Parameter-Efficient Coffee Leaf Disease Classification Using Quaternion-Valued EfficientNets

André Y. S. Okimoto[1], Arnaldo Matuchewski[1], Allan Junior[1], Hendrio Braganca[1], Italo Pelição Caliari[1], Antonio Fontenele[1], Wington L. Vital[1]

[1]Núcleo de Capacitação em Inteligência Artificial (NCIA), Fundação Paulo Feitoza - FPFtech, Av. Danilo de Matos Areosa, 1170, Distrito Industrial, Manaus, Amazonas, 69075-351, Brazil.

{andre.okimoto, arnaldo.252334, allan.252317, hendrio.braganca, italo.caliari, antonio.fontenele, wington.vital}@fpf.br

Abstract. *Deep Learning has enabled the automated identification of biotic stresses in coffee crops; however, conventional Convolutional Neural Networks (CNNs) process RGB channels as independent feature maps, potentially overlooking cross-channel chromatic relationships. In this work, we investigate a quaternion-valued EfficientNet-B0 (Q-EfficientNet) for coffee leaf disease classification using the Symptom dataset derived from BRACOL. The proposed architecture represents color information in a unified hypercomplex domain and implements quaternion-valued depthwise operations through Kronecker-product formulations, allowing efficient integration within standard deep learning frameworks. Experimental results show that the proposed model achieves competitive classification performance while requiring only 1.7 million trainable parameters, compared to 5.9 million parameters for its real-valued counterpart. These findings indicate that quaternion-valued representations can provide a compact and effective alternative for modeling multidimensional color information in agricultural computer vision tasks, highlighting their potential for deployment in resource-constrained environments.*

1. Introduction

Coffee cultivation is one of the pillars of the global agricultural economy, requiring increasingly precise methods for monitoring crop health and productivity. In recent years, Deep Learning techniques have been successfully applied to the automated identification of biotic stresses and maturation in coffee crops [Esgario et al. 2020]. However, the processing of color images through traditional Convolutional Neural Networks (CNNs) often neglects the intrinsic correlation between color channels (RGB), treating them as independent components in the real number domain. This simplification can result in a loss of structural and chromatic information essential for the accurate classification of leaf pathologies.

To overcome this limitation, the field of hypercomplex neural networks is a possibility, utilizing higher-dimensional algebras to represent multidimensional data in a more integrated manner. Quaternion-based neural networks, for instance, allow each pixel to be treated as a single four-dimensional vector, preserving phase and color relationships more efficiently than real-valued models. The development of Vector-Valued EfficientNets (V-EfficientNets) has yielded significant results in medical datasets [Neto and Valle 2025a].

Nonetheless, the application of these architectures to computer vision problems in other fields, such as agriculture, remains largely unexplored.

This paper proposes a study of a real-valued (traditional) and a quaternion-valued EfficientNet applied to an agricultural image classification task. Unlike previous works focused on medical datasets, this study demonstrates that quaternion-valued EfficientNets can provide superior performance and parameter efficiency in agricultural scenarios, even when compared against transfer learning-based real-valued models. Quaternion-valued neural networks can achieve comparable or superior performance while requiring fewer parameters [Parcollet et al. 2020]. The study employs the Symptom Dataset, composed of regions of interest extracted from the Brazilian Arabica Coffee Leaves (BRACOL) dataset. Experimental results show that the quaternion-valued EfficientNet outperforms the traditional model across all evaluated metrics (accuracy, precision, recall, and F1-score), despite the real-valued model being trained using transfer learning.

However, important differences between medical and agricultural imaging must be considered when evaluating the applicability of quaternion-valued architectures across these domains. The transition of hypercomplex models from the medical domain to agriculture is non-trivial due to substantial environmental and biological differences. While medical color variations often arise from controlled imaging conditions and artificial staining procedures, agricultural chromatic variations reflect dynamic physiological responses to biotic stresses, resulting in complex non-linear patterns. To address these challenges, adapting the color representation to the quaternion domain becomes essential. In particular, specialized HSV-based mappings can decouple chromatic information from illumination intensity, enabling a more meaningful quaternion encoding. This tailored representation allows the architecture to efficiently capture complex chromatic variations associated with biotic stresses while maintaining robustness to environmental noise, thereby motivating an independent validation in agricultural scenarios.

