



几种具有编解码结构的深度学习模型在建筑物综合中的应用与比较

杜佳威¹ 武芳¹ 行瑞星¹ 李彩霞¹ 李靖涵²

1 信息工程大学地理空间信息学院,河南 郑州,450001

2 61363部队,陕西 西安,710054

摘要:深度学习技术促使诸多领域研究取得突破性进展,基于深度神经网络的地图综合研究备受期待。将建筑物综合过程抽象解释为编解码过程,系统地研究基于编解码结构的深度学习神经网络在建筑物综合中的应用。首先,利用空间划分与矢量-栅格数据转换相结合的方式构建样本和样本集;然后,利用样本集训练基于编解码结构的深度学习神经网络,实现建筑物综合学习泛化并测试、评估其效果;最后,搭建5种代表性的基于编解码结构的深度学习神经网络,分析比较各模型在建筑物综合中的应用效果。实验结果表明,基于编解码结构的深度学习神经网络能够从建筑物综合样本中学习或推理出部分建筑物综合知识和综合操作,且5种模型中Pix2Pix更适用于建筑物综合的学习模拟。

关键词:地图综合;建筑物;深度学习;神经网络;样本

中图分类号:P283

文献标志码:A

长期以来,从人工综合过程中抽象出地图综合约束,并设计符合约束的综合算法是实现自动综合的常用策略。而人工综合中不可避免的主观理解和创作^[1]使地图综合约束难以全部清晰、准确地描述。利用机器学习技术从综合成果数据中学习地图综合的“自下而上”^[2]的研究思路逐步开展。传统机器学习方法相对简单、学习能力有限,可用于模拟能够被抽象为分类、聚类、优化任务的部分综合过程。例如,针对选取、化简、概括、移位4种常见地图综合算子^[3],产生了以支持向量机、 k 邻近、浅层神经网络等分类器为基础的自动选取方法^[4-8],以聚类算法为基础的自动概括方法^[9-11],以遗传算法、蚁群算法等优化算法为代表的自动化简^[12-14]、自动移位方法^[15-16]。然而,地图综合过程十分复杂,不仅包含除分类、聚类、优化等简单任务外更加复杂的制图风格迁移与主观综合创作,还需要多种综合操作的协同处理。随着传统机器学习向深度学习进阶,模型学习复杂任务能力显著增强,可尝试模拟更加复杂的综合过程。例如,文献[17]尝试利用生成式对抗网络模拟从原始矢量数据到多尺度瓦片地图的制

图风格迁移。

基于深度学习的建筑物综合研究刚刚起步。文献[18]利用深度卷积自编码器编码建筑物,基于建筑物编码进行模板匹配,实现单个建筑物的化简。但此方法受限于模板库内模板种类,只能实现有限的建筑物形态变换。文献[19-20]利用全卷积网络模拟单个建筑物综合过程,丰富了建筑物综合中形态变换的可能性,但模型预测结果中存在轮廓模糊或缺失等建筑物表达问题。文献[21]利用残差Unet优化了建筑物综合结果的表达效果。这些研究均采用了具有编解码结构的深度学习神经网络,这是因为编解码过程与人工地图综合过程具有一定的相似性和可类比性。具体而言,建筑物综合过程中需要从综合前数据中抽象出综合知识、约束,然后结合这些知识、约束进行综合操作得到综合结果;具有编解码结构的神经网络先利用编码器从原始数据中编码特征,然后利用解码器从特征编码中解码出目标结果。随着深度学习的发展,基于编解码结构的深度学习神经网络也不断丰富。本文研究基于编解码结构的深度学习神经网络在建筑物综合中的应用,给出相应

收稿日期:2021-01-22

项目资助:国家自然科学基金(41471386, 41801396)。

第一作者:杜佳威,博士,研究方向为自动制图综合与空间数据智能处理。whdxdjw@126.com

通讯作者:武芳,博士,教授,博士生导师。wufang_630@126.com

的样本构建方法和训练测试过程,通过实验测试分析比较几种代表性的基于编解码结构的深度神经网络在建筑物综合中的应用效果。

1 基于编解码结构的深度神经网络

1.1 经典的基于编解码结构的深度神经网络

编解码结构是人工神经网络中常用的结构之一,广泛应用于特征提取、语义分割、图像翻译等诸多领域。编解码结构由编码器和解码器两部分构成:编码器从输入数据中提取特征编码,解码器对特征编码进行解码得到输出结果。经典的编解码深度神经网络(encoder-decoder deep neural network, E-DDNN)的编码器由卷积层和下采样层构成,解码器由卷积层和上采样层构成,如图 1 所示。编码器通过卷积运算提取特征,通过池化运算对特征进行下采样,重复此过程得到特征编码;解码器通过插值运算对特征编码上采样,通过卷积运算压缩特征,重复此过程得到输出结果。此外,也可采用将插值和卷积合并的反卷积运算。

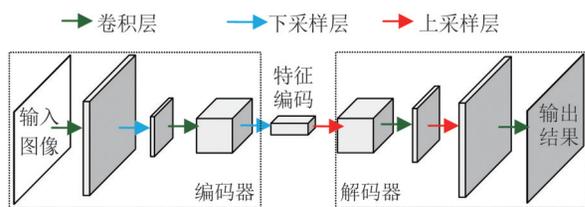


图 1 经典的 E-DDNN
Fig.1 A Classical E-DDNN

1.2 几种基于编解码结构的深度神经网络变体

随着神经网络研究的丰富,新的神经网络结构与编解码结构结合使 E-DDNN 出现变体,以下 4 种最具代表性:

1)具有跳跃结构的 E-DDNN。这种变体是在经典的 E-DDNN 中加入跳跃连接,将编码器提取的特征合并到解码器上,解码器利用特征编码的解码特征和编码器编码特征预测输出结果,如图 2(a)所示。这种跳跃结构能够将浅层特征与深层特征结合,以兼顾全局特征和局部特征预测输出结果。文献[22]提出的 Unet 就是一种具有跳跃结构的 E-DDNN。

2)残差 E-DDNN。编码器与解码器间的跳跃连接能够增加捷径,丰富待学习特征。在此基础上,在连续的每两层间增加跳跃连接即可构建残差块。残差块能够避免图像特征信息损失,降低深度网络退化风险。残差 E-DDNN 利用残差

块提取特征编码,再利用上采样和残差块压缩特征得到输出,如图 2(b)所示。

3)稠密 E-DDNN。利用跳跃连接能够改变原有编解码结构的拓扑连接,使 E-DDNN 中的连接关系更加稠密,使提取的特征信息得到更加充分的利用。文献[23]提出的 Unet++ 就是一种极具代表性的 E-DDNN,该模型仅通过各层间短连接即可有效融合深浅层特征,如图 2(c)所示。

4)与生成式对抗网络(generative adversarial network, GAN)结构结合的 E-DDNN。GAN 由生成器和判别器组成,如图 2(d)所示。生成器能够预测输出结果,判别器能够判别真实结果与预测结果,生成器与判别器竞争使得生成器预测结果令判别器无法区分。E-DDNN 常作为 GAN 的生成器,GAN 的目标是学习区分真假的损失函数并用此训练 E-DDNN 生成器,使其损失函数最小化。文献[24]将 Unet 作为生成器,构建了常用的一对一的 GAN 模型——Pix2Pix。

