

Data Augmentation and Convolutional Network Architecture Influence on Distributed Learning

Influência do Aumento de Dados e da Arquitetura de Redes Convolucionais no Aprendizado Distribuído

Victor Forattini Jansen¹, Emanuel Teixeira Martins¹, Yasmin Souza Lima¹, Flávio de Oliveira Silva^{2,3}, Rodrigo Moreira¹, Larissa Ferreira Rodrigues Moreira^{1*}

Abstract: Convolutional Neural Networks (CNNs) have proven to be highly effective in solving a broad spectrum of computer vision tasks, such as classification, identification, and segmentation. These methods can be deployed in both centralized and distributed environments, depending on the computational demands of the task. While much of the literature has focused on the explainability of CNNs, which is essential for building trust and confidence in their predictions, there remains a gap in understanding their impact on computational resources, particularly in distributed training contexts. In this study, we analyze how CNN architectures primarily influence model accuracy and investigate additional factors that affect computational efficiency in distributed systems. Our findings contribute valuable insights for optimizing the deployment of CNNs in resource-intensive scenarios, paving the way for further exploration of variables critical to distributed learning.

Keywords: Distributed Learning — Rice Classification — Data Augmentation — CNN — Factorial Design.

Resumo: As Redes Neurais Convolucionais (CNNs) têm se mostrado altamente eficazes na solução de uma ampla gama de tarefas de visão computacional, como classificação, identificação e segmentação. Esses métodos podem ser implementados tanto em ambientes centralizados quanto distribuídos, dependendo das exigências computacionais da tarefa. Embora grande parte da literatura tenha se concentrado na explicabilidade das CNNs, o que é essencial para gerar confiança em suas previsões, ainda há uma lacuna no entendimento de seu impacto nos recursos computacionais, especialmente em contextos de treinamento distribuído. Neste estudo, analisamos como as arquiteturas de CNN influenciam principalmente a precisão do modelo e investigamos fatores adicionais que afetam a eficiência computacional em sistemas distribuídos. Nossos resultados oferecem insights valiosos para otimizar a implementação de CNNs em cenários com alta demanda de recursos, abrindo caminho para a exploração de variáveis críticas para o aprendizado distribuído.

Palavras-Chave: Aprendizado Distribuído — Classificação de Arroz — Aumento de Dados — CNN — Análise Fatorial.

¹ Institute of Exact and Technological Sciences, Federal University of Viçosa (UFV), Brazil

² Faculty of Computing (FACOM), Federal University of Uberlândia (UFU), Brazil

³ Department of Informatics, University of Minho (UMINHO), Portugal

*Corresponding authors: {victor.jansen, emanuel.martins, yasmin.lima, rodrigo, larissa.f.rodrigues}@ufv.br, flavio@di.uminho.pt

DOI: <http://dx.doi.org/10.22456/2175-2745.143508> • Received: 23/10/2024 • Accepted: 08/12/2024

CC BY-NC-ND 4.0 - This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

1. Introduction

The increasing complexity of deep learning models, particularly Convolutional Neural Networks (CNNs), has heightened the demand for optimizing distributed training methods. Distributed training is essential for handling

large datasets and deep architectures efficiently, enabling better utilization of computational resources across multiple machines. This study focuses on evaluating how key factors — such as network depth and data augmentation (DA) — affect both performance and resource utilization

in distributed environments [1].

Managing computational resources efficiently is critical during the training phase, as both time and resources are valuable, especially with complex architectures. The selection of a CNN architecture has a direct influence on hardware usage metrics [2]. For example, the number of layers and the complexity of numerical computations in the network significantly affect GPU, CPU usage, memory consumption, and network traffic [3]. However, the specific impact of these factors in a distributed setting has not been fully explored. Despite there are efforts towards employed Convolutional Neural Network (CNN) to enrich disease classification on images, there is a lack of studies which dive into CNN hardware demands footprint.

