



(12) 发明专利申请

(10) 申请公布号 CN 112329771 A

(43) 申请公布日 2021.02.05

(21) 申请号 202011201983.7 *G06N 3/08* (2006.01)

(22) 申请日 2020.11.02 *G06T 7/00* (2017.01)

*G06T 7/11* (2017.01)

(71) 申请人 元准智能科技(苏州)有限公司

地址 215000 江苏省苏州市常熟市沙家浜  
镇常昆工业园区南新路11号

申请人 苏州市建筑科学研究院集团股份有  
限公司

(72) 发明人 赵力 程荣 张亦明

(74) 专利代理机构 南京苏科专利代理有限责任  
公司 32102

代理人 蒋慧妮

(51) Int. Cl.

*G06K 9/32* (2006.01)

*G06K 9/62* (2006.01)

*G06N 3/04* (2006.01)

权利要求书2页 说明书5页

(54) 发明名称

一种基于深度学习的建筑材料样本识别方法

(57) 摘要

本发明提供了一种基于深度学习的建筑材料样本识别方法,包括模型训练阶段及样本识别阶段,模型训练阶段包括制作建材样本数据集、构建样本识别的多尺度信息融合卷积神经网络;对样本数据集数据增强获得最佳模型性能;样本识别阶段包括将处理好的建材样本图像输入模型,进行特征提取生成最佳尺寸特征图,将生成特征图修正作为ROI,根据不同尺度ROI传入ROI pooling层中,并映射为同一尺寸proposal,投射在原始建材样本图像上生成proposal feature map,进行BBOX和CLS分支处理等步骤生成位置精准的建材检测框并识别出样本的材料性能状态。本方法通过多尺度的特征图进行信息提取,很好的学习到不同尺度的目标特征信息,具有良好的识别性能及泛用性,在建设工程领域有着广阔的应用前景。

1. 一种基于深度学习的建筑材料样本识别方法,其特征在于:包括模型训练阶段及样本识别阶段,

所述模型训练阶段包括如下步骤:

S1、进行建材样本采集、标注,制作建材样本数据集;所述数据集包含每种类别的建筑材料样本,并切分为训练集、测试集和评估集;

S2、构建建材样本识别的多尺度信息融合卷积神经网络;

S3、对S1中的样本数据集进行数据增强:以获得最佳的模型性能;

所述样本识别阶段包括如下步骤:

S1、将处理好的建材样本图像输入模型,通过imagenet上预训练好,并去顶层的Resnet进行特征提取,并生成最佳尺寸的特征图;

S2、将S1中生成的特征图分别经过RPN2,3,4,5,生成不同尺寸的候选anchor;对anchor面积进行设定,所有RPN生成的anchor统一采用1:1、1:2、2:1的长宽比,生成多个候选anchor后RPN根据真实标注,利用二分类和bounding box regression功能筛选出包含目标最完整的anchor并进行修正,作为ROI;

S3、不同尺度的ROI,使用不同的残差卷积模块所输出的特征图作为ROI pooling层的输入,大尺度的ROI采用深层卷积模块的输出特征图,使用ROI leave作为采用相应层卷积模块输出的判别标准:

$$ROI_{leave} = \left\lfloor K_0 + \log_2 \left( \frac{\sqrt{wh}}{224} \right) \right\rfloor, \text{其中, } w, h \text{ 为ROI区域的长宽, } K_0 \text{ 为基准leave};$$

S4、将S3生成的ROI传入ROI pooling层中,ROI pooling将多尺度的ROI统一映射为同一尺寸大小的proposal,并投射在原始的建材样本图像上生成proposal feature map,方便后续的BBOX和CLS分支进行处理;

S5、CLS分支对proposal feature map,通过全连接层与softmax计算每个样本所属的类别,输出最高的类别概率作为置信度;

S6、BBOX分支利用bounding box regression功能对proposal区域进行修正,生成位置更加精准的建材检测框并识别出建材样本的材料性能状态。

2. 如权利要求1所述的一种基于深度学习的建筑材料样本识别方法,其特征在于:所述模型训练阶段S3包括如下步骤:

S31、利用多种数据增强的方法组合构建出不同尺度和场景下的建材样本图像,对现有数据进行扩充,以模拟复杂的识别场景,提高模型对细节特征信息的学习,增强模型的泛用性;

S32、设置初始权重为imagenet上预训练的权重,初始学习率为0.001,学习率衰减指数为0.1,batch\_size为16,并输入图像尺寸;

S33、损失函数上,RPN系列模块采用二分类损失和回归损失;CLS分支采用多分类损失,BBOX分支采用回归损失;

S34、训练中采用SGD随机梯度下降优化器,在训练集、测试集上训练直至模型性能达到最佳。

3. 如权利要求2所述的一种基于深度学习的建筑材料样本识别方法,其特征在于:所述S31中数据增强的方法包括:

S311、Random Erasing:

- (1) IRE,在整个目标图像上随机选择遮挡位置;
- (2) ORE,在目标的bounding-box区域内随机选择遮挡位置;
- (3) 将IRE和ORE二者组合;

S312、Hide and Seek:

将图片切分为S\*S个网格,每个网格按概率进行隐藏,在训练中对同样的图片,每个批次都隐藏不同的网格组;

S313、Grid Mask:

为避免S31和S32中1存在过度删除区域,造成完整目标被删除或者上下文信息缺失的问题;通过Grid Mask对x,y,r,d四个参数进行设定:

$$k = \frac{\text{sum}(M)}{H \times W} = 1 - (1 - r^2)$$

其中,r为mask的尺寸,M为保留像素数,H、W为图像尺寸;x、y为在图像上随机生成的区域坐标;未遮挡区域的值为1,遮挡区域的值为0,生成一个和原图相同分辨率的mask,然后将该mask与原图相乘得到一个图像;

S314、Mixup

对图像进行混类增强,将不同类之间的图像进行混合;其算法可归结为:

$$\tilde{x}_1 = \lambda x_1 + (1 - \lambda) x_2$$

$$\tilde{x}_2 = \lambda x_1 + (1 - \lambda) x_2$$

其中 $x_1$ 、 $x_2$ 为不同的图像的像素, $\lambda$ 为权重;

S315、Cutmix

随机将一部分区域裁剪掉,并随机填充训练集中的其他数据的区域像素值。

4.如权利要求1所述的一种基于深度学习的建筑材料样本识别方法,其特征在于:所述样本识别阶段S1进行样本提取后生成最佳特征图的步骤包括:

S11、将Resnet中最后若干个残差卷积模块,记为 $\{C_1, C_2, C_3, \dots, C_n\}$ ,分别提取残差模块的输出特征图记为 $\{P_1, P_2, P_3, \dots, P_n\}$ ;

S12、将最深层的特征图 $P_5$ 经过1进行2倍的最近邻上采样;

S13、提取 $C_n$ 相邻的残差卷积模块 $C_{n-1}$ 的输出特征图 $P_{n-1}$ ,进行1\*1的卷积降维处理;

S14、然后通过对对应部位的像素值相加的方法,与特征图 $P_n$ 进行融合;

S15、融合特征图经过3\*3卷积减少上采样所带来的混叠效果;

S16、迭代上述过程,直至生成最佳尺寸特征图。

## 一种基于深度学习的建筑材料样本识别方法

### 技术领域

[0001] 本发明属于建设工程领域,具体涉及一种基于深度学习的建筑材料样本识别方法。

### 背景技术

[0002] 随着数据信息处理需求的增加和人工智能技术的迅速发展,人们开始尝试使用何种基于机器学习或深度学习的方法对建筑材料样本进行识别,如聚类神经网络,支持向量机,小波变换神经网络等浅层网络算法。但上述浅层网络算法需要各种复杂算法才能从回波信息中提取决定样品识别特征信息;计算复杂度和消耗计算资源较高,因此通用性较低。卷积神经网络是深度学习领域重要的模型之一,其网络结构对存在平移、反转、仿射变换等特性的图像数据具有高度不变性,因此,近年来卷积神经网络在计算机视觉中的各个领域被广泛使用并取得了优良成果。但是传统的单线性卷积神经网络在识别过程中,只针对最后一层输出,即:单一尺度的特征图进行信息提取,显然无法良好的学习到不同尺度的目标特征信息;因此在复杂的建筑材料场景中很难实现良好的识别性能。

