



(12)发明专利申请

(10)申请公布号 CN 111325165 A

(43)申请公布日 2020.06.23

(21)申请号 202010118553.2

(22)申请日 2020.02.26

(71)申请人 中南大学

地址 410011 湖南省长沙市岳麓区麓山南路932号

(72)发明人 陶超 鲁伟鹏 齐霁 李海峰

(74)专利代理机构 北京律诚同业知识产权代理有限公司 11006

代理人 祁建国

(51) Int. Cl.

G06K 9/00(2006.01)

G06K 9/42(2006.01)

G06K 9/62(2006.01)

G06N 3/04(2006.01)

G06N 3/08(2006.01)

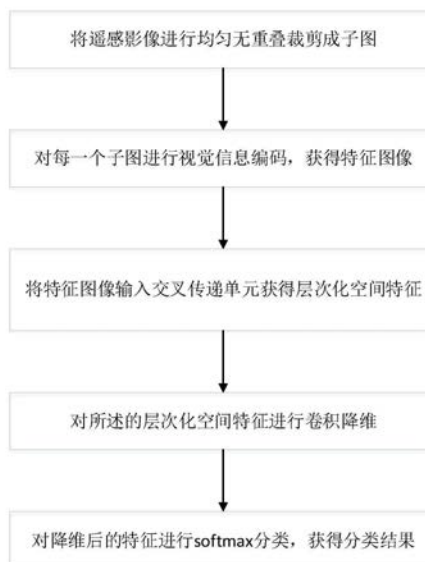
权利要求书1页 说明书5页 附图2页

(54)发明名称

考虑空间关系信息的城市遥感影像场景分类方法

(57)摘要

本发明公开了考虑空间关系信息的城市遥感影像场景分类方法,包括以下步骤:将遥感影像进行均匀无重叠裁剪成子图;对每一个子图进行视觉信息编码,获得特征图像 F_V ;将 F_V 输入交叉传递单元获得层次化空间特征;对所述的层次化空间特征进行卷积降维;对降维后的特征进行softmax模型分类,获得分类结果。本发明方法综合考虑两种空间关系信息:局部空间信息和长程空间信息在分类中的作用,在交叉传递单元中设计了三种路径的信息融合,从而获得了较好的城市遥感影像场景分类结果。



1. 考虑空间关系信息的城市遥感影像场景分类方法,其特征在于,包括以下步骤:

步骤1,将遥感影像进行均匀无重叠裁剪成子图;

步骤2,对每一个子图进行视觉信息编码,获得特征图像 F_V ;

步骤3,将特征图像 F_V 输入交叉传递单元获得层次化空间特征;

步骤4,对所述的层次化空间特征进行卷积降维;

步骤5,对降维后的特征进行softmax模型分类,获得分类结果;

步骤1中所述的裁剪过程为:尺寸 $M \times N$ 的大范围遥感影像 I 被均匀无重叠的滑动裁剪成 $m \times n$ 个子图,每个子图 $p_{i,j}$ 的尺寸为 $\frac{M}{n} \times \frac{N}{n}$,其中, $p_{i,j}$ 在 I 中的行列号 (i,j) 作为空间信息保存;

步骤2中所述的视觉信息编码,采用预训练好的深度卷积模型将对每一个子图 $p_{i,j}$ 进行视觉信息编码,将 $p_{i,j}$ 转化为一个向量 $f_{V_{i,j}}$,遥感影像 I 被转换为特征图像 F_V :

$$F_V = \begin{pmatrix} f_{V_{1,1}} & \cdots & f_{V_{1,n}} \\ \vdots & \ddots & \vdots \\ f_{V_{m,1}} & \cdots & f_{V_{m,n}} \end{pmatrix};$$

所述的交叉传递单元用于局部空间信息与长程空间信息的提取与融合,所述的局部空间信息的提取公式为: $F_R = \text{Conv}(F_V) = F_V * W + b$,其中, F_R 表示提取分析的空间关系信息, W 表示卷积核, b 表示偏置量,星号 $*$ 表示卷积计算,所述的长程空间信息采用基于循环神经网络的ReNet模块进行提取。