This paper examines the impact of factors such as CNN architecture and the introduction of Data Augmentation (DA) on various response variables, including Graphics Processing Unit (GPU), Central Processing Unit (CPU) usage, memory consumption, network packets, and accuracy. To further explore these relationships, we conducted experiments based on a factorial design within a distributed training environment, offering novel insights into how CNN architectures and DA influence the aforementioned metrics – an aspect not yet extensively explored in the state of the art. While our experiments use a rice disease classification dataset as a case study – because of the quality of the images and the diseases classification, the primary objective is to understanding the broader impact of CNN architecture and DA in distributed training scenarios [4].

The remainder of this paper is organized as follows. Section 2 surveys the related work. Section 3 presents our proposed approach to explore data augmentation and deep learning architecture influence on distributed learning. Section 4 describes our results and discusses the main findings. Finally, Section 5 presents the conclusion and opportunities for future work.

2. Related Work

In recent years, significant advances have been made worldwide in applying deep learning and computer vision techniques to security[5] and health [6,7] agricultural crops [8], particularly for classifying diseases in plant leaves. This section reviews recent advancements in computer vision techniques to address this challenge by focusing on the impacting resources. In the state-of-the-art, there are limited approaches that measure the impact that Convolutional Neural Networks (CNNs) and DA techniques impose on the underlying hardware, particularly from the point of view of network consumption [9, 10].

Petchiammal et al. [4] developed the Paddy Doctor dataset, which contains 16.225 classified images across 12 disease labels and one healthy label. The photos were

collected from real paddy fields, annotated under professional agricultural supervision, and used as a benchmark in different CNN models, such as VGG16, MobileNet, Xception, and Resnet34, with data augmentation (rotation, shear intensity, zoom, width, and height shift, and horizontal flip). In our study, we use this pre-classified dataset to evaluate the impacts of distributed learning on the classification of rice leaf disease.

Aggarwal et al. [11] proposed a federated transfer learning (F-TL) framework to address the challenges of rice-leaf disease classification, particularly in scenarios where data privacy and distributed data sources are concerned. The study presents an unbalanced dataset with four diseases that were separated on both IID (independent and identically distributed) and non-IID (non-identically distributed) and trained in models such as EfficientNetB3 and MobileNetV2, demonstrating strong performance across IID and non-IID datasets.

Ni et. al [12] introduced an improved model based on the RepVGG architecture, integrated with the Efficient Channel Attention (ECA) mechanism. The proposed model, RepVGG-ECA, is designed to improve classification through the attention mechanism. The study used the Paddy Doctor dataset under data augmentation such as inversion, saturation modification, contrast, and adding blur to focus on the diseases. Comparing the accuracy, macro-f1, macro-precision and macro-recall, they found out that their introduced model achieved 97.06% accuracy, outperforming other models such as ResNet34 and ShuffleNetV2.

Yang et al. [13] proposed the DGLNet, a lightweight network specifically designed for identifying rice diseases. Their network integrates a Global Attention Module (GAM) - which is engineered to capture critical information in complex and/or noisy environments - and a Dynamic Representation Module (DRM), that is designed to improve feature representation through a self-developed for-dimensional flexible convolution (4D-FConv). The study used two datasets, including the Paddy Doctor, achieving 99.71% of accuracy.

Senthly et al. [14] compared the performance of 11 deep CNN and 2 shallow CNN models with different architectures for disease detection. In study, deep features extracted from these models were combined with transfer learning and an SVM for classification tasks. The ResNet50 model, when paired with SVM, achieved the highest accuracy and F1 score of 98.38%. Similarly, the SVM classifier, when using features obtained from the AlexNet CNN, reached an accuracy of 96.8%. When analyzing shallow CNN architecture as MobileNetV2, the results was comparable to ResNet50, achieving 97.96% of accuracy.

While previous works have primarily focused on evaluating models to enhance classification through computer

vision, many have not sufficiently explored how these models impact hardware usage. To fill this gap, our paper studies the interplay between network depth and distributed training, analyzing hardware and training metrics to determine how the CNN architecture and data augmentation impact the process.