2. Related Work

The paper [Rodrigues and Valle 2025] presents a study on Vector-Valued Neural Networks (V-nets), explaining their mathematical foundations and implementing the architecture. The problem selected to evaluate the model was a medical classification dataset for detecting acute lymphoblastic leukemia. The work highlights the superior performance of quaternion- and octonion-based neural networks compared to the real-valued model. Along the same lines, the researchers in [Neto and Valle 2025a] proposed the V-EfficientNet models explored in this work. These models were evaluated using the same medical classification dataset. In the comparison between the different algebras used in V-EfficientNets, the research showed better performance for quaternion and hyperbolic quaternion algebras, outperforming the real-valued approach.

The concept of algebraic flexibility is explored by [Grassucci et al. 2024] with the introduction of Parameterized Hypercomplex Neural Networks (PHNNs). By utilizing Kronecker products to learn interaction rules between channels directly from the data, PHNNs allow operation in arbitrary n-dimensional domains. This approach enables the processing of RGB images in their natural domain ($n = 3$) with a 66% reduction in parameters, overcoming the limitations of traditional quaternion networks that require padding extra channels.

Expanding this discussion to other architectures and algebras, [Vieira and Valle 2022] established a general framework for hypercomplex Extreme Learning Machines (ELMs). The study demonstrated that models based on Cayley-Dickson algebras can outperform both real-valued and quaternion-based models in color image processing tasks. This generalization reinforces the effectiveness of higher-dimensional number systems in capturing correlations between color channels.

A major limitation in the existing literature is that standard real-valued architectures treat color channels (RGB) as independent feature maps, failing to capture the intrinsic cross-channel correlations that characterize plant disease manifestations. While some recent studies have explored hypercomplex neural networks to preserve these chromatic interactions, they predominantly rely on dense, computationally heavy architectures designed for high-performance medical servers [Vieira and Valle 2022]. Consequently, there is a critical gap regarding efficient hypercomplex models suitable for resource-constrained edge devices in agriculture. Current literature lacks architectures that simultaneously leverage the algebraic benefits of quaternion representations for multidimensional color features and the structural efficiency required for practical field deployment under environmental noise. This work directly addresses this limitation by developing and validating a lightweight, quaternion-valued depthwise separable framework.

Although previous works on BRACOL-derived datasets have explored color spaces such as HSV to improve classification performance, these approaches typically treat color channels independently or in their raw scalar form. In contrast, the proposed method introduces a structured embedding of the HSV space into a quaternion-compatible representation, enabling a unified modeling of chromatic dependencies.

3. Quaternion algebra

Quaternions are a 4D hypercomplex algebra denoted by $\mathbb{Q}$. The elements of a quaternion can be written as

$$\boldsymbol{x} = x_0\boldsymbol{e}_0 + x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + x_3\boldsymbol{e}_3, \tag{1}$$

where $\boldsymbol{e}_0 \equiv 1_{\mathbb{Q}}$ acts as the scalar unit, and $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, $\boldsymbol{e}_3$ are the hyperimaginary units. The quaternion product is associative and anticommutative, and it is particularly useful to describe rotations in the 3D Euclidean space $\mathbb{R}^3$. Formally, we have

$$\boldsymbol{e}_1^2 = \boldsymbol{e}_2^2 = \boldsymbol{e}_3^2 = -1, \qquad \boldsymbol{e}_0\boldsymbol{e}_2 = \boldsymbol{e}_3, \qquad \boldsymbol{e}_2\boldsymbol{e}_1 = -\boldsymbol{e}_3. \tag{2}$$

In addition to the related work in section 2 quaternions are among the most well-known hypercomplex algebras and have long been studied in the context of signal processing and multidimensional data representation. Early investigations into neural models based on quaternion algebra were already explored for signal processing applications [Sangwine 1996]. In recent years, the growing development of deep learning methods has renewed interest in quaternion-valued neural architectures. In particular, quaternion convolutional neural networks have been proposed to efficiently model multidimensional signals, demonstrating promising performance in applications such as automatic speech recognition [Parcollet et al. 2019]. Furthermore, quaternion neural models have also been successfully applied in computer vision tasks, including color image classification, where the RGB channels can be naturally encoded as a single quaternion-valued

representation, allowing the network to capture inter-channel correlations more effectively [Yang et al. 2019].

