

# Convolution Neural Network-Based System for Detecting Cracks on Exterior Wall

Hui Xiong<sup>1,2\*</sup>, Peifeng Liang<sup>1</sup>, Junjian Huang<sup>1</sup>, Min Hu<sup>1</sup>

<sup>1</sup>Guangdong Guangye Kaiyuan Technology Co., Ltd., Guangzhou Guangdong

<sup>2</sup>Guangdong Jinglong Construction Group Co., Ltd., Guangzhou Guangdong

Email: \*375610596@qq.com

Received: Nov. 28<sup>th</sup>, 2018; accepted: Dec. 12<sup>th</sup>, 2018; published: Dec. 19<sup>th</sup>, 2018

---

## Abstract

Based on the mathematical model of convolutional neural network, the database is built by taking images of external walls by unmanned aerial vehicle; a crack detection system for external wall and its facing material is established by software and hardware. The system can effectively identify the severity, general or slight damage degree of external wall cracks, and the effective identification rates are 86%, 91% and 97% respectively.

## Keywords

Convolutional Neural Network, Crack Detection, Image Processing

---

# 基于卷积神经网络的高楼外墙裂缝检测系统

熊 辉<sup>1,2\*</sup>, 梁培锋<sup>1</sup>, 黄俊健<sup>1</sup>, 胡 敏<sup>1</sup>

<sup>1</sup>广东广业开元科技有限公司, 广东 广州

<sup>2</sup>广东景龙建设集团有限公司, 广东 广州

Email: \*375610596@qq.com

收稿日期: 2018年11月28日; 录用日期: 2018年12月12日; 发布日期: 2018年12月19日

---

## 摘 要

基于卷积神经网络的数学模型, 通过无人机拍摄外墙图像建立数据库, 本文结合软硬件建立了一种外墙及饰面材料的裂缝检测系统, 能有效地识别外墙裂缝的严重、一般或轻微三种毁坏程度, 且有效识别率分别为86%, 91%, 97%。

\*通讯作者。

文章引用: 熊辉, 梁培锋, 黄俊健, 胡敏. 基于卷积神经网络的高楼外墙裂缝检测系统[J]. 软件工程与应用, 2018, 7(6): 273-282. DOI: 10.12677/sea.2018.76031

## 关键词

卷积神经网络, 裂缝检测, 图像处理

Copyright © 2018 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

一般而言, 建筑物因日久而逐渐老化, 墙缝吸收水分湿气, 经历温差变化所产生的热胀冷缩的效应, 加上地震或台风的破坏, 导致外墙或其饰面材料易被破坏、龟裂。若外墙饰面材料剥落, 极大可能造成意外伤亡。因此, 若能在外墙或饰面材料剥落之前提前预警, 则可将意外伤亡的概率降至最低。有鉴于此, 本文之主要目的, 在于提供一套有效的高楼外墙裂缝检测系统, 让原本繁杂、危险且高成本的例行性工作, 通过信息技术而变得简易而节省。

本文拟以小型无人机搭配智能运作系统, 能够在地震或者平时辨识饰面材料的损毁程度, 以期能让工人减少攀高的次数, 进行快速检测, 识别具有潜在危险的饰面材料位置, 且评估饰面材料退化的程度。为了实现上述目的, 本文以机器学习作为解决问题的主要手段, 以深度学习为核心, 使用卷积神经网络(convolutional neural network, CNN)技术将辨识功能实施在嵌入式开发板上, 让人能够以简易方式辨识大楼外墙饰面材料破裂的程度。

## 2. 研究方法

利用无人机对建筑外墙进行拍摄, 再将搜集回来的图像数据输入到嵌入式开发板, 进行损坏程度的辨认, 自动判断是否需要维修或更换。要想将无人机所搜集而成的数据分类并进行分析, 为此需先在嵌入式开发板(本文采用 Jetson TX2)与计算机主机内同时建立 Tensorflow [1]的环境并使用 Keras [2]作为架设模块的基础, 利用 Keras 设立 CNN 模型且以 Tensorflow 作为后端进行训练。在训练前先数据集分成几个类别(即标签化)并进行数据分类处理, 将劣化程度相近的照片归成一类; 利用不同的数据进行训练, 再根据训练适当调整参数, 让模块评估达到最低的错误率; 最后将训练完成的模块储存起来放进 Jetson TX2 里, 就能进行饰面材料损坏程度的判断了。