3. Influence Assessment Method

In this study, we employed a 2^2 factorial design, a specific type of Analysis of Variance (ANOVA), to systematically investigate the effects of two independent factors on the performance of CNN in a classification task. According to Table 1, the factors considered were DA and CNN Architectures, each evaluated at two levels: with-DA (1) and without-DA (-1) for the Data Augmentation factor, and shallow-CNN (1) and deep-CNN (-1) for the CNN Architectures factor. The factorial design allows for the examination of the main effects of each factor and the interaction effect between the factors.

Table 1. Factorial Design.

Factors	Levels	Response Variable				
		GPU	Network Packets	CPU	Memory Consumption	Accuracy
A: Data Augmentation	with-DA (1)	y ₁	y ₂	y ₃	y ₄	y ₅
	without-DA (-1)					
B: CNN Architecture	shallow-CNN (1)					
	deep-CNN (-1)					

This study aims to comprehend the scientific basis for the impact of two key elements, DA and CNN architecture, on the hardware used in distributed learning, particularly in quantitative analysis and the interplay between these two components. Given the above problem statement, we formulated the factorial experiment combinations according to Table 1. In our experimental design, we identified six (6) response variables that enable us to assess the impact of hardware on distributed learning environments as in Table 1. For shallow-CNN, we refer to those with few convolutional layers, whereas deep-CNN refers to a larger number of layers.

Mathematically, the experiment can be modeled using the following linear model:

$$Y = \mu + \tau_A + \tau_B + \tau_{AB} + \varepsilon$$

where Y represents the observed outcome (e.g., accuracy or F1-Score), μ is the overall mean, τ_A and τ_B are the main effects of DA and CNN Architectures, respectively, τ_{AB} is the interaction effect between the two factors, and ε represents the random error term. The main effects and interaction influence are estimated using the following contrasts:

$$\tau_A = \frac{1}{2} [(Y_{A_1B_1} + Y_{A_1B_0}) - (Y_{A_0B_1} + Y_{A_0B_0})]$$

$$\tau_B = \frac{1}{2} [(Y_{A_1B_1} + Y_{A_0B_1}) - (Y_{A_1B_0} + Y_{A_0B_0})]$$

$$\tau_{AB} = \frac{1}{2} [(Y_{A_1B_1} - Y_{A_1B_0}) - (Y_{A_0B_1} - Y_{A_0B_0})]$$

Our design provides a comprehensive understanding of how the different levels of DA and CNN Architectures influence the distributed learning hardware, individually and in combination. The interaction term, in particular, reveals whether the effect of one factor depends on the level of the other factor. In Fig. 1, we present our method to measure our factor influence on hardware in a distributed training setup.

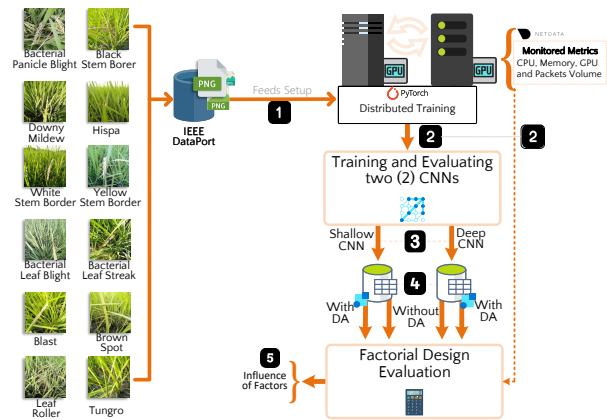


Figure 1. Proposed method.

The Paddy Doctor dataset, depicted in left side of Fig. 1, is a visual image dataset designed for automated paddy disease classification. It comprises 16,225 annotated images of paddy leaves, categorized into 13 classes: 12 diseases (e.g., Bacterial Leaf Blight, Brown Spot, Tungro) and normal leaves. Collected from real paddy fields and annotated by an agricultural officer, this dataset is valuable for developing and benchmarking deep learning models for paddy disease detection [4].

