PAPER • OPEN ACCESS

Lightweight deep CNN models for identifying drought stressed plant

To cite this article: MH. Kamarudin and Zool H. Ismail 2022 *IOP Conf. Ser.: Earth Environ. Sci.* **1091** 012043

Lightweight deep CNN models for identifying drought stressed plant

MH. Kamarudin*[1], Zool H. Ismail[1]

[1] Malaysia-Japan International Institute of Technology, Universiti Teknologi Malaysia, Jalan Sultan Yahya Petra, Kuala Lumpur 54100, Malaysia

*Corresponding author e-mail: hider@graduate.utm.my

Abstract. Drought is one of the most severe climatological disasters that has negative impact on agricultural production around the world. Over the years, computer vision technology has been used in conjunction with machine learning applications to replace traditional destructive and time-consuming methods for real-time monitoring of drought-affected plant. Deep learning (DL) techniques have gained a stellar reputation in image classification recently, with convolutional neural network (CNN) emerging as the industry standard. However, the size of deep CNN models is frequently large due to massive number of parameters and field application is often not feasible due to limited storage and computational resources. Several lightweight CNN models have been selected based on the number of network parameters of less than 6M and were trained and tested. The EfficientNet model has achieved a classification accuracy of 88.12 and 88.97 percent for identifying severe drought, mild drought, and no drought plants on visible and near-infrared images respectively. The findings of this study can be used to assist in the development of automated early detection of drought stressed plant with model sizes suitable for real-time plant diagnosis on mobile or embedded devices.

1. Introduction

Droughts are becoming more common around the world as a result of global climate change. Drought conditions that last for an extended period can have a detrimental impact on agriculture, prompting a red alert on food security. As irregular rainfall limits water availability, most plant crops suffered from water stress, the most critical of plant abiotic stresses that can affect plant growth. An early detection system for water stressed plants is crucial for ensuring sustainable agricultural productivity. Traditionally, visual characteristics such as leaf yellowing and wilting have been manually monitored as indicators of drought conditions. However, the technique is inefficient and unsuitable for early identification of drought-stressed plants. More capable techniques, such as quantifying leaf and stem water potential at midday [1], have been introduced; however, the methods are time consuming and destructive. Recently, the development of plant sensing equipment utilising remote sensing to assess environmental and plant physiological changes in a fast and nod-destructive manner has seen a significant leap in progress [2].

Computer vision has been widely adopted as a low-cost system for identifying drought-affected plants [3]. The Internet of Things (IoTs) and robotics platforms have been used to create an automated system based on image analysis that extracts information such as morphological features (size, shape, texture), spectra (colour, temperature, humidity), and time data (growth rate, development, dynamic change of spectral and morphological modes) that show strong correlations with the plant water status [4]. Plant phenotyping using computer vision has also been implemented in drought resistance plant breeding studies [5]. Despite this, the technology is still in research and development stage and has seen little use in practise due to operational and processing difficulties. The high-throughput data of plant images necessitates fast processing mechanism to determine the useful information that reflects the plant's water stress status.

Machine learning (ML) has quickly become the standard approach in computer vision applications due to its ability to effectively process large amounts of information in a non-linear framework. Gutiérrez, Diago, Fernández-Novales and Tardaguila [6] have used machine learning models in thermal imaging analysis to assess water stress status in vineyards. Zhuang, Wang, Jiang, Li and Gong [7] have proposed a maize water stress detection model based on a supervised learning algorithm of Gradient Boosting Decision Tree (GBDT) using features extracted from plant colour images. However, conventional machine learning techniques require manual extraction of image features, which are then fed into the ML model to identify drought stressed plant using pattern recognition. This requires specific technical knowledge and is currently the limiting factor influencing the model accuracy [8]. When compared to traditional machine learning, the advanced ML technique of deep learning (DL) provides the benefits of automated feature extraction and analysis with better performance. Convolutional Neural Networks (CNNs) have been developed for quick application and can provide a much-detailed representation of the image features.