上述 CNN 是一种由类神经网络为基础设计而成的模型, 其可以通过减少神经元与撷取特征改变类神经网络的运算量。CNN 是目前深度学习领域中最广泛使用的一种模型, 擅长处理图像数据。CNN 与图像处理技术不同, 传统的图像处理是针对输入的图像进行分析与加工(如形态学处理、撷取边缘、图像切割、增加、修复等), 而 CNN 则采用不同策略, 将大量的图片输入到类神经网络之中, 通过图层的特征撷取, 最后加以分类以识别图像。相较于传统的图像处理方法, CNN 能够处理复杂的图像, 并具备学习的能力, 亦可用于预测。目前的 CNN 的主流设计包括 LeNet-5 [3]、AlexNet [4]、Network-in-Network [5] 与 GoogLeNet [6]等。其中, AlexNet 采用最大池化(max pooling)技巧, 导入 ReLU (Rectified Liner Units) 机制, 让神经元处理非线性情境的“梯度消失”问题; AlexNet 在处理大量的数据与撷取较多的特征时, 在识别的正确率及整体的训练速度有较佳的表现。

因此, 在多方比较后, 本文决定采用 AlexNet 作为实现目的的主要模型。AlexNet 设计将图片分解成像素单位, 并将每个像素输入类神经网络, 通过连续的卷积程序作特征撷取, 并通过池化(pooling)程序

凸显每一个区域的最显著特征，此举可减少神经元的数量并降低所需的运算量。通过输入 ImageNet 的图像数据库作训练[7]，该模型能够从 120 万张照片作为训练样本、5 万张当测试组，从中分辨 1000 种类型的图片。

在 CNN 模型中，卷积程序主要是通过图片与卷积核作内积运算来撷取所需的特征。例如，肖进胜等针对现有的基于 CNN 的图像超分辨率算法参数较多、计算量较大、训练时间较长、图像纹理模糊等问题，结合现有的图像分类网络模型和视觉识别算法对其提出了改进[8]。周敏等将深度卷积神经网络应用于遥感图像飞机目标分类问题，按大致 4:1 的比例分为训练集和测试集，并对训练集进行合理扩充，针对遥感图像与飞机分类的特殊性，结合深度学习卷积神经网络的相关理论，设计了一个 5 层卷积神经网络[9]。葛芸等人将大规模数据集 ImageNet 上预训练的 4 种不同卷积神经网络用于遥感图像检索，在 UC-Merced 和 WHU-RS 数据集上得到了一系列实验结果[10]。傅隆生等人提出一种基于 LeNet 卷积神经网络的深度学习模型进行多簇猕猴桃果实图像的识别方法[11]。

如图 1 所示，红色框内的矩阵(维度  $3 \times 3$ )与维度  $3 \times 3$  卷积核进行内积而得的纯量值 3 填入等号右方的  $2 \times 2$  矩阵对应位置之中，此即为特征值。通过红色框窗不断地往右往下位移(此例为 1)再与同一卷积核作内积，可得到其他特征值并逐一填入等号右方的矩阵；该矩阵纪录撷取得到的特征值，将保留到下一层作后续计算。

$$\begin{array}{|c|c|c|c|} \hline 1 & 1 & 0 & 1 \\ \hline 1 & 0 & 1 & 0 \\ \hline 1 & 0 & 1 & 0 \\ \hline 0 & 1 & 1 & 0 \\ \hline \end{array} * \begin{array}{|c|c|c|} \hline 1 & 0 & 1 \\ \hline 0 & 1 & 0 \\ \hline 1 & 0 & 1 \\ \hline \end{array} = \begin{array}{|c|c|} \hline 3 & 3 \\ \hline 4 & 2 \\ \hline \end{array}$$

