Enhanced bi-LSTM for Modeling Nonlinear Amplification Dynamics of Ultra-Short Optical Pulses

Karina Saraeva † and Anastasia Bednyakova * †

Physics Department, Novosibirsk State University, Pirogova Str. 2, Novosibirsk 630090, Russia; k.saraeva@g.nsu.ru
* Correspondence: anastasia.bednyakova@gmail.com
† These authors contributed equally to this work.

Abstract: Fiber amplifiers are essential devices for optical communication and laser physics, yet the intricate nonlinear dynamics they exhibit pose significant challenges for numerical modeling. In this study, we propose using a bi-LSTM neural network to predict the evolution of optical pulses along a fiber amplifier, accounting for the dynamically changing gain profile and the Raman scattering. The neural network can learn information from both past and future data, adhering to the fundamental principles of physics governing pulse evolution over time. We conducted experiments with a diverse range of initial pulse parameters, covering the variation in the ratio between dispersion and nonlinear length, ranging from 0.25 to 250. This deliberate choice has resulted in a wide variety of propagation regimes, ranging from smooth attractor-like to noise-like behaviors. Through a comprehensive evaluation of the neural network performance, we demonstrated its ability to generalize across the various propagation regimes. Notably, our results showcase a relative speedup of 2000 times for evaluating the intensity evolution map using our proposed neural network compared to the NLSE numerical solution employing the split-step Fourier method.

Keywords: long short-term memory; recurrent neural network; Raman scattering; gain-guiding nonlinearity; fiber amplifier; numerical simulation

1. Introduction

A fiber amplifier is a crucial component of a laser system. The main challenges to deal with in active fibers are managing significant nonlinear phase accumulation without wave breaking and amplifying ultrashort pulses that are affected by strong gain shaping. Recently, there has been growing interest in a new regime for amplifying linearly chirped asymmetric pulses with gain-guiding nonlinearity (GGN), which was demonstrated in a research study by a group from Cornell University [1]. It is worth noting that this regime governs the pulse evolution in the symmetric arms of the Mamyshev oscillator [2], making it possible to achieve record-breaking peak power levels. The GGN regime is a nonlinear amplification process that occurs when high-power picosecond pulses are propagated, where the width of the spectrum is comparable to or exceeds the width of the amplification profile. This amplification process results in intricate nonlinear dynamics that lead to pulse asymmetry and the formation of a nonlinear attractor [3]. The dynamically changing amplification profile plays a crucial role in shaping the nonlinear attractor. Therefore, numerical simulations must use a complex model that considers the evolution of amplification along the fiber and its wavelength dependence.

Conventional numerical modeling presents substantial challenges for practical applications, primarily due to the time-consuming computations required for each new set of system parameters. Real experimental conditions are often difficult to fully parameterize, leading to the necessity of making assumptions and neglecting the physical model description.

Given that a fiber amplifier is the most computationally demanding component of a laser system in numerical modeling tasks, its application in real-time experimental scenarios
becomes challenging. One potential solution involves employing neural networks to predict the evolution of intensity profiles along the fiber [4–7]. Neural networks accelerate the modeling process by reducing the number of computational operations and overcoming the limitations associated with numerical simulations that rely on approximations and discretizations. Additionally, they possess the ability to generalize information, enabling the derivation of solutions from imperfect and noisy experimental data in cases in which a precise consideration of all factors influencing the experiment proves unfeasible.

In most studies employing deep learning techniques for fiber optics applications, modern architectures are rarely utilized. Instead, linear perceptrons are widely used; these do not account for temporal context and are suitable only for classification and for predicting the output pulse profile. Since the task of modeling pulse propagation through the fiber is entirely equivalent to forecasting time series, a more effective solution can be achieved by employing recurrent neural networks [8,9]. These networks, equipped with internal memory, efficiently leverage the preceding stages of pulse evolution to predict the subsequent steps. Importantly, when trained with a dataset generated through comprehensive numerical modeling, the PI-RNN becomes attuned to the physical principles governing pulse evolution in optical fibers. Other deep learning algorithms may not be able to incorporate such physics-informed features.

