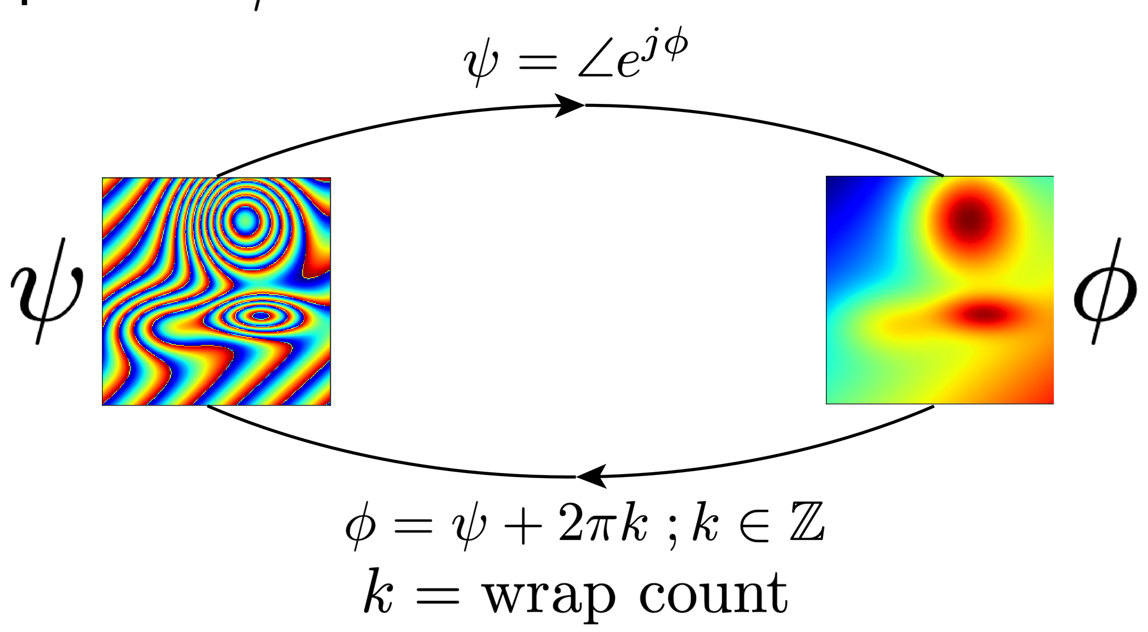
Paper ID : 5122
46th IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2021
Toronto, Ontario, Canada

A JOINT CONVOLUTIONAL AND SPATIAL QUAD-DIRECTIONAL LSTM NETWORK FOR PHASE UNWRAPPING

Malsha Perera*, Ashwin De Silva*
Department of Electronics and Telecommunication Engineering, University of Moratuwa, Sri Lanka

Introduction

Objective of phase unwrapping : Recover the true phase ϕ from wrapped phase ψ



Since $\psi = \angle e^{j\phi}$, the family denoted by $\phi + 2\pi\pi; n \in \mathbb{Z}$ will give rise to the same wrapped phase ψ

Phase unwrapping becomes challenging at the presence of **noise**, **phase discontinuities**, and **rapid variation of phase.**

Phase unwrapping (PU) problem is prevalent in applications such as quantita--tive susceptibility mapping in MRI, synthetic aperture radar interferometry, and fringe projection techniques.

Previous Work

1. Path Following Algorithms (QGPU, Branch-Cut Algorithms, etc.)
2. Minimum Norm Approaches
3. CNN based Approaches
   - Formulating PU as a Semantic Segmentation Task (PhaseNet 2.0)
     $$\phi^* = \psi + 2\pi f(\psi;\theta^*) \text{ where } \theta^* = \underset{\theta}{\operatorname{argmin}}\, \mathcal{L}(f(\psi;\theta), k)$$
   - Formulating PU as a Regression Task (Ryu et al.)
     $$\phi^* = g(\psi;\theta^*) \text{ where } \theta^* = \underset{\theta}{\operatorname{argmin}}\, \mathcal{L}(g(\psi;\theta), \phi)$$

Shortcomings : High Computational Cost (2), Low Noise Robustness (1), Difficulty of CNNs to model global spatial dependencies (3), High Data-intensiveness (3), Inappropriate Loss Functions (3)

