

基于ResNet和RF-Net的遥感影像匹配

廖明哲, 吴谨, 朱磊

引用本文:

廖明哲, 吴谨, 朱磊. 基于ResNet和RF-Net的遥感影像匹配[J]. 液晶与显示, 2020, 35(9): 972-980.

LIAO Ming-zhe, WU Jin, ZHU Lei. Remote sensing images matching based on ResNet and RF-Net[J]. *Chinese Journal of Liquid Crystals and Displays*, 2020, 35(9): 972-980.

在线阅读 View online: <https://doi.org/10.37188/YJYXS20203509.0972>

您可能感兴趣的其他文章

Articles you may be interested in

一种用于图像拼接的改进ORB算法

Improved ORB algorithm used in image stitching

液晶与显示. 2018, 33(6): 520-527 <https://doi.org/10.3788/YJYXS20183306.0520>

一种改进的SURF彩色遥感图像配准算法

Improved SURF algorithm for color remote sensing image registration

液晶与显示. 2017, 32(2): 144-152 <https://doi.org/10.3788/YJYXS20173202.0144>

结合 MSER 与 HSOG 的目标局部特征提取

Target local feature extraction combined MSER and HSOG

液晶与显示. 2016, 31(11): 1070-1078 <https://doi.org/10.3788/YJYXS20163111.1070>

一种多特征相结合的三维人脸关键点检测方法

A key point detection method of 3D face based on multi-feature

液晶与显示. 2018, 33(4): 306-316 <https://doi.org/10.3788/YJYXS20183304.0306>

基于谱残差和梯度纹理融合特征的舰船检测

Ship detection based on spectral residual and gradient texture fusion features

液晶与显示. 2019, 34(8): 803-815 <https://doi.org/10.3788/YJYXS20193408.0803>

文章编号:1007-2780(2020)09-0972-09

基于 ResNet 和 RF-Net 的遥感影像匹配

廖明哲^{1,2}, 吴 谨^{1,2*}, 朱 磊^{1,2}

(1. 武汉科技大学 信息科学与工程学院, 湖北 武汉 430081;

2. 冶金自动化与检测技术教育部工程中心, 湖北 武汉 430000)

摘要:针对 Receptive Fields Network(RF-Net)中存在网络较浅、缺乏深层语义信息的问题,提出了一种基于 Residual Network(ResNet)和 RF-Net 的改进网络用于遥感影像匹配。首先,通过对真实遥感影像进行裁剪、光照变换和仿射变换处理,得到图像对并计算同一序列中不同图像间的单应性矩阵,构建了一个遥感影像数据集。然后,提出了一种双通道的网络结构用于关键点检测,该双通道网络由 Receptive Fields Detection(RF-Det)和 ResNet 构成,前者提取含有细节信息的浅层特征图,后者提取含有语义信息的深层特征图。此外,采用特征描述子提取网络 L2-Net,得到 128 维特征向量用以描述关键点。最后,分别采用最近邻、带阈值的最近邻和最近邻距离比的策略对特征描述子进行匹配。实验结果表明,该网络在仅含光照变换、仅含仿射变换和同时包含这两种变换的遥感影像数据集上的匹配得分,比 RF-Net 分别提高了 0.002, 0.117, 0.104, 在关键点检测和匹配精度方面具有更优的性能。

关键词: 图像匹配; 关键点检测; 特征描述子提取; 仿射变换; 特征向量

中图分类号: TP751.1 **文献标识码:** A **doi:** 10.37188/YJYXS20203509.0972

Remote sensing images matching based on ResNet and RF-Net

LIAO Ming-zhe^{1,2}, WU Jin^{1,2*}, ZHU Lei^{1,2}

(1. School of Information Science and Engineering,

Wuhan University of Science and Technology, Wuhan 430081, China;

2. Engineering Research Center of Metallurgical Automation and Measurement Technology,
Ministry of Education, Wuhan 430000, China)

Abstract: Aiming at the problems of shallow network and lack of high-level semantic information in Receptive Fields Network(RF-Net), a new network based on Residual Network(ResNet) and RF-Net is proposed for the matching of remote sensing images. Firstly, the new remote sensing image datasets are setup, which consist of images and homography matrices. The images are obtained by cropping, illumination changing and affine transforming of the original remote sensing images. The matrices are obtained by calculating the homography between different images of one sequence. Secondly, a dual-channel network structure is proposed for the keypoints detection. Specifically, the dual-channel network composed by Receptive Fields Detection(RF-Det) and ResNet, respectively extracts

收稿日期: 2020-02-16; **修订日期:** 2020-05-05.

