域  $l_2$ 、预测和背景非重合区域  $l_4$ 。在式 (3) 中, 参数  $\beta \in [0.01, 0.1]$  控制代价函数第一项和第二项之间的平衡, 在本函数中更加关注那些分类错误区域的损失, 所以, 此  $\beta$  参数设置参照 Sensitivity Specificity Loss<sup>[3]</sup> 而定, 具体将其置为 0.02,  $l_1$ 、 $l_2$ 、 $l_3$ 、 $l_4$  详细的计算方式在后边给出。

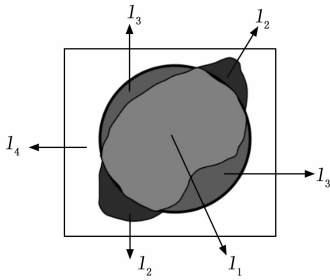


图 1 真值内部区域和真值外部区域

其一针对预测和真值重合区域来说, 即图 1 当中的  $l_1$  区域, 在该区域中一般其预测概率已经达到了很大的值, 说明该区域预测概率已经基本达到预测标准, 但该部分仍存在损失, 需要通过公式  $\log$  或者标签和预测差值的方法得到其损失。在此公式中使用的是  $\log$  计算方法, 因为在实验中发现使用此计算方法稍微优于另外一种, 分析原因是  $\log$  相比于另外一种计算方式对于那些困难样本的惩罚更加严重。该区域损失的计算如式 (4) 所示:

$$l_1 = - \frac{\sum_{c=1}^C \sum_{i=1}^N (g_i^c \cdot s_{\max\_i}^c) \cdot \log s_i^c \cdot (1 - s_i^c)^\alpha}{\sum_{c=1}^C \sum_{i=1}^N (g_i^c \cdot s_{\max\_i}^c) + \epsilon} \quad (4)$$

在式 (4) 中, 如果标签  $c$  是像素  $i$  的正确分类, 则  $g_i^c$  是二值指标;  $s_i^c$  是对应的预测概率;  $s_{\max\_i}^c$  是对  $s_i^c$  在分类维度上求最大值的位置得到的结果;  $\alpha$  为惩罚参数;  $\epsilon$  为防止分母为零而设置的极小的数字。( $g_i^c \cdot s_{\max\_i}^c$ ) 是为了定位这部分区域的元素;  $\log s_i^c$  就是常见的损失计算;  $(1 - s_i^c)^\alpha$  代表了像素注意力机制, 当  $s_i^c$  接近 1 时, 系数  $(1 - s_i^c)^\alpha$  会变得非常小, 同样当  $\alpha$  取得越大时,  $(1 - s_i^c)^\alpha$  会进一步变小, 即简单样本对损失结果的影响非常小。反过来说, 即困难样本在损失结果中所占比重越大, 代价函数于更注重困难上, 所以说这也相当于一种注意力机制。

其二针对预测和背景重合区域来说, 即图 1 当中的  $l_2$  区域, 在该区域中一般其预测概率达到了很大的值, 说明该区域预测概率同预测标准差异较大, 所以本文希望该区域概率变得很小, 更加接近 0, 其单个像素点的预测概率值就是其单个像素点的损失。该区域损失的计算如式 (5) 所示:

$$l_2 = \frac{\sum_{i=1}^N \sum_{c=1}^C [(s_{\max\_i}^c - g_i^c) \cdot s_{\max\_i}^c] \cdot s_i^c \cdot (s_i^c)^\alpha}{\sum_{c=1}^C \sum_{i=1}^N [(s_{\max\_i}^c - g_i^c) \cdot s_{\max\_i}^c] + \epsilon} =$$

$$\frac{\sum_{i=1}^N \sum_{c=1}^C [(s_{\max\_i}^c - g_i^c) \cdot s_{\max\_i}^c] \cdot (s_i^c)^{\alpha+1}}{\sum_{c=1}^C \sum_{i=1}^N [(s_{\max\_i}^c - g_i^c) \cdot s_{\max\_i}^c] + \epsilon} \quad (5)$$

在式 (5) 中,  $[(s_{\max\_i}^c - g_i^c) \cdot s_{\max\_i}^c]$  是为了定位这部分区域的元素,  $s_i^c$  是单个像素损失值, 至于在公式后边又乘以  $(s_i^c)^\alpha$ , 同式 (4) 中的  $(1 - s_i^c)^\alpha$  是一样的原理, 目的是使该区域代价函数的计算更加关注该区域的困难样本。

其三针对预测和真值非重合区域来说, 即图 1 当中的  $l_3$  区域, 在该区域中一般其预测概率非常小, 说明该区域预测概率同预测标准差异较大, 所以本文希望该区域概率变得较大一些, 更加接近 1, 需要通过公式  $\log$  或者标签和预测差值的方法得到单个像素损失值, 具体计算本文是使用了  $\log$  计算方法, 同式 (4) 中是一致的。该区域损失的计算如式 (6) 所示:

$$l_3 = - \frac{\sum_{c=1}^C \sum_{i=1}^N [(g_i^c - s_{\max\_i}^c) \cdot g_i^c] \cdot \log s_i^c \cdot (1 - s_i^c)^\alpha}{\sum_{c=1}^C \sum_{i=1}^N [(g_i^c - s_{\max\_i}^c) \cdot g_i^c] + \epsilon} \quad (6)$$

