MRC UNIVERSITY OF ALBERTA

3vG TERRA INSIGHTS

IGARSS Pasadena, California 2023

# Self-supervised Residual Distribution Learning for SAR Interferometric Phase Filtering and Coherence Estimation

*Xinyao Sun,* **Zheng Wang***, *Ryan Goldsbury, Irene Cheng*

*21 July, 2023*

# LOGISTICS

- Background - Phase Filtering and Coherence Estimation
- Methodology
  - Mixture Density Network – Distribution Learning
  - Random Masking – Self-supervised Learning
  - Monte Carlo Sampling – Coherence Estimation
- Experimental Results
  - Synthetic Data
  - Real-world Data
- Discussions and Conclusions

# Background - Coherence Estimation and Phase Filtering

**Sample coherence is used as the estimate of the theoretical coherence.**

- Complex coherence Δ of two zero-mean complex signals z1 and z2 for stationary processes:

$$\Delta = \frac{E(z_1 z_2^*)}{\sqrt{E\left(|z_1|^2\right)}\sqrt{E\left(|z_2|^2\right)}} \qquad (1)$$

- Coherence estimate with *L* samples:

$$\delta = \frac{\sum_{i=1}^{L} z_{1i} z_{2i}^*}{\sqrt{\sum_{i=1}^{L} |z_1|^2}\sqrt{\sum_{i=1}^{L} |z_2|^2}} \qquad (2)$$

- **Sample coherence** estimate from solely interferometric phase:

$$\gamma = \frac{1}{L}\left|\sum_{i=1}^{L} e^{j\cdot\phi_i}\right| \qquad (3)$$

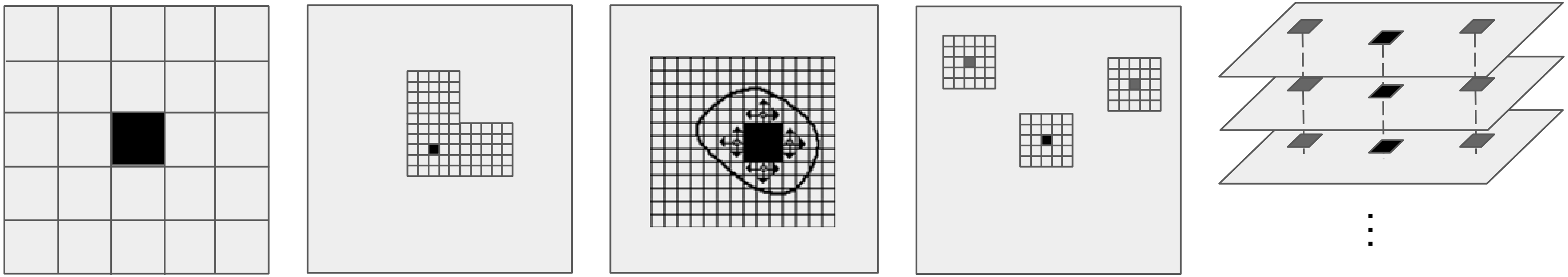
- **Filtered** interferometric **phase**:

$$\phi = \arctan\left(\frac{\sum_{i=1}^{L} \sin\phi_i}{\sum_{i=1}^{L} \cos\phi_i}\right) \qquad (4)$$

**Reference**: Touzi R, Lopes A, Bruniquel J, Vachon PW. Coherence estimation for SAR imagery. IEEE transactions on geoscience and remote sensing. 1999 Jan;37(1):135-49.

# Background - Coherence Estimation and Phase Filtering

**Various methods have been proposed to perform the (adaptive) selection of the L samples for coherence estimation and phase filtering.**

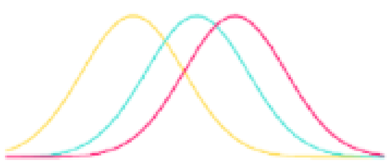


Boxcar window | Oriented windows | Region growing | Patchwise non-local | Pointwise non-local

**Reference**: Deledalle CA, Denis L, Tupin F, Reigber A, Jäger M. NL-SAR: A unified nonlocal framework for resolution-preserving (Pol)(In) SAR denoising. IEEE Transactions on Geoscience and Remote Sensing. 2014 Sep 19;53(4):2021-38.

