A Quaternion-Valued Variational Autoencoder

Eleonora Grassucci, Danilo Comminiello, Aurelio Uncini

Dept. Information Engineering, Electronics and Telecommunications (DIET), Sapienza University of Rome, Italy

IEEE ICASSP 2021 | Toronto, Canada

6-11 June 2021

Sapienza
Università di Roma

Background

State-of-the-art generative models are able to reach impressive results at the cost of millions of parameters which require huge computational resources.

Real-valued networks process image channels as independent elements, not considering intra-channels relations and correlation.

Motivations: making deep models more accessible

Quaternion neural networks (QNNs) [1, 2, 3] allow to reduce the number of parameters by sharing quaternion-weight components through multiple quaternion-input parts.

QNNs process image channels as a single entity and grasp internal latent information, preserving intra-channels relations, thanks to the Hamilton product.

[1] P. Arena, L. Fortuna, L. Occhipinti, and M. G. Xibilia, "Neural networks for quaternion-valued function approximation," in *IEEE Int. Symp. on Circuits and Syst. (ISCAS)*, (London, UK), pp. 307–310, May 1994

[2] C. Gaudet and A. Maida, "Deep quaternion networks," in *IEEE Int. Joint Conf. on Neural Netw. (IJCNN)*, July 2018

[3] T. Parcollet, M. Morchid, and G. Linarès, "A survey of quaternion neural networks," *Artif. Intell. Rev.*, Aug. 2019

Our lead character

The core of the quaternion-valued domain $\mathbb{H}$ is the quaternion number:

$$q = q_a + q_b\hat{\imath} + q_c\hat{\jmath} + q_d\hat{\kappa} = q_a + \overline{q}. \tag{1}$$

The imaginary units comply with the property:

$$\hat{\imath}^2 = \hat{\jmath}^2 = \hat{\kappa}^2 = -1 \tag{2}$$

Quaternions are not commutative under the operation of vector multiplication:

$$\hat{\imath}\hat{\jmath} = -\hat{\jmath}\hat{\imath} \quad , \quad \hat{\jmath}\hat{\kappa} = -\hat{\kappa}\hat{\jmath} \quad , \quad \hat{\kappa}\hat{\imath} = -\hat{\imath}\hat{\kappa}. \tag{3}$$

Hamilton product for quaternion convolutions

In deep quaternion neural networks, the quaternion convolution is performed through the Hamilton product:

$$\begin{aligned}\mathbf{W} \otimes \mathbf{q} = &\left(\mathbf{W}_a\mathbf{q}_a - \mathbf{W}_b\mathbf{q}_b - \mathbf{W}_c\mathbf{q}_c - \mathbf{W}_d\mathbf{q}_d\right) \\ &+ \left(\mathbf{W}_a\mathbf{q}_b + \mathbf{W}_b\mathbf{q}_a + \mathbf{W}_c\mathbf{q}_d - \mathbf{W}_d\mathbf{q}_c\right)\hat{\imath} \\ &+ \left(\mathbf{W}_a\mathbf{q}_c - \mathbf{W}_b\mathbf{q}_d + \mathbf{W}_c\mathbf{q}_a + \mathbf{W}_d\mathbf{q}_b\right)\hat{\jmath} \\ &+ \left(\mathbf{W}_a\mathbf{q}_d + \mathbf{W}_b\mathbf{q}_c - \mathbf{W}_c\mathbf{q}_b + \mathbf{W}_d\mathbf{q}_a\right)\hat{\kappa}\end{aligned} \tag{4}$$

The quaternion convolution allows to capture internal latent relations within the features of a quaternion.

The forward phase for a generic quaternion fully connected layer can be defined as:

$$\mathbf{y} = \alpha\left(\mathbf{W} \otimes \mathbf{x} + \mathbf{b}\right) \tag{5}$$

where $\mathbf{y}$ is the output of the layer, $\mathbf{b}$ is the quaternion-valued bias offset and α is any quaternion split activation function:

$$\alpha\left(q\right) = f\left(q_a\right) + f\left(q_b\right) + f\left(q_c\right) + f\left(q_d\right). \tag{6}$$

Deep QCNN may also involve other operations in the quaternion domain, like pooling and batch normalization [4].

