Towards ML-Based Diagnostics of Focused Laser Pulse

Y. R. Rodimkov1*, V. D. Volokitin1,2**, I. B. Meyerov1,2***, and E. S. Efimenko3****

(Submitted by A. B. Muravnik)

1Department of Mathematical Software and Supercomputing Technologies, Lobachevsky University, Nizhny Novgorod, 603950 Russia
2Mathematical Center, Lobachevsky University, Nizhny Novgorod, 603950 Russia
3Institute of Applied Physics of the Russian Academy of Sciences, Nizhny Novgorod, 603950 Russia

Received September 3, 2022; revised October 22, 2022; accepted November 8, 2022

Abstract—Currently, machine learning (ML) methods are widely used to process the results of physical experiments. In some cases, due to the limited amount of real experimental data, ML-models can be pre-trained on synthetic data simulated based on the analytical theory and then fine-tuned using real experimental data. A limitation of this approach is the presence of the latent parameters of the analytical model, which values are difficult or impossible to estimate. Setting these parameters incorrectly may induce a dataset shift even when applied to synthetic data. To overcome this problem, we train the ML-model on a dataset with randomly varied latent parameters of the analytical model to force the ML-model to concentrate on more general patterns that depend weakly on the latent parameters. We applied this approach to the problem of tight focusing of a laser pulse with the complex structure of the wavefront. We observed good accuracy of reconstructing of the tilt parameters when training and testing the ML-model on datasets generated for different values of the latent parameters. This confirms that the ML-model was able to select relevant information without over-fitting for specific features inherent in certain values of the latent parameters. We believe that this approach will enrich possible applications of ML-methods to an experimental diagnostics of laser pulses.

DOI: 10.1134/S199508022301033X

Keywords and phrases: tight focusing, deep learning, convolutional neural network, dataset shift.

1. INTRODUCTION

With the progress in optics technologies in the last decades, laser facilities can now achieve unprecedented intensities. Femtosecond laser pulses focused to spot size of the order of a few micrometers can now routinely achieve intensities above $10^{22}$ W/cm$^2$ [1–3], with a record intensity even exceeding $10^{23}$ W/cm$^2$ [4]. These ultra-relativistic laser pulses allow for multiple applications such as particle acceleration in plasmas [5, 6] and production of $\gamma$-rays [7], or exploring quan-to-electrodynamic effects in laser-plasma interaction in the high-field regime [8].

Such intensities are reached by the OPCPA technique, which successively stretches and compresses the pulse for further amplification [9]. Any imperfection in this process in the laser chain can thus introduce spatio-temporal couplings (STC), i.e. correlations between the longitudinal and transverse intensity profiles, which ultimately reduces the intensity at focus [10, 11]. STC tends to become stronger for more powerful laser systems — corresponding to an increasing waste of the pumping energy, which becomes greatly harmful for applications.

*E-mail: rodimkov@bk.ru
**E-mail: valyav95@mail.ru
***E-mail: meerov@vmk.unn.ru
****E-mail: evgeny.efimenko@ipfran.ru
Characterization of the laser pulse is thus crucial both for optimizing the laser profile in experiments and for its accurate numerical description without overestimation of the peak intensity—leading to better agreement with the experimental results and thus ultimately to a deeper understanding of the underlying physics.

Spatio-temporal characterization of the ultra-intense laser pulses is however not straightforward. Direct detection is precluded due to the high frequency of the pulse, and a combination of spectroscopy and interferometry has to be used. 3D characterization is thus a trade-off between cost, time of acquisition and resolution of the data. Recently, experimental techniques leading to a complete spatio-temporal characterization of ultra-intense laser pulses have been developed, such as TERMITE [12] or INSIGHT [13], which have been applied to top-class PW-laser facilities [14]. Such measurements are however obtained through extensive manipulations, and effort are also pursued for simplifying this process through the combination of experimental measurements and machine learning. First attempts in this sense were realized in the 90’s, in which neural network [15] and genetic algorithm [16] were used with frequency-resolved optical gating (FROG) for reconstructing the pulse phase. More recently, a neural network trained on simulated data retrieved the pulse phase even in the presence of high noise, and lowered the required knowledge about the relation between the pulse and its measurement [17]. Other works have also proposed to reconstruct the pulse phase with neural network from dispersion scan traces [18], 2D intensity patterns [19] or with a multimode fiber [20].

Training data for machine learning (ML) models can be collected experimentally or through simulation. Quite often the amount of experimental data is not sufficient for training the ML-model from scratch. In this case the training data can be numerically simulated on the basis of some analytical theory, and then fine-tuned on a real experimental data. This approach, although seems very natural, requires good correspondence between the analytical theory and experiment. The later can become very problematic, because the analytical model may have hidden or latent parameters which affect the results, but can not be directly measured or estimated in the experiment. The number of such latent parameters for ML-models can be large. Setting parameters incorrectly during data collection can cause a ML-model to break when applied to real data due to dataset shift.