3.1. The Kronecker Product

To improve computational efficiency, the matrix $M_L(A) \in \mathbb{R}^{nM\times nL}$ may be evaluated through the use of the Kronecker product. For two real matrices $A = (a_{ij}) \in \mathbb{R}^{N\times M}$ and $B = (b_{ij}) \in \mathbb{R}^{P\times Q}$, their Kronecker product is defined as

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1M}B \\ \vdots & \ddots & \vdots \\ a_{N1}B & \cdots & a_{NM}B \end{bmatrix} \in \mathbb{R}^{NP\times MQ}. \tag{19}$$

By exploiting this operation, one can express a wide range of algebraic and geometric transformations in a unified framework, thereby enabling a more compact representation of complex interactions and simplifying the analysis of their underlying structural properties

$$M_L(A) = \sum_{k=1}^{n} \begin{bmatrix} a_{11k}P_k^T & \cdots & a_{1Lk}P_k^T \\ \vdots & \ddots & \vdots \\ a_{M1k}P_k^T & \cdots & a_{MLk}P_k^T \end{bmatrix} = \sum_{k=1}^{n} A_k \otimes P_k^T, \tag{20}$$

where the decomposition $A = \sum_{k=1}^{n} A_k e_k$ is considered, with each matrix $A_k \in \mathbb{R}^{M\times L}$ for $k = 1, \ldots, n$. Consequently, the vector-valued matrix multiplication $C = AB$ can be rewritten as

$$C = \varphi^{-1}\left(\left[\sum_{k=1}^{n} A_k \otimes P_k^T\right]\varphi(B)\right). \tag{21}$$

In the construction above, the general case of an algebra was considered; however, in the case of quaternion algebras, the matrices A_k and P_k^T are chosen as follows.

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{A_1} \otimes \underbrace{[P_x]}_{s_1} \underbrace{\begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{A_2} \otimes \underbrace{[P_x]}_{s_2} \underbrace{\begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix}}_{A_3} \otimes \underbrace{[P_y]}_{s_3} \underbrace{\begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}}_{A_4} \otimes \underbrace{[P_z]}_{s_4}.$$

These choices reflect the intrinsic structure of quaternion multiplication, allowing the algebra to be represented faithfully in matrix form, which will be used in the dense and convolutional layers of the EfficientNet-B0 architecture.

4. EfficientNet Models

The first generation of EfficientNet models, referred to in this work as EfficientNetV1, was introduced by Tan and Le at ICML 2019 [Tan and Le 2019]. This family is built upon the baseline architecture EfficientNet-B0, which was obtained through a neural architecture search procedure aimed at achieving a favorable balance between predictive accuracy and computational cost. The larger variants, EfficientNet-B1 through EfficientNet-B7, are derived from this baseline model using a compound scaling approach that systematically expands the network dimensions.

The main building block of the EfficientNetV1 models is the mobile inverted bottleneck (MBConv) block. The MBConv, which includes a depthwise convolution operation, was initially proposed as part of the MobileNetV2 architecture and also features a squeeze and-excitation block [Hu et al. 2018]

Tan and Le have improved the EfficientNetV1 models at ICML 2021 [Tan and Le 2021]. The improved EfficientNet models, titled EfficientNetV2, have a faster training speed and better pa rameter efficiency than the previous models. The baseline EfficientNetV2 model has also been designed using neural ar chitecture search, but jointly optimizing performance, training speed, and parameter size. Moreover, besides the MBConv block, the search space has been enriched with additional operations, including the Fused-MBConv block.

The Fused-MBConv block, proposed by Gupta and Tan [Gupta and Tan 2019], replaces the depthwise and expansion convolutions of the original MBConv block with a single standard convolutional layer. As observed by Tan and Le, depthwise convolutions tend to be computationally expensive in the early stages of a network, while remaining highly effective in deeper layers. Consequently, EfficientNetV2 employs Fused-MBConv blocks in the initial layers and standard MBConv blocks in the later stages of the architecture.