在式 (6) 中,  $[(g_i^c - s_{\max\_i}^c) \cdot g_i^c]$  是为了定位这部分区域元素,  $(1 - s_i^c)^\alpha$  同式 (4) 中是一致的。

其四针对预测和背景非重合区域来说, 即图 1 当中的  $l_4$  区域, 在该区域中一般其预测概率非常小, 说明该区域预测概率同预测标准差异较小, 但本文希望该区域概率变得更小, 更加接近 0, 其单个像素点的预测概率值就是其单个像素点的损失。该区域损失的计算如式 (7) 所示:

$$l_4 = \frac{\sum_{c=1}^C \sum_{i=1}^N \{1 - [g_i^c + (s_{\max\_i}^c - g_i^c) \cdot s_{\max\_i}^c]\} \cdot s_i^c \cdot (s_i^c)^\alpha}{\sum_{c=1}^C \sum_{i=1}^N \{1 - [g_i^c + (s_{\max\_i}^c - g_i^c) \cdot s_{\max\_i}^c]\} + \epsilon} = \frac{\sum_{c=1}^C \sum_{i=1}^N \{1 - [g_i^c + (s_{\max\_i}^c - g_i^c) \cdot s_{\max\_i}^c]\} \cdot (s_i^c)^{\alpha+1}}{\sum_{c=1}^C \sum_{i=1}^N \{1 - [g_i^c + (s_{\max\_i}^c - g_i^c) \cdot s_{\max\_i}^c]\} + \epsilon} \quad (7)$$

在式 (7) 中,  $\{1 - [g_i^c + (s_{\max\_i}^c - g_i^c) \cdot s_{\max\_i}^c]\}$  是为了定位这部分区域的元素,  $(s_i^c)^\alpha$  同式 (5) 中是一致的。

## 2.2 GR Loss 函数中采用 $\log s_i^c$ 和 $(g_i^c - s_i^c)$ 对比

GR Loss 代价函数在公式中采用了  $\log s_i^c$  的计算方式, 因为通过实验表明, 使用  $\log s_i^c$  的计算方式得到的分割结果略微优于  $(g_i^c - s_i^c)$  的分割效果。从图 2 当中的 Mean IoU 以及图 3 当中的 Mean Class Accuracy 的对比实验都可以看出  $\log s_i^c$  的分割数据明显优于  $(g_i^c - s_i^c)$ , 本文分析原因是  $\log s_i^c$  相比于  $(g_i^c - s_i^c)$  对于图像分割中的那些困难样本的惩罚更加严重。

## 2.3 GR Loss 函数中参数 $\alpha$ 的确定

通过对 Generalized Region Loss 代价函数的参数  $\alpha$  进行测试, 从图 4 当中的对比实验可以明显看出, 当  $\alpha = 2, 3, 4$  时, 该三条曲线增长趋势基本一致; 从图 5 的对比实验可以看出, 当  $\alpha = 2, 3, 4$  时, 该三条曲线增长方式同样基本一致。

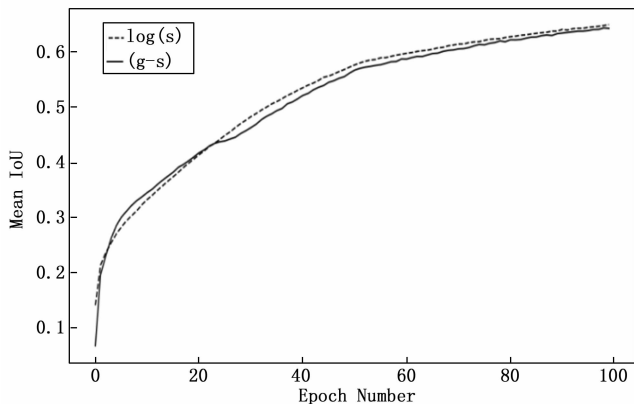


图 2 使用 FCN 网络结构训练时, GR Loss 代价函数使用  $\log s_i^c$  和  $(g_i^c - s_i^c)$  对 Mean IoU 的影响

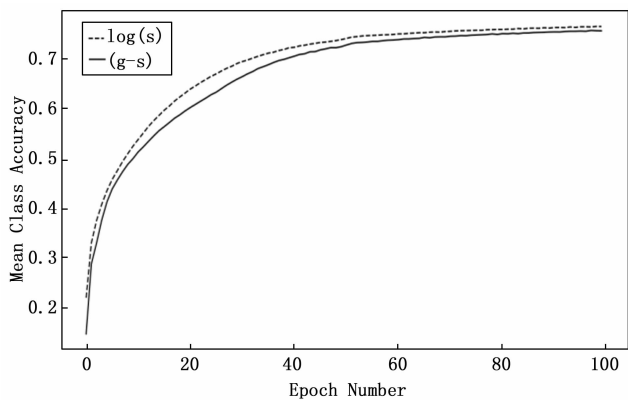


图 3 使用 FCN 网络结构训练时, GR Loss 代价函数使用  $\log s_i^c$  和  $(g_i^c - s_i^c)$  对 Mean Class Accuracy 的影响