Methodology

Regression formulation : $\phi^* = g(\psi;\theta^*)$ where $\theta^* = \underset{\theta}{\operatorname{argmin}}\, \mathcal{L}_c(g(\psi;\theta), \phi)$

here, $\mathcal{L}_c = \lambda_1 \mathcal{L}_{var} + \lambda_2 \mathcal{L}_{tv}$ (problem-specific composite loss function)

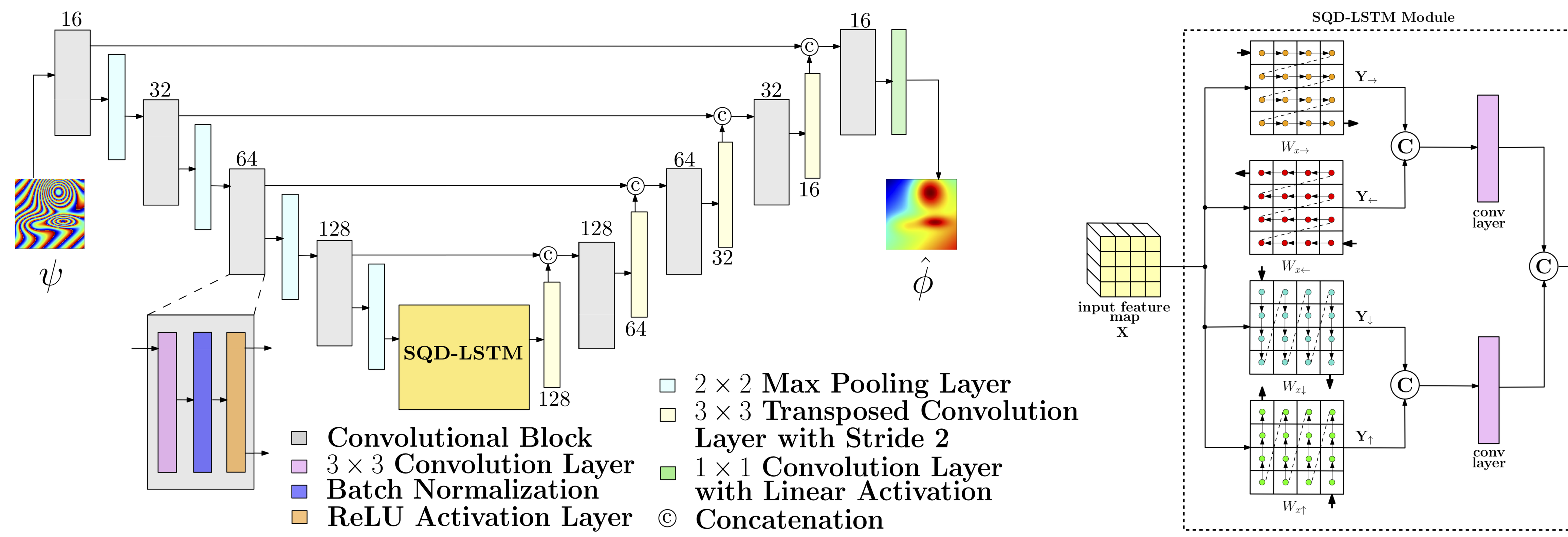
Variance of Error Loss $\rightarrow$ $\mathcal{L}_{var} = \mathbb{E}[(\hat{\phi} - \phi)^2] - (\mathbb{E}[(\hat{\phi} - \phi)])^2$

Total Variation of Error Loss $\rightarrow$ $\mathcal{L}_{tv} = \mathbb{E}[|\hat{\phi}_x - \phi_x| + |\hat{\phi}_y - \phi_y|]$

Here, ϕ is the true phase and $\hat{\phi} = g(\psi;\theta)$ is the predicted phase

The proposed $\mathcal{L}_c$ allows for **multiple solutions at convergence** while enforcing the **similarity between predicted phase and true phase.**

The function g is estimated by the following **joint convolutional and spatial quad-directional LSTM (SQD-LSTM) network.**



