Flooding based mobilenet v3 identiﬁes cucumber disease leaves in1 fuzzy scenes2

Liu Yiminga, Wang Zhengleb, Wang Rujiac and \*Gao Hongjub a

School of Computer Science(National Pilot Software Engineering School),Beijing University of Posts and4

Telecommunications,100876,China;liuyimingbyr@bupt.edu.cn5 b

College of Information and Electrical Engineering, China Agricultural University, Beijing 100083,6

China;wangzhengle@cau.edu.cn;hjgao@cau.edu.cn7 c

School of Information and Communication Engineering,School of Computer Science,Beijing University of Posts and8

Telecommunications,100876,China;WNHwrj@bupt.edu.cn9 A R T I C L E I N F O

Keywords:

Flooding;Lightweight;Mobilenet v3;Cucumber disease;Crop question and answer

11 A B S T R A C T2

Domestic cucumber production is declining due to various pathologic diseases, but the technol-14

ogy of plant pathologic detection is not mature and requires high labor costs. In addition, since15

the planting site is usually a high-density scene, most photos taken are shot from various angles,16

and the background is messy, resulting in poor detection reliability. The rise of online question-17

and-answer systems is an inspiration. This paper wants to establish an online QA system. Farm-18

ers can upload cucumber pictures by taking photos, and the system can quickly identify and judge19

with high accuracy. In this paper, the crawler program is used to collect many cucumber leaf20

image data in batches on an agricultural website, and simple preprocessing is carried out. With21

a lightweight and fast MobileNetv3 network structure, it can quickly and accurately complete22

the seven kinds of cucumber leaf disease classiﬁcation. The optimal network model is achieved23

by selecting appropriate parameters, optimizer, and batch capacity through the single variable24

method. In addition, a new training strategy of data set loss -ﬂooding method was introduced25

in this paper, replacing the strategy of ﬂooding after the ﬂooding threshold was reached, which26

ﬁnally achieved 88.3% accuracy. Finally, two public data sets of PlantVillage and apple disease27

were selected for another experiment. The accuracy was up to 99% and 98.1%, respectively,28

which proved the universality of the strategy proposed in this paper. In this paper, the code29

will be open source in https://github.com/YiQuanMarx/Agricultural\_Diseases\_Dentiﬁcation for30 reference.31

CRediT authorship contribution statement33

Liu Yiming: Responsible for paper experiment conception, data processing, main experiment realization, data34

processing, picture drawing and paper writing and polishing. Wang Zhengle: Participate in the preliminary research35

of the paper, responsible for the partial realization of the paper experiment and the preparation of the paper. Wang36

Rujia: Responsible for data processing, drawing and editing of the paper.\*Gao Hongju: Responsible for framing37

the paper, providing data, guiding the writing of the paper, and polishing the paper.38 1. Introduction39

The cucumber, Cucumis sativus, is a widely cultivated creeping vine in the gourd family that usually bears cylindri-40

cal fruits and is used as a vegetable. According to statistics, in 2019, the world produced 88 million tons of cucumbers41

and gherkins, of which China accounted for 80 percent. However, global production of cucumbers is declining as42 various diseases plague them.43

Traditional disease detection methods require manual inspection of diseased leaves through visual cues, which is44

easy to lead to low detection eciency and poor reliability due to human error. In addition, this labor-intensive task is45

ORCID(s):0000-0002-3006-5919(L. Yiming)

short author name: Preprint submitted to Elsevier Page 1 of 14

complicated and time-consuming by the large area to be detected and the millimeter-scale size of the early symptoms46

to be detected. Compounding the problem is a lack of expertise among farmers, and not enough agricultural experts47

able to spot these diseases also hinder overall harvests. Therefore, if available to farmers tools and techniques exist,48

early detection and classiﬁcation of cucumber diseases can signiﬁcantly alleviate these problems. The emergence of49

an online question-and-answer system provides a suitable treatment method. We can allow farmers to upload photos50

of their cucumber leaves through mobile devices such as mobile phones. After receiving the images, the system will51

process and analyze the pathological results of the cucumbers. The online question-and-answer system has been used52

in all aspects of life, including various shopping software, customer service, online hospital consultation, etc. However,53

the current online question-and-answer system is mainly based on manual identiﬁcation processing, and agricultural54

manual inspection is unsuitable for this situation. Therefore, this paper needs to seek a kind of online question-and-55 answer system without a manual.56

Currently, there are few methods for the pathological analysis of cucumber, including molecular analysis, spectral57

analysis, volatile organic compound analysis, etc. However, these methods are expensive and dicult to apply on a58

commercial scale. Martinelli et al.(2015) In this respect, computer vision has great inherent potential: symptoms of59

crop diseases often cause a feature on plant leaves that can be detected with image-based techniques and appropriate60

strategies. Crop diseases are detected and identiﬁed by analyzing the images’ color, texture, and shape of diseased61 leaves. Benfenati et al.(2021)62

However, many things could still be improved with the current approach. The ﬁrst problem is that existing methods63

need to correctly identify fruit leaf diseases in the Chinese region because all current practices are trained only on the64