Our study presents the results of employing a physically informed recurrent neural network (PI-RNN) for forecasting the nonlinear evolution of the spectral and temporal pulse intensity along an active optical fiber. We have chosen the range of initial pulse parameters that covers the variation in the relation between dispersion and nonlinear length from 0.25 to 250. This choice has led to a wide variety of propagation regimes, from smooth attractor-like modes to noise-like ones. Training the RNN within this range of parameters requires generalization across various propagation modes, which is a challenging task. We demonstrate that a single PI-RNN, trained on numerical simulation results, can accurately and rapidly reproduce the intricate dynamics of a nonlinear attractor within a fiber amplifier across a wide range of initial parameters. Building upon the findings presented in [8], in which the focus was on substituting the nonlinear Schrodinger equation (NLSE)-based numerical modeling of a passive fiber with LSTM predictions, we successfully developed an architecture capable of simulating the propagation through an active fiber with a more complicated physical model involved. The novelty of our study lies in the fact that, in contrast to the majority of existing works that employ deep learning methods to explore dynamics in passive fibers, we have effectively trained PI-RNN to predict nonlinear amplification, considering a dynamically changing gain profile and the Raman scattering. Most works applying RNNs to dynamic prediction typically present only a few evolution heatmaps, hindering an accurate assessment of real predictive ability. To address this limitation, we provide comprehensive error maps that illustrate predictive performance across the parameter domain. Additionally, we delve into the capabilities and adaptability of the technique for constructing autoregressive predictions employing a cold-start initialization.

2. Numerical Model of the Amplifier

We consider pulse evolution in a typical, highly doped, ytterbium fiber amplifier. A Gaussian pulse at 1028 nm is launched into an Yb-doped fiber amplifier with a 6-µm core diameter, which is co-pumped at 976 nm. As the pulse propagates, it accumulates a pronounced nonlinear phase, resulting in a significant broadening of the spectra. When the spectrum is broadened to match the width of the gain spectrum, the proper parameters of the input pulse facilitate an evolution towards a nonlinear attractor in the GGN amplification regime. It is also worth noting that, apart from this, there exists a wide diversity of different pulse propagation regimes, often accompanied by the formation of noisy Raman pulses. For modeling highly nonlinear propagation of ultrashort pulses inside the amplifier, a complex numerical model considering a dynamically changing amplification profile and Raman scattering is required.
The numerical model employed in simulations comprises a system of coupled equations governing pulsed signal generation and continuous-wave pump [10–13]:

$$\frac{\partial A_s(z,t)}{\partial z} = -\beta_2 \frac{\partial^2 A_s(z,t)}{\partial t^2} + \int_{-\infty}^{\infty} \frac{g_s(\omega,z)}{2} A_s(z,\omega) \exp(-i\omega t) d\omega +$$

$$i\gamma \left( 1 + \frac{i}{\omega_0} \frac{\partial}{\partial t} \right) \left( A(z,t) \int_{-\infty}^{\infty} R(t') |A(z,t-t')|^2 dt' \right)$$

(1)

$$\frac{\partial P_p(z)}{\partial z} = g_p(z) P_p(z),$$

(2)

where $A_s(z,t)$ is the slowly varying envelope associated with the signal, $P_p(z)$ is the average power of continuous-wave pump, $\beta_2$ is the group velocity dispersion, $\gamma$ is the Kerr nonlinearity, $g_s$ and $g_p$ are signal and pump gain/loss coefficients, correspondingly. The response function $R(t) = (1-f_R)\delta(t) + f_R h_R(t)$ includes both instantaneous electronic and delayed Raman contributions [14]. We used the Hollenbeck vibrational model [15] to describe the Raman response function $h_R$. The spectral window considered in the model extended from 865 to 1260 nm with the central wavelength at 1028 nm. The temporal window was equal to 150 ps.

The wavelength dependence of the gain is considered in the frequency domain, where the optical field $\tilde{A}(z,\omega)$ is multiplied by the gain profile $g_s(\omega,z)$. Each spectral component of the gain $g_s(\lambda_i,z)$ ($i = 1, \ldots, N_\omega$, where $N_\omega$ — is the number of the discreet frequencies in simulations) and the pump gain/loss coefficients at each step along the fiber were found based on the rate equations in the stationary case $dN_2/dt = 0$:

$$g_s(\lambda_i,z) = \sigma_{21}^{\text{e}}(\lambda_i) \rho_s(\lambda_i) N_2(z) - \sigma_{12}^{\text{e}}(\lambda_i) \rho_s(\lambda_i) N_1(z), i = 1, \ldots, N_\omega$$

