

RESEARCH

Open Access



# Pseudo high-frequency boosts the generalization of a convolutional neural network for cassava disease detection

Jiayu Zhang<sup>1</sup>, Chao Qi<sup>1</sup>, Peter Mecha<sup>1</sup>, Yi Zuo<sup>1</sup>, Zongyou Ben<sup>1</sup>, Haolu Liu<sup>2</sup> and Kunjie Chen<sup>1\*</sup>

## Abstract

Frequency is essential in signal transmission, especially in convolutional neural networks. It is vital to maintain the signal frequency in the neural network to maintain the performance of a convolutional neural network. Due to destructive signal transmission in convolutional neural network, signal frequency downconversion in channels results into incomplete spatial information. In communication theory, the number of Fourier series coefficients determines the integrity of the information transmitted in channels. Consequently, the number of Fourier series coefficients of the signals can be replenished to reduce the information transmission loss. To achieve this, the ArsenicNetPlus neural network was proposed for signal transmission modulation in detecting cassava diseases. First, multiattention was used to maintain the long-term dependency of the features of cassava diseases. Afterward, depthwise convolution was implemented to remove aliasing signals and downconvert before the sampling operation. Instance batch normalization algorithm was utilized to keep features in an appropriate form in the convolutional neural network channels. Finally, the ArsenicPlus block was implemented to generate pseudo high-frequency in the residual structure. The proposed method was tested on the Cassava Datasets and compared with the V2-ResNet-101, EfficientNet-B5, RepVGG-B3g4 and AlexNet. The results showed that the proposed method performed 95.93% in terms of accuracy, 1.2440 in terms of loss, and 95.94% in terms of the F1-score, outperforming the comparison algorithms.

**Keywords:** Disease detection, Pseudo high-frequency, Multi-attention, Instance batch normalization, Fourier analysis

## Introduction

Cassava (*Manihot esculenta Crantz*) is one of the most common crops widely grown throughout the world and is a major staple food crop, feeding approximately 800 million people worldwide in Africa (55.5%), Asia (30.2%), Americas (14.3%) and Oceania (0.1%) [1, 2]. Cassava is used as fodder and starch to develop ethanol fuel and as an industrial raw material. During food crises, research and exploration of cassava disease diagnosis using vision algorithms have helped people manage the crises and ensure that no unnecessary losses to crops occur.

There are more than 30 known cassava leaf diseases [3], of which four diseases, named cassava bacterial blight (CBB), cassava brown streak (CBS), cassava mosaic (CMD) and cassava green mottle (CGM) are extremely damaging to cassava and are the main ones which will cause cassava yield reduction.

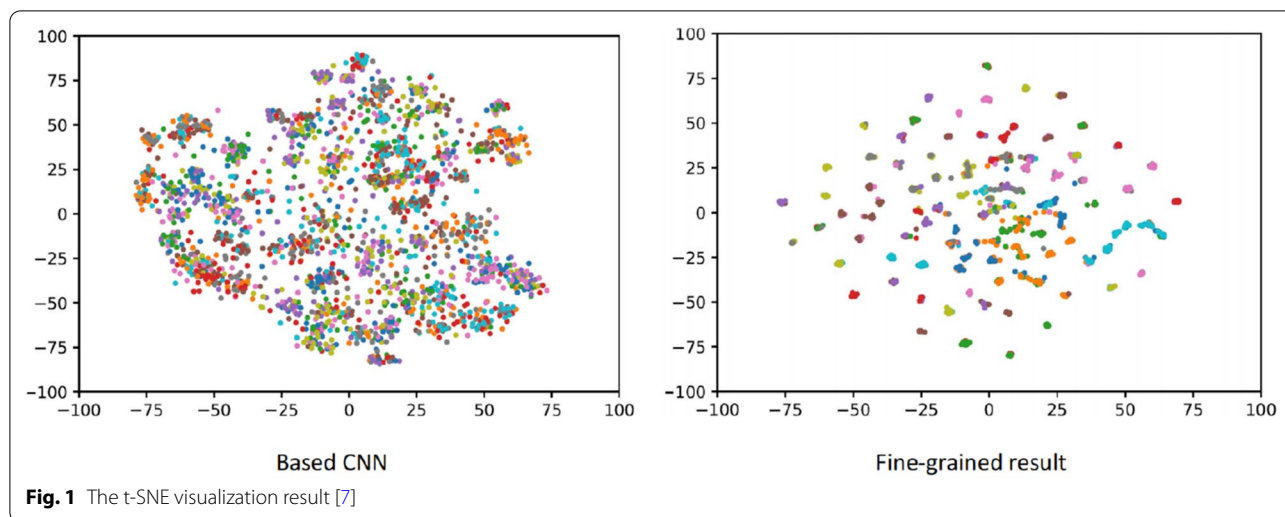
Growing cassava on small, and large scales across Southeast Asia and Africa has been challenging. The primary challenge is that cassava plants are vulnerable to a broad range of diseases as well as lesser-known viral strains. The incidence of epidemics of cassava mosaic virus has increased for decades in East Africa, especially the brown streak virus (CBS), leading to losses of 47% of production and US\$60 million per annum (*in lost yield*) and causing local famine. This has resulted in significant

\*Correspondence: kunjiechen@njau.edu.cn

<sup>1</sup> College of Engineering, Nanjing Agricultural University, Nanjing, China  
Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.



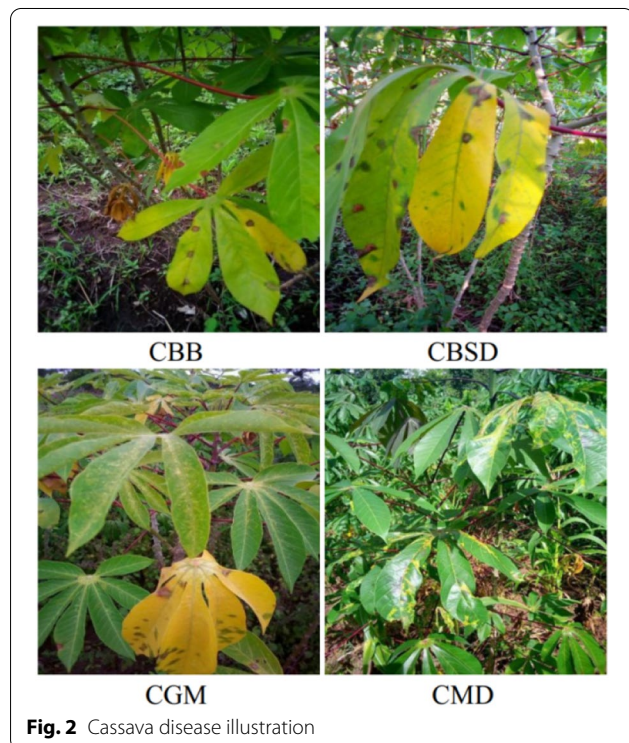
investments in plant breeding programs to overcome this issue [4]. Cassava bacterial blight disease (CBB) is a major constraint on cassava cultivation worldwide, and losses have exceeded 50–75% in regions where highly susceptible cultivars are grown [5]. To recognize disease rapidly, researchers have been exploring effective means of detecting diseases in cassava using visual algorithms.

Plant disease detection is a branch of fine-grained problems that can be expressed using the t-SNE (*t-Distributed Stochastic Neighbor Embedding*) algorithm [6]

to indicate the class separability and compactness in features extracted from a convolutional neural network [7]. The t-SNE visualization result is illustrated in Fig. 1. However, different from the clear background of images in the common fine-grained dataset, the cassava disease images in paper were captured in a real scenario with significant disorder texture, similar colour distribution, and irregular gradient disturbance. With the rapid development of the technology, the fine-grained research has been considered a high-performance feature descriptor for the encoder of the neural network, such as the EfficientNets algorithm [8]. Cassava diseases are shown in Fig. 2.

Ai et al. [9] utilized the Inception-ResNet-V2 model to recognize diseases in an efficient approach. The researchers used the competition disease leaf dataset to find the most efficient model by using an image dataset of 47,363 images for 27 disease-related 10 crop varieties. The based-inception algorithm structures exhibit excellent performance for fine-grained tasks based on transfer learning [10]. Fu et al. [11] proposed an algorithm to introduce the attention proposal sub-network (APN) as the local attention mechanism for convolutional neural networks for fine-grained tasks. The APN algorithm eliminates useless information and pays more attention to local responses. Fine-grained technology is essential for the development of neural networks, especially in person re-identification technology [12, 13].