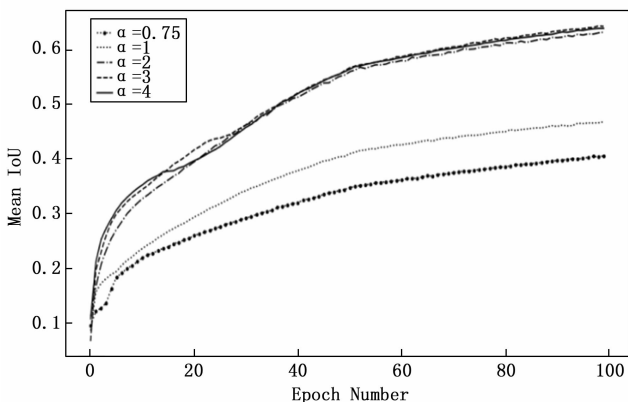


图 4 使用 FCN 网络结构训练时, GR Loss 代价函数  $\alpha$  置为不同参数对 Mean IoU 的影响

仔细观察图 5 可以看到,  $\alpha = 4$  的曲线会略微高于  $\alpha = 3$  的曲线, 但  $\alpha = 4$  时公式的计算复杂度也同样高于  $\alpha = 3$  的计算复杂度。当  $\alpha = 0.75, 1$  时, 从图 4 和图 5 来看, 该两条曲线的增长趋势过于缓慢, 在网络训练过程中会消耗比较多的时间复杂度和空间复杂度, 因此在本文不作考虑。

所以综上所述当  $\alpha = 3$  时, Mean IoU 以及 Mean Class

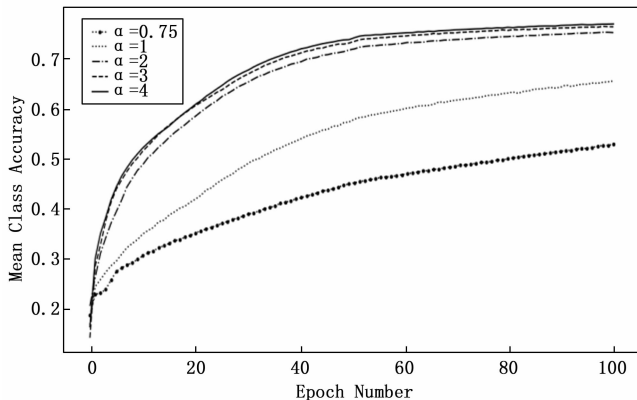


图 5 使用 FCN 网络结构训练时, GR Loss 代价函数  $\alpha$  置为不同参数对 Mean Class Accuracy 的影响

Accuracy 效果提升都是比较明显的, 因此在后续实验中 Generalized Region Loss 代价函数中的  $\alpha$  采用此参数。

### 3 代价函数在图像分割中的应用

#### 3.1 数据集描述

CamVid<sup>[22]</sup>数据集是一种由剑桥大学公开发布的城市道路场景的数据集, 同时它也是第一个具有目标类别语义标签的视频集合。该数据集包含了 700 多张精准标注的图片用于强监督学习, 可分为训练集、验证集和测试集。同时, 在 CamVid 数据集中通常使用 11 种常用的类别来进行分割精度的评估。分别为: 道路 (Road)、交通标志 (Symbol)、汽车 (Car)、天空 (Sky)、行人道 (Sidewalk)、电线杆 (Pole)、围墙 (Fence)、行人 (Pedestrian)、建筑物 (Building)、自行车 (Bicyclist)、树木 (Tree)。

#### 3.2 评估指标

为了评估分割精度, 本文使用了在图像分割中 4 个常用的评估指标, 主要包括有 Dice similarity coefficient (DSC), Intersection over Union (IoU), 精确率 (Precision) 和召回率 (Recall)。DSC、IoU、Precision 和 Recall 分别根据等式 (8) ~ (11) 定义:

$$DSC = \frac{2TP}{2TP + FP + FN} \quad (8)$$

$$IoU = \frac{TP}{TP + FP + FN} \quad (9)$$

$$Precision = \frac{TP}{TP + FP} \quad (10)$$

$$Recall = \frac{TP}{TP + FN} \quad (11)$$

#### 3.3 数据集处理

网络部署训练过程中, 在 FCN 网络框架下, 本文将 CamVid 数据集原图和标签统一中心裁剪为  $[3 \times 256 \times 480]$  大小的图像, 为的是可以一定程度上减少网络训练时间; 并且采用了  $[[0.485, 0.456, 0.406], [0.229, 0.224, 0.225]]$  的数据归一化处理, 从而消除奇异样本产生异常数据导致的不良影响; 优化器为 Adam 且初始学习率 lr 置为  $10^{-6}$ , 并且学习率伴随着训练次数的增加逐渐减小为之

前的 0.5 倍, 减小周期为 50 轮。而在 U-Net 网络框架下, 为适应网络模型结构的输入和输出, 本文将 CamVid 数据集原图统一中心裁剪为  $[3 \times 572 \times 572]$  大小的图像, 标签中心裁剪为  $[3 \times 388 \times 388]$  大小的图像, 其它参数像数据归一化、优化器、学习率等同 FCN 训练时保持一致。

本文实验环境为 Windows10 操作系统, 编程语言为 Python3.6, 深度学习框架为 PyTorch, 硬件环境 CPU 为 2.90 GHz Intel (R) Xeon (R) W-2102 (8 G 内存), GPU 为 NVIDIA Tesla T4 (12 G 显存)。