(3)

$$g_p(\lambda_i) = \sigma_{21}^{\text{p}} \rho_s(\lambda_i) N_2(z) - \sigma_{12}^{\text{p}} \rho_p N_1(z),$$

(4)

$$\frac{dN_2(z)}{dt} = \left( \sigma_{12}^{\text{p}} \frac{P_p(z)}{N_p} + \sum_{k=1}^{K} \left( \sigma_{21}^{\text{e}}(\lambda_k) \rho_s(\lambda_k) \frac{P_p(\lambda_k, z)}{N_p} \right) \right) N_1(z) -$$

$$- \left( \sigma_{21}^{\text{p}} \frac{P_p(z)}{N_p} + \sum_{k=1}^{K} \left( \sigma_{21}^{\text{e}}(\lambda_k) \rho_s(\lambda_k) \frac{P_p(\lambda_k, z)}{N_p} \right) \right) N_2(z)$$

(5)

$$N_1(z) = N - N_2(z),$$

where, $N_1, 2$ are population densities in the ground and excited energy levels correspondingly, $N = 4.8 \cdot 10^{15}$ m$^{-1}$ is the total number of Yb-ions integrated over the fiber mode cross-section, $P_p(\omega_k, z) = |\tilde{A}(z,\omega_k)|^2$ is the signal power at the frequency $\omega_k$ and position $z$ along the fibre, and $T = 850$ μs is the fluorescence lifetime. The effective pump absorption and emission cross sections at pump wavelengths of 976 nm are $\sigma_{12}^{\text{p}} = 2.5 \cdot 10^{-25}$ m$^2$ and $\sigma_{21}^{\text{p}} = 2.44 \cdot 10^{-27}$ m$^2$. The absorption and emission cross-section spectra in the considered spectral window are described by $\sigma_{12}^{\text{e}}(\lambda_i)$ and $\sigma_{21}^{\text{e}}(\lambda_i)$. The normalized pump and signal power distributions through the fiber cross-section are marked $\rho_{p,s} = \Gamma_{p,s} / \pi a^2$, where $a = 3$ μm is the core radius of a single-mode fiber, $\Gamma_{p,s}(\Gamma_s)$ corresponds to the modal overlap factor between the pump (signal) mode and the ion distribution. $\Gamma_{p} = 1$ for core pumping, $\Gamma_s = 1 - \exp(-2a^2 / w^2)$, $w$ is the 1/e electric field radius of the equivalent Gaussian spot.

We used the open-source Pyofss library [16] for numerical modeling; it has a newly added module that enables parallel computing with a Raman influence [17]. We also added our own modules for parallel amplification computing based on the Yb-coupled equations described above.

3. The Architecture of a Recurrent Neural Network

The neural network should not predict the dynamics in a particular propagation regime but should also guess the regime to be predicted. Predicting the spectrum intensity is complicated by the substantial broadening during propagation and its high modulation.
In terms of the temporal intensity prediction, the primary challenge lies in accurately forecasting the gain variation throughout the evolution. To our knowledge, there have been no attempts to simulate a fiber amplifier using a physically accurate gain model with a neural network in the field of fiber optics up to the present date.

Similarly to the study [8], we utilised a single-layer LSTM neural network architecture as a baseline. Since the baseline architecture proved to be insufficient for the dataset used, we improved the proposed architecture by incorporating stacked LSTM layers and a bidirectional cell structure. All the outputs from the LSTM cells, instead of just the last cell output, were then fed through several dense layers to refine the results further. We implemented our neural network using the PyTorch Python library [18].

The recurrent neural network’s architecture is depicted in Figure 1.

Figure 1. Scheme of the recurrent neural network used consisting of a recurrent (biLSTM) part and a fully connected part.

Here, we employ two neural networks with similar structures to predict spectral and temporal intensity evolution independently.

4. Data Preparation and Training Process

The RNN was trained with synthetic data generated using the model described in Section 2. The chosen range of initial pulse parameters spans the variation in the relationship between dispersion and nonlinear length, ranging from 0.25 to 250. The initial pulse intensity is uniformly variable, ranging from 100 W to 1000 W, while the pulse width varies logarithmically from 0.1 to 10 ps. The explored parameter space encompasses a wide diversity of pulse propagation regimes, ranging from GGN amplification to pulse amplification, accompanied by the formation of noise Raman pulses in high-intensity cases [19]. We selected a length of 7 m for the optical fiber, a choice deemed sufficient for stabilizing a nonlinear attractor, as suggested by Sidorenko et al. [1]. The training dataset comprises 879 examples of pulse evolution within the specified range of initial parameters. The RNN is trained to forecast the evolution profile at fiber intervals of 46 mm, requiring 150 steps to predict the evolution along the entire length of the fiber. No preprocessing was applied to the data, except for reducing the resolution using linear interpolation along the temporal (spectral) coordinate, from 16384 to 500 points. The choice of this dimensionality reduction was based on the resolution needed to display fine spectral modulation dynamics and the computational resources available for training the neural network, and it can be varied for different problems.

