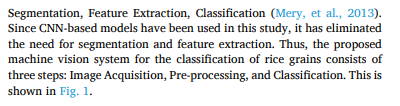
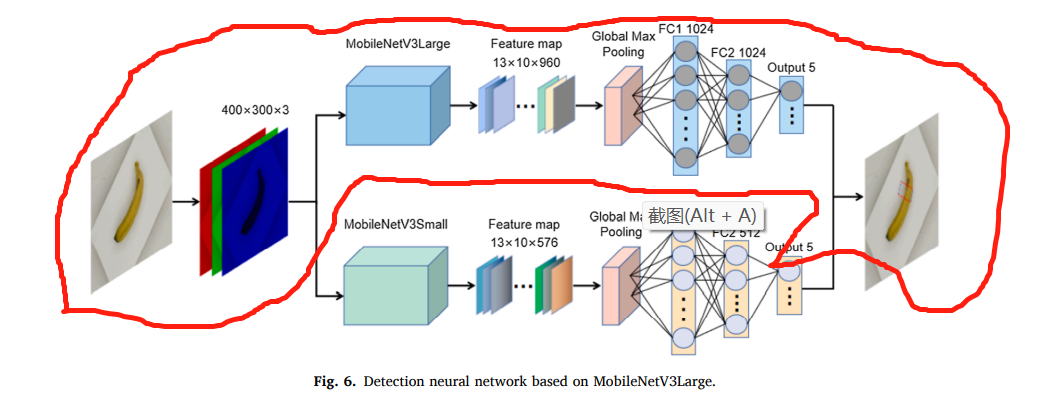
# Improved mobilenet\_v3 identifies cucumber pest and disease leaves in fuzzy scenes



# Introduction

黄瓜，Cucumis sativus，是葫芦科中广泛种植的蔓生藤本植物，通常结圆柱形果实，用作蔬菜。据统计，2019年，世界黄瓜和小黄瓜产量为8800万吨，其中中国占80%。然而，由于黄瓜受到各种疾病的困扰，其全球产量正在下降。

传统的疾病检测方法需要通过视觉线索对患病叶片进行人工检查，由于人为错误，这容易导致检测效率低和可靠性差。此外，这一劳动密集型任务因要探测的大面积区域以及要探测的早期症状的毫米级大小而变得复杂耗时。使问题更加严重的是，农民缺乏专业知识，没有足够能够发现这些疾病的农业专家也阻碍了整体的收成。因此，黄瓜病害的早期检测和分类如果是可以在工具和技术上实现提供给农民，那么可以大大缓解上述所有问题。

目前，在这方面提出的方法很少，包括分子分析、光谱分析、挥发性有机化合物分析等。但是这些方法都很昂贵，并且难以在商业操作规模中应用。2在这方面，计算机视觉具有内在的巨大潜力:作物疾病的症状通常会在植物叶片上引起一种特征，这种特征可以通过基于图像的技术和适当的策略检测出来。通过分析图像中病叶的颜色、纹理和形状来检测和识别作物病害。3

2. Martinelli F, Scalenghe R, Davino S, et al. Advanced methods of plant disease detection. A review[J]. Agronomy for Sustainable Development, 2015, 35(1): 1-25.

3. Benfenati A, Causin P, Oberti R, et al. Unsupervised deep learning techniques for powdery mildew recognition based on multispectral imaging[J]. arXiv preprint arXiv:2112.11242, 2021

但在目前的方法中仍存在许多问题。第一个问题是，现有的方法不能正确识别中国地区的水果叶病害，因为所有当前的实践仅在PlantVillage数据集上训练，该数据集基于来自美国和瑞士农场的图像。由于叶片形状、品种和环境因素的不同，水果的病害也不同于其他地区。此外，由于80%左右的黄瓜产自中国，用于黄瓜叶片检测模型训练的广泛使用的数据集很少。因此，中国农民很难获得一种检测准确率高的黄瓜病害检测技术。我们迫切需要开发一种新的数据集来检测中国地区黄瓜叶片的疾病，以便中国农民可以在早期阶段确定黄瓜的疾病，增加他们的收入，促进国家的经济发展。另一个问题是，广泛用于训练模型的数据集大多是由专业专家和摄影师拍摄的。但在现实中，这些照片大多是由农民拍摄的，他们无法拍摄出完美的照片进行分析，照片可能有各种背景、颜色和大小。因此，有必要在包含非专业叶子图像的数据集中训练模型。最后一个问题是，在农业等实际应用中，大部分中国农民不具有高精度的设备，一般使用移动端设备如手机等比较多。因此我们需要专门为小内存和低计算能力的设备，定制的小而低延迟的模型，同时病理检测结果还具有相当的精度。

