



CV Image Segmentation Model Combining with Local and Global Features of the Target

Xiaohui Li and Xili Wang

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

June 19, 2020

结合目标局部和全局特征的 CV 遥感图像分割模型

摘要 随着遥感卫星技术的发展, 高分辨率遥感影像不断涌现。从含有较多信息、背景复杂的遥感影像中自动提取目标成为一个亟待解决的难题。传统的图像分割方法主要依赖图像光谱、纹理等底层特征, 容易受到图像中遮挡和阴影等的干扰。为此, 针对特定的目标类型, 提出结合目标局部和全局特征的 CV(Chan Vest)遥感图像目标分割模型, 首先, 采用深度学习生成模型--卷积受限玻尔兹曼机建模表征目标全局形状特征, 以及重建目标形状; 其次, 利用 Canny 算子提取目标边缘信息, 经过符号距离变换得到综合了局部边缘和全局形状信息的约束项; 最终, 以 CV 模型为图像目标分割模型, 增加新的约束项得到结合目标局部和全局特征的 CV(Chan Vest)遥感图像分割模型。在遥感小数据集 Levir-oil drum、Levir-ship 和 Levir-airplane 上的实验结果表明: 所提模型不仅可以克服 CV 模型对噪声敏感的缺点, 且在训练数据有限、目标尺寸较小、遮挡及背景复杂的情况下依然能完整、精确地分割出目标。

关键词 图像分割; 形状先验; 卷积受限玻尔兹曼机; 深度学习; CV 模型

CV Image Segmentation Model Combining with Local and Global Features of the Target

Abstract With the development of remote sensing satellite technology, high-resolution remote sensing images continue to emerge. Automatic target extraction from remote sensing images with more information and complex background has become an urgent problem to be solved. The traditional image segmentation method mainly depends on the image spectrum, texture and other underlying features, in image segmentation tasks, shadow, occlusion and various interference in the image increase the difficulty of segmentation, and lead to unsatisfied results. For this reason, according to the specific target type, a CV (ChanVest) image segmentation model combining with local and global features of the target is proposed. Firstly, the deep learning generation model-CRBM(Convolution Restricted Boltzmann Machine), is used to represent the global shape features of the target and to reconstruct the shape of the target. Secondly, the edge information of the target is extracted by Canny operator, and a new shape constraint term which combines the local edge and global shape information is obtained by symbolic distance transformation is proposed. Finally, the CV model is used as the image target segmentation model, and new constraints are added to obtain the CV (ChanVest) remote sensing image segmentation model which combines the local and global features of the target. The experimental results on remote sensing dataset Levir-oil drum, Levir-ship and Levir-airplane show that the proposed model can not only overcome the noise sensitivity of CV model, but also segment the target completely and accurately in the case of limited training data, small target size, occlusion and complex background.

Key words image segmentation; shape prior; convolutional restricted boltzmann machine; deep learning; CV model

收稿日期: ; 修回日期:

基金项目: 国家自然科学基金(41471280, 61701290, 61701289)

Supported by: National Natural Science Foundation of China(41471280, 61701290, 61701289)

作者简介: 李晓慧 (1992-), 女, 硕士研究生, 主要研究方向为机器学习、图像处理。E-mail: lixiaohuihl@163.com.

导师简介: 汪西莉 (1969-), 女, 博士, 教授, 博士生导师, 主要研究领域为智能信息处理、模式识别、图像处理。E-mail: wangxili@snnu.edu.cn.

1 引言

随着遥感技术的发展, 高分辨率的遥感影像不断涌现, 从影像中自动提取目标引起了众多学者的广泛研究。与普通光学图像相比, 遥感图像的场景、目标分布等都更为复杂, 针对遥感影像复杂背景下的小目标分割^[1-2]更为困难。在采用分割提取目标的方法中, 基于水平集的 CV^[3]图像分割方法最大的优势在其处理拓扑变化的能力。但面对复杂的遥感图像时, 仅依靠图像颜色、纹理等底层信息, 当图像中存在目标部分被遮挡、目标与背景杂糅等情况时, 往往得不到正确的分割结果。因此, 在基于水平集的 CV 模型中引入与目标相关的形状先验信息, 有利于辅助模型分割出完整的目标。但如何准确且灵活地表示目标形状是困难的。就形状建模的方法来说, 主要分为三种: (1)基于映射的统计形状建模方法^[4-6], 即通过映射得到形状的特征向量表达形状特征空间, 如主成分分析和核主成分分析。然而对于复杂的形状, 这类方法不能灵活地表达。(2)基于标记点统计特征的形状建模方法^[7-9], 即通过标记目标轮廓中的关键点, 对其统计建模来表达全局形状特征。这些标记点都是手工标记的, 通过统计的方法建模形状, 建模过程较为复杂且不适合复杂形状。(3)基于深度学习的形状建模方法^[10-12], 用深度生成式模型从大量的目标形状样本中学习其结构并建模表示形状, 模型不仅可以自动提取训练集的目标形状特征, 而且可以灵活地表示形状及生成形状。

考虑到在采用深度学习模型获取目标形状的过程中, 图像归一化的操作可能造成目标形状细节特征的丢失, 本文在重建的目标形状基础上, 采用 Canny^[13]算子从原图中提取目标边缘信息, 结合局部边缘和全局形状构造目标形状约束项, 提出了结合目标局部和全局特征的 CV 图像分割模型。具体来说: 采用深度学习的卷积受限玻尔兹曼机模型 (Convolutional Restricted Boltzmann Machine, CRBM)^[14]来提取目标全局形状特征, 并结合 Canny 算子提取的原图像的边缘信息, 经过符号距离变换得到含有边缘约束的目标形状, 以此为先验信息引入到 CV 模型中指导曲线演化, 得到正确的分割结果。模型在训练数据有限, 背景复杂, 形态各异, 尺度变化较大且

成像面积小于 2000 像素的遥感影像目标提取中取得了理想的分割结果。

2 CV 模型

CV 模型是由 Chan 和 Vest 提出的图像分割模型。该模型将图像分割表达为能量函数最小化问题, 用 I 表示待分割的灰度图像, $I(x, y)$ 的值表示 I 中任意像素点 (x, y) 的灰度值。演化曲线 C 将 I 划分成两个部分, C_1 、 C_2 分别表示演化曲线 C 内部和外部的灰度信息的均值。 $L(C)$ 和 $A(inC)$ 的值分别表示演化曲线 C 的长度以及演化曲线内部的区域面积。 λ_1 、 λ_2 、 μ 、 ν 是各项的系数。得到 CV 模型的能量泛函为:

$$E^{CV} = \lambda_1 \int_{in} |I(x, y) - C_1|^2 dx dy + \lambda_2 \int_{out} |I(x, y) - C_2|^2 dx dy + \mu L(C) + \nu A(inC) \quad (1)$$

在式(1)中, 前两项使得演化曲线 C 不断向目标轮廓靠近, 合称为保真项。第三项用于约束演化曲线 C , 确保获得的曲线足够短, 称为长度约束项。第四项用于约束演化曲线, 避免曲线震荡, 称为面积项。只有当保真项的值为零, 即演化曲线在目标边界上时能量函数才能达到最小, 实现对图的分割。