Image                      Convolutional kernel

**Figure 1.** Convolution operation example (the 2-by-2 matrix on the right of the equal sign is the eigenvalue extracted from the convolution kernel according to the original image)

**图 1.** 卷积运作范例(等号右方  $2 \times 2$  矩阵为依原图像通过卷积核所撷取之特征值)

卷积层找出特征后，即能作为输入变量进行一般的类神经网络进行分类模型训练，但网络结构愈趋复杂，样本数若不够庞大，则容易产生“过度拟合”的问题。因此引入“池化”或局部取样(subsampling)的技巧，亦即在卷积核中再通过  $n \times n$  的小区域进行汇总，以凸显该区域的最显著特征。关于池化的技巧，可在很多应用文献中找到，此不赘述。

AlexNet 使用双 GPU (Graphics Processing Unit, 图像处理器)来训练类神经网络，因此将整个网络分散在两个 GPU 执行，再通过现行 GPU 架构的特性，能够让 GPU 直接存取另一个 GPU 的内存而不需通过主机的内存进行存取。这一特性可用来实现平行运算。

由上可知，AlexNet 的深度学习技术可以将数据通过多个处理层的线性或非线性转换，自动萃取出足以代表数据特性的特征。深度学习具有自动萃取特征的能力，因此也是一种特征学习的技术。

### 3. 图像预处理

通过无人机收集外墙或饰面材料裂缝图像数据并作处理，要先建构一个处理图像的环境，将数据通过  $224 \times 224$  的方式撷取，撷取完毕之后再将图像输入到 PC 主机端并预处理图像。之所以选择  $224 \times 224$  作为处理单位，是多次运算后基于预算量与检测率综合考虑而定的。图像预处理的目的是，在于将无人机拍摄的 RGB 三信道原始图像转为颜色相对单纯、但仍能保留原图像重要信息之形式，以利于机器学习软

件系统能更有效地捕捉其重要特征，而不至于因图像 RGB 颜色差异而干扰学习。之所以这样，是因为在研究进行过程中，多次实验显示机器学习的效果不如预期，后经详细检视发现，AlexNet 误将壁砖颜色视为需学习的特征之一，但壁砖颜色在本研究并非关键特征，因此先对图像执行预处理。

图像预处理的步骤与算法如下：

1) 灰度化处理(Gray-level Transformation): 公式为

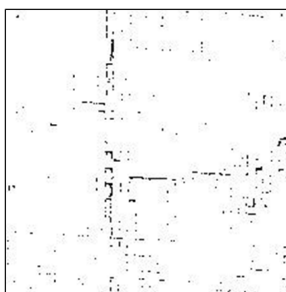
$$Gray = 0.299 \times R + 0.587 \times G + 0.114 \times B,$$

如图 2 所示，右图为左图灰度化处理后的示意图。

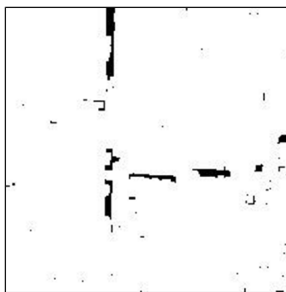


**Figure 2.** The original image and its grayscale processing results  
**图 2.** 原始图像及其灰度化处理结果

2) 图像二值化：二值化主要是将灰度图转化为黑白图，这个程序主要是将研究者感兴趣的部分与不感兴趣的部分区隔出来；通过设定灰度阈值，能将复杂的环境去除掉并且只显示裂缝的部分。利用计算软件 MATLAB (版本 2018a)的自适应(adaptive)二值化函数，针对不同的图像选择各自合适的阈值，可得到较理想的二值化结果。如图 3 所示，为图 2 的自适应二值化结果。