The dataset utilized for testing the model consisted of displaced grid points within the same initial pulse parameter interval. To ensure the construction of a reliable model capable of accurately predicting the evolution of different propagation regimes on a uniform grid, it
is essential to employ test and train datasets of equal size and distribution. Therefore, maintaining a 1:1 test-to-train ratio is imperative to effectively assess the prediction performance of the neural network within the chosen parameter range for this task.

The data for training and testing the neural network are prepared using the sliding window method illustrated in Figure 2a. This technique facilitates the division of the data into smaller segments. The first ten intensity profiles are interpreted as the neural network input, with the subsequent one serving as its target output. The window then slides by one point over the fiber length and repeats the process until the end of the training evolution. After creating 10+1 pairs, it is essential to shuffle these pairs from all the training evaluations to ensure a stable learning process. The prepared data are then fed into the neural network during the training process.

To find the global minimum in the loss function and ensure effective training, we employed several optimization and learning stabilization techniques. These included the Adam optimizer, hyperparameter tuning, and a learning rate scheduler. By utilizing these techniques, we aimed to enhance the training efficiency and stability of our model. Model training was performed on a local server using an NVIDIA RTX 4090 graphics processing unit (GPU).

5. Results

We used an autoregression approach to reconstruct the prediction. This method allows the neural network to forecast the data evolution for any number of steps forward by sequentially feeding its output back into its input.

5.1. Metrics Used for Tracking RNN Performance

Here, we outline the metrics used to assess the final predictive performance of the trained neural network. The network is trained to predict a single pulse profile by leveraging the evolution dynamics extracted from a sequence of profiles using the mean squared error (MSE) loss:

\[
MSE(I, \hat{I}) = \frac{1}{N} \sum_{i=1}^{N} (I_i - \hat{I}_i)^2,
\]

where \( I \) represents the temporal (spectral) intensity array at a fixed z coordinate along the fiber, and \( N \) denotes the size of the temporal (spectral) domain.

In autoregression prediction, the model effectively handles newly self-generated data that were not part of the training sample. To evaluate the final result, a different metric was employed. The normalized root mean square error (NRMSE) metric, widely utilization...
in similar tasks, provides a robust accuracy estimation for predicting errors in individual intensity profiles.

$$\text{NRMSE}(I, \hat{I}) = \sqrt{\frac{\sum_{i=1}^{N_{\text{domain}}} (I_i - \hat{I}_i)^2}{\sum_{i=1}^{N_{\text{domain}}} (\hat{I}_i)^2}}$$ (7)

However, when applied to the entire evolution map, the NRMSE tends to overestimate errors in the case of low-energy pulses with narrow spectra and underestimate errors in high-energy propagation regimes. Additionally, interpreting the quality of the prediction is not straightforward. In an attempt to address these issues, we propose using another version of the normalized MSE—the peak signal-to-noise ratio (PSNR) metric [20]:

$$\text{PSNR}(I_{\text{map}}, \hat{I}_{\text{map}}) = 20 \cdot \log_{10} \left( \frac{\max(\hat{I}_{\text{map}})}{\sqrt{\text{MSE}(I_{\text{map}}, \hat{I}_{\text{map}})}} \right),$$ (8)

where $I_{\text{map}}$ is a 2D array that represents the evolution of a temporal (spectral) intensity array along the spatial coordinate $z$.

The PSNR is frequently employed for assessing the quality of reconstructed images and offers several advantages over the NRMSE metrics, including a decibel scale and normalization based on the maximum intensity.

5.2. Forecasting of Temporal and Spectral Intensity Evolution

We assessed the prediction error of the PI-RNN model using Equation (7) and depicted the temporal and spectral intensity maps in Figures 3a and 4a, respectively. The plots illustrate the smooth evolution of errors along the fiber for all initial pulse parameters, with no discernible discontinuities or outliers within the prediction maps.

