About the equivalence between complex-valued and real-valued fully connected neural networks - application to PolInSAR images

- J. A. Barrachina
- C. Ren
- C. Morisseau
- G. Vieillard
- J-P. Ovarlez

SONDRA

CentraleSupélec

ONERA

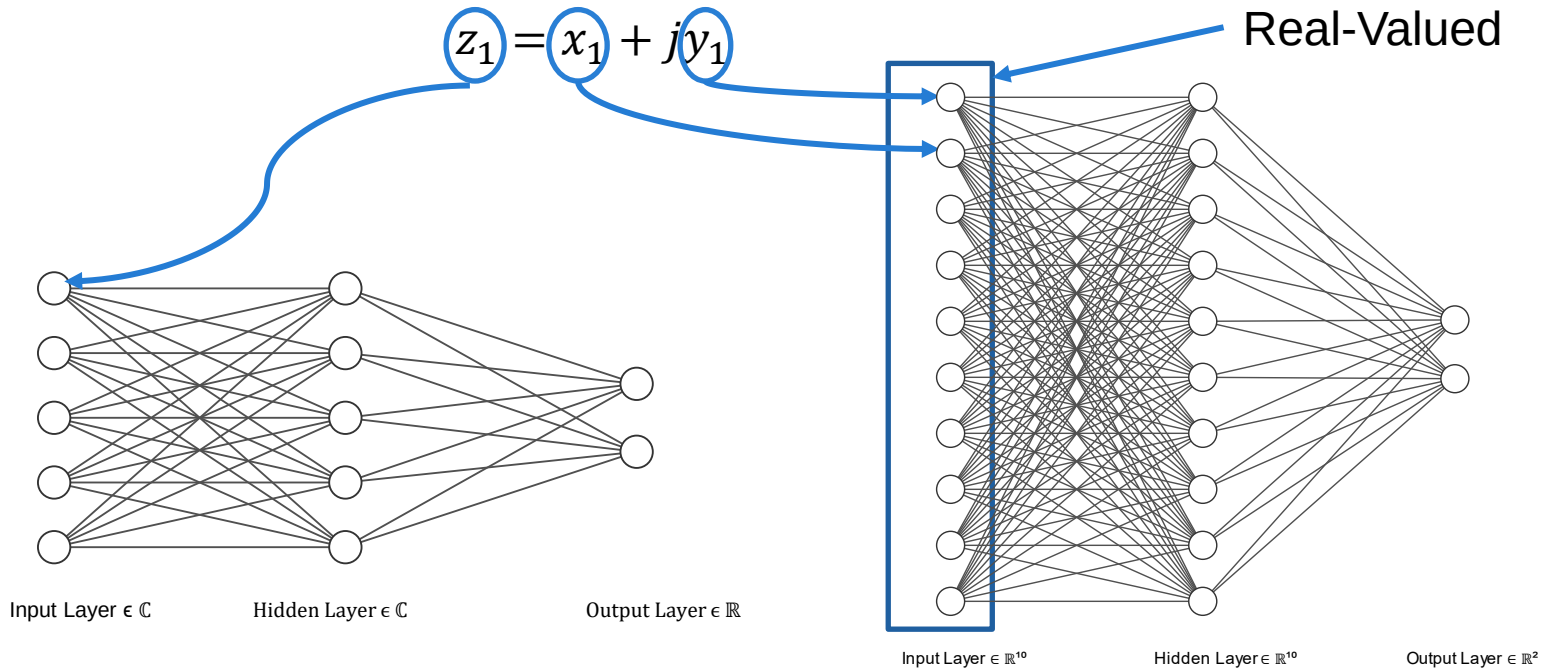
THE FRENCH AEROSPACE LAB

DGA

Liberté • Égalité • Fraternité
RÉPUBLIQUE FRANÇAISE

MINISTÈRE DES ARMÉES

Complex-Valued Neural Networks



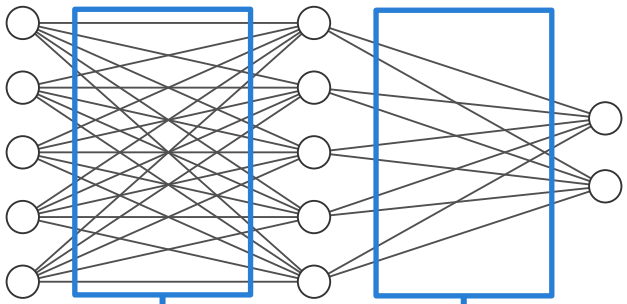
- Complex Input
- Complex Weights
- Complex activation functions $\begin{cases} f(z) = g(x) + jh(y) \text{ (Type A)} \\ f(z) = g(|z|)e^{jarg(z)} \text{ (Type B)} \end{cases}$