**Figure 3.** Adaptive binarization results of gray scale graph  
**图 3.** 灰度图的自适应二值化结果



**Figure 4.** Result of open operation  
**图 4.** 开运算结果

3) 形态学运算处理: 形态学用于二值化后的图像, 通过利用数学处理来凸显图像的形状或特征(如边界或细线等)。形态学最基本的运算包括膨胀(dilation)、侵蚀(erosion)、开运算(opening)、闭运算(closing)。本文使用开运算, 主要是希望能够强化裂缝的轮廓并且滤掉部分杂点。结果如图 4 所示。

## 4. 数据分类与参数调整

### 4.1. 图像数据收集

通过无人机以拍摄的方式收集原始图像数据, 未处理过的数据是整张未经撷取的图片, 如图 5 所示, 其中的红色小框为待处理的撷取范围的示意图。



**Figure 5.** Original image of crack in exterior wall  
**图 5.** 大楼磁砖裂缝原始图像



**Figure 6.** A schematic diagram of a crack  
**图 6.** 某处裂缝示意图

原始图像在被处理前, 先刻意摒除容易造成干扰的因素(如窗户、树木等与壁砖劣化无关的部位), 再一律以  $224 \times 224$  的分辨率撷取原始图片的磁砖裂缝部分, 撷取后将会获得数百张类似的局部图片, 图 6 为其中之一。

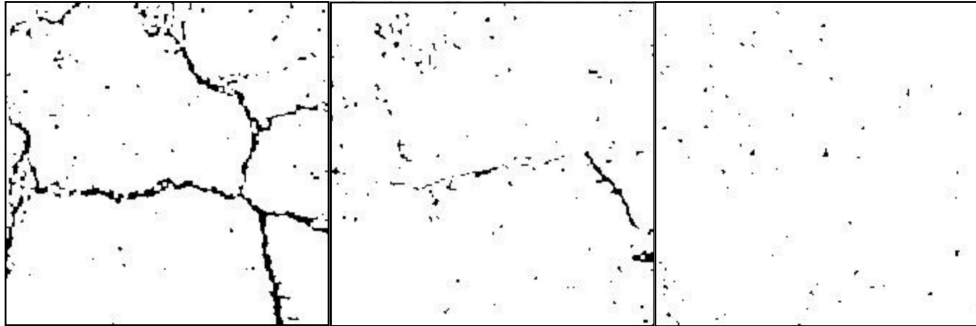
本文的外墙裂缝图像采集于城市公共建筑, 如图书馆、体育馆、商业大厦等, 图 7 是其中的几个裂缝实例。



**Figure 7.** Four examples of cracks in exterior wall  
**图 7.** 四个外墙(瓷砖)裂缝实例

## 4.2. 数据分类

所有图片进行预处理程序之后, 再按卷标分类, 比如分为严重、一般和轻微三个等级。图片类别卷标的范例如图 8 所示。设定卷标完成之后, 将数据分别放置于训练集与测试集两个文件夹内以利于后续存取。

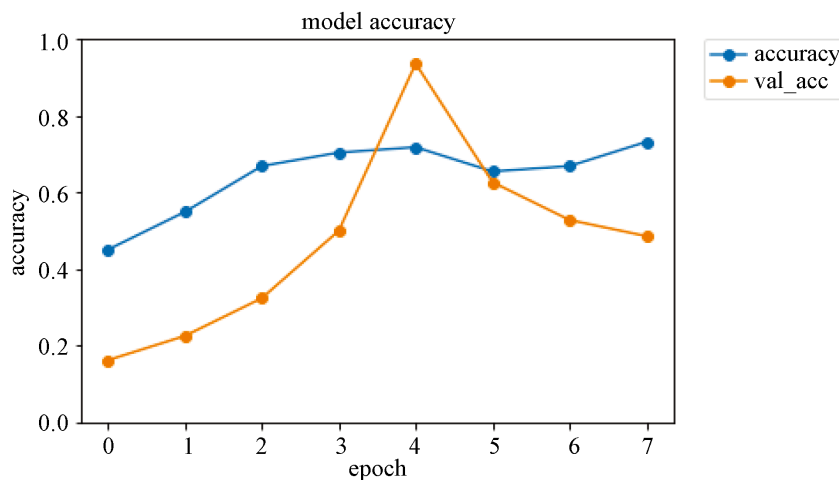


**Figure 8.** Volume labels are set to severe, general, and minor levels from left to right  
**图 8.** 从左至右依次将卷标设为严重、一般、轻微级别