2. 根据权利要求1所述的的城市遥感影像场景分类方法,其特征在于,所述的交叉传递单元的输入为特征图像 F_V ,输出为层次化空间特征 F_E ,采用三条路径进行信息的提取分析与传递:第一条路径先对 F_V 进行局部空间信息的提取,而后进行长程空间信息的提取;第二条路径先对 F_V 先进行长程空间信息的提取,而后进行局部空间信息的提取;第三条路径则直接将 F_V 传递到交叉传递单元末端的捷径,不进行额外处理,由此,输出层次化空间特征 F_E 表示为:

$$F_E = \tanh(\text{ReNet}^2(\text{Conv}^1(F_V))) + \text{Conv}^2(\text{ReNet}^1(F_V)) + F_V$$

其中 \tanh 为双曲正切函数, $\text{ReNet}^1, \text{ReNet}^2$ 表示两个不同的ReNet模块, Conv^1 和 Conv^2 表示两个不同的卷积模块。

3. 根据权利要求1或2所述的的城市遥感影像场景分类方法,其特征在于,步骤3中所述的特征图像经过三个串联的交叉传递单元,得到的层次化空间特征记为 F_M ;

在步骤4中,采用尺寸为 1×1 的卷积层 $\text{Conv}_{1 \times 1}$ 对 F_M 进行卷积降维;

在步骤5中,采用softmax模型进行分类, $p_{i,j}$ 的分类结果 $C_{i,j}$ 表示为:

$$C_{i,j} = \text{argmax}(\text{softmax}(\text{conv}_{1 \times 1}(F_M)_{i,j}))$$

其中, $\text{argmax}(x)$ 表示向量 x 最大分量所对应的维度。

4. 根据权利要求1所述的的城市遥感影像场景分类方法,其特征在于,所述ReNet模块可以从上下左右四个方向沿像素的行列进行长程空间信息的提取分析。

5. 根据权利要求1所述的的城市遥感影像场景分类方法,其特征在于,所述softmax模型训练过程中,损失函数为交叉熵损失,采用反向传播方法来优化模型参数。

考虑空间关系信息的城市遥感影像场景分类方法

技术领域

[0001] 本发明属于遥感影像分析和深度学习技术领域,涉及大尺度高分辨率遥感影像的场景分类方法,具体涉及考虑空间关系信息的城市遥感影像场景分类方法。

背景技术

[0002] 随着遥感影像分辨率的不断提高,影像的分析已经从面向像素和面向对象阶段进入了面向场景阶段。城市场景的准确提取分类对城市规划、交通运输等方面发挥着重要的作用。自2012年以来深度学习在数字图像处理中的迅速发展,国内外许多研究人员已经开始将深度学习技术与遥感影像分析分类结合。如:Liu等通过在卷积神经网络中引入多级金字塔结构,提出了SPP-Net+MKL模型,实现了多尺度的遥感影像分析(参考文献:LIU Q, HANG R, SONG H, et al. 2017. Learning multiscale deep features for high-resolution satellite image scene classification. IEEE Transactions on Geoscience and Remote Sensing[J], 56: 117-126.); Cheng等人则通过设计特殊的损失函数,提出了Discriminative CNNs 模型,实现了对视觉上相似的不同遥感影像场景的分类(CHENG G, YANG C, YAO X, et al. 2018. When deep learning meets metric learning: Remote sensing image scene classification via learning discriminative CNNs. IEEE Transactions on Geoscience and Remote Sensing[J], 56: 2811-2821.)。与自然图片不同的是,遥感影像之间具有更加丰富的空间关系信息。由于自然图像之间几乎没有空间关系,因此基于自然图像分析的深度学习模型不能对于这种空间关系信息进行建模分析。同时,在建立用于模型训练的数据集时,往往影像与影像之间的信息会被抛弃。这两者使得深度学习模型在遥感影像上的总体表现不如其在自然图像上的表现。

发明内容

[0003] 本发明的目的在于提供考虑空间关系信息的城市遥感影像场景分类方法,本发明方法综合考虑两种空间关系信息:局部空间信息(regional spatial relationship)和长程空间信息(long-range spatial relationship)在分类中的作用,从而获得了较好的城市遥感影像场景分类结果。

[0004] 本发明的目的是这样实现的,包括以下步骤:

[0005] 步骤1,将遥感影像进行均匀无重叠裁剪成子图;

[0006] 步骤2,对每一个子图进行视觉信息编码,获得特征图像 F_v ;

[0007] 步骤3,将 F_v 输入交叉传递单元获得层次化空间特征;

[0008] 步骤4,对所述的层次化空间特征进行卷积降维;