## 4 实验结果与分析

### 4.1 使用 FCN 网络结构

在图 6 和图 7 当中使用 FCN 网络结构进行测试实验, 通过对用于图像分割的主流 12 种代价函数进行测试, 包括 CE Loss、Wce Loss、Topk Loss、Focal Loss、DPCE、dice Loss、Lovasz Loss、Generalized Dice Loss、Tversky loss、Focal Tversky Loss、Asymmetric Similarity Loss、Sensitivity Specificity Loss。因测试曲线过多, 在图中不易区分, 我们只将表现最好的代价函数曲线同其它代价函数曲线进行区分制线。

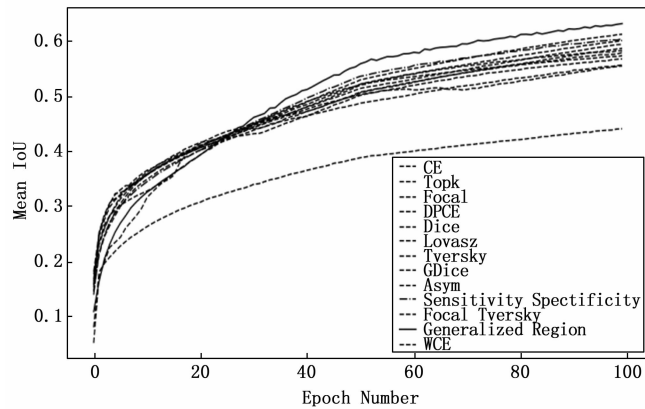


图 6 使用 FCN 网络结构训练时, 改变其代价函数, 将其它 12 种代价函数同 GR Loss 代价函数进行对比, 对比它们之间的 Mean IoU

观察实验结果, 从图 6 当中的 Mean IoU 来看, 虽然 Mean IoU 在 30 轮之前使用 Generalized Region Loss 代价函数提升精度的趋势有缓慢, 但在训练 30 轮之后提升速度的趋势明显优于其它代价函数, 并且 Generalized Region Loss 在训练 80 轮左右达到的分割准确率在使用其它代价函数时至少需要 100 轮训练才能达到, 促进了网络的收敛, 大大节省了训练时间。

从图 7 当中的 Mean Class Accuracy 来看, 新提出的代价函数同 Sensitivity Specificity Loss 在 15 轮和 65 轮左右都出现了交点, 在 15 轮之前, Sensitivity Specificity Loss 提升速度略微高于 Generalized Region Loss; 15 轮到 65 轮之间, Sensitivity Specificity Loss 提升速度却又低于 Generalized Region Loss; 65 轮之后 Sensitivity Specificity Loss 提升速

度却又略微高于 Generalized Region Loss, 可以看出新的代价函数与其也是不分伯仲。

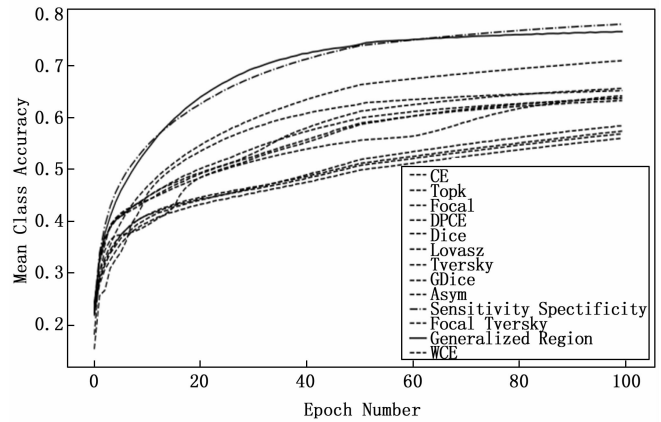


图 7 使用 FCN 网络结构训练时, 改变其代价函数, 将其它 12 种代价函数同 GR Loss 代价函数进行对比, 对比它们之间的 Mean Class Accuracy

### 4.2 使用 U-Net 网络结构

在图 8 和图 9 当中使用 U-Net 网络进行测试实验, 由于图 6 和图 7 已经对 12 种代价函数在 FCN 网络上做了测试, 为了避免资源浪费, 所以图 8 和图 9 是以图 6 和图 7 为依据, 从 12 种代价函数中挑选出 4 个表现良好的, 明显可以促进网络收敛的代价函数进行测试, 主要包括有 CE Loss、Topk Loss、Focal Loss、Sensitivity Specificity Loss。