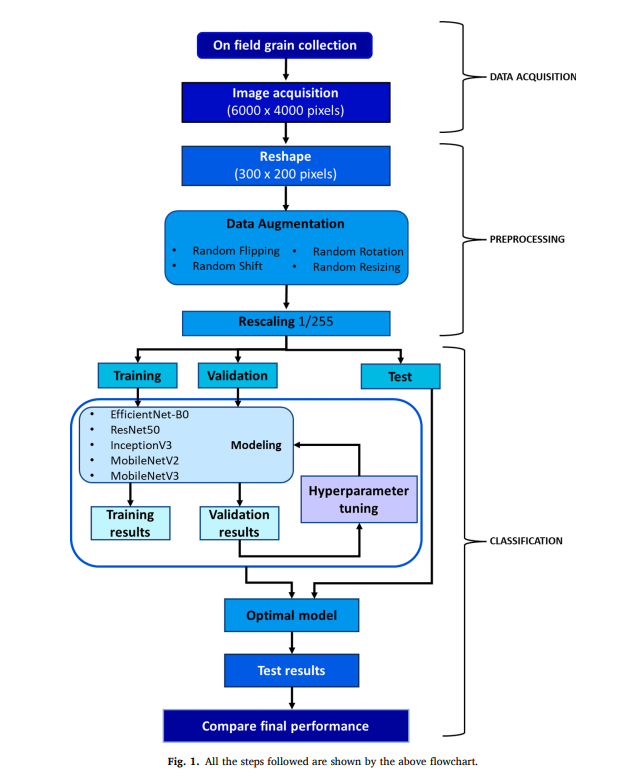
尽管先前的工作在其真实栽培条件图像数据集上实现了高分类精度，但是仍然存在几个问题。首先，基于深度学习的疾病诊断方法需要大量的训练图像。与其他一般的计算机视觉任务不同，标记疾病数据集需要专业的背景知识，而农民很难掌握这些知识。此外，为了收集大型数据集的完美图像，植物必须生长在严格控制的环境中，这是劳动密集型的，并且非常昂贵。其次，过拟合问题在植物诊断任务中特别严重，因为与疾病相关的线索通常不清楚，并且诸如图像背景的其他因素通常对最终决策有显著影响。不仅如此，由于数据集中潜在的相似性而导致的过度拟合通常会导致另一个数据集(例如，来自其他农场的图像)的精度显著下降。例如，在来自广角图像的黄瓜疾病诊断中，在同一农场中的诊断性能在F1分数中显示为86.0%，但是在不同农场中下降到20.7% 21。为了解决过拟合的问题，Saikawa等人22提出了一种从感兴趣区域(RoI)中去除背景的方法作为预处理步骤，结果表明他们可以将精度提高12.2%。但他们也指出，这种方法需要大量昂贵的屏蔽数据，可能会消除对诊断很重要的周围信息。23

22. Saikawa T, Cap Q H, Kagiwada S, et al. AOP: An anti-overfitting pretreatment for practical image-based plant diagnosis[C]//2019 IEEE International Conference on Big Data (Big Data). IEEE, 2019: 5177-5182.

23. Cap Q H, Uga H, Kagiwada S, et al. Leafgan: An effective data augmentation method for practical plant disease diagnosis[J]. IEEE Transactions on Automation Science and Engineering, 2020.

以往研究所采用的数据集大部分都来自公开的数据集PlantVillage，这个数据集的图像标准规范，背景简单清晰，细节拍摄到位。但是实际生活中，大部分农民都采用手机进行拍摄，没有良好的设备拍摄清晰度足够高的照片，并且由于手机型号规格的不同，所得图像的大小和清晰度也不同，需要我们进行进一步处理。不仅如此，样本图像会在不破坏农作物的前提下，在农田直接进行拍摄，所以图像背景大都复杂多变，拍摄角度多样，大大增加了我们病理判断的难度。

简单背景病理判断不适用于实际农业生活，在本文中，我们在一个农业论坛上随机收集黄瓜叶片图像，这意味着这些图像来自中国各地，在植物病理学家的帮助下，这些图像被贴上标签，成为本次实验的数据集。数据集被分成80%用于训练，20%用于测试。如图，我们提出了用一种轻量级的、快速的MobileNetv3网络结构，能够快速准确地完成黄瓜叶部病害的七种分类。不仅如此，本文还通过对优化器、批处理量的选择，还有flooding方法的运用，有效地抑制了过拟合，达到了较高地准确率。



# 二 related work

人工智能(AI)、机器学习(ML)和计算机视觉(CV)技术的最新进展开辟了新的可能性，为通过自动识别相关特征在作物检测中利用光学传感器的数据铺平了道路。通过在农业中采用新的设备、技术和算法，深度学习是智能农业的核心4。深度学习被广泛用于解决复杂的问题，如特征提取、变换、模式分析和图像分类，这有助于大力发展、控制和提高农业生产。

在过去的几十年中，许多类型的深度学习架构被提出用于植物疾病分类，产生了许多针对真实栽培条件的植物疾病诊断系统。

在【6】中，Mohanty等人使用大型CNN(Google net和Alexnet)对14种作物的26种疾病进行了分类，该分类在由PlantVillage知识库【7】中的54，306幅，带标签的患病和健康植物叶片的彩色图像，组成了公共数据并进行了训练。经过训练的模型在保留测试集上达到了99.35%的准确率，证明了智能手机与计算机视觉相结合来辅助植物疾病诊断方法的可行性。

6. Mohanty S P, Hughes D P, Salathé M. Using deep learning for image-based plant disease detection[J]. Frontiers in plant science, 2016, 7: 1419.

7. Hughes DP, Salathé M (2015) An open access repository of images on plant health to enable the development of mobile disease diagnostics. CoRR abs/1511.08060

在【8】中Sladojevic等人使用深度学习框架CaffeNet，提出了建立植物病害识别模型的新方法。开发的模型能够从健康叶片中识别出13种不同类型的植物疾病，并能够将植物叶片与其周围环境区分开来。用从网上下载的4483幅(增加到30，880幅)图像进行模型的训练，PlantVillage数据集用于评估所提出技术的性能。在开发的模型上的实验结果实现了91%和98%之间的精度，对于单独的类测试，平均为96.3%。

8.Sladojevic S, Arsenovic M, Anderla A, et al. Deep neural networks based recognition of plant diseases by leaf image classification[J]. Computational intelligence and neuroscience, 2016, 2016.

