



[首页](#)[资讯](#)[深度资源](#)[产业视频](#)[GMIS峰会](#)[AI 商用搜索](#)

[登录](#)/[注册](#)

SEARCH

轻量级深度卷积神经网络在农业机器人中的应用

By [机器之心](#) 2017年6月01日 00:15

最近，深度学习正随着训练过程简单化和准确化而变得越来越流行。对于旨在将研究转换为工业应用的研究者而言，这是一个至关重要的需求，例如机器人。

在本文中，作者提出了学习轻量级模型的一个非常新颖的方式，在实现精度超过 90% 的同时让训练速度快了一个数量级，让参数也少了一个数量级。作者主要使用了以下三个步骤：

- 1). 将预先训练好的模型适应于手头的任务中。
- 2). 在适应的模型中使用模型压缩技术来学习到一个具有更少参数的轻量级深度卷积神经网络 (DCNN)
- 3). 作者将 K 个轻量级模型结合起来作为一个混合模型来提高轻量级模型的性能

最后，这个方法被用在了农业机器人中，并实现了良好的性能。

简介

如今农业机器人的使用在稳步增长，例如图 1 所示的 AgBot II [1]，它帮助进行杂草检测和分类，以及 Harvey[2]，可以用来进行检测和分割作物。目前的杂草分割方法是把形状和像素统计特征相结合，使用一个随机森林分类器来做分类。因为机器人平台的资源是有限的，所以很难部署深度卷积神经网络，并且用有限的数据去训练深度神经网络也是很困难的。使用最先进的网络所需要的计算能力是那些常见机器人负担不起的。这篇文章在复杂度和准确度之间做了一个折中，并通过以下步骤解决了这个问题。首先，开发者把一个预先训练好的模型 Inception-v3[3] 适应在这个任务中。然后，他们使用模型压缩和「蒸馏」技术实现了指数级减少的参数。最后，基于之前的工作 [4]，作者将一组 K 个轻量级模型结合成了一个混合模型来提升性能。



Fig. 1. Top is an image of AgBot II which performs weed management and bottom is Harvey a robotic sweet pepper harvester.

图 1：上边是 AgBot II，它在执行杂草管理；下面是 Harvey，它是一个机器人甜椒收割机

这个方法在杂草分割上实现了可观的结果。Adapted-IV3 模型将准确率从 85.9% 提升到了 90.3。并且，对于 $K=4$ 的轻量级深度卷积神经网络，它能够在使用更少的参数和更快的帧率时实现 90.3% 的准确率，

作者提供了一份详实的综述，论述了这个新颖思想的历史以及导致这个思想的方法。他们还提供了对目前的学习特征和模型压缩这两个技术的趋势的优劣分析。这不是这篇论文的重点，所以在这里不会深入细节。如果你对这篇论文感兴趣，它值得一读。

提出的方法

新方法主要是一个被分为三步的过程，这个过程在速度和内存大小之间做了精确的折中。这个方法被用来有效地解决诸如 AgBot II 这种机器人平台的杂草分割问题。滑动窗口的维度在彩色图像上是 $81 \times 81 \times 3$ （因为彩色图像是三通道的），窗口中间的像素就被判断为要么是作物，要么是杂草。由于大多数机器人视觉问题都是稀疏问题，所以全连接神经网络（FCN）没有被使用。这篇论文中的稀疏问题指的是杂草分割仅仅需要对包含植被的像素做出分类，所以作者做出了关于他们为何没有使用全连接神经网络的解释，因为全连接神经网络是针对密集决策（对每一个像素）来设计和训练的。作者用图 2 解释了什么是稀疏问题。