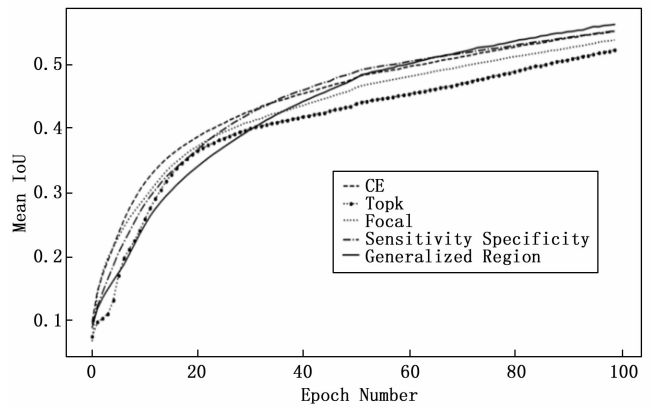


图 8 使用 U-Net 网络结构训练时, 改变其代价函数, 将其它 4 种代价函数同 GR Loss 代价函数进行对比, 对比它们之间的 Mean IoU

观察实验结果, 从图 8 当中的 Mean IoU 和图 9 当中的 Mean Class Accuracy 来看, 将其它 4 种代价函数同新改进的 Generalized Region Loss 进行对比, 实验结果表明, 虽然图 8 当中的 Mean IoU 指标在 70 轮之前使用新的代价函数提升趋势有些缓慢, 但训练 70 轮之后提升趋势明显优于其它 4 种代价函数, 以及新的代价函数在训练 90 轮左右达到的分割精度在使用其它代价函数时至少需要 100 轮训练才能达到, 同样促进了网络的收敛, 大大节省了训练时间。

表 1 CamVid 数据集上,使用 FCN+5 种代价函数对应的 11 种分类精度,黑体代表每种类别精度的最大值

| Loss Function      | Sky            | Building       | Pole         | Road           | Sidewalk       | Tree           | SignSymbol     | Fence          | Car            | Pedestrian     | Bicyclist      |
|--------------------|----------------|----------------|--------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| CE                 | 0.963 1        | 0.931 3        | 0.032 6      | 0.985 3        | 0.783 6        | 0.782 8        | 0.329          | 0.285 2        | 0.837 2        | 0.199 4        | 0.111 2        |
| Topk               | 0.963 1        | <b>0.931 9</b> | 0.037 1      | <b>0.985 8</b> | 0.788 4        | <b>0.784 1</b> | 0.333 3        | 0.296 1        | 0.833 8        | 0.215 1        | 0.132 5        |
| Focal              | <b>0.963 4</b> | 0.931 2        | 0.062 3      | 0.984 9        | 0.793 5        | 0.783 5        | 0.376 7        | 0.302          | 0.840 4        | 0.243 7        | 0.130 6        |
| SS                 | 0.942 8        | 0.880 7        | <b>0.593</b> | 0.968 8        | 0.876 9        | 0.768 3        | <b>0.867 6</b> | <b>0.436 8</b> | <b>0.918 8</b> | <b>0.787 7</b> | <b>0.533 8</b> |
| Generalized Region | 0.949 8        | 0.877 2        | 0.530 2      | 0.973 1        | <b>0.891 0</b> | 0.746 4        | 0.824 9        | <b>0.438 4</b> | <b>0.922 6</b> | 0.759 5        | 0.503 1        |

表 2 CamVid 数据集上,使用 U-Net+5 种代价函数对应的 11 种分类精度,黑体代表每种类别精度的最大值

| Loss Function      | Sky            | Building       | Pole         | Road           | Sidewalk       | Tree           | SignSymbol     | Fence          | Car            | Pedestrian     | Bicyclist      |
|--------------------|----------------|----------------|--------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| CE                 | 0.977 5        | 0.955 6        | 0.147 7      | 0.988 4        | 0.888 5        | <b>0.822 6</b> | 0.616 6        | 0.351 9        | 0.918 5        | 0.399 8        | 0.054 2        |
| Topk               | 0.982 8        | 0.812 9        | 0.174 9      | 0.988 3        | 0.898 2        | 0.816 1        | 0.598 3        | 0.312 3        | 0.913          | 0.418 7        | 0.008          |
| Focal              | 0.978 9        | <b>0.960 0</b> | 0.119 1      | <b>0.989 4</b> | 0.884 4        | 0.815 9        | 0.498          | 0.259 7        | 0.910 4        | 0.273 2        | 0.186 0        |
| SS                 | 0.943 4        | 0.872          | <b>0.877</b> | 0.961 3        | 0.912 2        | 0.782 2        | <b>0.891 9</b> | 0.440 1        | 0.942 6        | <b>0.822 5</b> | <b>0.508 0</b> |
| Generalized Region | <b>0.985 6</b> | 0.825 9        | 0.719 6      | 0.984 8        | <b>0.921 3</b> | 0.698 1        | 0.861 2        | <b>0.447 9</b> | <b>0.946 0</b> | 0.806 6        | 0.472 0        |

从图 9 当中的 Mean Class Accuracy 来看,新的代价函数的提升趋势会略微低于 Sensitivity Specificity Loss,但从总体来看,Generalized Region Loss 函数也有其优异的地方。

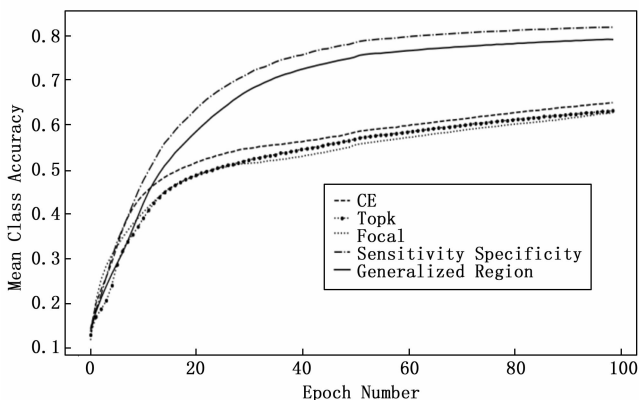


图 9 使用 U-Net 网络结构训练时,改变其代价函数,将其它 4 种代价函数同 GR Loss 代价函数进行对比,对比它们之间的 Mean Class Accuracy

