IDENTIFICATION OF DEEP NUTRIENT DEFICIENCIES IN WINTER RYE USING UAV ACQUIRED RGB IMAGES THOUGH MULTIPLE TRANSFER LEARNING BASED CONVOLUTIONAL NEURAL NETWORKS

Paul Uhrich, Matjaz Cigler, Ian Saari, Marie Howel, Md Jaber Al Nahian

University of Calgary, 2024

ABSTRACT

In the pursuit of advancing precision agriculture, this study leverages Unmanned Arial Vehicle (UAV) imagery and Convolutional Neural Networks (CNNs) enhanced by transfer learning to identify nutrient deficiencies in winter wheat and winter rye. Recognizing the critical impact of nutrient management in crop yield and health, we propose an approach that utilizes UAV images as an input to various pre-trained CNN models. Each model, fine-tuned through transfer learning techniques, is adept at detecting and classifying varieties of common nutrient deficiencies. Results demonstrate a promising accuracy rate dependent on the base CNN and show the potential for use of the model as a reliable tool for early diagnosis and intervention in nutrient management practices. This not only contributes to the field of agricultural technology, but also to more sustainable farming methodologies.

Index Terms — CNN, transfer learning, nutrient deficiencies, neural networks, deep learning, precision agriculture, UAV imagery.

1. INTRODUCTION

The advent of deep learning has revolutionized precision agriculture and the way that crop health is monitored and managed. As shown in this paper winter rye is heavily influenced by adequacy of nutrient supply. Nutrient deficiencies, therefore, pose a significant threat to both yield and nutritional value, necessitating timely and accurate detection [1].

This paper utilizes the DND-Diko-WWWR dataset, and the long-term fertilizer experiment (LTFE) Dikopshof primarily in the works of precision agriculture and their association with deep learning technologies. The integration of these technologies aims to transcend traditional monitoring methods, offering a non-invasive scalable solution to identify nutrient deficiencies in winter wheat and winter rye.

This report intends to further the LTFE research with the use of various prevalent pre-trained convolutional neural networks and the existing UAV imagery to discern the best applicable existing model and the most performant hyperparameters.

The ensuing sections will detail the methodology employed, the results obtained, and the implications of this research on the broader context of agricultural technology, sustainable farming, and deep learning.

2. RELATED WORK

The intersection of deep learning techniques and agricultural sciences has seen significant developments, particularly in the field of precision agriculture. A cornerstone of this evolution is the work conducted under the Long-Term Fertilizer experiment (LTFE) at Dikopshof, spearheaded by the University of Bonn and funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy. [1] [2] The LTFE's focus on non-invasive nutrient deficiency diagnosis represents a pivotal step towards sustainable and precision agriculture. This section reviews the existing literature and contributions from the LTFE and related works on the DND Dinko dataset.

The LTFE at Dikopshof was established to explore the impacts of various fertilizer treatments on crop health and yield. Its goals include improving nutrient management practices, minimizing environmental impacts, and enhancing crop yield and quality through precision agriculture techniques.

One of the significant advancements from the LTFE is the development of methodologies for non-invasive nutrient deficiency diagnosis. By leveraging UAV imagery and (CNNs), researchers have been able to identify and classify nutrient deficiencies in crops with remarkable accuracy. This approach not only aids in the early detection and management of such deficiencies but also contributes to reducing the overuse of fertilizers. [3]

In part of the 8th workshop on Computer Vision in Plant Phenotyping and Agriculure (CVPPA) the LTFE Dikopshof offed up a challenge to properly identify winter wheat and winter rye using UAV-based RGB images [5]. Within this competition, applicants were invited to utilize deep learning methodologies to classify nutrient deficiences using a supplied set of images, known as the DND-Diko-WWWR dataset [6].

Different groups of researchers tried to solve this problem using different deep learning and transfer learning models. Zhang et al. [19] proposed the DEEM (DividE and

EnseMble) technique to classify nutrient deficiency of winter wheat and rye from drone images. They divided the dataset into different groups based on the collection date of data and developed separate models for each group of data. A self-supervised pseudo-labelling approach was utilized to label the test data with the help of ensemble learning and iteratively include the labeled test images into the training dataset. They reported that their proposed technique secured first place in the CVPPA challenge.