# Methodology

**This study aims to address the issue from a different perspective**:



$N(\mu_{Real}, \sigma^2_{Real})$, $N(\mu_{Imag}, \sigma^2_{Imag})$

Monte Carlo Sampling

$$\phi = \arctan\left(\frac{\mu_{Imag}}{\mu_{Real}}\right)$$

$$\gamma = \frac{1}{L}\left|\sum_{i=1}^{L} e^{j\cdot\phi_i}\right|$$

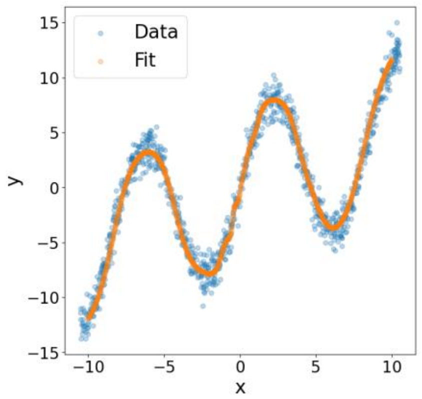
- **Mixture Density Network (MDN)** – Learn the residual distributions, i.e. $N(\mu_{Real}, \sigma^2_{Real})$, $N(\mu_{Real}, \sigma^2_{Imag})$ of phase noise from a large receptive field (e.g. 128x128).

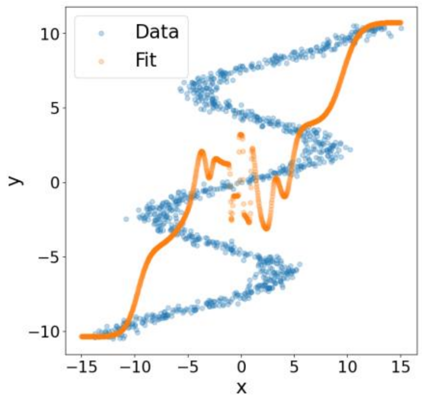
- **Monte Carlo Sampling** – Generate L (e.g. 200) samples from the learned distributions.

- **Random Masking** – Self-supervised learning.

# Methodology – Mixture Density Network (MDN)

- Curve fitting problem




One/many-to-one fitting

One-to-many fitting

$y_{true}(x)=7\sin(0.75x)+0.5x+\boldsymbol{\varepsilon}$

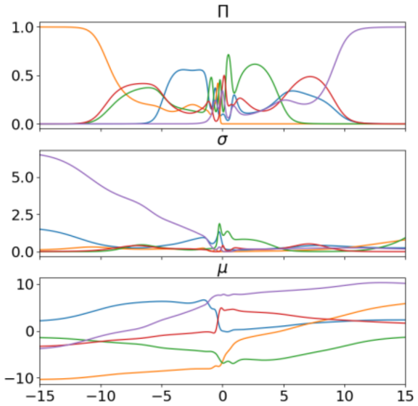
$x_{true}(y)=7\sin(0.75y)+0.5y+\boldsymbol{\varepsilon}$

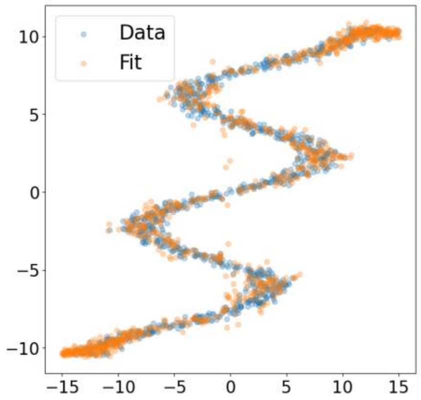
**Reference**: Bishop, C.M., 1994. Mixture density networks.

- MDN predicts a class of probability distributions called Mixture of Gaussians or Gaussian Mixture Models, where the output value is modelled as a weighted sum of multiple ($K$) Gaussians:

$$P(y|x) = \sum_{k=1}^{K} \Pi_k(x) G(y,\, \mu_k(x),\, \sigma_k(x)),$$

where $\Pi_k$ is the weight, $G$ a Gaussian function at given $\mu_k$ mean, $\sigma_k$ standard deviation.



# Methodology – Learning Strategies

**Depending on the targets/labels, there are different learning strategies for image denoising.**

- Noise2Truth Learning (Supervised)

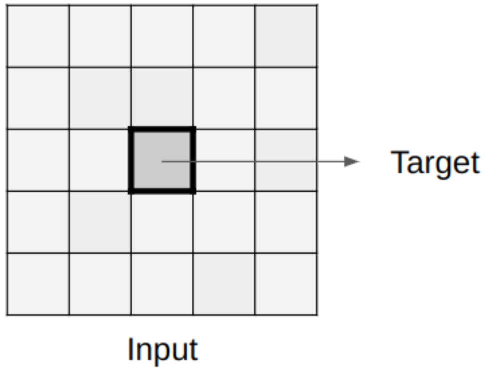
$$\underset{\theta}{\operatorname{argmin}} \sum_i L\left(f_\theta(\hat{x}_i),\, y_i\right)$$