[1] Kuroe et al. *"On Activation Functions for Complex-Valued Neural Networks"* 2003

Complex-Valued NN: Real-equivalent model

What is a real-equivalent network? [1]

Complex Network

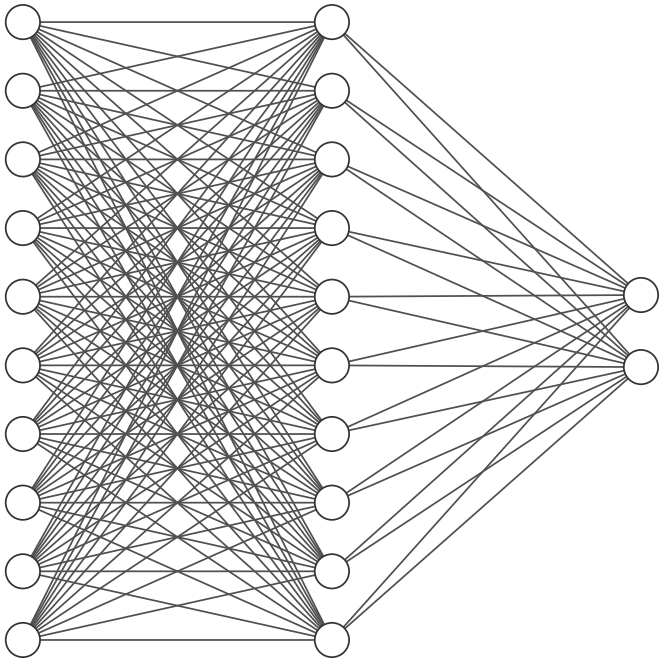


Input Layer $\in \mathbb{C}$ Hidden Layer $\in \mathbb{C}$ Output Layer $\in \mathbb{R}$

$$\begin{cases} 5 * 2 = 10_{\mathbb{C}} = 20_{\mathbb{R}} \\ 10 * 2 = 20_{\mathbb{R}} \end{cases}$$

$$\begin{cases} 5 * 5 = 25_{\mathbb{C}} = 50_{\mathbb{R}} \\ 10 * 10 = 100_{\mathbb{R}} \end{cases}$$