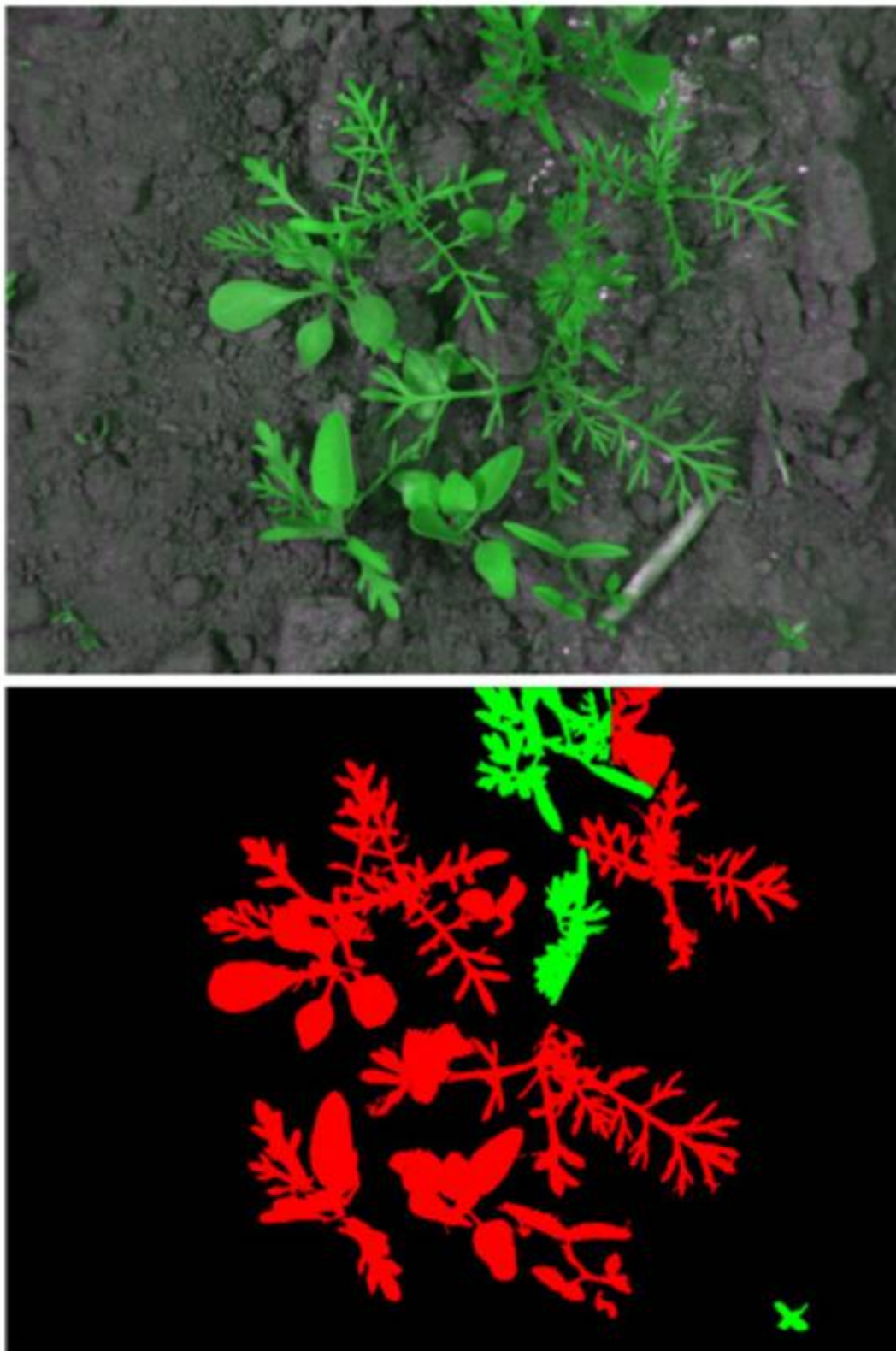


Fig. 2. An example of the challenging weed vs crop segmentation problem. Top is the original image and bottom is the associated ground truth where green represents the crop and red represents the weed.

图 2：一个具有挑战性的杂草 vs 作物的分割问题。上面的图片是原始图像，下面是对应的事实，其中绿色代表作物，红色代表杂草

现在我要简要地描述一下这三个摘要过程。

A. 迁移学习：适应复杂的预训练网络

就像在这个例子中一样，如果你只有有限的数据来训练一个神经网络，最好的方法就是拿一个预先训练好的模型并将其适应在你的任务中。有名的几个常被使用的预训练网络是：VGGnet、GoogleNet 等等。在这篇论文中，作者使用了 GoogleNet 的最新版本-- Inception-v3，与其他模型相比，它是一个更小的模型。随后作者将原始图像进行了重采样，以让它与 Inception-v3 模型所需要的尺寸相匹配。