在【9】中，Karthik等人提出了两级深度学习番茄叶部病害检测技术。第一层架构应用残差学习来学习[分类](https://www.sciencedirect.com/topics/computer-science/classification" \o "Learn more about classification from ScienceDirect's AI-generated Topic Pages)的重要特征。第二层架构在[残差深度网络](https://www.sciencedirect.com/topics/computer-science/deep-residual-network" \o "Learn more about residual deep network from ScienceDirect's AI-generated Topic Pages)之上应用注意力机制。使用Plant Village Dataset数据集进行实验，该数据集包含三种病害，即早疫病、晚疫病和叶霉病。作者利用了CNN使用注意力机制在各种处理层次结构中学习的特征，并在五重交叉验证中对验证集的总体准确率达到了98%。

9. Karthik R, Hariharan M, Anand S, et al. Attention embedded residual CNN for disease detection in tomato leaves[J]. Applied Soft Computing, 2020, 86: 105933.

在【10】中张等为了提高作物病害叶片识别模型的准确性和病害叶片的定位，提出了一种改进的快速RCNN来检测健康番茄叶片和四种病害。首先，作者使用深度残差网络代替VGG16进行图像特征提取，以获得更深层次的疾病特征。其次，使用k-means聚类算法对包围盒进行聚类，再根据聚类结果改进锚定。改进的锚定框架趋向于数据集的真实边界框。最后，作者用三种不同的特征提取网络进行了k-means实验。实验结果表明，改进后的作物叶部病害检测方法比原来的快速RCNN识别准确率提高了2.71%，检测速度加快。

10. Zhang Y, Song C, Zhang D. Deep learning-based object detection improvement for tomato disease[J]. IEEE Access, 2020, 8: 56607-56614.

Patrick et al .在12中，作者提出了一种新的计算机视觉系统，可以自动识别几种疾病，检测以前未发现的疾病，并估计每片叶子的严重程度。模型的训练和测试使用了PlantVillage番茄数据集的九种番茄疾病的几个修改版本，并显示了不同的叶子属性如何影响疾病检测。

12. Wspanialy P, Moussa M. A detection and severity estimation system for generic diseases of tomato greenhouse plants[J]. Computers and Electronics in Agriculture, 2020, 178: 105701.

Kawasaki等人17训练了一个三层卷积神经网络，该网络可以自动获取分类所需的特征，并获得高分类性能，以在真实农场的图像上诊断三类黄瓜疾病，其中目标对象具有复杂的背景。在4重交叉验证策略下，作者的模型达到了94.9%的平均准确率。

17. Kawasaki Y, Uga H, Kagiwada S, et al. Basic study of automated diagnosis of viral plant diseases using convolutional neural networks[C]//International symposium on visual computing. Springer, Cham, 2015: 638-645.

DeChant等人18提出了一种自动系统，该系统是由几层卷积神经网络（CNN）构成，用于在从玉米植株田地获取的图像上识别大斑叶枯病病变，并在测试集上实现96.7%的准确度。

18. DeChant C, Wiesner-Hanks T, Chen S, et al. Automated identification of northern leaf blight-infected maize plants from field imagery using deep learning[J]. Phytopathology, 2017, 107(11): 1426-1432.

以上这些研究通过各种卷积神经网络，都获得了很高的判断准确度，但都是基于背景清晰的规范图像，一旦在背景模糊的前提下，其精确度会大大降低，无法满足需求。

在[1]中，叶中华等人就研究了在真实农业生产环境中，通过不同模型的对比改进，最终采用SSD目标检测模型，实现对复杂背景农作物图像病害区域的预测，实验结果表明，最终模型在测试集的平均精度均值达到83. 90%。

[1]叶中华,赵明霞,贾璐.复杂背景农作物病害图像识别研究[J].农业机械学报,2021,52(S1):118-124+147.

在【2】中，研究自然水下环境下基于机器视觉的海参自动分割。海参通常生活在光照、能见度和稳定性普遍不可控的环境中，这导致海参的水下图像失真、模糊、衰减严重。通过建立基于边缘信息的矩形轮廓作为身体分割的初始轮廓，该方法可用于海参实时监测和捕捞。

Xi Qiao, Jianhua Bao, Lihua Zeng, Jian Zou, Daoliang Li,An automatic active contour method for sea cucumber segmentation in natural underwater environments,Computers and Electronics in Agriculture,Volume 135,2017,Pages 134-142,ISSN 0168-1699,

【3】为了解决复杂背景图像和图像目标的分割误差，考虑到模糊集在图像处理上的优势，提出了最大模糊相关准则。将图像转换到模糊域，然后计算最大模糊相关准则。最后，对树叶和水果农业图像进行实验，结果表明模糊集能够检测出更多的边缘细节。

Ronghua Gao, Huarui Wu,Agricultural image target segmentation based on fuzzy set,Optik,Volume 126, Issue 24,2015,Pages 5320-5324,ISSN 0030-4026,

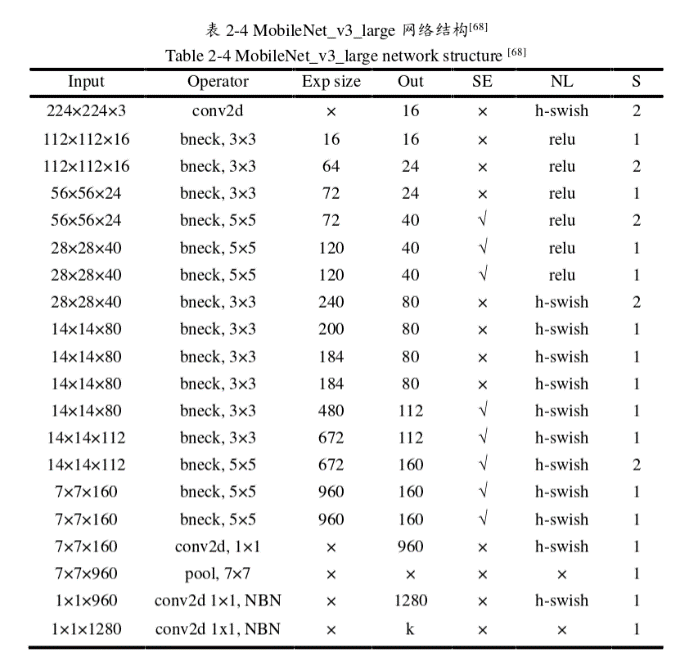
# 三 method

## 3.1 MobileNet\_v3

MobileNetV3 is also a lightweight network. MobileNetV3 uses a network architecture search (NAS) to search the global network structure by optimizing each network block, supplemented by the NetAdapt algorithm. This technique can efficiently determine an optimal model for a given hardware platform. In addition, MobileNetV3 uses the hswish activation function to improve accuracy (Howard et al., 2019 ).