Real Network

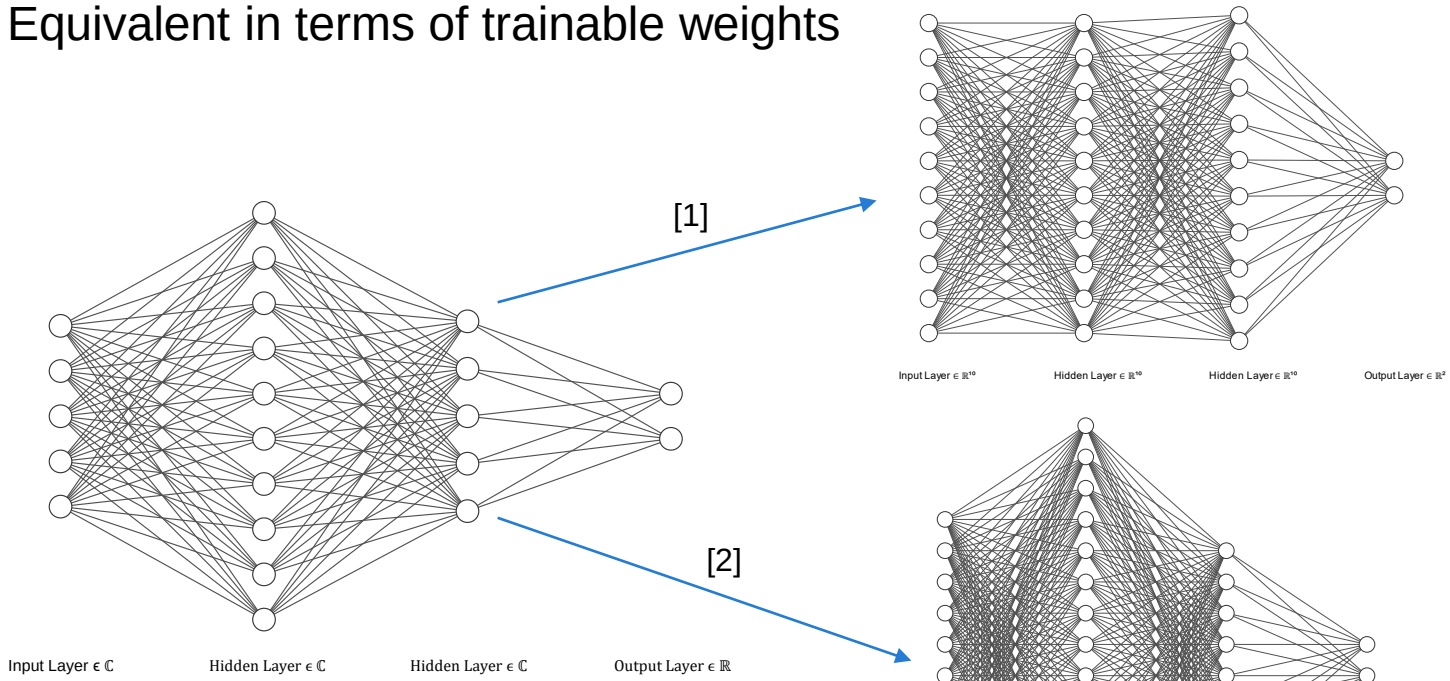


Input Layer $\in \mathbb{R}^{10}$ Hidden Layer $\in \mathbb{R}^{10}$ Output Layer $\in \mathbb{R}^{2}$

[1] Mönning et al. *"Evaluation of Complex-Valued Neural Networks on Real-Valued Classification Tasks"* 2018.

Complex-Valued NN: Real-equivalent model

Equivalent in terms of trainable weights



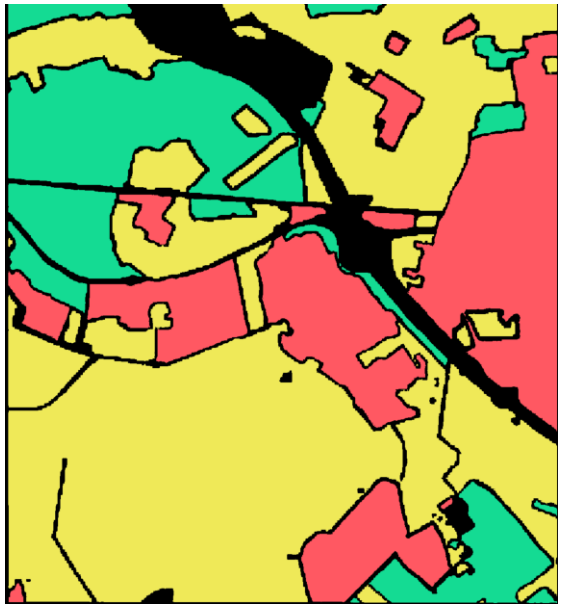
$$N_i^{\mathbb{R}} = r\, N_i^{\mathbb{C}}, \qquad 1 \leq r < 2$$

Solution to: $tp = rN_0N_1^{\mathbb{C}} + \sum_{i=1}^{K-1} r^2 N_i^{\mathbb{C}} N_{i+1}^{\mathbb{C}} + rN_K^{\mathbb{C}} N_L$

[1] Mönning et al. *"Evaluation of Complex-Valued Neural Networks on Real-Valued Classification Tasks"* 2018.