Many studies have used deep CNN models to detect drought affected plants, with such applications being solved either by finely tuned CNNs or by training from scratch. Chandel, Chakraborty, Rajwade, Dubey, Tiwari and Jat [9] tested three different pre-trained deep CNNs, AlexNet, GoogLeNet, and Inception V3 to identify drought stress in maize, okra, and soybean in which GoogLeNet performance was found to be superior compared to the other models. Zhuang, Wang, Jiang and Li [10] used CNN to identify drought stress in maize by learning features of the leaf phenotype. In their most recent study, Islam and Yamane [11] created HortNet417v1, a deep CNN model for detecting drought stress in pot-grown peach plants. Aside from exceptional performance, deep model architectures typically include large network parameters, which necessitates large storage capacity devices. Deep learning models are also extremely dependable on large amounts of data and involve significant computational resources, often supplied by the graphics processing units (GPUs). The application of such technology in the agricultural field may result in significant hardware costs. Cloud computing may be the best alternative option, but it requires a consistent internet connection, which is simply not available in some croplands.

Lightweight deep learning model structures for image classification have been developed over the years for small and mobile devices, including edge computing applications. These models have comparatively simpler and more efficient network constructions with fewer parameters, consuming less memory as well as computing power. The objective of this research is to investigate the ability of several lightweight deep CNN models, namely MobileNet, Mobile NasNet, and EfficientNet, to identify drought-affected plants. The goal is to create a real-time embedded image processing and plant diagnosis system for water stress detection using limited computer hardware.

2. Materials and devices

2.1. Plant images dataset

Plant images were obtained from the Donald Danforth Plant Science Centre's standard publicly available dataset [12]. The dataset includes images of plant shoot area of ten Setaria grass lines (*S. viridis* (accession A10), *S. italica* (accession B100), and eight randomly selected RILs derived from a cross of *S. viridis* and *S. italica*). Setaria grass is a model plant that has been used in many drought-related studies to analyse the plant phenotypes [13].

Four water treatments were performed on full-water capacity (FC): 100% FC, 66% FC, 33% FC, and 0% FC imposed 17 days after planting (DAP) and maintained for another 17 days. Images prior to 17 DAP were excluded from the sample due to low to none biomass production. Plants that received no water after 17 DAP died within 7 days thus 0% FC images were also omitted from the sample. Because early treatment has no discernible effect on the plants, the sample images were taken after 2 days of treatment, from 21 DAP until 33 DAP. One top view and four side view images were taken for each plant. We combined all four-sided images and omitted the top view images in the sample. These images were sorted by drought severity of severe drought, mild drought, and no drought.

The dataset was created using visible (VIS) and near-infrared (NIR) images, two types of images commonly used in plant water stress analysis studies. Visible image of red, green, and blue (RGB), is the most widely used imagery for measuring plant's morphological properties collected using optical device that detects light wavelengths between 400 and 700 nm. The rate of NIR absorption is the highest in the spectral range of 1400 to 1450 nm and is highly correlated with plant moisture content. We used both images to determine which image provided the best accuracy for identifying drought-stressed plants using lightweight CNN models. As NIR image is essentially a grayscale image, it should be faster to compute than an RGB image. The RGB image resolution is 2,454 x 2,056 pixels, while the NIR image resolution is 150 x 200 pixels. Number of images in each sample class are shown in Table 1.

Table 1. Number of images in each sample class

Drought severity	Number of images	
	RGB	NIR
Severe drought	2,281	2,280
Mild drought	2,286	2,285
No drought	2,382	2,300

2.2. Image pre-processing

Image cropping was performed as a pre-processing step to remove any boxes and pot pixels. This is done to eliminate the unnecessary background and focus only on the plant features, as shown in figure 1 and 2. The cropping process also reduces the size of the images, which speeds up the computation. Images for this study were taken in a laboratory setting with a white background and adequate lighting. Because the background is white, no plant segmentation is required. We trained with a single image of a plant shoot because we believe that the plant structure would provide morphological information that aids in detecting drought-stressed plants. The images were downscaled to fit the size specification of various models, and the intensity values were adjusted to fill the entire intensity range [0, 255].

(a) **(b)** **(c)**

Figure 1. RGB plant images (a) severe drought, (b) mild drought (c) no drought

(a) **(b)** **(c)**

Figure 2. NIR plant images (a) severe drought, (b) mild drought (c) no drought

2.3. Lightweight deep CNNs

The architectures of different lightweight CNN models can be defined by the number of convolutional layers, activation function of each layer, and the hidden units of each layer. Predefined models of MobileNet (versions 1 and 2), MnasNet and EfficientNet were evaluated for their accuracy in identifying drought stressed plant, as shown in Table 2. These models were chosen based on a small number of parameters (fewer than 6 million) to reduce computational complexity and memory cost.

