engineering proceedings

Proceeding Paper

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

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

[1] Guinee Biomedical Maintenance, R&D Center, Bentourayah 3137, Republic of Guinea; jordanesikati@gmail.com
[2] Department of Electronics Engineering, Pusan National University, Busan 46241, Republic of Korea
[3] CAS Corporation, Headquarters, R&D Center, Yangju 11415, Republic of Korea
* Correspondence: krxsange@pusan.ac.kr; Tel.: +82-1094011957

† Presented at the 10th International Electronic Conference on Sensors and Applications (ECSA-10), 15–11 30 November 2023; Available online: https://ecsa-10.sciforum.net/.

Abstract: When it comes to growing lettuce, specific nutrients play vital roles in its growth and development. These essential nutrients include full nutriments (FN), nitrogen (N), phosphorus (P), and potassium (K). Insufficient or excess levels of these nutrients can have negative effects on lettuce plants, resulting in various deficiencies that can be observed in the leaves. To better understand and identify these deficiencies, a deep learning approach is employed to improve these tasks. For the study, YOLOv8 Nano, a lightweight deep network, is chosen to classify the observed deficiencies in lettuce leaves. Several enhancements to the baseline algorithm are made, the backbone is replaced with VGG16 to improve the classification accuracy, and depthwise convolution is incorporated into it to enrich the features while keeping the head unchanged. The proposed network, incorporating these modifications, achieved superior classification results with a top-1 accuracy of 99%. This performance outperformed other state-of-the-art classification methods, demonstrating the effectiveness of the approach in identifying lettuce deficiencies. The objective of the research was to improve a baseline algorithm that could achieve the classification task above 85% of top-1 accuracy, with a FLOP inferior to 10G, and classification latency below 170 ms per image.

Keywords: Lettuce nutrient deficiency; Classification; Deep learning; YOLOV8 Nano

Citation: Sikati, J.; Nouaze, J.C. YOLO-NPK: A Light Deep Network for Lettuce Nutrients Deficiency Classification Based on Improved YOLOv8 Nano. **2023**, *5*, x. https://doi.org/10.3390/xxxxx

Academic Editor(s):

Published: 15 November 2023

1. Introduction

Lettuce (Lactuca sativa) is a widely cultivated leafy vegetable with significant economic and dietary importance. Adequate nutrient supply, particularly Nitrogen (N), Phosphorus (P), and Potassium (K) is essential for optimal lettuce growth and quality. Nitrogen is a primary component of chlorophyll and essential for photosynthesis. Nitrogen deficiency in lettuce results in stunted growth, pale leaves, and reduced leaf size, affecting the overall yield and nutritional content of lettuce, as well as its susceptibility to diseases [1]. Phosphorus is crucial for energy transfer in plants and plays a key role in root development. Lettuce plants deficient in phosphorus exhibit poor root growth, delayed maturity, and smaller heads. Phosphorus deficiency can also lead to decreased nutrient uptake, negatively impacting overall plant health [2]. Potassium is vital for maintaining plant turgor, enzyme activation, and disease resistance. Lettuce plants with potassium deficiency display wilted leaves, necrosis at leaf margins, and reduced resistance to pathogens [3]. Potassium deficiency can reduce the lettuce's marketability due to decreased visual appeal [4].

This paper is structured as follows: Section 2 relates the previous research on lettuce deficiencies, Section 3 presents the materials and methods, Section 4 discusses the experimental results of the proposed method, and finally, Section 5 provides the conclusions of this article and future work.

2. Related Work

In recent years, there has been a growing interest in the development of deep learning-based approaches for the diagnosis and early detection of nutrient deficiencies in lettuce plants. Watchareeruetai et al. introduced in 2018 an image analysis method for identifying nutrient deficiency in plants based on their leaves using convolutional neural networks [5], setting the stage for subsequent research in this area. In addition, a deep convolutional neural network for image-based diagnosis of nutrient deficiencies in plants grown in Aquaponics is proposed by Taha et al. in 2022 [6]. Furthermore, Lu et al., in 2023 introduced a lettuce pant trace-element-deficiency symptom identification via machine vision methods [7]. Collectively, these studies represent significant contributions to the field of lettuce NPK deficiency detection and illustrate the increasing reliance on deep learning methodologies for precision agriculture applications. Continued research in this area is crucial to developing sustainable agricultural practices that can meet the increasing demand for high-quality lettuce. In this way a deep learning approach called YOLO-NPK based on YOLOv8 Nano Classification algorithms [8,9] is employed in this study, to classify those deficiencies. The objective of the research is to improve a baseline algorithm that could achieve the classification task above 85% of Top-1 Accuracy, with a FLOP inferior to 10G, and classification latency below 170 ms per image.

