



CNN-based neural network model for amplified laser pulse temporal shape prediction with dynamic requirement in high-power laser facility

LU ZOU,^{1,2} YUANCHAO GENG,² BINGGUO LIU,¹ FENGDONG CHEN,¹
WEI ZHOU,² ZHITAO PENG,² DONGXIA HU,² QIANG YUAN,²
GUODONG LIU,^{1,3} AND LANQIN LIU^{2,4}

¹*Harbin Institute of Technology, Harbin 150001, China*

²*Laser Fusion Research Center, China Academy of Engineering Physics, Mianyang 621900, China*

³*lgd@hit.edu.cn*

⁴*lylymao2013@163.com*

Abstract: The temporal shape of laser pulses is one of the essential performances in the inertial confinement fusion (ICF) facility. Due to the complexity and instability of the laser propagation system, it is hard to predict the pulse shapes precisely by pure analytic methods based on the physical model [Frantz-Nodvik (F-N) equation]. Here, we present a data-driven model based on a convolutional neural network (CNN) for precise prediction. The neural network model introduces sixteen parameters neglected in the F-N equation based models to expand the representation dimension. The sensitivity analysis of the experimental results confirms that these parameters have different degrees of influence on the temporal output shapes and cannot be ignored. The network characterizes the whole physical process with commonality and specificity features to improve the description ability. The prediction accuracy evaluated by a root mean square of the proposed model is 7.93%, which is better compared to three optimized physical models. This study explores a nonanalytic methodology of combining prior physical knowledge with data-driven models to map the complex physical process by numerical models, which has strong representation capability and great potential to model other measurable processes in physical science.

© 2022 Optica Publishing Group under the terms of the [Optica Open Access Publishing Agreement](#)

1. Introduction

Inertial confinement fusion (ICF) is a possible way to harness thermonuclear fusion energy by using a high-power laser driver to bombard the deuterium-tritium target [1–3]. ICF laser driving device is a huge, complex, and highly systematic optical engineering. It focuses multiple laser beams on the target simultaneously so that the high-energy and high-power laser interacts uniformly with the material to produce a high-temperature and high-pressure environment. The power balance between different laser beams is directly related to the success of the ICF experiment, which is ensured by precisely controlling the temporal shapes of the laser pulses. The main amplifier system of the laser device provides most of the power gain of the laser pulses but is also the primary source of waveform distortion [4]. Due to the expensive cost of the main amplifier operation, it is required to predict the output laser performance accurately for control by offline numerical simulation instead of experimental trials [5]. Besides, the temporal shapes are diverse for each trial shot, which causes difficulties in modeling, prediction, and control.

The existing prediction methods rely on physical models of the optical transmission and amplifying process. National Ignition Facility (NIF) [6–9] of Lawrence Livermore National Laboratory developed the Laser Performance Operations Model (LPOM) software for physical modeling [10,11], which describes the pulse amplification process by Frantz-Nodvik saturated gain model [12] based on precise measurements. The overall agreement with the measurements

is better than 5% for the time-resolved power with 1ns boxcar smoothing [13,14]. For OMEGA EP laser system, PSOPS, a semi-analytic model was developed [15]. It can forward and backward predict the performance of beamlines accurately in real-time with multi-pass amplification model using measured small signal gain and appropriate saturation fluence. In another Chinese high-power laser device [16], the SG99 was developed as a precise analysis and simulation model [17]. It derives the recurrence relation of the particles on upper and lower energy levels based on the velocity equation and reduces the gain recovery effect and the thermal effect of the degraded energy level in multi-pass amplification. However, it uses an iterative algorithm with a complex model, so the computation takes a longer time. Moreover, the Laser Performance Operation Simulation System (LPOSS) was developed as an efficient real-time computational model [18,19]. It gives out the trend of the waveform based on optimized Frantz-Nodvik saturated gain model. However, there is still a gap between the theoretically calculated pulse shape and measured data of the main amplifier due to the unstable device state, complex environmental factors, and varying uneven near-field distribution [20]. Therefore, the theoretical model is not accurate enough in practice but is challenging to correct.

Due to the limitations of the analytical approach, some empirically found relevant influences are difficult to incorporate into the physical model. The prediction model used now is a stable model for a single shot, but hard to describe time-varying trends for long-time operations. A data-driven approach allows an implicit model to express the physical process with a complex scenario. A large amount of data is used to train and find suitable parameters to characterize the model. There are several similar applications of data-driven models in ICF facilities, but most of them are for image processing [21]. The convolutional neural network (CNN) is a typical model which has a strong ability to extract features and accounts for the structure and correlation of data [22]. It is adopted to classify the microscopic damages for optical components repairing [23,24], detect the laser-induced optical defects online [25], and select capsule mandrels best suited for capsule production [26]. Sequence prediction problems, such as lifetime prediction, often use CNN to extract features and coupled RNN (LSTM or SRU) for calculation [27–29]. But, CNN can also perform calculations on time series data such as temporal convolutional neural (TCN) network [30–32]. Meanwhile, according to the convolution theorem, convolution can better characterize the variation in the frequency domain, which assists in accelerating the computation or processing of detailed data [33–37]. The models above make the computational process robust, repeatable, efficient, and operator-independent, but most studies are for 2-dimensional images or 3-dimensional image sequences. In this research, inspired by LPOM, LPOSS, and CNNs, we employ one-dimensional (1D) CNN to predict temporal pulse shape and introduce prior empirical knowledge to steer the direction of network modeling for solving the regression problem.