### 4.3 使用 FCN 和 U-Net 时, CamVid 数据集的 11 种类别精度对比分析

通过在 FCN 和 U-Net 两种网络框架上,将表现良好的 4 种代价函数 CE Loss、Topk Loss、Focal Loss、Sensitivity Specificity Loss 以及新的代价函数 Generalized Region Loss 进行对比分析,从表 1 和表 2 可以看出,Generalized Region Loss 代价函数对 Sidewalk、Fence、Car 分割精度明显优于其他代价函数,尤其针对 Sidewalk 和 Fence 两类分割小目标来说分割精度明显提高。

### 4.4 使用 FCN 和 U-Net 时,四项指标对比分析

从表 3 和表 4 来看,使用 Generalized Region Loss 代价函数在 DSC、IoU、Recall 几项指标上表现得比较优异,明显优于其它 4 种代价函数,而在 Precision 这项指标上会略微逊色于其它代价函数,可见 Generalized Region Loss 代价函数对误报损失 FP (False Positives) 不太敏感。

表 3 使用 5 种代价函数在 FCN 上训练的结果,黑体数字表示每个指标的最高值和次高值

| Loss function | DSC            | IoU            | Precision      | Recall         |
|---------------|----------------|----------------|----------------|----------------|
| CE            | 0.573 1        | 0.593 4        | 0.336 8        | 0.305 4        |
| Focal         | 0.593 2        | <b>0.611 7</b> | <b>0.346 4</b> | 0.316 0        |
| SS            | <b>0.637 0</b> | 0.603 2        | 0.336 0        | <b>0.408 1</b> |
| Topk          | 0.599 0        | 0.600 3        | <b>0.348 1</b> | 0.319 1        |
| GR Loss       | <b>0.653 1</b> | <b>0.631 0</b> | 0.345 5        | <b>0.403 8</b> |

表 4 使用 5 种代价函数在 U-Net 上训练的结果,黑体数字表示每个指标的最高值和次高值

| Loss function | DSC            | IoU            | Precision      | Recall          |
|---------------|----------------|----------------|----------------|-----------------|
| CE            | 0.654 6        | 0.552 7        | <b>0.353 8</b> | 0.365 3         |
| Focal         | 0.653 3        | 0.538 7        | <b>0.357 7</b> | 0.357 8         |
| SS            | <b>0.677 2</b> | <b>0.553 0</b> | 0.351 8        | <b>0.423 0</b>  |
| Topk          | 0.641 0        | 0.523 5        | 0.351 4        | 0.356 9         |
| GR Loss       | <b>0.657 2</b> | <b>0.582 9</b> | 0.347 9        | <b>0.4117 4</b> |

### 4.5 不同代价函数的分割结果

本文使用了 4 种代价函数对 CamVid 数据集进行了实际的测验分割,具体分割结果如图 10 所示。本文随机挑选了 CamVid 数据集预测图中的 4 张不同道路场景图进行结果分析,横向来看,第一列是道路真实场景图;第二列是道路真实标签图;第三列是将代价函数置为 CE Loss 后的分割掩码图像,可以看到该类型代价函数在道路 (Road) 以及行人 (Pedestrian) 上的分割效果明显不如其它 3 种;第四张是将代价函数置为 Focal Loss 后的分割掩码图像,可以看到该类型代价函数在道路 (Road) 以及行人 (Pedestrian) 上的分割效果略微优于使用 CE Loss 的分割效果,分割轮廓略微准确;第五张是将代价函数置为 Sensitivity Specificity Loss 后的分割掩码图像,可以看到该类型代价函数在交通标志 (Symbol) 上的分割效果明显不如其它 3 种;而第六张是将代价函数置为 Generalized Region Loss 后的分

割掩码图像, 可以看到该类型代价函数不论是在在道路 (Road)、行人 (Pedestrian) 或是交通标志 (Symbol) 上的分割效果明显优于其它 3 种代价函数, 分割轮廓相对来说比较清晰, 在 CamVid 数据集上的整体分割效果已经基本达到了令人满意的程度。