3. Materials and Methods

3.1. Data Acquisition and Augmentation Strategy

The lettuce NPK dataset [10] was acquired on Kaggle. This dataset contains images of the following lettuce deficiency category together with Fully Nutritional lettuce: Fully Ntrutional (FN) 12 images, Nitrogen Deficient (-N) 58 images, Phosphorus Deficient (-P) 66 images, and Potassium Deficient (-K), 72 images. The images in this dataset were done in a controlled environment for a project on hydroponic lettuce deficiencies. The idea was to build a system that recognizes the lettuce deficiencies from the captured images and provides the classification of these deficiencies in hydroponics and other applications (Figure 1).

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

Augmentation techniques have been used to increase the training set and the validation set. The following pre-processing was applied to each image: auto-orientation of pixel data (with EXIF-orientation stripping) and resizing to 640 × 640 (Stretch). Furthermore, the successive augmentation was applied to create augmented versions of each source image: 50% probability of horizontal flip, 50% probability of vertical flip, equal probability of one of the following 90-degree rotations: none, clockwise, counter-clockwise, upside-down, randomly crop between 0 and 20 percent of the image, random shear of between -

15° to +15° horizontally and -15° to +15° vertically. 3192 samples were obtained from augmentation, with −K 1175, −N 975, −P 847, and FN 195. Therefore, the dataset was split into 70% for the training and 30% for the validation.

3.2. VGG16 (Visual Geometry Group 16) Feature Extractor

VGG16 (Visual Geometry Group 16) is a convolutional neural network (CNN) architecture for deep learning that was developed by the Visual Geometry Group at the University of Oxford [11]. It is part of the VGG family of models and is known for its simplicity and effectiveness in image classification tasks. It consists of 16 weight layers, including 13 convolutional layers and 3 fully connected layers. The architecture uses 3 × 3 convolutional filters with a stride of 1 and 2 × 2 max-pooling layers with a stride of 2. Also, it is characterized by its deep architecture, with small 3x3 convolutional filters stacked multiple times. This depth helps the network learn complex hierarchical features from images. The network uses 3 × 3 convolutional filters with a stride of 1 and "same" padding, which means the spatial dimensions of the feature maps do not change after convolutions. Rectified Linear Units (ReLU) are used as the activation function in the network, helping with the vanishing gradient problem and improving training.

3.3. Depthwise Convolution

Depthwise convolution is a specific type of convolutional operation used in deep learning and convolutional neural networks (CNNs). It is a fundamental building block for various lightweight and efficient neural network architectures, particularly those designed for mobile and edge devices [12]. Depthwise convolution differs from standard convolution in how it processes input channels. In a standard convolution operation, a kernel (also called a filter) slides through the entire input volume, considering all input channels simultaneously. In contrast, in depthwise convolution, each input channel is convolved with a separate kernel. This means that if you have k input channels and k separate kernels, each kernel is responsible for convolving with its corresponding input channel. It significantly reduces the number of parameters in the model compared to standard convolution. This reduction in parameters can lead to models that are more memory-efficient and faster to compute, making them suitable for resource-constrained environments. Often used in conjunction with pointwise convolution (1 × 1 convolution). This combination is referred to as a depthwise separable convolution. In depthwise separable convolution, the depthwise convolution layer is followed by a 1 × 1 pointwise convolution layer. The pointwise convolution combines the information from the separate channels produced by the depthwise convolution. Lastly, it maintains the spatial dimensions (width and height) of the input, but it can change the number of channels (depth). This contrasts with standard convolution, which can also change spatial dimensions. So, it is particularly efficient when dealing with low-level features in an image, where inter-channel correlations are not as significant. By separating the channels, it reduces computational complexity.

3.4. YOLOv8 (You Only Look Once version 8)

The YOLO (You Only Look Once) series [8,13–18] refers to a family of real-time object detection models that have been widely used in computer vision and deep learning. YOLO was initially introduced by Redmon et al. [9] in 2016 and has since seen several iterations, each with improvements and enhancements [8,9]. The primary idea behind YOLO is to perform object detection in a single forward pass of a neural network, making it very efficient and suitable for real-time applications. YOLOv8, developed by Ultralytics, represents the most recent iteration of the YOLO series. As an advanced and state-of-the-art model, it extends upon the achievements of its predecessors by introducing novel features and enhancements, resulting in elevated levels of performance, adaptability, and resource efficiency. YOLOv8 boasts comprehensive support for a wide spectrum of vision-based artificial intelligence tasks, encompassing detection, segmentation, pose estimation, tracking, and classification. This versatility empowers users to harness the diverse capabilities of YOLOv8 across a multitude of applications and domains.