B. 模型压缩：训练轻量级深度卷积神经网络

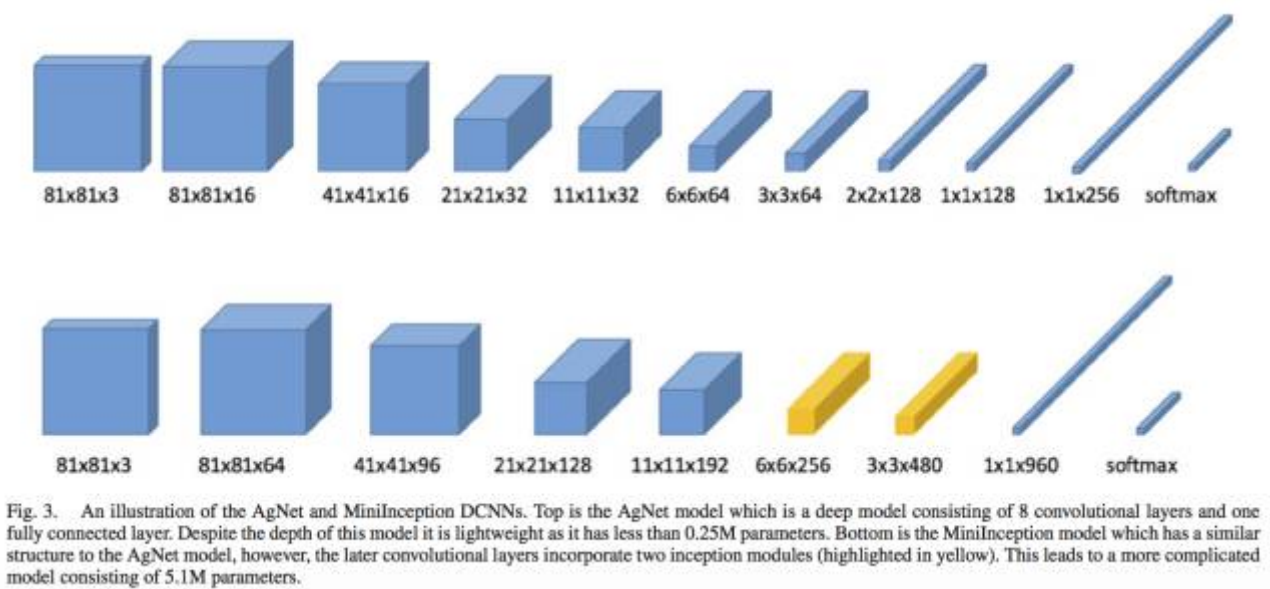


图 3：对 AgNet 和 MiniInception DCNNs 的描述。上面是一个 AgNet 模型，它是一个由 8 个深度卷积层和一个全连接层组成的深度模型。模型的深度是轻量级的，并且拥有少于 25 万的参数。下面是 MiniInception 模型，它与 AgNet 有着类似的结构，然而，后者的卷积层结合了两个感知模型 (图中高亮的部分)。这使得模型会更加复杂，拥有 510 万个参数。

尽管 Inception-v3 拥有 25 万个参数（这已经远远少于其他模型了），但是它仍然需要模型压缩技术。这里，将过程 A 中的模型作为「教师网络」，然后从中学习到一个轻量级的「学生深度卷积神经网络」。为了训练这个模型，使用了学生网络和事实之间的损失函数，以及教师网络和学生网络的逻辑输出之间的 L2 损失。作者考虑了两种潜在的结构，第一个包含 8 个卷积层和 1 个全连接层，它与 AlexNet 类似，你可以在图 3 中看到细节。第二个网络结构包含 4 个卷积层，并在后面还跟随着 2 个感知模型和一个全连接层 (类似于 GoogleNet)，你可以在图 3 的底部看到。图 4 中描述了 4 个子模型。你可以在图 4 的解释中看到细节。