将上式中演化曲线 C 表示为水平集函数 $\phi(x, y)$, 令 $\phi(x, y)$ 表示图像 I 中任意点 (x, y) 与演化曲线 C 的最小欧式距离值。且设定当 $\phi(x, y)$ 在图像的内部时水平集函数 $\phi(x, y) > 0$, 当 $\phi(x, y)$ 恰好在目标边界上时, $\phi(x, y) = 0$ (这一步骤即是初始化水平集), 否则水平集函数 $\phi(x, y) < 0$ 。于是公式(1)改写为:

$$E^{CV}(C_1, C_2, \phi) = \lambda_1 \int_{\Omega} |\mu_0(x, y) - C_1|^2 H(\phi(x, y)) dx dy + \lambda_2 \int_{\Omega} |\mu_0(x, y) - C_2|^2 (1 - H(\phi(x, y))) dx dy + \mu \int_{\Omega} \delta(\phi(x, y)) |\nabla \phi(x, y)| dx dy + \nu \int_{\Omega} H(\phi(x, y)) dx dy \quad (2)$$

其中海氏(Heaviside)函数 $H(z)$ 和狄拉克(Dirac)函数 $\delta(z)$ 分别为:

$$H(z) = \begin{cases} 1 & \text{if } z \geq 0 \\ 0 & \text{if } z < 0 \end{cases}, \delta(z) = \frac{d}{dz} H(z) \quad (3)$$

根据变分原理, 使用欧拉-拉格朗日(Euler-Lagrange)方法推导出水平集函数的演化方程为:

$$\begin{cases} \frac{\partial \phi}{\partial \alpha} = \delta(\phi) \left\{ \mu f - v - \lambda_1 [\mu_6(x, y) - C_1]^2 + \lambda_2 [\mu_6(x, y) - C_2]^2 \right\} \\ \phi(0, x, y) = \phi(x, y) \\ C_1 = \frac{\int_{\Omega} I(x, y) H[\phi(x, y)] dx dy}{\int_{\Omega} H[\phi(x, y)] dx dy} \\ C_2 = \frac{\int_{\Omega} I(x, y) \{1 - H[\phi(x, y)]\} dx dy}{\int_{\Omega} \{1 - H[\phi(x, y)]\} dx dy} \end{cases} \quad (4)$$

其中 f 表示演化曲线的曲率，通过迭代水平集函数的演化方程使得 CV 模型的能量函数值最小，得到分割结果。

3 深度生成形状模型

3.1 卷积受限玻尔兹曼机

为了在 CV 分割模型中引入有效的目标形状先验知识，本文采用两层的卷积受限玻尔兹曼机深度学习模型来建模和获取目标的全局形状特征。该模型是在 RBM 的基础上，引入卷积神经网络(Convolutional Neural Network, CNN)^[15] 中“卷积权值共享”思想衍生的生成模型。如图 1 所示，CRBM 模型是包含一个输入层 v 和一个隐层 h 的两层模型，由于 CRBM 模型中卷积操作增加了局部感受野和权值共享的特点，且 CRBM 的输入是二维的，使得模型能够较好地获取输入数据的二维空间结构信息和图像的局部信息。

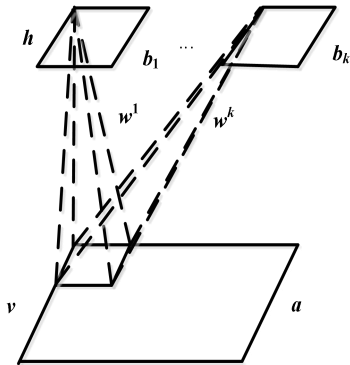


图 1 CRBM 模型结构图

Fig.1 The structure of CRBM model

在 CRBM 模型中用 $\theta = \{w^k, a, b_k\}$ 表示模型参数， K 表示卷积核个数， w^k 表示与隐层 h 相连的第 k 个卷积核的值， a 表示输入层 v 的偏置， b_k 表示隐层 h 中第 k 个特征图的偏置。且令 v_{ij} 表示输入层 v 中节点 (i, j) 的单元的状态， h_{ij}^k 表示与隐层 h 相连的第 k 个特征图中节点 (i, j) 的

单元状态。 \bullet 表示矩阵的乘积， \sim 表示将矩阵按顺时针方向旋转 180 度， N_v 表示输入层大小， N_h 表示隐层 h 中特征图的大小， N_w 表示卷积核大小， $*$ 表示卷积操作。该模型所有的节点的状态均为二值变量，且模型中每个输入层节点只和隐层的节点相关，和其他输入层的节点是独立的，即输入层节点的状态只受隐层节点的影响，反之隐层节点也是，只受输入层节点的影响。通过定义可得该模型对应的能量函数：

$$E(v, h; \theta) = -\sum_{k=1}^K h^k \bullet (\tilde{w}^k * v) - \sum_{k=1}^K b_k \sum_{i,j=1}^{N_h} h_{ij}^k - a \sum_{i,j=1}^{N_v} v_{ij} \quad (5)$$

故此，该模型的联合概率分布为：

$$\begin{cases} P(v, h; \theta) = \frac{1}{Z} \exp[-E(v, h; \theta)] \\ Z(\theta) = \sum_{v, h} e^{-E(v, h; \theta)} \end{cases} \quad (6)$$

输入层单元 v_i 和隐层单元 h_{ij}^k 的激活概率分别表示为：

$$P(v_{ij} = 1 | h) = g\left(\left(\sum_k w^k * h^k\right)_{ij} + a\right) \quad (7)$$

$$P(h_{ij}^k = 1 | v) = g\left(\left(\sum_k \tilde{w}^k * v\right)_{ij} + b_k\right) \quad (8)$$

3.2 卷积受限玻尔兹曼机模型训练

模型结构确定后，利用训练样本训练模型，其目的是确定模型的参数。本文训练 CRBM 模型时以二值图像作为输入，0 表示背景，1 表示目标，即模型中所有的节点状态均为二值变量。

CRBM 模型采用对比散度算法(Contrast Divergence, CD)^[16] 训练模型，以训练集中的样本作为模型的初始状态，并将训练集中的样本记为 $x = \{x_0, x_1, \dots, x_m\}$ ，模型初始值记为 v_0 ，用训练样本初始化模型可视层单元的状态，即令 $v_0 = x_0$ 。首先是根据公式(8)确定隐层单元的状态，即由 $P(h | v)$ 求得 h ；其次根据公式(7)重构输入层单元的状态，即由 $P(v | h)$ 求得 v 。不断重复上述 k 次，得到 x_0 对应的重构结果 v_k 。使用梯度下降法更新 CRBM 模型参数 θ ，故此，基于 CD 算法的参数 θ 的梯度更新公式为：

$$\begin{aligned} CD_k(\theta, v_0) = & -\sum_n P(h | v_0) \frac{\partial E(v_0, h)}{\partial \theta} \\ & + \sum_n P(h | v_k) \frac{\partial E(v_k, h)}{\partial \theta} \end{aligned} \quad (9)$$

其中 $k = 1$ ，该算法的伪代码表 1 所示。

表 1 基于 CD—1 算法的 CRBM 训练算法

Tab.1 A training algorithm for CRBM based CD-1

输入：训练样本 $x = \{x_0, x_1, \dots, x_m\}$ 卷积核的个数 k ，卷积核的大小 $N_w \times N_w$ 。输入层大小 N_v ，隐层中特征图的大小 N_h