基金项目: 国家自然科学基金青年基金(No.61502358, No.61702384)

Supported by the Youth Fund Project of National Natural Science Foundation of China(No.61502358, No. 61702384)

* 通信联系人, E-mail: wujin@wust.edu.cn

receptive feature maps with detail information, and the deep layer feature maps with semantic information. The L2-Net is introduced at descriptor extraction module to obtain the 128-dimension feature vectors as description of keypoints. Lastly, the strategies of the nearest neighbor, nearest neighbor with a threshold and nearest neighbor distance ration are used to match the descriptors. Experiments on remote sensing images show that the matching scores of the proposed network are 0.002, 0.117, 0.104, higher than RF-Net, when testing on different datasets. It outperforms on keypoints detection and matching accuracy.

Key words: image matching; keypoints detection; feature descriptor extraction; affine transform; feature vectors

1 引言

遥感影像包含来自卫星、地面和航空等不同来源的图像数据,通过对这些图像进行匹配,可以建立起不同图像间的空间对应关系,为检测地球表面环境变化,应用于地质勘测和军事侦察等领域提供服务。

图像匹配使来自相同或者不同传感器、具有不同视角和不同时相的多幅同一场景图像,在空间上实现几何对齐^[1]。图像匹配的常用方法可以分为3类:基于灰度和模板的方法、基于域变换的方法和基于特征的方法。其中,基于特征的匹配方法鲁棒性好,适用于较多的应用场景^[2]。典型的基于特征的方法包括两部分:带有属性的关键点检测和特征描述子提取。

基于特征的深度学习图像匹配是目前非常重要的匹配技术,在深度学习图像匹配框架中,有些方法仅用于检测关键点,有些方法仅用于提取特征描述子。关键点检测部分构建响应图,特征信息丰富的响应图有利于检测到更多的关键点;特征描述子提取部分端到端地训练特征描述子^[3],特征描述子是用于描述图像中某个关键点的特征向量。关键点检测和特征描述子提取基于不同目标进行优化,两个部分的增益无法直接叠加,在同一个通道中同时训练这两个部分难以取得更好的效果^[4]。因此,如何联合训练关键点检测和特征描述子提取,使两个部分彼此更好地协作,是深度学习图像匹配需要解决的问题。

LIFT^[5]是最早实现关键点检测和特征描述子提取联合训练的网络之一,该网络以尺度不变特征变换(Scale Invariant Feature Transform, SIFT)^[6]中特征点所在的图像块作为输入,其关

键点检测的效果与 SIFT 算法类似,鲁棒性较好,但无法准确提取边缘光滑目标的特征点。与 LIFT 方法不同,SuperPoint^[7]采用了自监督训练的特征点检测和特征描述子提取方式,通过 VGG^[8]提取特征,关键点检测部分需要在合成图像数据集上进行预训练,且整个网络也需要在合成变换后的图像上进行训练。LF-Net^[4]使用 Siamese 结构^[9],无需任何手工方法的帮助,其通过深层特征提取网络产生特征图,可从输入图像中提取出感受野(特征图上像素点在输入图像上映射的区域)较大的深层特征,但损失了浅层特征。RF-Net^[10]在 LF-Net 的基础上提出了一个新的基于感受野的关键点提取方法,保留了浅层特征,得到信息更丰富的尺度空间和响应图,在特征描述子提取模块上采用了与 Hard-Net^[11]一致的网络结构,并提出了一个通用的损失函数项,用以解决图像的刚性变换后,像素点位置偏移所造成的负面影响。

网络的深度对于学习表征能力更强的特征至关重要,深层特征包含有丰富的语义信息,但随着网络深度的增加,会造成梯度消失或者梯度爆炸,以及网络退化的问题。ResNet^[12]为当前应用最为广泛的 CNN 特征提取网络之一,为了解决网络的退化问题,引入了基本的残差学习;为了克服梯度消失和梯度爆炸的问题,使用了批归一化^[13],并将激活函数置换为线性整流函数^[14]Rectified Linear Unit(ReLU)。具有强表征能力的深层特征可以更好地定位显著性目标的关键点位置,由 ResNet 获取的深层特征图有助于关键点的检测。