Oberpfaffenhofen PolInSAR

Oberpfaffenhofen PolInSAR database
European Space Agency (esa)



A Built-up Area B Wood Land C Open Area

Experimental Data

$$S^{(i)} = \left(S_{HH}^{(i)}, \sqrt{2}S_{HV}^{(i)}, S_{VV}^{(i)}\right)^T, i \in [1,2]$$

$$k^{(i)} = \frac{1}{\sqrt{2}}\left(S_{HH}^{(i)} + S_{VV}^{(i)}, S_{HH}^{(i)} - S_{VV}^{(i)}, 2S_{HV}^{(i)}\right)^T$$

Interferometric Polarimetric Coherency matrix

$$T = \frac{1}{n}\sum_{j}^{n} k_j k_j^H, \qquad k = [k^{(1)T}, k^{(2)T}]$$

- $\mathbb{C}^{6\times 6}$ Hermitian
- Real-valued diagonal
- Total 21 values

Circularity: Model Architecture

Complex-Valued Multi-Layer Perceptron

	CVNN	RVNN
Output Size	3	3
Input Size	21	42
First Hidden Layer	100	*
Second Hidden Layer Size	50	*
Dropout	50%	50%
Activation Function	ReLU Type A [1]	ReLU
Output Activation	Softmax over absolute value	Softmax

* Equivalent RVNN: 2 Hidden Layer

Model:

- Loss: Categorical cross-entropy
- Weight initialization: Glorot uniform
- SGD (Stochastic Gradient Descent)
  - Learning rate 0.1
  - Wirtinger Derivative

Dataset:

- 8% train set
- 2% validation set
- 90% test set

Simulation:

- 100 trials each model
- 300 epochs
- Batch size 100

[1] Kuroe et al. *"On Activation Functions for Complex-Valued Neural Networks"* 2003

Oberpfaffenhofen PolInSAR: Results

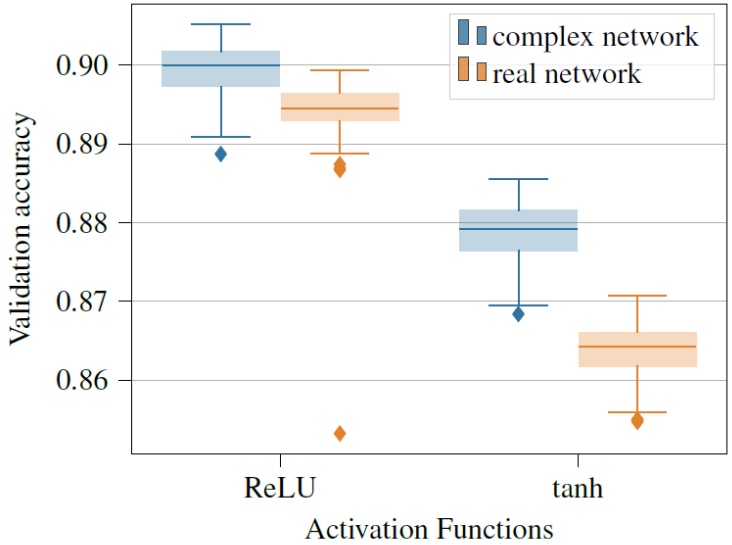
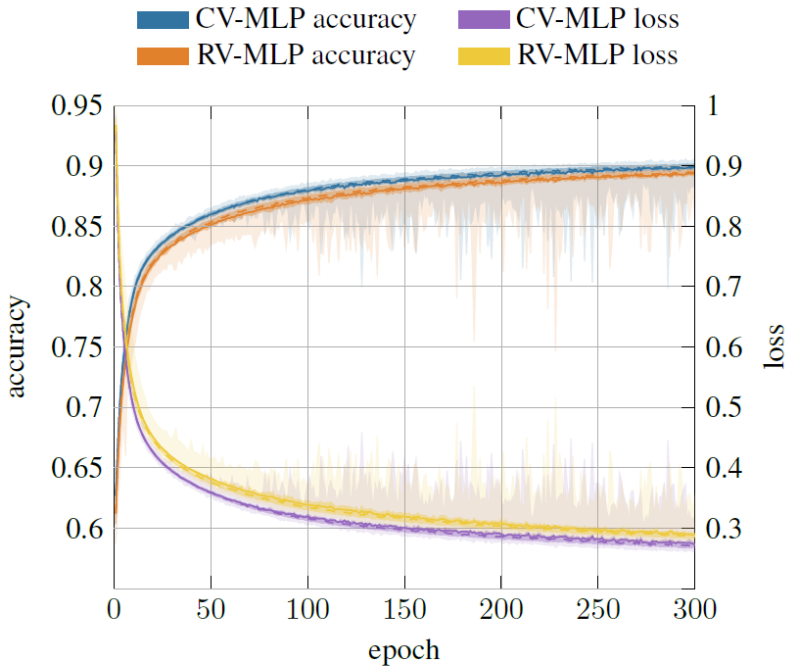


