Computers and Electronics in Agriculture 150 (2018) 26–32

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

Deep recursive super resolution network with Laplacian Pyramid for better agricultural pest surveillance and detection

Yi Yue*,[1], Xi Cheng*,[1], Di Zhang[1], Yunzhi Wu, Yang Zhao, Yiqiong Chen, Guohua Fan, Youhua Zhang*

College of Information and Computer, Anhui Agricultural University, China

ARTICLE INFO

Keywords:
Agricultural pest
Super resolution
Object detection
Deep convolutional neural networks
Laplacian Pyramid

ABSTRACT

Computer vision technologies greatly improved the efficiency of recognizing and controlling of agricultural pests. However, the density of cameras deployed in the farmland is usually sparse and the images or videos of agricultural pests collected are often obscure. This always results in low resolution of pests in the pictures, making them difficult to observe and monitor. In addition, the existing object detection method is not satisfactory for the detection of small targets with low pixel resolution. Therefore, it is necessary to restore and upsample the collected images so as to improve the recall rate of the detection. In this work, we proposed a novel super resolution model based on deep recursive residual network. Compared with the traditional interpolation methods and the models with shallow convolutional neural networks, the method we proposed is more powerful in image reconstruction and achieves the state of the art performance. The experimental results show that our method greatly improved the recall rate of pest detection by 202.06%. In addition, compared with image upscaling methods such as Bicubic Interpolation and Super-Resolution Convolutional Neural Network (SRCNN), our method is average 111.31% and 41.89% improved respectively. The model we put forward could reduce the density of the camera layout of the agricultural Internet of Things (IOT) monitoring systems and reduce cost of infrastructure, which is of high practical value.

1. Introduction

In order to improve the efficiency of agricultural production, it is of great significance to monitor, recognize and control agricultural production processes and agricultural pests with computer vision and deep learning. Nowadays, agricultural management tasks via computer vision is becoming a hot research topic which greatly improved the speed compared to manual methods. These tasks including classification (Huang and Du, 2008; Du et al., 2006), detection (Mundada and Gohokar, 2013) and segmentation (Wang et al., 2010) for plants and pests. Many machine-learning based pest recognition methods were proposed in the recent years. An insect classifier was developed based on multitasking sparse coding and multiple kernel learning by Xie et al. (2015) and an insect taxonomic system was constructed using sparse coding based spatial pyramid was developed subsequently (Xie et al., 2016). Zhu and Zhang (2010) researched on color features of insects and proposed an insect identification method based on color histogram and Gray-level co-occurrence matrix. Along with the development of deep learning (Krizhevsky et al., 2012), features of pests could be automatically extracted by convolutional neural networks (CNN) and the recognition performance in further enhanced. Cheng et al. (2017) proposed a deep residual learning method with residual convolutional neural networks and achieved high recognition accuracy. Clear and high-resolution images not only make the farmland easy to observe, but also make the pest recognition systems achieve high recognition accuracy. However, in fact, the arrangement of the cameras in the farmland is relatively sparse, even if the pests were captured by these cameras, they will be very small in pixel and have low resolution. Although the existing object detection method is quite mature for the detection of large objects such as people, pets and vehicles, the detection for smaller pixel scales still could not reach a satisfactory level. In pest identification tasks, it is unrealistic to place surveillance cameras close to target pests, and these videos or images captured by agricultural Internet of Things (IOT) monitoring equipment placed in fixed locations are often not sufficiently clear. The pest morphology is usually relatively small. It means that the existing object detection methods could not completely and accurately detect all the pests in the monitor.

Single image super resolution is an important low-level computer

* Corresponding authors.
E-mail addresses: yyyue@ahau.edu.cn (Y. Yue), opteroncx@ahau.edu.cn (X. Cheng), zhangyh@ahau.edu.cn (Y. Zhang).
[1] These authors have contributed equally to this work.

https://doi.org/10.1016/j.compag.2018.04.004
Received 4 February 2018; Received in revised form 22 March 2018; Accepted 3 April 2018
0168-1699/

vision task, which is aimed at restoring unclear and low-resolution images to high-resolution images. Super resolution has a very high application prospect in medical images, photo repair, video games and many other fields. Yang et al. (2010, 2012) developed an image super-resolution system via sparse encoding. Later, the A + and IA models was proposed by Timofte et al. (2014, 2016), using anchored neighborhood regression to build the super-resolution model. To further improve the image reconstruction quality, an image upscaling model was proposed based on self-exemplars by Huang et al. (2015). The Super-Resolution Convolutional Neural Network (SRCNN) model was proposed by Dong et al. (2016), which for the first time uses convolutional neural networks for single-image super-resolution. For fast super resolution on videos, a model of efficient sub-pixel convolutional neural network was proposed by Shi et al. (2016), which realizes the ability of real-time processing. For multiscale model training, the Laplacian Pyramid based LapSRN model was proposed by Lai et al. (2017). As depth deepens, the performance of neural networks continues to grow (Simonyan and Zisserman, 2014), the VDSR with very deep structure was proposed by Kim et al. (2016a,b). Neural network structures have become more diverse in recent years, and the Deeply-Recursive Convolutional Network (DRCN) (Kim et al., 2016a,b) with recursive structures was subsequently proposed. In addition, with the advent of Generative Adversarial Networks(GANs) (Ledig et al., 2016; Kaae Sønderby et al., 2016) and Perceptual Loss (Johnson et al., 2016), reconstructed images are further enhanced in terms of visual perception.