PlantVillage dataset, which is based on images from farms in the United States and Switzerland. Fruit diseases also65

dier from other regions due to dierences in leaf shape, variety, and environmental factors. In addition, as about 80%66

of cucumbers are produced in China, there are few widely used data sets for training cucumber leaf detection models.67

Therefore, it is dicult for Chinese farmers to obtain cucumber disease detection technology with high accuracy. We68

urgently need to develop a new data set to detect diseases in cucumber leaves in regions of China so that Chinese farmers69

can identify diseases in cucumbers early, increase their income and boost the country’s economic development.70

Another problem is that professional experts and photographers mostly take the data sets widely used in training71

models. However, most of these photos are taken by farmers who cannot get the perfect shot for analysis, which can72

come in various backgrounds, colors, and sizes. Therefore, it is necessary to train the model in a data set containing73

non-specialized leaf images. The last problem is that most Chinese farmers need high-precision equipment for practical74

applications such as agriculture and generally use mobile terminal devices such as mobile phones. Therefore, we need75

small, low-latency models explicitly tailored for devices with small memory and low computing power. At the same76

time, the results of pathological tests are accurate.77

short author name: Preprint submitted to Elsevier Page 2 of 14

Although previous work achieved high classiﬁcation accuracy on its data set of images of natural cultivation con-78

ditions, several problems still need to be solved. First, deep learning-based disease diagnosis methods require many79

training images. Unlike other general computer vision tasks, labeling disease data sets requires specialized background80

knowledge dicult for farmers to master. Also, to collect perfect images of large data sets, plants must be grown in81

tightly controlled environments, which is labor-intensive and very expensive. Second, overﬁtting problems are partic-82

ularly acute in plant diagnostic tasks because clues related to the disease are often unclear, and other factors, such as83

the image’s background, often signiﬁcantly impact the ﬁnal decision. Not only that, but overﬁtting due to potential84

similarities in the data set often results in a signiﬁcant decline in the accuracy of another data set (for example, images85

from other farms). For example, in cucumber disease diagnosis from wide-angle images, diagnostic performance on86

the same farm showed 86.0% in F1 scores but decreased to 20.7% in dierent farms. To solve the overﬁtting problem,87

Saikawa et al. Saikawa et al.(2019) proposed a method to remove background from the region of interest (RoI) as a88

pre-processing step. The results showed that they could improve accuracy by 12.2%. However, they also point out that89

this approach requires much more expensive masking data, potentially eliminating surrounding information essential90

to a diagnosis. Cap et al.(2020)91

This paper is expected to use a lightweight and fast mobile ETV3 network structure to make it suitable for mobile92

terminal device recognition processing. The optimal network model is achieved by selecting appropriate parameters,93

optimizer, and batch capacity through the single variable method. To further improve its accuracy. With the epoch94

increase, if the loss on the test set reaches a certain threshold and continues training to reduce loss, overﬁtting will95

occur. This paper considers introducing a new training strategy for data set loss to replace the strategy to reduce96

loss after reaching the threshold to solve the overﬁtting problem on the test set. The ﬂooding method is expected to97

improve the situation. Finally, it can achieve high accuracy and better apply to the daily agricultural life of cucumber98 pathological judgment.99

2. Related work100

Recent advances in artiﬁcial intelligence (AI), machine learning (ML), and computer vision (CV) technologies101

have opened up new possibilities, paving the way for the use of data from optical sensors in crop detection by automat-102

ically identifying relevant features. Deep learning is at the heart of intelligent farming through adopting new devices,103

technologies, and algorithms in agriculture 4. Deep learning is widely used to solve complex problems, such as feature104

extraction, transformation, pattern analysis, and image classiﬁcation, which helps signiﬁcantly develop, control, and105

improve agricultural production.106

Over the past few decades, many types of deep learning architectures have been proposed for plant disease classi-107

ﬁcation, resulting in several plant disease diagnosis systems tailored to real cultivation conditions.108

short author name: Preprint submitted to Elsevier Page 3 of 14

In Prasanna Mohanty et al.(2016), Mohanty et al., using a large CNN(Google net and Alexnet), classify 26 diseases109

in 14 crops in 54, PlantVillage Repository Hughes et al.(2015),306 labeled color images of diseased and healthy plant110

leaves formed public data and were trained. The trained model achieved 99.35 percent accuracy on the retention test111

set, demonstrating the feasibility of combining smartphones with computer vision to aid in plant disease diagnosis112 methods.113

In Sladojevic et al.(2016), Sladojevic et al. used the deep learning framework CaeNet to propose a new method114

to establish a plant disease recognition model. The developed model was able to identify 13 dierent types of plant115

diseases from healthy leaves and was able to distinguish plant leaves from their surroundings. The model was trained116

with 4483(increased to 30,880) images downloaded from the Internet, and the PlantVillage dataset was used to evaluate117

the performance of the proposed technique. Experimental results on the developed model achieved an accuracy of118

between 91% and 98%, with an average of 96.3% for individual class tests.119

In Karthik et al.(2020), Karthik et al. proposed a two-stage deep-learning technique for tomato leaf disease de-120

tection. The ﬁrst architecture applies residual learning to learn essential features of classiﬁcation. The second layer121

architecture applies the attention mechanism to the deep residual network. The experiment was conducted using the122