Using deep neural networks, significant applications can be implemented in plant disease detection tasks. Various technologies have been utilized in neural networks to pursue high-performance results. These technologies include transfer learning, multi-task learning, meta learning [14], fine-tuning methods [14], ensemble learning [15], knowledge distillation [16], and loss function



fusion [17]. Several applications have been used in the literature. For instance, Tetila et al. [18] proposed a neural network algorithm to automatically recognize soybean leaf diseases based on unmanned aerial vehicle (UAV) images. The result of this automatic algorithm had 99.04% in terms of its accuracy based on the fine-tuning method. However, the number of images was too low to provide many features of disease detection in real scenarios. The performance of this neural network was based on transfer learning to fine-tune the neural network weight. MobileNet [19], a lightweight-class CNN-based algorithm [20], achieved an accuracy of 94% in cassava disease diagnosis. This algorithm was pretrained on the COCO dataset. Singh et al. [21] proposed a preprocessing algorithm to process images of mango leaf datasets and proposed a customized algorithm to detect the anthracnose disease in mango leaves with the dropout algorithms. As stated in a study by Li et al. [22], the variance shift in dropout was different from batch normalization, which illuminated an applicable case for plant disease detection. Many studies have been conducted to find an appropriate expression for features to make up for the limitations of Batch Normalization. Nonetheless, more research is needed [23–27]. For example, background images in the research of Singh et al. [21] were not clearly captured in a field the fields scenario. This may make neural networks unsuitable for detecting leaf diseases. Yuan et al. [28] proposed a spatial pyramid-oriented encoder-decoder method cascade with a convolutional neural network for crop disease segmentation to locate the infected regions of leaves. This disease segmentation algorithm was 90% accurate based on K-fold cross-validation. The number of parameters and the inference time may not be considered in many research explorations but can be considered in the deployment stage. Zhang et al. [29] proposed the global pooling dilated convolutional neural network to detect cucumber leaf disease. The researchers used the inception block to develop high-level feature maps based on the AlexNet structure and replaced the fully connected layer with a global pooling layer to reduce the network parameters. The results showed that the AlexNet neural network was a classical algorithm. However, the spatial dimension decreases, as each convolutional layer or block is followed by a sub-sampling layer [30]. Therefore, Han et al. [31] argued that in deep CNNs, a drastic increase in the feature-map depth and, at the same time, the loss of spatial information limits the learning ability of CNNs. Reyes et al. [32] used a pre-trained convolutional neural network using 1.8 million images and a fine-tuning strategy to transfer the learned recognition ability from the general domain to the specific challenge of the plant recognition task. Lee et al. [33] proposed a deep learning approach to quantify discriminatory leaf.

Thai et al. [34] proposed a vision transformer (ViT) [35] to detect the early leaf disease. It was a expensive method for plant disease detection, however, its a powerful solution for early leaf disease detection. De et al. [36] apply Faster Region-based Convolutional Neural Network (F-RCNN) to detect and recognize tomato plant leaf disease. Zhang et al. [37] improve F-RCNN by replacing VGG16 with a depth residual network resulting in 2.71% higher recognition accuracy compared with previous work. RepVGGS may be an excellent solution in F-RCNN. The reparametrization method can be utilized to boost the generalization of VGG neural networks. Sun et al. [38] used data enhancement and image segmentation for tea images and achieved higher accuracy through frequently adjusting iteration times and learning rates. Zhou et al. [39] proposed a deep residual dense networks to obtain higher accuracy in classifying tomato leaf diseases using fewer parameters. Oyewola et al. [40] proposed the detection of cassava mosaic disease using deep residual convolutional neural networks with different computation block.

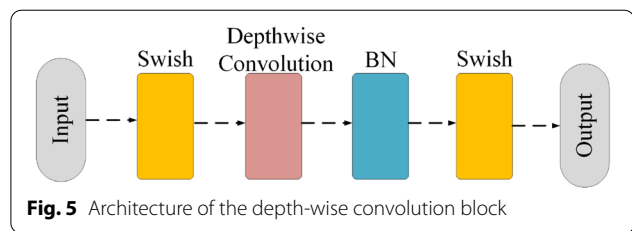
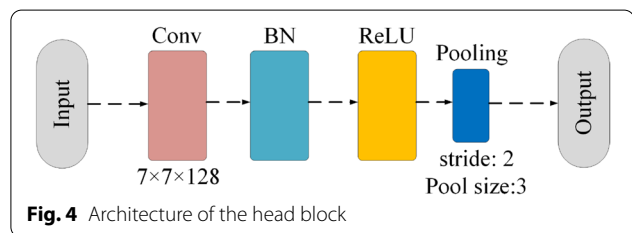
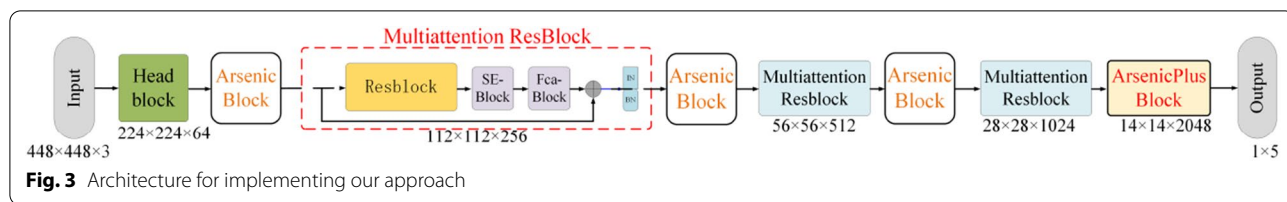
In a third generation neural network [41], variation of light in an image has an essential property in feature description. The texture information expresses the high-frequency component in the images [42]. In the study of Wang et al. [43], a high-frequency component is known to boost the generalization performance in a convolutional neural network. As mentioned above, the proposed multi-attention mechanism proposed maintained the long-term dependency of the feature maps in neural network channels. To comply with the constraints of the *Nyquist–Shannon* sampling theorem, the Arsenic Block was proposed to downconvert the signal frequency in channels of the neural network. The pseudo high-frequency component was utilized to maintain the number of Fourier series coefficients of signals in neural network channels.

The field images are utilized in this paper to overcome implicit obstacles in the field [44].

### The proposed method

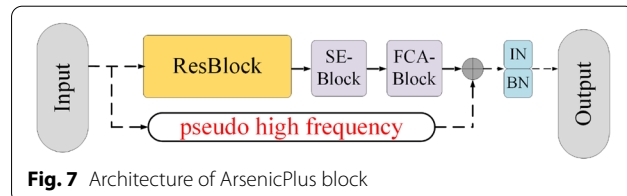
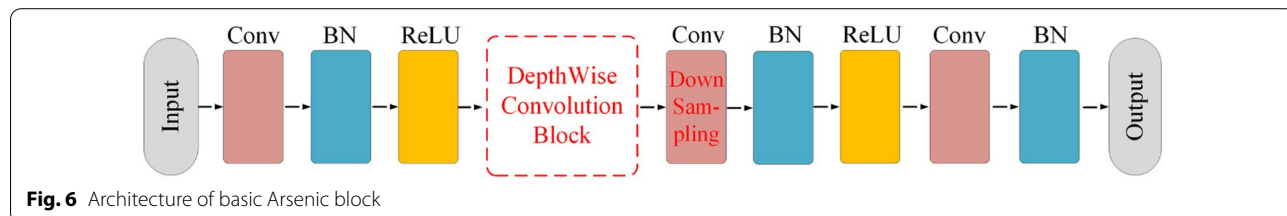
A large dataset may cause a lower angular frequency of the kernel function. Consequently, based on the property of the convolution, the high-frequency of the convolution kernel function is maintained, and more information can be maintained in the neural network channels. Thus, more effective information can be saved in the filter operations. The effective information can be expressed as an objective function of the input signal in the mathematical expression.

Angular frequency is essential for maintaining feature long-term dependency to keep the objective function with arbitrary small loss in a convolutional neural



network. When the angular frequency of the convolution kernel function refers to  $\omega_{kernel} \rightarrow 0$ , and  $\omega_{kernel} \neq 0$ , the objective function of the input signal refers to  $S$ , and the frequency of  $\omega_S = \rho$ , the most ideal case is  $\frac{\omega_S}{\omega_{kernel}} = C, C \in N^+$ . When the angular frequency of the Fourier series coefficients of the convolution kernel function refer to  $\omega_{kernel} \rightarrow 0$  and the action scope of the Fourier series was  $\infty = \lim_{\omega \rightarrow 0} \frac{2\pi}{\omega}$ , all objective functions of the input signals will be maintained in this convolution operation. Kernel functions uniform convergence to a good kernel function is stated in [Appendix](#).

Considering the indicators of *GFLOPs* and *Parameters* in the neural network, v2-ResNet-101 was utilized as the baseline. The pipeline is illustrated in [Fig. 3](#). [Figure 4](#) shows the head block utilized to capture contour information at the beginning of the network. The depth-wise



convolution block of the essential component of the Arsenic basic block is illustrated in [Fig. 5](#). The Arsenic block is illustrated in [Fig. 6](#).

ArsenicNet is composed of a multi-attention ResBlock and Arsenic block. The multi-attention ResBlock was modified with a pseudo high-frequency component to give the ArsenicPlus block. The ArsenicPlus block was the basic component in stage 4 [45] of ArsenicNetPlus. The other stages in ArsenicNetPlus were maintained in the architecture of ArsenicNet without being modified. The architecture of the ArsenicPlus block is illustrated in [Fig. 7](#).

### Keeping long-term dependency based on multi-attention component

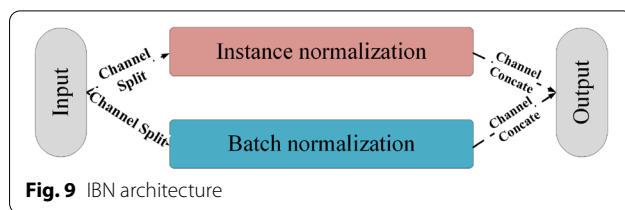
In the communication theorem, the greater the numbers of Fourier series coefficients, the clearer the information transmitted in channels. This idea can be transferred from communication theory to neural networks. More numbers of Fourier series coefficients of the signals help boost the generalization of the neural networks.