In order to enhance the pest detection and recognition systems, low resolution pest images need to be upscaled. Therefore, this work proposed a fast and accuracy super resolution algorithm based on deep recursive residual network and Laplacian Pyramid to restore low resolution pest images taken from farmland.

2. Proposed method

2.1. Single-scale super-resolution model structure

The overall structure of our proposed super resolution model is shown in Fig. 1, the model used local residual learning and global residual learning. In each block, we follow the design of the residual block of He et al. (2016), stacking two convolutional layers and one excitation layer, in addition we add a skip connection to form the local residual learning. Global residual learning is mainly used for image reconstruction, which will be described in detail in Section 2.3.

For deep networks, it is a big challenge in how to conduct effective training because of the tendency of gradient vanishing or gradient explosion. Inspired by DenseNet (Huang et al., 2017) and MemNet (Tai et al., 2017), we build the model with densely connected structure. The low-level convolution feature is passed to the upper layer and concatenated with it, superimposing the channel, acting as a signal gain. In Fig. 1, features from low levels are stacked to the output of each block and each recursive unit. In the experiment we also use a 1 × 1 convolution layer to construct the transition layer, which remaps the superimposed multichannel feature map back to the specified number of channels and then outputs to the deconvolution module for upsampling. The model used in the experiment contains 8 units and each unit contains 2 recursive blocks. In addition, if the model needs to be extended, the model depth calculation can be expressed by formula (1):

$$\text{Depth} = BU*(R*3+3) \tag{1}$$

In the formula, B represents the number of branches of various super resolution scales in Laplacian Pyramid, U represents the number of units and R represents the number of recursive blocks in each unit.

The activation function in the network structure uses Leaky ReLU (Maas et al., 2013; Xu et al., 2015). Compared with the traditional Sigmoid, it can effectively prevent the gradient vanishing phenomenon in the gradient transmission, which is beneficial to the training of deep stacking convolutional networks. Leaky ReLU is multiplied by a factor p for values less than 0 as opposed to ReLU ignoring these values directly. Leaky ReLU can be expressed by the following formula.

$$y_i = \begin{cases} px_i & (x \leqslant 0) \\ x_i & (x > 0) \end{cases} \tag{2}$$

Each Block can be expressed as the formula below:

$$x_{i+1} = W_2 LReLU(W_1 x_i) + x_i \tag{3}$$

where x_i is the output of the previous block and x_{i+1} is the output of the current block, W_1 and W_2 represent the two convolution layers as shown in Fig. 1 and LReLU is the Leaky ReLU used as the activation function in our model.

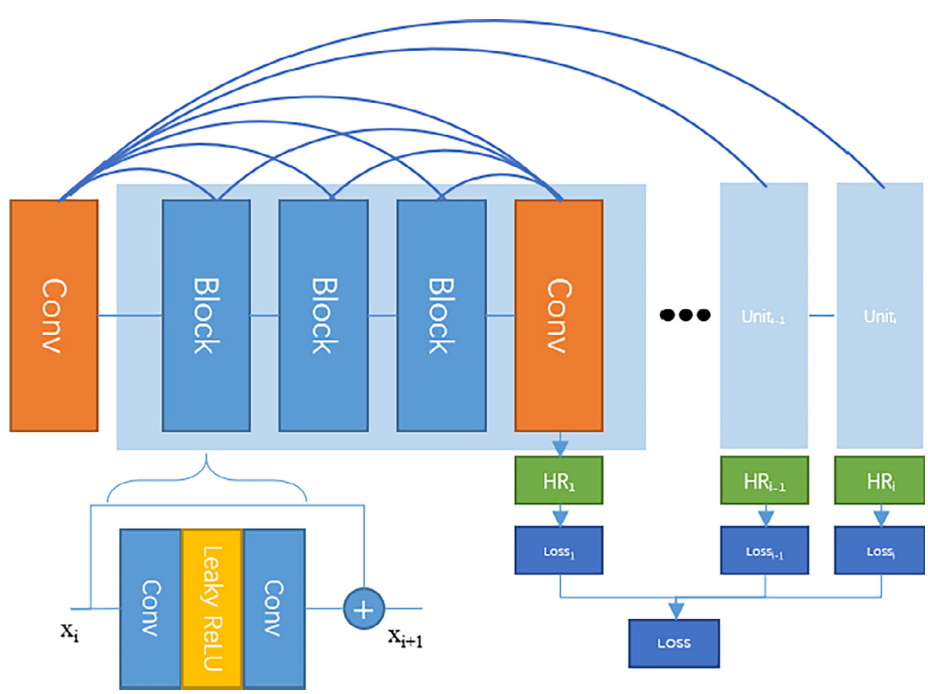

Fig. 1. Overview of the model structure.