3.5. YOLO-NPK

To enhance classification accuracy, the VGG16 feature extractor is integrated into the backbone of YOLOv8n-cls (YOLOv8 Nano Classification). Furthermore, depthwise convolution is introduced within the feature extractor to facilitate feature reuse and empower the deep network to extract more complex and richer features. The diagram below provides an overview of the proposed approach for classifying lettuce deficiencies. The proposed feature extractor receives a 640 × 640 RGB deficient lettuce image as input and extracts richer features. The classification head fuses the learned feature and performs a classification task, returning a classification result as output (Figure 2).

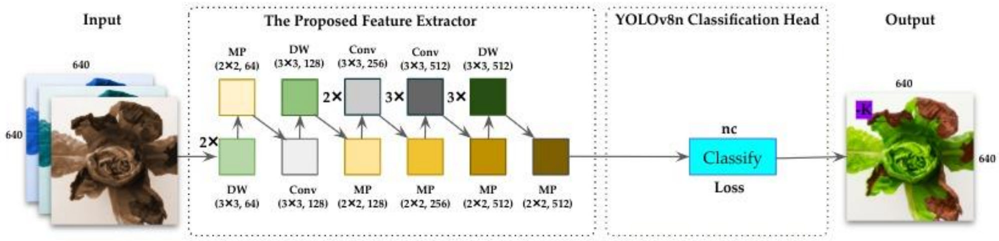

Figure 2. The architecture of YOLO-NPK. Conv, DW, MP, and nc respectively stand for convolution, depthwise convolution, max-pooling layer, and number of classes. The original backbone of YOLOv8n-cls has been replaced with the proposed feature extractor, and the classification head remains unchanged.

4. Results and Discussion

4.1. Experimental Setup

The experiments were carried out on a computer equipped with the following specifications: an Intel® Core™ i5-11400H processor 11th Generation with 64-bit architecture, running at 2.70GHz and featuring a dodecore CPU. Additionally, the computer was equipped with an NVIDIA GeForce RTX 3050 GPU. The model received input images sized at 640 x 640 pixels. However, due to constraints on GPU memory, the batch size was set to 8 during training. The training process spanned 116 epochs and commenced with an initial learning rate of 0.01, which was later adjusted to a final learning rate of 0.1. Moreover, specific hyperparameters were set as follows: a momentum of 0.937 and a weight decay of 0.0005. During the warmup epoch, warmup momentum, and warmup bias learning rate stages, the values were configured at 3.0, 0.8, and 0.1, respectively. The optimizer employed for training the models was Stochastic Gradient Descent (SGD). Data augmentation techniques were used proportionally such as mosaic, paste-in, and scaling while training the deep network to avoid unbalanced classes. The early stop mechanism was employed to overcome overfiting.

In the context of classification accuracy, *Top-1 Accuracy* refers to the proportion of correctly classified samples where the model's top prediction matches the true label. It can be mathematically expressed as follows:

$$Top1\ Accuracy = \frac{Number\ of\ Correct\ Predictions}{Total\ Number\ of\ Predictions} \times 100 \tag{1}$$

In this expression, *Number of Correct Predictions* is the count of instances where the model's top prediction matches the true class labels, and *Total Number of Predictions* is the total number of instances or samples in the dataset. The result is typically expressed as a percentage to represent the accuracy rate. *Top-1 Accuracy* is a common metric used to evaluate the performance of classification models, where only the highest-confidence prediction is considered for each sample.

4.2. Ablation Study

Several components of the YOLOv8n-cls backbone were modified to obtain the desired results. The overall structure of the backbone was replaced by the VGG16 feature to improve the classification accuracy, and the depthwise convolutional layers were inserted along the feature extractor to allow efficient memory computation and better reuse of features. These operations have shown interesting improvement. Table 1 provides details on these diverse modifications.