Table 2. Lightweight CNN models evaluated in the study

Model	Parameters	Size (MB)	Depth	Time (ms) per inference step	
				CPU	GPU
MobileNet	4,253,864	16	88	22.6	3.44
MobileNetV2	3,538,984	14	88	25.9	3.83
NasNet mobile	5,326,716	23	-	27.04	6.70
EfficientNet	5,330,571	29	-	46.0	4.91

2.3.1. MobileNet

MobileNet [14] is a lightweight architecture optimized for mobile and embedded vision systems. It divides the convolution into a depth-wise separable convolution followed by a pointwise convolution to construct a lightweight deep CNN model. The parameters are much lower than those of established models such as VGG [15], but the accuracy is comparable when trained on well-known public datasets.

2.3.2. MobileNetV2

MobileNetV2 [16] is an updated version of MobileNet that improves its efficiency and effectiveness in terms of accuracy and speed. MobileNetV2 enhances the state-of-the-art performance of mobile models on a variety of tasks and benchmarks, as well as across a range of model sizes. In contrast to traditional residual models, the architecture is based on an inverted residual structure, with the input and output of the residual block being thin bottleneck layers.

2.3.3. MnasNet

Mobile NasNet or MnasNet [17] is a CNN architecture designed for mobile devices with limited computing power. Based on the Neural Architecture Search (NAS) approach, lightweight deep model is automatically designed and optimized using reinforcement learning [18]. It outperforms MobileNetV2 on ImageNet classification and COCO detection tasks, performing 1.5 times faster. The model's computational cost and parameters can be easily scaled to address a wide range of problems.

2.3.4. EfficientNet

EfficentNet's architecture [19] employs mobile inverted bottleneck convolution (MBConv), which is similar to MobileNetV2 and MnasNet but slightly larger. EfficientNet used compound scaling to create a series of EfficienNets that were both more accurate and larger in size. Only the baseline architecture was used in this study.

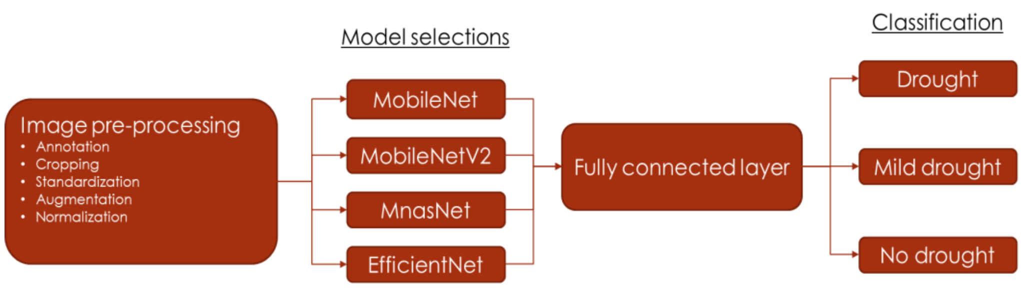

Figure 3. Schematic representation of the flow of work

2.4. Training and validation

Seventy percent of the sample images were used for training, twenty two percent for cross-validation, and ten percent for testing. To avoid overfitting, validation data was used to tune network parameters and hyperparameters. To obtain a generalized measure of classification accuracy, unseen test data was used. Flipping, adjusting aspect ratio, and intensity transformation were all part of the augmentation process as in figure 3. To evaluate training and testing accuracies, Adam optimizer was used. The learning rate was fixed to 0.001, and the models were built using the Tensorflow [13] framework in python platform. To reduce the risk of model overfitting, the callbacks function was used to perform early stopping. Fine tune layer starts at layer 134 for MobileNets (versions 1 and 2), layer 50 for NasNet, and layer 100 for EfficientNet.