RF-Net 采用公共数据集训练,用于遥感影像匹配时只能检测到较少的关键点,误匹配率高;且关键点检测模块中网络较浅,感受野受限,缺乏高

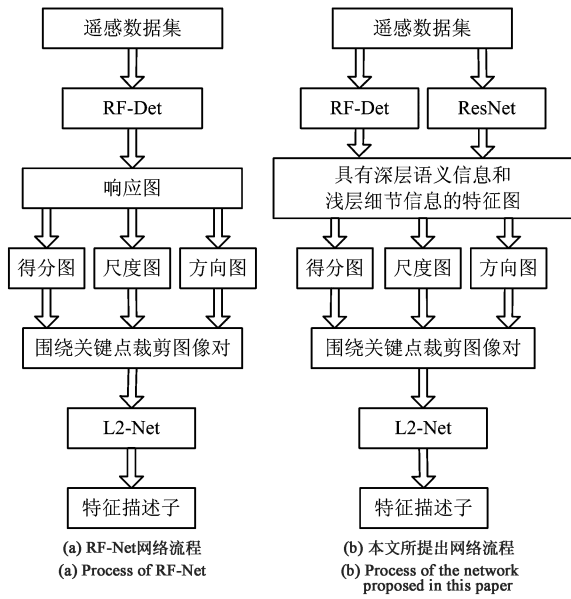


图 1 网络流程示意图

Fig.1 Diagram of network flow

层语义信息。本文基于 ResNet 和 RF-Net 提出了一个新的双通道网络,用于检测图像关键点和提取关键点的特征描述子。RF-Net 的网络流程如图 1(a)所示,基于 ResNet 和 RF-Net 的双通道网络流程如图 1(b)所示。通过对真实遥感影像进行光照变换、仿射变换,以及计算不同影像间的单应性矩阵,构建遥感影像匹配数据集,并采用该数据集对网络进行训练和测试。

2 ResNet

ResNet 提出残差网络结构,采用批归一化,并将激活函数设置为 ReLU,在增加网络深度的同时,避免了因梯度消失或爆炸带来的性能下降。本文所提出的改进网络通过 ResNet 获取带有丰富语义信息的深层特征,能更好地定位图像中显著性目标的关键点位置。

2.1 残差学习

随着网络层数的增加,整体模型的表征能力更强,但训练精度反而下降,出现网络退化问题。残差学习可用于解决深层网络普遍存在的退化问题,残差函数的表达式为:

$$F(x) = H(x) - x, \quad (1)$$

式中, x 为网络的输入, $H(x)$ 为多个非线性网络层拟合成的基础映射。

基本的残差单元可表示为:

$$y_i = x_i + F(x_i, w_i), \quad (2)$$

$$x_{i+1} = g(y_i), \quad (3)$$

式中, x_i 、 x_{i+1} 和 w_i 分别表示第 i 个残差块的输入、输出和权重, F 表示残差函数, g 表示 ReLU 激活函数。

由上式可得,网络层数加深的过程可以表示为:

$$x_L = x_i + \sum_{j=i}^{L-1} F(x_j, w_j). \quad (4)$$

通过损失函数进行反向传播,从而更新参数。

如果损失函数用 Loss 表示,则有:

$$\frac{\partial \text{Loss}}{\partial x_i} = \frac{\partial \text{Loss}}{\partial x_L} * \frac{\partial x_L}{\partial x_i} = \frac{\partial \text{Loss}}{\partial x_L} (1 + \frac{\partial F(x)}{\partial x_i}). \quad (5)$$

综上,当 $F(x) = 0$ 时,叠加的非线性网络层被构造为恒等映射,通过网络学习到的特征与网络输入类似。在实际网络训练环境下,即 $F(x) \neq 0$ 时,网络层学习到新的特征,性能更优。

2.2 基于短接的恒等映射

由多个非线性网络层组成的结构,在输入和输出间引入一条直接的关联通道,使得网络参数层集中学习残差,如图 2 所示。残差块可表示为:

$$Y = g[F(x, \{w_i\}) + x], \quad (6)$$

式中, x 为网络输入, Y 为网络输出。 $F(x, \{w_i\})$ 表示所学习的残差映射, g 表示 ReLU 激活函数。

残差块的组成较为灵活。图 3(a)为基本残差块结构,由两个网络层组成;图 3(b)为瓶颈结

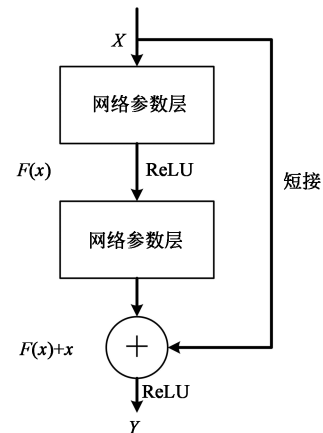


图 2 残差结构

Fig.2 Structure of ResNet

构,由3个网络层组成,分别对应 1×1 、 3×3 、 1×1 卷积,其无参数的恒等短接减少了计算参数,可更有效地训练网络。

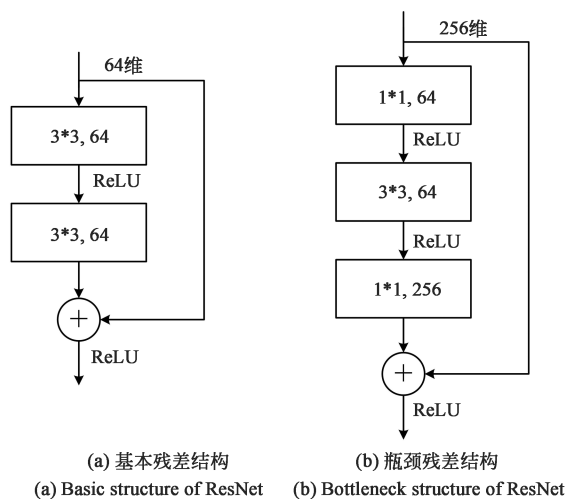


图3 残差结构
Fig.3 Structure of ResNet

以图3(a)为例,该残差块由2个网络层组成,即 $F = w_2 \sigma(w_1 x)$, σ 表示ReLU激活函数,为了简化结构省略了偏差项。 $F + x$ 由短接的加法来表示,加法后再进行另一非线性操作。