Subsequently, the multi-attention component was proposed to maintain long-term dependency in feature maps. The SE-block [46] and FCA-block [47] were utilized as the multi-attention structure. The multi-attention is a product of two linear transformation coefficients, and the architecture is shown in [Fig. 8](#).



**Boosting the generalization using instance batch normalization**

The instance batch normalization (IBN) [48] is a special algorithm that can be applied to convolutional neural networks. It is a combination of the instance normalization and batch normalization [49]. The architecture of the IBN is illustrated in Fig. 9.



**Fig. 9** IBN architecture

**Complying with the restrictions in down-sampling limitation**

In the *Noisy-channel coding theorem*, which demonstrates that if the transmission rate  $R \leq$  capacity  $C$ , there exists an encode mode to transmit information with minimum error probability. The correlation in bandwidth  $B$ , capacity  $C$ , and white Gaussian noise is stated as follows:

$$C = B \log_2 \left( 1 + \frac{S}{N} \right), \tag{1}$$

where  $C$  refers to the capacity of channels,  $B$  refers to the bandwidth,  $S$  refers to the signal power, and  $N$  refers to the noise power. The *Noisy-channel coding theorem* is applicable to digital signals and analogue signals. If the noise in the convolutional neural network can be controlled and it approaches to zero, then  $C\epsilon \rightarrow 0$ ; thus  $\frac{S}{T} \rightarrow \infty$ . The *Noisy-channel coding theorem* can be rewritten as follows:

$$C = B \log_2(1 + \infty). \tag{2}$$

Based the curve of  $\log_2 \infty$ , the asymptotic value  $C$  can be calculated based on bandwidth  $B$ , where  $C$  is the actual coding capacity. The coding capacity is an unknown parameter in a convolutional neural network.

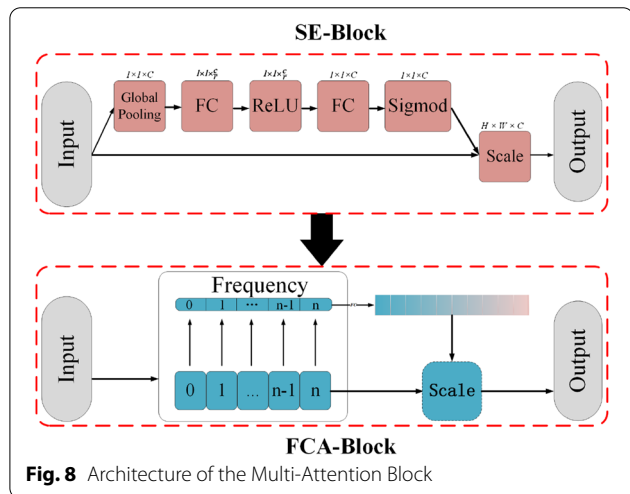
In a convolutional neural network, the sampling frequency did not comply with the definition of the

*Nyquist–Shannon sampling theorem*. To comply with the restrictions of the down-sampling limitation, the Arsenic block plays two roles in the proposed neural network. First, it cleans the aliasing signals in feature maps. Second, it down-converts the frequency to comply with the down-sampling frequency. However, the feature descriptor was not an arbitrarily small loss coding operator, when the neural network did not use transfer learning.

Based on the abovementioned information, the signal frequency in the neural network channels will eventually meet the down-sampling frequency limitation. A weaker signal frequency was proven to exist in the stage 4 of the convolutional neural network, which is illustrated in Table 1. Thus, to repair the weakness signals, the ArsenicPlus block (Fig. 7) was utilized in stage 4 of the proposed method. The results in Table 1 were evaluated using 7-fold cross-validation.

The *Nyquist–Shannon sampling theorem* was found to be applicable to convolutional neural networks and was named ArsenicNet. To evaluation the generalization of ArsenicNet, the Fine-Grained Visual Classification of Aircraft (FGVC-Aircraft) dataset was utilised in this paper. The FGVC-Aircraft dataset was cited in over 1000 papers, and was utilised as a benchmark dataset in over 200 papers [50].

As statistic in Table 6, the ArsenicNet-3 (based ResNet50) has achieved 84.70% in terms of accuracy, that is 5.9% higher than the experimental consequence of ResNet-50 method in the study of Lee et al. [51] in terms accuracy. Therefore, ArsenicNet is mentioned in this paper as the basic neural network of the ArsenicNetPlus neural network.



**Fig. 8** Architecture of the Multi-Attention Block

**Table 1** Results of ArsenicNet method (based on V2-ResNet-101)

Method	Accuracy (%)	Recall (%)	Precision (%)	F1-Score (%)
ArsenicNet-3	95.55	95.52	95.59	95.56
ArsenicNet-1	95.03	94.97	95.06	95.01
ArsenicNet-2	95.43	95.40	95.47	95.44
ArsenicNet-4	95.36	95.34	94.95	94.91

The suffix number of ArsenicNet is the number of Arsenic blocks in the neural network which are only conducted before down-sampling

### Building pseudo high-frequency residual structure

As stated in the studies of Wang et al. [43], high-frequency plays an important role in convolutional neural networks. Unfortunately, a signal with destructive transmission in a convolutional neural network causes a high-frequency loss, and the signals fall to an extreme weakness state.

The extremely weakness signals can not provide an accurate representation of the objective function in the source signal. We put forward a new concept: pseudo-high frequencies, and invented a method by adding pseudo-high frequencies to extreme weakness signals to maintain the integrity of the signal as much as possible, which can be used to solve such questions. Consequently, it is difficult to reconstruct the extreme weakness signals to the original signals. The pseudo high-frequency approximates the original signals rather than restoring the original signals. The equations of the pseudo high-frequency component are as follows:

1. Initialize an offset template. Set Matrix  $M \in C^{m \times n}$ ,  $C_{init}^{m \times n} = 1.0$ , and update the value of Matrix  $M$  via backpropagation.

$$M_{init} = 1.0, trainable = True. \quad (3)$$

2. Matrix  $M$  is used as the exponent of the input tensor:

$$T_{i,j} = x_{i,j}^{M_{i,j}}, \in [0, 1, \dots, channel], \quad (4)$$

where  $i$  refers to the width index of feature map  $x$ ,  $j$  refers to the height index of feature map  $x$ , and refers to the index of feature map channels.

$$f(x) = \int_{-\infty}^{\infty} F(jn\omega) e^{jn\omega x} d_n \rightarrow f(x)^{M_{i,j}} = \int_{-\infty}^{\infty} [F(nj\omega)^{M_{i,j}}] e^{(jn\omega \times M_{i,j})} d_n. \quad (5)$$

This is a stretching operation in the frequency domain; nevertheless, the nonlinear phase spectra change causes distortion of the signal distribution. Hence, this pseudo high-frequency residual operation was utilised only once in the ArsenicPlus block (Fig. 7) of stage 4 of the proposed neural network to replenish the pseudo high-frequency in the weak signals.

## Experiments

### Cassava datasets

There are 21,393 images maintained in the original cassava leaf disease dataset. The origin dataset was not kept balanced for data distribution to categories, and the most imbalance categories of CMD disease and CBB disease had 13,158 images and 1086 images, respectively. The imbalanced data distribution was an obstacle for plant

disease detection training. The imbalanced distribution may cause the the primary performance to tilt to the most images of the categories.

This network dataset has a significant number of imprecise images. To avoid the problem of image pollution, downstream projects cause downstream of models such as costly iterations, discard, and harm to communities [52]. Three main problems images were removed [53]. The three problems are shown as follows:

- (1) Unmaintained attributes: The unclear and low-quality images of cassava leaf disease. It is difficult to clearly distinguish regions of disease in these images.
- (2) Typing error: Labeling errors were present. The origin cassava leaf disease dataset includes not only cassava leaves but also cassava fruits, magazine covers and other unrelated material.
- (3) Inaccurate data: Losing focus. Losing focus will cause high-frequency component loss in images. The high-frequency component of the images was an essential component to boost the generalization in a convolutional neural network. Thus, the inaccurate data will destroy the downstream project.

Based on the abovementioned items, there were found more than 1000 healthy category images with niduses, which is an unacceptable rate of disease diagnosis errors in medical image diagnosis. To maintain balance among categories, the Gaussian noise, horizontal flipping, cutting-out, and vertical flipping were used to conduct augmentation. The 20,000 colour images were randomly combined into five balanced categories, and the CMD category in this paper was selected from 13,158 images from raw data with random. The preprocessed images with a resolution of  $448 \times 448 \times 3$  pixels, and the details are presented in Table 2.

There are approximately 3400 bad lighting and back-light cassava images, accounting for 17% of the image dataset in this paper, and partially obstructed in approximately 2000 images, accounting for 10% of the dataset.

### Experimental parameters and methods

#### Experimental parameters and methods for performance comparison

The proposed method was trained on the cassava dataset using the following settings: an stochastic gradient descent (SGD) optimizer [54] was used with an initial learning rate of 0.2, decay of 0.96 in every epoch, momentum of 0.9, weight decay of  $1e-5$ , and batch normalization momentum of 0.9. The coefficient of L2 regularization in descriptor is set to  $1e-5$ . The Hard-Sigmoid function in SE-Block reduces the computing cost of the neural network.