According to Fig. 1, step one (1) of our method is to feed our distributed training environment with dataset images. In the step two (2), we have parallel procedures flow, while our distributed testbed carried out training over data, we measured and collect resources metrics to measure the the impact. We employed the monitoring tool NetData to record the hardware consumption of the CPU, memory, network packets, and GPU. Step three (3) refers to the experiment factors combinations with shallow-CNN (MobileNetV2-100 with bath normalization) and deep-CNN (MobileOne-S1). Step four (4) refers to the level combinations regarding the use of DA or not.

Our employed data augmentation pipeline (step 4 of Fig. 1) includes random rotations ($\pm 5^\circ$), affine transformations with shear (0.2) and slight translations ($\pm 5\%$), random resized cropping (80-100%) of the original image size), horizontal flipping, and color jittering. These

enhancements may improve model generalization by introducing variability in spatial orientation, scale, color, and positioning while preserving the essential features of the images. In step five (5) we carried out ANOVA to assess the factors and levels on different variables of response.

4. Evaluation and Discussion

We conducted experiments using the hardware specifications listed in Table 2 to measure the factors influencing hardware resources in a distributed training scenario on top of Ubuntu 20.04 Long Term Evolution (LTE). Specifically, our distributed learning environment consists of two servers exchanging training gradients and CNN weights over a Local Area Network (LAN) with a 1 Gbps interface, which was negotiated to operate at 100 Mbps. We employed the Torch Distributed Data Parallel as the backend of distributed training.

Table 2. Experimental Testbed.

Server	CPU	RAM	GPU	Ethernet
#1	Intel(R) Core(TM) i5-4430 3.00GHz	32 GB	GeForce RTX 4060 Ti 8 GB	1Gbps
#2	Intel(R) Core(TM) i5-4430 CPU 3.00GHz	16 GB	GeForce GTX 1050 Ti 4 GB	1Gbps

Considering the training and validation performance, we computed the accuracy and loss graphs for both the training and testing phases in our distributed setup. As shown in Fig.2 (a), (b), (c), and (d), empirical evidence suggests that for all experiment combinations, there was learning progress over the epochs in our proposed evaluation of the impact of factors such as DA and CNN architecture, specifically on Server #1. Although we standardized the number of epochs to 100 in our distributed scenario, for graphical presentation, we applied early stopping in the training process, as shown in Fig.2(c), and 2(d), concluding at epoch 20 due to the absence of further improvements.

Similarly, we assessed the behavioral performance of the model trained on Server #2. Empirical evidence also showed that the model used in our performance evaluation was able to learn correctly over epochs. This is evident in Fig.2 (e), (f), (g), and (h), where all four experimental combinations showed a decrease in loss with exponential decay and a sigmoid increase in accuracy.

Regarding the central objective of our evaluation, we conducted a factorial analysis based on the ANOVA method considering the experimental variations in Table 1. The results of the experiments are reported in Table 3, showing the average measurements for each response variable. We assessed the average (%) consumption of Y_{GPU} , $Y_{NetworkPackets}$, Y_{CPU} , Y_{Memory} , and $Y_{Accuracy}$ for the four experimental combinations. Factors such as Y_{Memory} , although there were experimental variations, maintained its

average consumption.

Table 3. Average of Response Variable in Factorial Experiments.

Experiment	Response Variable				
	Y_{GPU} (%)	$Y_{NetworkPackets}$ (Pkts/s)	Y_{CPU} (%)	Y_{Memory} (%)	$Y_{Accuracy}$ (%)
with-DA (1) and shallow-CNN (1)	95,12	19994,50	51,15	81,70	98,71
without-DA (-1) and shallow-CNN (1)	97,18	15698,97	47,38	81,75	99,60
with-DA (1) and deep-CNN (-1)	97,21	19973,00	47,03	80,45	94,09
without-DA (-1) and deep-CNN (-1)	98,29	10526,36	43,85	81,45	96,58

On the other hand, the response variable $Y_{NetworkPackets}$ exhibited non-uniform behavior. We found that the experimental combination with-DA imposes a higher volume of network packets in distributed training. Fig. 3 shows the grouped violin plot of two categorical variables: with-DA and without-DA. As observed, for both neural network architectures (deep and shallow), the introduction of DA results in a higher volume of network packets, as indicated by the accumulation of instances in the plot.