The second-best submission method used the Swin Transformer V2 model as a baseline model [20]. They utilized the pre-trained weights of the ImageNet dataset and applied different state-of-the art augmentation techniques. They also used Test Time Augmentation technique as a post processing step to get the better result.

Liu and Liang [21] secured third place in the challenge. They also used Swin Transformer V2 model. Kuzu et al. [22] created a pipeline for detecting nutrient deficiencies that uses soft vegetation indices taken from RGB photos as well as a modified vision transformer named ViT-hybrid.

3. MATERIALS AND METHODS

The DND-Diko-WWWR dataset has 1800 RGB images of each winter wheat, and winter rye, classified with one of 7 types of fertilizer treatment ranging from completely unfertilized, to single deficiencies of nitrogen, phosphorus, potassium, lime, and some with an additional application of mineral fertilizer and farmyard manure [19]. This dataset represented equal samples of each subtype at three different dates throughout the experimental period. Of each respective dataset, 1332 images were categorized. Of the categorized data, these were each split into training, validation, and test datasets in a 70%, 15% and 15% split respectively.

Fig. 1. Example image from the DND-Diko-WWWR Dataset [1]

The intent of this research aims to identify the most performant model for nutrient deficiencies utilizing the DND-Diko-WWWR dataset, utilizing existing PyTorch models with default weights. All processing was completed on the University of Calgary Teaching and Learning Cluster (TALC) using Python v3.12.

Given the detailed nature of the RGB images in the DND-Diko dataset, preference was given to models with higher native resolution initial training weights. Three types of models were used to develop the array of final models produced. First ResNext was used, ResNext is a model that is developed by Microsoft and is a trusted reference for image classification models [16]. Second was SwinV2, this model is a transformer that uses the Residual post normalization approach [15]. This model brings diversity into the set to ensure that a range of models was used and not just traditional CNN's. Finally, EfficientNet was used. EfficientNet is most interesting in its ingenious use of compound coefficient to scale up the model [18]. This creates a model that is more accurate as well as efficient. All the models selected for use in this analysis included ResNext, SwinV2, EfficientNetB0, EfficientNetB1, EfficientNetB2, EfficientNetB3, EffifientNetB4, EfficientNetB5, EfficientNetB6, EfficientNetB7, as well as a native built CNN using the DND-Diko dataset [4], [5], [6], [7], [8], [9], [10], [11], [13], [14]. Given the maximum GPU processing capacity of TALC, results using EfficientNet B2 – B7 were limited.

For each of the models, training was completed on the 70% split, using the following batch sizes, and learning rates; Batch Sizes = [8,16,32,64,128], Learning Rates = [0.05, 0.01, 0.005, 0.001, 0.0005 0.0001], and validated on the second 15% split. Data was transformed prior to training, normalizing to the specified average, standard deviations of each pre-trained model as well as resizing, cropping, and interpolating on the same basis. Each model was optimized with the ADAM optimizer at the specified learning rate and learning rate scheduler with step size = 7 and gamma = 0.1. Models were initially set to a max of 30 epochs, and a patience set of 5. None of the models that provided competitive accuracy ran for the full 30 epochs and would patience out relatively early on.

Following completion of each model run, all applicable hyperparameter sets were logged and maintained with ideal model weights saved for further testing analysis. Each result set from each run for each hyperparameter set was analyzed for each model, and the lowest loss model was then selected for final testing for comparison against other models.

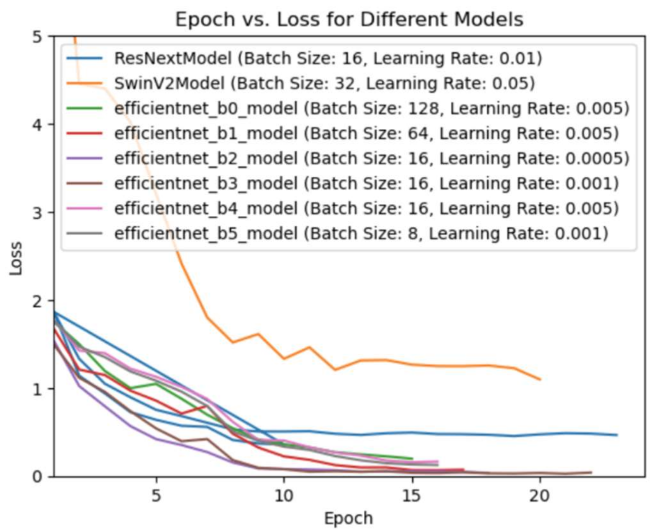

Fig. 2. Model Hyperparameter tuning visualization (EfficientNet b1, Learning Rate: 0.005, Batch Size: 64).