Inspired by the efficiency and compound scaling strategy of EfficientNets, which balances network width, depth, and input resolution, this work investigates a quaternion-valued version of this architecture for coffee leaf disease classification. Unlike conventional real-valued neural networks, where inter-channel correlations are learned implicitly during training, quaternion-valued representations model color channels jointly as a single hypercomplex entity, naturally preserving their relationships throughout the network. Following the motivation of exploiting multidimensional data in a more structured manner, as discussed in [Neto and Valle 2025b], we employ a Quaternion-EfficientNet to capture inter-channel dependencies more effectively while retaining the parameter efficiency and scalability that have made EfficientNets successful in image classification tasks.

5. Experimental Results

This section presents the evaluation of a real-valued and the proposed quaternion-valued EfficientNet-B0 models using the Symptom dataset derived from the BRACOL database. The models are evaluated in terms of classification performance using standard metrics such as accuracy, precision, recall, and F1-score. Additionally, we analyze their computational complexity and parameter efficiency.

5.1. The Symptom dataset

The dataset used in this study is derived from the BRACOL dataset, which contains images of healthy and diseased leaves captured via smartphones in Espírito Santo, Brazil [Esgario et al. 2020]. Leaf samples collected from the field were photographed in a semi-controlled setting using a white background. The images were captured using several different smartphone models. The specific problem addressed in this work involves cropped leaf images focused on the regions of interest, referred to as the "Symptom dataset".

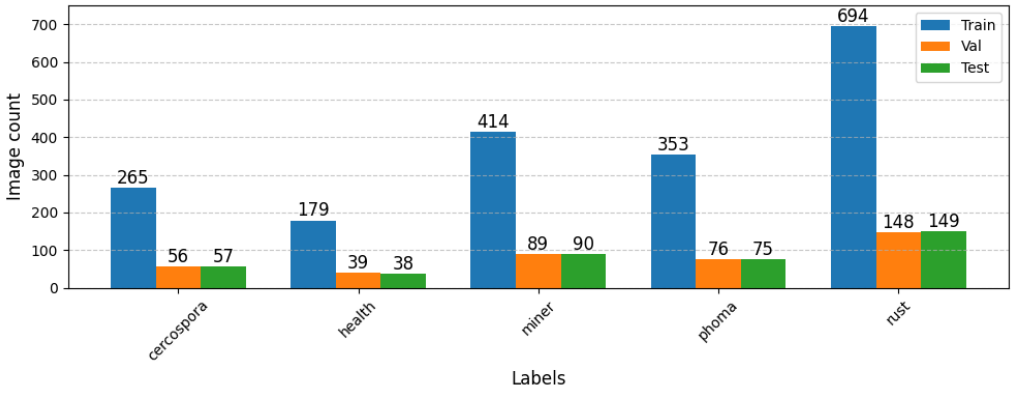


Figure 1. Distribution of the symptom dataset in train, validation and test splits.

The Figure 2 presents examples of the five classes within this dataset (health, leaf miner, rust, brown leaf spot and cercospora), along with a sample image. The dataset consists of 1,905 training images, 408 validation images, and 409 test images. Figure 1 illustrates the class imbalance in the dataset, which was mitigated using the strategies discussed in the following subsection.



(a) (b) (c) (d) (e)

Figure 2. Examples of area of interest in coffee leaves, affected or not by different biotic stresses: health (a), leaf miner (b), rust (c), brown leaf spot (d) and cercospora (e).

5.2. Experimental Setup

To establish a mathematically and methodologically baseline, both the real-valued and quaternion-valued EfficientNet-B0 architectures were trained from scratch using random weight initialization. This approach ensures a fair comparison, particularly since pre-trained weights for the quaternion-valued model are unavailable for transfer learning. All four experimental configurations (real-valued on RGB/HSV and quaternion-valued on RGB/HSV) shared the exact same hyperparameters, optimization settings, and learning rate schedule, as detailed in Table 1.

The models were trained for 100 epochs. To prevent evaluating sub-optimal or overfitted models at the final epoch, a checkpointing callback monitored validation accuracy, saving only the weights from the epoch that achieved the highest validation performance. These optimal weights were subsequently loaded to calculate all evaluation metrics on the test partition. The training pipeline for all models incorporated online data augmentation, consisting of random rotation (15%), translation (10%), and horizontal image flipping. To address dataset class imbalance, class weights calculated dynamically

from the training split distribution were applied to the loss function, encouraging the models to better learn features from minority classes.