Plant Village Dataset, which contained three diseases: early blight, late blight, and leaf mold. The author takes advan-123

tage of the features CNN uses attention mechanism to learn in various processing hierarchies, and the overall accuracy124

of the veriﬁcation set reaches 98% in ﬁve-fold cross-validation.125

In Zhang et al.(2020), Zhang et al. proposed an improved fast RCNN to detect healthy tomato leaves and four126

diseases to improve the accuracy of the crop disease leaf recognition model and location of disease leaves. First, the127

author used a deep residual network instead of VGG16 for image feature extraction to obtain deeper disease features.128

Secondly, a k-means clustering algorithm is used to cluster the bounding box, and then anchoring is improved according129

to the clustering results. The improved anchoring framework is the genuine bounding box of the data set. Finally, the130

author conducts a k-means experiment with three feature extraction networks. The experimental results show that the131

improved method is 2.71% more accurate than the original fast RCNN, and the detection speed is faster.132

Patrick et al. In Wspanialy and Moussa (2020), the authors propose a new computer vision system that can auto-133

matically identify several diseases, detect previously undetected diseases, and estimate the severity of each leaf. The134

model was trained and tested using several modiﬁed versions of nine tomato diseases from the PlantVillage tomato135

dataset and showed how dierent leaf attributes aect disease detection.136

Kawasaki et al. Kawasaki et al.(2015) trained a three-layer convolutional neural network, which can automat-137

ically acquire features required for classiﬁcation and obtain high classiﬁcation performance to diagnose three types138

of cucumber diseases on real farm images where the target object has a complex background. Under the four-fold139

cross-validation strategy, the average accuracy of the model achieved 94.9%.140

short author name: Preprint submitted to Elsevier Page 4 of 14

DeChant et al. DeChant et al.(2017) proposed an automatic system consisting of several layers of convolutional141

neural networks (CNN) for identifying large spot blight lesions on images obtained from maize plant ﬁelds and achiev-142

ing an accuracy of 96.7% on the test set.143

The above studies obtained high judgment accuracy through various convolutional neural networks, but they were144

all based on standardized images with transparent backgrounds. Once the background was blurred, the accuracy would145

be signiﬁcantly reduced, which could not meet the requirements.146

In Zhonghua et al.(2021), Ye Zhonghua et al. studied the real agricultural production environment and ﬁnally147

adopted the SSD target detection model through the comparison and improvement of dierent models to realize the148

prediction of crop image disease regions with complex backgrounds. The experimental results showed that the average149

accuracy of the ﬁnal model in the test set reached 69.894%.150 3. Dataset and method151

3.1. Dataset152

Most of the data sets used in previous studies are from the public data set PlantVillage, which has standard image153

speciﬁcations, simple and clear background, and accurate shooting details. However, the simple background patho-154

logical judgment does not apply to agricultural life.155

In this paper, we used the crawler program written to collect a large number of cucumber leaf image data in batches156

on an agricultural website, which means that these images come from all over China, and most of these images are157

randomly taken by farmers with mobile terminal devices. In real life, most farmers use mobile phones to shoot, so158

there is no suitable equipment to shoot photos with high enough resolution. Moreover, due to the dierent models and159

speciﬁcations of mobile phones, the size and resolution of the images are also dierent, which requires us to process160

them further. Moreover, sample images will be shot directly on farmland without destroying crops, so the background161

of images is complex and changeable, and the shooting angles are diverse, as shown in Figure1.162

With the help of plant pathologists, these images were labeled and became the data set for the experiment. The163

data set consisted of 2392 images, of which 80% were used for the training set and 598 images, or 20%, were used for164

the test set. As shown in the ﬁgure, we propose a lightweight and fast MobileNetv3 network structure that can quickly165

and accurately complete the classiﬁcation of seven kinds of cucumber leaf diseases. The seven pathologic conditions166

are downy mildew, powdery mildew, bacterial angular leaf spot, target leaf spot, gummy stem blight, fusarium wilt,167

and anthracnose. Therefore, the machine vision system proposed in this paper for cucumber pathological diagnosis168

consists of three steps: image acquisition, preprocessing and classiﬁcation, and network model optimization.shown in Figure2.170

short author name: Preprint submitted to Elsevier Page 5 of 14

short author name: Preprint submitted to Elsevier Page 6 of 14 3.2. MobileNet v3171

MobileNetV3 is also a lightweight network. MobileNetV3 uses a network architecture search (NAS) to search the172

global network structure by optimizing each network block, supplemented by the NetAdapt algorithm. This technique173

can eciently determine an optimal model for a given hardware platform. In addition, MobileNetV3 uses the h-swish174

activation function to improve accuracy Howard et al.(2019)175

In contrast to other classiﬁcation models, it operates a single convolution at each depth of the input image rather176

than combining and ﬂattening all the depths of the input, which is achieved by depth-oriented separable convolution.177