**Table 2** Dataset analysis

Category	Base	Noise	Horizontal	Vertical	Coutout	Total
CBB	986	985	986	986	57	4000
CBSD	1772	–	1772	456	0	4000
CGM	1861	28	1861	250	0	4000
CMD	4000	–	–	–	0	4000
Healthy	1054	838	1054	1054	0	4000

**Table 3** ArsenicNetPlus method versus other methods on cassava dataset using 7-fold cross-validation

Method	Accuracy (%)	Recall (%)	Precision (%)	Loss	F1-score (%)
ArsenicNetPlus	95.93	95.91	95.98	1.244	95.94
V2-Resnet-101	86.90	86.80	87.02	0.779	86.92
EfficientNet-B5	92.43	92.37	92.46	1.469	92.42
AlexNet	62.46	61.98	62.94	2.718	62.46
RepVGG-B3g4	93.08	93.07	93.14	0.347	93.11

Categorical cross-entropy was utilized as the loss function in this paper. This experiment utilized 7-fold cross-validation to obtain the representativeness result.

The proposed network was compared with EfficientNet-B5 [8], RepVGG-B3g4 [55], V2-Resnet-101 [45], and AlexNet [56]. As stated in the study of Ferentions [57], the VGG nuclear network and AlexNet accuracy have been ranked as first and second over other neural networks. The classical neural network VGG was modified to a new structure named RepVGG.

#### Experimental parameters and methods for ArsenicNet

The parameters and methods used in the experiment are consistent with those mentioned above. The proposed network was compared with the ArsenicNet neural network to verify the effectiveness of the pseudo high-frequency component.

## Results and discussion

### Classic algorithm comparison results

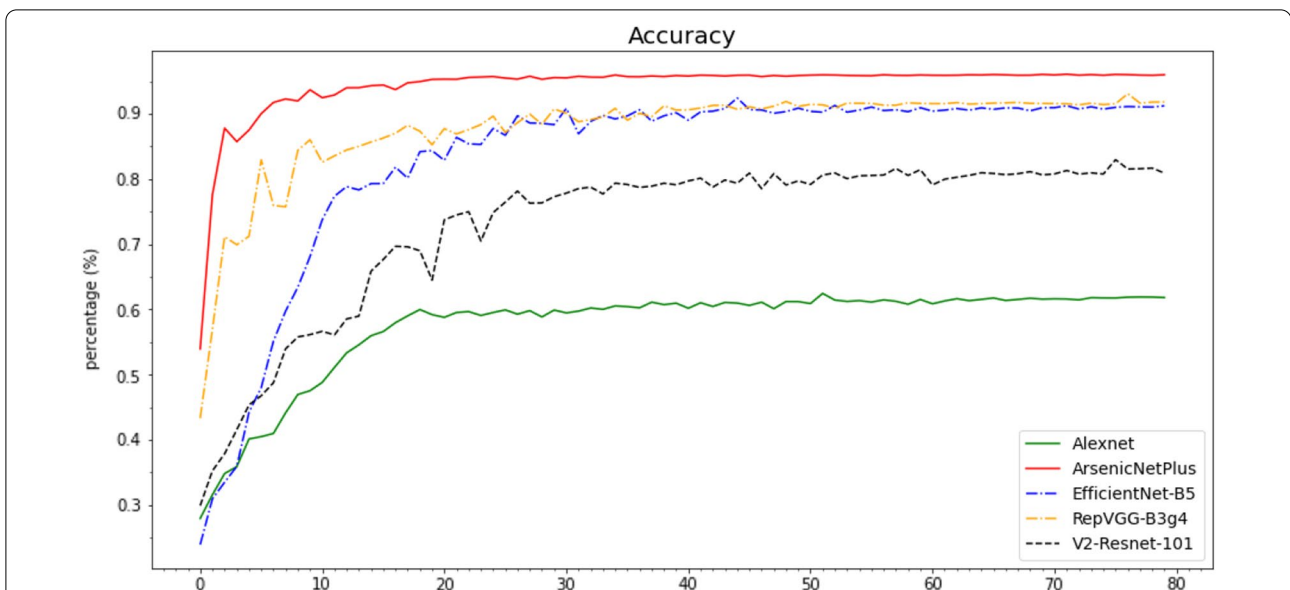
In this section, several classical algorithms including V2-ResNet-101, EfficientNet-B5, AlexNet, and RepVGG-B3g4 were compared with ArsenicNetPlus. Notably, this comparison did not use transfer learning and ensemble learning. The comparison results are illustrated in Table 3. The experimental software platform is TensorFlow framework 2.4.1, and the hardware is AMD Ryzen 7 3800XT @3.89GHz with a NVIDIA GeForce RTX 3090.

**Table 4** Accumulative confusion matrix of ArsenicNetPlus in 7-fold

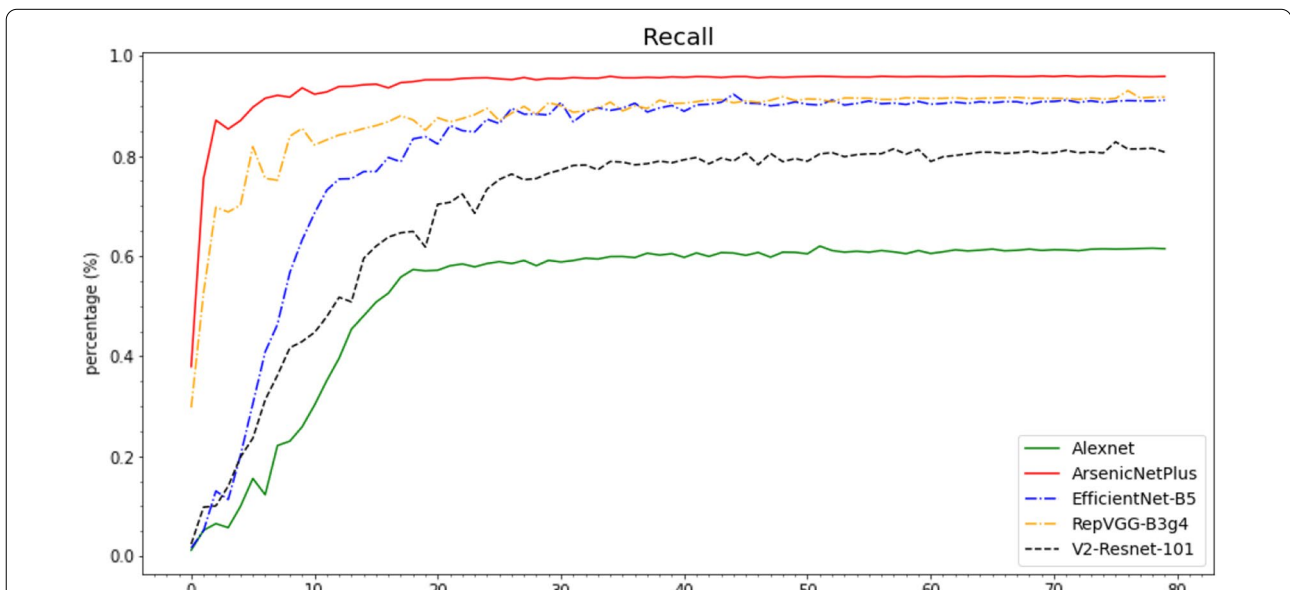
CLASS	CBB	CBSD	CGM	CMD	Healthy
CBB	3820	173	0	7	0
CBSD	221	3730	0	49	0
CGM	2	0	3751	0	247
CMD	9	69	0	3922	0
Healthy	0	0	105	0	3895

The above-mentioned classical methods were not have an indicator for the the extreme weakness signals, and not have the ability to repair the extreme weakness signals. The Arsenic block can be utilized as an indicator to check the extreme weakness signals, and ArsenicNetPlus block can be utilised in the extreme weakness stage to boost performance.

The Accuracy (Fig. 10), Recall (Fig. 11), Precision (Fig. 12), and F1-Score (Fig. 13) curves of ArsenicNetPlus are similar. The formulas for accuracy, recall, precision, and F1-score are as follows:  $accuracy = \frac{TP+TN}{TP+TN+FP+FN}$ ,  $recall = \frac{TP}{TP+FN}$ ,  $precision = \frac{TP}{TP+FP}$ , and  $F1 - score = \frac{2*precision*recall}{precision+recall}$ , respectively. The fluctuations of the aforementioned indicators had a narrow interval and were smoother than those of the other comparison algorithms used. The validation loss function (Fig. 14) of the ArsenicNetPlus neural network had a fast gradient descent rate similar to curves of  $y = x^{-\frac{1}{t}}$ ,  $-\frac{1}{t} = C, x \in [0, \infty]$ . The accumulative confusion matrix of ArsenicNetPlus is shown in Table 4.