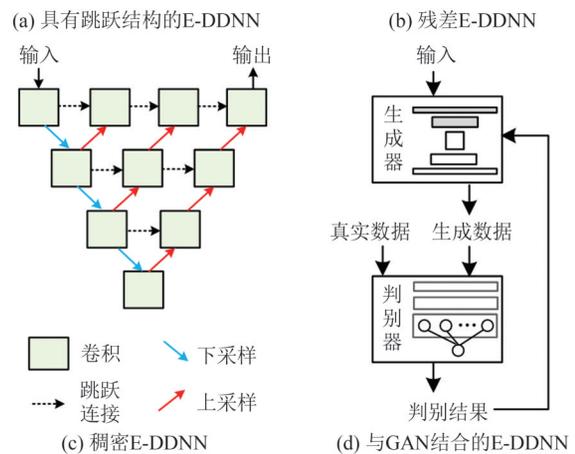
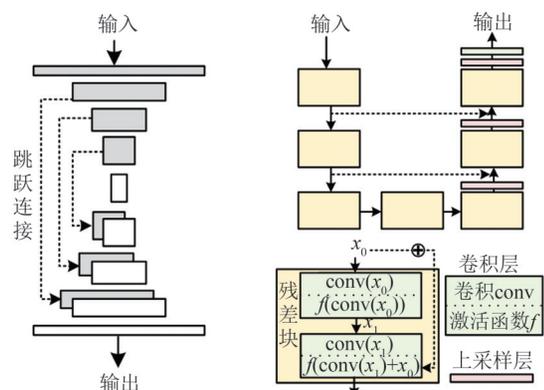


图 2 E-DDNN 的几种变体
Fig.2 Some Variants of E-DDNN

2 基于 E-DDNN 的建筑物综合

2.1 技术路线

深度学习模型是模拟建筑物综合的核心,本

文采用E-DDNN学习模拟从综合前到综合后的建筑物综合过程。现有的E-DDNN并不适用于直接处理综合前后的矢量建筑物数据,先要从已有的综合前后的矢量建筑物中提取、构建能够被E-DDNN学习处理的栅格样本;再将提取的栅格样本划分为训练集、验证集和测试集,利用训练集和验证集训练E-DDNN,利用测试集测试E-DDNN在模拟建筑物综合中的效果,具体技术路线如图3所示。

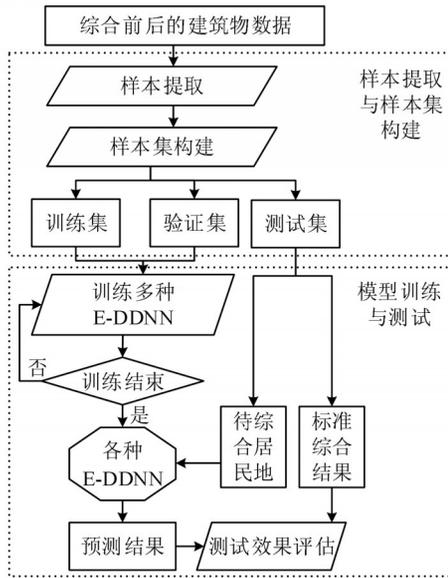


图3 基于E-DDNN的建筑物综合流程

Fig.3 Flowchart of Building Generalization Based on E-DDNN

设 I_a 和 I_b 分别为综合前和综合后的建筑物图像,E-DDNN前 n_1 层为编码器, $n_1 \sim n_2$ 层为解码器,第 i 层激活函数、卷积核、偏置分别记为 f_i, k_i, b_i ; *为卷积运算,Pool为池化函数,Unpool为反池化函数,Entropy为求熵函数; λ 为学习率, θ 为特征编码。则利用经典E-DDNN学习、模拟建筑物综合的计算过程如下:

1)随机初始化 $\{k_i\}$ 和 $\{b_i\}$ 。

2) I_a 的编码计算:

(1)令 $I=I_a, i \in (0, n_1]$;

(2)卷积、池化,即 $I_i=f(k_i \times I) + b_i; I_i=\text{Pool}(I_i)$;

(3)令 $I=I_i$,重复步骤(2)~(3)至 n_1 终止,则 $\theta=I$ 。

3) θ 的解码计算:

(1)令 $I=\theta, j \in (n_1, n_2]$;

(2)卷积、反池化,即 $I_i=f(k_i \times I) + b_i; I_i=\text{Unpool}(I_i)$;

(3)令 $I=I_i$,重复步骤(2)~(3)至 n_2 终止,则

$I'_b=I$ 。

4)损失值计算,即 $\Delta=\text{Entropy}(I_b, I'_b)$ 。

5)权重和偏置的调整:

(1)令 $i=n_2$;

(2)计算新的参数值,即 $k_i=k_i-\lambda; b_i=b_i-\lambda$;

(3)反向调整所有参数值。令 $i=i-1$,利用反向传播方法^[25]依次调整权重、偏置值,重复步骤(2)~(3),直至 $i=1$ 为止。

虽然不同E-DDNN模型的结构不同,模拟建筑物从综合前到综合后的图像变换原理不尽相同,但是基本思路大体一致。具体而言,首先,利用编码器从综合前的建筑物图像中提取编码;然后,利用解码器将特征编码恢复为栅格图像预测综合结果;接着,将预测的综合结果与该区域标准综合结果比较并计算损失值;最后利用梯度下降算法^[25]和反向传播算法^[25]调节E-DDNN的模型参数。通过相同综合尺度的多对综合前、综合后的建筑物图像训练E-DDNN,优化E-DDNN模型参数,使其能够模拟从综合前到综合后的图像变换。类比于人工综合,训练后的E-DDNN能够根据综合前的建筑物提取与综合相关的知识(特征编码),并利用这些知识(特征编码)进行建筑物综合操作、绘制综合结果。

2.2 建筑物综合样本提取与样本集构建

构建适用于E-DDNN的建筑物综合样本和样本集需要顾及以下内容。

1)计算资源限制。建筑物是大比例尺地图中的常见要素,分布广泛且数量庞大。直接将海量建筑物作为样本供E-DDNN学习,将产生庞大的参数计算,需要极大的计算空间;建筑群间存在大面积空白区域,这些空白区域对建筑物综合影响甚微,导致计算资源浪费。此外,将广泛分布的大量建筑物作为样本还会增加学习难度,令E-DDNN难以有效学习模拟。因此,顾及计算资源限制、简化学习难度,需要对建筑物进行分割。

2)图像处理优势。顾及E-DDNN在图像处理中的良好应用,将分割后的矢量建筑物转换为栅格建筑物作为学习样本。栅格建筑物的像元大小和像元数量设置要恰当。若像元过小,则样本覆盖空间范围有限,难以充分表达建筑物分布特征及其影响的综合处理;若像元过大,则样本细节损失,难以充分反映建筑物局部特征及其影响的综合处理。若像元数量过少,则样本覆盖空间有限;若像元数量过多,则加重模型负荷。转换后的栅格建筑物还要进行归一化处理以适应