- Noise2Noise Learning

$$\underset{\theta}{\operatorname{argmin}} \sum_i L\left(f_\theta(\hat{x}_i),\, \hat{y}_i\right)$$

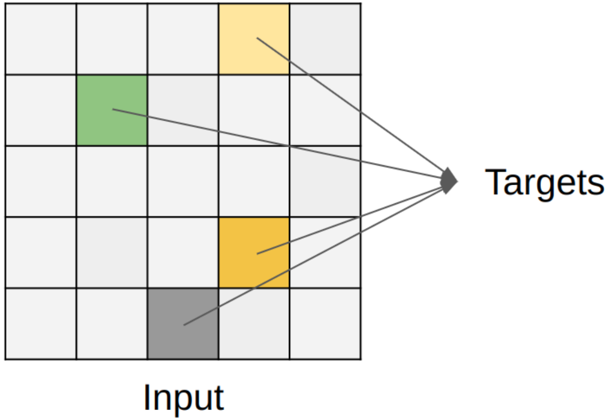
- Noise2void Learning (Self-supervised)

$$\underset{\theta}{\operatorname{argmin}} \sum_i L\left(f_\theta(\hat{x}_i),\, \hat{y}_i\right)$$




**References**: Lehtinen, J., Munkberg, J., Hasselgren, J., Laine, S., Karras, T., Aittala, M. and Aila, T., 2018. **Noise2Noise**: Learning image restoration without clean data. arXiv preprint arXiv:1803.04189.
Krull, A., Buchholz, T.O. and Jug, F., 2019. **Noise2void**-learning denoising from single noisy images. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 2129-2137).

# Methodology – Random Masking




**Random masking is similar to denoising/masked autoencoder.**

- Definition of variables

Observation: $\alpha$,
Additive noise: $\boldsymbol{\varepsilon}$,
Residual/noise estimate: $\omega$, which follows $N(\mu_\omega, \sigma^2_\omega)$.

- **Training**

Assuming 20% pixels are randomly masked.

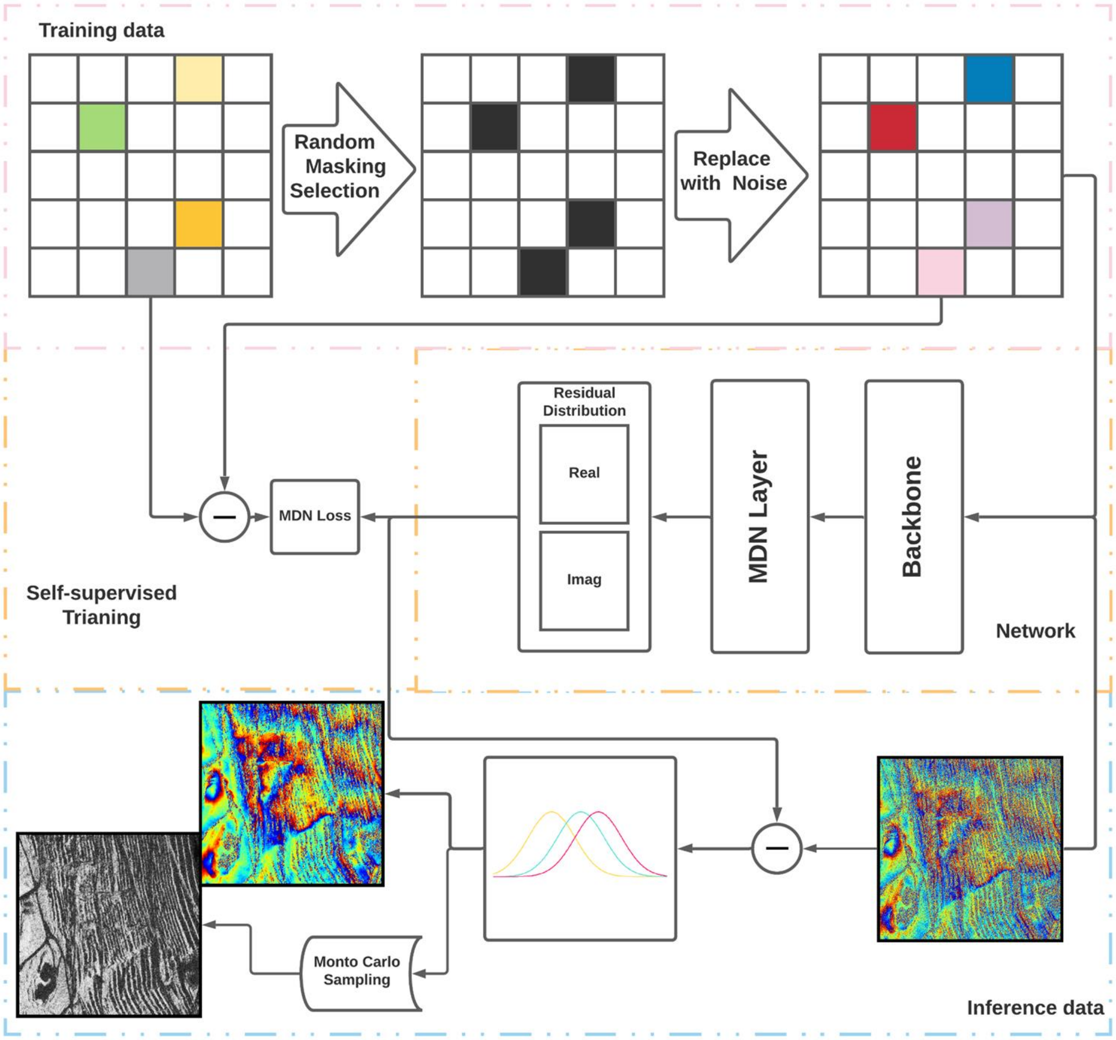
Input: 80% $\alpha$ + 20% $\boldsymbol{\varepsilon}$.
Label: 20% $\alpha$ + 80% zeros.
Out: 100% $\omega$.
Loss: -logProb($\boldsymbol{\varepsilon}$ - $\mu_\omega$, $\sigma_\omega$, $\alpha$) with 20% masked pixels.