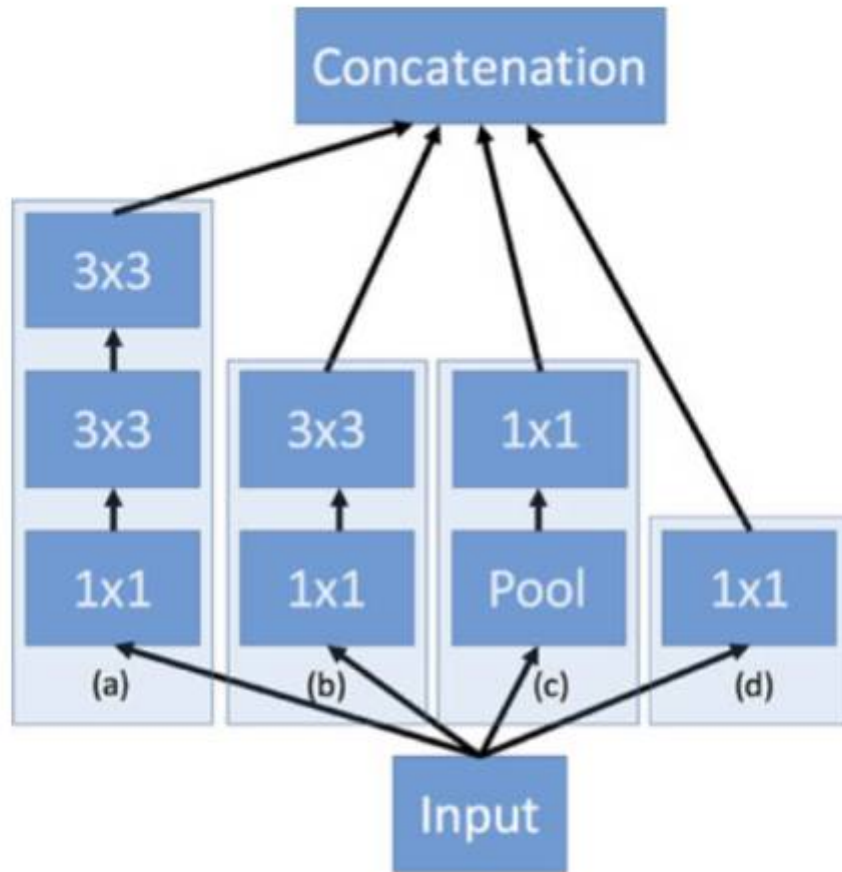


Fig. 4. An illustration of the Inception module that we use, it consists of four sub-modules (a)–(d). The kernel size for each of the sub-modules is smaller than the input, however, by concatenating them at the end the output feature vector is larger. Module (a) acts as a 5×5 convolution that incorporates downsampling of 2, it is calculated efficiently by stacking two 3×3 convolutions. Module (b) is a 3×3 convolution that downsamples by a factor of 2 and (c) first pools to downsample by a factor of 2 followed by a 1×1 convolution. Module (d) is a 1×1 convolution that downsamples the image by a factor of 2.

图 4：我们所用的感知模型，它包括四个子模型 (a)-(d)

参数的细节如表 1 所示。

表 I：用在子模型中的两个感知网络的参数数量

TABLE I
THE NUMBER OF PARAMETERS FOR THE DIFFERENT PARTS OF THE
SUB-MODULES FOR THE TWO INCEPTION MODULES USED FOR
THE MININCEPTION DCNN

		Sub-Module			
		(a)	(b)	(c)	(d)
Inception Module 1	(1×1)	16	96	32	64
	(3×3)	32	128		
	(3×3)	32			
Inception Module 2	(1×1)	32	128	64	128
	(3×3)	96	192		
	(3×3)	96			

For Sub-module (a) the Number of Parameters are (from Top to Bottom) for the *1times1* Layer, and then for the two 3×3 Layers. For the Sub-module (b) the Number of Parameters are for the *1times1* Layer, and then the 3×3 Layer. For Sub-modules (c) and (d) the Number of Parameters are for the 1×1 Layers.

C. 深度卷积神经网络的混合和组合

在这篇论文中，通过结合这些轻量级模型提升了性能。通过平均决策，它可以以较低的成本提升性能。首先，基于最大逻辑值计算占用概率：

$$\alpha_k^t = \frac{\exp \{C_k\}}{\sum_{j=1}^K \exp \{C_j\}} \quad (1)$$

然后，如公式 2 所示，我们得到了最终的分类，它是在对应的占用概率下的加权和。

$$c_n = \sum_{k=1}^K \alpha_k^t c_{k,n}, \quad (2)$$

作者利用这个方法解决了机器人平台中的杂草分割问题。并且这个模型在 tensorflow 中得到了实现。Adam 优化器的参数如下所示

- 学习率 (learning rate): $\gamma = 1e-4$, $\epsilon = 0.1$
- 批处理大小 (batch size) : $b = 60$
- 下降率 (drop out rate) :50%.