梯度下降方法,减少数值问题和收敛问题。

3)数据增强需求。为丰富样本多样性和差异性,样本构建过程中还要包含能够衍生出更多差异样本的数据增强方法。

4)样本集间互不重叠。训练集、验证集和测试集间互不重叠确保了验证集之于训练集、测试集之于训练集的未知性,进而确保了模型泛化的有效性。因此,需要将获得的所有样本按一定比例分割、组织成互不重叠的训练集、验证集和测试集。

基于以上考虑,从矢量建筑物数据中提取栅格样本、构建样本集的具体过程如下:

(1)确定矢量-栅格转换的相关参数。首先,确定像元大小,记为 p_i 。为保证地图要素清晰可视,需满足 $p_i \leq \epsilon/\tau_1$,其中, ϵ 为地图最小可分辨距离, τ_1 为综合前的比例尺。令 $p_i = \epsilon/\tau_1$ 能够覆盖更大范围,尽可能充分地表达建筑物间的空间关系。然后,确定栅格图像行、列像素数量,记为 p_n 。

(2)对综合前后的矢量数据进行规则格网分割。以综合前建筑群包络矩形左下顶点为起点 (x,y) ,以 $d_i = p_i \times p_n$ 为边长,逐行、逐列依次构建规则格网覆盖包络矩形。利用构建的规则格网分割综合前、综合后的矢量数据,分割结果记为 $\{V_1(i,j), V_2(i,j)\}_m$ 。其中, (i,j) 为格网行列索引, $V_1(i,j)$ 和 $V_2(i,j)$ 分别为包含于格网 (i,j) 内综合前和综合后的部分矢量数据, m 为格网数量。

(3)获取栅格样本。首先,删除不包含建筑物的 $V_1(i,j)$ 及其对应的 $V_2(i,j)$;然后,利用步骤(1)中的相关参数,将剩余 $V_1(i,j)$ 、 $V_2(i,j)$ 转换为栅格图像;最后,对图像像素值归一化处理得到栅格样本 $\{R_1(i,j), R_2(i,j)\}$ 。依此步骤遍历 $\{V_1(i,j), V_2(i,j)\}_m$,得到栅格样本集合,记为 $S = \{R_1(i,j), R_2(i,j)\}_n$, n 为样本数量。

(4)实现样本增强。改变规则格网构建的起点 (x,y) 至 $(x-d_x, y-d_y)$,对于 $d_x \in (0, d_i)$ 且 $d_y \in (0, d_i)$,重复步骤(2)~(3)都能得到新的样本集合。样本增强后,得到的所有样本集合记为 $\{S_p\}$ 。

(5)构建训练集、验证集和测试集。将 S_p 内的样本按8:1:1依次分割为 S_{p_1} 、 S_{p_2} 和 S_{p_3} 3个样本集,即 $S_p = \{S_{p_1}, S_{p_2}, S_{p_3}\}$ 。则 $\{S_p\} = \{\cup\{S_{p_1}, S_{p_2}, S_{p_3}\}\} = \{\cup\{S_{p_1}\}, \cup\{S_{p_2}\}, \cup\{S_{p_3}\}\} = \{\Phi, \Psi, \Omega\}$ (Φ 、 Ψ 、 Ω 分别为最终得到的训练集、验证集和测试集),完成样本集组织构建。

2.3 模型的训练和测试

E-DDNN通过训练自适应调整模型参数,学

习从综合前到综合后的建筑物综合变换;训练结束后,E-DDNN参数固化,通过测试评判模型对建筑物综合的模拟效果。训练集用于训练过程中E-DDNN的参数调整,验证集用于初步训练模型的验证并为当前模型训练提供评估,测试集用于已训练好模型的效果测试。设计模型预测的综合结果与标准综合结果差异的度量指标(E_1),用于不同E-DDNN训练过程和测试效果的评估、比较。 E_1 的计算公式为:

$$E_1 = \frac{\sum_{i=1}^{p_n} \sum_{j=1}^{p_n} |R_2[q, i, j] - R_3[q, i, j]|}{p_n^2} \quad (1)$$

式中, $R_2[q, i, j]$ 、 $R_3[q, i, j]$ 分别表示集合内第 q 张综合结果图中第 i 行、 j 列处像素的标准值和预测值。

模型训练和测试的具体过程如下。

1)依次从 Φ 中取 n_b 个样本作为一个批次输入E-DDNN,利用选定的梯度下降算法和学习率自适应调整模型参数,计算并记录训练集样本的 E_1 。重复此步骤,直至 Φ 中所有样本都参与过模型训练时进入下一步。

2)固化E-DDNN参数,将 Ψ 中的样本输入模型,计算并记录验证集样本的 E_1 ,进入下一步。

3)重复步骤1)~步骤2)过程 n_c 次,完成E-DDNN训练,进入下一步。

4)固化E-DDNN参数,进行模型测试。将 Ω 中的样本输入模型,计算并记录测试集样本的 E_1 。利用查准率、查全率和F1分数3种指标评价模型预测精度,分别记为 E_2 、 E_3 、 E_4 ,计算公式为:

$$E_2 = \frac{\text{NUM}(R_2[q] \cap R_3[q])}{\text{NUM}(R_2[q])} \quad (2)$$

$$E_3 = \frac{\text{NUM}(R_2[q] \cap R_3[q])}{\text{NUM}(R_3[q])} \quad (3)$$

$$E_4 = \frac{2 \times E_2 \times E_3}{E_2 + E_3} \quad (4)$$

式中, $R_2[q]$ 、 $R_3[q]$ 分别表示测试集内第 q 张标准综合结果、预测综合结果中构成建筑物的像素;函数NUM用于统计像素个数。

由于地图综合过程中进行了取舍操作,分割后的局部综合结果中可能不包含建筑物,必须对式(2)~式(4)中的特殊情况加以约定。

1)若 $\text{NUM}(R_2[q])=0$ 且 $\text{NUM}(R_3[q])=0$,则 $E_2=1$ 、 $E_3=1$ 和 $E_4=1$;

2)除1)外,若 $\text{NUM}(R_2[q])=0$ 或 $\text{NUM}(R_3[q])=0$,则 $E_2=0$ 、 $E_3=0$ 和 $E_4=0$;

3)除1)、2)外,若 $E_2=0$ 且 $E_3=0$,则 $E_4=0$ 。

3 建筑物综合实验与分析

3.1 实验设置

实验数据来源于某地约 $44.1\text{ km}\times 38.6\text{ km}$ 范围内1:1万和1:5万的矢量地图数据,如图4所示。其中,1:1万数据中包含综合前建筑物