As shown in Table 4 we advanced in our analysis by evaluating the influence of the factors DA (τ_A) and CNN architecture (τ_B) both individually and in combination (τ_{AB}) on the response variables. The factor τ_B (CNN architecture) has an influence of 48.64% on the response variable Y_{GPU} , followed by an influence of 46.83% from the factor τ_A . Regarding the response variable $Y_{NetworkPackets}$, we validated the initial hypothesis regarding the influence of introducing DA in training. We found that introducing DA in distributed learning results in a 77.92% influence on the volume of packets transmitted during deployment.

Table 4. Comparison of the influence of factors on different response variables.

Factors Iteration	Influence Measurement (%)				
	Y_{GPU}	$Y_{NetworkPackets}$	Y_{CPU}	Y_{Memory}	$Y_{Accuracy}$
τ_A	46,83	77,92	45,07	24,53	15,86
τ_B	48,64	11,13	54,61	54,75	80,60
τ_{AB}	4,53	10,94	0,32	20,72	3,54

In distributed training with , increasing the data volume led to a 27.37% rise in network packet transmission with data augmentation (DA) and a shallow CNN, and an 89.73% increase with DA and a deep CNN. This rise in network packet transmission is due to the need for frequent gradient synchronization across GPUs. Differences in GPU performance, such as between the RTX 1080 and RTX 4060, can further amplify this communication load, as the faster GPU must wait for the slower one to synchronize.