F 与 x 的维度必须保持一致,可改变 x 的维度以匹配 F 。

$$y = F(x, \{w_i\}) + w_s x, \quad (7)$$

式中, w_s 表示第 s 个残差块的权重。

本文采用ResNet-50的网络结构用以提取深层特征,通过图3(b)中的瓶颈结构搭建整个网络,其中输出维度为256,512,1 024,2 048的瓶颈结构的个数分别为3,4,6,5。

3 基于 ResNet 和 RF-Net 的网络

基于ResNet和RF-Net的网络由关键点检测和特征描述子提取两个部分构成。关键点检测部分,构建了由两组卷积神经网络融合的双通道关键点检测网络,通过深层网络通道获取包含丰富语义信息的高层特征,并与浅层网络通道中感受野递增的低层特征融合;特征描述子提取部分采用L2-Net^[15],对输入的图像对进行训练,提取特征描述子。整个网络的结构如图4所示。在训练期间,关键点网络输入数据集影像,输出得分图 S 、方向图 Θ 和尺度图 $\bar{S}$,分别代表关键点的空间位置、方向属性以及尺度信息。通过这些图截取关键点图像对,并将其送入特征描述子网络,生成

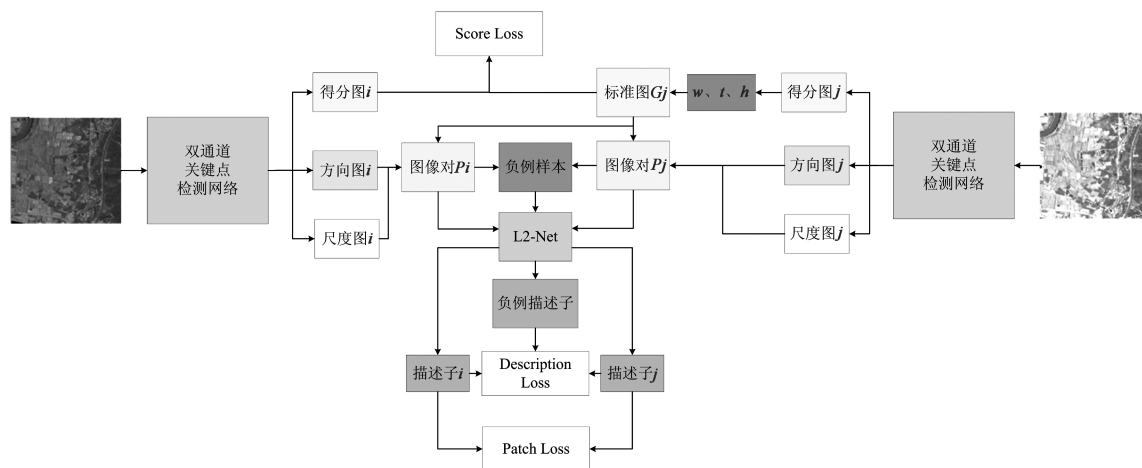


图4 基于双通道的网络框架
Fig.4 Network structure of dual-channel

固定长度的特征向量,用于匹配。

3.1 双通道网络结构

本文的关键点检测部分,采用双通道的方式构建网络,两个通道分别为RF-Det感受野递增的浅层特征提取通道和ResNet的深层全局信息

特征提取通道。前者提取的浅层特征具有丰富的细节信息,后者提取的深层特征包含更具代表性的全局信息。连接两种不同的特征,提取出最具视觉区分度的特征图,如图5所示。第一个通道由 3×3 卷积、实例归一化(Instance Normalization)