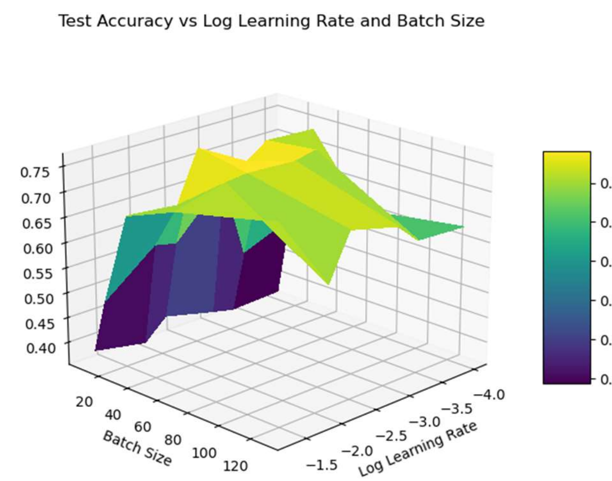

Fig. 3. Model Hyperparameter tuning visualization (EfficientNet b1, Learning Rate: 0.005, Batch Size: 64).

The peak model accuracy and loss following this step was used to compare models against each other. Here we found that the EfficientNet B4 model, (the highest resolution model we could run on TALC) when properly tuned yielded our best results.

All code utilized for this report can be found in the project repository, found at https://github.com/mzcigler14/ENEL645_Project.

4. RESULTS AND DISCUSSION

The initial hypothesis was that aside from hyperparameter tuning, model resolution and model complexity would play the largest part in determining the overall accuracy and loss of the resultant image classification. It was because of this reason that the EfficientNet model series, with 8 distinct different models and initial weight sets was chosen [12]. Each model in the series had increasingly higher image input resolution and model complexity.

In the process of completing the models on TALC, we did reach the theoretical computational limitations of the system, and were unable to compelete transfer learning model training sets in their entirity on EfficientNet B6 and EfficientNet B7. Thus for the purposes of our reports, these results will be completely disqualified from the results, or if shown should not imply qualitative results. Additionally, we were only able to complete selected hyperparameter sets (lower batch sizes) on all EfficientNet Models other than B0 (this represents models with input images of resolution 256x256 and higher).

As shown in figure 3, we can see that the generalizatoin for hyperparameter tuning for this application showed significant improvements when increasing batch sizes from the lower of 8 to 16/32, however thereafter saw diminishing, or even decreasing returns. Learning Rate had less of an influence, with a small bias towards lower learning rates. This behavior was consistent throughout all models with some variance.

When seeing a comparison of tuned model hyperparameters for various models and configurations (shown below), the EfficientNet models all performed similarly, even though the higher resolution models could not be trained at large batch sizes. This leads to the hypothesis that if a more powerful GPU was used to train the high resolution models, they would yeild the best results. Further even after ideal tuning for SwinV2 and ResNext models, their accuracy and loss was still a less than ideal than any of the EfficientNet models.

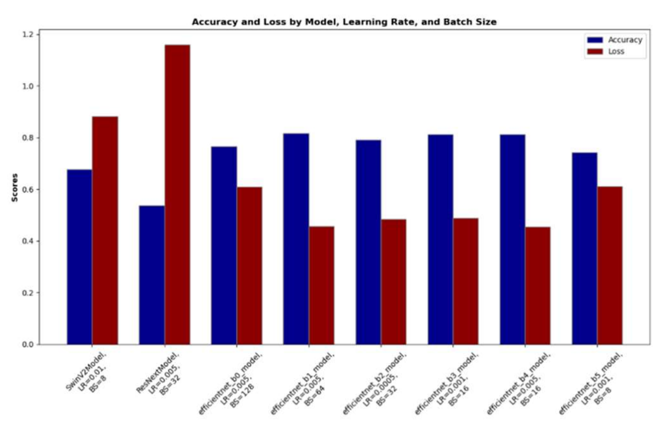

Fig. 4. Peak results of various CNN's following Hyperparameter tuning.