Howard, A., Sandler, M., Chu, G., Chen, L.C., Chen, B., Tan, M., Adam, H., 2019. Searching for mobilenetv3. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 1314–1324.

与其他分类模型相比，它在输入图像的每个深度上操作单个卷积，而不是组合和展平输入的所有深度，通过深度方向可分离卷积实现这一点。这种深度卷积将卷积过程分为两层，一层用于滤波，一层用于合并。这种组合减小了模型的大小。MobileNetv3由4个2D卷积层、2个(112x122)瓶颈层、2个(56×56)瓶颈层、3个(28×28)瓶颈层、7个(14×14)瓶颈层、2个(7×7)瓶颈层组成，其中Swish和Relu用作激活。在两个密集层之前使用一个汇集层(7x7)。挤压和激励层也包括在内，使其速度更快，重量更轻。在创建输出要素地图时，这种添加会为通道分配不相等的权重。最后，应用具有1024个单元的密集层来获得特征向量。下图为MobileNet\_v3\_large的网络结构图，表中Input为输入图像尺寸，Operator为卷积层或倒残差结构，Exp size、Out分别为倒残差结构第一层与最后一层卷积核的个数，SE为是否采用SE模块，NL为倒残差结构第一层与第二层采用的激活函数，S为倒残差结构深度卷积层的步长。



## 3.2 随机失活法（Dropout）

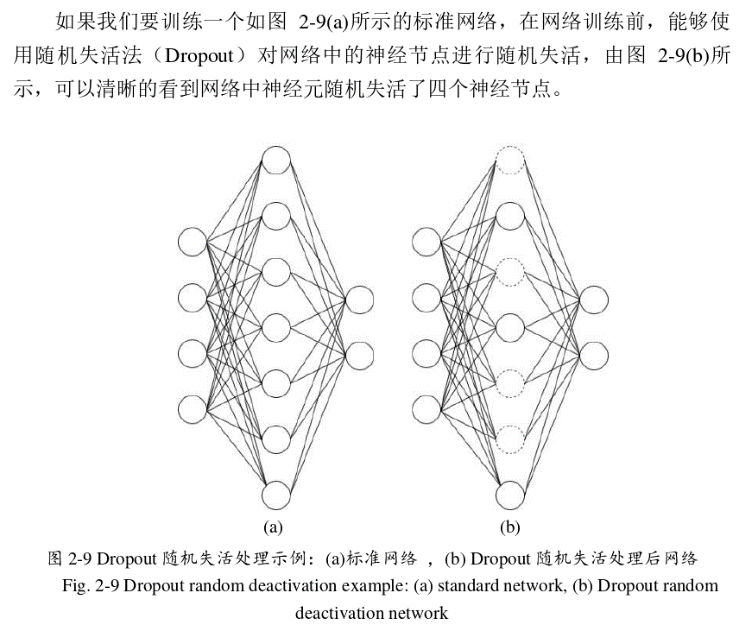
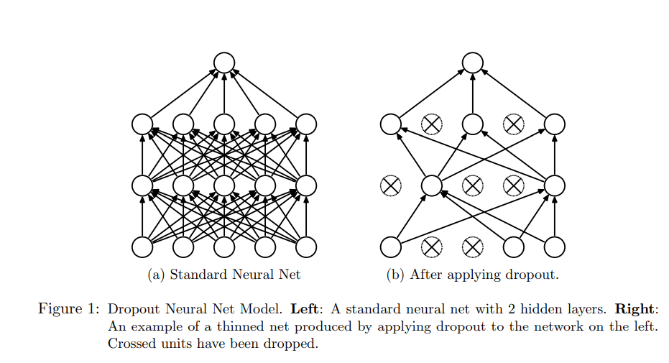
当样本数目不够大时，容易造成模型泛化能力不好，出现这种情况被称为模型过拟合。在2012年，Hinton 等人[1]设计出了随机失活法（Dropout）算法，此方法能够有效的抑制模型的过拟合，提高泛化能力。

随机失活法指的是关闭神经网络中的单元(隐藏的和可见的)。随着网络继续训练，在数据中具有更好表示的特征主导学习过程，并且与这些特征相关联的权重变得更强。随机失活法提升卷积神经网络的性能是通过抑制每层滤波器而实现，也就是网络模型在对样本图像进行特征提取的过程中，随机除去图像部分特征信息以此提升CNN模型的泛化能力。

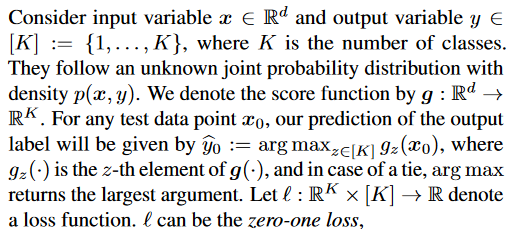
[1] Hinton G E, Srivastava N, Krizhevsky A, et al. Improving neural networks by preventing co-adaptation of feature detectors[J]. arXiv preprint arXiv:1207.0580, 2012.