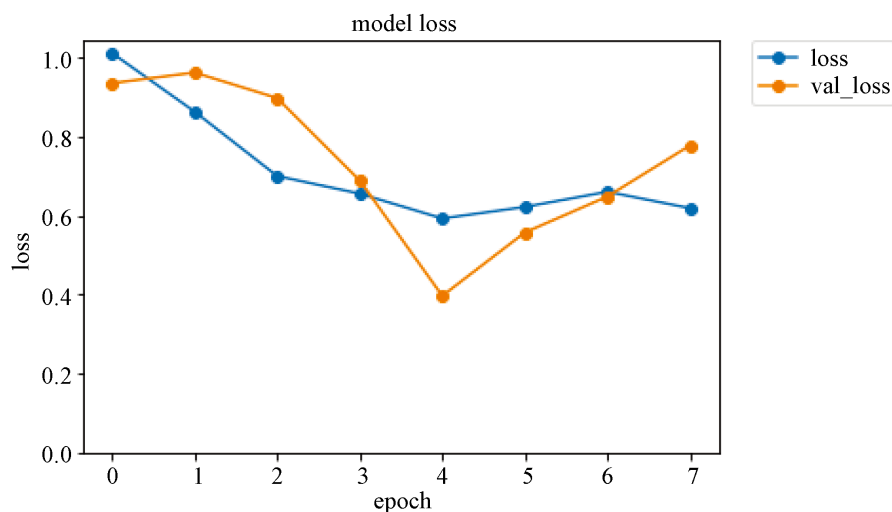
鉴于计算软件 MATLAB (版本 2018a) 具有强大的图像处理能力, 可以考虑用其对图像进行批量预处理。在预处理完成后, 开始建构 CNN, 开发环境需安装 Python、Tensorflow 与 Keras, 这样的组合为目前最常见。为了加速整个训练的流程, 同时在 PC 端使用 GPU 进行训练; 最后选择使用 AlexNet 作为训练的核心网络。

## 4.3. 训练与调整

通过输入不同的数据集与校正参数, 来训练 AlexNet 并获取适当的模型。基本上, 可以通过观察训练结果学习曲线(accuracy)与损失曲线(loss)的形态来决定参数是否可行。如图 9 所示, 为显示训练集与验证集的学习曲线, 分别用蓝色(accuracy)与橘色(val\_acc)标识, 数值越靠近 1.0 则代表模型训练越好越正确。图 10 的损失曲线则代表着模型的错误率, 数值越低则代表着图片的判别错误的机率越低。图的横轴 Epoch 代表整个数据集输入到网络的训练周期。