**Fig. 10** Accuracy of algorithms on cassava dataset using 7-fold cross-validation



**Fig. 11** Recall of algorithms on cassava dataset using 7-fold cross-validation

**Ablation experiment with pseudo high-frequency component**

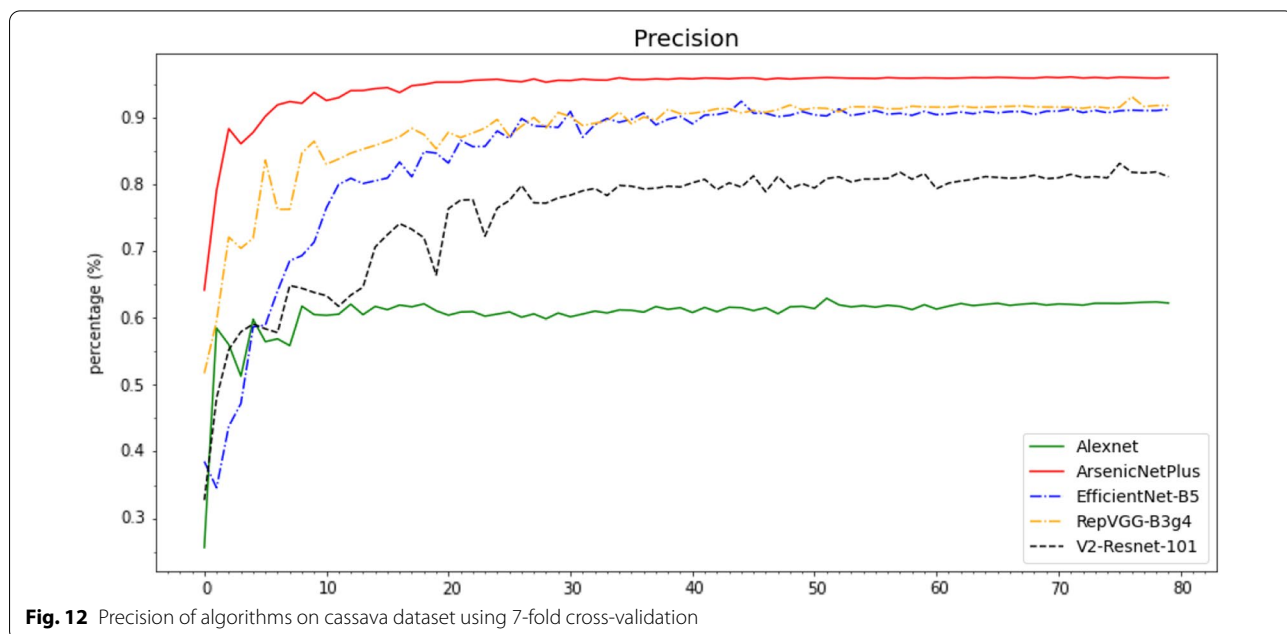
The best performance of ArsenicNetPlus and the ArsenicNet neural network on the cassava dataset using 7-fold cross-validation is shown in Table 5.

The comparison loss curves of ArsenicNet-3 and ArsenicNetPlus are shown in Fig. 15, and the accuracy

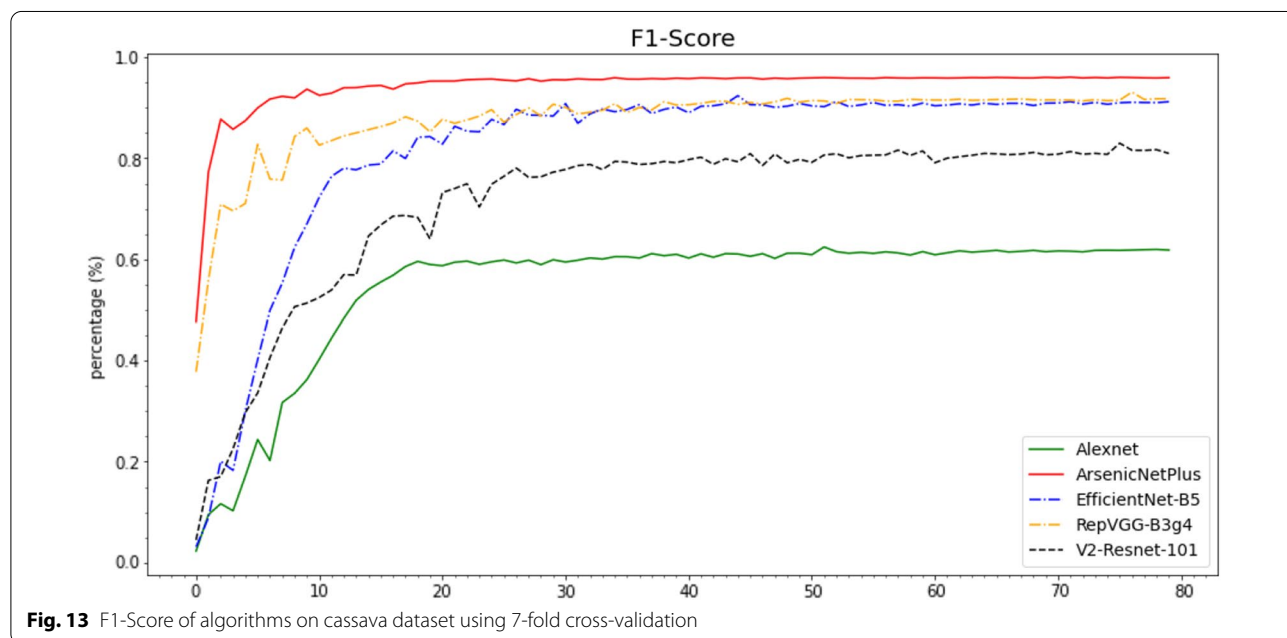
curves is shown in Fig. 16. The comparison curves of Recall, Precision, and F1-Score were similar to the Accuracy comparison curve (Fig. 16).

The results of ArsenicNet and ArsenicNetPlus were carried out in the same software environment, with the same training strategy and the same training hyperparameters.





**Fig. 12** Precision of algorithms on cassava dataset using 7-fold cross-validation

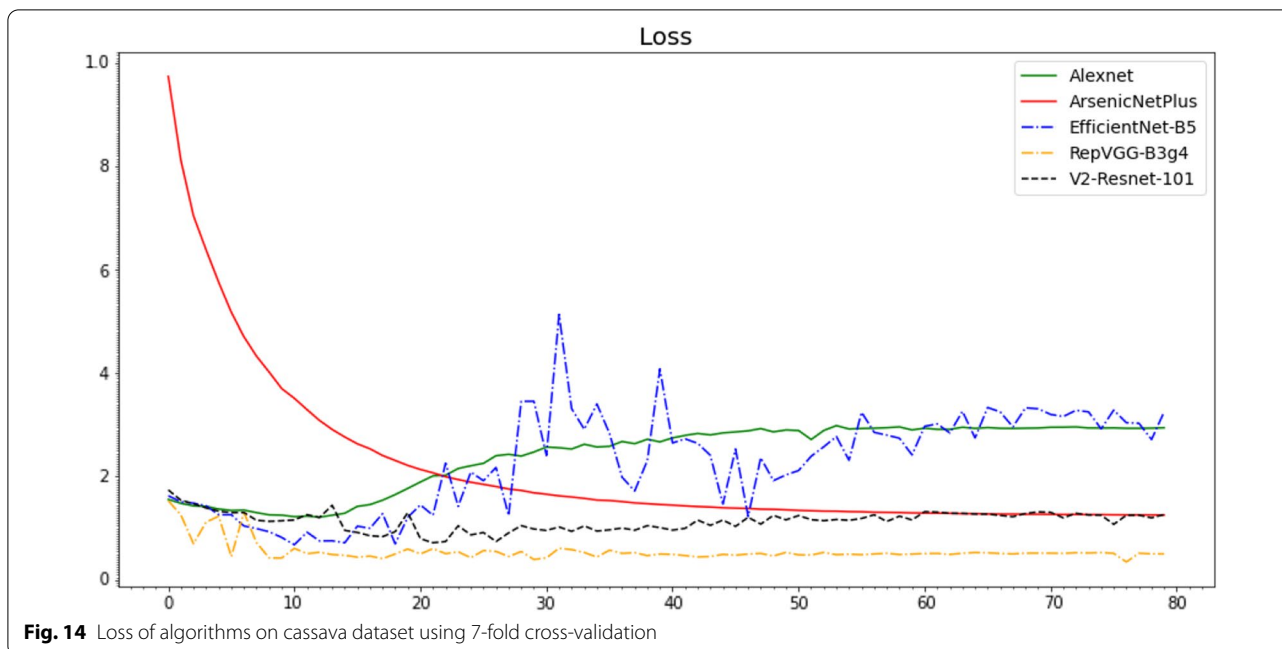


**Fig. 13** F1-Score of algorithms on cassava dataset using 7-fold cross-validation

**Benchmark dataset performance**

In the fine-grained research field, the Fine-Grained Visual Classification of Aircraft (FGVC-Aircraft) dataset [58] was a classical fine-grained categorization dataset. We used the FGVC-Aircraft dataset to evaluate the performance of our proposed algorithm and prove the effectiveness of our proposed fine-grained algorithm.

This evaluation was executed based on the manufacturer data format. To keep the image distribution balanced, a series of augmentation methods, including horizontal flipping, vertical flipping, horizontal vertical flipping, image offsetting, shift scaling and rotation, and Gaussian noise addition, were used to enlarge the number of images. As a result, the dataset contained 30



**Fig. 14** Loss of algorithms on cassava dataset using 7-fold cross-validation

categories, and each category contained 1467 images for training. The benchmark results are shown in Table 6.