输出：权重矩阵 w^k ，输入层的偏置 a ，隐层的偏置 b_k

初始化：令输入层的单元状态 $v_0 = x_0$ ， w^k 、 a 和 b_k 的初始值随机初始化。

```

for i=1:k
  for  $\forall i, j \in N_h$ 
    计算  $P(h_{ij}^k = 1 | v_0)$ ，即  $P(h_{ij}^k = 1 | v) = g((\sum_k \tilde{w}^k * v)_{ij} + b_k)$ 

    从条件分布  $P(h_{ij}^k | v_0)$  中采样  $y_{ij}^k \in \{0, 1\}$ 
  end for
end for
 $h_1 = y$ 
for i=1:k
  for  $\forall i, j \in N_v$ 
    计算  $P(v_{ij} = 1 | h_1)$ ，即  $P(v_{ij} = 1 | h) = g((\sum_k w^k * h^k)_{ij} + a)$ 

    从条件分布  $P(v_{ij} | h_1)$  中采样  $x_{ij}^k \in \{0, 1\}$ 
  end for
end for
 $v_1 = x$ 
for i=1:k
  for  $\forall i, j \in N_h$ 
    计算  $P(h_{ij}^k = 1 | v_1)$ ，即  $P(h_{ij}^k = 1 | v) = g((\sum_k \tilde{w}^k * v)_{ij} + b_k)$ 

    从条件分布  $P(h_{ij}^k | v_1)$  中采样  $y_{ij}^k \in \{0, 1\}$ 
  end for
end for
 $h_2 = y$ 
更新参数：
 $a = a + g \sum (v_1 - v_0)$ 
 $b^k = b^k + g \sum (h_2 - h_1)$ 
 $w^k = w^k + g(\tilde{h}_2 * v_1 - \tilde{h}_1 * v_0)$ 

```

3.3 卷积受限玻尔兹曼机模型生成形状

在确定模型参数之后，采用多步吉布斯采样的^[7]方法从模型中生成形状。其核心思想是通过模拟的联合概率分布直接推导出条件分布，多次反复直到采样得到的样本与模型所定义的概率分布的样本非常接近。CRBM 模型的采样过程如图 2 所示，为了生成近似服从 CRBM 模型所定义的样本，本文实验采用二值图像初始化模型输入层的单元状态，并执行 Gibbs 采样得到输入层单元的状态(即生成样本)。由于 CRBM 的输入是二

维的，使得模型能够更好地获取输入样本的二维空间结构信息和图像的局部信息，从而使得模型的生成形状更加接近真实的形状数据。

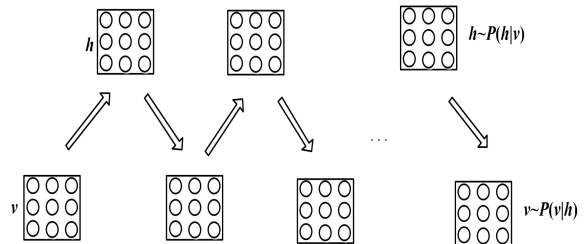


图 2 CRBM 模型采样过程图
Fig.2 The sampling process of CRBM.

4 结合目标局部和全局特征的 CV 遥感图像分割模型(CLG-CV)

4.1 边缘信息

边缘信息作为图像最基本特征之一，是图像分割依赖的重要特征。Canny 算子用一阶导数的极大值表示边缘，可以很好的描述图像目标的外部形状和空间结构的轮廓特性。如图 4-2 是 Levir 数据集^[18]中 Levir-oil drum001、Levir-ship001、Levir-airplane001 提取边缘信息的距离变换示例，图 3(a)表示原始图像，图 3(b)表示原始图像边缘提取结果图，将边缘提取结果图 3(b)距离变换得到表示边缘形状的距离函数图图 3(c)，即令边缘像素点的值为 1，其它像素取其最近边缘像素点之间的欧式距离 $\min dist(i, j)$ 。由此边缘距离函数 ϕ^B 可定义为：

$$\phi^B = \begin{cases} 1 & \text{if } i \in \text{edge} \\ \min dist(i, j) & \text{if } i \notin \text{edge}, j \in \text{edge} \end{cases} \quad (10)$$

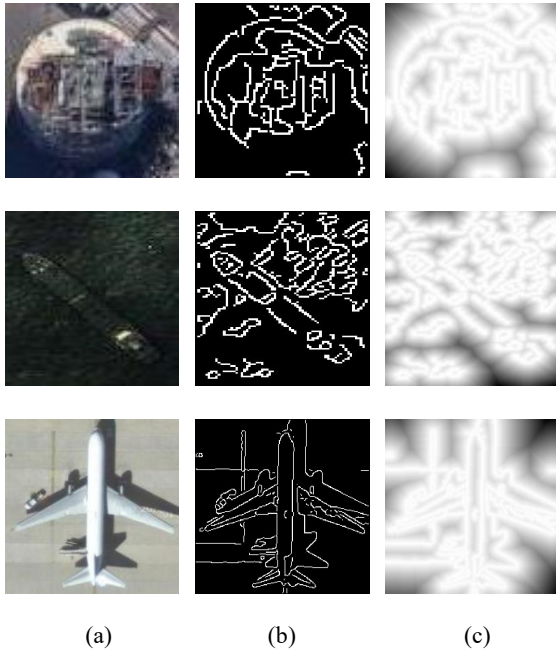


图 3 距离变换示例图，(a)原始图像，(b)边缘提取图，(c)边缘距离函数图。

Fig.3 Example diagram of distance transformation, (a) Original image, (b)Result of edge extract, (c)Edge distance function.

4.2 形状约束项

为了融合深度学习模型中的生成形状和图像边缘信息，本文将深度学习模型中的生成形状如图 4(a)，用距离函数表示模型生成形状的结果图 4(b)，使得形状边界处像素点的值为 1，其它像素取其最近边界点之间的欧式距离

$\min dist(i, j) < 1$ 。由此生成形状距离函数 ϕ^S 可定义为：

$$\phi^S = \begin{cases} 1 & \text{if } i \in \text{boundary} \\ \min dist(i, j) & \text{if } i \notin \text{boundary}, j \in \text{boundary} \end{cases} \quad (11)$$

将生成形状距离函数 ϕ^S 与边缘距离函数 ϕ^B 作点乘运算，即将图 4(b)以及图 3(c)这两幅图中对应位置像素的欧式距离值做乘运算，将运算结果作为输出图像相应位置像素的欧式距离值，最终得到结合边缘和深度学习的目标形状图 4(c)。该方法不仅有效的去除了由 Canny 算子提取的边缘信息中虚假目标边界，而且也改善了深度学习生成模型-CRBM 生成形状时，目标边缘细节丢失的情况。并将结合边缘和深度学习的目标形状表示为符号函数(特殊的水平集函数)^[19]如图 4(d)，即设定边界上像素点的值为 0，其他像素取其最近边界上像素点之间的最小欧式距离。为方便描述，用符号距离函数 ψ^{BS} 表示结合边缘和深度学习的目标形状，用水平集距离函数 ϕ 表示演化曲线，通过定义得到新的形状约束项：

$$E^{BS} = \int_{\Omega} |H(\phi) - H(\psi^{BS})|^2 dx dy \quad (12)$$

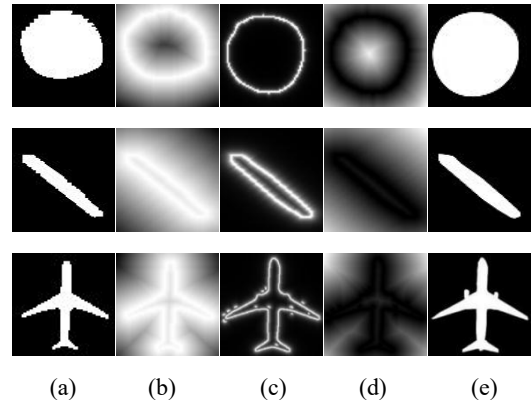


图 4 距离变换示例图，(a)生成形状，(b)生成形状距离函数，(c)结合边缘和深度学习的目标形状，(d)结合边缘和深度学习的目标形状的符号距离函数，(e)Ground-truth 图。