Hyperparameter	**Value**
Optimizer	RMSprop
Loss Function	Sparse Categorical Crossentropy
Batch Size	32
Initial Learning Rate	10^{-3}
Momentum	0.9
Rho (ρ)	0.9
Weight Decay (L_2 regularizer)	5×10^{-5}
Learning Rate Scheduler	2.4-epoch linear warmup + exponential decay (decay factor of 0.96 every 2.4 epochs)
Data Augmentation	Random Rotation (15%), Translation (10%), and Horizontal Flip
Training Epochs	100

Table 1. Hyperparameters for EfficientNet-B0 training from scratch.

In this work, different color representations are considered in order to evaluate the influence of input geometry on model performance. In particular, we distinguish between the standard RGB representation, which serves as a baseline sensor-driven encoding, and alternative perceptually and geometrically motivated color spaces, such as HSV, which introduce nonlinear and partially periodic structure into the input space. This distinction is directly reflected in the experimental setup, where both RGB- and HSV-based inputs are evaluated under identical training conditions for real-valued and quaternion-valued EfficientNet-B0 architectures. This modeling choice leads to two distinct preprocessing pipelines, described below.

To use the real-valued EfficientNet-B0, it was only necessary to resize the images to 224×224, as this is the input size chosen for the models. However, the quaternion-valued model requires a different preprocessing pipeline, as described in Figure 3. In addition to resizing the image, the HSV color space was adopted. The image is transformed into four channels according to the representation $(\cos(H)S,\ \cos(H)V,\ \sin(H)S,\ \sin(H)V)$, where (H), (S), and (V) denote the Hue, Saturation, and Value components, respectively. This transformation is motivated by both the geometric structure of the HSV color space and the algebraic structure of quaternion-valued neural networks. Since the Hue component is an angular variable defined on the circle (S^1), its direct use as a scalar variable introduces an artificial discontinuity arising from the identification $0 \equiv 2\pi$. To preserve the circular topology of the hue component, the mapping

$$H \mapsto (\cos(H), \sin(H)) \tag{3}$$

is employed, embedding the hue information into $\mathbb{R}^2$ while preserving its periodic nature. Consequently, chromatically similar colors remain close in the transformed representation, including colors located near the discontinuity of the angular coordinate system.

The transformed representation combines the embedded hue information with the saturation and value components, yielding a four-dimensional real-valued vector. Since

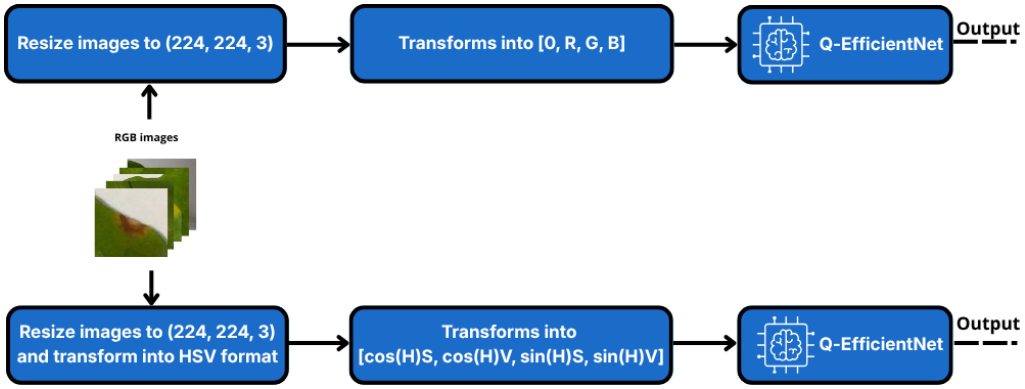


Figure 3. Classification pipeline using the Quaternion-valued model.

quaternion-valued convolutions operate on four-dimensional entities, this representation can be naturally identified with an element of the quaternion algebra. Therefore, the color information is encoded in a format compatible with quaternion-valued processing while preserving the geometric relationships induced by the HSV color space. Furthermore, the transformation can be interpreted as the mapping

$$\Phi : (H, S, V) \longmapsto (\cos(H)S,\ \cos(H)V,\ \sin(H)S,\ \sin(H)V), \tag{4}$$

which provides a continuous representation of the HSV feature space in $\mathbb{R}^4$. Except for degenerate achromatic cases associated with vanishing saturation, the mapping preserves the information contained in the original HSV representation and does not introduce additional features. Rather, Φ reorganizes the same color information into a structure suitable for quaternion-valued computation. Consequently, any performance differences observed in the experiments cannot be attributed to an increase in the amount of information provided to the model. Instead, they reflect the ability of quaternion-valued operations to exploit the correlations among the transformed components through the algebraic structure of quaternion multiplication.