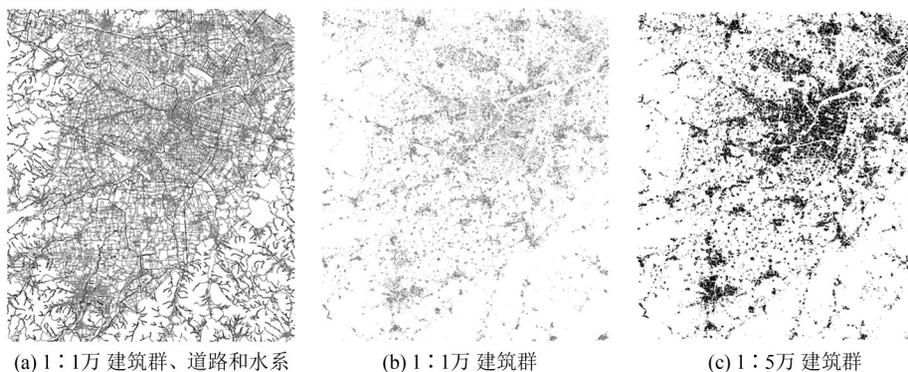


图4 实验数据
Fig.4 Test Data

基于 ArcEngine 二次开发,利用§2.2中方法构建样本集:根据参考文献[3],令 $\epsilon=0.2\text{ mm}$,确定像素大小 $p_r=2\text{ m}$;统计综合后1:5万数据中建筑物包络矩形最长边长,如图5所示,其中综合后99.59%的建筑物能完全包含于 $512\text{ m}\times 512\text{ m}$ 的矩形区域,可以令 $p_n=256$;构建规则网格分割矢量数据并转换为栅格样本,得到 $S_1=\{R_1(i, j), R_2(i, j)\}_{5182}$ 和 $S_2=\{P_1(i, j), R_2(i, j)\}_{5182}$,其中, $R_1(i, j)$ 中只包含建筑物, $P_1(i, j)$ 中包含建筑物、道路网和水系要素;为丰富样本类型和数量,令 $d_x=d_y=256\text{ m}$,得到 $S'_1=\{R'_1(i, j), R'_2(i, j)\}_{5185}$ 和 $S'_2=\{P'_1(i, j), R'_2(i, j)\}_{5185}$;组织 S_1 和 S'_1 、 S_2 和 S'_2 分别得到 $\{\Phi_1, \Psi_1, \Omega_1\}$ 、 $\{\Phi_2, \Psi_2, \Omega_2\}$ 两组样本集。构建的样本均为灰度图像,并利用灰色和黑色区分综合前后的建筑物,样本示例如图6所示。

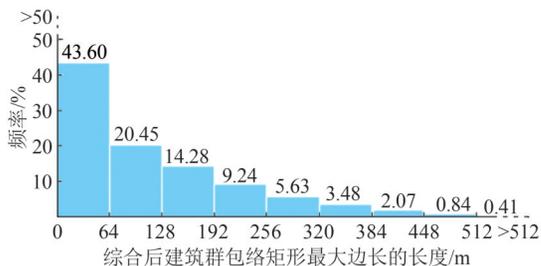


图5 综合后建筑群包络矩形最大边长的频率分布
Fig.5 Frequency of the Longest Edge Length of the Envelope of Generalized Resident at the Scale of 1:50 000

89 539个,1:5万数据中包含综合后建筑物21 732个,涵盖多种形态及空间分布的建筑物,涉及选取、化简、合并、典型化等诸多综合操作。

顾及原始数据中道路、水系对建筑物综合的潜在影响,构建两组样本集提供给EDnet、Unet、ResUnet、Unet++、Pix2Pix这5种E-DDNN模型学习。一组输入样本只包含建筑物,另一组输入样本包含建筑物、道路、水系。

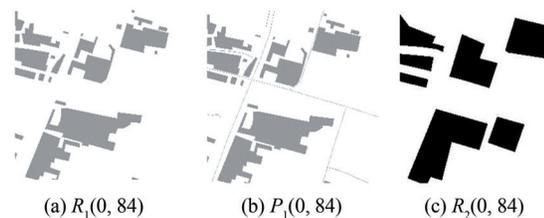


图6 样本示例
Fig.6 A Typical Sample

在图形处理器(graphic processing unit, GPU)(RTX 2070, 6 GB)+Tensorflow环境下通过Python编程搭建了5种E-DDNN模型(EDnet、Unet、ResUnet、Unet++、Pix2Pix),分别对应经典E-DDNN模型及其4种变体。其中,EDnet为本实验搭建的经典E-DDNN,包含4个下采样层、4个上采样层和8个卷积层。本文搭建的Unet、ResUnet、Unet++、Pix2Pix也较文献[21-24]稍加改动:(1)采用边界补零策略,使编解码结构的输入、输出等大,与EDnet一致;(2)上采样、下采样次数与EDnet一致。几种模型均采用Adam算法作为梯度下降算法,学习率为0.000 2。顾及实验硬件条件,令 $n_b=16$ 、 $n_e=50$,利用§2.3中方法分别基于两组样本集对5种E-DDNN模型进行训练和测试。