Fig.4 Example diagram of distance transformation, (a)Generated shape, (b)Generated shapedistance function, (c) Target shape based on edge and deep learning, (d) Sign distance function of target shape based on edge and deep learning, (e)Ground-truth.

4.3 CLG-CV 模型

本文所提的 CLG-CV 模型在传统的 CV 模型的能量函数中做了以下改进：首先舍弃了面积项对演化曲线的约束。其次融合边缘和卷积受限玻尔兹曼机的目标形状信息，定义新的形状约束项 E^{BS} ，对演化曲线加以约束。最后结合 CV 模型

的能量项 E^{CV} 和形状约束能量项 E^{BS} 得到新的能量泛函 E 。故 CLG-CV 模型的能量函数为：

$$\begin{aligned} E &= E^{CV} + \lambda E^{BS} \\ &= E^{CV}(C_1, C_2, \phi) + \lambda E^{BS}(\phi, \psi^{BS}) \\ &\quad + \lambda_2 \int_{\Omega} |\mu_0(x, y) - C_2|^2 \{1 - H_e[\phi(x, y)]\} dx dy \quad (13) \\ &\quad + \mu \int_{\Omega} \delta_e[\phi(x, y)] |\nabla \phi(x, y)| dx dy \\ &\quad + \lambda \int_{\Omega} |H(\phi) - H(\psi^{BS})|^2 dx dy \end{aligned}$$

其中 λ 是形状先验项系数。通过对公式(11)对应的欧拉-拉格朗日方程进行求解, 可得到如下演化方程:

$$\frac{\partial \phi}{\partial t} = \delta_e(\phi) \left\{ \mu \operatorname{div} \frac{\nabla \phi}{|\nabla \phi|} - \lambda_1 (I - C_1)^2 + \lambda_2 (I - C_2)^2 \right. \\ \left. + [H(\phi) - H(\psi^{BS})] \right\} \quad (14)$$

最后通过迭代水平集函数的演化方程使得 CV 模型的能量函数值达到最小, 曲线停止演化, 得到最终的分割结果。E^{CV} 没有介绍

4.4 分割算法步骤

本文主要针对少于 2000 像素的小目标遥感图像来提取目标, 其分割任务更具有挑战性。所提模型通过施加包含全局和局部的目标特征的约束, 使分割得到的目标在完整性和边缘细节保留方面都体现得更好, 在复杂背景下也能得到正确的分割结果。分割算法步骤如下:

- 1.粗分割: 给定待分割灰度图像 P , 建立 CV 模型对应的能量函数 $E^{CV}(\phi)$, 将其最小化, 得到粗分割结果 P^* 。
- 2.训练 CRBM 模型: 把一批已归一化至 32×32 大小的 ground truth 图作为训练集, 送入到 CRBM 模型中训练。
- 3.生成形状: 将 P^* 归一化至 32×32 大小, 并将其作为模型的输入, 经过采样得到重建结果 S , 将 S 扩大至待分割图像 P 大小。由式(11)计算生成形状距离函数 ϕ^S 。
- 4.边缘提取: 用 Canny 算子提取原图像 P 的边缘信息, 并进行距离变换, 由式(10)计算边缘距离函数 ϕ^B 。
- 5.定义形状约束项: 将步骤(3)中得到的 ϕ^S 与步骤(4)中得到的 ϕ^B 做点乘运算得到融合边缘信息和深度学习的目标形状, 并将其表示成符号距离函数 ψ^{BS} , 由式(12)得到形状约束项 E^{BS} 。
- 6.CLG-CV 图像分割: 结合形状约束项 E^{BS} 和能量项 E^{CV} , 由式(13)得到新的能量函数 E , 求解对

应欧拉-拉格朗日方程, 使能量函数达到最小值, 得到最终的分割结果。

5 数据集、实验参数配置、对比方法和评价指标

5.1 数据集

Levir 数据集来源于 Google Earth, 包含 RGB 图像、标记图像(含有检测框的图像)两类图像, 每副彩色图像的尺寸均为 800×600 , 空间分辨率 0.2m。该数据集中的目标为油桶、飞机、船舰三类, 目标面积小于 2000 像素。从标记图像中随机选择 500 张图像作为实验对象。

采用三阶段方法制作新的小目标遥感数据集。第一阶段, 将所有标记图像按照标记信息裁剪图像, 得到仅含一个特定类别的图像 2485 副。其中油桶 651 幅, 船舰 492 幅, 飞机 1342 幅。第二阶段将一阶段中所有图像, 按类别制作作为 Levir-oil drum、Levir-ship 和 Levir-airplane 三个数据集。第三阶段, 从 Levir-oil drum 数据集中选取 600 幅图像作为训练集, 51 幅图像作为测试集。从 Levir-ship 数据集中选取 400 幅图像作为训练集, 92 幅图像作为测试集。从 Levir-airplane 数据集中选取 1000 幅图像作为训练集, 342 幅飞机图像作为测试集。数据集中的部分图像如图 5。

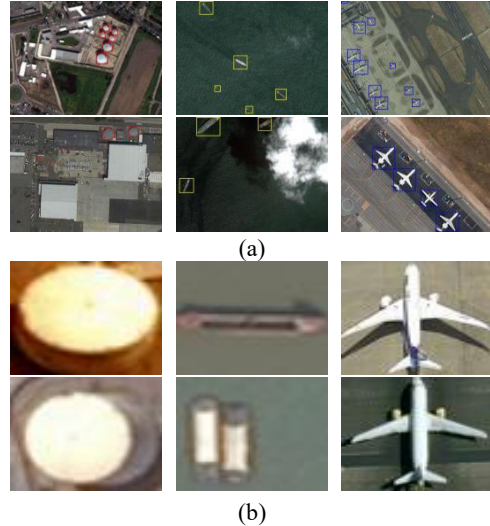


图 5 Levir 数据集, (a)被标记图像, (b)预处理后的图像。
Fig.5 Levir data set, (a)Marked images, (b)Preprocessed images.