This deep convolution divides the convolution process into two layers, one for ﬁltering and one for merging. This178

combination reduces the size of the model. MobileNetv3 consists of 42D convolution layers,2(112x122) bottleneck179

layers,2(56◊56) bottleneck layers,3(28◊28) bottleneck layers,7(14◊14) bottleneck layers, and 2(7◊7) bottleneck180

layers, in which Swish and Relu are used for activation. Use a pooling layer (7x7) before two dense layers. Extrusion181

and excitation layers are also included, making it faster and lighter. This addition assigns unequal weights to channels182

when creating a map of output elements. Finally, a dense layer with 1024 units is applied to obtain the feature vector.183

The following Table1 is the network structure diagram of MobileNet v3 large. Input in the table is the size of the input184

image. The operator is the convolution layer or the reciprocal residual structure, Exp size and Out are the numbers of185

convolution kernels of the ﬁrst and last layer of the reciprocal residual structure, respectively, and SE is whether the186

SE module is used. NL is the activation function used in the ﬁrst and second layers of the reciprocal residual structure,187

and S is the step size of the deep convolution layer of the reciprocal residual structure.188 3.3. Flooding189

In this paper, the superiority of the network model is judged mainly by the loss size. Firstly, the generation modeof the loss function is introduced. This paper’s experiment’s loss function adopts the cross entropy loss function toclassify the pathology of cucumber leaves into seven categories: C =7 and batch capacity N =12. The calculation

formula of the loss function is as follows.1:

l(p, q)= L =

l1,, lN

Ò, lm =\*

C

c=1

wc log

exp

xm,c

≥C

i=1 exp

xm,i

ym,c (1)

Where x is the input,y is the target, w is the weight, and l is the loss function value.190

The loss value of each data sample is calculated through the cross-entropy Loss function. Then the total lossfunction of an epoch is added and calculated according to the batch size to obtain the loss value of the image in the

short author name: Preprint submitted to Elsevier Page 7 of 14 Table 1

MobileNet\_v3\_large network structure

Input Operator Exp size Out SE NL S224ù224ù3 conv2d ù16ù h-swish 2

112ù112ù16 bneck,3ù31616ù relu 1

112ù112ù16 bneck,3ù36424ù relu 2

56ù56ù24 bneck,3ù37224ù relu 1 56ù56ù24 bneck,5ù57240

˘

relu 2

28ù28ù40 bneck,5ù512040

˘

relu 1

28ù28ù40 bneck,5ù512040

˘

relu 1

28ù28ù40 bneck,3ù324080ù h-swish 2

14ù14ù80 bneck,3ù320080ù h-swish 1

14ù14ù80 bneck,3ù318480ù h-swish 1

14ù14ù80 bneck,3ù318480ù h-swish 1

14ù14ù80 bneck,3ù3480112ù h-swish 1 14ù14ù112 bneck,3ù3672112

˘

h-swish 1

14ù14ù112 bneck,5ù5672160

˘

h-swish 2

7ù7ù160 bneck,5ù5960160

˘

h-swish 1

7ù7ù160 bneck,5ù5960160

˘

h-swish 1

7ù7ù160 conv2d,1ù1ù960ù h-swish 1 7ù7ù960 pool,7ù7ùùùù1

1ù1ù960 conv2d 1ù1, NBN ù1280ù h-swish 1

1ù1ù1280 conv2d 1ù1, NBN ù k ùù1

experiment in Chapter 4. That is,2:

l(p, q)=

N

m=1

lm (2)

In this paper, the size of the loss function is taken as the benchmark for the superiority of the network model. In the191

follow-up experiments, we will ﬁnd that the network model we used has an overﬁtting phenomenon. A loss evaluation192

strategy needs to be replaced. After reaching a certain threshold, the strategy does not take simple loss decline as the193

training orientation. In this way, the loss on the test set shows a relatively ﬂat trend, and then the rising speed of the194

loss on the test set will be reduced, and even a secondary decline may occur. Finally, the accuracy is further improved195

to some extent. We needed a way to solve this problem, and the ﬂooding method came into being. Ishida et al.(2020)196

Consider input variable p À Jd and output variable q À[C]:={1,,C}, where C is the number of classes. They

follow an unknown joint probability distribution with density p(p, q). We denote the score function by f : Jd ô JC

. For any test data point p0, our prediction of the output label will be given by öq0:= argmaxzÀ[C] fz

p0

, where

fz() is the z -th element of f (), and in case of a tie, arg max returns the largest argument. Let l : JC ù[C]ô J

denote a loss function. l can be the zero-one loss, where w :=

w1,,wC

ÒÀ JC , or a surrogate loss such as the softmax cross-entropy loss,3:

lCE

w, z®

:=\* log

exp

wz®

≥

zÀ[C] exp

wz

.(3)

For a surrogate loss l , we denote the classiﬁcation risk. The goal of multi-class classiﬁcation is to learn f thatminimizes the classiﬁcation error J01(f ). In optimization, we consider the minimization of the risk with a almost

surely dierentiable surrogate loss J (f ) instead to make the problem more tractable. Furthermore, since p(p, q) isusually unknown and there is no way to exactly evaluate J (f ), we minimize its empirical version calculated from the training data instead4:

öJ (f ):=1

m

m

i=1

l

f

pi

, qi

(4)

where

pi, qi

m

i=1 are i.i.d. sampled from p(p, q). We call öJ the empirical risk.197

Definition1. The ﬂooded empirical risk is deﬁned as 4 õJ (f )=öJ (f )\* b+ b (5)

Note that when b =0, then õJ (f )=öJ (f ). The gradient of õJ (f )w.r.t. model parameters will point to the same direction198

as that of öJ (f ) when öJ (f )> b but in the opposite direction when öJ (f )< b . This means that when the learning199

objective is above the ﬂood level, we perform gradient descent as usual (gravity zone), but when the learning objective200

is below the ﬂood level, we perform gradient ascent instead (buoyancy zone).Pushing the parameters towards a more201

stable region keeps the convergence of the loss function near a threshold value, which improves the generalization202

performance and better resists perturbations.203 4. Experiment204

In this experiment, the image was preprocessed ﬁrst, and the PyTorch framework was used to scale the image to205

448ù448 for data standardization. In this paper, Mobilenet v3 network model was selected as well as optimizer ASGD,206

the learning rate was set to 0.001, the L1 regularity coecients were all 0.01, the batch size was 12, and 300 rounds207

of iterative training were conducted on the training set and the test set respectively. In order to prevent overﬁtting in208

the experiment, we also apply the algorithm of Dropout to randomly inactivate the neural nodes in the network before209

network training, reduce the interdependence between neurons, and thus ensure the extraction of important features210

that are independent of each other and improve the generalization ability of the model. As shown in the ﬁgure, it can211

be seen that neurons randomly deactivate seven neural nodes in the network.212

short author name: Preprint submitted to Elsevier Page 8 of 14 4.1. Contrast test213

We selected seven mainstream network models and MobileNet v3 network for a comparison test on the same214

cucumber pathological leaf image data set. The experimental data accurately reﬂected the superiority of MobileNet215 v3 network.216

In this paper, Alexnet, Resnet, VGG, Ecientnet v2, Ecientnet v3, Ecientnet v7, Mobilenet v2, and Mobilenet217

v3 leaf pathological recognition models were trained, and image training set and test set were used to test and compare218

them. This way, the network model performance’s superiority is tested and further optimized. The ﬁgure4 shows the219

experimental results, in which vg19 represents VGG model, alex represents Alexnet model, re50 represents Resnet220

model, mob3 represents Mobilenet v3 model, mob2 represents mobilenet v2 model. e7 represents the Ecientnet v7221

model, e3 represents the Ecientnet v3 model, and e2 represents the Ecientnet v2 model.222

The loss function and accuracy of the training set of Alexnet model converge well. When epoch=100, they begin223

to converge and gradually become stable. However, the eect on the test set could be better, loss and accuracy ﬂuctuate224

considerably, and the convergence eect could be better. Compared with other models, its loss in the test set is higher,225

its accuracy is lowest, and its performance is poor.226

The training set of the VGG model converges quickly, and the loss image of the data set begins to converge in227

the 70th round of iteration. The accuracy image of the training set and the test set converge faster and in about 30228

iterations. However, the degree of ﬁtting in the test set is not high, and the test set loss and accuracy image of the VGG229

model show no convergence trend. The average accuracy is less than 70%.230

The training set of Resnet begins to converge when the number of training rounds is around 40, and the convergence231

speed is breakneck. When testing the test set, we found that the convergence ﬁtting degree is considerably high, and232

the maximum accuracy is as high as 81.4%. However, the ﬂuctuation range of loss and accuracy of the test set is larger233

than that of Alexnet network model numerically, and the loss function also appears to be an overﬁtting phenomenon.234

The results of Ecientnet v2 and Ecientnet v7 models are the same. The loss and accuracy of the test set tend235

to converge, and the ﬁtting degree is higher than that of the training set. However, the degree of overﬁtting of the loss236

image of the test set is too high and ﬂuctuates wildly.237

The Ecientnet v3 model, where the precision image of the test set begins to converge around the 150th iteration238

round, is the slowest of all models. The test set of Ecientnet v3 shows a convergence eect, but a severe overﬁtting239

phenomenon occurs, and the ﬂuctuation is the largest from the experimental results.240

In the MobileNet v2 model, the training set starts to converge from the 30th iteration, the loss ﬁnally keeps ap-241

proaching 0, and the accuracy also keeps increasing with the training rounds. The resulting trend of the test set also242

roughly ﬁts the training set, but there is an overﬁtting phenomenon. The generalization degree is shallow, and the243 ﬂuctuation degree is enormous.244

short author name: Preprint submitted to Elsevier Page 9 of 14

Considering the ﬁtting eect of each model test set comprehensively, the test set result trend of Mobilenet v3245

network is consistent with the training set trend, and its maximum accuracy is relatively the highest, reaching 81.3% or246

above. Moreover, Mobilenet v3 converges 66% faster than other networks due to its lightweight framework. Therefore,247

we ﬁnally chose Mobilenet v3 network for the next optimization experiment.248 4.2. The choice of optimizer249