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

VGG16	Depthwise Convolution	Top-1 Accuracy (%)	FLOPs (G)	CPU Latency (ms)
		93	3.3	19.8
✓		97.5	14.5	68.3
	✓	95.2	2.4	18.2
✓	✓	99	9.2	64.1

4.3. Classification Performance

The performance of YOLO-NPK was measured on the validation set, which represents all the classes. Notably, it shows acceptable results in terms of classification. The model performs efficiently on the FN set and achieves good classification results on other classes(-N, -P, and -K). Figure 3 shows the confusion matrix of the proposed model. Figure 4 shows the classification output of YOLO-NPK.

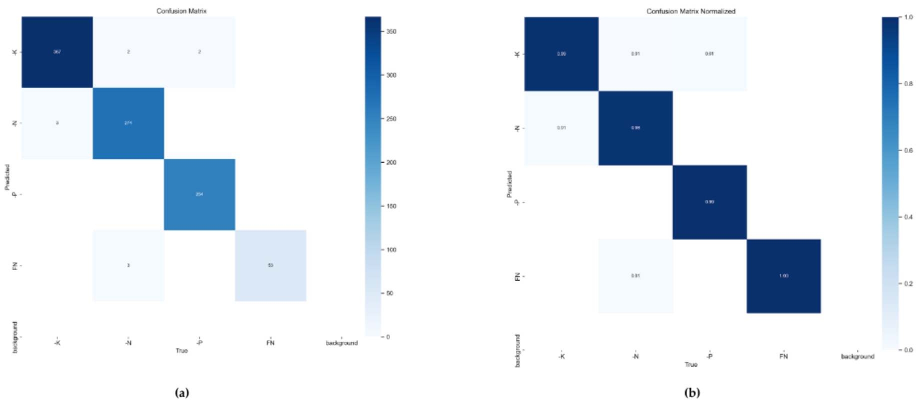

Figure 3. The confusion matrix of YOLO-NPK. (a) confusion matrix. (b) confusion matrix normalized Fully Nutritional lettuce (FN); Phosphorus Deficiency (-P); Nitrogen Deficiency (-N); (**d**) Potassium Deficiency (-K). True represents the ground truth in the dataset, predict is the classification result, and the background is the images that were missed by the model.

This proves the learning capability of the proposed method. More details are provided in Table 2.

Table 2. Classification performance of YOLO-NPK. FN, -N, -P, and -K respectively represent Full Nutritional, Nitrogen deficiency, Phosphorus deficiency, and Potassium deficiency.

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 %

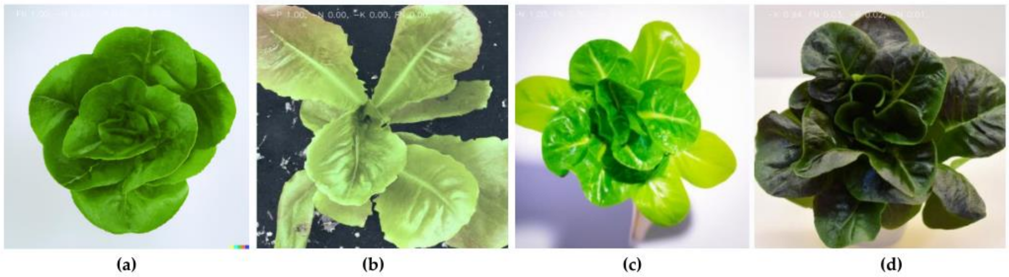

Figure 4. The classification output of YOLO-NPK. (**a**)Fully Nutritional lettuce (FN); (**b**) Phosphorus Deficiency (-P); (**c**) Nitrogen Deficiency (-N); (**d**) Potassium Deficiency (-K).

4.4. Comparison of State-of-the-art Methods

The proposed method, YOLO-NPK, was compared with different state-of-the-art methods. The proposed one has shown better classification accuracy. The Top-1 Accuracy reached 99%, The FLOP 9.2G, and the classification latency per image 64.1 ms. This respects the guidelines fixed before the experiments (Top-1 accuracy above 85%, FLOP under 10G, and Latency below 170 ms). Other methods respected the FLOP and Latency conditions, but could not fulfill the Top-1 Accuracy expectation, proving the efficiency and the robustness of the proposed model. Table 3 gives details on those comparisons.

Table 3. Comparison of the 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

5. Conclusions and Future Work

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.