Figures 3b and 4b, calculated using Equation (8), illustrate the comprehensive error maps across the plane of initial pulse parameters. These maps reveal the impact of initial pulse parameters on the overall PSNR error along the fiber. The interpretation of the error diagrams is as follows: the prevailing trend indicating the most significant errors corresponds to situations with numerous high-intensity modulation peaks or noise-like generation. Specifically, in the prediction of the temporal evolution, the highest accumulated error is associated with the Raman generation. When predicting the spectral evolution, the highest error is associated with regions exhibiting significant nonlinear phase accumulation before transitioning to a smooth attractor-like regime or regions characterized by the emergence of a noisy Raman pulse.

Figure 3. Temporal error maps for the test dataset, showing a 140-step prediction using 10 inputs for initiation. (a) Illustration of the NRMSE evolution in predicting temporal intensity depending on the distance along the fiber; (b) dependency of the PSNR metric on various initial parameters; points labeled in red correspond to examples of propagation regimes shown in Figure 5.

Figures 3b and 4b, calculated using Equation (8), illustrate the comprehensive error maps across the plane of initial pulse parameters. These maps reveal the impact of initial pulse parameters on the overall PSNR error along the fiber. The interpretation of the error diagrams is as follows: the prevailing trend indicating the most significant errors corresponds to situations with numerous high-intensity modulation peaks or noise-like generation. Specifically, in the prediction of the temporal evolution, the highest accumulated error is associated with the Raman generation. When predicting the spectral evolution, the highest error is associated with regions exhibiting significant nonlinear phase accumulation before transitioning to a smooth attractor-like regime or regions characterized by the emergence of a noisy Raman pulse.
Figure 4. Spectral error maps for the test dataset, showing a 140-step prediction using 10 inputs for initiation. (a) Illustration of the NRMSE evolution in predicting spectral intensity depending on the distance along the fiber; (b) Dependency of the PSNR metric on various initial parameters; points labeled in red correspond to examples of propagation regimes shown in Figure 6.

We conducted tests on various data preprocessing methods, including normalization and logarithmic transformation, which have demonstrated effectiveness for passive fibers, as reported in previous works [8,9]. However, we observed that these methods diminished the information content of the spectral intensity evolution within the active fiber. Recognizing the significance of spectral intensity amplitude as an additional feature aiding the neural network in precise predictions and evolution stage determination through autoregression, we opted to use the spectral intensity data in the original linear scale without any preprocessing.

To compare pulse propagation maps predicted by the PI-RNN with numerical modeling using the NLSE, we picked three distinctive regimes from the data. Three main types of behaviors include a smooth GGN regime resulting in a formation of a nonlinear attractor, a transient regime, and a regime showing a pronounced influence of the stimulated Raman scattering on a pulse amplification. Figure 5 shows examples of the temporal evolutions, along with their locations in the error map displayed with red labeled points in Figure 3b. Figure 6 shows typical spectral intensity propagation regimes, with their locations in the error map in Figure 4b. The Raman scattering manifests itself as the generation of a noise-like pulse with energy comparable to the main pulse, downshifted by about 13.2 THz in frequency (Figure 6c) [21].

The PI-RNN demonstrates its capability to model the temporal and spectral evolution for any point within the training parameter area.

Autoregressive reconstruction of the evolution map along the 7-meter-long fibre, as presented in the Figures 5 and 6, takes about 0.05 s with PI-RNN. This is approximately 700 times faster than the fastest paralleled numerical NLSE model when using the NVIDIA RTX 4090 GPU and 2000 times faster than the conventional CPU-based numerical model. This notable difference is attributed to the reduction in the number of numerical operations required for each step along the fiber, a decrease in the total number of steps needed to obtain a solution, and the lower temporal (spectral) resolution required for the computations. The final PI-RNN model has an estimated tens of millions of trainable parameters.
5.3. Resistance to Noise

The neural network, trained on undisturbed synthetic data, was found to be robust to external noise in the test dataset. The PI-RNN can capture the pulse propagation regime up to the signal-to-noise ratio values of 20 dB. Specifically, when given an initial pulse with added ‘white’ Gaussian noise, the neural network predicts the output pulse with noise suppression. This feature allows for greater practical significance of this study in the future, as it can handle realistic experimental data that may contain noise or other imperfections.