After selecting Mobilenet v3 as the ﬁnal experimental network, this paper optimizes it. The ﬁrst is the selection of250

the optimizer. The optimizer is used to update and calculate network parameters that aect model training and model251

output to approximate or reach the optimal value, thereby minimizing (or maximizing) the loss function. Choosing252

an appropriate optimizer can make our network model reach convergence faster and achieve better accuracy. On the253

same cucumber pathological leaf image data set, Mobilenet v3 network was selected in this paper, and the learning254

rate was set to 0.001, the regularity coecients of L1 and L2 were both 0.01, and 300 rounds of iterative training255

were conducted on the training set and test set respectively. In this paper, ASGD, SGD, RMSprop, RAdam, NAdam,256

AdamW, Adamax, Adadelta, and Agagrad are selected for comparison, and loss functions and accurate images of the257

training set and test set are obtained, as shown in the ﬁgure.5258

As shown in the ﬁgure, although the four optimizers, RAdam, NAdam, Adam, and RMSprop, all have a conver-259

gence trend at last, their loss value on the training set is very high, and their highest accuracy is not more than 60%.260

Compared with other optimizers, the eect could be better, and they are unsuitable for this paper’s network model.261

AdamW optimizer performs well in the training set. The convergence rate of the loss function and accuracy image262

is the fastest compared with other optimizers. However, its performance in the test set could be better. The loss function263

and accuracy image have large ﬂuctuations, and its maximum accuracy is at most 70%.264

Adamax optimizer begins to converge after 100 iteration rounds of the training set. Its loss function image value265

in the test set is higher than that of the ten optimizers, and its accuracy is low, with an average accuracy of less than266 65%.267

The data set images of Adadelta and Adagrad optimizers almost coincide. Both the training set and the test set268

converge. The loss value gradually decreases with the increase in the number of iterations, which is the lowest among269

the ten optimizers. The accuracy also increases with the number of iterations, reaching a high accuracy of 83%.270

However, its convergence speed could be faster. The training set begins to converge in the 250th iteration round, and271

the test set begins to converge in the 150th iteration round, which takes the longest time.272

The data set images of the two optimizers, ASGD and SGD, almost coincide, converge in both the training set and273

the test set, and the convergence is faster. The training set begins to converge in the 80th iteration round, and the test274

set begins to converge in the 20th iteration. The accuracy of the test set peaked at 81.4%. Numerically, the ASGD has275

short author name: Preprint submitted to Elsevier Page 10 of 14 Table 2

Use eﬀect of diﬀerent batch sizes

batch\_size max\_acc(%) mean\_acc(%) sstd\_acc(%) max\_loss mean\_loss std\_loss

479.87064.19710.9161367.639367.173181.912

881.79475.4923.227178.613125.19620.972

1282.11576.5682.962135.26176.6199.009

1681.54276.8483.556123.97655.9406.853

2082.45176.6264.069118.69944.2326.210

2481.75176.8544.498117.16636.6385.916

2882.24576.5275.140113.50231.2515.801

3282.23076.4865.303112.95827.3935.897

less ﬂuctuation than the SGD optimizer.276

Therefore, after comparing convergence speed,ﬁtting degree, accuracy, loss function size, and other aspects,277

ASGD has a higher peak value, faster convergence, and minor ﬂuctuation. In this article, the ASGD is chosen as278 the ﬁnal optimizer.279

subsectionThe choice of batch size In this comparison experiment, Mobilenet v3 network model was selected, the280

optimizer was ASGD, the learning rate was set to 0.001, the regularity coecients of L1 and L2 were both 0.01, and281

300 rounds of iterative training were conducted on the training set and test set respectively. Batch size was selected as282

4,8,12,16,20,24,28, and 32. The images and data results were obtained as shown in the ﬁgure below.6283

Batch size =4, as the batch size value is too small, the gradient of each layer has high randomness and takes much284

time. The resulting image also ﬂuctuates, and the ﬁnal precision eect is considerably poor, resulting in an underﬁtting285

phenomenon. The convergence eect is not good enough.286

It can be seen from the ﬁgure6 that the convergence speed increases with the increase in batch size. According287

to the numerical results,2 with the increase of batch size, the maximum loss function, average loss function, and the288

standard deviation of the loss function, namely the volatility, of the test set gradually decrease. However, after batch289

size =12, each loss value changes little with the increase in batch size. In addition, the maximum accuracy after290

convergence increases weakly and sometimes even regresses. Moreover, after batch size =12, the standard deviation291

of the accuracy of the test set began to rise continuously, indicating that the model’s generalization ability declined.292

Before batch size =12, the test set’s accuracy increases while the loss ﬂuctuation decreases. When batch size =12,293

the standard deviation of loss is minimum, and the anti-aliasing eect is best. At the same time, the accuracy of the294 test set increased to 82.1%.295