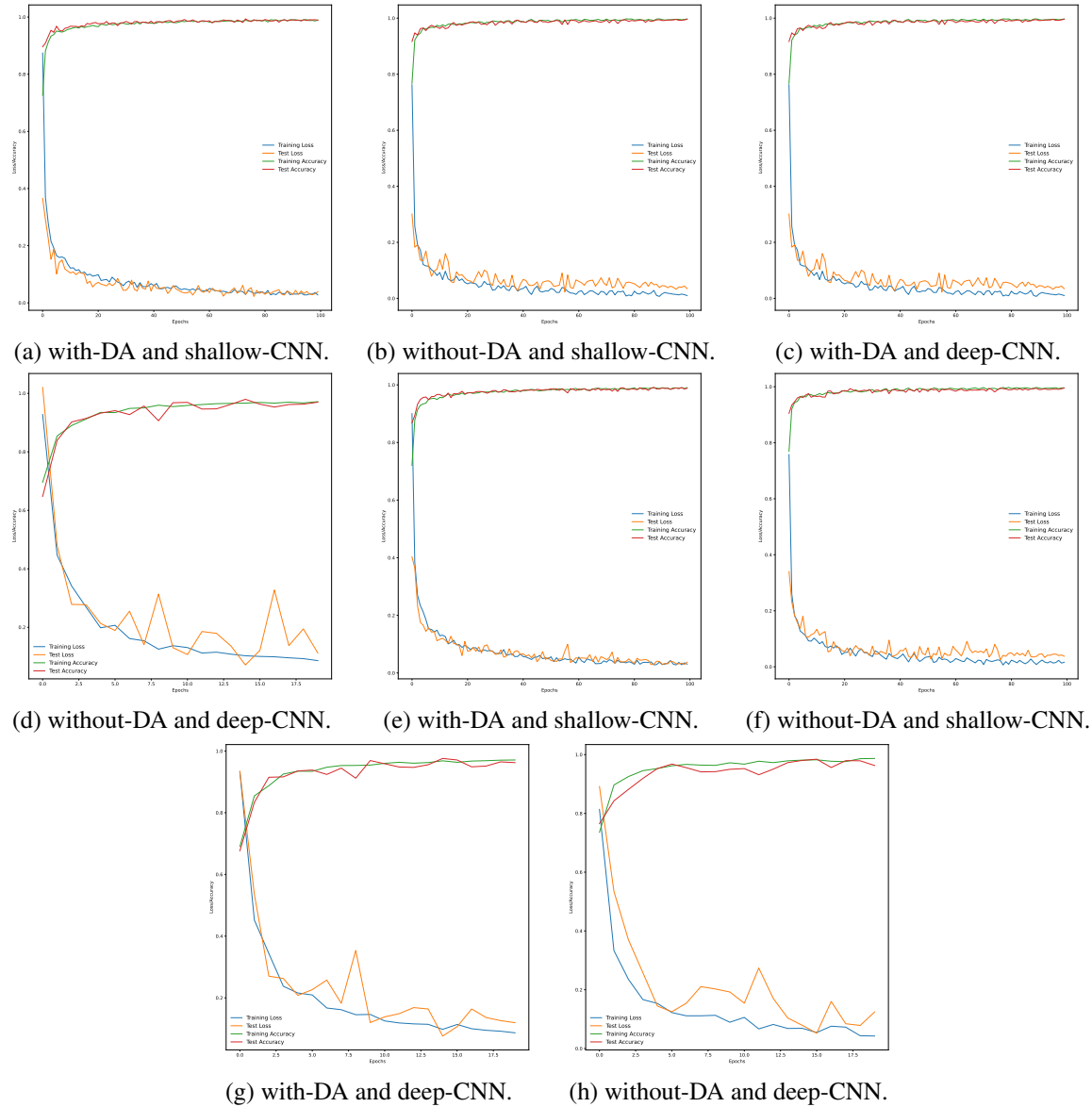


Figure 2. Training and Testing Metrics for Server #1 (a), (b), (c), (d) and Server #2 (e), (f), (g), (h).

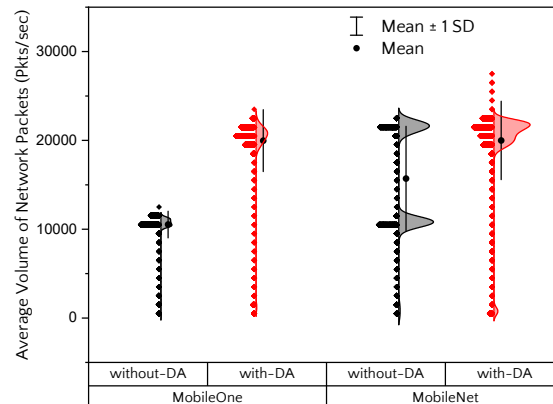


Figure 3. DA is affecting the distribution of average packet volume.

Limitations. The inclusion of other data augmentation methods may change the volume of network packet transmission in distributed training, potentially affecting scalability and efficiency. Similarly, the early stopping applied during training may have restricted further insights into long-term model performance trends, limiting the comprehensive evaluation across all epochs.

5. Concluding Remarks

This paper evaluated the impact of distributed learning on different response variables using CNN and a computer vision dataset of rice diseases. Understanding these behavioral nuances is crucial, as it provides reliability and predictability when deploying such technologies in production environments. We observed that, while state-of-the-art research focuses on model explainability, it often overlooks the impact of these models on different response variables, especially in the context of computer vision.

Among our findings, we assessed that introducing DA in distributed training has a significant effect on the underlying infrastructure. For future work, we are developing analyses that consider other factors, such as CNN parameters and different datasets. This study offers valuable insights for the deployment phases of CNN models in real-world environments, where neglecting these effects can significantly impact energy consumption and data center performance.

Acknowledgments

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001. Rodrigo Moreira gratefully acknowledges the financial support of FAPEMIG (Grant #APQ00923-24). We also acknowledge the financial support of the FAPESP MCTIC/CGI Research project 2018/23097-3 - SFI2 - Slicing Future Internet Infrastructures.

Author contributions

VFJ, ETM, and YSL: Software, Investigation, Validation, Visualization, Writing - original draft. FOS, LFRM, and RM: Conceptualization, Methodology, Visualization, Supervision, Funding acquisition, Resources, Writing - Review & Editing, and Project administration.