1. 杂草分割的准确度

作者使用了公开的训练数据及「Crop/Weed Field Image Dataset (CW-FID)」，它包括 20 张训练数据和 40 张测试数据。作者还将结果与 Haug 和 Ostermann 的工作 [5] 做了比较，Haug 和 Ostermann 在特征上训练了一个随机森林。我们可以在表 2 中看到细节结果，它表明文中提出的所有深度学习方案都要好于之前提出的方法。Adapted-IV3 模型得到了最高的准确率，大约是 93.9%，然而 Haug 和 Ostermann 达到的准确率是 85.9%。

表II：不同的模型在 CW-FID 数据集上的准确率

TABLE II
CLASSIFICATION ACCURACY FOR THE VARIOUS MODELS ON THE CROP/WEED FIELD IMAGE DATASET (CWFID) [7]

	Accuracy	Num. Params
Adapted-IV3	93.9%	25M
Mix-MiniInception ($K = 4$)	90.7%	20.4M
Mix-MiniInception ($K = 2$)	90.5%	10.2M
MiniInception	89.9% \pm 0.1%	5.1M
Mix-Agnet ($K = 8$)	90.5%	2.0M
Mix-Agnet ($K = 4$)	90.3%	1.0M
Mix-Agnet ($K = 2$)	89.7%	0.5M
AgNet	88.9% \pm 0.4%	0.25M
Shape+Stat. [7]	85.9%	N/A

Also presented is the model complexity in terms of number of parameters (Num. Params). The accuracy for AgNet is presented as the mean accuracy and standard deviation of the 8 trained AgNet models.

2. 轻量级深度卷积神经网络的混合

这里，作者还结合了多个轻量级的模型，例如 Mix-AgNet 和 Mix-MiniInception。与使用集合模型相比，MixDCNN 方法提供了 0.15% 的性能提升。对单 GPU 而言，也没有增加模型复杂度。在表 II 中，可以看到，通过增加 K（模型的数量），全局性能得到了提升，但是相对性能却降低了。K=4 的 AgNet DCNNs 在模型复杂度和准确度之间是一个不错的折中。在图 5 中你可以看到可视化的结果，描述了 3 个深度卷积神经网络模型：AgNet、MixAgNet（K = 4）以及 Adapted-IV3。

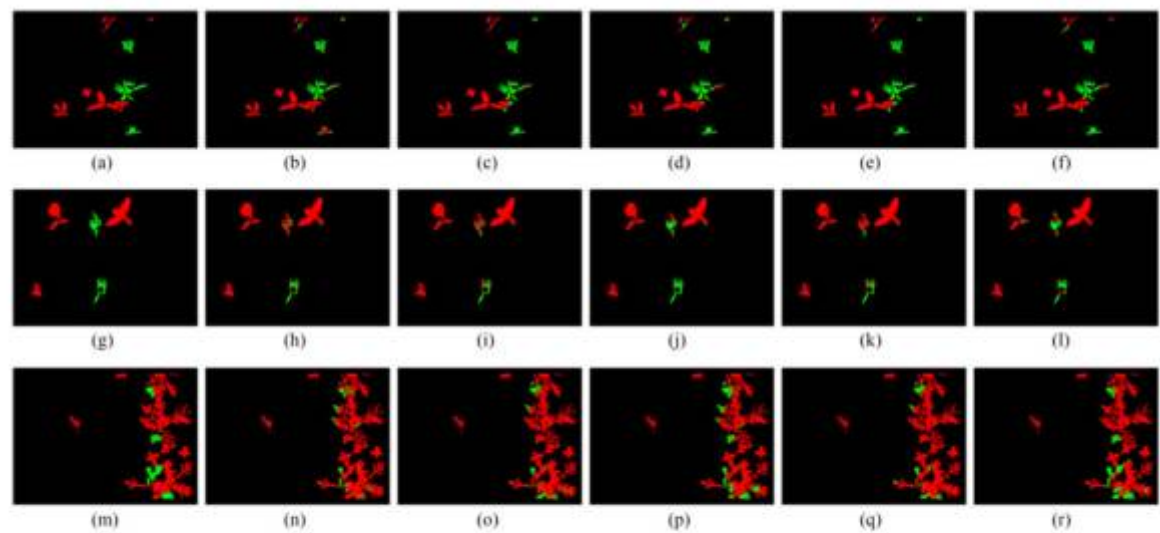


Fig. 5. Example output of the three different deep learning approaches, red represents the weeds and green represents the crop. From left to right is results for the ground truth, AgNet, MiniInception, Mix-AgNet ($K = 4$), Mix-MiniInception ($K = 2$) and the Adapted-IV3 models. Each row represents the result for a different image, where the last row represents a failure case for the lightweight models.