3.2 5种模型效果比较

5种模型训练过程中训练集、验证集输出结果 E_1 的平均值(记为 $AVE(E_1)$)如图7、图8所示,

模型测试结果 E_1, E_2, E_3, E_4 的平均值 $AVE(E_1), AVE(E_2), AVE(E_3), AVE(E_4)$ 如表 1 所示。

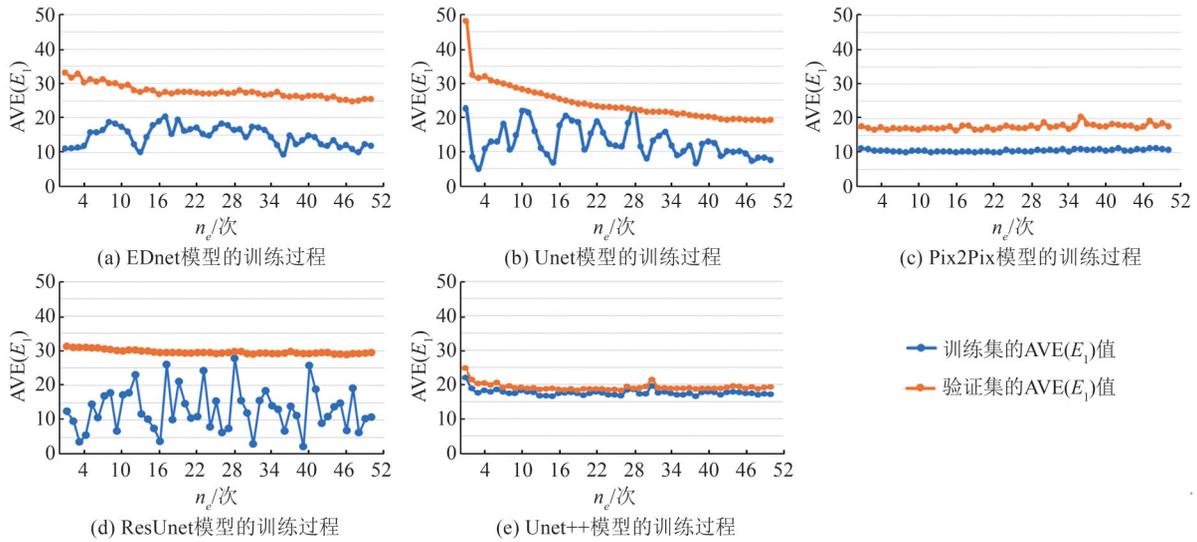


图 7 基于 $\{\Phi_1, \Psi_1\}$ 的模型训练过程

Fig. 7 Training Processes of Different Models with $\{\Phi_1, \Psi_1\}$

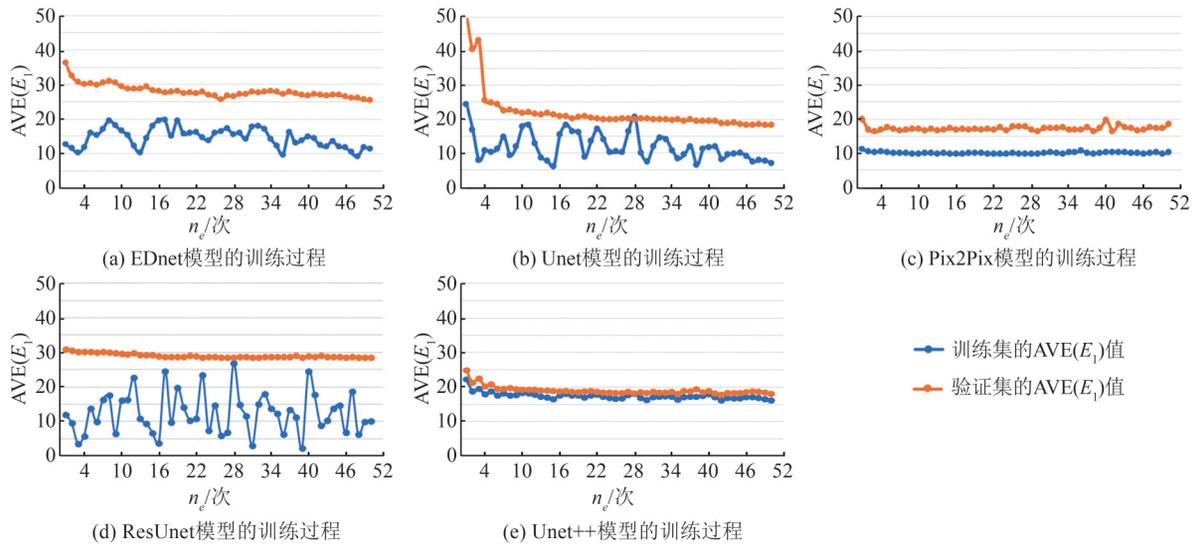


图 8 基于 $\{\Phi_2, \Psi_2\}$ 的模型训练过程

Fig. 8 Training Processes of Different Models with $\{\Phi_2, \Psi_2\}$

表 1 测试结果评价

Tab. 1 Evaluation of Test Results Generalized by Different Models

评价指标	不同模型对 Ω_1 的测试效果					不同模型对 Ω_2 的测试效果				
	EDnet	Unet	ResUnet	Unet++	Pix2Pix	EDnet	Unet	ResUnet	Unet++	Pix2Pix
$AVE(E_1)$	23.644	17.498	16.338	17.727	16.543	25.470	16.928	14.722	14.289	17.936
$AVE(E_2)$	0.503	0.608	0.678	0.651	0.596	0.541	0.619	0.627	0.632	0.568
$AVE(E_3)$	0.623	0.718	0.644	0.614	0.612	0.617	0.725	0.684	0.650	0.615
$AVE(E_4)$	0.534	0.627	0.643	0.611	0.584	0.554	0.644	0.635	0.618	0.567

从训练过程看,随着循环次数的增多,各模型训练结果的 $AVE(E_1)$ 趋于稳定,且模型验证结果的 $AVE(E_1)$ 也都趋于稳定,模型收敛且能够被泛化。训练结束(即 $n_c=50$)时,各模型训练结果的 $AVE(E_1)$ 基本在 10 左右,各模型对训练集的

模拟能力相差不大;各模型验证结果的 $AVE(E_1)$ 却差异明显,Unet、Pix2Pix、Unet++ 验证结果的 $AVE(E_1)$ 较 EDnet、ResUnet 相对更小,表现出更好的泛化准确性。特别是 Pix2Pix 和 Unet++ 通过较少的训练循环就可达到较高的验证准确度,

表现出更强的学习、泛化能力。

训练后的模型在一定程度上具备模拟建筑物综合的能力,将其应用于测试集后对测试结果进行比较分析:

1)定量比较。各模型测试结果的量化评价结果见表1,整体上看,EDnet的测试准确度明显低于其他4种模型的测试准确度,而其他4种模型测试结果的量化评价指标相差不大。其中,ResUnet和Unet++的查准率、查全率和F1分数的平均值相对较大,Unet和Pix2Pix相对较小,由此可知训练后的ResUnet和Unet++模型能够更准确地预测构成综合后建筑物的黑色像素;对于 $\{\Phi_1, \Psi_1, \Omega_1\}$,ResUnet和Pix2Pix测试结果的 $AVE(E_1)$ 值相对较小,对于 $\{\Phi_2, \Psi_2, \Omega_2\}$,ResUnet和Unet++测试结果的 $AVE(E_1)$ 值相对较小,由此可知ResUnet、Unet++、Pix2Pix对测试集预测的综合结果与标准综合结果间差异相对较小。

2)定性比较。定量评价是对预测结果像素的准确性进行评价,建筑物作为独立且完整的地

理对象,形态表达的合理性十分重要。测试结果示例如图9所示。其中,EDnet预测的综合结果十分模糊,几乎不能辨别所表达建筑物的外部轮廓与整体形态;Unet预测的结果中建筑物整体形态得到相对清晰表达,但建筑物的边缘轮廓仍然模糊或缺失;ResUnet和Unet++预测的综合结果中建筑物轮廓相对清晰,但建筑物内部存在较多空洞,建筑物间也存在模糊粘连,建筑物表达的独立性和完整性仍有欠缺;Pix2Pix预测结果中建筑物边缘轮廓清晰,建筑物形态得到独立、完整表达。

综上所述,经典的EDnet模型结构简单、深度较浅,难以模拟复杂的建筑物综合过程;深度加深并添加跳跃连接的Unet能够初步模拟复杂的建筑物综合过程;ResUnet、Unet++、Pix2Pix分别引入不同结构优化Unet模型,使得预测综合结果的准确性与合理性得到明显提升。其中,Pix2Pix模型表现尤其突出,预测结果中建筑物形态表达更加合理,更适用于建筑物综合的学习模拟。