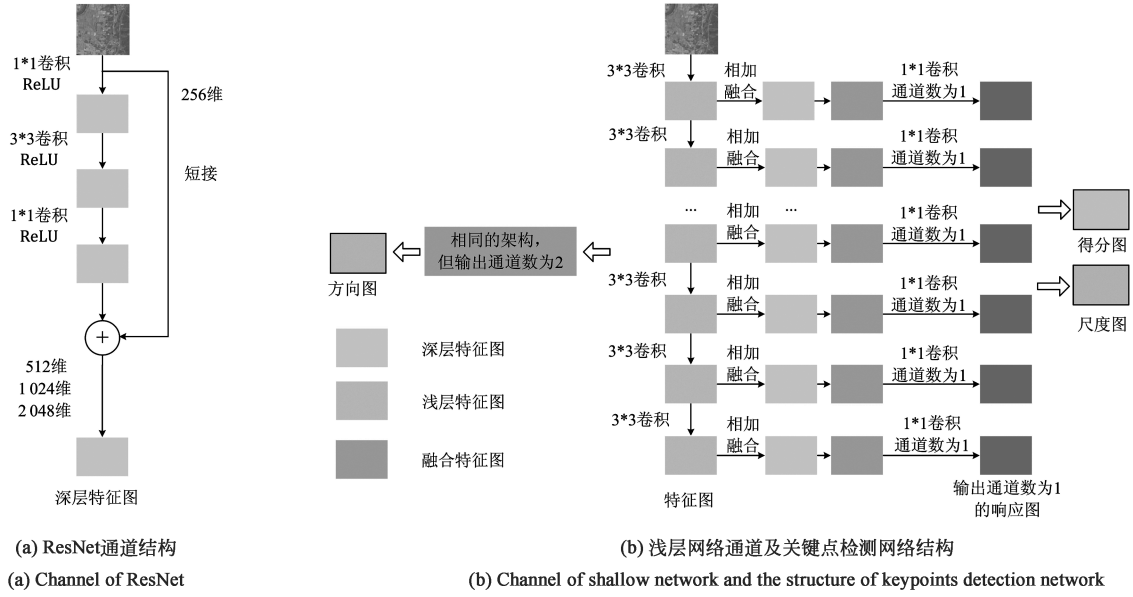


图 5 双通道网络结构

Fig.5 Structure of dual-channel network

正则函数和 ReLU 激活函数通过特征金字塔 FPN (Feature Pyramid Networks)^[16] 的卷积方式构成, 第二个通道选用 ResNet-50。浅层特征和深层特征融合后, 对侧边输出使用 1×1 卷积和实例正则化函数 (Instance Normalization) 以生成所需的多尺度响应图。本文提出的改进网络得到的响应图所需感受野较小, 但包含丰富的深层语义信息和浅层细节信息, 有助于检测到更多显著性目标的关键点。

3.2 关键点检测

选择高响应像素作为关键点, 响应图 h_n 表示多个尺度上的像素响应, 本文设计了一个类似于 RF-Net 和 LF-Net 的关键点检测结构。

分别进行两次 Softmax 逻辑回归以产生得分图。第一次 Softmax 采用 $15 \times 15 \times N$ 的滑动窗口来产生不失真的响应图 $\hat{h}_n$, 随后通过第二次 Softmax 将所有的 $\hat{h}_n$ 融合成最后的得分图 S 。

$$P_n = \text{Softmax}(\hat{h}_n) \quad , \quad (8)$$

$$S = \sum_n \hat{h}_n \odot P_n \quad , \quad (9)$$

式中, $\odot$ 为哈达码矩阵, P_n 表示像素点为关键点的可能性^[10]。

对于方向图 $\{\theta_n\}$, 其值分别表示方向的 sine 和 cosine, 通过 $\arctan$ 函数计算角度。同样, 将 θ_n 融合成最后的方向图 Θ 。

$$\Theta = \sum_n \theta_n \odot P_n \quad , \quad (10)$$

同样的操作也用于生成尺度图 $\bar{S}$ 。

$$\bar{S} = \sum_n \bar{S}_n \odot P_n \quad , \quad (11)$$

式中, $\bar{S}_n$ 表示 h_n 的感受野范围。

3.3 特征描述子提取网络

本文采用 L2-Net 用于提取特征描述子。该网络由 6 个 3×3 卷积层和 1 个 8×8 卷积层组成, 在卷积层后都会加上 ReLU 激活函数, 并进行批归一化。描述子提取网络最后一个卷积层的滤波器尺寸过大, 对网络会造成负担, 但可以获取包含丰富信息的 128 维张量, 并转化为特征向量用以描述关键点。

3.4 损失函数

关键点检测网络对关键点的位置、方向和尺度进行预测, 其损失函数由得分图损失 (Score Loss) 和图像对损失 (Patch Loss) 组成。

得分图损失是指将图像对 I_i 和 I_j 输入网络得到得分图 S_i 和 S_j , 通过 S_j 产生参考标准图像 (Ground Truth, GT) G_j , 计算 S_i 和 G_j 间的均方误差 (Mean Squared Error, MSE)。 G_j 是指从形变后的 S_j 中提取 K 个关键点, 并使用高斯卷积 ($\sigma = 0.5$) 得到干净的 GT。

$$G_j = g[t\{w(S_j)\}] \quad , \quad (12)$$

式中, w 、 t 、 g 分别表示形变, 选取关键点和高斯卷积的过程, 如图 4 所示。得分图损失的计算公

式如下:

$$L_{\text{score-loss}}(S_i, G_j) = |S_i - G_j|^2 \quad (13)$$

图像对损失用于优化检测网络,使相关响应区域剪裁下的图像对尽可能相似。从 G_j 中选择 K 个关键点,然后将其空间坐标形变至输入图像 I_j ,关键点由空间位置坐标以及预测得到的方向和尺度信息组成。将响应的图像对 $p_i^k, \hat{p}_j^k$ 送入特征描述子提取网络,产生 $D_i^k, \hat{D}_j^k$ 。则有:

$$L_{\text{patch-loss}} = \frac{1}{K} \sum_{k=1}^K \sqrt{2 - 2 D_i^k \hat{D}_j^k} \quad (14)$$