In this article, we demonstrate an approach to reconstructing the physical parameters of a tightly focused laser pulse based on the energy flux distribution in the case of the latent parameters uncertainty. In order to make the laser pulse model closer to experimental conditions we impose a spectral-dependent tilt on its wavefront, and the properties of this tilt act as the latent parameters of the model. The resulting analytical model is rather simplified representation of the real experimental conditions, nevertheless it can help to find better approaches for the problem. This parameterized model with randomly varied parameters was used for synthetic data generation. We use this data to train the ML-model to solve the inverse problem of reconstructing pulse parameters. In order to study the influence of the choice of the latent parameters on accuracy of the ML-model we study the generalization ability of this model by training and testing on different subsets of the latent parameters values. By doing this we demonstrate tolerance of the ML-model to wrong choice of the latent parameters during training, that can be useful for the application of ML-methods to real experimental data.

In a physical experiment a diagnostics can be made based on a beam image taken by CCD camera, which registers the total energy flux. The information obtained from this image can be used for adjusting laser setup, manually or by means of adaptive optics. To model this approach numerically we simulated the propagation of a tightly focused laser pulse to the focal plane, where the cumulative energy flux was calculated to be used for the ML-model training.

We numerically solved the system of Maxwell equations in vacuum

$$\nabla \times \mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t}, \quad \nabla \times \mathbf{B} = \frac{1}{c} \frac{\partial \mathbf{E}}{\partial t}$$
in Cartesian coordinates using a spectral approach [21]. This set of equation does not include currents, because we solve initial problem by setting the initial field distribution far from the focal plane.

The modeling was performed by means of the Hi-Chi framework [22] with an implementation of the PSATD method [23]. This method operates in a Fourier space and uses analytical integration over time making it dispersion-free. This property allows accurate modeling of propagation of a tightly focused ultrashort laser pulse even with resolution of 8–10 points per wavelength. Moreover, this method is absolutely stable and allows using arbitrary time step, thus the propagation of a tightly focused laser pulse to a focal plane was done in one iteration. The latter is very important for collecting ML-dataset that consists of thousands of examples. The periodic boundary conditions inherent to FFT-based solvers were used together with mapping to periodic space approach [22] to reduce the size of the grid.

The initial distribution of laser field was set as the pulse with a spherical wavefront, propagating along the $x$-axis, so any field component $u$ in the far field zone can be set as

$$u(x, y, z) = u_l(r(x, y, z)) u_1(\alpha(x, y, z)),$$

where $r = |\mathbf{r}| = |\mathbf{R} - \mathbf{R}_0|$ is the distance from the center of the spherical wavefront $\mathbf{R}_0$ to the point $\mathbf{R}$ and $\alpha = \arcsin \sqrt{y^2 + z^2}/r$ is the angle between the $x$-axis and vector $\mathbf{r}$. The electromagnetic fields are defined as $\mathbf{E}(\mathbf{R}) = u(\mathbf{R}) \mathbf{s}_0$ and $\mathbf{B}(\mathbf{R}) = u(\mathbf{R}) \mathbf{s}_1$ with the normalized vectors $\mathbf{s}_0 = e_y \times e_y \times \mathbf{r}$ and $\mathbf{s}_1 = e_y \times \mathbf{r}$.

The longitudinal $u_l$ and transverse $u_t$ profiles are defined as follows

$$u_l(x) = \sin \left( \frac{2\pi x}{\lambda_0} + \varphi \right) \exp \left( -\frac{x^2}{2L^2} \right) \Pi(x, -4L, 4L),$$

$$u_t(\alpha) = \Pi \left( \alpha, 0, \theta - \frac{\varepsilon}{2} \right) + \cos^2 \left( \frac{\pi(\alpha - \theta + \varepsilon/2)}{2\varepsilon} \right) \Pi \left( \alpha, \theta - \frac{\varepsilon}{2}, \theta + \frac{\varepsilon}{2} \right),$$

where $\lambda_0$ is the central wavelength, $L = 4\lambda_0$ is the length of the pulse and $\Pi$ is the rectangular function, with $\Pi(x, x_{\min}, x_{\max}) = 1$ if $x$ is in $[x_{\min}, x_{\max}]$ and 0 otherwise. In this case $u_l$ corresponds to a transverse flattop profile with no intensity outside of the opening angle $\theta$ with the edge smoothing angle $\varepsilon$. The opening angle $\theta$ is connected to the F-number of the focusing optics by $\theta = \arctan(1/2F)$. An additional phase $\varphi$ is used to model the imperfections of the wavefront that occur in experimental conditions.