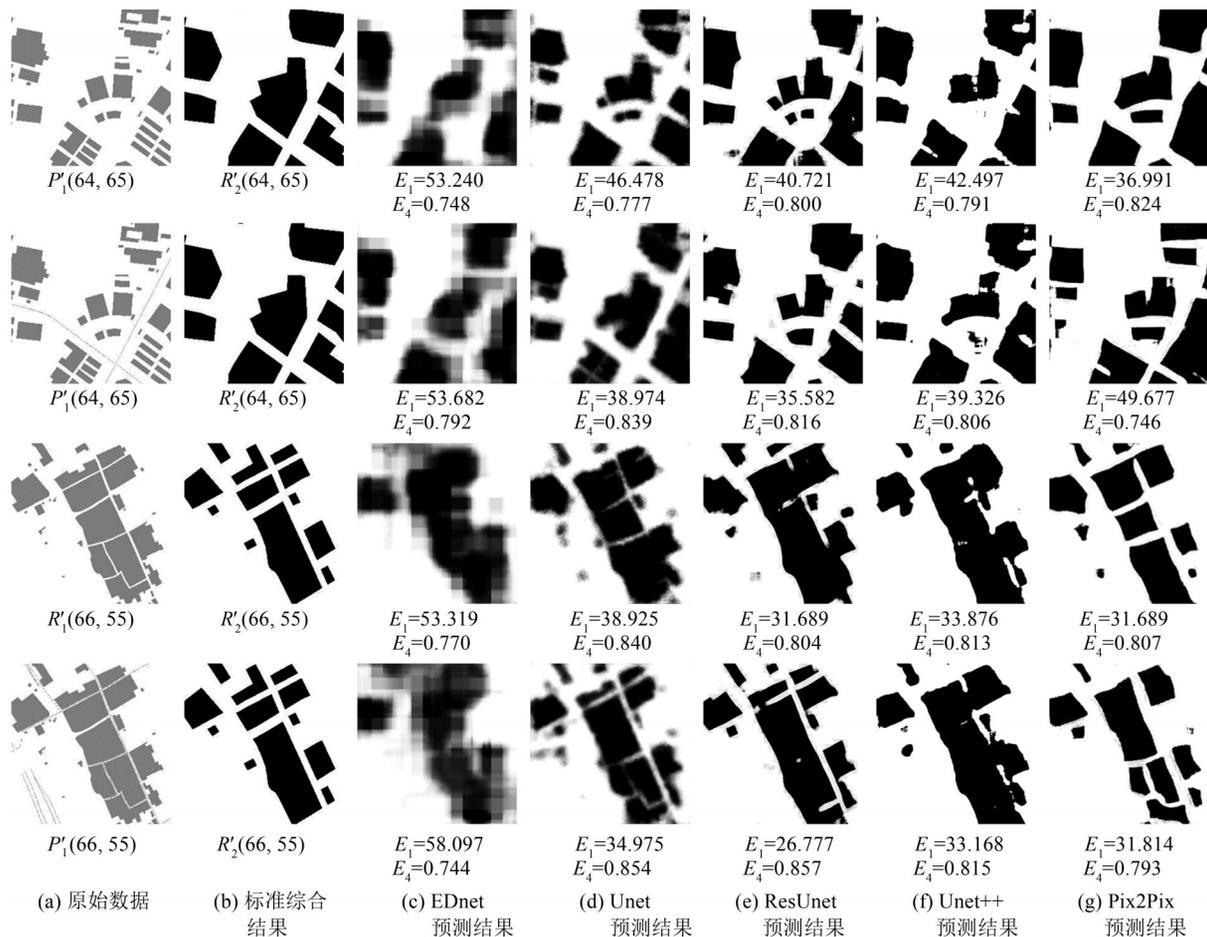


图9 不同模型测试结果示例

Fig.9 Test Results of Different Models

3.3 建筑物综合效果讨论

利用几种 E-DDNN 模型从综合前后的制图成果中学习建筑综合知识、综合操作,践行了“自下而上”^[2]的研究思路。分析综合结果可知:

1) 训练后的深度学习模型能够进行建筑物选取、化简、合并、典型化、夸大等操作。图 10 所示的 Pix2Pix 模型生成器预测的 $R'_1(66,44)$ 综合结果隐含了对原始建筑物的选取、合并、典型化,预测的 $R'_1(64,67)$ 综合结果隐含了对原始建筑物的合并、化简。

2) 深度学习模型也能从数据中学习并利用与建筑物综合相关的潜在知识。分析表 1 中 E_1 和

E_4 值可知,基于 $\{\Phi_2, \Psi_2, \}$ 训练的 EDnet、Unet、ResUnet、Unet++ 总体上比基于 $\{\Phi_1, \Psi_1, \}$ 训练的 EDnet、Unet、ResUnet、Unet++ 具有更高的测试准确性。也就是说,基于道路、水系、建筑物训练的 E-DDNN 用于建筑物综合比基于建筑物单一要素训练的 E-DDNN 模型更加准确,这与建筑物综合需要考虑近邻水系、道路的影响具有一致性。通过目视判读也可从 Pix2Pix 预测的综合结果中发现这种一致性,如图 10 中 $P'_1(64,67)$ 的预测结果和 $P'_1(66,53)$ 的预测结果都顾及了邻近道路对建筑物综合的影响,较 $R'_1(64,67)$ 的预测结果和 $R'_1(66,53)$ 的预测结果更加合理。

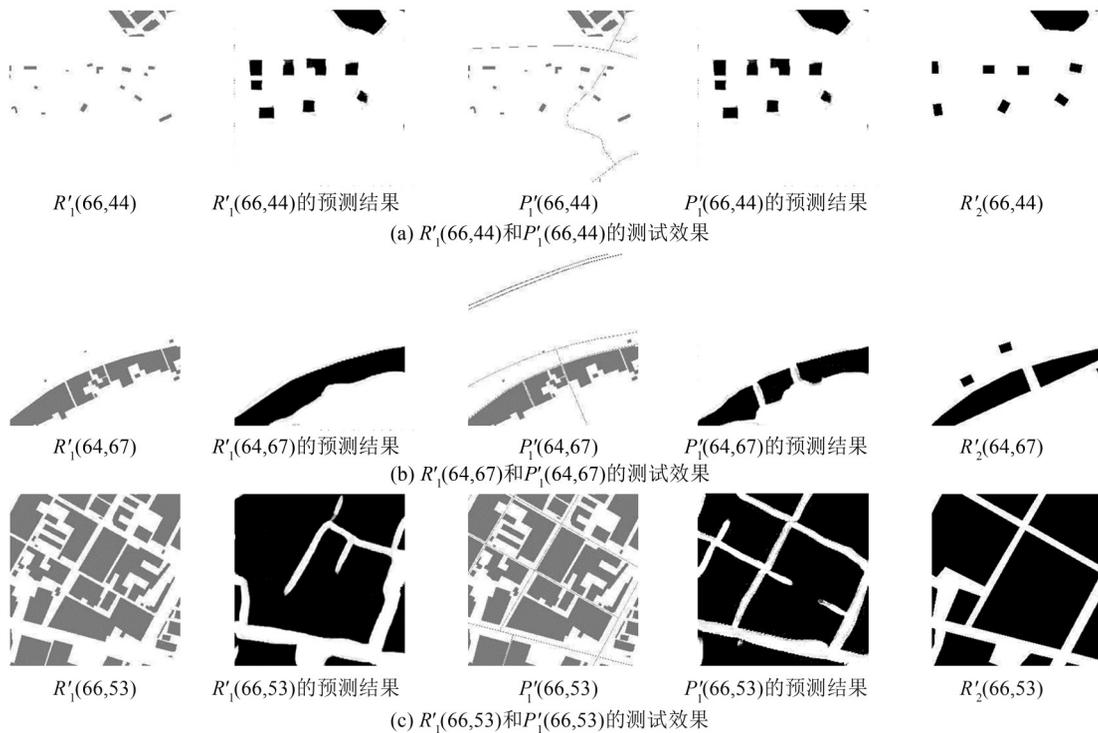


图 10 Pix2Pix 模型测试结果
Fig.10 Test Results of Pix2Pix Model

3) 这种完全“自下而上”的研究思路过于依赖深度学习模型和实验样本,虽然实验训练的深度学习模型具有一定的地图综合能力,但还存在不足。如图 11 所示,由于缺少属性信息, $R_1(65,$