表 2 列出了 Levir-oil drum、Levir-ship 和 Levir-airplane 这三个数据集中的目标尺寸分析, 表明该数据集在图像分割方面具有较大的挑战。

表 2 Levir 数据集图像中目标尺寸分析表
Tab.2 Analysis of targets in LEVIR base on scale

Scale (pixel)	Scale <1000	Scale 1000-1500	Scale 1500-2000	Total -
Levir-oil drum	156	342	153	651
Levir-ship	271	96	125	492
Levir-airplane	692	386	264	1342

5.2 实验参数配置

电脑配置为 Intel(R)Xeon(R) CPU E5-2690, 2.6GHz, 256GB RAM, 实验环境为 Win10 系统下安装的 Matlab R2016a。CV 模型分割实验中, 使用数据集中的灰度图像。设定初始演化曲线是通过位于图像中心的矩形表示, 步长 $\Delta t = 0.1$, λ_1 与 λ_2 取值均为 1, 迭代次数为 100。在 Canny 算子提取边缘实验中, 直接调用 Matlab 中的 Canny 函数, 不需单独为不同的数据集设置参数。

参考文献[20]并结合实际设置模型参数。针对 Levir 数据集中的油桶类, CRBM 的参数设置如下: 卷积核的大小为 3×3 , 卷积核个数为 20, 学习率为 0.01, 迭代次数为 1500。针对 Levir 数据集中的船舰类, CRBM 的参数设置如下: 卷积核的大小为 4×4 , 卷积核个数为 20, 学习率为 0.01, 迭代次数为 2000。针对 Levir 数据集中的飞机类, CRBM 的参数设置如下: 卷积核的大小为 3×3 , 卷积核个数为 20, 学习率为 0.005, 迭代次数为 3000。

5.3 对比方法

为了验证本文 CLG-CV 模型的分割性能, 分别与 CV、CG-CV 两种分割模型进行实验对比。

这三种分割模型, 就分割原理来说, CV 模型的分割原理最简单, 该方法是通过统计中心区域与局部邻域灰度特性对比度值, 寻找遥感影像中的目标区域信息, 然后在通过空间关系剔除疑似的噪声点, 使得演化曲线向着目标轮廓不断逼近, 最终得到正确的分割结果。优点是对灰度均匀的图像能自动提取出完整的目标, 缺点是针对复杂的遥感影像中目标背景区与目标区杂糅在一起的情况, 很难取得较好的分割结果。

CG-CV 分割模型, 是以传统 CV 模型为基础的衍生模型, 首先, 采用深度学习生成模型--卷积受限玻尔兹曼机建模表征目标全局形状特征, 以及重建目标形状。其次, 利用重建目标形状, 做符号距离变换, 得到结合目标全局形状信息的约束项。最终, 以 CV 模型为图像目标分割模型,

增加新的约束项得到结合目标全局特征的 CG-CV 遥感图像分割模型。该模型是在传统的 CV 模型的能量函数中做了以下改进: 首先舍弃了面积项对演化曲线的约束。其次通过卷积受限玻尔兹曼机生成形状, 并定义新的形状约束项 E^S 。最后结合 CV 模型的能量项 E^{CV} 和形状约束能量项 E^S 得到新的能量泛函 E 。由此新的 CG-CV 模型的能量函数公式为:

$$\begin{aligned}
 E &= E^{CV} + \lambda E^S \\
 &= E^{CV}(C_1, C_2, \phi) + \lambda E^S(\phi, \phi^S) \\
 &= \lambda_1 \int_{\Omega} |\mu_0(x, y) - C_1|^2 H(\phi(x, y)) dx dy \\
 &\quad + \lambda_2 \int_{\Omega} |\mu_0(x, y) - C_2|^2 (1 - H(\phi(x, y))) dx dy \\
 &\quad + \mu \int_{\Omega} |\delta(\phi(x, y))| |\nabla \phi(x, y)| dx dy \\
 &\quad + \lambda \int_{\Omega} |H(\phi) - H(\phi^S)|^2 dx dy
 \end{aligned} \tag{15}$$

优点是克服了传统 CV 分割模型对噪声敏感的缺点, 缺点是在采用深度学习模型 CRBM 获取目标形状的过程中, 图像归一化的操作造成目标形状细节特征的丢失。

5.4 评价指标

本文采用全局精度 (Global acc) 和交并比 (IOU) 两种评价标准来定量评价分割结果的正确性和完整性, 定义如下:

$$Global \ acc = \frac{TP + TN}{TP + TN + FN + FP} \tag{16}$$

$$IOU = \frac{TP}{TP + FN + FP} \tag{17}$$

TP : true positive 目标正确分类的像素数目

TN : true negative 背景正确分类的像素数目

FN : false negative 目标分为背景的像素数目

FP : false positive 背景分为目标的像素数目

6 实验结果及分析

6.1 Levir 数据集生成形状实验

分别使用 Levir-oil drum、Levir-ship 和 Levir-airplane 三个数据集来训练 CRBM 模型。通过度量生成的形状图像与原始形状图像之间的欧式距离差异, 来定量的分析 CRBM 模型对不同形状的建模效果差异, 该值越小表明模型生成的形状越接近真实的形状, 生成的形状越接近真实的形状。CRBM 模型在不同数据集上的训练时间以及平均欧氏距离度量值如表 3 所示, 通过对比

三个数据集在 CRBM 的训练时间可以发现，Levir-ship 数据集的训练时间最短，这是由于 Levir-ship 数据集训练样本的数量最少。且针对不同的数据集在 CRBM 模型上的生成形状不论是训练集还是测试集的平均欧氏距离都较小，这是由于 CRBM 中的卷积操作能够很好的提取输入样本的局部信息，从而使得生成形状的结果更接近真实的形状。

表 3 模型训练时间及生成形状结果的平均欧氏距离度量
Tab.3 Ttraining time of model and average Euclidean distance measure of shape results

数据集	训练时长(s)	训练集	测试集
Levir-oil drum	927.82	3.3475	3.6956
Levir-ship	627.61	3.7843	4.1538
Levir-airplane	1720.4	4.1912	4.5390

图 6 分别展示了 Levir-oil drum、Levir-ship 和 Levir-airplane 数据集的测试集中的五幅图像生成形状结果。其中(a)是输入图像，(b)是 CRBM 生成形状的结果图。模型生成形状的时间都在 0.25s 左右。由图可见，就生成形状的效果而言油桶的生成形状结果最好，重构出来的形状能很好的保留目标油桶的形状特征，符合真实的形状。从 Levir-ship 数据集生成的生成形状结果可以看出，对于包含多个目标船舶的图像，CRBM 模型依然能够很好的建模形状。这是 CRBM 模型中的“卷积操作”使得模型能够很好的提取输入样本的全局特征，可以灵活地表示形状及生成形状。从 Levir-airplane 数据集生成的生成形状结果可以看到，由于目标飞机的形态结构较为复杂，使得 CRBM 生成的飞机形状与真实的飞机形状仍有较大的差异，如第 3、5 幅图像结果所示，生成的飞机形状丢失较多机翼和发动机的细节特征。

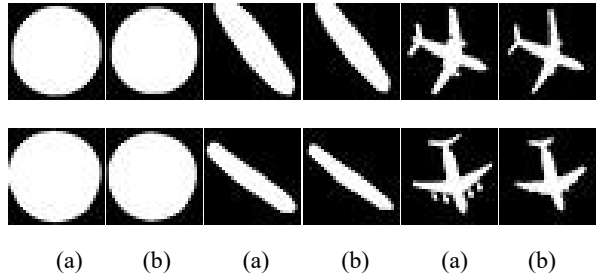
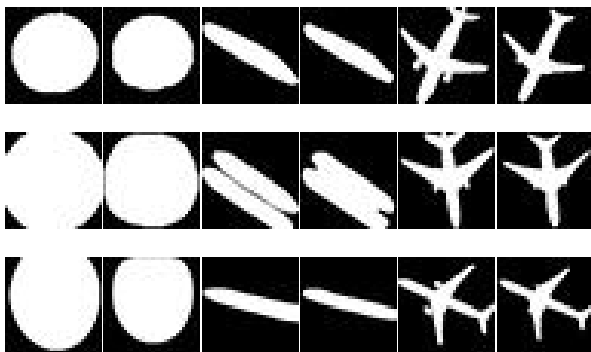


图 6 不同数据集的生成形状结果，(a)输入图像，(b)CRBM 生成形状结果。

Fig.6 Results of shape on different data set, (a)Input image, (b)CRBM shaps.

