Supplementary Material – Speckle Tracking in X-ray Phase-Contrast Imaging with Attention-based Neural Networks

Simón González[*], Carlos Ávila , Manuel F. Sánchez

Physics Department, Universidad de los Andes, Bogotá, Colombia

[*]s.gonzalez38@uniandes.edu.co

ABSTRACT

Supplementary material to the article Speckle Tracking in X-ray Phase-Contrast Imaging with Attention-based Neural Networks is presented here. A detailed schematic of the CNN used is shown, as well as details of the model training. In addition, the method used to integrate the differential phase signal and comments on the geometry of the hyperboloid used as an evaluation sample are explained.

I. DETAILED MODEL ARCHITECTURE

The structure of the network is presented in Fig. 1. Each convolutional block within the model is labeled with the shape of the output provided by the block. The gray dotted lines represent the skip connection between the encoder and the decoder. This operation consists of a concatenation of outputs, so the number of input channels to the decoder is the sum of those of the previous output and that of the encoder. On the other hand, the blue dotted lines represent residual connections inside each residual block with CBAM. These represent the addition operation of the block input with the nonlinear output of the block, so the number of channels is not altered in this process.

II. TRAINING DETAILS

As mentioned in the main manuscript, the data were normalized within the training loop to handle data with a mean of 0 and a standard deviation of 1, to avoid working with minimal gradients. For this, each data f belonging to a group F, either inputs, d_x or d_y, is normalized as follows:

$$\hat{f} = \frac{f - mean(F)}{std(F)} \tag{1}$$

On the other hand, the network weights were initialized with the Kaiming [1] method. This initialization helps to avoid both gradient vanishing and explosive gradients. Since our activations are ReLU units, the weights associated with each neuron have zero mean and standard deviation $sqrt2/n$ with n the number of neurons in the convolutional layer. Also, the biases were initialized as zero.

For the layers with Batch Normalization, the scale weights were initialized as 1 and the biases as 0, so that they act at the beginning only as normalization. This facilitates stable learning, since the weights will be updated gradually, avoiding abrupt jumps in these values.

The loss function implemented was the End Point Error. Which is defined as:

$$EPE = \frac{1}{B * H * W} \sum_b \sum_h \sum_w \sqrt{\left(\vec{d}(x,y)^{bhw}_{pred} - \vec{d}(x,y)^{bhw}_{GT} \right)^2} \tag{2}$$
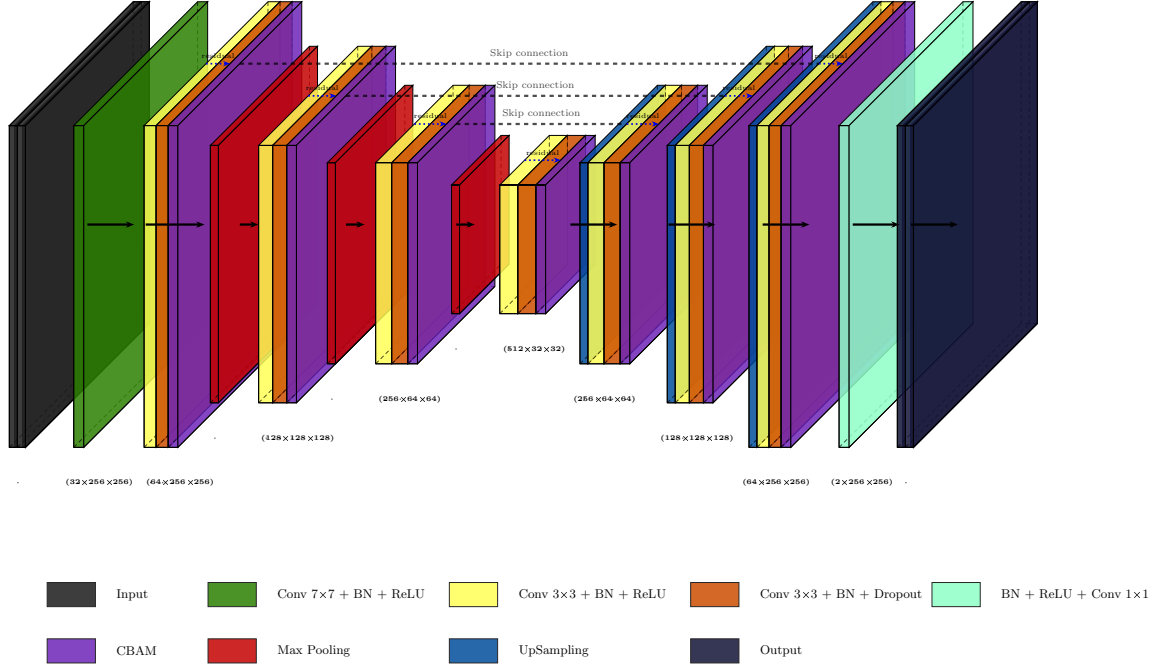
Fig. 1: Schematic representation of the proposed CNN architecture. The network follows a U-Net structure with encoder, bottleneck, and decoder blocks. Each residual block incorporates a Convolutional Block Attention Module (CBAM). The color legend identifies the different components and operations used throughout the architecture.