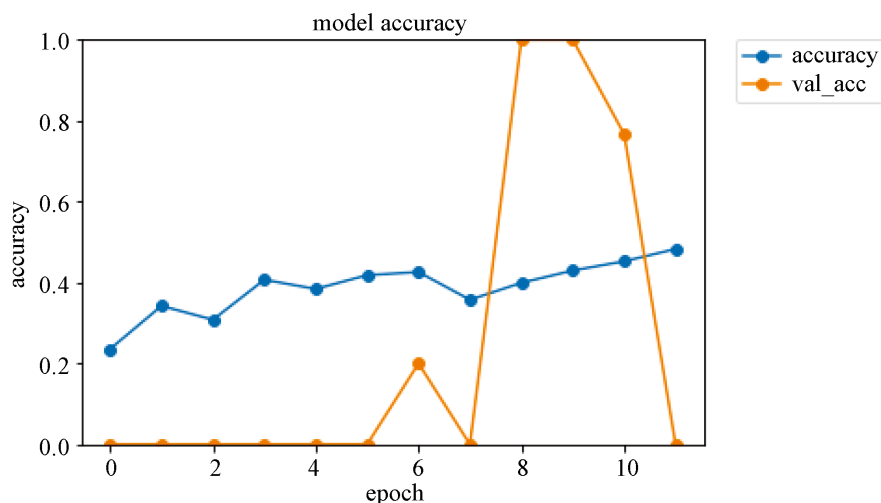


**Figure 9.** Accuracy learning curve  
**图 9.** Accuracy 学习曲线图



**Figure 10.** Loss curve  
**图 10.** Loss 损失曲线图

验证集与训练集的资料都是从训练集的图片分类而得，验证集跟训练集的多寡会影响模型训练出来的结果。通过上述曲线，可作两个观察。第一，如图 11 所示，若测试集的曲线慢慢成长，但验证集的曲线并无明显增幅或发生弹跳(epoch ≥ 5 的情境下)，则表明出现了过度拟合(over-fitting)；其发生的主要原因，是学习过程太久或训练的样本太少；一旦发生这种情形，需尝试改变数据集，例如增加图片或者是调整参数。另一种如图 12 所示，损失曲线的训练集曲线随着 epoch 增加而慢慢的下降，但验证集的曲线未见相应的增长或呈现弹跳现象；这种情况可能是因为图片的卷标搞错，导致模型无法正确学习，可以返回去检查图片的卷标是否存在分类错误。



**Figure 11.** Over-fitting curve  
**图 11.** 过度拟合曲线图

参数调整是使用类神经网络训练必须面对的课题，本研究的核心参数为“学习率” $\gamma$ ，在此可以通过观察 Loss 损失曲线图来调整；另外，还可以调整训练集与验证集的分配比例，通过这一调整可以让模型更加正确地学习。

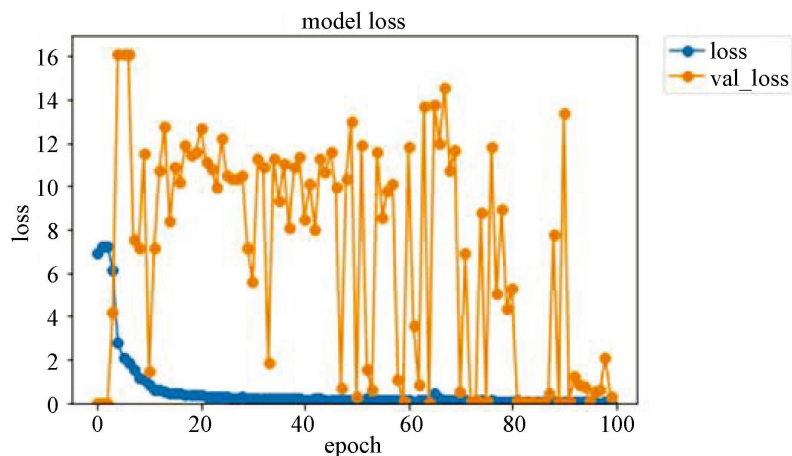


Figure 12. Curve of uncorrectly marked volume

图 12. 卷标有问题的曲线图

## 5. 图像处理的结果

### 5.1. 实验执行步骤

为了测试方法的效能，实验执行步骤如下：

- 1) 将测试数据转换成  $224 \times 224$  维度，无论任何大小都须转换；
- 2) 确认测试数据的完整性；
- 3) 辨识图片；
- 4) 统计验证结果；
- 5) 比较每个模型的错误率。

### 5.2. 按批次测试

将测试数据分成 5 个批次，都标示正确的标签，每一批测试图片为 20 张，所以总共有 100 笔的测试资料。通过统计预测的结果是否符合卷标，与观察模型的学习曲线来验证模型可用性。表 1 为每一个批次测试集类别的张数，通过输入不同类型的测试集来进行验证，从而避免了图像选择的特定性；其中，数据集四(DataSet4)和数据集五(DataSet5)特定类别的测试张数为零，最主要的用意是为了确认特定类别的错误张数不会偏高。