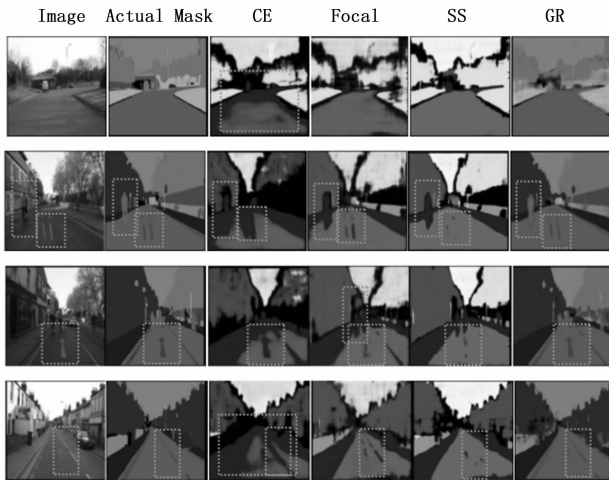


图 10 使用 CE、Focal、SS 以及 GR 在 CamVid 数据集上训练 150 轮后的实际分割结果图

## 5 结束语

本文提出了一种聚焦于解决图像分割中困难样本的代价函数, 主要采用了两种策略: ①通过在损失计算中引入  $s_{\max\_i}^c$  参数, 可以更好地定位损失区域; ②对预测图像和标签图像分为四区域细分计算其损失, 使损失对困难样本的度量更加精细化。实验结果表明, 本文提出的代价函数始终提高了分割质量, 具有更好的召回精度平衡, 与现有的用于图像分割的代价函数相比取得了最优的结果。但由于 Generalized Region Loss 函数对误报率敏感度不高并且增加了算法的复杂度, 后续工作将从简化损失复杂度和网络结构的角度出发进一步提升图像的分割精度。

### 参考文献:

- [1] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation [C] // Proceedings of the IEEE conference on computer vision and pattern recognition, 2015: 3431 - 3440.
- [2] 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. Springer, Cham, 2015: 234 - 241.
- [3] BADRINARAYANAN V, KENDALL A, CIPOLLA R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation [J]. IEEE transactions on pattern analysis and machine intelligence, 2017, 39 (12): 2481 - 2495.
- [4] JADON S. A survey of loss functions for semantic segmentation [C] // 2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB). IEEE,

2020: 1 - 7.

- [5] HASHEMI S R, SALEHI S S M, ERDOGMUS D, et al. Asymmetric loss functions and deep densely-connected networks for highly-imbalanced medical image segmentation: Application to multiple sclerosis lesion detection [J]. IEEE Access, 2018, 7: 1721 - 1735.
- [6] UHRIG J, CORDTS M, FRANKE U, et al. Pixel-level encoding and depth layering for instance-level semantic labeling [C] // German conference on pattern recognition. Springer, Cham, 2016: 14 - 25.
- [7] DROZDZAL M, VORONTOV E, CHARTRAND G, et al. The importance of skip connections in biomedical image segmentation [M] // Deep learning and data labeling for medical applications. Springer, Cham, 2016: 179 - 187.
- [8] RAHMAN M A, WANG Y. Optimizing intersection-over-union in deep neural networks for image segmentation [C] // International symposium on visual computing. Springer, Cham, 2016: 234 - 244.
- [9] SALEHI S S M, ERDOGMUS D, GHOLIPOUR A. Tversky loss function for image segmentation using 3D fully convolutional deep networks [C] // International workshop on machine learning in medical imaging. Springer, Cham, 2017: 379 - 387.
- [10] BERMAN M, TRIKI A R, BLASCHKO M B. The lovasz-softmax loss: A tractable surrogate for the optimization of the intersection-over-union measure in neural networks [C] // Proceedings of the IEEE conference on computer vision and pattern recognition, 2018: 4413 - 4421.
- [11] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [C] // Proceedings of the IEEE International Conference on Computer Vision, 2017: 2980 - 2988.
- [12] KERVADEC H, BOUCHTIBA J, DESROSIERS C, et al. Boundary loss for highly unbalanced segmentation [C] // International Conference on Medical Imaging with Deep Learning, PMLR, 2019: 285 - 296.
- [13] KARIMI D, SALCUDEAN S E. Reducing the hausdorff distance in medical image segmentation with convolutional neural networks [J]. IEEE Transactions on Medical Imaging, 2019, 39 (2): 499 - 513.
- [14] WANG S, YU L, LI K, et al. Boundary and entropy-driven adversarial learning for fundus image segmentation [C] // International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2019: 102 - 110.
- [15] ABRAHAM N, KHAN N M. A novel focal tversky loss function with improved attention u-net for lesion segmentation [C] // 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019). IEEE, 2019: 683 - 687.
- [16] HASHEMI S R, SALEHI S S M, ERDOGMUS D, et al. Asymmetric loss functions and deep densely-connected networks for highly imbalanced medical image segmentation: Application to multiple sclerosis lesion detection [J]. IEEE Access, 2018, 7: 1721 - 1735.