6.2 Levir 数据集分割实验

图 7(I)、7(II)和 7(III)展示了 Levir-oil drum、Levir-ship 和 Levir-airplane 三个数据集的训练集和测试集中的五幅图像分割结果图。其中(a)表示原始图像，(b)表示 CV 模型分割结果图，(c)表示 CG-CV 模型分割结果图，(d)表示 CLG-CV 模型分割结果图，(e)表示 Ground-truth 图。

图 8(I)、8(II)和 8(III)分别是对应图 7(I)、7(II)和 7(III)中不同模型在 Levir-oil drum、Levir-ship 和 Levir-airplane 三个数据集的训练集和测试集中的五幅图像分割结果评价。其中全局精度(Global acc)和交并比(IOUS)分别衡量了分割的完整性和正确性，其值越高越好。由图可见 Levir-oil drum 集中，存在原始图像是目标油桶被周围目标背景阴影遮挡的情况(4、5)、原始图像是目标油桶图像较复杂的情况(3、6、7)、以及原始图像的尺寸小于 32×32 的情况(1、2、10)。Levir-ship 数据集中目标船舶个数不同、形态各异，且目标船舶周围都存在一定的背景影响。例如第 1、3、5、8 幅原始图像中目标船舶与背景海面相近且原始图像尺寸均小于 32×32 。第 4 幅原始图像中目标船舶的个数为两个的情况。第 6、7、10 幅原始图像中是目标船舶较小的情况。相较于传统的 CV 模型，由于 CG-CV 模型和 CLG-CV 模型引入先验知识指导分割，使得 CG-CV 模型和 CLG-CV 模型能有效的克服目标与背景杂糅、目标遮挡等噪声干扰的影响，有效的提升了模型的分割精度。并且所提模型 CLG-CV 在每一幅图上的度量指标能够达到最高，分割的结果更符合实际，效果更好。这是由于在 CLG-CV 模型中将边缘信息引入模型的形状约束项中，加强了目标形状的边缘细节特征，能更好的约束演化曲线向目标轮廓不断逼近，得到正确的分割结果。

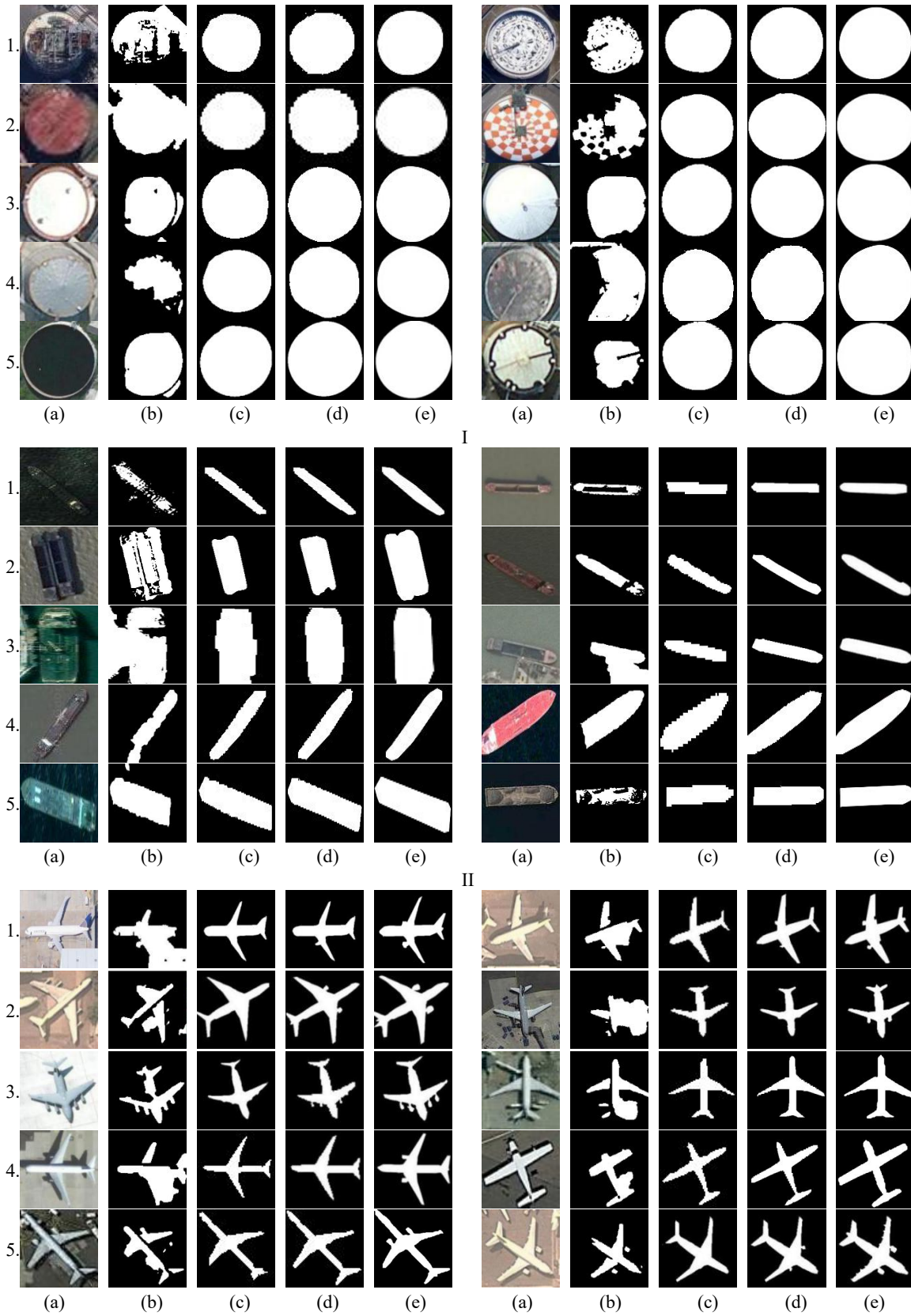
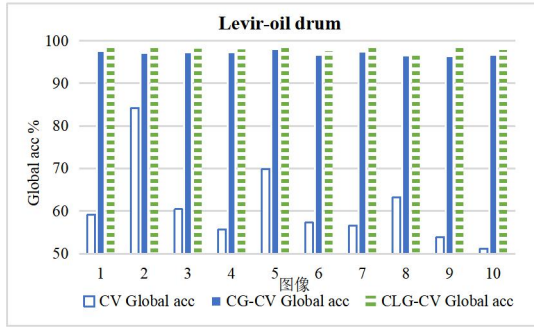
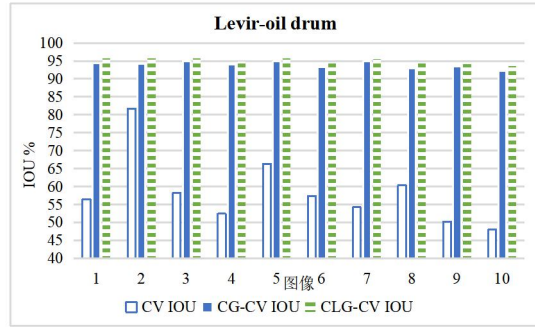