61) 中邻近道路的重要孤立建筑物(红圈内)本应被夸大保留,却被 Pix2Pix 模型删除;由于对合并和典型化算子差异学习不足, $R_1(65,28)$ 中建筑物本应被典型化,却被 Pix2Pix 合并。

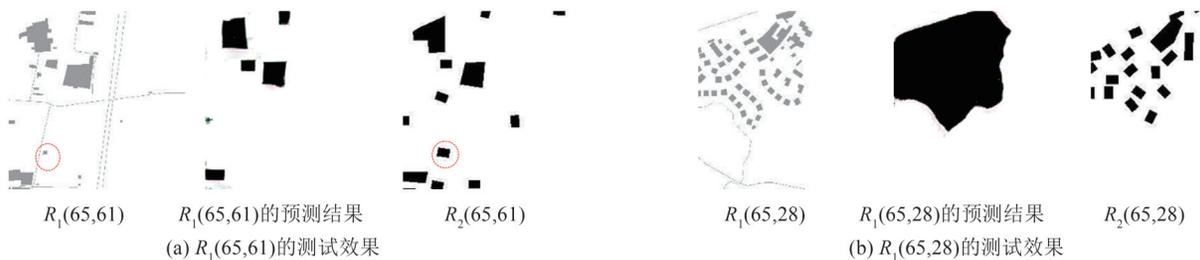


图 11 基于 Ω_2 的 Pix2Pix 模型测试结果
Fig.11 Test Results of Pix2Pix Model Based on the Ω_2

分析实验结果可知,通过以下改进能够优化基于深度学习的建筑物综合研究:

1)深度学习模型是影响综合效果的关键因素,实验结果表明结构更优化的E-DDNN用于模拟建筑物综合效果更好。因此,通过增加模型深度、优化模型结构、改进模型超参数等能进一步优化综合效果。

2)待学习任务的复杂性是影响综合效果的直接因素,实验结果表明E-DDNN能够学习合并、化简、典型化等建筑物综合算子,但多算子选择、协同效果并不理想。因此,通过有选择地学习综合算子以简化待学习的综合任务能进一步优化综合效果。

3)样本是影响综合效果的重要因素,实验结果表明利用包含道路网、水系等信息量更丰富的样本训练的建筑物综合模型效果更好。因此,通过增加样本量、丰富样本多样性、增加样本属性特征、完善样本包含空间数据的完整性和连续性等能进一步优化综合效果。

4)先验知识是影响综合效果的客观因素,实验结果表明,单纯依靠E-DDNN预测的建筑物综合结果可能出现与建筑物综合、表达规律不一致的问题。通过在E-DDNN中引入必要的先验知识,或对综合结果进行滤波、去噪、直角化等后处理,使其与先验知识相一致等,都能进一步优化综合效果。

4 结 语

本文基于空间分割和矢量-栅格数据转换思路构建了建筑物综合样本集,利用多种E-DDNN对建筑物综合进行学习模拟,实验测试了EDnet、Unet、ResUnet、Unet++、Pix2Pix 5种E-DDNN在建筑物综合中的应用效果,得出如下结论:结构相对简单、深度较浅的E-DDNN难以直接用于模拟建筑物综合这一复杂的图像变换,增加深度、优化结构后的E-DDNN可用于建筑物综合的学习模拟;E-DDNN通过训练能够从建筑物综合成果数据中学习部分综合知识和综合操作,但预测结果与实际居民地综合要求仍有差距;实验测试的5种模型中,Pix2Pix表现突出,更适用于建筑物综合研究。此基础上通过优化模型、调整超参数、丰富样本等方式提升建筑物综合准确性的研究有待进一步展开。

参 考 文 献

- [1] Alexander K. Trust Me, I'm a Cartographer: Post-Truth and the Problem of Acritical Cartography [J]. *The Cartographic Journal*, 2017, 54 (3):193-195
- [2] Wu Fang, Gong Xianyong, Du Jiawei. Overview of the Research Progress in Automated Map Generalization [J]. *Acta Geodaetica et Cartographica Sinica*, 2017, 46(10): 1645-1664 (武芳, 巩现勇, 杜佳威. 地图制图综合回顾与展望[J]. *测绘学报*, 2017, 46 (10): 1645-1664)
- [3] Wang Jiayao, Fan Yi' ai, Han Tongchun, et al. *Cartographic Generalization Theory of General Map* [M]. Beijing: Surveying and Mapping Press, 1993 (王家耀, 范亦爱, 韩同春, 等. 普通地图制图综合原理[M]. 北京:测绘出版社, 1993)
- [4] Wang Jiayao, Tian Zhen. The Cartographic Generalization of Soundings on Chart by Artificial Neural Network Techniques [J]. *Acta Geodaetica et Cartographica Sinica*, 1999, 28(4): 335-339 (王家耀, 田震. 海图水深综合的人工神经网络方法[J]. *测绘学报*, 1999, 28(4): 335-339)
- [5] Shao Lixia, He Zongyi, Ai Zixing, et al. Automatic Generalization of River Network Based on BP Neural Network Techniques [J]. *Geomatics and Information Science of Wuhan University*, 2004, 29(6): 555-557 (邵黎霞, 何宗宜, 艾自兴, 等. 基于BP神经网络的河系自动综合研究[J]. *武汉大学学报·信息科学版*, 2004, 29(6): 555-557)
- [6] Xie Limin, Qian Haizhong, He Haiwei, et al. A Habitation Selection Method by Using Case-Based Reasoning [J]. *Acta Geodaetica et Cartographica Sinica*, 2017, 46(11): 1910-1918 (谢丽敏, 钱海忠, 何海威, 等. 基于案例推理的居民地选取方法[J]. *测绘学报*, 2017, 46(11): 1910-1918)
- [7] Lee J, Jang H, Yang J, et al. Machine Learning Classification of Buildings for Map Generalization [J]. *ISPRS International Journal of Geo-Information*, 2017, 6(10): 309-324
- [8] Zhou Q, Li Z. A Comparative Study of Various Supervised Learning Approaches to Selective Omission in a Road Network [J]. *The Cartographic Journal*, 2017, 54(3): 254-264
- [9] Cheng Boyan. *Research on ANN-based Map Intelligent Generalization for Buildings* [D]. Chengdu: University of Electronic Science and Technology of China, 2014 (程博艳. 基于神经网络的地图建筑物要素智能综合研究[D]. 成都:电子科技大学, 2014)
- [10] Zhang L, Deng H, Chen D, et al. A Spatial Cognition-Based Urban Building Clustering Approach and