The experiment in this paper is carried out under a blurred background image, so we need as much generalization296

ability as possible. Moreover, the model proposed in this paper should apply to mobile terminal devices, should be as297

lightweight as possible, and need to select the smallest batch size value possible. Therefore, from the perspective of298

background requirements and image data analysis, batch size =12 was selected as the optimal experimental parameter299

short author name: Preprint submitted to Elsevier Page 11 of 14 Table 3

Use eﬀect after diﬀerent values of parameter b in ﬂooding

b max\_acc(%) mean\_acc(%) sstd\_acc(%) max\_loss mean\_loss std\_loss0.34981.64476.0252.980139.20763.7065.395

0.29181.60876.5083.076136.61065.3885.288

0.25282.06876.3443.080138.87967.1565.405

0.31183.30878.2632.658135.18764.8815.256

0.29781.64176.3152.980139.06565.0605.400

0.33080.89076.3953.248138.47564.6705.387

0.29781.37576.4582.953138.85864.9295.319

0.29681.56876.3853.146137.58065.5045.320

0.27482.55776.4973.249136.90965.8205.494

0.17182.27176.7173.136136.26969.9506.173

0.23282.82076.6103.306136.02167.7365.490

0.26581.78676.7023.238134.72766.2505.396

0.25681.82876.2173.118138.51066.8545.473

0.20781.63476.4963.263137.23769.3575.813

0.22382.58876.7823.382136.99567.8195.571 in this paper.300

4.3. Flooding301

In Chapter 3, we introduced the basic principle and used a ﬂooding mode. By changing the loss function and302

adding a threshold, the loss eventually ﬂuctuated around the threshold. Flooding allows us to directly select the level303

of training loss, which is dicult to achieve with other regularizers. There was an overﬁtting phenomenon in the loss304

result in images of the Mobilenet v3 experiment mentioned above. In this section,ﬂooding was used to realize the305

secondary decrease of data set loss and prevent overﬁtting.306

In this experiment, the optimizer used ASGD, the learning rate was 0.001, the regularization coecients of L1307

and L2 were 0.01, and 300 iteration experiments were conducted. The loss threshold is set with 15 dierent values for308

comparative analysis of images and data.309

As seen from the image,7 after ﬂooding was added, the overﬁtting rising trend of the loss function image in the test310

set was eectively suppressed. When b =0.310,0.348, and 0.290, the ﬂooding not only resulted in good inhibition311

but also resulted in secondary descending, which solved the overﬁtting problem.312

According to a series of comparisons of table data, after adding ﬂooding to 3, the mean test set accuracy increased313

by 0.2%, and the maximum test set accuracy increased by 0.5%. The ﬁnal goal of this paper is to select the test set314

with the highest accuracy to achieve the best pathological recognition eect of the cucumber leaf image. The ﬁnal315

selection threshold is 0.310, at which time the overﬁtting of the loss function is well suppressed, and the accuracy is316 317

We compared the two experiments without ﬂooding with the ﬂooding method. The results are shown in the ﬁgure.8318

Methods with ﬂooding tend to improve test accuracy compared to baseline methods without ﬂooding. Continue to train319

short author name: Preprint submitted to Elsevier Page 12 of 14

the model without ﬂooding until, eventually, the loss function may continue to rise and accuracy may decline. However,320

according to the results, the ﬁnal model has good predictive performance when there is ﬂooding, which means that321

ﬂooding helps improve test accuracy in later training. During training with ﬂooding, test losses became lower and322

ﬂatter. On the other hand, the training loss reached a secondary decline and continued to ﬂoat around the ﬂooding323 threshold with stability.324

4.4. Discussion325

In this study, the data set was replaced with PlantVillage public data set and another public apple disease data set326

in China to conduct the pathological judgment experiment of apple leaves. In this experiment,10 rounds of iterative327

experiments were conducted. The experimental results are shown in the ﬁgure below, where the apple curve represents328

the experimental results using the apple disease data set. The plant curve represented the experimental results using329

the PlantVillage public data set.330

As can be seen from the image results,9 the loss function of the training set and the test set is constantly close to331

0, and the accuracy also increases with the increase of iteration rounds. The accuracy of PlantVillage public data set332

after applying the strategy in this paper is as high as 99%, and the accuracy of the apple disease data set is also as high333

as 98.1%, which is far higher than the 76.5% accuracy of Zhou Minmin’s apple-leaf-disease-detection-system based334

on transfer learning.Minmin (2019) It is proved that compared with the existing strategies, the proposed strategies are335

universal, accurate, and less time-consuming and can better meet the needs of Chinese farmers for crop pathological336 judgment in today’s society.337

5. Conclusion338

In today’s society, the rise of the online QA system has brought great convenience to people’s lives, but it is not339

widely used in agriculture. The pathological judgment of agricultural plants is an essential part of agricultural planting340

life. Today’s crop pathological judgment mostly requires high labor costs, low detection eciency, and poor reliability341

because of its dense growing environment and chaotic background.342

To solve these problems, this paper proposed a Mobilenet v3 based on ﬂooding to identify crop leaves in fuzzy343

scenes. It satisﬁed the requirement of mobile terminal using a lightweight framework and could quickly and accurately344

judge crop pathological conditions through farmers’ shooting pictures.345

In this paper, cucumber leaf images were randomly collected from a Chinese agricultural website and labeled.346