5.4. Autoregression Problems

Overfitting poses a significant challenge in autoregressive prediction problems. In autoregressive predictions, the neural network establishes a feedback loop, using slightly perturbed input data influenced by its own prediction errors to make subsequent predictions. This feedback loop complicates the monitoring and prevention of overfitting, since the neural network is only trained to predict only one step initially and is not explicitly trained for autoregressive prediction. Even though the model has high accuracy
for one-step prediction on a validation dataset, it may face issues when estimating autoregressive predictions on a test dataset. The slightly undertrained model seems to be more robust to overfitting the numerical modeling data when predicting the evolution autoregressively.

Figure 6. Comparison between the spectral intensity evolution calculated using the NLSE and the prediction made with PI-RNN. The prediction is built 140 steps ahead along the fiber using 10 consecutive pulses as input. (a) $P_0 = 432\, \text{W}, T_0 = 0.16\, \text{ps}$, (b) $P_0 = 810\, \text{W}, T_0 = 2.19\, \text{ps}$, (c) $P_0 = 460\, \text{W}, T_0 = 8.9\, \text{ps}$.

5.5. Cold Start Problems

Cold start is a method of evolution reconstruction that starts with a single initial pulse profile, which is then fed to all RNN inputs. This method facilitates the reconstruction of the autoregressive evolution map using only a single pulse profile, thereby simplifying its application for various tasks.

To improve the model’s performance in cold start prediction scenarios, we used a specific approach to prepare the training data. The main concept involves incorporating “cold start data” into the training sample, as shown in Figure 2b. “Cold start data” are a synthetic type of data that mimic the model’s task of reconstructing evolution from one tiled initial profile. The basic idea is to replace the first $n$ evolution profiles, where $n$ is the number of inputs for the recurrent neural network, with the initial profile. This approach aids the model in improving predictions by learning from these artificial examples.
We obtained results with comparable accuracy for predicting cold start temporal and spectral evolutions as the prediction based on the starting pulse sequence.

6. Conclusions

In this paper, we explored the potential of using PI-RNN to predict the evolution of spectral and temporal pulse intensity along the fiber amplifier—a computationally challenging task in nonlinear optics. We introduced an updated PI-RNN architecture designed to learn the complex dynamics of the optical field from a large dataset generated via numerical simulations and conducted a thorough evaluation of its performance. We found that the PI-RNN can accurately and precisely estimate the evolution map, outperforming the conventional split-step numerical solution of the NLSE by a significant margin—2000 times faster. This improved speed persists even when parallelized with a GPU, resulting in a 660-fold faster computation.

Furthermore, the PI-RNN adeptly performs interpolation and extrapolation of the field evolution along the fiber with reasonable accuracy. It exhibits adaptability to various grid sizes along the z-coordinate. Notably, a single neural network proved capable of capturing diverse propagation regimes, making it a universal tool for investigating dynamics within the chosen parameter space. The PI-RNN approach can achieve further enhancement by integrating experimental data into the training set. This can not only improve the time performance but also enhances its descriptive capabilities, surpassing numerical modeling.

We have also detailed the challenges encountered during the PI-RNN training process for autoregressive prediction tasks. One significant drawback identified in reconstructing the entire evolution map using PI-RNN is its cold start prediction performance, heavily dependent on the data sample provided. To address this issue, we introduce a novel approach to preparing training data, resulting in improved cold start execution. This modification leads to notable performance enhancements, particularly in predicting the temporal intensity map using a single input profile.

In conclusion, we assert that the PI-RNN has proven to be a promising technique for predicting the evolution of pulse intensity along an active fiber, demonstrating substantial advantages over traditional numerical methods.

Author Contributions: Conceptualization, A.B.; methodology, K.S. and A.B.; software, K.S.; validation, A.B. and K.S.; investigation, K.S. and A.B.; writing—original draft preparation, K.S. and A.B., visualization, K.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Ministry of Science and Higher Education of the Russian Federation (Project No. FSUS-2021-0015).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

- NLSE: Nonlinear Schrödinger equation.
- LSTM: Long short-term memory.
- MSE: Mean square error.
- PSNR: Peak signal-to-noise ratio.
- GPU: Graphics processing unit.
References

1. Sidorenko, P.; Fu, W.; Wise, F. Nonlinear ultrafast fiber amplifiers beyond the gain-narrowing limit. *Optica* 2019, 6, 1328–1333. [CrossRef] [PubMed]


Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.