Previous works on coffee leaf analysis have also explored HSV-based representations for preprocessing, segmentation, and mobile disease detection. Esgario et al. used the saturation channel of the HSV color space to support region-of-interest extraction from coffee leaf images [Esgario et al. 2020], while Manso et al. investigated HSV and YCbCr color spaces in a smartphone-based application for coffee leaf rust and leaf miner detection [Manso et al. 2019].

5.3. Results and Discussion

To evaluate the generalization and representation capacity of hypercomplex architectures, the performance of the trained real-valued and quaternion-valued models was assessed on the test split ($N = 409$). Table 2 summarizes the global classification metrics and trainable parameters for all four experimental configurations. The experimental results demonstrate that the quaternion-valued architectures consistently outperform their real-valued counterparts across all color spaces.

Specifically, the quaternion-valued RGB model achieved the highest overall accuracy of **0.9560**, macro F1-score of **0.9501**, and weighted F1-score of **0.9558**, compared to the real-valued RGB model which reached an accuracy of 0.9487. A similar trend

Model Configuration	Parameters	Accuracy	Weighted F1-Score	Macro F1-Score
Real-valued RGB	5.86M	0.9487	0.9482	0.9416
Real-valued HSV	5.86M	0.9487	0.9491	0.9403
Quaternion-valued RGB	**1.72M**	**0.9560**	**0.9558**	**0.9501**
Quaternion-valued HSV	**1.72M**	0.9535	0.9537	0.9477

Table 2. Global performance comparison of real-valued and quaternion-valued EfficientNetV2-B0 models.

is observed in the HSV space, where the quaternion-valued model reached an accuracy of **0.9535** and weighted F1-score of **0.9537**, outperforming the real-valued HSV model which scored 0.9487 and 0.9491, respectively.

Crucially, this performance improvement was achieved with significantly lower model capacity. While the standard real-valued architectures require approximately **5.86 million** trainable parameters, the scaled quaternionic models (using an architecture channel scaling factor of 0.25) require only **1.72 million** parameters—representing a **3.4-fold reduction in parameter count**. This indicates that the algebraic coupling of weight matrices in hypercomplex convolutions acts as a strong regularizer that preserves spatial-color dependencies, making it highly parameter-efficient and less prone to overfitting on small datasets.

A detailed class-by-class analysis for the RGB and HSV networks is shown in Table 3 and Table 4, respectively. Although the real-valued models are capable of identifying healthy leaves with high precision, they show limitations in distinguishing between highly correlated symptoms of disease classes under small data regimes.

	Quaternion-valued RGB			Real-valued RGB			
Class	**Prec.**	**Rec.**	**F1**	**Prec.**	**Rec.**	**F1**	**Support**
Health	0.97	0.97	0.97	1.00	0.97	0.99	38
Miner	0.92	0.98	0.95	0.88	1.00	0.94	90
Rust	0.99	0.97	0.98	0.99	0.99	0.99	149
Phoma	0.96	0.96	0.96	0.96	0.88	0.92	75
Cercospora	0.92	0.86	**0.89**	0.92	0.84	0.88	57
Accuracy	–	–	**0.9560**	–	–	0.9487	409
Macro Avg	0.95	0.95	**0.9501**	0.95	0.94	0.9416	409
Weighted	0.96	0.96	**0.9558**	0.95	0.95	0.9482	409

Table 3. Detailed class-wise metrics comparison in the RGB color space.