Fig. 2. Box Plot Comparison between activation functions



ig. 3. ReLU validation loss and accuracy per epoch, solid ne represents the mean.

Oberpfaffenhofen PolInSAR

ReLU	CV-MLP	RV-MLP	polar-RV-MLP
median	**90.00 ± 0.07**	89.45 ± 0.06	84.93 ± 0.04
mean	**89.92 ± 0.09**	89.40 ± 0.13	84.93 ± 0.04
IQR	89.73 – 90.17	89.30 – 89.45	84.93 – 85.03
range	88.87 – 90.52	85.32 – 89.94	84.50 – 85.32

Table I: Validation Accuracy (%)

Cao et al. “Pixel-Wise PolSAR Image Classification via a Novel Complex-Valued Deep Fully “ 2019

90.02±0.56

89.36±1.30

A Built-up Area

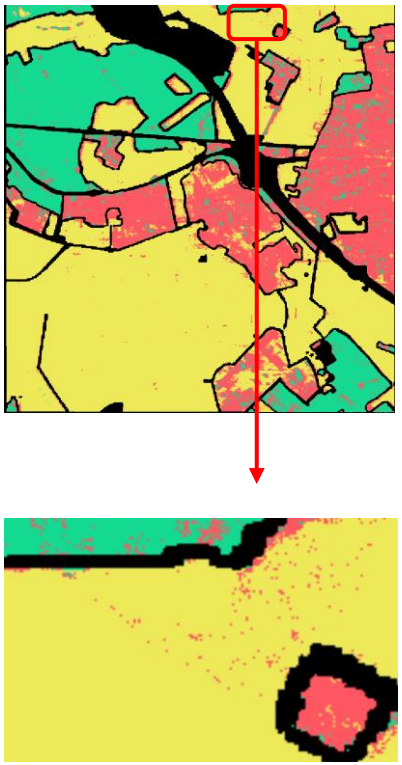
B Wood Land

C Open Area

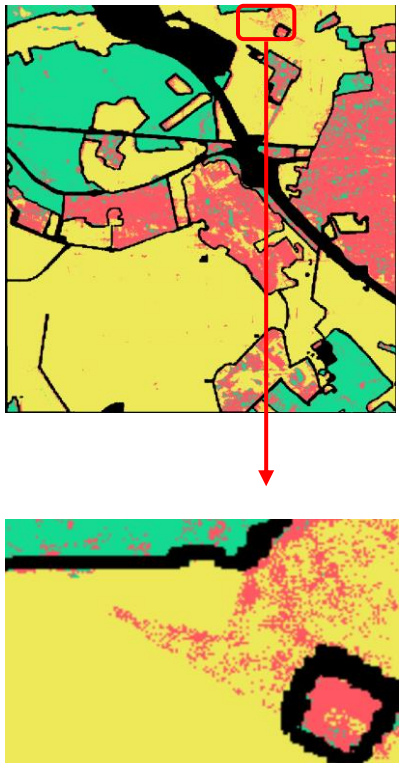
	CV-MLP			RV-MLP		
	A	**B**	**C**	**A**	**B**	**C**
A	**81.25**	8.62	10.13	**81.01**	9.19	9.80
B	6.86	**92.83**	0.30	7.33	**92.39**	0.28
C	4.11	0.21	**95.67**	4.97	0.24	**94.79**

Table 2. Confusion matrix of mean accuracy (%)

CVNN



RVNN



Ground Truth

Thank you!