综上,关键点检测网络损失函数为:

$$L_{\text{det}} = L_{\text{score-loss}} + L_{\text{patch-loss}} \quad (15)$$

描述子损失 (Description Loss) 从 Hard-Net^[11] 中引入特征描述子损失函数 L_{des} ,用于最大化距离最近正例和距离最近反例之间的距离,使特征描述子训练更加稳定。

$$L_{\text{des}}(\tau_{\text{pos}}, \tau_{\text{ng}}) = \frac{1}{K} \sum_{k=1}^K \max(0, 1 + \tau_{\text{pos}} - \tau_{\text{ng}}) \quad (16)$$

$$\tau_{\text{pos}}(D_i^k, \hat{D}_j^k) = d(D_i^k, \hat{D}_j^k) = \sqrt{2 - 2 D_i^k \hat{D}_j^k} \quad (17)$$

$$\tau_{\text{ng}} = \min[d(D_i^k, \hat{D}_j^n), d(D_i^m, \hat{D}_j^k)] \quad (18)$$

式中, $\hat{D}_j^n$ 是 D_i^k 最近的非匹配特征描述子, D_i^m 是 $\hat{D}_j^k$ 最近的非匹配特征描述子。

4 实验结果

4.1 训练数据以及评估标准

针对现有匹配网络缺少遥感领域数据进行训练,造成在遥感影像上测试效果较差的问题,本文构建了一个遥感影像数据集(Remote-B),对真实遥感影像进行裁剪、光照变换和仿射变换,生成了包括亮度变化和视点变化的另外 9 幅图像,并计算原始图像和每幅生成图像之间的单应性(位置关系)矩阵,所有图像和单应性矩阵构成了一个序列。数据集 Remote-B 由 56 个不同的序列、560 幅影像组成,数据集部分图像如图 6 所示。此外,同时构建了仅含光照变换的 42 个序列、294 幅影像的数据集(Remote-illu),以及仅含仿射变换的

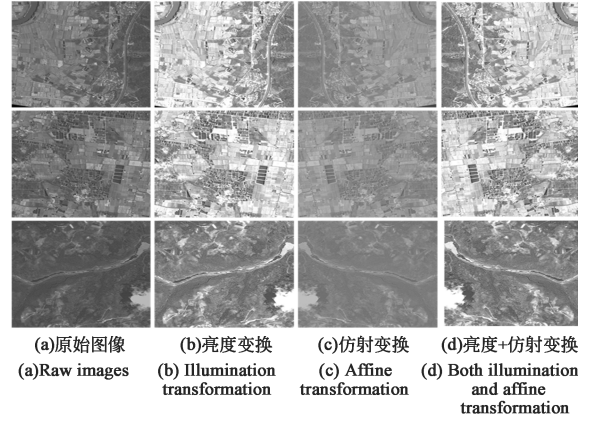


图 6 遥感图像数据集部分图像

Fig.6 Part of images in remote sensing images datasets

48 个序列、336 幅影像的数据集(Remote-view)用于测试。

匹配的标准取决于匹配策略,本文分别采用 3 种匹配策略,计算匹配分数,用以定量评估^[17]。

策略一:最近邻 NN(Nearest Neighbor)。在该标准下,每个描述子只能有一个匹配。两个区域 A 和 B ,当且仅当其描述子 D_B 和 D_A 为最近邻描述子, A 与 B 匹配;

策略二:带阈值的最近邻 NNT(Nearest Neighbor with a Threshold)。两个区域 A 和 B ,当其描述子 D_B 和 D_A 为最近邻描述子且两者之间的距离小于阈值 t 时, A 与 B 匹配;

策略三:最近邻距离比 NNR(Nearest Neighbor distance Ratio)。两个区域 A 和 B , D_B 和 D_C 分别是 D_A 的最近邻和次近邻描述子,当 $\|D_A - D_B\| / \|D_A - D_C\| < t$ 时, A 与 B 匹配。

这 3 种匹配策略用以衡量匹配的精度和关键点数量。在训练过程中,所有学习过的描述子均采用 L2 正则化,其距离范围是 $[0, 2]$ 。NNT 和 NNR 的阈值 t 分别设置为 1 和 0.7。

4.2 训练过程

本文在构建的遥感影像数据集 Remote-B 上进行训练和测试,将数据集按 9:1 的比例划分为训练集和测试集,500 幅遥感影像用于训练,60 幅遥感影像用于测试。