The proposed pulse definition does not account for complex structure of wavefront inherent to real experimental conditions. In order to model complex wavefront structure in a simplified manner we introduce three spectral components by applying mask functions $f_i(k)$ (see Fig. 1) to the spectra of the longitudinal field $u_{li}(k) = u_l(k) f_i(k)$, where $u_l(k) = \int_{-\infty}^{\infty} u_l(x) e^{-ikx} dx$ is the spatial Fourier spectra of the longitudinal profile $u_l(x)$ and $k$ is the wavenumber. Three spectral components are separated by the two wavenumbers $k_1$ and $k_2$ such as $k_1 > k_0 > k_2$ with $k_3 = 2\pi c/\lambda_i$, where $c$ is the speed of light. These two boundary wavenumbers $k_{1,2}$ are varied in the range $[k_{i_{\min}}, k_{i_{\max}}]$ to parameterize the laser pulse definition. For convenience, later in the text we operate with corresponding wavelengths $\lambda_{1,2}$ as boundary wavelengths and treat them as the latent parameters of the model.

![Fig. 1. Mask functions $f_i(k)$ used to create three spectral components with different tilt properties.](image)
For each of these three spectral component we add a linear tilt by adding the additional phase \( \varphi_i \) in the longitudinal profile \( u_i(x) \) in the Eq. (2)

\[
\varphi_i = a_i y_i' = y \cos \theta_i - z \sin \theta_i,
\]

where \( a_i \) is a constant characterizing the amplitude of the tilt and \( \theta_i \) is an angle characterizing the direction of the tilt in the transverse plane. Further in this paper \( a_i \) and \( \theta_i \) are referred to as the tilt amplitude and the tilt angle, respectively.

After applying the mask function we calculate longitudinal profile \( u_{i,1} \) for each spectral component by performing inverse Fourier transform \( u_{i,1}(x) = 1/2\pi \int_{-\infty}^{\infty} u_i(x)e^{ikx} \, dk \) and calculate the corresponding spectral component initial distribution \( u_i \) according to Eq. (1): \( u_i(x, y, z) = u_{i,1}(r) \, u_i(\alpha) \).

After initialization in the far field zone each component is propagated to the focal plane using the spectral solver. Using the linearity of Maxwell’s equations, we can then get the total field at the focal position as the sum of fields of all spectral components as \( u_{\text{tot}}(R) = \sum_i A_i u_i(R) \), with \( A_i \) defining the relative spectral component amplitude compared with the central component corresponding to the mask function \( f_0 \) (i.e., \( A_0 = 1 \)). These relative amplitude parameters \( A_{1,2} \) act as another pair of latent parameters of our laser pulse model.

Finally, we calculate the cumulative energy flux in the focal plane

\[
I(y, z) = \frac{1}{4\pi} \int_{-\infty}^{\infty} E_{\text{tot}}(x, y, z) \times B_{\text{tot}}(x, y, z) \cdot \mathbf{e}_x \, dx.
\]

Summarizing, in our simplified model the energy flux depends on following parameters: the boundary wavelengths \( \lambda_{1,2} \), the relative amplitudes \( A_{1,2} \), the tilt amplitudes \( a_i \) and the tilt angles \( \theta_i \) with \( i \in \{0, 1, 2\} \). The last six parameters \( a_i \) and \( \theta_i \) were predicted by the ML-model, while \( \lambda_{1,2} \) and \( A_{1,2} \) acted as the latent parameters of the pulse model. Data was generated using central wavelength \( \lambda_0 = 800 \text{ nm} \), F-number of the focusing optics \( F=1.5 \), the pulse is initialized at \( 65\lambda_0 \) from the spherical wavefront center. Size of the grid was \( 150 \times 150 \times 60 \) with \( \Delta x = \Delta y = 0.3\lambda_0 \) and \( \Delta z = 0.1\lambda_0 \). Other parameters were set in the following ranges: \( \lambda_1 \in [600 \text{ nm}, 750 \text{ nm}] \), \( \lambda_2 \in [850 \text{ nm}, 1000 \text{ nm}] \), \( A_{1,2} \in [0.5, 2] \), \( a_{0,1,2} \in [0, 0.6] \) and \( \theta_{0,1,2} \in [0, 2\pi] \).