Additionally, we had attempted to make a novel CNN trained exclusively on the DND-Diko Dataset, however found that due to the lack of extensive training from excessive images for patterns and shapes such as any ImageNet based CNNs, the results from this model were significantly dispersant when compared to the pretrained varietals shown in the results here.

Further, as we were intending to test EfficientNet up to B7, we hoped to further explore the limitations of higher resolution models. However, the Teaching and Learning Cluster did not have the computational power to process these models. While we did observe the pattern of increases in accuracy with higher resolution models the materiality of them may not have validated the additional computational power for utilization in actual application.

Based on the above, the best model for our application was the EfficientNet B4 model, with a batch size of 16, and a learning rate of 0.005, rewarding us with a test accuracy of just higher than 81%, and a loss of 0.454.

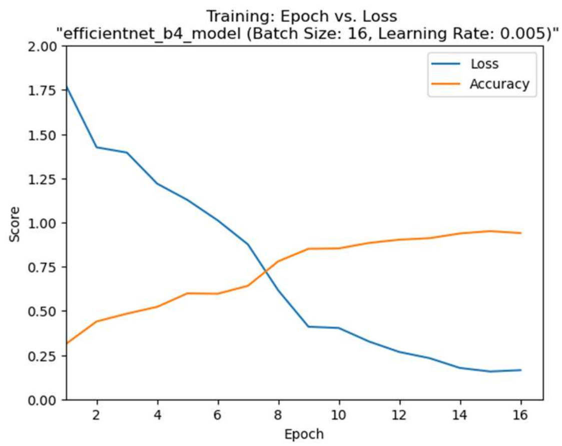

Fig. 5. Confusion Matrix of EfficientNet B4 CNN for Deep Nutrient Deficiencies in Winter Rye

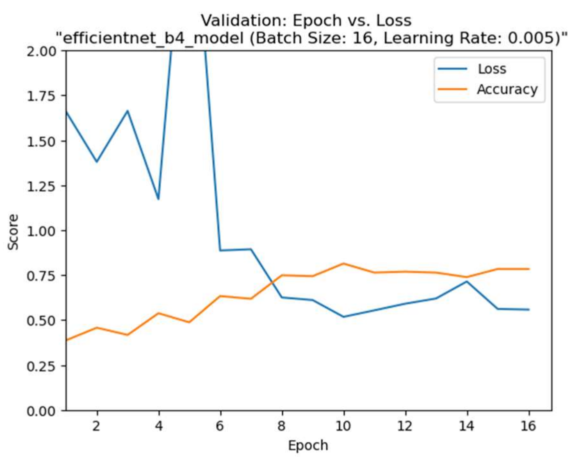

Fig. 6. Confusion Matrix of EfficientNet B4 CNN for Deep Nutrient Deficiencies in Winter Rye

The model's accuracy and loss over the development of the model can be seen in figures 5 and 6.

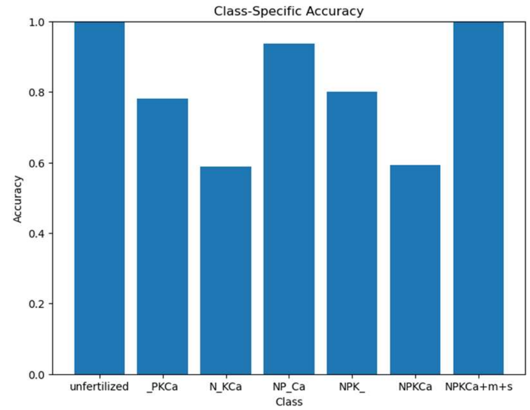

Fig. 7. EfficientNet B4 Test Class-Specific Accuracies

As shown in figure 5, we can clearly see that we could very accurately determine if a crop was completely unfertilized or entirely fertilized, however specifying precise deep nutrient deficiencies was slightly less accurate, though still within reason acceptable range.