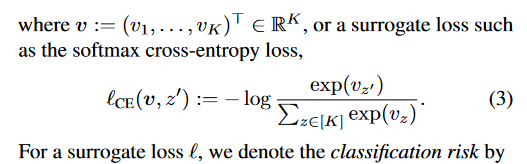
在Dropout层，不同的神经元组合被关闭，这代表了一种不同的结构，所有这些不同的结构使用一个的子数据集并行地带权重训练，而权重总和为1。如果Dropout层有 n 个神经元，那么会形成 2n 个不同的子结构。在预测时，相当于集成这些模型并取均值。这种结构化的模型正则化技术有利于避免过拟合。Dropout有效的另外一个视点是：由于神经元是随机选择的，所以可以减少神经元之间的相互依赖，从而确保提取出相互独立的重要特征。

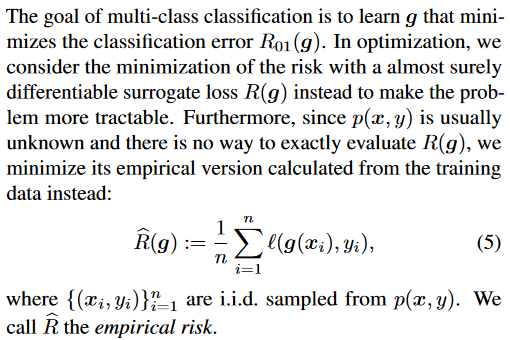
例如我们要训练一个如图所示的标准网络，在网络训练前，能够使用随机失活法（Dropout）对网络中的神经节点进行随机失活，变成土（b），可以清晰的看到网络中的神经元随机失活了4个神经节点。

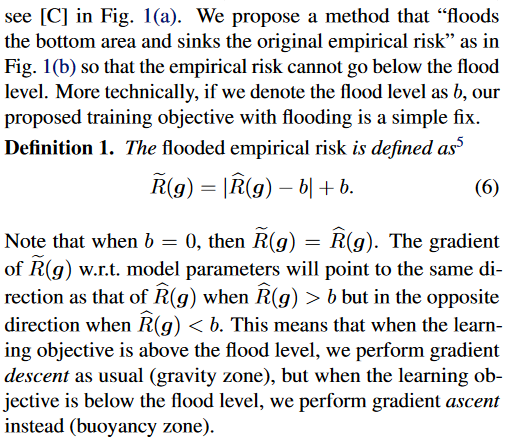
 

## 3.3 Flooding









推动参数往更平稳的区域走，让损失函数收敛保持在一个阈值附近，这提高了泛化性能，更好地抵抗扰动。本文采取穷举法来搜索最佳阈值（？）

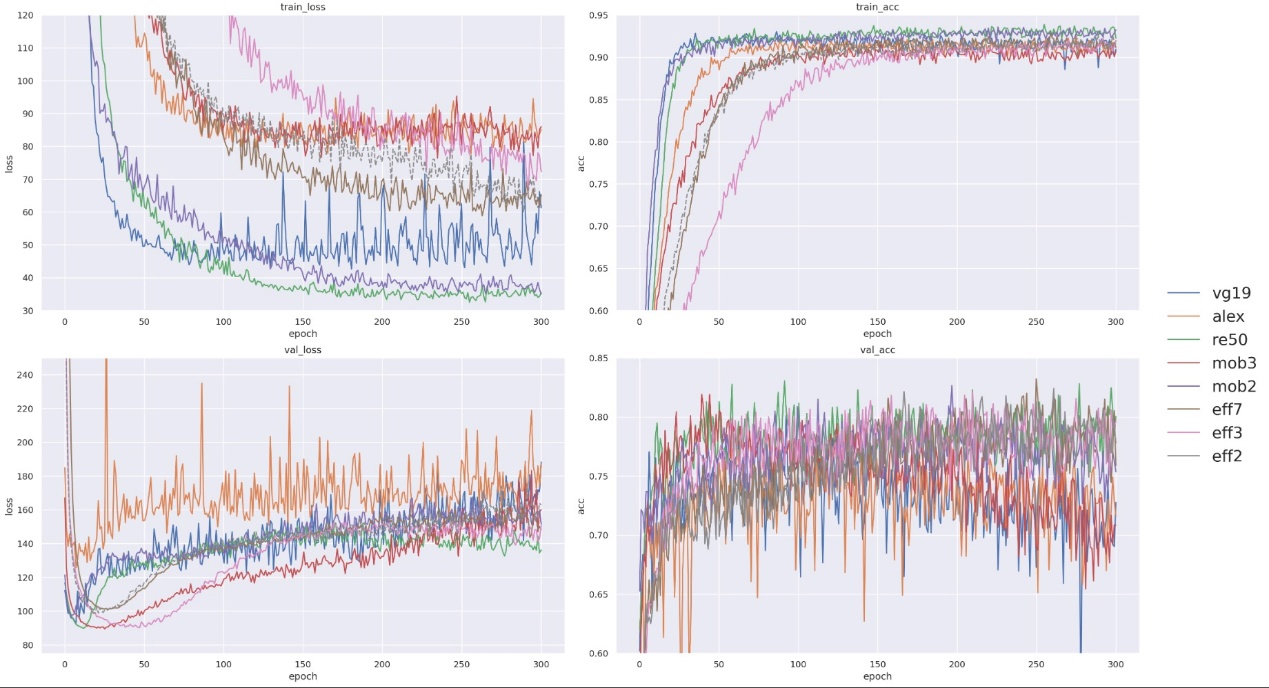
# 四 experiment

## 4.1 对比试验

实验一：八个算法对比，mobile v3. + ASGD ，loss，acc

在相同的黄瓜病理树叶图像数据集上，我们选择了7种当今主流的网络模型与MobileNet\_v3网络进行对比试验，从实验数据精确反映出MobileNet\_v3网络的优越性。