3. Results and discussion

Table 3 shows that the deep CNN models were able to efficiently classify plants that had been subjected to three different levels of drought stress with respect to the accuracy and their training epochs. With 0.6875, MobileNet version 1 has the lowest accuracy. For RGB images, MobileNetV2 has an accuracy of 0.7687 and MnasNet has an accuracy of 0.8187. For RGB images, EfficientNet has the highest accuracy of 0.8812. EfficientNet also provides the best results for NIR images, with 0.8897 accuracy, followed by MnasNet model with 0.7364 accuracy, MobileNetV2 with 0.7396 accuracy, and MobileNetV1 with 0.7083 accuracy.

Table 3. Accuracy results on test dataset

Model	**Test accuracy**		**Epochs**	
	RGB	NIR	RGB	NIR
MobileNet	0.6875	0.7083	46	74
MobileNet-V2	0.7687	0.7396	67	87
NasNet mobile	0.8187	0.7364	51	74
EfficientNet	**0.8812**	**0.8897**	103	75

Overall, EfficientNet produced the best results for identifying drought-stressed plants. This could be due to EfficientNet's deeper structure being the largest in comparison to the other models. MnasNet, which had comparable parameters and size to EfficientNet, also performed well. The two MobileNet versions had lower accuracy, which could be attributed to smaller parameters and sizes. According to table 1, EfficientNet can be used in mobile devices to detect drought-stressed plants. When more data is available, the model's ability to scale up is also advantageous. In addition, when compared to other models, the results showed that EfficientNet transferred well to other datasets. Although an upscaled model can improve accuracy, the trade-off between accuracy and computational cost can be overcome by introducing an appropriate upscale coefficients multiplier.

Although the use of NIR images may simplify the process of detecting drought stressed plants through fast model convergence as referred to epochs in table 2, the results show that VIS images perform better when deep CNN models are used. This is to be expected, as colour is one of the dominant features retrieved by the CNN alongside shape and texture for drought-stressed plant identification [10]. This demonstrates that high accuracy can be obtained without the use of specialised cameras to detect drought-stressed plants. Nonetheless, both VIS and NIR systems could be used concurrently to improve detection accuracy. It should be noted that the pre-trained models are mostly trained on the publicly available ImageNet dataset, which contains colour images. This may result in a slight bias toward using colour images over grayscale images. To avoid this issue, it is recommended that the CNN model be trained from scratch using the NIR image rather than fine-tuning from the existing models.

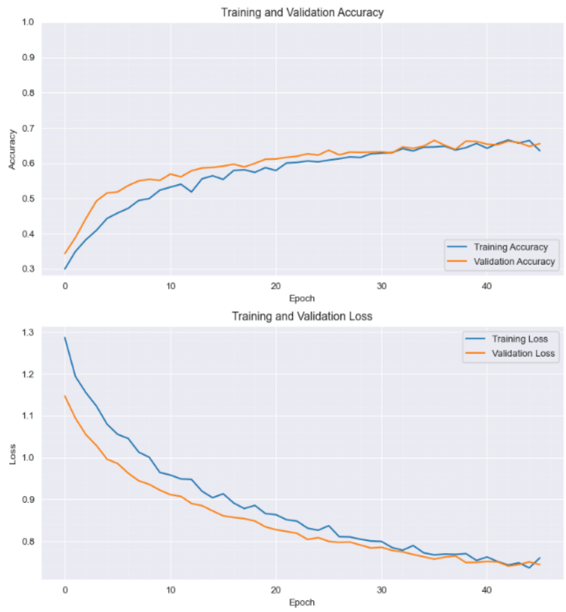

Figure 4. MobileNet training and validation accuracy and loss versus epochs for RGB image.

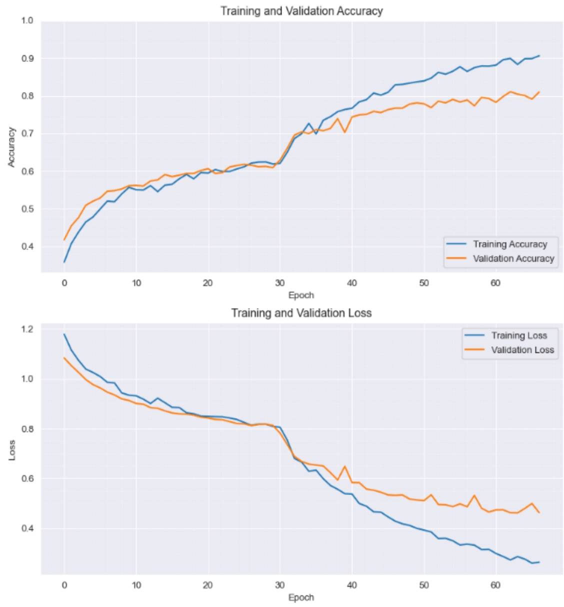

Figure 5. MobileNet-V2 training and validation accuracy and loss versus epochs for RGB image.