The ArsenicNetPlus (based ResNet50) has achieved 86.59% in terms of accuracy, that is 7.79% higher than the experimental consequence of [51] in terms accuracy, and improve 1.89% in terms of accuracy than ArsenicNet-3.

**Comparison of existing methods for cassava disease detection**

ArsenicNetPlus was an end to end neural network algorithm. In comparison to other cassava leaf disease detection methods (Table 7), ArsenicNetPlus has not utilised transfer learning, ensemble learning or fine-tuning methods. The comparison of existing approaches for cassava disease detection was shown in Table 7.

**Discussion**

To verify the performance of the algorithm proposed in this paper, the other four algorithms are compared in Table 8. The

**Table 5** Comparison results for cassava dataset using 7-fold cross-validation

Indicator	ArsenicNet	ArsenicNetPlus
Accuracy	95.82%	96.27%
Recall	95.80%	96.25%
Precision	95.87%	96.30%
Loss	1.721	1.360
F1-score	95.83%	96.27%

proposed algorithm achieved the highest accuracy among the comparison algorithms used [20, 59]. As a comparison algorithm, the traditional machine learning methods used by Emuoyibofarhe et al. [66] had a weaker encoding performance in complex contexts than ArsenicNetPlus.

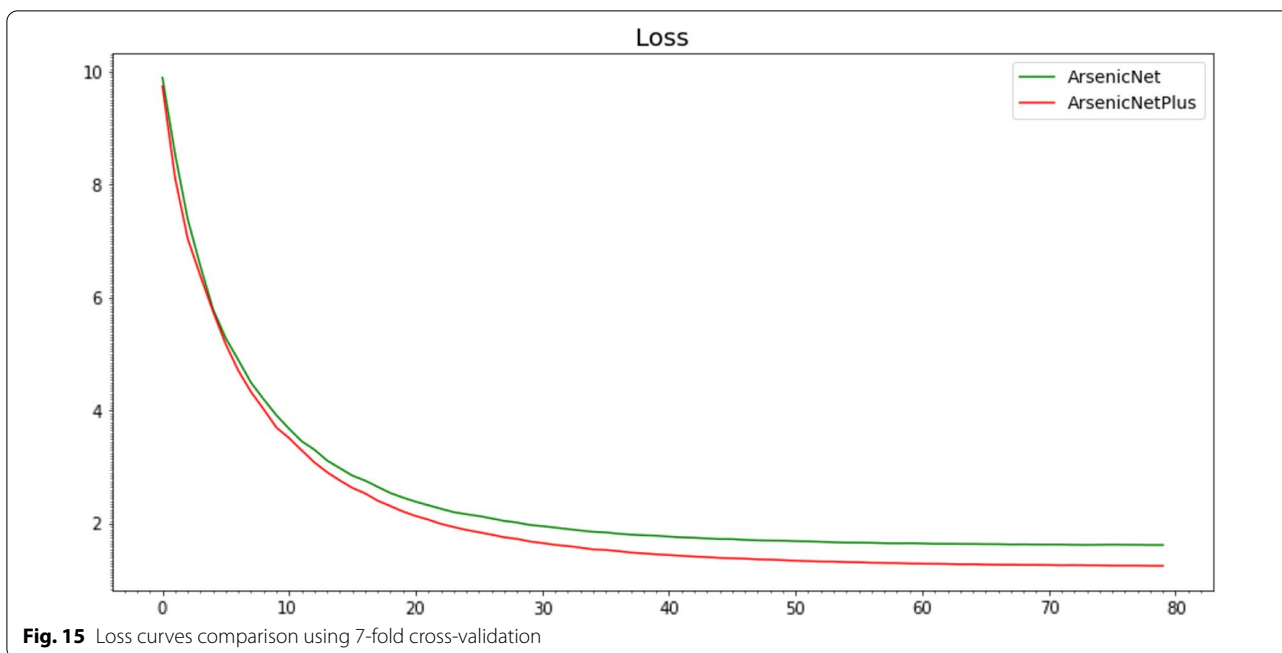
**Conclusion**

A signal frequency was continued to down-convert the convolutional neural network, and the objective function in signals was lost in extreme weakness. Thus, the pseudo high-frequency component can be utilized to approximate the destination function to boost the generalization performance.

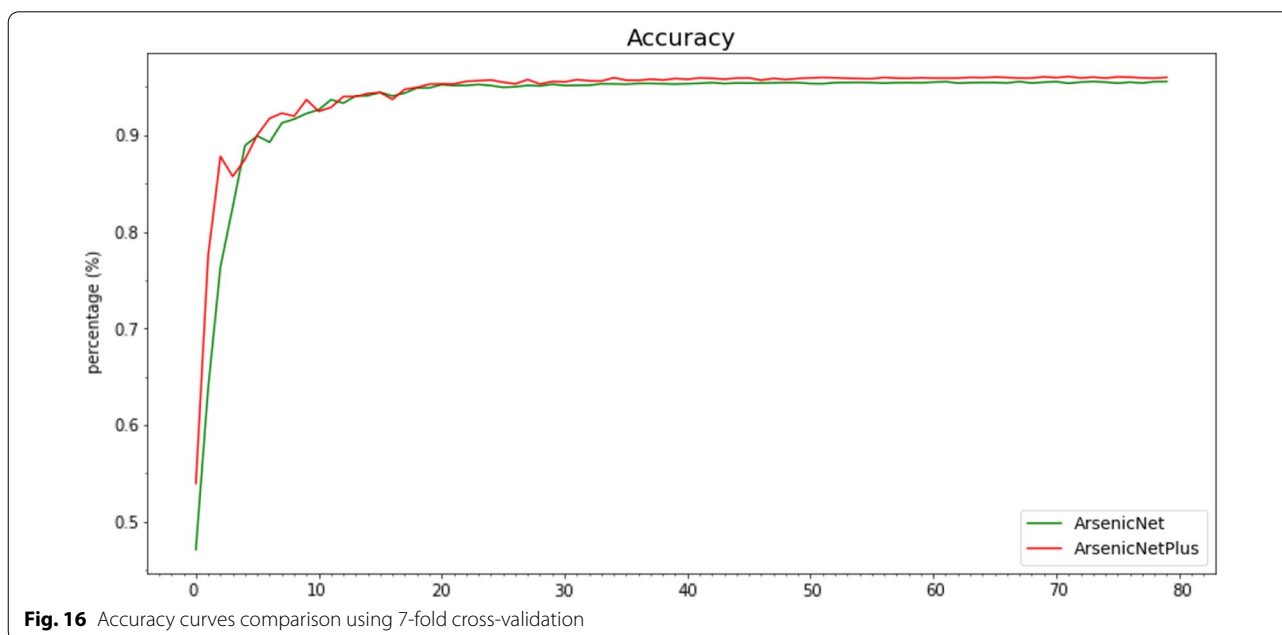
A clear difference can be found in the loss curves in Fig. 15, where the loss values for ArsenicNetPlus are lower than that for ArsenicNet. Correspondingly, the accuracy of the ArsenicNetPlus was higher than ArsenicNet in Fig. 16. The performance of ArsenicNetPlus on the FGVC-Aircraft dataset demonstrates (Table 6) that pseudo high-frequency can improve the generalisation ability of the neural network.

Consequently, the pseudo high-frequency component is useful in two ways:

- (1) The ability to maintain high frequency in feature maps is an important factor that impacts the generalization performance of the neural network.



**Fig. 15** Loss curves comparison using 7-fold cross-validation



**Fig. 16** Accuracy curves comparison using 7-fold cross-validation

(2) The pseudo high-frequency is an approximate approach to replenish the high-frequency in a weaker state of a convolutional neural network.

than that of RepVGG-B3g5. Thus, our next work will be devoted to modifying the loss function to make it converge faster, to boost the performance of the proposed neural network [67].

In contrast, the proposed method has a higher initial loss value and its loss function converges more slowly

**Table 6** Results of our proposed method

Method	Accuracy (%)	Pretraining
ResNet-50 [51]	78.8	Yes
ArsenicNet-1 (Based ResNet-50)	83.05	No
ArsenicNet-2 (Based ResNet-50)	83.68	No
ArsenicNet-3 (Based ResNet-50)	84.70	No
ArsenicNet-4 (Based ResNet-50)	84.13	No
ArsenicNetPlus (Based ResNet-50)	86.59	No

**Table 7** Comparison of existing approaches for cassava disease detection

Research	Methodology
Ramcharan et al. [20]	MobileNet
Sambasivam et al. [59]	custom 7-layer CNN
Ayu et al. [60]	MobileNetV2
Sangbamrung et al. [61]	Faster R-CNN and CNN
Oyewola et al. [40]	Deep residual CNN
Ramcharan et al. [62]	InceptionV3
Abayomi-Alli et al. [63]	MobileNetV2
Zhao et al. [64]	Residual learning with attention mechanism
Ravi et al. [65]	Ensemble learning, EfficientNets, and attention mechanism

**Table 8** Comparison of the results

Research	Accuracy (%)	Pretraining
ArsenicNetPlus	96.27	No
Ramcharan et al. [20]	93	Yes
Sambasivam et al. [59]	93	Yes
Emuoyibofarhe et al. [66]	72.75	No