本文训练了alexnet、resnet、VGG、efficient\_v2、efficient\_v3、efficient\_v7、mobilenet\_v2、mobilenet\_v3八种树叶病理识别模型，使用图像训练集和测试集对其进行测试对比，以此检验网络模型性能的优越性，并对此进行进一步的优化。本次对比实验优化器选择ASGD，学习率设为0.001，L1正则系数为0.01，L2正则系数0.01，训练集进行300轮迭代训练，对测试集进行300轮的训练



Alexnet模型的训练集的损失函数和准确度都收敛效果很好，在epoch=60时开始收敛，逐渐趋于稳定。但是在测试集上的效果不佳，loss与精确度的波动都很大，收敛效果也不好。

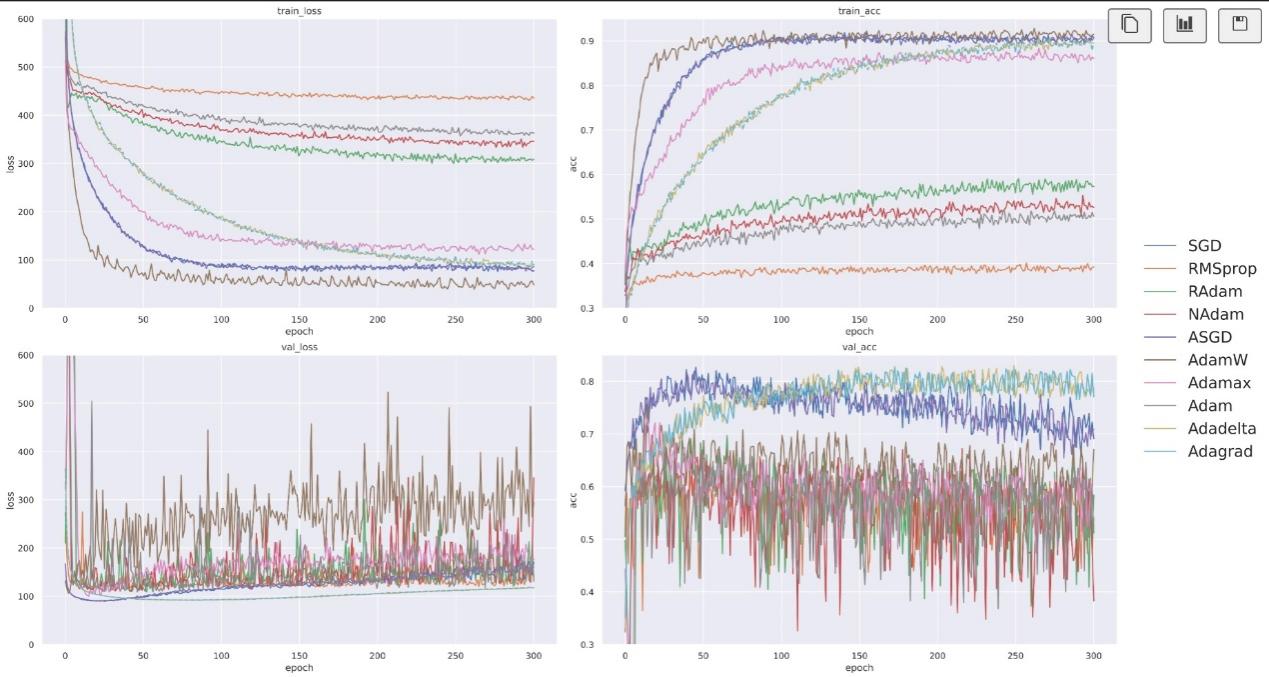
Resnet的训练集在训练轮数40左右时开始收敛，收敛速度很快，但测试集的loss和精确度的波动幅度从数值上看比alexnet网络模型的更大，损失函数还出现了过拟合现象。两种模型的拟合效果均不佳。

VGG模型的训练集收敛速度很快，在第20轮迭代时开始收敛；efficient\_v2模型稍慢，在第70迭代伦茨开始收敛。但相比于上两个模型VGG和efficient\_v2的测试集的波动更大，VGG模型的测试集loss和准确度图像基本没收敛趋势，最终效果不好。Efficent\_v2的测试集呈现了收敛效果，但是出现了过拟合现象。

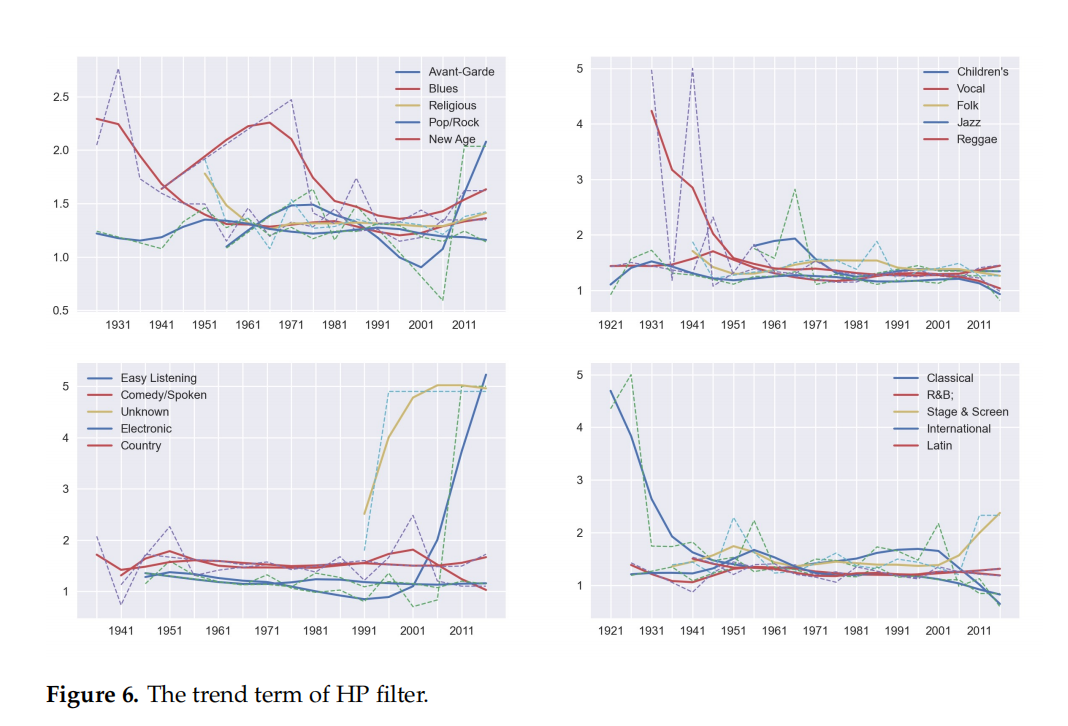
Efficient系列模型的测试集的loss和准确度都趋于收敛，与训练集相比你和程度也较高，但其测试集的loss图像过拟合程度过高，且波动较大。MobileNet\_v2模型的测试集loss和准确率图像波动很大，基本没有收敛趋势，是明显不适用的网络模型。MobileNet\_v2模型，训练集从第50迭代开始收敛，loss最终不断趋近于0，准确率也随着训练轮次不断升高。测试集的结果趋势也与训练集的大致拟合 ，虽然存在过拟合现象，但其从数值上来说，loss的标准差值最小。