A dataset with complex image background was constructed, and seven kinds of cucumber leaf pathologic judgments347

were made. Through the control variable method, the network model, the optimizer, and batch size, three rounds of348

experiments were compared and analyzed to achieve the optimal network model. In this paper,ﬂooding method was349

short author name: Preprint submitted to Elsevier Page 13 of 14

used to replace an evaluation strategy of loss. The accuracy of the test set was increased by 0.5% again, reaching the350

highest 88.3%. Finally, two public data sets of PlantVillage and apple disease were selected for the experiment again.351

The accuracy was up to 99% and 98.1%, respectively, which proved the universality of the proposed strategy and its352 353

References354

Benfenati, A., Causin, P., Oberti, R., Stefanello, G.,2021. Unsupervised deep learning techniques for powdery mildew recognition based on355

multispectral imaging. arXiv preprint arXiv:2112.11242.356

Cap, Q.H., Uga, H., Kagiwada, S., Iyatomi, H.,2020. Leafgan: An eective data augmentation method for practical plant disease diagnosis. IEEE357

Transactions on Automation Science and Engineering .358

DeChant, C., Wiesner-Hanks, T., Chen, S., Stewart, E.L., Yosinski, J., Gore, M.A., Nelson, R.J., Lipson, H.,2017. Automated identiﬁcation of359

northern leaf blight-infected maize plants from ﬁeld imagery using deep learning. Phytopathology 107,1426-1432.360

Howard, A., Sandler, M., Chu, G., Chen, L.C., Chen, B., Tan, M., Wang, W., Zhu, Y., Pang, R., Vasudevan, V., et al.,2019. Searching for361

mobilenetv3, in: Proceedings of the IEEE/CVF international conference on computer vision, pp.1314-1324.362

Hughes, D., Salathé, M., et al.,2015. An open access repository of images on plant health to enable the development of mobile disease diagnostics.363

arXiv preprint arXiv:1511.08060.364

Ishida, T., Yamane, I., Sakai, T., Niu, G., Sugiyama, M.,2020. Do we need zero training loss after achieving zero training error? arXiv preprint365 arXiv:2002.08709.366

Karthik, R., Hariharan, M., Anand, S., Mathikshara, P., Johnson, A., Menaka, R.,2020. Attention embedded residual cnn for disease detection in367

tomato leaves. Applied Soft Computing 86,105933.368

Kawasaki, Y., Uga, H., Kagiwada, S., Iyatomi, H.,2015. Basic study of automated diagnosis of viral plant diseases using convolutional neural369

networks, in: International symposium on visual computing, Springer. pp.638-645.370

Martinelli, F., Scalenghe, R., Davino, S., Panno, S., Scuderi, G., Ruisi, P., Villa, P., Stroppiana, D., Boschetti, M., Goulart, L.R., et al.,2015.371

Advanced methods of plant disease detection. a review. Agronomy for Sustainable Development 35,1-25.372

Minmin, Z.,2019. Research on Android Detection System of Apple Leaf Disease Based on Transfer Learning. Master’s thesis. Northwest AF373 University.374

Prasanna Mohanty, S., Hughes, D., Salathe, M.,2016. Using deep learning for image-based plant disease detection. arXiv e-prints , arXiv-1604.375

Saikawa, T., Cap, Q.H., Kagiwada, S., Uga, H., Iyatomi, H.,2019. Aop: An anti-overﬁtting pretreatment for practical image-based plant diagnosis,376

in:2019 IEEE International Conference on Big Data (Big Data), IEEE. pp.5177-5182.377

Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D., Stefanovic, D.,2016. Deep neural networks based recognition of plant diseases by leaf378

image classiﬁcation. Computational intelligence and neuroscience 2016.379

Wspanialy, P., Moussa, M.,2020. A detection and severity estimation system for generic diseases of tomato greenhouse plants. Computers and380

Electronics in Agriculture 178,105701.381

Zhang, Y., Song, C., Zhang, D.,2020. Deep learning-based object detection improvement for tomato disease. IEEE Access 8,56607-56614.382

Zhonghua, Y., Mingxia, Z., Lu, J.,2021. Research on crop disease image recognition with complex background. Transactions of the Chinese383

Society for Agricultural Machinery .384

short author name: Preprint submitted to Elsevier Page 14 of 14

short author name: Preprint submitted to Elsevier Page 15 of 14 List of Figures385

1 Cucumber complex background data set presentation ..........................16386

2 Article algorithm .............................................17387

3 A comparison of the neural network before and after using dropout ...................18388

4 Comparison Experiment between mobilenet v3 and Mainstream Image Classiﬁcation Algorithms ...19389

5 Comparison Experiment between ASGD and Mainstream Optimizer Experiment ...........20390

6 Comparative Experiments with Dierent batch Sizes ..........................21391

7 Comparative Experiment on Dierent Values of Parameter b After Using ﬂooding ...........22392

8 Comparative Experiment Before and After Flooding ...........................23393 9 Eect experiment of mobile v3 based on ﬂooding on dierent data sets .................24394