[0009] 步骤5,对降维后的特征进行softmax模型分类,获得分类结果;

[0010] 步骤1中所述的裁剪过程为:尺寸 $M \times N$ 的大范围遥感影像 I 被均匀无重叠的滑动裁剪成 $m \times n$ 个子图,每个子图 $p_{i,j}$ 的尺寸为 $\frac{M}{n} \times \frac{N}{n}$,其中, $p_{i,j}$ 在 I 中的行列号 (i, j) 作为空间信息保存;

[0011] 步骤2中所述的视觉信息编码,采用预训练好的深度卷积模型将对每一个子图 $p_{i,j}$ 进行视觉信息编码,将 $p_{i,j}$ 转化为一个向量 $f_{V_{i,j}}$,遥感影像I被转换为特征图像 F_V :

$$F_V = \begin{pmatrix} f_{V_{1,1}} & \cdots & f_{V_{1,n}} \\ \vdots & \ddots & \vdots \\ f_{V_{m,1}} & \cdots & f_{V_{m,n}} \end{pmatrix};$$

[0012] 所述的交叉传递单元用于局部空间信息与长程空间信息的提取与融合,所述的局部空间信息的提取公式为: $F_r = \text{Conv}(F_V) = F_V * W + b$,其中, F_r 表示提取分析的空间关系信息, W 表示卷积核, b 表示偏置量,星号*表示卷积计算,所述的长程空间信息采用基于循环神经网络的ReNet模块进行提取。

[0013] 优选地,所述的交叉传递单元的输入为特征图像 F_V ,输出为层次化空间特征 F_E ,采用三条路径进行信息的提取分析与传递:第一条路径先对 F_V 进行局部空间信息的提取,而后进行长程空间信息的提取;第二条路径则相反,对 F_V 先进行长程空间信息的提取,而后进行局部空间信息的提取;第三条路径则直接将 F_V 传递到交叉传递单元末端的捷径,不进行额外处理,由此,输出层次化空间特征 F_E 表示为:

$$[0014] \quad F_E = \tanh(\text{ReNet}^2(\text{Conv}^1(F_V)) + \text{Conv}^2(\text{ReNet}^1(F_V)) + F_V)$$

[0015] 其中 \tanh 为双曲正切函数, ReNet^1 , ReNet^2 表示两个参数不同的ReNet模块, Conv^1 和 Conv^2 表示两个参数不同的卷积模块。

[0016] 优选地,步骤3中所述的特征图像经过三个串联的交叉传递单元,得到的层次化空间特征记为 F_M ;步骤4中采用尺寸为 1×1 的卷积层 $\text{Conv}_{1 \times 1}$ 对 F_M 进行卷积降维;

[0017] 在步骤5中,采用softmax模型进行分类, $p_{i,j}$ 的分类结果 $C_{i,j}$ 表示为:

$$[0018] \quad C_{i,j} = \text{argmax}(\text{softmax}(\text{conv}_{1 \times 1}(F_M)_{i,j}))$$

[0019] 其中 $\text{argmax}(x)$ 表示向量 x 最大分量所对应的维度。

[0020] 优选地,所述ReNet模块可以从上下左右四个方向沿像素的行列进行长程空间信息的提取分析。

[0021] 本发明方法针对现阶段的遥感影像分析方法无法分析影像之间的空间关系,设计了一个可以进行不同影像间空间关系提取分析的模型,模型可大致分为两个部分:视觉信息提取编码部分和空间关系信息融合部分,通过本发明方法能够实现对遥感影像进行较好场景分类,比如对遥感影像中的商业、工业、住宅、机构用地的区分。

附图说明

[0022] 图1本发明方法的流程示意图;

[0023] 图2本发明实施例中ReNet模块结构图;

[0024] 图3本发明实施例中交叉传递单元结构示意图;

[0025] 图4本发明实施例的数据处理流程示意图。

具体实施方式

[0026] 下面结合实施例和附图对本发明作进一步的说明,但不以任何方式对本发明加以限制,基于本发明教导所作的任何变换或替换,均属于本发明的保护范围。

[0027] 针对现阶段的遥感影像分析方法无法分析影像之间的空间关系,设计了一个可以

进行不同影像间空间关系提取分析的模型,模型可大致分为两个部分:视觉信息提取编码部分和空间关系信息融合部分,通过本发明方法能够实现对遥感影像进行较好场景分类,比如对遥感影像中的商业、工业、住宅、机构用地的区分。