图 7 不同模型的分割结果, (I)Levir-oil drum 数据集的分割结果, (II)Levir-ship 数据集分割结果, (III)Levir-airplane 数据集的分割结果, (a)原图像, (b)CV 分割结果, (c)CG-CV 分割结果, (d)CLG-CV 分割结果, (e)Ground-truth 图。
 Fig.7 Different model segmentation results, (I) segmentation results of Levir-oil drum, (II) segmentation results of Levir-ship, (III) segmentation results of Levir-airplane, (a)RGB image, (b)CV model segmentation results, (c)CG-CV model segmentation results, (d)CLG-CV model segmentation results, (e)Ground-truth.

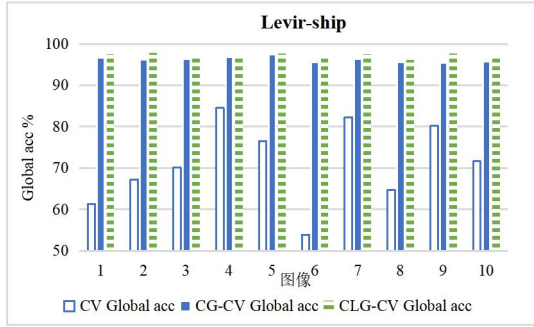


(a)

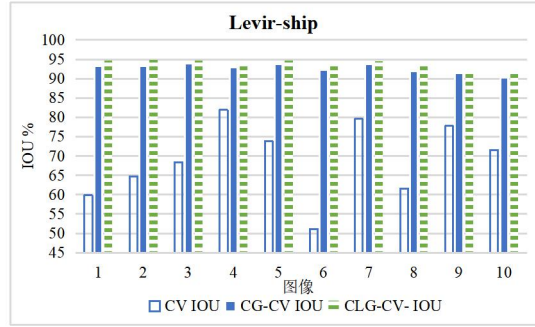


(b)

I

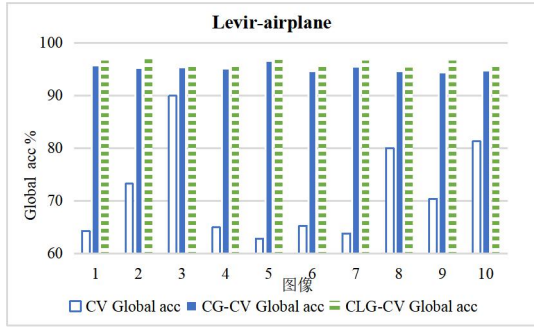


(a)

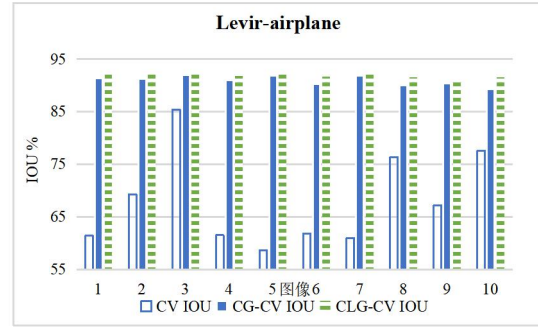


(b)

II



(a)



(b)

III

图 8 不同模型的分割评价结果, (I)Levir-oil drum 数据集的分割结果评价, (II)Levir-ship 数据集分割结果评价, (III)Levir-airplane 数据集的分割结果评价, (a)不同模型的全局精度结果, (b)不同模型的交并比结果。

Fig.8 The segmentation results of each model, (I)evaluation of segmentation results on Levir-oil drum, (II) evaluation of segmentation results on Levir-ship, (III) evaluation of segmentation results on Levir-airplane, (a)the Global acc values of each model, (b)the IOU values of each model

表 4 不同模型在 Levir 测试集上的分割结果评价

Tab.4 The segmentation results of each model on Levir text set

	CV		CG-CV		CLG-CV	
	Avg Global acc	Avg IOU	Avg Global acc	Avg IOU	Avg Global acc	Avg IOU
Levir-oil drum	60.495	57.407	97.271	94.391	98.654	95.328
Levir-ship	69.382	66.182	96.082	93.586	97.936	94.140
Levir-airplane	87.753	83.375	95.453	91.742	96.628	92.425