3. RESULTS

3.1. Methodology

The problem under consideration is the inverse problem of predicting of six parameters (the tilt angle and the tilt amplitude for three spectral components) based on the cumulative energy flux in the focal plane. The analytical model has latent parameters, which can not be directly reconstructed from the physical experiment. In our rather simplified model such parameters are the boundary wavelengths \( \lambda_{1,2} \) and the relative amplitudes of the spectral component with the lowest and highest wavelengths, respectively, \( A_{1,2} \). Determination of these parameters in the physical experiment, in which the ML-model will be used after training on synthetic data, is a difficult task. Thus, methodologically, the work is divided into two stages. At the first stage, the inverse problem of reconstructing the tilt parameters \( a_i \) and \( \theta_i \) is solved. In order to do this a suitable architecture of the ML-model is chosen, and the best way for reconstructing the parameters is searched for. At the second stage, after the inverse problem is solved, the generalization ability of the proposed ML-model is analyzed. To do this, the ML-model is trained on a subset of latent parameters values, and then tested on the entire available set of latent parameter values, i.e. the test set includes values of latent parameters that were not present in the training set. Such numerical experiments will help to answer the main question of interest of this paper, whether the trained ML-model is able to generalize the extracted information to unknown values of latent parameters or not. Hereinafter, experiments are understood as numerical experiments, unless explicitly stated otherwise. The data and scripts required to reproduce the numerical results may be downloaded from https://github.com/hi-chi/Machine-Learning (the relevant examples are located in the “Focus ML” folder).
$k_1 = 1, k_2 = 3$  \quad  k_1 = 3, k_2 = 1  \quad  A_1 = 1.1, A_2 = 2.0  \quad  A_1 = 1.4, A_2 = 0.8$

Fig. 2. Examples of the cumulative energy flux distribution for different values of latent parameters. (a, b) $\lambda_1$ and $\lambda_2$ are different, $A_{1,2}$ are fixed; (c, d) $A_1$ and $A_2$ are different, $\lambda_{1,2}$ are fixed. Parameters $a_i$ and $\theta_i$ are fixed for (a, b) and (c, d).

3.2. Methods and Metrics

The cumulative energy flux in the focal plane can be treated as a single-channel image, see Fig. 2. Neural networks are one of the most versatile methods for analyzing data with the same type of continuous features. For image analysis, the best results at the moment are shown by convolutional neural networks [24, 25] and transformers [26, 27]. It is known that transformers are difficult to train, they are very demanding on the amount of data, and also require pre-training. At the moment, in computer vision problems, there are not a large number of successful applications of this method without model pre-training [28, 29]. Therefore, convolutional neural networks were chosen for the ML-based analysis of the flux.

After several experiments a convolutional neural net architecture was chosen that consists of two consecutive blocks. Each block included two convolutional layers and a max pooling layer, with the number of convolutions 16 and 32 in each block, respectively. These two blocks were followed by four fully connected layers with the number of neurons 1024, 512, 128, 32, and finally there was a layer that predicts 6 tilt parameters. ReLU was used as an activation function on each convolutional and fully connected layer, except for the last one. For training, the root-mean-square error and the Adam [30] optimizer with default parameters in the Keras framework were used. The model was trained for 30 epochs.

To assess the quality of ML-models in our study we computed the Mean Absolute Percentage Error (MAPE) and the coefficient of determination ($R^2$):

$$\text{MAPE} = \frac{100}{n} \sum_{i=1}^{n} \frac{|\hat{y}_i - y_i|}{\max(y)}, \quad R^2 = 1 - \sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{(\bar{y} - y_i)^2},$$

where $\hat{y}_i$ and $y_i$ are the true and the predicted value of the $i$th object, respectively, and $\bar{y}$ is the value of the predicted parameter averaged over the training set.

For the reconstructed parameters a linear transformation to the range from $-1$ to 1 was used. Each energy flux profile was normalized in the range from zero to one. In all experiments Keras and Tensorflow frameworks were used to train neural networks.

3.3. The Solution of the Inverse Problem

Our goal was to reconstruct the tilt amplitude $a_i$ and the tilt angles $\theta_i$ for three spectral components of a laser pulse based on the profile of the cumulative energy flux in the focal plane. The study of the data and experiments showed that in order to achieve good accuracy the reconstructed parameters should be carefully chosen. First, the tilt angles $\theta_i$ have an obvious periodicity with a period of $2\pi$, and therefore require a special treatment. There are different approaches to this problem, for example, using a special loss function that takes into account the periodicity of angles, or reconstructing trigonometric functions of angles instead of angles themselves. Second, when the tilt amplitude $a_i$ is small, the dependence of the flux on the corresponding tilt angle $\theta_i$ becomes weak, i.e. specific angle values are indistinguishable for the ML-model, which leads to a strong increase in the angles-related error. This is caused by the fact that the tilt angle and amplitude are not used separately in computation of the additional phase $\phi_i$ (see Eq. (3)), but as the product of the tilt amplitude and the trigonometric functions...
Table 1. Average values and standard deviations of the coefficient of determination ($R^2$) and the mean error (MAPE) for reconstructing of derived parameters $p_i$ and $q_i$ by trained convolutional network. Metrics were averaged over 10 runs with different splits into training and test samples in a 90/10 ratio.