Table 1. Three test sets

表 1. 三类测试集

测试集 \ 卷标	DataSet1	DataSet2	DataSet3	DataSet4	DataSet5
轻微(slight)	5	5	10	10	0
一般(fair)	5	10	5	0	10
严重(serious)	10	5	5	10	10

#### 5.2.1. 未经预处理的测试集

未经过预处理的训练集的学习曲线(正确率数值)与损失曲线(错误率数值)分别如图 13 和图 14 所示。由图 13 可见，验证结果与实际的训练差异太大；由图 14，可以观察到错误率也居高不下。因此，未经



预处理的测试集的正确识别率差强人意，这也可以观察到预处理模型必要性。此外，也可以很明显地了解到，在数据量少的情况下分成三类，再加上每个类别的共通性又不大的情况下，导致整个模型无法正确地训练。

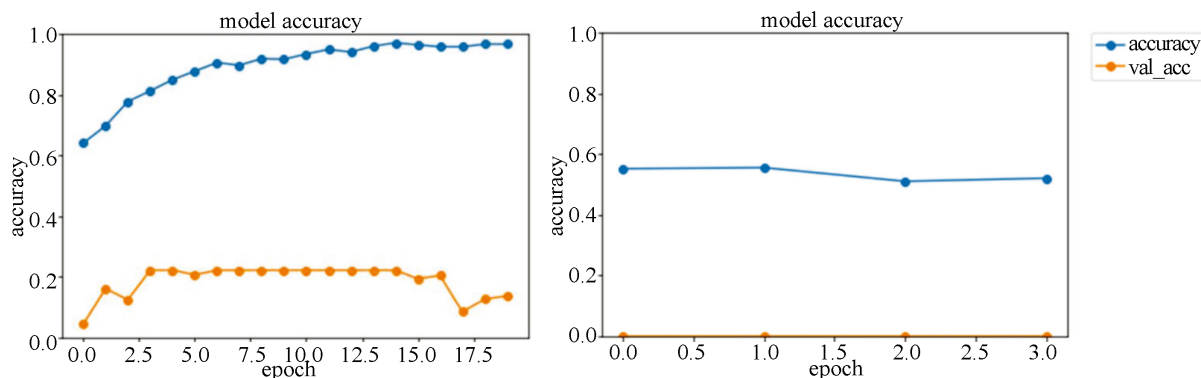


Figure 13. Unprocessed learning curve  
图 13. 未预处理的学习曲线

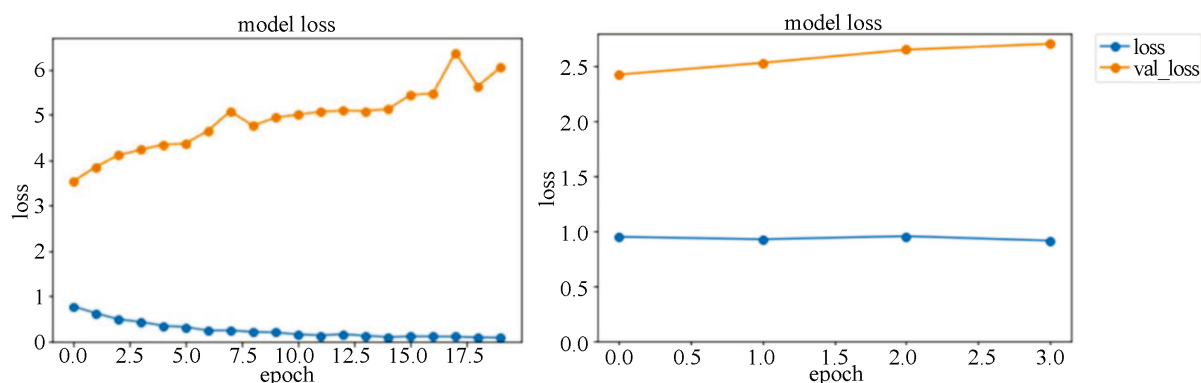


Figure 14. Unprocessed loss curve  
图 14. 未预处理的损失曲线

### 5.2.2. 经过预处理的训练集

经过预处理后，三类卷标的正确识别率得到了有效的提高，对比结果如表 2 所示。显然，经过预处理图像，其识别率大大高于未经预处理的图像。

Table 2. Comparison of identification rates of three types of test sets  
表 2. 三类测试集识别率比较