### Appendix

A good kernel function can be calculated via Cesàro summation. The derivation is:

$$\begin{aligned}
 \sigma_n f(x) &= \frac{\frac{1}{2\pi} \sum_{-\pi}^{\pi} f(y)g_1(x-y) + \dots + \frac{1}{2\pi} \sum_{-\pi}^{\pi} f(y)g_n(x-y)}{N} \\
 &= \frac{\frac{1}{2\pi} \sum_{-\pi}^{\pi} \{f(y) \times [g_1(x-y) + \dots + g_n(x-y)]\}}{N} \\
 &= \frac{1}{2\pi} \sum_{-\pi}^{\pi} \left\{ f(y) \times \frac{[g_1(x-y) + \dots + g_n(x-y)]}{N} \right\}
 \end{aligned} \tag{6}$$

where the  $g$  function refers to  $\{g_n\}_{n=1}^{\infty}$ , a series of kernel functions;  $f$  refers to a function of period  $2\pi$ ;  $n$  refers to the number of kernel functions in the series; and  $\sigma_n f(x)$  refers to the convolution result. In the complex field, the addition of the two functions can be expressed as follows:

$$F\{f_1(t)\} = F_1(jn\omega_1), F\{f_2(t)\} = F_2(jn\omega_2), \tag{7}$$

$$f(t) = f_1(t) + f_2(t), \tag{8}$$

$$F(jn\omega) = F_1(jn\omega_1) + F_2(jn\omega_2). \tag{9}$$

The function  $F(jn\omega)$  was a result of replenishing the Fourier series coefficients.

The Cesàro sum operation was a special method to convert the non-good kernel functions to the good kernel functions. The good kernel function was a property in mathematics and physics, a manifestation of a form of function. However, the Cesàro sum operation is still an idea in convolutional neural networks.

### Abbreviations

UAV: Unmanned aerial vehicle; CNN: Convolutional neural network; t-SNE: t-Distributed stochastic neighbor embedding; SGD: Stochastic gradient descent; CBB disease: Cassava bacterial blight disease; CBSD disease: Cassava brown streak disease; CGM disease: Cassava green mottle disease; CMD disease: Cassava mosaic disease; IBN: Instance batch normalization; SVM: Support vector machine; GFLOPs: Giga floating point operations [67]; FGVC-Aircraft: Fine-grained visual classification of aircraft.

### Acknowledgements

We appreciated Kunjie Chen professor for guide the manuscript.

### Author contributions

JZ: conceptualization, methodology, software, investigation, formal analysis, writing—original draft; CQ: data curation, writing—original draft; PM: data curation, visualization, polishing the writing; YZ: resources, supervision; ZB: software, validation; HL: visualization, writing—review and editing; KC: conceptualization, funding acquisition, resources, supervision, writing—review and editing. All authors read and approved the final manuscript.

### Authors' information

Jiayu Zhang was born in Xuzhou, Jiangsu, China in 1993. He received the M.Sc. degree in software engineering from the Hangzhou Normal University, in 2019. He is currently pursuing the Ph.D. degree in agricultural electrification and automation at Nanjing Agricultural University. His research interests include machine vision, deep learning and digital image processing.

Chao Qi is a doctoral student at Nanjing Agricultural University, majoring in agricultural electrification and automation, with research interests in image processing and machine learning, especially for deep learning techniques. He received MS degree in 2019 at Nanjing Agricultural University, with his research interests in image processing and machine learning, focusing mainly on digital image processing techniques.

Peter Mecha is a doctoral student at Nanjing Agricultural University, China. He also works as an assistant lecturer at Egerton University, Kenya. He is currently involved in designing heat pump drying processes for various vegetables like mushrooms and day lily flowers among others. Additionally, he is working on applications of deep learning and image processing in agricultural processing to help solve food security problems.

Yi Zuo is a doctoral student at Nanjing Agricultural University, majoring in agricultural electrification and automation, with research interests in image processing and machine learning, especially for deep learning techniques. He received MS degree in 2021 at Nanjing Agricultural University, with his research interests in image processing and machine learning, focusing mainly on digital image processing techniques.

Zongyou Ben is a doctoral student at Nanjing Agricultural University, majoring in agricultural mechanization engineering, with research interests in deep learning, digital image processing. He received MS degree in 2016 at Nanjing Tech University, with his research interests in fluid machinery design.



Haolu Liu an assistant researcher at Nanjing Institute of Agricultural Mechanization, Ministry of Agriculture and Rural Affairs, is mainly engaged in the field of multi-source information fusion and intelligent chassis of agricultural machinery. He received his M.Eng. degree from Nanjing Agricultural University in 2019 and is currently pursuing his Ph.D. degree in Agricultural Mechanization Engineering at the Chinese Academy of Agricultural Sciences. Kunjie Chen is a postdoctoral fellow at The University of Reading, UK. He was awarded a grant from the China Scholarship Council to do visiting research at the School of Biology and Food, University of Reading, UK. He is currently the executive director of the China Livestock Processing Research Association and the director of the Livestock Processing Engineering Committee. He is mainly engaged in the research of processing and quality testing of agricultural and livestock products.

#### Funding

Not applicable.

#### Availability of data and materials

The origin dataset can be found in <https://www.kaggle.com/competitions/cassava-leaf-disease-classification/data>. There are 21,394 images in original cassava leaf disease dataset. The original dataset was annotated by experts at the Uganda National Crops Resources Research Institute (NaCRRI) in collaboration with the AI lab at Makerere University, Kampala. The cassava dataset used in this study can be found at the following link: [https://pan.baidu.com/s/1thrlr\\_OuB3gzYSPT317gtg](https://pan.baidu.com/s/1thrlr_OuB3gzYSPT317gtg) (Password: abcd).

#### Declarations

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup>College of Engineering, Nanjing Agricultural University, Nanjing, China. <sup>2</sup>Nanjing Institute of Agricultural Mechanization, Ministry of Agriculture and Rural Affairs, Nanjing, China.

Received: 6 August 2022 Accepted: 4 December 2022

Published online: 14 December 2022

#### References

- Chisenga SM, Workneh TS, Bultosa G, Alimi BA. Progress in research and applications of cassava flour and starch: a review. *J Food Sci Technol*. 2019;56(6):2799–813.
- Zhang L, Zhang J, Wei Y, Hu W, Liu G, Zeng H, Shi H. Microbiome-wide association studies reveal correlations between the structure and metabolism of the rhizosphere microbiome and disease resistance in cassava. *Plant Biotechnol J*. 2021;19(4):689–701.
- Legg J, Kumar LTM, Tripathi L, Ferguson M, Kanju E, Ntawuruhunga P, Cuellar W. Cassava virus diseases: biology, epidemiology, and management. *Adv Virus Res*. 2015. <https://doi.org/10.1016/bs.aivir.2014.10.001>.
- Burns A, Gleadow R, Cliff J, Zacarias A, Cavagnaro T. Cassava: the drought, war and famine crop in a changing world. *Sustainability*. 2010;2(11):3572–3607.
- Wydra K, Verdier V. Occurrence of cassava diseases in relation to environmental, agronomic and plant characteristics. *Agric Ecosyst Environ*. 2002;93(1–3):211–26.
- Van Der Maaten L. Accelerating t-sne using tree-based algorithms. *J Mach Learn Res*. 2014;15(1):3221–45.
- Behera A, Wharton Z, Hewage PR, Bera A. Context-aware attentional pooling (cap) for fine-grained visual classification. In: Proceedings of the AAAI conference on artificial intelligence, vol. 35; 2021. p. 929–37.
- Tan M, Le Q. Efficientnet: rethinking model scaling for convolutional neural networks. In: International conference on machine learning, PMLR; 2019. p. 6105–14.
- Ai Y, Sun C, Tie J, Cai X. Research on recognition model of crop diseases and insect pests based on deep learning in harsh environments. *IEEE Access*. 2020;8:171686–93.
- Plested J, Shen X, Gedeon T. Rethinking binary hyperparameters for deep transfer learning. In: International conference on neural information processing. Springer; 2021. p. 463–75.
- Fu J, Zheng H, Mei T. Look closer to see better: recurrent attention convolutional neural network for fine-grained image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2017. p. 4438–46.
- Ye M, Shen J, Lin G, Xiang T, Shao L, Hoi SC. Deep learning for person re-identification: a survey and outlook. *IEEE Trans Pattern Anal Mach Intell*. 2021;44(6):2872–93.
- Li W, Zhu X, Gong S. Harmonious attention network for person re-identification. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2018. p. 2285–94.
- Weiss K, Khoshgoftar TM, Wang D. A survey of transfer learning. *J Big Data*. 2016;3(1):1–40.
- Sagi O, Rokach L. Ensemble learning: a survey. *Wiley Interdiscip Rev Data Min Knowl Discov*. 2018;8(4):1249.
- Lin Y-K, Wang C-F, Chang C-Y, Sun H-L. An efficient framework for counting pedestrians crossing a line using low-cost devices: the benefits of distilling the knowledge in a neural network. *Multimed Tools Appl*. 2021;80(3):4037–51.
- Chang D, Ding Y, Xie J, Bhunia AK, Li X, Ma Z, Wu M, Guo J, Song Y-Z. The devil is in the channels: mutual-channel loss for fine-grained image classification. *IEEE Trans Image Process*. 2020;29:4683–95.
- Tetila EC, Machado BB, Menezes GK, Oliveira AdS, Alvarez M, Amorim WP, Belete NADS, Da Silva GG, Pistori H. Automatic recognition of soybean leaf diseases using uav images and deep convolutional neural networks. *IEEE Geosci Remote Sens Lett*. 2019;17(5):903–7.
- Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto M, Adam H. Mobilenets: efficient convolutional neural networks for mobile vision applications. arXiv. 2017. <http://arxiv.org/abs/1704.04861>.
- Ramcharan A, McCloskey P, Baranowski K, Mbilinyi N, Mrisho L, Ndalahwa M, Legg J, Hughes DP. A mobile-based deep learning model for cassava disease diagnosis. *Front Plant Sci*. 2019;10:272.
- Singh UP, Chouhan SS, Jain S, Jain S. Multilayer convolution neural network for the classification of mango leaves infected by anthracnose disease. *IEEE Access*. 2019;7:43721–9.
- Li X, Chen S, Hu X, Yang J. Understanding the disharmony between dropout and batch normalization by variance shift. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition; 2019. p. 2682–90.
- Liang S, Huang Z, Liang M, Yang H. Instance enhancement batch normalization: an adaptive regulator of batch noise. In: Proceedings of the AAAI conference on artificial intelligence, vol. 34; 2020. p. 4819–27.
- Gao S-H, Han Q, Li D, Cheng M-M, Peng P. Representative batch normalization with feature calibration. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition; 2021. p. 8669–79.
- Yao Z, Cao Y, Zheng S, Huang G, Lin S. Cross-iteration batch normalization. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition; 2021. p. 12331–40.
- Benz P, Zhang C, Karjauv A, Kweon IS. Revisiting batch normalization for improving corruption robustness. In: Proceedings of the IEEE/CVF winter conference on applications of computer vision; 2021. p. 494–503.
- Awais M, Iqbal MTB, Bae S-H. Revisiting internal covariate shift for batch normalization. *IEEE Trans Neural Netw Learn Syst*. 2020;32(11):5082–92.
- Yuan Y, Xu Z, Lu G. Spedccnn: spatial pyramid-oriented encoder-decoder cascade convolution neural network for crop disease leaf segmentation. *IEEE Access*. 2021;9:14849–66.
- Zhang S, Zhang S, Zhang C, Wang X, Shi Y. Cucumber leaf disease identification with global pooling dilated convolutional neural network. *Comput Electron Agric*. 2019;162:422–30.
- Khan A, Sohail A, Zahoora U, Qureshi AS. A survey of the recent architectures of deep convolutional neural networks. *Artif Intell Rev*. 2020;53(8):5455–516.