- **Inference**

No random masking or additive noise.

Clean interferogram
= input - output
= 100% $\alpha$ - 100% $\omega$

**Reference**: He K, Chen X, Xie S, Li Y, Dollár P, Girshick R. Masked autoencoders are scalable vision learners. InProceedings of the IEEE/CVF conference on computer vision and pattern recognition 2022 (pp. 16000-16009).


Training data
Random Masking Selection
Replace with Noise
Residual Distribution
Real
Imag
MDN Layer
Backbone
MDN Loss
Self-supervised Trianing
Network
Monto Carlo Sampling
Inference data

# Experimental Results – Synthetic Data

**Additive noise**: σ=0.5 radian

**Network**: UNet

**Training**:

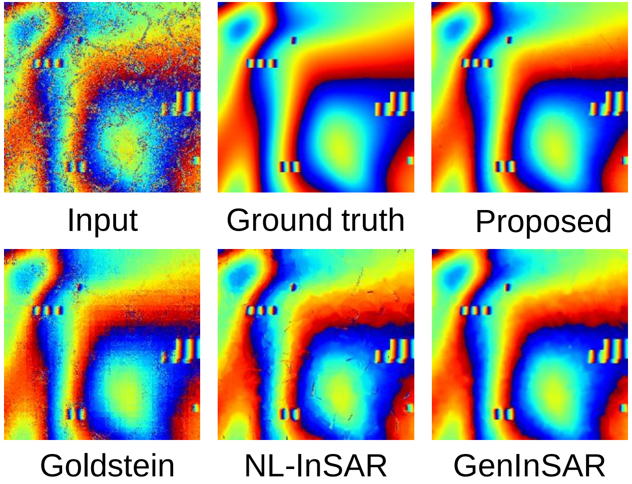
- 240 interferograms (size: 1000*1000)
- 40.6k data samples (patch size: 128*128)
- 30 training epochs (batch size: 16).
- AdamW optimization (lr=1e-3, betas=(0.9, 0.999), eps=1e-8)

**Random Masking**:

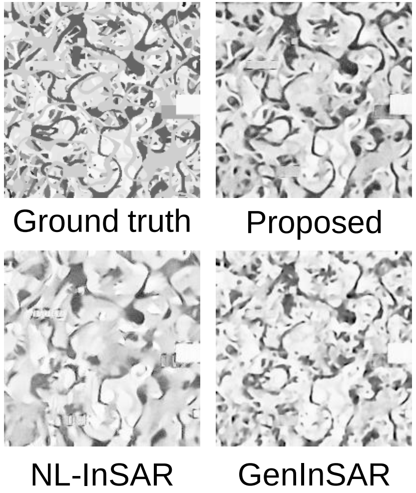
- Mask ratio: 20%-30%
- Replacement for the masked pixels: uniform random noise (-pi, pi]