The computational complexity of the models was evaluated through the number of floating-point operations (FLOPs). Calculations were performed using the TensorFlow Profiler, considering a single inference pass (forward pass) with the model in evaluation mode. The models were tested with a fixed input of 224×224 pixels, with the distinction that the quaternion-valued model uses four channels, while the real-valued model uses three. The results are shown in Table 5. The real-valued model performs fewer floating-point operations, at 1.4553 GFLOPs per input, compared to 2.1021 GFLOPs for the quaternion model.

This behavior is consistent with recent theoretical results on the computational

	Quaternion-valued HSV			Real-valued HSV			
Class	**Prec.**	**Rec.**	**F1**	**Prec.**	**Rec.**	**F1**	**Support**
Health	1.00	0.97	0.99	0.95	0.97	0.96	38
Miner	0.94	0.92	0.93	0.95	0.93	0.94	90
Rust	0.99	0.99	0.99	0.98	0.98	0.98	149
Phoma	0.92	0.92	0.92	0.97	0.93	0.95	75
Cercospora	0.88	0.93	**0.91**	0.84	0.89	0.86	57
Accuracy	–	–	**0.9535**	–	–	0.9487	409
Macro Avg	0.95	0.95	**0.9477**	0.94	0.94	0.9403	409
Weighted	0.95	0.95	**0.9537**	0.95	0.95	0.9491	409

Table 4. Detailed class-wise metrics comparison in the HSV color space.

complexity of hypercomplex-valued neural networks [Niemczynowicz and Kycia 2026]. In particular, for a hypercomplex algebra of dimension a, the FLOP complexity of dense layers scales as $O(b \cdot a \cdot n \cdot u)$, whereas the corresponding real-valued formulation scales as $O(b \cdot n \cdot u)$. Similarly, hypercomplex-valued convolutional layers exhibit an additional multiplicative factor proportional to a, arising from the algebraic expansion of hypercomplex multiplications into structured real-valued operations and tensor contractions over the algebra dimension.

More precisely, convolutional layers satisfy the complexity relation $O\big(b \cdot a \cdot \prod_{l=1}^{k} L_l \cdot n \cdot F\big)$ for the hypercomplex case, compared to $O\big(b \cdot \prod_{l=1}^{k} L_l \cdot n \cdot F\big)$ in the real-valued setting. In both dense and convolutional layers, the factor a reflects the intrinsic expansion of hypercomplex arithmetic into multiple real-valued operations, which increases the number of underlying computations per parameter.

In the case of quaternion-valued networks ($a = 4$), this algebraic expansion introduces a constant-factor increase in FLOP complexity, leading to higher computational cost despite the reduced number of learnable parameters. Therefore, the observed overhead is not implementation-dependent but a direct consequence of the algebraic structure governing quaternion arithmetic.

This demonstrates how the quaternion model can achieve superior performance in metrics while having fewer parameters $(1.700.000)$ than its real-valued counterpart (5.9 million), despite performing more operations in return. From a theoretical perspective, this trade-off between parameter efficiency and computational cost is explained by the structure of hypercomplex arithmetic, where the computational complexity scales with the dimension of the underlying algebra. In particular, this induces additional real-valued operations per parameter across both dense and convolutional layers, as a consequence of the algebraic expansion of quaternion multiplications. This highlights the need for optimized implementations of hypercomplex-valued networks in order to reduce the computational gap with real-valued models, a direction that has been explored in recent works on alternative hypercomplex algebras such as tessarines [Cariow and Cariowa 2025].

6. Conclusion and future works

This work presents an investigation of quaternion-valued EfficientNets for the classification of biotic stresses in coffee crops. The scope of the research focused on the use of the Symptom dataset, which consists of images of healthy and diseased leaves, to

EfficientNet-B0	GFLOPs	Parameters (M)
Quaternion-valued	2.1021	1.7
Real-valued	1.4553	5.9

Table 5. Comparison of complexity based on GFLOPs (10^9 FLOPs).

evaluate the effectiveness of these networks in the automated identification of agricultural pathologies.

Final results indicated that the quaternion-valued architectures consistently outperformed their real-valued counterparts across all evaluated metrics. To ensure a fair comparison, both models were trained from scratch, without relying on transfer learning. The quaternion-valued RGB model achieved the highest accuracy of 0.9560, compared to 0.9487 for the real-valued RGB model. Furthermore, the quaternion version proved to be highly parameter-efficient, utilizing only 1.72 million parameters (a 3.4-fold reduction) compared to the 5.86 million parameters of the real-valued network. However, as a consequence of the algebraic expansion of hypercomplex arithmetic, it presents higher computational complexity, requiring 2.1021 GFLOPs per input versus 1.4553 GFLOPs for the real-valued model.