| Metrics   | $p_0$        | $p_1$        | $p_2$        | $q_0$        | $q_1$        | $q_2$        |
|-----------|--------------|--------------|--------------|--------------|--------------|--------------|
| $R^2$     | 0.991 ± 0.001| 0.944 ± 0.008| 0.980 ± 0.003| 0.991 ± 0.002| 0.948 ± 0.007| 0.981 ± 0.003|
| MAPE, %   | 2.71 ± 0.21  | 5.70 ± 0.42  | 3.77 ± 0.32  | 2.74 ± 0.24  | 5.66 ± 0.45  | 3.78 ± 0.33  |

of the tilt angle. Experiments showed that instead of reconstructing six original parameters $a_i$ and $\theta_i$ it makes sense to reconstruct six derived parameters that are the combination of the original ones $p_i = a_i \sin \theta_i$, $q_i = a_i \cos \theta_i$. This transformation can be interpreted as a transition from polar coordinates to Cartesian coordinates. It should be noted that the proposed transformation is one-to-one, and the original parameters $a_i$ and $\theta_i$ can be reconstructed from derived parameters, but the accuracy of angle reconstructing for the close to zero tilt amplitude can still be poor.

The ML-model architecture and training parameters are described in Section 3.2. For the ML-model training an experimental dataset of 50000 samples was numerically generated. The experimental results showed that the parameters $p_i$, $q_i$ are reconstructed with an accuracy close to ideal (see Table 1). It should be noted that the parameters $p_1$ and $q_1$ are reconstructed worse than other parameters, although the accuracy is still high. This effect is systematic and was observed in all experiments. Apparently, this result is due to the fact that these parameters correspond to the spectral component with the shortest wavelength, information about which is worse extracted from the cumulative energy flux. The accuracy of reconstructing of the parameters related to the central spectral component $p_0$ and $q_0$ is the best, which can be explained by the fact that both its amplitude and the central wavelength are fixed. Thus, the use of a convolutional network makes it possible to solve the inverse problem with a very high accuracy.

3.4. Analysis of the Generalization Ability of the ML-Model

3.4.1. The main idea. In the previous section, we demonstrated the high accuracy of solving the inverse problem using the convolutional neural network trained on model data. At the same time, the application of ML-models trained on synthetic data to the real physical experiment can be problematic, because the wrong choice of values of the latent parameters of the analytical model during training may significantly affect the accuracy of the ML-model on real experimental data, which is explained by a change in the data distribution. In this section we address the ability of the ML-model to cope with such negative effects using synthetic test dataset with other values of the latent parameters.

For a better understanding of the importance of a correct choice of the latent parameters values, it makes sense to visually evaluate the influence of the choice of boundary wavelengths $\lambda_{1,2}$ and relative amplitudes $A_{1,2}$ on the cumulative energy flux profile in the considered analytical model (see Fig. 2). Local changes in the energy distribution can be noticeable, for example the distribution at the bottom in Fig. 2 (a) is different from the case with the same parameters except $\lambda_{1,2}$ in Fig. 2 (b). Similar differences can be seen if compare Fig. 2 (c) and (d), where all parameters are the same except relative amplitudes $A_{1,2}$. These differences show the data variability, but what is more important, is that the general shape of the energy flux distribution in the profile is preserved, which indicates that a significant part of the relevant information is contained in the tilt angles and amplitudes.

This fact demonstrates the possibility of information generalization for different values of latent parameters, although the good result is not guaranteed. It is known that ML-models can overemphasize irrelevant information. For example, in adversarial attacks, when a change in the data is selected in a special way to change the prediction of the network [31, 32]. There are also examples when irrelevant features were changed in the data, which led to a dataset shift problem and a decrease in the accuracy of the model [33, 34]. In this regard, it seems important to evaluate the influence of the choice of latent parameters on the accuracy of the previously considered ML-model, as well as the ability of the ML-model to extract relevant information even in the case of previously unseen values of the latent parameters.

Based on this, the methodology for further experiments is following: the previously considered convolutional network is trained from scratch on a subset of the data with a reduced set of latent parameters.
parameter values, and then it is tested on a test dataset that includes all values of latent parameters, including those not present in the training dataset. By doing this we try to quantify the effect of wrong choice of latent parameters values during training on the final accuracy of the ML-model. The experiments are divided into two parts. In the first part, the boundary wavelengths $\lambda_{1,2}$ act as the studied latent parameters, and in the second part, the same is done for the relative amplitudes $A_{1,2}$.

### 3.4.2. Generalization for different wavelengths.

In the first experiment, the ML-model was trained on a newly generated dataset with 50000 samples with the values of the boundary wavelengths $\lambda_1 \in \{650\,\text{nm}, 700\,\text{nm}\}$ and $\lambda_2 \in \{900\,\text{nm}, 950\,\text{nm}\}$. The ML-model was tested on a dataset with 50000 samples from Section 3.3 with $\lambda_1 \in \{600\,\text{nm}, 650\,\text{nm}, 700\,\text{nm}, 750\,\text{nm}\}$ and $\lambda_2 \in \{850\,\text{nm}, 900\,\text{nm}, 950\,\text{nm}, 1000\,\text{nm}\}$. All other parameters were randomly changed. The model architecture and training parameters are the same as those used in Section 3.2. The results of the experiments are shown in Fig. 3.

