YOLO-NPK: A Light Deep Network for Lettuce Nutrients Deficiency Classification Based on Improved YOLOv8 Nano

The 10th International Electronic Conference on Sensors and Applications
ECSA-10
15–30 November 2023 | Online

Jordane Sikati [1] and Joseph Christian Nouaze [2,3,*]

[1] Guinee Biomedical Maintenance, R&D Center, Bentourayah 3137, Republic of Guinea; [2] Pusan National University, Department of Electronics Engineering, Busan 46241, Republic of Korea; [3] CAS Corporation, Headquarters, R&D Center, Yangju 11415, Republic of Korea; *Correspondence: Joseph Christian Nouaze (krxsange@pusan.ac.kr)

BACKGROUND:

Lettuce (Lactuca sativa) is a widely cultivated leafy vegetable of great economic and dietary significance. Adequate nutrient supply, particularly Nitrogen (N), Phosphorus (P), and Potassium (K), is crucial for optimal lettuce growth and quality. Deficiencies in these nutrients can lead to various issues, impacting yield, nutritional content, and disease resistance. Previous research has shown a growing interest in using deep learning-based approaches for early detection and diagnosis of nutrient deficiencies in lettuce plants, setting the stage for advancements in precision agriculture.

AIMS:

To explore the importance of Nitrogen, Phosphorus, and Potassium in lettuce cultivation, highlighting their roles and the consequences of deficiencies. Additionally, it discusses recent advancements in using deep learning techniques for the early detection of nutrient deficiencies in lettuce plants [1], with a focus on the YOLO-NPK approach, and sets the objective of improving classification accuracy, computational efficiency, and latency in identifying these deficiencies.

MATERIALS AND METHODS

Data Acquisition

The lettuce NPK dataset [16] from Kaggle consists of images representing various lettuce deficiency categories, including Fully Nutritional (FN) with 12 images, Nitrogen Deficient (-N) with 58 images, Phosphorus Deficient (-P) with 66 images, and Potassium Deficient (-K) with 72 images. These images were captured in a controlled hydroponic environment as part of a project aimed at developing a system for identifying and classifying lettuce deficiencies.