### 发明内容

[0003] 为了解决现有技术中的难题,本发明提出了一种基于深度学习的建筑材料样本识别方法。

[0004] 本发明的目的通过以下技术方案来实现:

[0005] 一种基于深度学习的建筑材料样本识别方法,包括模型训练阶段及样本识别阶段,

[0006] 所述模型训练阶段包括如下步骤:

[0007] S1、进行建材样本采集、标注,制作建材样本数据集;所述数据集包含每种类别的建筑材料样本,并切分为训练集、测试集和评估集;

[0008] S2、构建建材样本识别的多尺度信息融合卷积神经网络;

[0009] S3、对S1中的样本数据集进行数据增强:以获得最佳的模型性能;

[0010] 所述样本识别阶段包括如下步骤:

[0011] S1、将处理好的建材样本图像输入模型,通过imagenet上预训练好,并去顶层的Resnet进行特征提取,并生成最佳尺寸的特征图;

[0012] S2、将S1中生成的特征图分别经过RPN2,3,4,5,生成不同尺寸的候选anchor;对anchor面积进行设定,所有RPN生成的anchor统一采用1:1、1:2、2:1的长宽比,生成多个候选anchor后RPN根据真实标注,利用二分类和bounding box regerssion功能筛选出包含目标最完整的anchor并进行修正,作为ROI;

[0013] S3、不同尺度的ROI,使用不同的残差卷积模块所输出的特征图层作为ROI pooling层的输入,大尺度的ROI采用深层卷积模块的输出特征图,使用ROI leave作为采用相应层卷积模块输出的判别标准:

[0014]  $ROI_{leave} = \left\lfloor K_0 + \log_2 \left( \frac{\sqrt{wh}}{224} \right) \right\rfloor$ , 其中, w、h为ROI区域的长宽,  $K_0$ 为基准leave;

[0015] S4、将S3生成的ROI传入ROI pooling层中, ROI pooling将多尺度的ROI统一映射为同一尺寸大小的proposal, 并投射在原始的建材样本图像上生成proposal feature map, 方便后续的BBOX和CLS分支进行处理;

[0016] S5、CLS分支对proposal feature map, 通过全连接层与softmax计算每个样本所属的类别, 输出最高的类别概率作为置信度;

[0017] S6、BBOX分支利用bounding box regression功能对proposal区域进行修正, 生成位置更加精准的建材检测框并识别出建材样本的材料性能状态。

[0018] 优选地, 所述模型训练阶段S3包括如下步骤:

[0019] S31、利用多种数据增强的方法组合构建出不同尺度和场景下的建材样本图像, 对现有数据进行扩充, 以模拟复杂的识别场景, 提高模型对细节特征信息的学习, 增强模型的泛用性;

[0020] S32、设置初始权重为imagenet上预训练的权重, 初始学习率为0.001, 学习率衰减指数为0.1, batch\_size为16, 并输入图像尺寸;

[0021] S33、损失函数上, RPN系列模块采用二分类损失和回归损失; CLS分支采用多分类损失, BBOX分支采用回归损失;

[0022] S34、训练中采用SGD随机梯度下降优化器, 在训练集、测试集上训练直至模型性能达到最佳。

[0023] 优选地, 所述S31中数据增强的方法包括:

[0024] S311、Random Erasing:

[0025] (3) IRE, 在整个目标图像上随机选择遮挡位置;

[0026] (4) ORE, 在目标的bounding-box区域内随机选择遮挡位置;

[0027] (3) 将IRE和ORE二者组合;

[0028] S312、Hide and Seek:

[0029] 将图片切分为S\*S个网格, 每个网格按概率进行隐藏, 在训练中对同样的图片, 每个批次都隐藏不同的网格组;

[0030] S313、Grid Mask:

[0031] 为避免S31和S32中1存在过度删除区域, 造成完整目标被删除或者上下文信息缺失的问题; 通过Grid Mask对x, y, r, d四个参数进行设定:

$$[0032] \quad k = \frac{\text{sum}(M)}{H \times W} = 1 - (1 - r^2)$$

[0033] 其中, r为mask的尺寸, M为保留像素数, H、W为图像尺寸; x、y为在图像上随机生成的区域坐标; 未遮挡区域的值为1, 遮挡区域的值为0, 生成一个和原图相同分辨率的mask, 然后将该mask与原图相乘得到一个图像;

[0034] S314、Mixup

[0035] 对图像进行混类增强, 将不同类之间的图像进行混合; 其算法可归结为:

$$[0036] \quad \tilde{x}_1 = \lambda x_1 + (1 - \lambda) x_2$$

$$[0037] \quad \tilde{x}_2 = \lambda x_1 + (1 - \lambda) x_2$$

[0038] 其中 $x_1$ 、 $x_2$ 为不同的图像的像素， $\lambda$ 为权重；

[0039] S315、Cutmix

[0040] 随机将一部分区域裁剪掉，并随机填充训练集中的其他数据的区域像素值。

[0041] 优选地，所述样本识别阶段S1进行样本提取后生成最佳特征图的步骤包括：

[0042] S11、将Resnet中最后若干个残差卷积模块，记为 $\{C_1, C_2, C_3, \dots, C_n\}$ ，分别提取残差模块的输出特征图记为 $\{P_1, P_2, P_3, \dots, P_n\}$ ；

[0043] S12、将最深层的特征图 $P_5$ 经过1进行2倍的最近邻上采样；

[0044] S13、提取 $C_n$ 相邻的残差卷积模块 $C_{n-1}$ 的输出特征图 $P_{n-1}$ ，进行 $1 \times 1$ 的卷积降维处理；

[0045] S14、然后通过对对应部位的像素值相加的方法，与特征图 $P_n$ 进行融合；

[0046] S15、融合特征图经过 $3 \times 3$ 卷积减少上采样所带来的混叠效果；

[0047] S16、迭代上述过程，直至生成最佳尺寸特征图。

[0048] 本发明的有益效果体现在：本发明的方法通过多尺度的特征图进行信息提取，很好的学习到不同尺度的目标特征信息，具有良好的识别性能及泛用性，在建设工程领域有着广阔的应用前景。

### 具体实施方式

[0049] 以下结合实施例具体阐述本发明的技术方案，本发明揭示了一种基于深度学习的建筑材料样本识别方法，包括模型训练阶段及样本识别阶段，所述模型训练阶段包括如下步骤：

[0050] S1、进行建材样本采集、标注，制作建材样本数据集；所述数据集包含每种类别的建筑材料样本，并切分为训练集、测试集和评估集；

[0051] S2、构建建材样本识别的多尺度信息融合卷积神经网络；

[0052] S3、对S1中的样本数据集进行数据增强，以获得最佳的模型性能；

[0053] 具体的，

[0054] S31、利用多种数据增强的方法组合构建出不同尺度和场景下的建材样本图像，对现有数据进行扩充，以模拟复杂的识别场景，提高模型对细节特征信息的学习，增强模型的泛用性；

[0055] S32、设置初始权重为imagenet上预训练的权重，初始学习率为0.001，学习率衰减指数为0.1，batch\_size为16，输入图像尺寸为 $224 \times 224$ 。

[0056] S33、损失函数上，RPN系列模块采用二分类损失和回归损失；CLS分支采用多分类损失，BBOX分支采用回归损失。

[0057] S34、训练中采用SGD随机梯度下降优化器，在训练集、测试集上训练20epoch，直至模型性能达到最佳。

[0058] 其中，所述S31中增强的方法包括如下步骤：

[0059] S311、Random Erasing：

[0060] (5) IRE，在整个目标图像上随机选择遮挡位置；

[0061] (6) ORE，在目标的bounding-box区域内随机选择遮挡位置；

[0062] (3) 将IRE和ORE二者组合;

[0063] S312、Hide and Seek:

[0064] 图片切分为S\*S个网格,每个网格采用一定概率进行隐藏,在训练中对同样的图片,每个批次都隐藏不同的网格组;

[0065] S313、Grid Mask:

[0066] 为避免S31和S32中1存在过度删除区域,造成完整目标被删除或者上下文信息缺失的问题;通过Grid Mask对x,y,r,d四个参数进行设定:

$$[0067] \quad k = \frac{\text{sum}(M)}{H \times W} = 1 - (1 - r^2)$$

[0068] 其中,r为mask的尺寸,M为保留像素数,H、W为图像尺寸;x、y为在图像上随机生成的区域坐标;未遮挡区域的值为1,遮挡区域的值为0,生成一个和原图相同分辨率的mask,然后将该mask与原图相乘得到一个图像;

[0069] S314、Mixup

[0070] 对图像进行混类增强,将不同类之间的图像进行混合;其算法可归结为:

$$[0071] \quad \tilde{x}_1 = \lambda x_1 + (1 - \lambda) x_2$$

$$[0072] \quad \tilde{x}_2 = \lambda x_1 + (1 - \lambda) x_2$$

[0073] 其中 $x_1$ 、 $x_2$ 为不同的图像的像素, $\lambda$ 为权重。

[0074] S315、Cutmix

[0075] 随机将一部分区域裁剪掉,并随机填充训练集中的其他数据的区域像素值。

[0076] 所述样品识别阶段包括如下步骤:

[0077] S1、将处理好的建材样本图像输入模型,通过imagenet上预训练好并去顶层的Resnet进行特征提取。

[0078] S2、将Resnet中最后5个残差卷积模块,记为 $\{C_1, C_2, C_3, C_4, C_5\}$ ,分别提取5个残差模块的输出特征图记为 $\{P_1, P_2, P_3, P_4, P_5\}$ ;生成最佳尺寸的特征图;

[0079] 所述生成特征图的包括如下步骤:

[0080] S21、将最深层的特征图 $P_5$ 经过1进行2倍的最近邻上采样;

[0081] S22、提取 $C_5$ 相邻的残差卷积模块 $C_4$ 的输出特征图 $P_4$ ,进行 $1 \times 1$ 的卷积降维处理;

[0082] S23、然后通过对应部位的像素值相加的方法,与特征图 $P_5$ 进行融合,融合方式;

[0083] S24、融合特征图经过 $3 \times 3$ 卷积减少上采样所带来的混叠效果;

[0084] S25、迭代上述过程,直至生成最佳尺寸的特征图;

[0085] S3、上述步骤中生成的特征图分别经过RPN2,3,4,5,生成不同尺寸的候选anchor; anchor面积分别设定为 $32 \times 32$ 、 $64 \times 64$ 、 $128 \times 128$ 、 $256 \times 256$ ,所有RPN生成的anchor统一采用1:1、1:2、2:1的长宽比,生成多个候选anchor后RPN根据真实标注,利用二分类和bounding box regerssion功能筛选出包含目标最完整的anchor并进行修正,作为ROI;

[0086] S4、不同尺度的ROI,使用不同的残差卷积模块所输出的特征图层作为ROI pooling层的输入,大尺度的ROI采用深层卷积模块的输出特征图,使用ROI leave作为采用

某层卷积模块输出的判别标准:  $ROI_{\text{leave}} = \left\lfloor K_0 + \log_2 \left( \frac{\sqrt{wh}}{224} \right) \right\rfloor$ ,其中,w、h为ROI区域的长

宽,  $K_0$ 为基准leave, 小尺度ROI就用深度浅层卷积模块的输出特征图, 设定为5, 代表特征图  $P_5$ 尺寸。

[0087] S5、将S4生成的ROI传入ROI pooling层中, ROI pooling将多尺度的ROI统一映射为7\*7尺寸大小的proposal, 并投射在原始的建材样本图像上生成proposal feature map, 方便后续的BBOX和CLS分支进行处理;

[0088] S6、CLS分支对proposal feature map, 通过全连接层与softmax计算每个样本所属的类别, 输出最高的类别概率作为置信度;

[0089] S7、BBOX分支利用bounding box regression功能对proposal区域进行修正, 生成位置更加精准的建材检测框, 并识别出建材样本的材料性能状态。

[0090] 当然本发明尚有多种具体的实施方式, 在此就不一一列举。凡采用等同替换或者等效变换而形成的所有技术方案, 均落在本发明要求保护的范围之内。