As for practical deployment, the proposed Q-EfficientNet framework is highly suitable for real-time inference on resource-constrained edge devices, such as smartphone applications, autonomous tractors, or unmanned aerial vehicles, enabling direct field operations without continuous internet connectivity. To transition this model into an operational tool, three primary adaptations should be conducted: post-training quantization to optimize weight bit-width, integration of adaptive normalization layers to handle local exposure variations, and an edge-to-cloud continual learning pipeline to address real-world data drift.

For future work, we intend to explore the conversion of other architectures, such as ResNet50 and GoogLeNet, to the hypercomplex domain. Furthermore, investigating optimization methods to reduce the computational complexity (FLOPs) in quaternion-valued models emerges as a promising path to advance the study of hypercomplex-valued neural networks.

7. Acknowledgments

This research was conducted with the support of Fundação Amazônica de Amparo à Pesquisa e Desenvolvimento Tecnológico Desembargador Paulo dos Anjos Feitoza (FPFtech) and Foxconn Brasil.

References

Cariow, A. and Cariowa, G. (2025). Reduced-complexity algorithms for tessarine neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, 36(4):7751–7756.

Esgario, J. G., Krohling, R. A., and Ventura, J. A. (2020). Deep learning for classification and severity estimation of coffee leaf biotic stress. *Computers and Electronics in Agriculture*, 169:105162.

Grassucci, E., Zhang, A., and Comminiello, D. (2024). Phnns: Lightweight neural networks via parameterized hypercomplex convolutions. *IEEE Transactions on Neural Networks and Learning Systems*, 35(6):8293–8305.

Gupta, S. and Tan, M. (2019). Efficientnet-edgetpu: Creating accelerator-optimized neural networks with automl. *Google AI Blog*, 2(1).

Hu, J., Shen, L., and Sun, G. (2018). Squeeze-and-excitation networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7132–7141.

Manso, G. L., Knidel, H., Krohling, R. A., and Ventura, J. A. (2019). A smartphone application to detection and classification of coffee leaf miner and coffee leaf rust. *arXiv*.

Neto, G. V. and Valle, M. E. (2025a). V-efficientnets: Vector-valued efficiently scaled convolutional neural network models. In *2025 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8.

Neto, G. V. and Valle, M. E. (2025b). V-efficientnets: Vector-valued efficiently scaled convolutional neural network models. In *2025 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE.

Niemczynowicz, A. and Kycia, R. A. (2026). Fully tensorial approach to hypercomplex-valued neural networks. *Information Sciences*. In Press, Journal Pre-proof, Article 123796.

Parcollet, T., Morchid, M., and Linarès, G. (2020). A survey of quaternion neural networks. *Artificial Intelligence Review*, 53(4):2957–2982.

Parcollet, T., Ravanelli, M., Morchid, M., Linarès, G., Trabelsi, C., De Mori, R., and Bengio, Y. (2019). Quaternion convolutional neural networks for end-to-end automatic speech recognition. In *ICASSP*.

Rodrigues, C. and Valle, M. (2025). A study on convolutional vector-valued neural networks for color image classification. In *Anais do XXII Encontro Nacional de Inteligência Artificial e Computacional*, pages 1419–1430, Porto Alegre, RS, Brasil. SBC.

Sangwine, S. J. (1996). Quaternion neural networks for signal processing. *IEEE Workshop on Neural Networks for Signal Processing*.

Tan, M. and Le, Q. (2019). Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning*, pages 6105–6114. PMLR.

Tan, M. and Le, Q. (2021). Efficientnetv2: Smaller models and faster training. In *International conference on machine learning*, pages 10096–10106. PMLR.

Vieira, G. and Valle, M. E. (2022). A general framework for hypercomplex-valued extreme learning machines. *Journal of Computational Mathematics and Data Science*, 3:100032.

Yang, Y. et al. (2019). Quaternion convolutional neural networks for color image classification. *IEEE Access*.