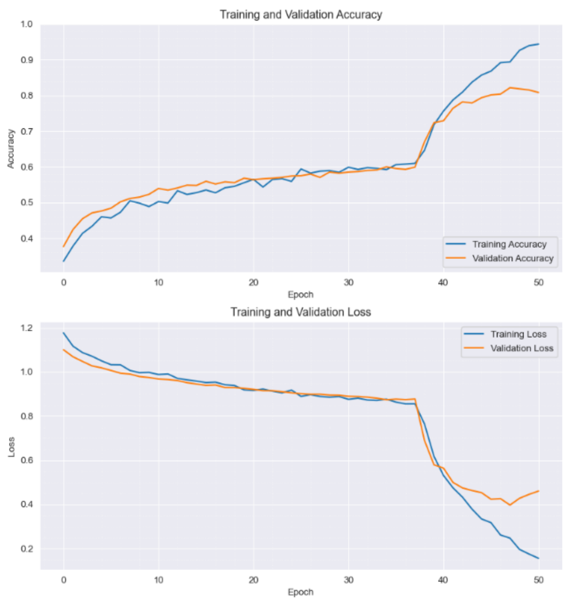

Figure 6. NasNet mobile training and validation accuracy and loss versus epochs for RGB image.

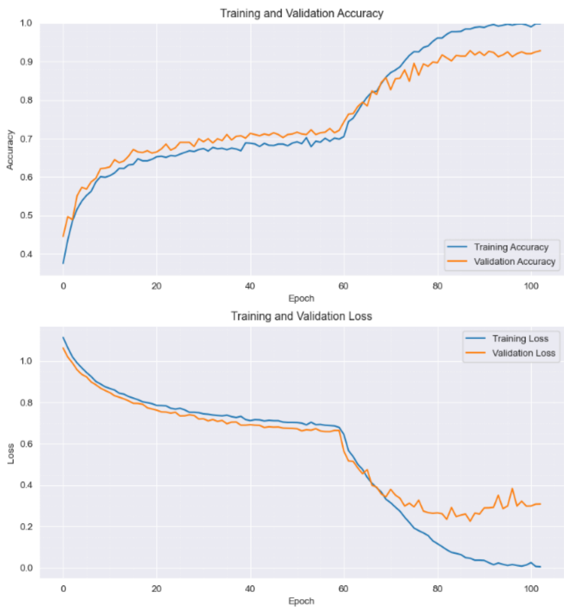

Figure 7. EfficientNet training and validation accuracy and loss versus epochs for RGB image.

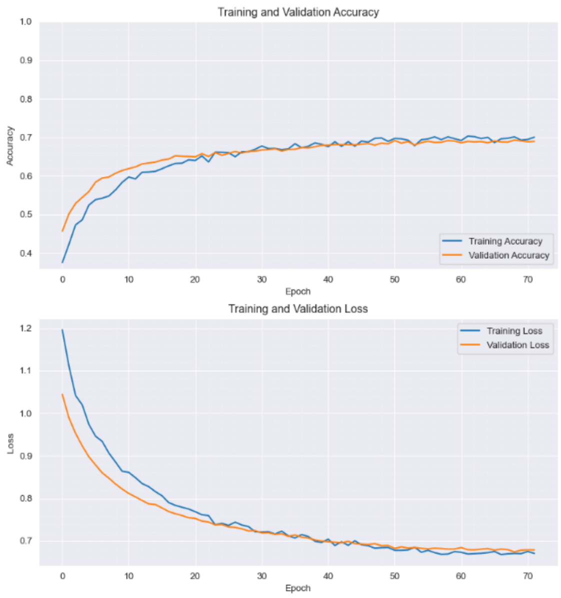

Figure 8. MobileNet training and validation accuracy and loss versus epochs for NIR image.