- Its Applications [J]. *International Journal of Geographical Information Science*, 2013, 27(4): 721-740
- [11] Allouche M K, Moulin B. Amalgamation in Cartographic Generalization Using Kohonen's Feature Nets [J]. *International Journal of Geographical Information Science*, 2005, 19(8/9): 899-914
- [12] Wu Fang, Deng Hongyan. Using Genetic Algorithms for Solving Problems in Automated Line Simplification [J]. *Acta Geodaetica et Cartographica Sinica*, 2003, 32(4): 349-355 (武芳, 邓红艳. 基于遗传算法的线要素自动化简模型[J]. 测绘学报, 2003, 32(4): 349-355)
- [13] Zheng Chunyan, Guo Qingsheng, Hu Huake. The Simplification Model of Linear Objects Based on Ant Colony Optimization Algorithm [J]. *Acta Geodaetica et Cartographica Sinica*, 2011, 40(5): 635-638 (郑春燕, 郭庆胜, 胡华科. 基于蚁群优化算法的线状目标简化模型[J]. 测绘学报, 2011, 40(5): 635-638)
- [14] He Haiwei, Qian Haizhong, Duan Peixiang, et al. Automatic Line Simplification Algorithm Selecting and Parameter Setting Based on Case-Based Reasoning [J]. *Geomatics and Information Science of Wuhan University*, 2020, 45(3): 344-352 (何海威, 钱海忠, 段佩祥, 等. 线要素化简及参数自动设置的案例推理方法[J]. 武汉大学学报·信息科学版, 2020, 45(3): 344-352)
- [15] Wilson I D, Ware M. Reducing Graphic Conflict in Scale Reduced Maps Using Genetic Algorithm [C]// The 7th ICA Workshop on Progress in Automated Map Generalization, Paris, France, 2003
- [16] Sun Yageng, Guo Qingsheng, Liu Yuangang, et al. A Real-Coded Genetic Algorithm Considering Gestalt Principles to Building Displacement [J]. *Geomatics and Information Science of Wuhan University*, 2015, 40(2): 269-273 (孙雅庚, 郭庆胜, 刘远刚, 等. 顾及格式塔原则的建筑物群移位实数编码遗传算法[J]. 武汉大学学报·信息科学版, 2015, 40(2): 269-273)
- [17] Kang Y, Gao S, Roth R E. Transferring Multiscale Map Styles Using Generative Adversarial Networks [J]. *International Journal of Cartography*, 2019, 5(2/3): 115-141
- [18] Ma Lei. A Simplification Model for Building Shapes Based on Machine Learning [D]. Lanzhou: Lanzhou Jiaotong University, 2018 (马磊. 基于机器学习的建筑物形状化简模型[D]. 兰州: 兰州交通大学, 2018)
- [19] Touya G, Zhang X, Lokhat I. Is Deep Learning the New Agent for Map Generalization [J]. *International Journal of Cartography*, 2019, 5(2/3): 142-157
- [20] Sester M, Feng Y, Thiemann F. Building Generalization Using Deep Learning [J]. *International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences*, 2018, 17(4): 565-572
- [21] Feng Y, Thiemann F, Sester M. Learning Cartographic Building Generalization with Deep Convolutional Neural Networks [J]. *ISPRS International Journal of Geo-Information*, 2019, 8(6): 258-278
- [22] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation [C]// Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 2015
- [23] Zhou Z, Siddiquee M M R, Tajbakhsh N, et al. U-Net++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation [J]. *IEEE Transactions on Medical Imaging*, 2020, 39(6): 1856-1867
- [24] Isola P, Zhu J, Zhou T, et al. Image-to-Image Translation with Conditional Adversarial Networks [C]// IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 2017
- [25] Goodfellow I, Bengio Y, Courville A. Deep Learning [M]. Massachusetts: MIT Press, 2016

Trial and Comparison of Some Encoder-Decoder Based Deep Learning Models for Automated Generalization of Buildings

DU Jiawei¹ WU Fang¹ XING Ruixing¹ LI Caixia¹ LI Jinghan²

¹ Institute of Geographical Spatial Information, Information Engineering University, Zhengzhou 450001, China

² Troops 61363, Xi'an 710054, China

Abstract: Objectives: The deep learning technology has made a breakthrough in many fields, and the automated map generalization based on the deep neural network is highly expected. Building generalization is a

typical generalization case and can be abstracted and interpreted as the process of encoding and decoding. It is worth systematically studying the simulation of building generalization with encoder-decoder based deep neural networks. **Methods:** Samples are drawn from vector buildings via the segmentation of spatial data and the conversion between vector data and raster data, which is based on a natural principle of objective generalization (i.e., the legibility principle drawn by scale). The encoder-decoder based deep neural networks are trained and tested with these samples to learn buildings generalization and to evaluate this application. Five representative models based on encoder-decoder structure are constructed for learning of generalizing buildings from 1 : 10 000 to 1 : 50 000. The experimental data cover 44.1 km \times 38.6 km and contain 89 539 and 21 732 buildings before and after generalization respectively. Five representative models are EDnet, Unet, ResUnet, Unet++ and Pix2Pix representing the basic multi-layers encoder-decoder neural network, the encoder-decoder neural network with skip connection, the residual encoder-decoder neural network, the dense encoder-decoder neural network, and the encoder-decoder neural network with generative adversarial mechanism respectively. The training process and test results of the three constructed models are compared, and effects of these five models while they are applied to the automated generalization of buildings are analyzed. **Results:** The pixel accuracies of predicted results during EDnet training and testing are lowest obviously, Unet takes the second lowest place, and the pixel accuracies of results generated by other models (ResUnet, Unet++ and Pix2Pix) are higher and relatively close. Additionally, there are some holes and fuzzy representations in generalized buildings predicted by ResUnet and Unet++, which does not conform with the regular representation of buildings. Generalization results predicted by the Pix2Pix show that most of generalized building edges are clear and generalized buildings can be represented independently and completely. **Conclusions:** The Pix2Pix is more suitable than the other four models for learning and simulating buildings generalization in our experiment. The encoder-decoder based deep neural networks are likely to deduce some knowledge and operators of buildings generalization by learning from samples of the original and generalized data. The encoder-decoder based deep neural networks have the potential in the learning of map generalization.

Key words: map generalization; buildings; deep neural network; samples

First author: DU Jiawei, PhD, specializes in automated cartographic generalization and intelligent spatial data processing. E-mail: whxd-jw@126.com

Corresponding author: WU Fang, PhD, professor. E-mail: wufang_630@126.com

Foundation support: The National Natural Science Foundation of China (41471386, 41801396).

引文格式: DU Jiawei, WU Fang, XING Ruixing, et al. Trial and Comparison of Some Encoder-Decoder Based Deep Learning Models for Automated Generalization of Buildings[J]. Geomatics and Information Science of Wuhan University, 2022, 47(7):1052-1062. DOI:10.13203/j.whugis20200143(杜佳威, 武芳, 行瑞星, 等. 几种具有编解码结构的深度学习模型在建筑物综合中的应用与比较[J]. 武汉大学学报·信息科学版, 2022, 47(7):1052-1062. DOI:10.13203/j.whugis20200143)