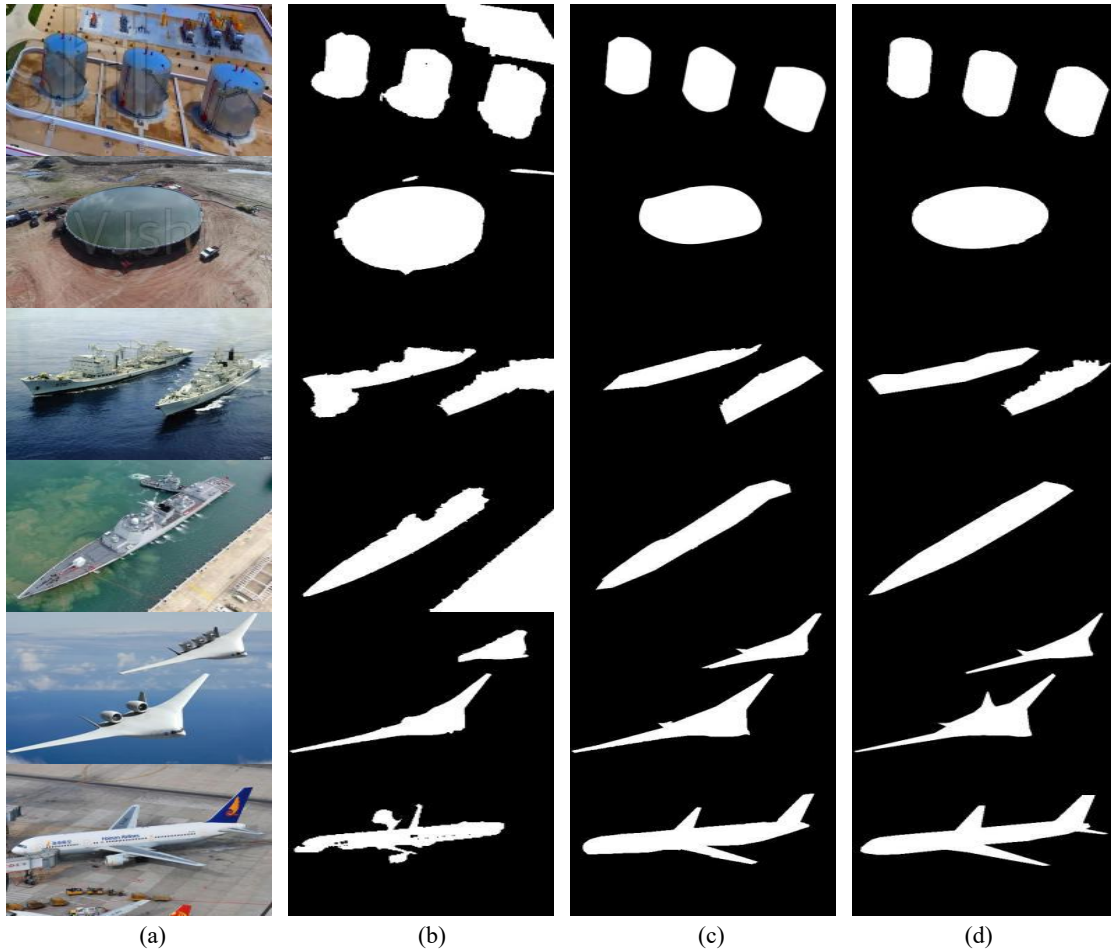


图9 其他图像上的分割结果图, (a)原图像, (b)CV分割结果, (c)CG-CV分割结果, (d)CLG-CV分割结果。

Fig.9The segmentation results on other images,(a)RGB image, (b)CV model segmentation results, (c)CG-CV model segmentation results, (d)CLG-CV model segmentation results.

在 Levir-airplane 数据集中, 由于原始图像目标飞机的形态较油桶及船舰更为复杂, 且原始图像中飞机周围都存在背景环境的影响。如第一幅图像中是目标飞机被云层遮挡的情况。可以看出无论是训练集还是测试集, 虽然 CLG-CV 模型在每副图像的性能评价价值最高。这是由于在模型中将边缘信息引入模型的形状约束项中, 加强了目标形状的边缘细节特征, 能更好的约束演化曲线向目标轮廓不断逼近, 得到正确的分割结果。但与标记图相比仍有一定的差距, 丢失了大量尾翼和飞机发动机的细节信息。这是由于建模目标形状以及生成目标时, 模型的输入尺寸为 32×32 , 在图像缩放的过程中丢失了细节信息。从而使得分割效果较差。

为了验证所提模型的分割性能, 表4给出了 CV、CG-CV 以及 CLG-CV 模型在 Levir 数据集的测试上的分割性能评价结果。相较于 CV 模型, CLG-CV 模型在 Levir-oil drum、Levir-ship 和 Levir-airplane 三个数据集上的平均 Global acc 值

分别提高至 98.654%、97.936%以及 96.628%。在 Levir-oil drum、Levir-ship 和 Levir-airplane 三个数据集的测试集上的平均 IOU 值分别提高至 95.328%、94.140%以及 92.425%。

6.3 其他图像分割实验

为了验证所提模型 CLG-CV 的推广性, 从网上选取与训练集中图像在背景、位置、数目、上均有差异的 256×256 大小的图像, 送入模型进行测试, 图9中(a)原始图像, (b)CV分割结果图, (c)CG-CV分割结果图, (d)CLG-CV分割结果图。

7 结论

为解决遥感影像复杂背景下, 远距离成像的小目标分割问题, 提出结合目标局部和全局特征的 CV 遥感图像分割(CLG-CV)。在小目标遥感数据上的实验结果表明, 相较于只结合卷积受限玻尔兹曼机 CG-CV 模型, CLG-CV 模型的分割效果最佳。这是由于 CLG-CV 在采用生成式模型获

取目标形状信息的同时,将原图像的边缘信息引入模型的形状约束项中,通过“点乘”的方式减少了边缘信息中虚假目标边界,弥补了深度学习生成模型生成形状时,丢失的目标边缘细节特征。图9的结果表明,与传统CV模型分割方法和CG-CV模型分割方法相比,所提模型在小目标遥以及其他尺寸较大的遥感影像分割应用中也依然能取得较好的分割结果。

参考文献

- [1] 王好贤,董衡,周志权.红外单帧图像弱小目标检测技术综述[J].激光与光电子学进展,2019,56(08):9-22.
- [2] 王好贤,董衡,周志权.红外单帧图像弱小目标检测技术综述[J].激光与光电子学进展,2019,56(08):9-22.
- [3] Chan T F, Sandberg B Y, Vese L A.Active Contours without Edges for Vector-Valued Images[J].Journal of Visual Communication and Image Representation,2000,11(2):130-141.
- [4] Scholkopf B, Smola A, Muller K R.Nonlinear Component Analysis as a Kernel Eigenvalue Problem[J].Neural Computation,996,10(5):1299-1319.
- [5] 杨建功,汪西莉,李虎.融合 Kernel PCA 形状先验信息的变分图像分割模型[J].中国图像图形学报,2015,20(8):1035-1041.
- [6] 田杰,韩冬,胡秋霞等.基于 PCA 和高斯混合模型的小麦病害彩色图像分割[J].农业机械学报,2014,45(7):267-271.
- [7] Amberg B, Vetter T.Optimal landmark detection using shape models and branch and bound[C]//Proceedings of the IEEE International Conference on Computer Vision,2011:455-462.
- [8] 郭靖,江洁,曹世翔.水平集分层分割遥感图像中的建筑物 [J]. 红外与激光工程,2014,43(4):1332-1337.
- [9] 董建园,郝重阳.基于统计先验形状的水平集图像分割综述 [J]. 计算机科学,2010,37(1):6-9.
- [10] 张娟,汪西莉,杨建功.基于深度学习的形状建模方法[J].计算机学报,2018,41(01):132-144.
- [11] Fangzheng D, Weihui D.Aircraft Reconstruction in High-Resolution SAR Images Using Deep Shape Prior[J].International of Geo-Information Washington,2017, 12(6):214-232.
- [12] Yuka K,Matvey S,Tsuhana C.In the shadows, shape priors shine: Using Occlusion to improve Multi-Region Segmentation[C]//Proceedings of IEEE Computer Society Conference on computer Vision and Pattern Recognition,2016:392-401.
- [13] Canny J.A computational approach to edge detection[J].Pattern Analysis and Machine Intelligence,1986:679-698
- [14] Norouzi M, Ranjbar M, Mori G.Stacks of convolutional restricted boltzmann machines for shift-invariant feature learning[C]//Proceedings of the IEEE Computer Vision and Pattern Recognition,2009:2735-2742.
- [15] Fukushima K, Miyake S, Ito T.Neocognitron:A Self-Organizing Neural Network Model for a Mechanism of Visual Pattern Recognition[J].IEEE Transactions on Systems Man and Cybernetics,1970,SMC-13(5):826-834.
- [16] Hinton G E, Geoffrey E.Training products of experts by minimizing contrastive divergence[J]. Neural computation,2002,14(8):1771-1800.
- [17] Walsh B.Markov chain Monte Carlo and Gibbs sampling[J].Notes,2004,91(8): 497-537.
- [18] Zou Z, Shi Z.Random Access Memories:A New Paradigm for Target Detection in High Resolution Aerial Remote Sensing Images[J].IEEE Transactions on Image Processing,2018,27(3):1100-1111.
- [19] Chan T F, Sadberg B Y, Vese L A.Active contours without Edges for vector-valued images[J].Journal of Visual Communication and Image Representation,2000,11(2):130-141.
- [20] 李晓慧,汪西莉.结合卷积受限玻尔兹曼机的 CV 分割模型 [J]. 激光与光电子学进展,2020,57(4):041018.