[4] R. Vecchi, S. Scardapane, D. Comminiello, and A. Uncini, "Compressing deep-quaternion neural networks with targeted regularisation," *CAAI Trans. Intell. Technol.*, vol. 5, pp. 172–176, Sept. 2020

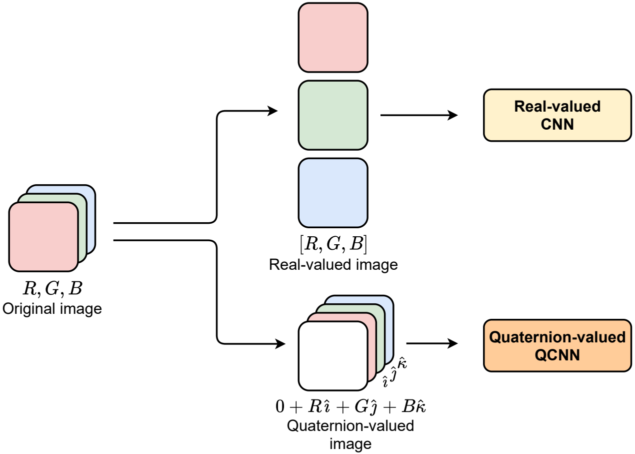

Figure 1: Image processing with real-valued CNN (top) and quaternion-valued QCNN (bottom).

Augmented second-order statistics

Second-order statistics in the quaternion domain involve the definition of the augmented covariance matrix $\tilde{\mathbf{C}}_{\mathbf{qq}}$. Further details can be found in the paper.

$$\tilde{\mathbf{C}}_{\mathbf{qq}} = \mathrm{E}\left\{\tilde{\mathbf{q}}\tilde{\mathbf{q}}^{\mathsf{H}}\right\} = \begin{bmatrix} \mathbf{C}_{\mathbf{qq}} & \mathbf{C}_{\mathbf{qq}^{\hat{\imath}}} & \mathbf{C}_{\mathbf{qq}^{\hat{\jmath}}} & \mathbf{C}_{\mathbf{qq}^{\hat{\kappa}}} \\ \mathbf{C}^{\mathsf{H}}_{\mathbf{qq}^{\hat{\imath}}} & \mathbf{C}_{\mathbf{q}^{\hat{\imath}}\mathbf{q}^{\hat{\imath}}} & \mathbf{C}_{\mathbf{q}^{\hat{\imath}}\mathbf{q}^{\hat{\jmath}}} & \mathbf{C}_{\mathbf{q}^{\hat{\imath}}\mathbf{q}^{\hat{\kappa}}} \\ \mathbf{C}^{\mathsf{H}}_{\mathbf{qq}^{\hat{\jmath}}} & \mathbf{C}_{\mathbf{q}^{\hat{\jmath}}\mathbf{q}^{\hat{\imath}}} & \mathbf{C}_{\mathbf{q}^{\hat{\jmath}}\mathbf{q}^{\hat{\jmath}}} & \mathbf{C}_{\mathbf{q}^{\hat{\jmath}}\mathbf{q}^{\hat{\kappa}}} \\ \mathbf{C}^{\mathsf{H}}_{\mathbf{qq}^{\hat{\kappa}}} & \mathbf{C}_{\mathbf{q}^{\hat{\kappa}}\mathbf{q}^{\hat{\imath}}} & \mathbf{C}_{\mathbf{q}^{\hat{\kappa}}\mathbf{q}^{\hat{\jmath}}} & \mathbf{C}_{\mathbf{q}^{\hat{\kappa}}\mathbf{q}^{\hat{\kappa}}} \end{bmatrix} \tag{7}$$

Second-order statistics and $\mathbb{Q}$-Properness

For $\mathbb{Q}$-proper distributions, $\tilde{\mathbf{C}}_{\mathbf{qq}}$ is a diagonal matrix:

$$\tilde{\mathbf{C}}_{\mathbf{qq}} = \mathrm{E}\left\{\tilde{\mathbf{q}}\tilde{\mathbf{q}}^{\mathsf{H}}\right\} = \begin{bmatrix} \mathbf{C}_{\mathbf{qq}} & 0 & 0 & 0 \\ 0 & \mathbf{C}_{\mathbf{q}^{\hat{\imath}}\mathbf{q}^{\hat{\imath}}} & 0 & 0 \\ 0 & 0 & \mathbf{C}_{\mathbf{q}^{\hat{\jmath}}\mathbf{q}^{\hat{\jmath}}} & 0 \\ 0 & 0 & 0 & \mathbf{C}_{\mathbf{q}^{\hat{\kappa}}\mathbf{q}^{\hat{\kappa}}} \end{bmatrix} \tag{8}$$

The diagonal contains the covariance matrices of the quaternion input and its involutions.