[0028] 如图1所示,本发明实施例包括以下步骤:

[0029] 步骤1,将遥感影像进行均匀无重叠裁剪成子图;

[0030] 步骤2,对每一个子图进行视觉信息编码,获得特征图像 F_V ;

[0031] 步骤3,将 F_V 输入交叉传递单元获得层次化空间特征;

[0032] 步骤4,对所述的层次化空间特征进行卷积降维;

[0033] 步骤5,对降维后的特征进行softmax分类,获得分类结果。

[0034] 步骤1中,本实施例为保留遥感影像之间的空间关系信息,一张尺寸为 $M \times N$ 的大范围遥感影像 I 被均匀无重叠的滑动裁剪成 $m \times n$ 个子图,每个子图 $p_{i,j}$ 的尺寸为 $\frac{M}{n} \times \frac{N}{n}$,其中, $p_{i,j}$ 在 I 中的行列号 (i,j) 作为空间信息保存。

[0035] 步骤2中,对于每一个子图 $p_{i,j}$,采用预训练好的深度卷积模型将对其进行视觉信息编码,将 $p_{i,j}$ 转化为一个向量 $f_{V_{i,j}}$,最后 I 将被转换为特征图像 F_V :

$$[0036] \quad F_V = \begin{pmatrix} f_{V_{1,1}} & \cdots & f_{V_{1,n}} \\ \vdots & \ddots & \vdots \\ f_{V_{m,1}} & \cdots & f_{V_{m,n}} \end{pmatrix}, f_{V_{i,j}} \in \mathbb{R}^c \quad (1)$$

[0037] 在获取到 F_V 后,可将每个 $f_{V_{i,j}}$ 视为一个像素,对 I 中的 $p_{i,j}$ 分类问题则被转换为对 F_V 的语义分割问题。考虑到场景分布具有面状分布(机场、居民区等)和线状分布(道路、河流等),本实施例在考虑空间信息时主要考虑两种空间关系信息:局部空间信息(regional spatial relationship)和长程空间信息(long-range spatial relationship)。对空间关系信息的建模包括以下三个方面:

[0038] 1、局部空间信息建模

[0039] 对于 F_V ,局部空间信息可理解为 $f_{V_{i,j}}$ 与其一定邻域范围内的向量的关系。卷积神经网络模型通过卷积操作,可以提取融合一定邻域内的信息,从而达到局部空间信息建模的目的。因此在局部空间信息分析中本发明将采用卷积模型。假设用 W 表示卷积核, b 表示偏置量, F_r 表示提取分析的空间关系信息,则一层卷积模型可表示为:

$$[0040] \quad F_r = \text{Conv}(F_V) = F_V * W + b \quad (2)$$

[0041] 其中星号(*)表示卷积计算。

[0042] 2、长程空间信息建模

[0043] ReNet模块结构图如图2所示,对于 F_V ,长程空间信息可理解为可 $f_{V_{i,j}}$ 与其所在行列向量的关系。循环神经网络在序列模型中有广泛的应用,其特殊的阶段信息处理结构可以对上下文信息进行综合的分析。考虑到同一行或同一列的 f_V 可以被视为一个序列数据,因此本实施例引入了基于循环神经网络的ReNet模块。ReNet模块可以从上下左右四个方向沿像素的行列进行长程空间信息的提取分析,实验表明,其在部分公开数据及上的表现可以达到卷积神经网络的水平 (VISIN F, KASTNER K, CHO K, et al. 2015. ReNet: A Recurrent Neural Network Based Alternative to Convolutional Networks. Computer

Science[J],25: 2983-2996.)。

[0044] 3、空间关系融合建模

[0045] 交叉传递单元结构示意图如图3所示， \oplus 表示相加，本实施例设计了交叉传递单元(crossing transfer unit,CTU)来实现局部空间信息与长程空间信息的融合。CTU以特征图像作为输入(F_V)与输出(F_E)，采用三条路径进行信息的提取分析与传递：第一条路径先对 F_V 进行局部空间信息的提取，而后进行长程空间信息的提取；第二条路径则相反，对 F_V 先进行长程空间信息的提取，而后进行局部空间信息的提取；第三条路径则直接将 F_V 传递到CTU末端的捷径，不进行额外处理，实验表明，增加一条直传路径可以加速模型的收敛速度(He K,Zhang X, Ren S, et al. Deep residual learning for image recognition[C] Proceedings of the IEEE conference on computer vision and pattern recognition.2016:770-778.)。最终的输出结果 F_E 可表示为：

$$[0046] \quad F_E = \tanh(\text{ReNet}^2(\text{Conv}^1(F_V)) + \text{Conv}^2(\text{ReNet}^1(F_V)) + F_V) \quad (3)$$