图 5：三种深度学习方法的例子：红色代表杂草，绿色代表作物。每一行从左到右依次是：实际结果、Agnnet 模型、MiniInception 模型、Mix-AgNet（K=4）、Mix-MiniInception（K=2）以及 Adapted-IV3 模型。每一行代表一张不同的图片，最后一行代表一个失败的案例。

3. 速度和模型复杂度

表 III 展示了每秒钟可以被处理的区域的数量。我们发现，模型越简单，每秒处理的区域数越多。复杂度会随着模型数量的增加而增加，这会降低每秒钟处理的区域数量。为了维持它们之间的平衡，我们选择了 K=4 时候的 Mix-AgNet 模型以及 K=2 时候的 Mix-MiniInception 模型。我们也可以发现，考虑到速度和复杂度，AgNet、Mix-AgNet 和 Mix-MiniInception 这三种方法也具有被部署到机器人平台上的潜力。

表 III：参数的数目以及每秒钟可以处理的区域的数量

TABLE III
NUMBER OF PARAMETERS (NUM. PARAMETER) AND TIMING IN TERMS OF
REGIONS PROCESSED PER SECOND (REG./SEC.)

	Num. Parameters	reg./sec.
Adapted-IV3	25M	155
Mix-MiniInception ($K = 4$)	20.4M	767
Mix-MiniInception ($K = 2$)	10.2M	1,437
MiniInception	5.1M	2,737
Mix-AgNet ($K = 8$)	2.0M	1,394
Mix-AgNet ($K = 4$)	1.0M	2,445
Mix-AgNet ($K = 2$)	0.5M	3,736
AgNet	0.25M	6,300

结论

这篇论文提供了一个可训练被用在机器人平台上的深度卷积神经网络的新方式。论文有着全面的文献综述和清晰的图表，这可以让读者更容易地理解背景信息以及实现这个方法的细节。尽管作者并没有告诉我们他们为什么要用这种方法，但是这篇论文给我们提供了一个思考和使用深度神经网络的新方法。正如作者在论文末尾建议的一样，我们可以尝试更多的方法来提升性能，但是这篇论文的价值在于它提供了可以被接受并且应用在日常生活中的研究结果。

参考文献

- [1] Bawden, Owen John. *Design of a lightweight, modular robotic vehicle for the sustainable intensification of broadacre agriculture*. Diss. Queensland University of Technology, 2015.
- [2] Lehnert, Christopher, et al. "Autonomous sweet pepper harvesting for protected cropping systems." *IEEE Robotics and Automation Letters* 2.2 (2017): 872-879.
- [3] Szegedy, Christian, et al. "Rethinking the inception architecture for computer vision." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016.
- [4] Ge, ZongYuan, et al. "Fine-grained classification via mixture of deep convolutional neural networks." *Applications of Computer Vision (WACV), 2016 IEEE Winter Conference on*. IEEE, 2016.
- [5] Haug, Sebastian, and Jörn Ostermann. "A crop/weed field image dataset for the evaluation of computer vision based precision agriculture tasks." *European Conference on Computer Vision*. Springer International Publishing, 2014.

更多有关GMIS 2017大会的内容，请「[点击这里](#)」查看机器之心官网 GMIS 专题↓↓↓



声明：本文由机器之心原创出品，版权归作者所有，转载请查看要求，机器之心对于违规侵权者保有法律追诉权。

[卷积神经网络农业机器人产业计算机视觉](#)



[提交评论](#)

登录后参与评论[去登录](#)

lyk_ffl_163_com

请问，该文章的原论文全名是什么？或者链接有吗？想阅读下原论文

2017-06-05 | 12:25 PM

[回复](#)

黄小天

不是一篇论文，请看一下参考文献。

2017-06-18 | 15:30 PM

[回复](#)



[关于我们寻求报道商务合作服务条款](#)

©2018版权所有 机器之心（北京）科技有限公司

京 ICP 备 12027496

全球人工智能信息服务

[友情链接](#)

[Synced Global](#)[机器之心](#) [Medium](#) [博客PaperWeekly](#)[网易智能](#)[动脉网](#)[硬蛋网](#)



联系电话：+86 010-57150141

联系邮箱：contact@jiqizhixin.com