Furthermore, it must also be noted that this represents individual image-based decision-making of the model. In a real-life application, given the relatively effective accuracy of the model, one could use aggregate model categorization to discern larger scale fertilizer application decisions. Incorporating this logic into a model, utilizing geo-spatial data could potentially increase the model materially by giving it the ability to make aggregated decisions based on numerous samples and taking a probabilistic approach to deep nutrient deficiency categorization.

Additionally, if given more information outside of image data, such as chronological information about the crops, base soil nutrient information, rainfall and weather conditions throughout the growing season, we would expect that the recognition of deep nutrient deficiencies to be significantly more accurate, allowing growers to utilize all easily available data to get confident and reliable categorization.

5. CONCLUSIONS

As a result of our analysis, we have observed that initial weights and model architectures play the largest role in determining a model's effectiveness in image categorization. Furthermore, we have also confirmed that such models are implicitly important in producing an ideal transfer learning convolutional neural network for image categorization in our application of deep nutrient deficiencies for winter rye.

Based on the limitations of our computational power, we found that higher resolution models are more effective at categorizing deep nutrient deficiencies and can do so reliably and effectively when tuned with optimum hyperparameters.

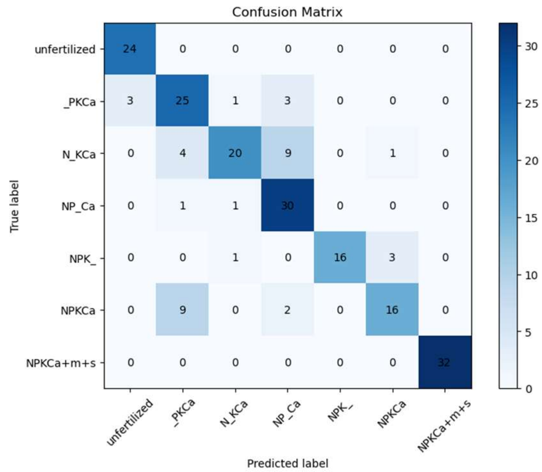

Fig. 8. Confusion Matrix of EfficientNet B4 CNN for Deep Nutrient Deficiencies in Winter Rye

Additionally, with our optimum model, EfficientNet B4 properly tuned with a batch size of 16 and learning rate of 0.005 and using an ADAM optimizer, gaining an average accuracy of 81% overall (max 100% full fertilization, min 60% nitrogen deficient fertilization) we can conclude that with images alone we could effectively categorize deep nutrient deficiencies.

These values could be improved with the help of higher computational processing power for higher resolution models, such as EfficientNet B7, and even further improved by creating composite models utilizing not only image categorization, but also geo-spatial information and aggregate image categorization, as well as other crop and region-specific information such as weather, or soil base nutrient data.

Overall, we have found our research to be successful in identifying ideal models for deep nutrient deficiency categorization, however recognizing that further work can be done to make them more robust and effective.

REFERENCES

[1] Dikipshof, DND-Diko-WWWR: RGB images of winter wheat and winter rye for nutrient deficiency recognition., Bonn, Germany: LTFE, 2023.

[2] J. Yi, G. Lopez, S. Hadir, J. Weyler, L. Klingbell, M. Deichmann, J. Gall and S. J. Seidel, "Non-invasive Diagnosis of Nutrient Deficiences in Winter Wheat and Winter Rye using UAV-Based RGB Images," SSRN, Bonn, Germany, 2023.

[3] J. Yi, L. Krusenbaum, P. Unger, H. Huging, S. J. Seidel, G. Schaaf and J. Gall, "Deep Learning for Non-Invasive Diagnosis of Nutrient Deficiencies in Sugar Beet Using RGB Images," National Library of Medicine, 2020.

[4] J. Yi, "CVPPA@ICCV'23: Image Classification of Nutrient Deficiences in Winter Wheat and Winter Rye," LTFE, 2023. [Online]. Available: https://codalab.lisn.upsaclay.fr/competitions/13833.