- [17] FEI L, XU Y, FANG X, et al. Low rank representation with adaptive distance penalty for semi-supervised subspace classification [J]. *Pattern Recognition*, 2017, 67: 252 - 262.
- [18] SUDRE C H, LI W, VERCAUTEREN T, et al. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations [M] // *Deep learning in medical image analysis and multimodal learning for clinical decision support*, Springer, Cham, 2017: 240 - 248.
- [19] YANG S, KWEON J, KIM Y H. Major vessel segmentation on x-ray coronary angiography using deep networks with a novel penalty loss function [C] // *International Conference on Medical Imaging with Deep Learning Extended Abstract Track*, 2019.
- [20] TAGHANAKI S A, ZHENG Y, ZHOU S K, et al. Combo loss: Handling input and output imbalance in multi-organ segmentation [J]. *Computerized Medical Imaging and Graphics*, 2019, 75: 24 - 33.
- [21] WONG K C L, MORADI M, TANG H, et al. 3D segmentation with exponential logarithmic loss for highly unbalanced object sizes [C] // *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, Cham, 2018: 612 - 619.
- [22] BROSTOW G J, SHOTTON J, FAUQUEUR J, et al. Segmentation and recognition using structure from motion point clouds [C] // *European conference on computer vision*, Springer, Berlin, Heidelberg, 2008: 44 - 57.
- [9] 苗少师. 导引头伺服机构谐振频率与去耦精度分析 [D]. 长沙: 国防科学技术大学, 2008.
- [10] 赵 丽. 某型号雷达导引头测试控制系统的设计 [D]. 西安: 西安交通大学, 2008.
- [11] 李正大, 陈光达, 马洪波. 结构因素对雷达伺服系统性能影响及其测试研究 [J]. *现代电子技术*, 2008, 31 (3): 33 - 36.
- [12] 程 鹏. 自动控制原理 [M]. 北京: 高等教育出版社, 2003.
- [13] 王海金. 基于模糊 PID 的雷达导引头控制系统研究 [D]. 北京: 中国航天科工集团第二研究院, 2007.
- [14] 李连升. 雷达伺服系统 [M]. 北京: 国防工业出版社, 1993.
- [15] QINGXIN H, HUI L. DSP control system of brushless DC motor without position sensor [C] // *2005 International Conference on Electrical Machines and Systems*. IEEE, 2005, 1: 371 - 375.
- [16] STAJIC D, PERIC N, DEUR J. Friction compensation methods in position and speed control systems [C] // *ISIE'99. Proceedings of the IEEE International Symposium on Industrial Electronics (Cat. No. 99TH8465)*. IEEE, 1999, 3: 1261 - 1266.
- [17] 章卫国, 杨向忠. 模糊控制理论与应用 [M]. 西安: 西北工业大学出版, 1999.
- [18] 毕永利, 刘 洵, 葛文奇, 等. 机载多框架陀螺稳定平台速度稳定环设计 [J]. *光电工程*, 2004, 31 (2): 16 - 18.
- [19] 肖英奎, 尚 涛, 陈殿生. 伺服系统实用技术 [M]. 北京: 化学工业出版社, 2004 (2): 60 - 80.
- [20] 夏福梯. 防空导弹制导雷达伺服系统 [M]. 北京: 宇航出版社, 1996.
- [21] 石辛民, 郝整清. 模糊控制及其 MATLAB 仿真 [M]. 北京: 清华大学出版社, 北京交通大学出版社, 2008.
- [22] 赵 超. 导引头稳定系统隔离度研究 [J]. *电光与控制*, 2008, 15 (7): 78 - 82.
- [23] DUPONT P E, DUNLAP E P. Friction modeling and control in boundary lubrication [C] // *1993 American Control Conference*. IEEE, 1993: 1910 - 1915.
- [24] PROCYK T J, MAMDANI E H. A linguistic self-organizing process controller [J]. *Automatica*, 1979, 15 (1): 15 - 30.
- [25] 吴 晔. 导引头天线机电伺服系统 [J]. *制导与引信*, 2001, 22 (4): 4 - 10.

(上接第 170 页)

由仿真结果可以看出, 模糊控制相比常规 PID 控制不受被控对象变化而影响, 具有较强的鲁棒性, 对于伺服系统实际工程应用有较大的帮助。

## 5 结束语

本文通过对伺服系统的设计及仿真, 首先建立起了满足实际工程需求的导引头伺服系统, 考虑到非线性摩擦因素对实际系统的影响, 利用力矩平衡的原理, 通过测量实际转台的数据建立起系统的摩擦模型, 并在此模型基础上进行 Matlab/Simulink 仿真后, 对比仿真曲线, 得出在考虑摩擦环节对伺服系统的影响时, 在阶跃信号引导时, 采用传统 PID 控制算法不能很好地跟踪输入信号, 模糊控制策略可以降低系统超调量, 响应时间也符合系统设计的要求; 在正弦信号引导时, 传统 PID 控制算法的结果中位置跟踪存在“平顶”现象, 波形失真严重, 而模糊控制策略可以有效地消除“平顶”等现象, 可以达到高精度跟踪的要求, 在可控范围内变化时, 模糊控制的效果相比传统 PID 更优越, 更适合实际工程应用。

## 参考文献:

- [1] 秦永元. 惯性导航 [M]. 北京: 科学出版社, 2006.
- [2] 姬 伟. 陀螺稳定光电跟踪平台伺服控制系统研究 [D]. 南京: 东南大学, 2006.
- [3] 周 涛, 朱景成. 机载光电跟踪平台伺服系统自抗扰控制 [J]. *光电工程*, 2011, 38 (4): 31 - 36.
- [4] FANSON J L, CAUGHEY T K. Positive position feedback control for large space structures [J]. *AIAA journal*, 1990, 28 (4): 717 - 724.
- [5] GARNELL P. *Guided Weapon Control Systems* [M]. Beijing Institute of Technology, 2004: 49 - 61.
- [6] RICHARD C. DORF, ROBERT H. Bishop. *Modern Control Systems* [M]. Pearson Education, 2002: 557 - 585.
- [7] 吴 鹏. 精确制导光电导引头稳定平台数字控制技术 [J]. *弹箭与制导学报*, 2002, 22 (3): 37 - 41.
- [8] ARMSTRONG-HELOUVRY B. Stick slip and control in low-speed motion [J]. *IEEE Transactions on Automatic Control*,