在训练时,将数据集中遥感影像尺度缩放,修改至 320×240 ,并进行灰度变换。对于描述子提取网络部分,围绕关键点裁剪 32×32 尺寸大小影像输入网络训练。在训练阶段,提取 $K = 512$ 个

关键点,但是在测试阶段,可以任选所需的关键点数量,通过描述子网络,得到 512 个 128 维特征向量。采用适应性矩估计 Adam(Adaptive moment estimation)^[18]进行优化,初始学习率设置为 0.1,同时训练两次描述子网络和一次关键点检测网络。

4.3 匹配结果

分别按照 SIFT 算法、RF-Net 和本文改进网络,对两幅场景相同、仿射变换不同的遥感影像进行匹配,图像间关键点的正确匹配如图 7 所示。

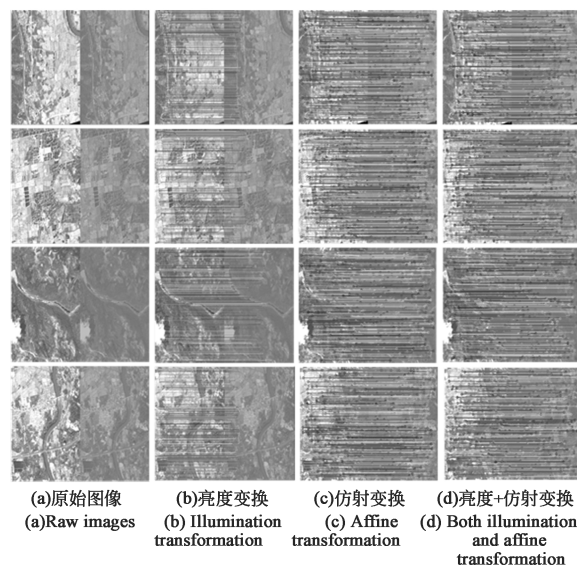


图 7 SIFT、RF-Net 及本文改进网络的匹配结果。

Fig.7 Matching result of SIFT, RF-Net, and the improved network proposed in this paper.

RF-Net 和本文改进网络的匹配效果明显优于传统的 SIFT 算法。一方面,从图中可以直观看到,RF-Net 和本文改进网络的关键点匹配对数明显多于 SIFT 算法,深度学习卷积神经网络具有强大的特征提取能力,有助于图像中关键点的检测;另一方面,在同一区域内,SIFT 算法的关键点更为密集,关键点距离过近会导致产生的图像对太相似而出现错误匹配,该结果可以验证稀疏关键点的匹配效果更好。

将 RF-Net 和本文提出的改进网络,在数据集 Remote-B、Remote-illu 和 Remote-view 上分别采用 NN、NNT 和 NNR 策略进行匹配。其匹配得分和平均匹配得分 (Mean Score) 分别如表 1、表 2 和表 3 所示。

表 1 Remote-B 测试结果定量分析

Tab.1 Test results on Remote-B

网络	策略			
	NN	NNT	NNR	Mean Score
RF-Net	0.134	0.431	0.372	0.312
本文方法	0.369	0.448	0.431	0.416

表 2 Remote-illu 测试结果定量分析

Tab.2 Test results on Remote-illu

网络	策略			
	NN	NNT	NNR	Mean Score
RF-Net	0.896	0.996	1.000	0.964
本文方法	0.903	0.996	1.000	0.966

表 3 Remote-view 测试结果定量分析

Tab.3 Test results on Remote-view

网络	策略			
	NN	NNT	NNR	Mean Score
RF-Net	0.289	0.647	0.654	0.530
本文方法	0.508	0.722	0.711	0.647

由表 1 可见,在 Remote-B 遥感数据集中,本文网络的平均得分比 RF-Net 高出 0.104,匹配精度更高。由表 2 可见,在 Remote-illu 数据集中,本文网络的匹配得分高出 0.002,在光照变换条件下本文提出的网络和 RF-Net 都取得了较高的匹配精度,但本文网络的匹配精度略优于 RF-Net。由表 3 可见,在 Remote-view 数据集中,本文网络的匹配得分比 RF-Net 高出 0.117,在仿射变换条件下匹配性能提升效果更为明显。

综上所述,本文所提出的改进网络比 RF-Net 具有更好的关键点提取和匹配精度性能,尤其具有更好的抗仿射变换鲁棒性能,可以更好地应用于遥感影像匹配。

5 结 论

遥感影像匹配可以将来自不同设备、不同时间和不同角度的影像数据联系起来,提取更为丰富的信息。基于深度学习的匹配方法,不仅可以提高图像匹配的准确率,也可以适应复杂应用场景的需

求,具有非常重要的研究意义。本文提出了一种基于 ResNet 和 RF-Net 的双通道检测网络,用以提取表征能力更强的深层特征,并融合感受野递增的浅层特征,提高了检测到的关键点数量和匹配的准确率。在与 RF-Net 的对比实验中可以看出,本文提