[5] PyTorch, "efficientnet_b0," Torch Contributors, 2017. [Online]. Available: https://pytorch.org/vision/main/models/generated/torchvision.models.efficientnet_b0.html#torchvision.models.efficientnet_b0. [Accessed 12 03 2024].

[6] PyTorch, "efficientnet_b1," Torch Contributors, 2017. [Online]. Available: https://pytorch.org/vision/main/models/generated/torchvision.models.efficientnet_b1.html. [Accessed 12 03 2024].

[7] PyTorch, "efficientnet_b2," Torch Contributors, 2017. [Online]. Available: https://pytorch.org/vision/main/models/generated/torchvision.models.efficientnet_b2.html. [Accessed 12 03 2024].

[8] PyTorch, "efficientnet_b3," Torch Contributors, 2017. [Online]. Available: https://pytorch.org/vision/main/models/generated/torchvision.models.efficientnet_b3.html. [Accessed 12 03 2024].

[9] PyTorch, "efficientnet_b4," Torch Contributors, 2017. [Online]. Available: https://pytorch.org/vision/main/models/generated/torchvision.models.efficientnet_b4.html. [Accessed 12 03 2024].

[10] PyTorch, "efficientnet_b5," Torch Contributors, 2017. [Online]. Available: https://pytorch.org/vision/main/models/generated/torchvision.models.efficientnet_b5.html. [Accessed 12 03 2024].

[11] PyTorch, "efficientnet_b6," Torch Contributors, 2017. [Online]. Available: https://pytorch.org/vision/main/models/generated/torchvision.models.efficientnet_b6.html. [Accessed 12 03 2024].

[12] PyTorch, "efficientnet_b7," Torch Contributors, 2017. [Online]. Available: https://pytorch.org/vision/main/models/generated/torchvision.models.efficientnet_b7.html. [Accessed 12 03 2024].

[13] A. Sarkar, "Understanding EfficientNet — The most powerful CNN architecture," Medium, 8 05 2021. [Online]. Available: https://arjun-sarkar786.medium.com/understanding-efficientnet-the-most-powerful-cnn-architecture-eaeb40386fad. [Accessed 11 03 2024].

[14] PyTorch, "resnext101_32x8d," Torch Contributors, 2017. [Online]. Available: https://pytorch.org/vision/main/models/generated/torchvision.models.resnext101_32x8d.html#torchvision.models.resnext101_32x8d. [Accessed 2 03 2024].

[15] PyTorch, "swin_v2_s," Torch Contributors, 2017. [Online]. Available: https://pytorch.org/vision/main/models/generated/torchvision.models.swin_v2_s.html#swin-v2-s. [Accessed 5 03 2024].

[16] S.-H. Tsang, "Review — Swin Transformer V2: Scaling Up Capacity and Resolution," Medium, 29 03 2023. [Online]. Available: https://sh-tsang.medium.com/review-swin-transformer-v2-scaling-up-capacity-and-resolution-401c28b02df8. [Accessed 5 03 2024].

[17] A. Erdoğan, "ResNeXt: A New Paradigm in Image Processing," Medium, 4 11 2023. [Online]. Available: https://medium.com/@atakanerdogan305/resnext-a-new-paradigm-in-image-processing-ee40425aea1f#:~:text=ResNeXt%20is%20a%20Convolutional%20Neural,Transformations%20for%20Deep%20Neural%20Networks.%E2%80%9D. [Accessed 2 03 2024].

[18] Q. V. Le and M. Tan, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," Cornell University, 2019.

[19] H. Zhang, X. Shen, H. Du, H. Chen, C. Liu, H. Sheng, Q. Xu, M.D. Khan, Q. Yu, T. Zhu, and S. Chapman, 2023. Divide and Ensemble: Progressively Learning for the Unknown. arXiv preprint arXiv:2310.05425.

[20] Y. Wang, 2023. Boosting High Resolution Image Classification with Scaling-up Transformers. arXiv preprint arXiv:2309.15277.

[21] X. Liu and G. Liang, Image Classification of Nutrient Deficiencies Technical Report, 2023.

[22] R.S Kuzu, V. Jancauskas, A. Zappacosta, and S. Ofori-Ampofo Classification of Nutrient Deficiencies in Winter Wheat and Winter Rye, 2023.