Where B is the batch size, (H, W) is the image shape, which in this case is (256,256) and $vecd_{pred}$, $vecd_{pred}$ are the predicted speckle pattern displacements and ground truth, respectively. That is, EPE is the Euclidean distance between the average displacement vectors per pixel within the batch.

III. DIFFERENTIAL PHASE SIGNAL INTEGRATION

The integration of the differential phase signal was done using the Fourier derivative theorem [2]. First, the complex gradient of the phase was constructed:

$$\overline{\phi}(x,y) = \frac{\partial}{\partial x}\phi(x,y) + i\,\frac{\partial}{\partial y}\phi(x,y) \tag{3}$$

Applying the 2D Fourier transform to both sides of the above equation yields:

$$\mathscr{F}\left(\overline{\phi}(x,y)\right) = 2\pi(i\,k_x - k_y)\mathscr{F}\left(\phi(x,y)\right) \tag{4}$$

where k_x and k_y are the coordinates x, y in the reciprocal space. The phase signal is finally:

$$\phi(x,y) = \mathscr{F}^{-1}\left[\frac{\mathscr{F}\left(\overline{\phi}(x,y)\right)}{2\pi(i\,k_x - k_y) + \epsilon}\right] \tag{5}$$

Where ϵ is a regularization parameter to avoid division by zero.

IV. Evaluation Sample: Hyperboloid Geometry

The evaluation geometry shown in the main manuscript is a hyperboloid sample made of PMMA. The equation of the constructed hyperboloid is:

$$\left(\frac{x}{R_x 2}\right)^2 + \left(\frac{y}{R_y}\right)^2 - \left(\frac{z}{R_z}\right)^2 = 1 \tag{6}$$

With $R_x = R_y = 2.2$mm, $R_z = 4.4$mm. A projection along the z-axis was taken in the wave-optics simulation, i.e., the X-rays propagate in that direction, as shown in Fig. 2.
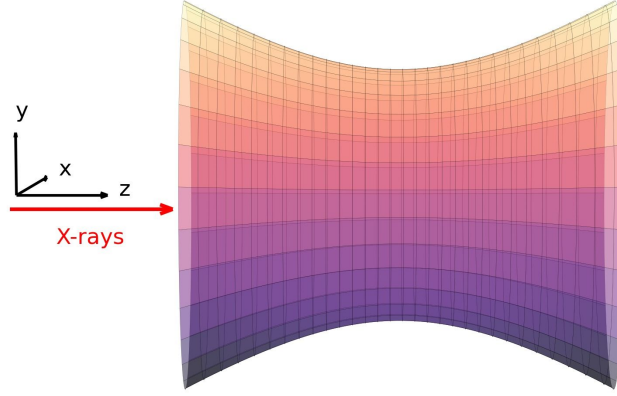

Fig. 2: Geometry of the hyperboloid used as test sample, highlighting the direction of X-ray propagation. The figure illustrates the orientation and spatial distribution of the structure relative to the beam path, which determines the local shifts in the speckle pattern.

References

[1] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," in *Proc. IEEE Int. Conf. Computer Vision (ICCV)*, 2015, pp. 1026–1034.

[2] C. Kottler, C. David, F. Pfeiffer, and O. Bunk, "A two-directional approach for grating-based differential phase contrast imaging using hard X-rays," *Optics Express*, vol. 15, no. 3, pp. 1175–1181, 2007.