[0047] 其中 \tanh 为双曲正切函数， $\text{ReNet}^1, \text{ReNet}^2$ 表示两个参数不同的ReNet模块， Conv^1 和 Conv^2 表示两个参数不同的卷积模块。

[0048] 经过三个串联的交叉传递单元，空间关系信息的建模所得到的结果记为 F_M ，本实施例采用一个尺寸为 1×1 的卷积层 $\text{Conv}_{1 \times 1}$ 对 F_M 进行卷积降维，采用softmax模型进行分类，最终 $p_{i,j}$ 的分类结果 $C_{i,j}$ 可表示为：

$$[0049] \quad C_{i,j} = \text{argmax}(\text{softmax}(\text{conv}_{1 \times 1}(F_M)_{i,j})) \quad (4)$$

[0050] 其中 $\text{argmax}(x)$ 表示向量 x 最大分量所对应的维度。

[0051] softmax模型训练过程中，损失函数为交叉熵损失，采用反向传播方法来优化模型参数。数据处理的基本流程如图4所示。

[0052] 实验中所采用的数据集为CSU-RESISC10数据集，该数据集预处理后的训练测试样本分布如表1所示。

[0053] 表1

场景类别	道路	商业区	工业区	住宅区	建设用地	机构用地	港口	水域	公共场所	机场
[0054] 测试集	17129	6768	1588	39806	530	1948	5331	12304	11587	3009
验证集	2480	512	506	5728	22	386	665	1765	2642	494

[0055] 对于CSU-RESISC10数据集中的每一张 2000×2000 的遥感影像 I ，首先将其均匀无重叠地裁剪为 20×20 个子图 $p_{i,j}$ ，每个子图地尺寸为 100×100 。

[0056] 对于每一个子图 $p_{i,j}$ ，使用在CSU-RESISC10上预训练的Xception模型将其进行视觉信息编码，将 $p_{i,j}$ 转化为一个2048维向量 $f_{v_{i,j}}$ 。最后 I 将被转换为特征图像 $F_V \in \mathbb{R}^{20 \times 20 \times 2048}$ 。为了减少计算量，在进行下一步计算之前，采用一个尺寸为 1×1 的卷积将 F_V 降至512维。

[0057] 预训练共训练50个批次，学习率为 10^{-5} ，衰减率为0.98，预训练的结果可参照表2第一行。

[0058] 在进行空间关系建模时，本发明通过增加补边和控制卷积核补偿，保持所有输出特征图像和输入特征图像的尺寸不变。为了充分提取 $f_{v_{i,j}}$ 之间的空间关系，本发明采用了3个CTU模块逐级递进地提取层次化空间特征。空间关系建模的最终输出结果为

$F_M \in \mathbb{R}^{20 \times 20 \times 512}$ 。

[0059] 本实施例最后按照公式(4)分类。

[0060] 模型训练时采用交叉熵作为损失函数,本发明模型共训练100批次,学习率为 10^{-5} ,衰减率为0.98。在训练约15次后,模型收敛。

[0061] 为验证本发明的有效性,除本实施例外,额外选择了SPP-Net+MKL、Discriminative CNNs和传统自然图片分类模型Xception(Chollet F.Xception: Deep learning with depthwise separable convolutions[C]Proceedings of the IEEE conference on computer vision and pattern recognition.2017: 1251-1258.)作为对比,在CSU-RESISC10数据集上进行分类实验,并选择F1分数和Kappa系数(κ)作为评价依据。