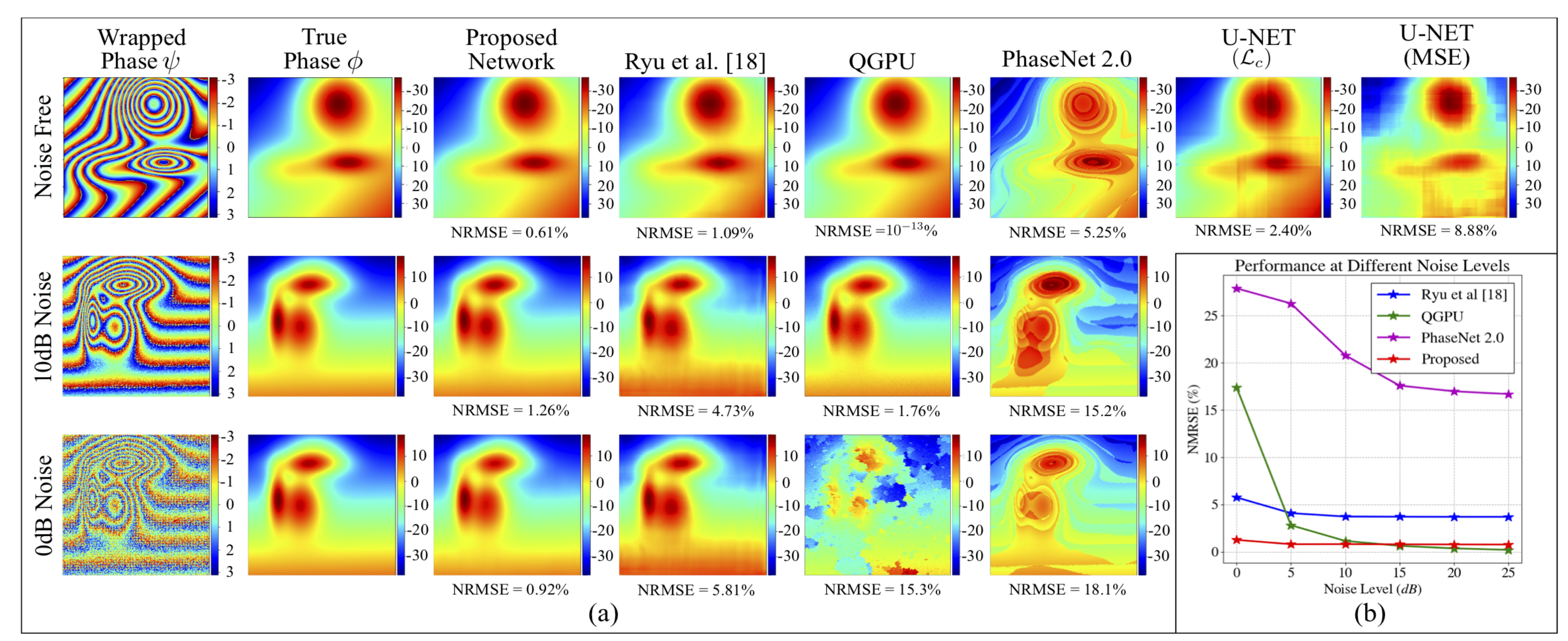
The SQD-LSTM block learns the **global spatial dependencies** within the phase images.

Experiments and Results

Phase images were simulated by adding gaussians of random shapes and positons, together with ramp phase.

Datasets : (1) Noise Free (2) Noisy (0, 5, 10, 20, and 60 dB noise levels)

Train-Test Split = 5000 : 1000



PU Performance Metric ⟶

$$\text{NRMSE} = \frac{\sqrt{\mathbb{E}[(\phi^* - \phi)^2]}}{\phi_{max} - \phi_{min}}$$

Table 1. Results

Method	Noise Free NRMSE	Noisy NRMSE	Computational Time (s)
UNET (MSE)	14.24%	-	0.234
UNET ($\mathcal{L}_c$)	2.75%	-	0.262
Ryu et al.[18]	2.23%	3.84%	0.687
PhaseNet 2.0 [8]	9.41%	17.53%	0.234
QGPU [5]	**10^{-13}%**	5.04%	35.42
Proposed Method	0.84%	**0.90%**	**0.054**

Conclusions

- Achieved state-of-the-art PU performance under severe noise levels
- Expends a signficantly less computational time during inference