卷标	预处理后	未预处理	成长幅度
轻微(slight)	97%	66%	46.97%
一般(fair)	91%	33%	175.76%
严重(serious)	86%	44%	95.45%

## 6. 结论与后续研究

本文通过设计、测试与实验，基于 CNN 的高楼外墙及饰面材料裂缝检测技术，建构一个能够快速检测的系统，让使用者能够有效地辨识裂缝的破裂程度。利用无人机，通过简易的操作，使得普通民众也

能快速处理,可大幅地减少工人攀爬高楼的频次,可大幅地减少人力的损耗。特别是针对地震等自然灾害现场,通过无人机搭配本文的识别系统,可达到安全检测、快速识别等目的。

本研究尚有若干方向可改进,包括但不限于:

- 1) 增加训练集的数据量,让辨识的正确率更加提升,或是缩减神经网络的架构;
- 2) 优化二值化的程序,更加精准地凸显人们感兴趣的特征;
- 3) 对于数据集的标签需要提出更标准化的评估分类指标;
- 4) 对特定材料(如玻璃幕墙、外墙石材等)进行识别或过滤,以避免辨识错误;
- 5) 实现更高度的自动辨识功能,能够从影片自动圈选出磁砖裂缝部分,以达到更方便的检视;
- 6) 可以将本识别系统直接内置于无人机,使其能够做到实时的辨识。

## 参考文献

- [1] Tensorflow. <https://www.tensorflow.org/>
- [2] Keras Documentaion. <https://keras.io/>
- [3] Léculn, Y., Bottou, L., Bengio, Y., et al. (1998) Gradient-Based Learning Applied to Document Recognition. *Proceedings of the IEEE*, **86**, 2278-2324. <https://doi.org/10.1109/5.726791>
- [4] Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2012) ImageNet Classification with Deep Convolutional Neural Networks. International Conference on Neural Information Processing Systems. Curran Associates Inc., 1097-1105.
- [5] Lin, M., Chen, Q. and Yan, S. (2013) Network in Network. *Computer Science*.
- [6] Schmidhuber, J. (2015) Deep Learning in Neural Networks: An Overview. *Neural Networks*, **61**, 85-117. <https://doi.org/10.1016/j.neunet.2014.09.003>
- [7] Vu, H.T. and Huang, C.C. (2017) A Multi-Task Convolutional Neural Network with Spatial Transform for Parking Space Detection. *IEEE International Conference on Image Processing*, IEEE, 1762-1766. <https://doi.org/10.1109/ICIP.2017.8296584>
- [8] 肖进胜, 刘恩雨, 朱力, 等. 改进的基于卷积神经网络的图像超分辨率算法[J]. 光学学报, 2017(3): 96-104.
- [9] 周敏, 史振威, 丁火平. 遥感图像飞机目标分类的卷积神经网络方法[J]. 中国图象图形学报, 2017, 22(5): 702-708.
- [10] 葛芸, 江顺亮, 叶发茂, 等. 基于 ImageNet 预训练卷积神经网络的遥感图像检索[J]. 武汉大学学报(信息科学版), 2018, 43(1): 67-73.
- [11] 傅隆生, 冯亚利, Elkamil, 等. 基于卷积神经网络的田间多簇猕猴桃图像识别方法[J]. 农业工程学报, 2018, 34(2): 205-211.

### 知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>  
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2325-2286, 即可查询
2. 打开知网首页 <http://cnki.net/>  
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: [sea@hanspub.org](mailto:sea@hanspub.org)