[0062] 对比试验中超参数保持与本发明方法一致。各方法在不同场景中的表现如表2所示。

[0063] 表2

方法	F1 分数										κ
	道路	商业区	工业区	住宅区	建设用地	机构用地	港口	水域	公共场所	机场	
Xception	0.8131	0.3922	0.3541	0.8640	0.3793	0.2838	0.8615	0.9380	0.8340	0.8421	0.7638
SPP-Net+MKL	0.8133	0.4293	0.4680	0.8734	0.3750	0.1746	0.8265	0.9109	0.8260	0.8566	0.7624
Discriminative CNNs	0.8434	0.3723	0.4912	0.8802	0.4000	0.2639	0.8239	0.9273	0.8422	0.8057	0.7731
本发明方法	0.8329	0.6030	0.7643	0.9014	0.4400	0.6218	0.9239	0.9598	0.8841	0.9648	0.8410

[0064] 实验结果显示,由于场景的复杂性,单一的遥感影像并不能很好地实现对商业、工业、住宅、机构用地的区分。用于对比的三种方法, κ 均小于0.78。由于将影像之间的空间关系纳入考虑,本发明方法相较于其他三组对比试验, Kappa的提升幅度分别为10.1%、10.3%和8.8%。

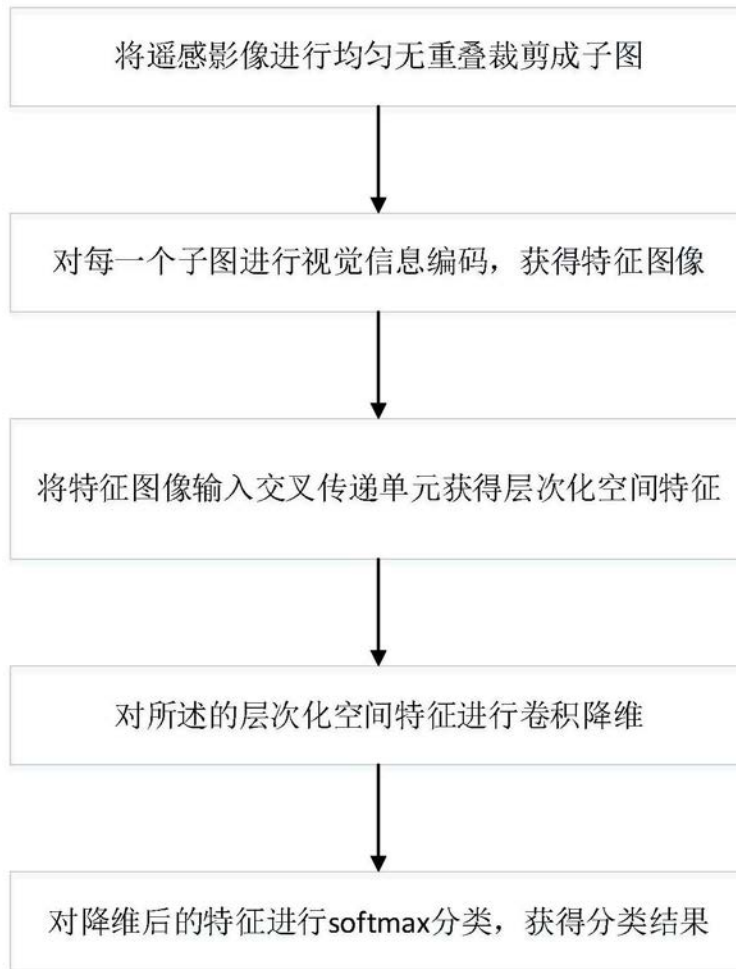


图1

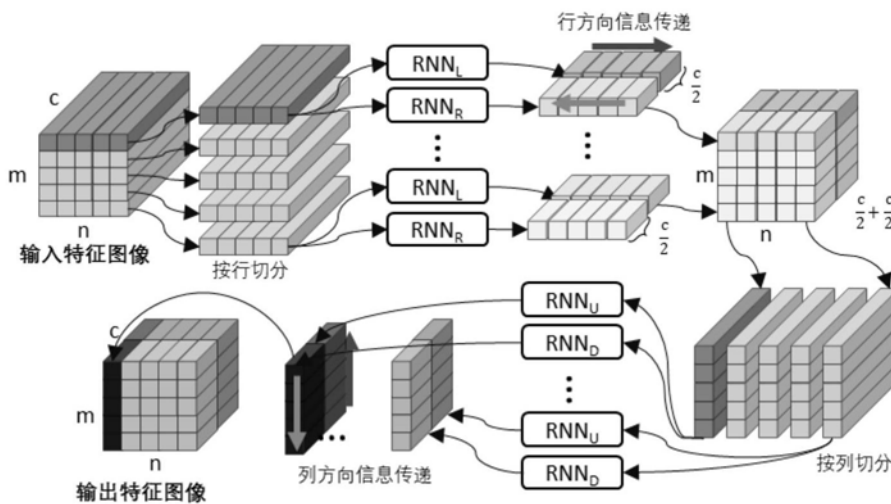


图2

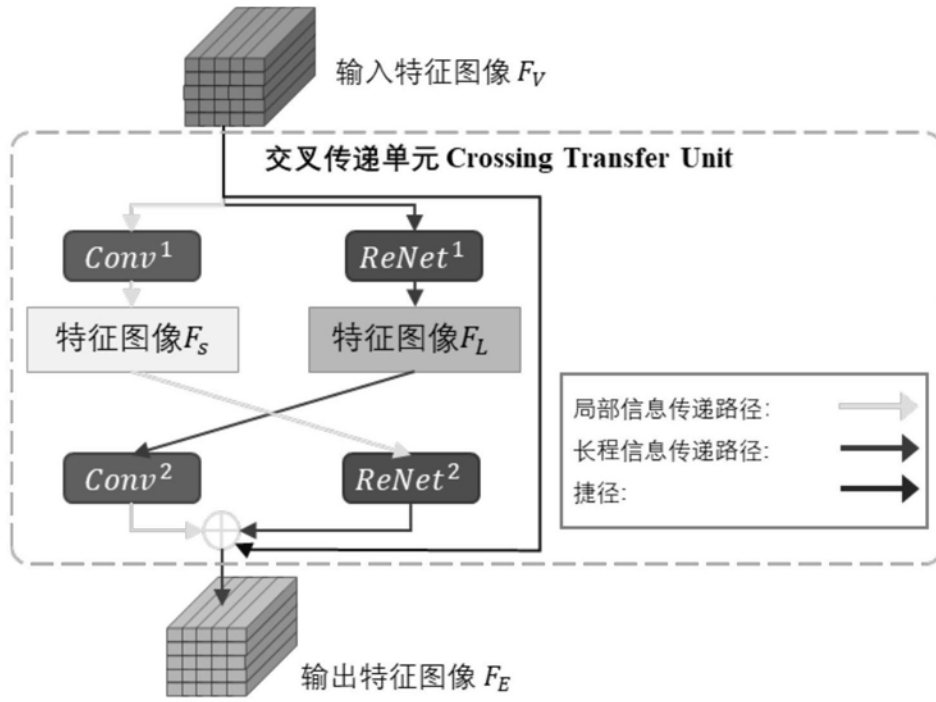


图3

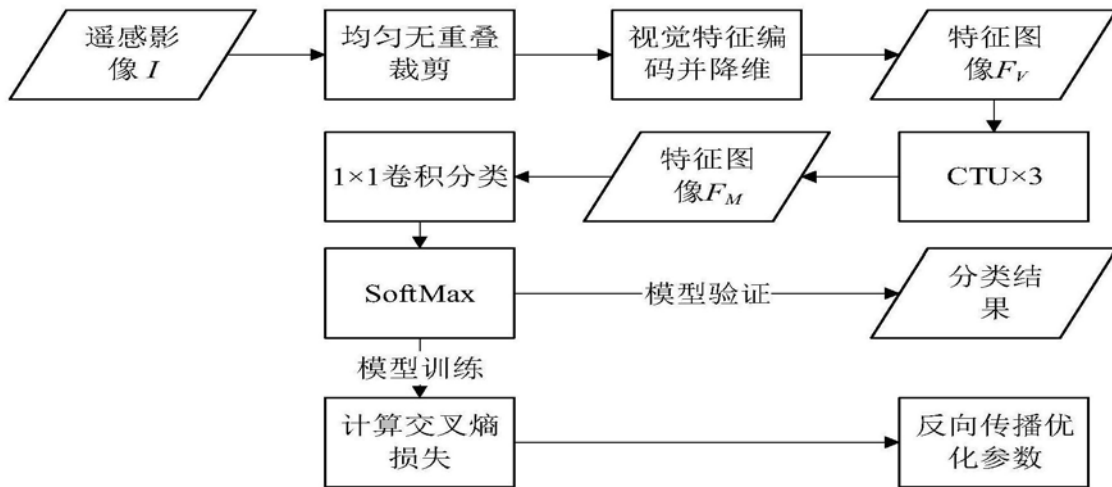


图4