**Testing**: 60 interferograms (1000*1000)

| Method | Phase RMSE (Radian) | Coherence RSME | Cosine Error |
|---|---|---|---|
| Goldstein | 1.26 | N/A | 0.048 |
| NL-InSAR | 0.85 | 0.159 | 0.014 |
| GenInSAR (Noise2void) | 0.687 | 0.138 | 0.005 |
| **Proposed** | **0.557** | **0.025** | **0.003** |



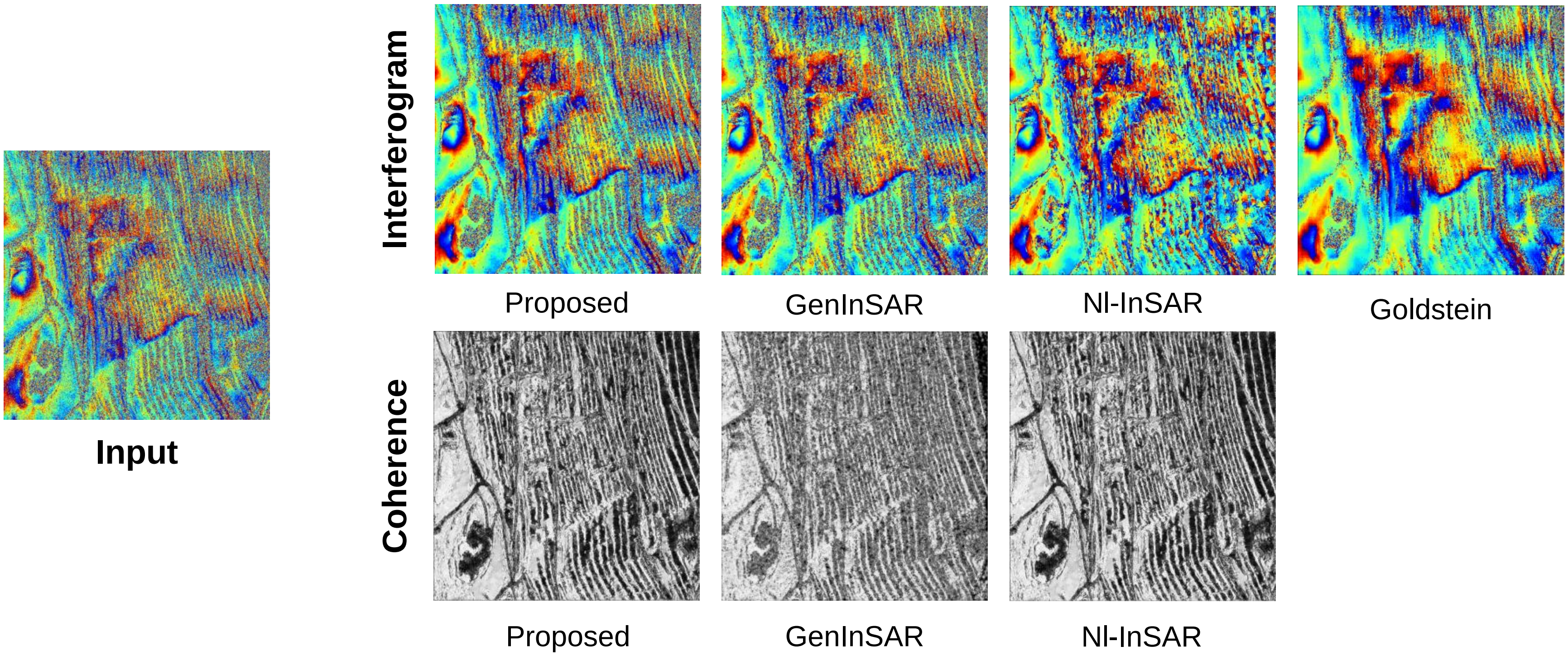

**Interferogram**




**Coherence**

# Experimental Results – Real-world Data

300 ifgs (size: 1000*1000) from multiple sensors; 50.7k data samples (patch size: 128*128)



# Discussions & Conclusions

- The proposed method is feasible and effective for InSAR phase filtering and coherence estimation.

- **Random masking** allows more pixels than the blind-spot scheme (used in GenInSAR/Noise2void) to contribute to the loss function and ensures the training is compatible with any modern fully convolutional image to image architecture.

- **Monte Carlo** sampling generates samples for coherence estimation, so that produces flawless [0, 1] coherence value outputs. On the other hand, GenInSAR used a handcrafted formula and the coherence values can be outside the range.

# THANK YOU

## Contact

*Zheng Wang, PHD*

*Scientific Developer II*

*Email: zwang@3vgeomatics.com*

*Website: https://3vgeomatics.com/*