Author Contributions: Conceptualization, J.S., and J.C.N.; validation, J.S., and J.C.N.; investigation, J.S., and J.C.N.; data curation, J.S., and J.C.N.; formal analysis, J.S., and J.C.N.; methodology, J.S., and J.C.N.; software, J.S., and J.C.N.; visualization, J.S., and J.C.N.; supervision, J.C.N.; writing—original draft preparation, J.S.; writing—review and editing, J.S., and J.C.N.; project administration, J.C.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable

Data Availability Statement: The Lettuce NPK dataset used in this project was provided by Kaggle.

The dataset is available here [10].

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Muñoz-Huerta, R.F.; Guevara-Gonzalez, R.G.; Contreras-Medina, L.M.; Torres-Pacheco, I.; Prado-Olivarez, J.; Ocampo-Velazquez, R.V. A Review of Methods for Sensing the Nitrogen Status in Plants: Advantages, Disadvantages and Recent Advances. *Sensors* **2013**, *13*, 10823–10843, doi:10.3390/s130810823.
2. Khan, F.; Siddique, A.B.; Shabala, S.; Zhou, M.; Zhao, C. Phosphorus Plays Key Roles in Regulating Plants' Physiological Responses to Abiotic Stresses. *Plants* **2023**, *12*, 2861, doi:10.3390/plants12152861.
3. Nouaze, J.C.; Kim, J.H.; Jeon, G.R.; Kim, J.H. Monitoring of Indoor Farming of Lettuce Leaves for 16 Hours Using Electrical Impedance Spectroscopy (EIS) and Double-Shell Model (DSM). *Sensors* **2022**, *22*, 9671, doi:10.3390/s22249671.
4. Davis, J.L.; Armengaud, P.; Larson, T.R.; Graham, I.A.; White, P.J.; Newton, A.C.; Amtmann, A. Contrasting Nutrient-Disease Relationships: Potassium Gradients in Barley Leaves Have Opposite Effects on Two Fungal Pathogens with Different Sensitivities to Jasmonic Acid. *Plant Cell Environ* **2018**, *41*, 2357–2372, doi:10.1111/pce.13350.
5. Watchareeruetai, U.; Noinongyao, P.; Wattanapaiboonsuk, C.; Khantiviriya, P.; Duangsrisai, S. Identification of Plant Nutrient Deficiencies Using Convolutional Neural Networks. *2018 International Electrical Engineering Congress (iEECON)* **2018**, 1–4, doi:10.1109/IEECON.2018.8712217.
6. Taha, M.F.; Abdalla, A.; ElMasry, G.; Gouda, M.; Zhou, L.; Zhao, N.; Liang, N.; Niu, Z.; Hassanein, A.; Al-Rejaie, S.; et al. Using Deep Convolutional Neural Network for Image-Based Diagnosis of Nutrient Deficiencies in Plants Grown in Aquaponics. *Chemosensors* **2022**, *10*, 45, doi:10.3390/chemosensors10020045.
7. Lu, J.; Peng, K.; Wang, Q.; Sun, C. Lettuce Plant Trace-Element-Deficiency Symptom Identification via Machine Vision Methods. *Agriculture* **2023**, *13*, 1614, doi:10.3390/agriculture13081614.
8. Jocher, G.; Chaurasia, A.; Qiu, J. YOLO by Ultralytics 2023.
9. Ultralytics Home Available online: https://docs.ultralytics.com/ (accessed on 5 September 2023).
10. Lettuce NPK Dataset Available online: https://www.kaggle.com/datasets/baronn/lettuce-npk-dataset (accessed on 5 September 2023).
11. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition 2015.
12. Chollet, F. Xception: Deep Learning with Depthwise Separable Convolutions. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); July 2017; pp. 1800–1807.
13. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection.; IEEE Computer Society, June 1 2016; pp. 779–788.
14. Redmon, J.; Farhadi, A. YOLO9000: Better, Faster, Stronger. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); July 2017; pp. 6517–6525.
15. Redmon, J.; Farhadi, A. YOLOv3: An Incremental Improvement 2018.
16. Jocher, G.; Stoken, A.; Borovec, J.; ChristopherSTAN; Changyu, L.; Hogan, A.; Laughing; NanoCode012; yxNONG; Diaconu, L.; et al. Ultralytics/Yolov5: V2.0 2020.
17. Li, C.; Li, L.; Jiang, H.; Weng, K.; Geng, Y.; Li, L.; Ke, Z.; Li, Q.; Cheng, M.; Nie, W.; et al. YOLOv6: A Single-Stage Object Detection Framework for Industrial Applications 2022.
18. Wang, C.-Y.; Bochkovskiy, A.; Liao, H.-Y.M. YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors 2022.