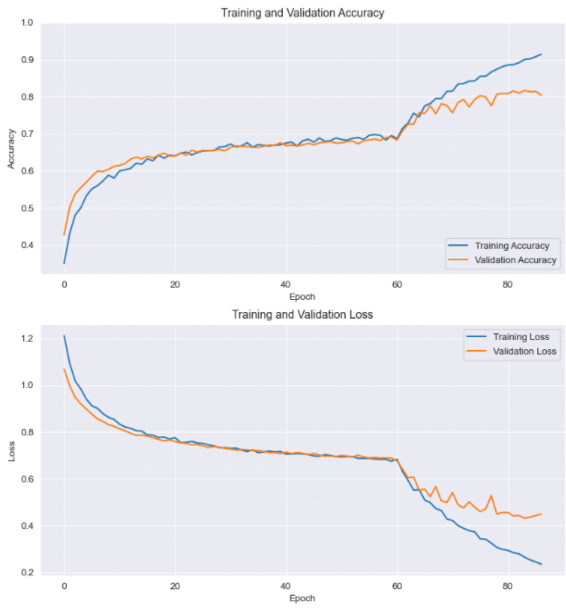

Figure 9. MobileNet-V2 training and validation accuracy and loss versus epochs for NIR image.

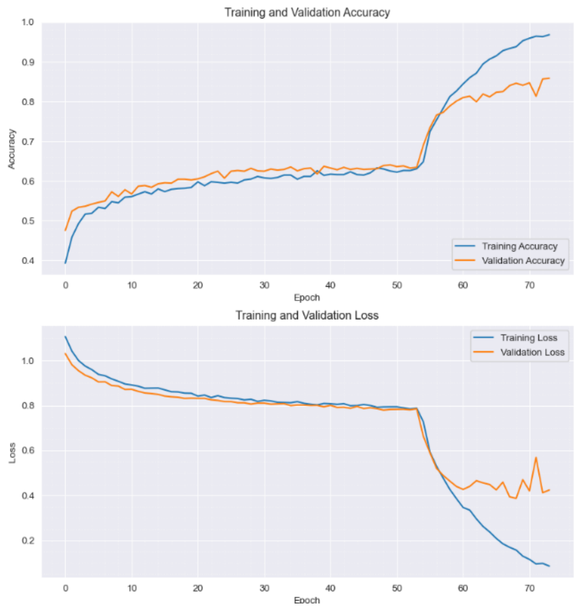

Figure 10. NasNet mobile training and validation accuracy and loss versus epochs for NIR image.

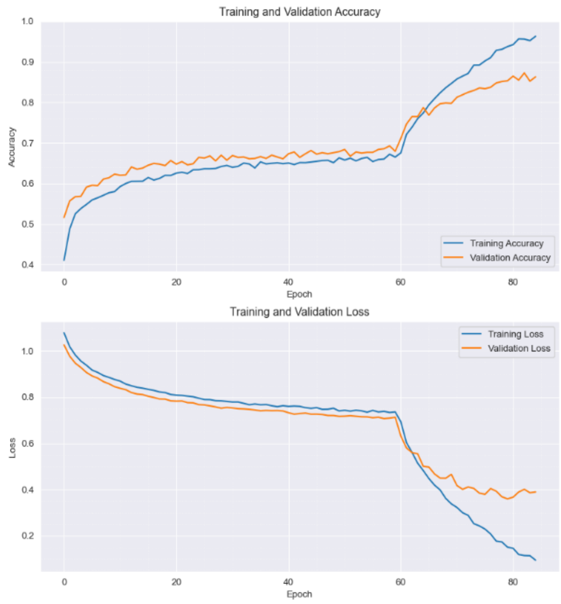

Figure 11. EfficientNet training and validation accuracy and loss versus epochs for NIR image.

4. Conclusion

This study investigated the ability of lightweight deep convolutional neural network models to identify drought-stressed plants, which can easily meet the design requirements for mobile and embedded vision applications. EfficientNet, the most successful model architecture, achieved success rates of 88.12 percent and 88.97 percent for RGB and NIR images, respectively. In the future, the trade-off between accuracy and computational cost may lead to the development of small and low-cost models, and appropriate models can be selected based on resource constraints.