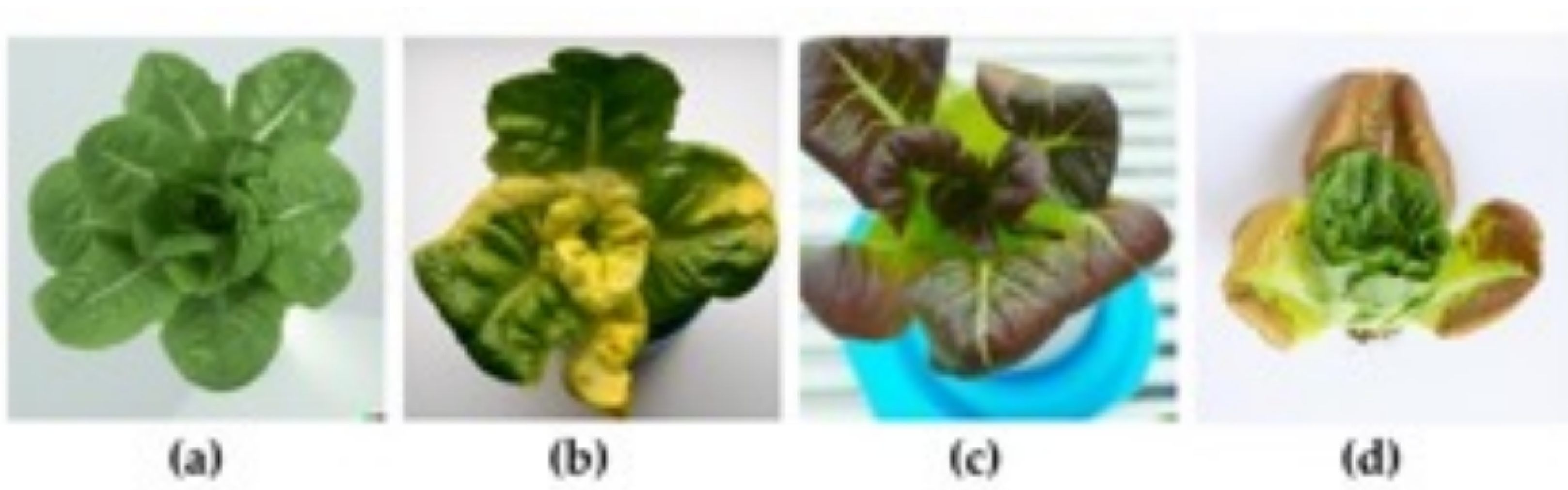

Figure 1: The dataset samples. (**a**)Fully Nutritional lettuce (FN); (**b**) Nitrogen Deficiency (-N); (**c**) Phosphorus Deficiency (-P); (**d**) Potassium Deficiency (-K).

VGG16 (Visual Geometry Group 16) Feature Extractor

VGG16, or Visual Geometry Group 16, is a highly effective and straightforw ard convolutional neural network (CNN) developed by the University of O xford. It's renowned for its 16 layers, including 13 convolutional and 3 fully connected layers, as well as its use of 3x3 convolutional filters with a stride of 1 and 2x2 max-pooling layers with a stride of 2. VGG16's deep architectur e, featuring repeated 3x3 convolutional filters and ReLU activation function s, enables it to learn intricate image features while maintaining spatial dime nsions through "same" padding.

Depthwise Convolution

Depthwise convolution is a key operation in deep learning and CNNs, espe cially in lightweight networks for mobile and edge devices. Unlike standard convolution, it processes input channels individually, reducing model para meters, memory usage, and computation time. It's often used with 1x1 poin twise convolution, creating a depthwise separable convolution that maintai ns spatial dimensions while reducing computational complexity. This appro ach is particularly efficient for low-level image features with less inter-chan nel correlation.

YOLOv8 Vs. YOLO-NPK

The YOLOv8 model, part of the YOLO (You Only Look Once) series, is kno wn for its efficiency in real-time object detection. It incorporates advanced t echniques like integrating the VGG16 feature extractor and depthwise conv olution to enhance feature extraction and classification accuracy, as seen in t he proposed approach for classifying lettuce deficiencies.

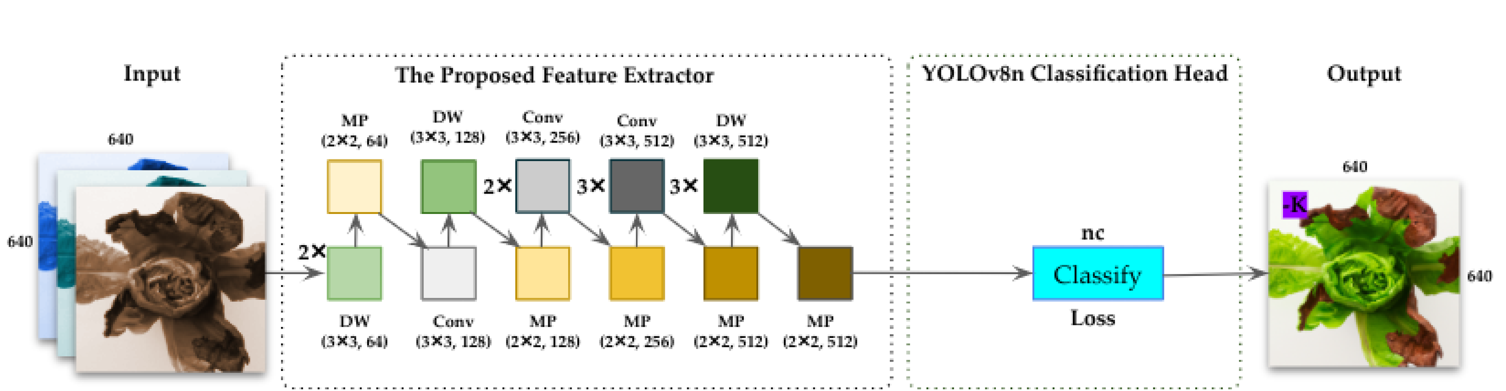

Figure 2: The architecture of YOLO-NPK. Conv, DW, MP, and nc respectively stand for convolution, depthwise convolution, max-pooling layer, and number of classes. The original backbone of YOLOv8n-cls has been replaced with the proposed feature extractor, and the classification head remains unchanged.