It can be seen that the accuracy of determining the parameters on data with the boundary wavelengths that are not included in the training set slightly decreases. The ML-model shows the worst result when the values of the boundary wavelengths become close to each other. In this case, the difference between the spectral components from different spectral ranges decreases, which complicates the reconstructing of parameter values.

### 3.4.3. Generalization for different amplitudes.

In the second experiment, the ML-model was trained on a newly generated dataset with 50000 samples with the values of the relative amplitudes $A_{1,2} \in \{0.95, 1.1, 1.25, 1.4\}$. The ML-model was tested on a dataset with 50000 samples from Section 3.3 with $A_{1,2} \in \{0.5, 0.65, 0.8, 0.95, 1.1, 1.25, 1.4, 1.55, 1.7, 1.85, 2\}$. All other parameters were randomly changed. The model architecture and training parameters are the same as those used in Section 3.2. The results of the experiments are shown in Fig. 4.
is similar: when the value of the parameter $A_1$ decreases (see Fig. 4b). In this case, the relative contribution of this spectral component to the energy becomes smaller, which complicates the determination of its parameters. For parameters $A_2$, we can see that the accuracy falls sharply if the value of the parameter $A_1$ decreases (see Fig. 4b). In this case, the relative contribution of this spectral component to the energy flux becomes smaller, which complicates the determination of its parameters. For parameters $A_2$, the situation is similar: when the value of the parameter $A_2$ takes on minimum values, the accuracy of reconstructing the parameters decreases. In the case of parameters $p_0$ and $q_0$, the accuracy deteriorates when the values of the parameters $A_1$ or $A_2$ are close to the maximum values. Thus, it can be summarized that in the case of a change in the relative amplitudes of various spectral components, the key factor is the distinguishability of the corresponding spectral component against the background of others. In other words, the relative energy contribution of the spectral component to the cumulative energy flux should be noticeable.

### 4. DISCUSSION

This article is devoted not only to solving the problem of reconstructing the tilt parameters of a laser pulse, but to a greater extent to studying the generalization ability of the ML-model. The results of the experiment, when the ML-model is trained on some values of latent parameters, and tested on other values, can be treated differently. From the point of view of latent parameters, the available information is extrapolated by the ML-model beyond the set of parameters values used for training. Generally speaking, the possibility of solving such type of problems with high accuracy is not guaranteed, because the data distribution is different on training and testing datasets, which generally leads to the dataset shift problem. Fortunately, in the problem under consideration, it is the values of the reconstructed tilt parameters that mainly determine the energy flux distribution, while the latent parameters lead only to...
its modification. This fact explains the good generalization ability of the model, which makes it possible to optimistically assess the possibility of its application in the case when the values of latent parameters are unknown and, as an ultimate case, to real experimental data.

In the context of applying the trained ML-model to real experimental data, it should be noted that in this work a simplified pulse model was used, when the entire spectrum of the laser pulse was divided only into three spectral components. In the case of adapting the model to a real physical experiment, when the wavefront has a complex structure, such a partition must be done into a larger number of spectral components. In this regard, it is important to note that in our numerical experiments it was observed a systematic effect of deterioration in the accuracy of reconstructing the parameters of spectral components with a small wavelength or relative amplitude. In addition, the effect of reducing the accuracy was observed in the case when the boundary wavelengths become close, i.e. one of the spectral ranges becomes very narrow. In all these cases, the influence of such spectral components on the energy flux decreases, which makes it difficult to reconstruct the corresponding tilt parameters. In addition, cross-influence of spectral components with close wavelengths on each other is possible, when the values of tilt parameters from one spectral component can be erroneously attributed by the ML-model to another. The probability of this kind of errors increases sharply with a decrease of the spectral component width, because the relative energy contribution to the cumulative energy flux will also decrease. This means that in the case of partitioning into a large number of spectral components, the characteristics of which may be close, the accuracy of simple ML-models may deteriorate significantly, which will require the use of more complex ML-models or approaches.

5. CONCLUSIONS

In this paper an ML-based approach for determining the parameters of the spectral-dependent tilt of a focused laser pulse based on the cumulative energy flux in the focal plane was proposed. A simplified formulation of the problem was given with the division of the spectrum of a laser pulse into three spectral components with an individual setting of the tilt parameters for each of them. It was shown that it is possible to obtain good accuracy of tilt parameters reconstructing using the convolutional neural network. In order to study the generalization ability of the proposed ML-model the methodology based on the study of the influence of the latent parameters choice on the accuracy of the ML-model was used. For this purpose separate datasets with different subsets of the latent parameters values were used during training and testing. Two experiments were carried out in which the boundary wavelengths and relative amplitudes acted as such latent parameters. In both experiments, it was shown that the network has good generalization ability and can reconstruct the tilt parameters with quite good accuracy even for the latent parameter values not included in the training set. Significant degradation of accuracy was observed only in cases where the relative energy contribution of the spectral component to the energy flux became small. These results indicate that the proposed approach is promising even in the presence of uncertainty in the choice of the latent parameter values that are typical for application of analytical models to a real experiment.