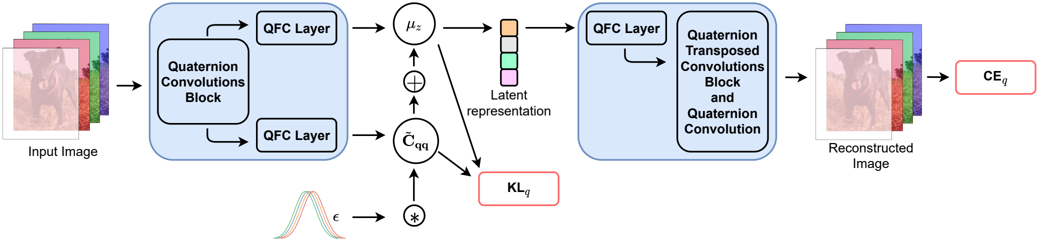

Figure 2: Quaternion-valued variational autoencoder architecture (QVAE). The quaternion encoder learns the quaternion mean μ_q and the augmented covariance matrix $\tilde{\mathbf{C}}_{\mathbf{qq}}$ to build the latent representation. The quaternion decoder reconstruct the 4-channel image. The PyTorch implementation of the QVAE is available online at https://github.com/eleGAN23/QVAE.

1 Experimental Results

Reconstruction task

Figure 3: Original test set and reconstructed samples sets from plain VAE and proposed QVAE.

Generation task

Figure 4: Generated fake image samples from the plain VAE and the proposed QVAE.

Objective metrics results

Table 1: Averaged results from objective metrics on reconstruction (SSIM, MSE) and generation (FID) tasks.

	SSIM↑	MSE↓	FID↓	# parameters↓
VAE	0.8492	0.0047	195.7	3,762,539
QVAE	**0.8941**	**0.0031**	**175.7**	**1,404,996**

Conclusions

Moving neural networks from the real domain to the quaternion one allows the network to process image channels as a whole element, capturing internal latent relations.

Quaternion layers reduces the number of parameters and memory consumption.

The plain QVAE shows promising results, generating better images with less than a half the number of parameters with respect to the real-valued counterpart.

Future works

- Expand QVAE for $\mathbb{Q}$-improper distributions.
- Test more complex variational autoencoder in the quaternion domain.
- Extend QVAE for other kind of signals such as audio.

References

[1] P. Arena, L. Fortuna, L. Occhipinti, and M. G. Xibilia, "Neural networks for quaternion-valued function approximation," in *IEEE Int. Symp. on Circuits and Syst. (ISCAS)*, (London, UK), pp. 307–310, May 1994.

[2] C. Gaudet and A. Maida, "Deep quaternion networks," in *IEEE Int. Joint Conf. on Neural Netw. (IJCNN)*, July 2018.

[3] T. Parcollet, M. Morchid, and G. Linarès, "A survey of quaternion neural networks," *Artif. Intell. Rev.*, Aug. 2019.

[4] R. Vecchi, S. Scardapane, D. Comminiello, and A. Uncini, "Compressing deep-quaternion neural networks with targeted regularisation," *CAAI Trans. Intell. Technol.*, vol. 5, pp. 172–176, Sept. 2020.

[5] X. Hou, L. Shen, K. Sun, and G. Qiu, "Deep feature consistent variational autoencoder," in *Applications of Computer Vision (WACV), 2017 IEEE Winter Conference on*, pp. 1133–1141, IEEE, 2017.

[6] A. Vahdat and J. Kautz, "NVAE: A deep hierarchical variational autoencoder," in *Neural Information Processing Systems (NeurIPS)*, 2020.

[7] D. Comminiello, M. Lella, S. Scardapane, and A. Uncini, "Quaternion convolutional neural networks for detection and localization of 3D sound events," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, (Brighton, UK), pp. 8533–8537, May 2019.

[8] E. Grassucci, D. Comminiello, and A. Uncini, "Quaternion-valued variational autoencoder," in *arXiv preprint arXiv:2010.11647v1*, 2020.

THANK YOU FOR YOUR ATTENTION

QUESTIONS?

Eleonora Grassucci

eleonora.grassucci@uniroma1.it