RESULTS

The experiments were conducted on a computer with specific hardware and software configurations, including an Intel® Core™ i5-11400H processor, an NVIDIA GeForce RTX 3050 GPU, and training with input images of size 640 x 640 pixels. The model underwent 300 epochs of training, utilizing various hyperparameters and data augmentation techniques. Top-1 Accuracy is a widely used metric in classification evaluation, representing the percentage of correctly classified samples based on the model's highest-confidence prediction.

Significant modifications were made to the YOLOv8n-cls backbone, includi ng replacing the backbone structure with VGG16 for improved classificatio n accuracy and adding depthwise convolutional layers to enhance memory efficiency and feature reuse. Table 1 summarizes these key modifications. Y OLO-NPK performs well on the validation set, particularly excelling in the FN class, and maintains strong performance in other classes (-N, -P, -K), de monstrating robust learning abilities (see Table 2). Compared to state-of-the -art methods, our approach outperforms with 99% Top-1 Accuracy, 9.2G F LOP, and 64.1 ms image classification latency, meeting our predefined criteria (Top-1 accuracy > 85%, FLOP < 10G, Latency < 170 ms). Other methods met the computational criteria but couldn't achieve the desired accuracy, highlighting our model's efficiency and resilience (details in Table 3).

Table 1: Ablation study on different modifications of YOLO-NPK

Table 2: Classification metrics. Comparison with different SOTA of YOLOv3, YOLOv4, YOLOv7 and the proposed method for the two studied classes (healthy and scab) that apple represents. The input image size is 416x416.

Classes	Images	Correctly Classified		Falsely Classified		Missed	
		Count	Rate	Count	Rate	Count	Rate
FN	53	53	100 %	0	0 %	0	0 %
-N	279	274	98.21 %	5	1.79 %	0	0 %
-P	256	254	99.22 %	2	0.78 %	0	0 %
-K	370	367	99.19 %	3	0.81 %	0	0 %

Figure 3: The classification output of YOLO-NPK. (**a**)Fully Nutritional lettuce (FN); (**b**) Phosphorus Deficiency (-P); (**c**) Nitrogen Deficiency (-N); (**d**) Potassium Deficiency (-K).

Table 3: Comparison of YOLO-NPK with the SOTA (State-Of-The-Art) method.

Methods	Images Size	Top-1 Accuracy (%)	FLOPs (G)	CPU Latency (ms)
SVM	640	85.3	12	141.6
VGG16	640	87.9	15.2	170.3
MobileNetV2	640	82.5	3.4	41.6
ShuffleNetv2	640	81.6	2.1	30.8
YOLOV8n-cls	640	93	3.3	19.8
YOLO-NPK	640	99	9.2	64.1

CONCLUSION

The study introduces YOLO-NPK, a lightweight deep neural network tailored for lettuce deficiency classification, building upon the foundation of YOLOv8 Nano Classification. This research aimed to enhance the baseline algorithm by introducing a custom feature extractor aligned with the study's needs. This goal was successfully met, achieving a Top-1 Accuracy exceeding 85%, maintaining a FLOP count under 10G, and ensuring a CPU latency below 170 ms per image, meeting the predefined objectives. Future plans involve integrating this solution into more complex systems for smart farming applications.

References

1. Nouaze, Joseph Christian, Jae Hyung Kim, Gye Rok Jeon, and Jae Ho Kim. "Monitoring of Indoor Farming of Lettuce Leaves for 16 Hours Using Electrical Impedance Spectroscopy (EIS) and Double-Shell Model (DSM)." *Sensors* 22, no. 24 (2022): 9671.