References

[1] SONG, J. et al. Dima-cnn: Deep learning toolkit-accelerated analytical modeling and optimization of cnn hardware and dataflow. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, p. 1–1, 2024.

[2] Rodrigues Moreira, L. F. et al. An Artificial Intelligence-as-a-Service Architecture for deep learning model embodiment on low-cost devices: A case study of COVID-19 diagnosis. *Applied Soft Computing*, v. 134, p. 110014, 2023. ISSN 1568-4946.

[3] WIEDEMANN, S.; MÜLLER, K.-R.; SAMEK, W. Compact and computationally efficient representation of deep neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, v. 31, n. 3, p. 772–785, 2020.

[4] PETCHIAMMAL et al. Paddy doctor: A visual image dataset for automated paddy disease classification and benchmarking. In: *Proceedings of the 6th Joint International Conference on Data Science & Management of Data (10th ACM IKDD CODS and 28th COMAD)*. New York, NY, USA: Association for Computing Machinery, 2023. (CODS-COMAD '23), p. 203–207. ISBN 9781450397971. Disponível em: <<https://doi.org/10.1145/3570991.3570994>>.

[5] COSTA, F. et al. Self-portrait to id document face matching: Cnn-based face verification in cross-domain scenario. In: *Anais do XVII Workshop de Visão Computacional*. Porto Alegre, RS, Brasil: SBC, 2021. p. 31–36. ISSN 0000-0000. Disponível em: <<https://sol.sbc.org.br/index.php/wvc/article/view/18885>>.

[6] ROCHA, E. L. da; RODRIGUES, L.; MARI, J. F. Maize leaf disease classification using convolutional neural networks and hyperparameter optimization. In: SBC. *Anais do XVI Workshop de Visão Computacional*. [S.l.], 2020. p. 104–110.

[7] BOUZON, M. et al. Automatic kidney stone detection using low-cost cnn with coronal ct images. In: *Anais do XVIII Workshop de Visão Computacional*. Porto Alegre, RS, Brasil: SBC, 2023. p. 24–29. ISSN 0000-0000. Disponível em: <<https://sol.sbc.org.br/index.php/wvc/article/view/27527>>.

[8] MOREIRA, R. et al. Agrolens: A low-cost and green-friendly smart farm architecture to support real-time leaf disease diagnostics. *Internet of Things*, v. 19, p. 100570, 2022. ISSN 2542-6605.

[9] PARK, S.; LEE, J.; KIM, H. Hardware resource analysis in distributed training with edge devices. *Electronics*, v. 9, n. 1, 2020. ISSN 2079-9292. Disponível em: <<https://www.mdpi.com/2079-9292/9/1/28>>.

[10] MOHAIDAT, T.; KHALIL, K. A survey on neural network hardware accelerators. *IEEE Transactions on Artificial Intelligence*, v. 5, n. 8, p. 3801–3822, 2024.

[11] AGGARWAL, M. et al. Federated transfer learning for rice-leaf disease classification across multiclient cross-silo datasets. *Agronomy*, v. 13,

n. 10, 2023. ISSN 2073-4395. Disponível em:
<<https://www.mdpi.com/2073-4395/13/10/2483>>.

[12] NI, H. et al. Classification of typical pests and diseases of rice based on the eca attention mechanism. *Agriculture*, v. 13, n. 5, 2023. ISSN 2077-0472. Disponível em: <<https://www.mdpi.com/2077-0472/13/5/1066>>.

[13] YANG, Y. et al. A lightweight rice disease identification network based on attention mechanism and dynamic convolution. *Ecological Informatics*, v. 78,

p. 102320, 2023. ISSN 1574-9541. Disponível em:
<<https://www.sciencedirect.com/science/article/pii/S1574954123003497>>.

[14] SETHY, P. K. et al. Deep feature based rice leaf disease identification using support vector machine. *Computers and Electronics in Agriculture*, v. 175, p. 105527, 2020. ISSN 0168-1699. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0168169919326997>>.