ACKNOWLEDGMENTS

The authors would like to thank Julien Ferri for useful discussions and other contributions. The authors acknowledge the use of computational resources provided by the Lobachevsky University and Joint Supercomputer Center of the Russian Academy of Sciences.

FUNDING

This research was funded by the Ministry of Science and Higher Education of the Russian Federation, agreement number 075-15-2020-808.
REFERENCES

1. V. Yanovsky, V. Chvykov, G. Kalinchenko, P. Rousseau, T. Planchon, T. Matsuoka, A. Maksimchuk, J. Nees, G. Cheriaux, G. Mourou, and K. Krushelnick, “Ultra-high intensity-300-TW laser at 0.1 Hz repetition rate,” Opt. Express 16, 2109–2114 (2008).

2. A. S. Pirozhkov, Y. Fukuda, M. Nishiuchi, H. Kiriyama, A. Sagisaka, K. Ogura, M. Mori, M. Kishimoto, H. Sakaki, N. P. Dover, K. Kondo, N. Nakannii, K. Huang, M. Kanasaki, K. Kondo, and M. Kando, “Approaching the diffraction-limited, bandwidth-limited Petawatt,” Opt. Express 25, 20486–20501 (2017).

3. G. Tiwari, E. Gaul, M. Martinez, G. Dyer, J. Gordon, M. Spinks, T. Toncian, B. Bowers, X. Jiao, R. Kupfer, L. Lisi, E. McCary, R. Roycroft, A. Yandow, G. D. Glenn, et al., “Beam distortion effects upon focusing an ultrashort petawatt laser pulse to greater than 10^{22} W/cm^{2},” Opt. Lett. 44, 2764–2767 (2019).

4. J. W. Yoon, Y. G. Kim, I. W. Choi, J. H. Sung, H. W. Lee, S. K. Lee, and C. H. Nam, “Realization of laser intensity over 10^{23} W/cm^{2},” Optica 8, 630–635 (2021).

5. A. Higginson, R. J. Gray, M. King, R. J. Dance, S. D. R. Williamson, N. M. H. Butler, R. Wilson, R. Capdessus, C. Armstrong, J. S. Green, S. J. Hawkes, P. Martin, W. Q. Wei, S. R. Mirfayzi, X. H. Yuan, et al., “Near-100 MeV protons via a laser-driven transparency-enhanced hybrid acceleration scheme,” Nat. Commun. 9 (1), 1–9 (2018).

6. A. J. Gonsalves, K. Nakamura, J. Daniels, C. Benedetti, C. Pieronek, T. C. H. de Raadt, S. Steinke, J. H. Bin, S. S. Bulanov, J. van Tilborg, C. G. R. Geddes, C. B. Schroeder, Cs. Tóth, E. Esarey, K. Swanson, et al., “Petawatt laser guiding and electron beam acceleration to 8 GeV in a laser-heated capillary discharge waveguide,” Phys. Rev. Lett. 122, 084801 (2019).

7. G. Sarri, D. J. Corvan, W. Schumaker, J. Cole, A. Di Piazza, H. Ahmed, C. Harvey, C. H. Keitel, K. Krushelnick, S. P. D. Mangles, Z. Najmudin, D. Symes, A. G. R. Thomas, M. Yeung, Z. Zhao, and M. Zepf, “Ultrahigh brilliance multi-MeV γ-ray beams from nonlinear relativistic Thomson scattering,” Phys. Rev. Lett. 113, 224801 (2014).

8. A. di Piazza, C. Müller, K. Z. Hatsagortsyan, and C. H. Keitel, “Extremely high-intensity laser interactions with fundamental quantum systems,” Rev. Mod. Phys. 84, 1177 (2012).

9. D. Strickland and G. Mourou, “Compression of amplified chirped optical pulses,” Opt. Commun. 55, 447–449 (1985).

10. K. T. Kim, C. Zhang, A. D. Shiner, S. E. Kirkwood, E. Frumker, A. Naumov, D. M. Villeneuve, and P. B. Corkum, “Manipulation of quantum paths for space-time characterization of attosecond pulses,” Nat. Phys. 9, 159–163 (2013).

11. C. Bourassin-Bouchet, M. M. Mang, F. Delmotte, P. Chavel, and S. de Rossi, “How to focus an attosecond pulse,” Opt. Express 21, 2506–2520 (2013).