Acknowledgments
This research was funded by Ministry of Higher Education and Universiti Teknologi Malaysia under research grant no. R.K130000.7843.5F348 and Q.K130000.2543.19H97. The authors gratefully thank Siti Aisyah Rohade for her assistance with the preparation of the manuscript.

ORCID iDs
MH. Kamarudin https://orcid.org/0000-0002-1187-4622

References
[1] Virnodkar S S, Pachghare V K, Patil V and Jha S K 2020 Remote sensing and machine learning for crop water stress determination in various crops: a critical review *Precision Agriculture* 1-35
[2] Ahmad U, Alvino A and Marino S 2021 A Review of Crop Water Stress Assessment Using Remote Sensing *Remote Sensing* **13** 4155
[3] Biabi H, Mehdizadeh S A and Salmi M S 2019 Design and implementation of a smart system for water management of lilium flower using image processing *Computers and Electronics in Agriculture* **160** 131-43
[4] Katsoulas N, Elvanidi A, Ferentinos K P, Kacira M, Bartzanas T and Kittas C 2016 Crop reflectance monitoring as a tool for water stress detection in greenhouses: A review *biosystems engineering* **151** 374-98
[5] Kim M, Lee C, Hong S, Kim S L, Baek J-H and Kim K-H 2021 High-Throughput Phenotyping Methods for Breeding Drought-Tolerant Crops *International Journal of Molecular Sciences* **22** 8266
[6] Gutiérrez S, Diago M P, Fernández-Novales J and Tardaguila J 2018 Vineyard water status assessment using on-the-go thermal imaging and machine learning *PloS one* **13** e0192037
[7] Zhuang S, Wang P, Jiang B, Li M and Gong Z 2017 Early detection of water stress in maize based on digital images *Computers and Electronics in Agriculture* **140** 461-8
[8] Chlingaryan A, Sukkarieh S and Whelan B 2018 Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review *Computers and electronics in agriculture* **151** 61-9
[9] Chandel N S, Chakraborty S K, Rajwade Y A, Dubey K, Tiwari M K and Jat D 2020 Identifying crop water stress using deep learning models *Neural Computing and Applications* 1-15
[10] Zhuang S, Wang P, Jiang B and Li M 2020 Learned features of leaf phenotype to monitor maize water status in the fields *Computers and Electronics in Agriculture* **172** 105347
[11] Islam M P and Yamane T 2021 HortNet417v1—A Deep-Learning Architecture for the Automatic Detection of Pot-Cultivated Peach Plant Water Stress *Sensors* **21** 7924
[12] Fahlgren N, Feldman M, Gehan M A, Wilson M S, Shyu C, Bryant D W, Hill S T, McEntee C J, Warnasooriya S N and Kumar I 2015 A versatile phenotyping system and analytics platform reveals diverse temporal responses to water availability in Setaria *Molecular plant* **8** 1520-35
[13] Feldman M J, Paul R E, Banan D, Barrett J F, Sebastian J, Yee M-C, Jiang H, Lipka A E, Brutnell T P and Dinneny J R 2017 Time dependent genetic analysis links field and controlled environment phenotypes in the model C4 grass Setaria *PLoS genetics* **13** e1006841
[14] Howard A G, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto M and Adam H 2017 Mobilenets: Efficient convolutional neural networks for mobile vision applications *arXiv preprint arXiv:1704.04861*
[15] Simonyan K and Zisserman A 2014 Very deep convolutional networks for large-scale image recognition *arXiv preprint arXiv:1409.1556*
[16] Sandler M, Howard A, Zhu M, Zhmoginov A and Chen L-C 2018 Mobilenetv2: Inverted residuals and linear bottlenecks. In: *Proceedings of the IEEE conference on computer vision and pattern recognition,* pp 4510-20
[17] Tan M, Chen B, Pang R, Vasudevan V, Sandler M, Howard A and Le Q V 2019 Mnasnet: Platform-aware neural architecture search for mobile. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition,* pp 2820-8

[18] Zoph B and Le Q V 2016 Neural architecture search with reinforcement learning *arXiv preprint arXiv:1611.01578*
[19] Tan M and Le Q 2019 Efficientnet: Rethinking model scaling for convolutional neural networks. In: *International Conference on Machine Learning*: PMLR) pp 6105-14