31. Han D, Kim J, Kim J. Deep pyramidal residual networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2017. p. 5927–35.
32. Reyes AK, Caicedo JC, Camargo JE. Fine-tuning deep convolutional networks for plant recognition. *CLEF*. 2015;1391:467–75.
33. Lee SH, Chan CS, Mayo SJ, Remagnino P. How deep learning extracts and learns leaf features for plant classification. *Pattern Recogn*. 2017;71:1–13.
34. Thai H-T, Tran-Van N-Y, Le K-H. Artificial cognition for early leaf disease detection using vision transformers. In: 2021 international conference on advanced technologies for communications (ATC), IEEE; 2021. p. 33–8.
35. Liu Z, Lin Y, Cao Y, Hu H, Wei Y, Zhang Z, Lin S, Guo B. Swin transformer: Hierarchical vision transformer using shifted windows. In: Proceedings of the IEEE/CVF international conference on computer vision; 2021. p. 10012–22.
36. De Luna RG, Dadios EP, Bandala AA. Automated image capturing system for deep learning-based tomato plant leaf disease detection and recognition. In: TENCON 2018—2018 IEEE region 10 conference, IEEE; 2018. p. 1414–9.
37. Zhang Y, Song C, Zhang D. Deep learning-based object detection improvement for tomato disease. *IEEE Access*. 2020;8:56607–14.
38. Xiaoxiao S, Shaomin M, Yongyu X, Zhihao C, Tingting S. Image recognition of tea leaf diseases based on convolutional neural network. In: 2018 international conference on security, pattern analysis, and cybernetics (SPAC), IEEE; 2018. p. 304–9.
39. Zhou C, Zhou S, Xing J, Song J. Tomato leaf disease identification by restructured deep residual dense network. *IEEE Access*. 2021;9:28822–31.
40. Oyewola DO, Dada EG, Misra S, Damaševičius R. Detecting cassava mosaic disease using a deep residual convolutional neural network with distinct block processing. *PeerJ Comput Sci*. 2021;7:352.
41. Maass W. Networks of spiking neurons: the third generation of neural network models. *Neural Netw*. 1997;10(9):1659–71.
42. Schwartz M. Harvard University Online Education, lecture 8: fourier transforms. <https://scholar.harvard.edu/files/schwartz/files/lecture8-fouriertransforms.pdf>. Accessed 5 Apr 2022.
43. Wang H, Wu X, Huang Z, Xing EP. High-frequency component helps explain the generalization of convolutional neural networks. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition; 2020. p. 8684–94.
44. Boulent J, Foucher S, Théau J, St-Charles P-L. Convolutional neural networks for the automatic identification of plant diseases. *Front Plant Sci*. 2019;10:941.
45. He K, Zhang X, Ren S, Sun J. Identity mappings in deep residual networks. In: European conference on computer vision. Springer; 2016. p. 630–45.
46. Hu J, Shen L, Sun G. Squeeze-and-excitation networks 2018; p. 7132–41
47. Qin Z, Zhang P, Wu F, Li X. Fcanet: frequency channel attention networks. *arXiv*. 2020. <http://arxiv.org/abs/2012.11879>.
48. Pan X, Luo P, Shi J, Tang X. Two at once: enhancing learning and generalization capacities via ibn-net 2018; p. 464–79.
49. Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. In: International conference on machine learning, PMLR; 2015. p. 448–56.
50. With Code team P. paperswithcode. 1 November 2022. <https://paperswithcode.com/dataset/fgvc-aircraft-1>.
51. Lee J, Won T, Lee TK, Lee H, Gu G, Hong K. Compounding the performance improvements of assembled techniques in a convolutional neural network. *arXiv*. 2020. <http://arxiv.org/abs/2001.06268>.
52. Sambasivan N, Kapania S, Highfill H, Akrong D, Paritosh P, Aroyo LM. Everyone wants to do the model work, not the data work: data cascades in high-stakes ai. In: Proceedings of the 2021 CHI conference on human factors in computing systems; 2021. p. 1–15.
53. Azeroual O. Data wrangling in database systems: purging of dirty data. *Data*. 2020;5(2):50.
54. Zhou P, Feng J, Ma C, Xiong C, Hoi SCH, et al. Towards theoretically understanding why sgd generalizes better than adam in deep learning. *Adv Neural Inf Process Syst*. 2020;33:21285–96.
55. Ding X, Zhang X, Ma N, Han J, Ding G, Sun J. Repvgg: making vgg-style convnets great again. *arXiv*. 2021. <http://arxiv.org/abs/2101.03697>.
56. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Adv Neural Inf Process Syst*. 2012;25:1097–105.
57. Ferentinos KP. Deep learning models for plant disease detection and diagnosis. *Comput Electron Agric*. 2018;145:311–8.
58. Maji S, Kannala J, Rahtu E, Blaschko M, Vedaldi A. Fine-grained visual classification of aircraft. Technical report. 2013. <http://arxiv.org/abs/1306.5151>.
59. Sambasivam G, Opiyo GD. A predictive machine learning application in agriculture: cassava disease detection and classification with imbalanced dataset using convolutional neural networks. *Egypt Inform J*. 2021;22(1):27–34.
60. Ayu H, Surtono A, Apriyanto D. Deep learning for detection cassava leaf disease. *J Phys Conf Ser*. 2021;1751:012072.
61. Sangbamrung I, Praneetpholkkrang P, Kanjanawattana S. A novel automatic method for cassava disease classification using deep learning. *J Adv Inform Technol*. 2020;11(4):241–248. <https://doi.org/10.12720/jait.11.4.241-248>.
62. Ramcharan A, Baranowski K, McCloskey P, Ahmed B, Legg J, Hughes DP. Deep learning for image-based cassava disease detection. *Front Plant Sci*. 2017;8:1852.
63. Abayomi-Alli OO, Damaševičius R, Misra S, Maskeliūnas R. Cassava disease recognition from low-quality images using enhanced data augmentation model and deep learning. *Expert Syst*. 2021;38(7):12746.
64. Lu J, Hu J, Zhao G, Mei F, Zhang C. An in-field automatic wheat disease diagnosis system. *Comput Electron Agric*. 2017;142:369–79.
65. Ravi V, Acharya V, Pham TD. Attention deep learning-based large-scale learning classifier for cassava leaf disease classification. *Expert Syst*. 2022;39(2):12862.
66. Emuoyibofarhe O, Emuoyibofarhe JO, Adebayo S, Ayandiji A, Demeji O, James O. Detection and classification of cassava diseases using machine learning. *Int J Comput Sci Softw Eng*. 2019;8(7):166–76.
67. Molchanov P, Tyree S, Karras T, Aila T, Kautz J. Pruning convolutional neural networks for resource efficient inference. *arXiv e-prints*. 2016;1611–06440. <http://arxiv.org/abs/1611.06440>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