综合考虑各个模型测试集的拟合效果，mobilenet v3这个网络的测试集结果趋势大体与训练集趋势一致拟合，且其精确度的最大值是相对最高的，达到了81.3%及以上。不仅如此，mobilenet v3因其轻量级框架的优势，其收敛速度比其它网络快66%左右。因此，我们最终选择mobilenet v网络进行下一步的优化实验

实验二：优化器adam,ASGD(六-八个优化器)...->ASGD



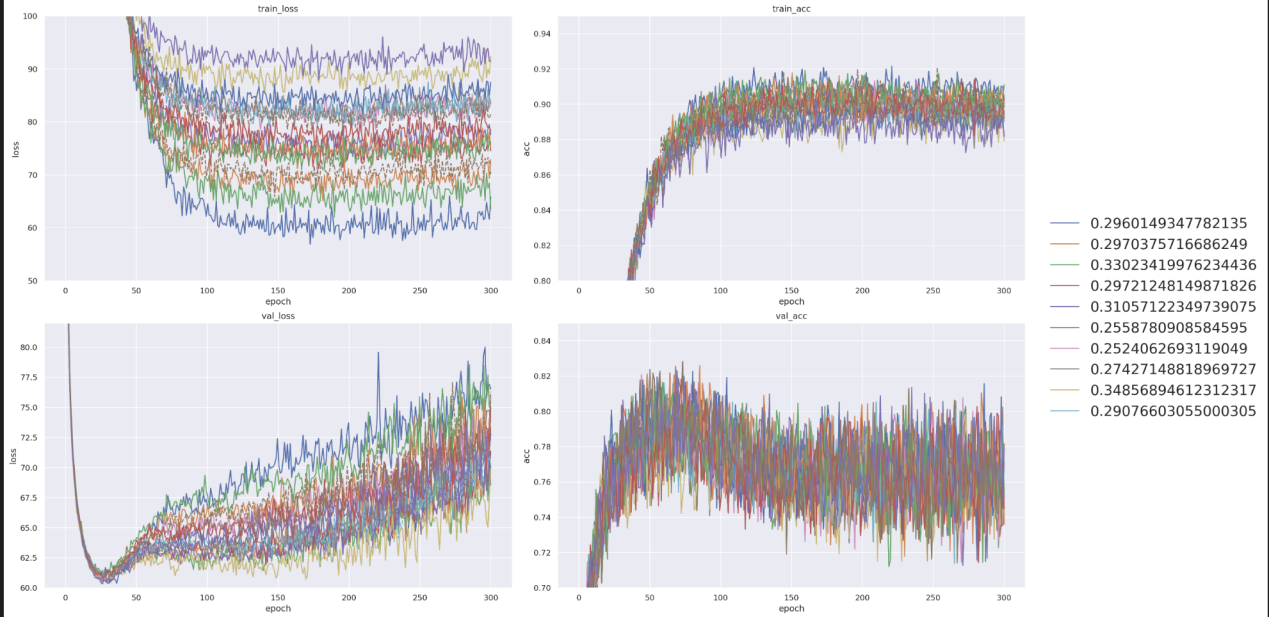
实验三：batch size:三个不同的batch\_size：8 16 32 64