12. G. Pariente, V. Gallet, A. Borot, O. Gobert, and F. Quéré, “Space-time characterization of ultra-intense femtosecond laser beams,” Nat. Photon. 10, 547–553 (2016).

13. A. Borot and F. Quéré, “Spatio-spectral metrology at focus of ultrashort lasers: A phase-retrieval approach,” Opt. Express 26, 26444–26461 (2018).

14. A. Jeantet, A. Borot, K. Nakamura, S. W. Jolly, A. J. Gonsalves, C. Toth, H. S. Mao, W. P. Leemans, and F. Quéré, “Spatio-temporal structure of a petawatt femtosecond laser beam,” J. Phys.: Photon. 1, 035001 (2019).

15. M. A. Krumbügel, C. L. Ladera, K. W. DeLong, D. N. Fittinghoff, J. N. Sweetser, and R. Trebino, “Direct ultrashort-pulse intensity and phase retrieval by frequency-resolved optical gating and a computational neural network,” Opt. Lett. 21, 143–145 (1996).

16. J. W. Nicholson, F. G. Omenetto, D. J. Funk, and A. J. Taylor, “Evolving FROGS: phase retrieval from frequency-resolved optical gating measurements by use of genetic algorithms,” Opt. Lett. 24, 490–492 (1999).

17. T. Zahavy, A. Dikopoltsev, D. Moss, G. I. Haham, O. Cohen, S. Mannor, and M. Segev, “Deep learning reconstruction of ultrashort pulses,” Optica 5, 666–673 (2018).

18. S. Kleinert, A. Tajalli, T. Nagy, and U. Morgner, “Rapid phase retrieval of ultrashort pulses from dispersion scan traces using deep neural networks,” Opt. Lett. 44, 979–982 (2019).

19. R. Ziv, A. Dikopoltsev, T. Zahavy, I. Rubinstein, P. Sidorenko, O. Cohen, and M. Segev, “Deep learning reconstruction of ultrashort pulses from 2D spatial intensity patterns recorded by an all-in-line system in a single-shot,” Opt. Express 28, 7528–7538 (2020).

20. W. Xiong, B. Redding, S. Gertler, Y. Brumberg, H. D. Tagare, and H. Cao, “Deep learning of ultrafast pulses with a multimode fiber,” APL Photon. 5, 096106 (2020).

21. I. Haber, R. Lee, H. Klein, and J. Boris, “Advances in electromagnetic simulation techniques,” in Proceedings of the 6th Conference on Numerical Simulation of Plasmas, Berkeley, CA (1973), pp. 46–48.

22. E. Panova, V. Volokitin, E. Efimenko, J. Ferri, T. Blackburn, M. Marklund, A. Muschet, A. Gonzalez, P. Fischer, L. Veisz, I. Meyerov, and A. Gonoskov, “Optimized computation of tight focusing of short pulses using mapping to periodic space,” Appl. Sci. 11, 956 (2021).
23. J. L. Vay, I. Haber, and B. B. Godfrey, “A domain decomposition method for pseudo-spectral electromagnetic simulations of plasmas,” J. Comput. Phys. 243, 260–268 (2013).
24. Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” Proc. IEEE 86, 2278–2324 (1998).
25. A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” Adv. Neural Inform. Process. Syst. 60 (6), 84–90 (2017).
26. A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, “Attention is all you need,” Adv. Neural Inform. Process. Syst. 30 (2017).
27. A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, “An image is worth 16x16 words: Transformers for image recognition at scale,” arXiv: 2010.11929 (2020).
28. M. Caron, H. Touvron, I. Misra, H. Jégou, J. Mairal, P. Bojanowski, and A. Joulin, “Emerging properties in self-supervised vision transformers,” in Proceedings of the IEEE/CVF International Conference on Computer Vision (2021), pp. 9650–9660.
29. H. Touvron, M. Cord, M. Douze, F. Massa, A. Sablayrolles, and H. Jégou, “Training data-efficient image transformers & distillation through attention,” in Proceedings of the International Conference on Machine Learning PMLR (2021), pp. 10347–10357.
30. D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” arXiv: 1412.6980 (2014).
31. I. J. Goodfellow, J. Shlens, and C. Szegedy, “Explaining and harnessing adversarial examples,” arXiv: 1412.6572 (2014).
32. X. Yuan, P. He, Q. Zhu, and X. Li, “Adversarial examples: Attacks and defenses for deep learning,” IEEE Trans. Neural Networks Learn. Syst. 30, 2805–2824 (2019).
33. M. Wang and W. Deng, “Deep visual domain adaptation: A survey,” Neurocomputing 312, 135–153 (2018).
34. G. Wilson and D. J. Cook, “A survey of unsupervised deep domain adaptation,” ACM Trans. Intell. Syst. Technol. 11 (5), 1–46 (2020).