出的改进网络在遥感影像数据集 Remote-B、Remote-illu 和 Remote-view 上的匹配得分,分别为 0.416、0.966 和 0.647,对仿射变换具有更好的鲁棒性,在关键点检测和匹配精度方面性能更优,可以有效地应用于遥感图像的关键点匹配。

参 考 文 献:

- [1] ZITOVÁ B, FLUSSER J. Image registration methods: a survey [J]. *Image and Vision Computing*, 2003, 21(11): 977-1000.
- [2] 余先川,吕中华,胡丹.遥感图像配准技术综述[J].*光学精密工程*, 2013, 21(11): 2960-2972.
YU X C, LV Z H, HU D. Review of remote sensing image registration techniques [J]. *Optics and Precision Engineering*, 2013, 21(11): 2960-2972. (in Chinese)
- [3] 贾迪,朱宁丹,杨宁华,等.图像匹配方法研究综述[J].*中国图象图形学报*, 2019, 24(5): 677-699.
JIA D, ZHU N D, YANG N H, *et al.* Image matching methods [J]. *Journal of Image and Graphics*, 2019, 24(5): 677-699. (in Chinese)
- [4] ONO Y, TRULLS E, FUA P, *et al.* LF-Net: learning local features from images [C]//*Proceedings of the 32nd International Conference on Neural Information Processing Systems*. Montreal, Canada: ACM, 2018: 6237-6247.
- [5] YI K M, TRULLS E, LEPETIT V, *et al.* LIFT: Learned invariant feature transform [C]//*Proceedings of the 14th European Conference on Computer Vision*. Amsterdam, the Netherlands: Springer, 2016: 467-483.
- [6] LOWE D G. Distinctive image features from scale-invariant keypoints [J]. *International Journal of Computer Vision*, 2004, 60(2): 91-110.
- [7] DETONE D, MALISIEWICZ T, RABINOVICH A. SuperPoint: self-supervised interest point detection and description [C]//*Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. Salt Lake City, UT, USA: IEEE, 2018.
- [8] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [C]//*Proceedings of 3rd International Conference on Learning Representations*. San Diego, CA, USA: ICLR, 2015: 1-13.
- [9] CHOPRA S, HADSELL R, LECUN Y. Learning a similarity metric discriminatively, with application to face verification [C]//*Proceedings of 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. San Diego, CA, USA: IEEE, 2005: 539-546.
- [10] SHEN X L, WANG C, LI X, *et al.* RF-Net: an end-to-end image matching network based on receptive field [C]//*Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach, CA, USA: IEEE, 2019: 8124-8132.
- [11] MISHCHUK A, MISHKIN D, RADENOVIC F, *et al.* Working hard to know your neighbor's margins: Local descriptor learning loss [C]//*Proceedings of the 31st Conference on Neural Information Processing Systems*. Long Beach, CA, USA: NeurIPS, 2017: 4826-4837.
- [12] HE K M, ZHANG X Y, REN S Q, *et al.* Deep residual learning for image recognition [C]//*Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, Nevada, USA: IEEE, 2016: 770-778.
- [13] IOFFE S, SZEGEDY C. Batch normalization: accelerating deep network training by reducing internal covariate shift [C]//*Proceedings of the 32nd International Conference on Machine Learning*. Lille, France: ACM, 2015: 448-456.
- [14] NAIR V, HINTON G E. Rectified linear units improve restricted boltzmann machines [C]//*Proceedings of the 27th International Conference on Machine Learning*. Haifa, Israel: ACM, 2010: 807-814.
- [15] TIAN Y R, FAN B, WU F C. L2-Net: Deep learning of discriminative patch descriptor in euclidean space [C]//*Proceedings of the 30nd IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, Hawaii,

- USA: IEEE, 2017: 6128-6136.
- [16] LIN T Y, DOLLÁR P, GIRSHICK R, *et al.* Feature pyramid networks for object detection [C]//*Proceedings of the 30nd IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, Hawaii, USA: IEEE, 2017: 936-944.
- [17] MIKOLAJCZYK K, SCHMID C. A performance evaluation of local descriptors [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, 27(10): 1615-1630.
- [18] KINGMA D, BA J. Adam: a method for stochastic optimization [C]//*Proceedings of the 3rd International Conference on Learning Representations*. San Diego, CA, USA: ICLR, 2015: 1-15.

作者简介:



廖明哲(1995—),男,湖北仙桃人,硕士研究生,2017年于武汉科技大学获得学士学位,主要从事图像处理和模式识别方面的研究。E-mail: liaomingzhe1995@163.com



吴 谨(1967—),女,安徽芜湖人,博士,教授,博士生导师,2006年于华中科技大学获得博士学位,主要从事计算机视觉和模式识别方面的研究。E-mail: wujin@wust.edu.cn