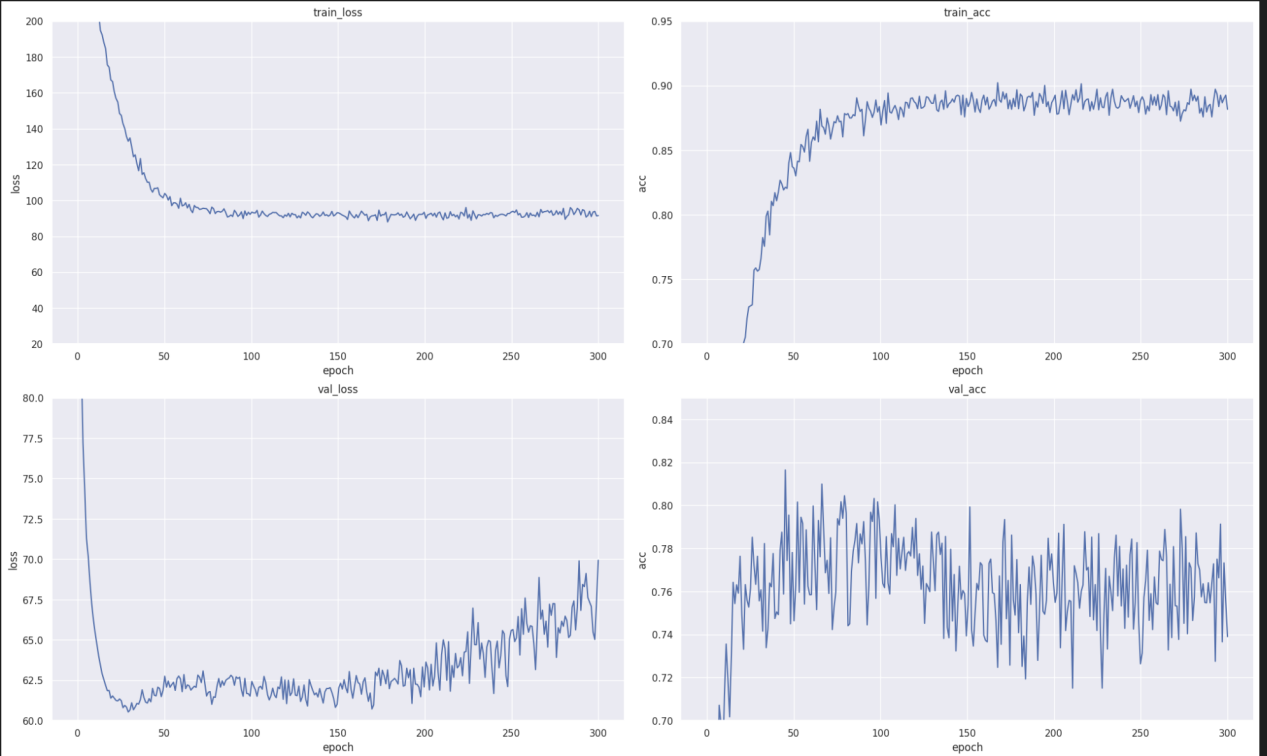
本次实验的优化器采用ASGD，学习率为0.001，L1和L2的正则化系数均为0.01，进行100次迭代实验。批处理量分别取6，16，32进行图像和数据的对比分析。

根据mobilenet\_v3网络的图像和数据结果分析，我们可以看出，mobilenet\_v3网络模型的实验结果非常优秀，在loss的波动减小的同时，测试集的精确度上升。当批处理量batch size=16时，loss的标准差最小，抗锯齿化效果最好。同时，测试集的精确度也上升到了83.2%。

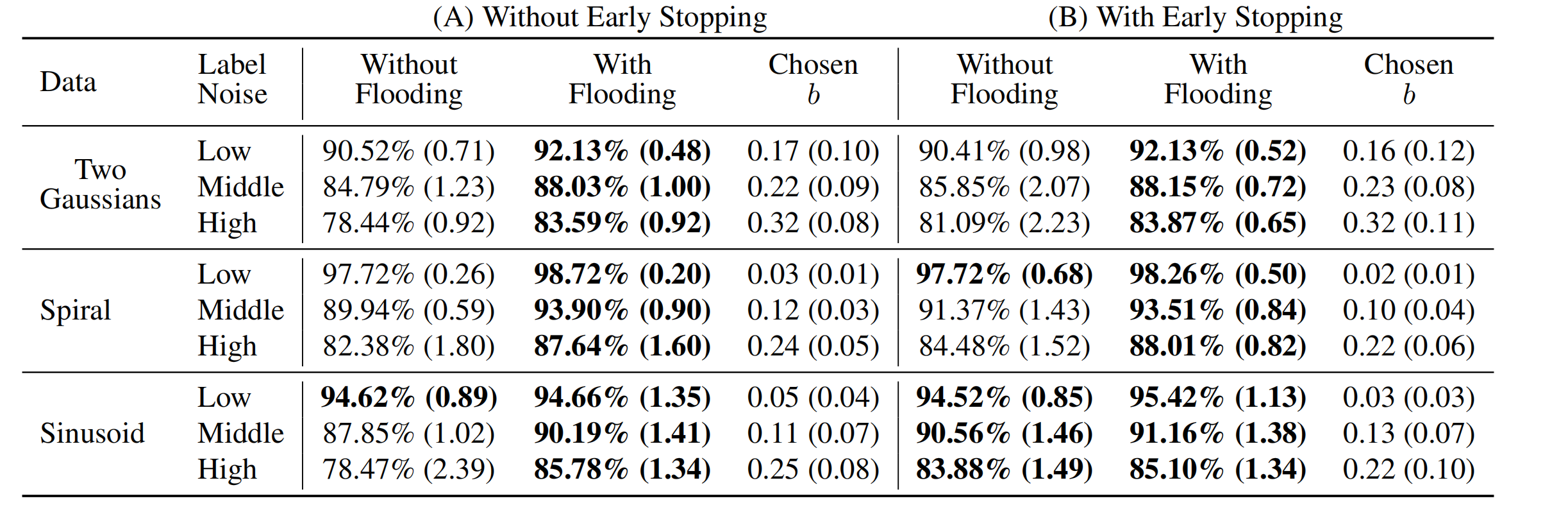
实验四：加了flooding：15个b



可以看到加flooding之后对val\_loss上升有明显的抑制作用



加了flooding后，val\_acc平均值能提高0.2%+，val\_acc最大值能提高0.5%。



在第三章我们简单介绍了flooding的基本原理和使用方式，即通过改变损失函数，加一个阈值，使loss最终在阈值附近波动。上述mobilenet\_v3实验的loss结果图像均出现了过拟合现象，本节将通过flooding实现数据集loss的二次下降，抗过拟合。

本次实验的优化器采用ASGD，学习率为0.001，L1和L2的正则化系数均为0.01，进行100次迭代实验。Loss的阈值分别设为0.28580141067504883、0.29054024815559387、0.33533796668052673进行图像和数据的对比分析。