In this paper, we propose and demonstrate a method for predicting the temporal shape of the laser output from the main amplifier based on a novel convolutional neural network. With the physical process analysis, 16 additional influence parameters neglected by the conventional models are introduced to describe the variation of the main amplifier state, including environmental, temporal, and near-field factors. A multi-branch CNN with additive and multiplicative skip connections is designed to represent the amplification and propagation process with the characteristics of physics and data. The prediction accuracy of the proposed method on the experimental data in the ICF device studied in this paper is RMS 7.93%, better than other optimized F-N based physical models.

2. Methods

2.1. Main amplifier system and physic-based models

The main amplifier of the high-power laser facility discussed in this paper consists of a cavity amplifier, cavity spatial filter, boost amplifier, and transport spatial filter, which is shown in Fig. 1. The gain medium shown as dark blue oblique lines is Nd: glass disks pumped by light before

experiments. When the laser transmits through the main amplifier, it gains the energy stored in the upper laser level in the gain medium before the arrival of the input signal. The laser pulse passes through the oscillator-amplifier medium several times along the directed red lines at the top of the figure. The leading edge of the pulse encounters more reversed particles than the trailing edge, so the pulse changes temporal shape as it propagates through the amplifier and the peak moves forward.

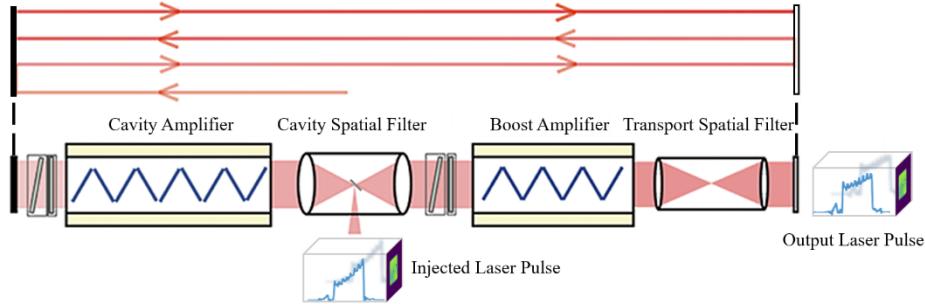


Fig. 1. Structure of the main amplifier.

In the gain process, the rate of change of the inverse population density n and the photon density ϕ can be expressed as

$$\frac{\partial n}{\partial t} = -n\sigma\phi c - \frac{n}{\tau_f} + W_p(n_g - n), \quad (1)$$

$$\frac{\partial \phi}{\partial t} + c \frac{\partial \phi}{\partial z} = c\sigma\phi n - \frac{\phi}{\tau_c} + S, \quad (2)$$

where σ is the stimulated emission cross section, c is the light speed, τ_f is the spontaneous emission time, W_p is the pump rate, n_g is the population density of the ground level in the four level gain medium, τ_c is the decay time for photons in the optical resonator, and S is the rate at which spontaneous emission is added to the laser emission [12].

With the simplifying assumptions that 1) the pump rate is much faster than the transit time of the laser pulse, 2) the spontaneous emission time is much longer than the pulse width of the laser, 3) the effect of fluorescence and pumping is ignored during the pulse duration, 4) a square pulse is injected into the amplifier, Frantz-Nodvik solved the rate equations. The Frantz-Nodvik (F-N) solution describing the amplification and distortion of the temporal shape of laser pulses is expressed as

$$P_o(t) = \frac{\exp\left(\frac{E_i(t)}{E_s}\right)}{\frac{1}{G_0} + \exp\left(\frac{E_i(t)}{E_s}\right) - 1} P_i(t), \quad (3)$$

where $P_i(t)$ and $P_o(t)$ are input and output temporal pulse shapes, $E_i(t)$ is the time dependent input energy $E_i(t) = \int_0^t P_i(t') dt'$, E_s is the saturation energy, and G_0 is the small-signal gain.

For slow change of the facility state caused by the long time operation, an optimized dynamic G_0 calculation method is adopted in LPOSS [38]

$$G_{n+1} = \frac{1}{p[m(n+1)]} \frac{\sum_{i=1}^n G_i p[m(i)] q_i}{\sum_{i=1}^n q_i}, \quad (4)$$

where G_i is the small signal gain of the i th shot, n is the amount of historical shots used to update the small signal gain, q_i is the weight of historical shots, p is the correction factor dependent on the number of shots on the same day m .

The FN equation is a relatively rough model. Since it does not consider the losses in the transmission process, an optimized F-N equation introducing losses is proposed in LPOSS [18]

$$P_o(t) = \frac{G_0 k(t)}{\left\{1 + A[k(t) - 1] \frac{G_0 - 1}{\ln G_0}\right\} \left\{1 + A[k(t) - 1] \frac{G_0 - 1}{\ln G_0} + G_0[k(t) - 1]\right\}} P_i(t), \quad (5)$$

where A is the loss factor and $k(t) = \exp\left(\frac{E_i(t)}{E_s}\right)$.

For the multi-pass amplification in the facility, an optimized iterative amplification model is proposed in PSOPS [15]

$$P_k(t) = \beta^2 G_k(t) P_{k-1}(t), \quad (6)$$

$$G_k(t) = \frac{1}{1 - \left\{1 - \left[\frac{1}{G_0} \exp\left(-\frac{\gamma(R)E_k(t)}{E_s}\right)\right]\right\}}, \quad (7)$$

$$E_k(t) = \int_0^t \beta P_{k-1}(t') dt', \quad (8)$$

where β is the per-disk-surface loss factor, $\gamma(R)$ is the empirical scaling factor on the pulse width [39].

All these methods have made some improvements on the original F-N equation and achieved good results in their respective devices. But they remain the fitting of key parameters of the improved model, characterizing the gain capability with a small number of averaged factors. These methods work well for analyzing the gain capacity of historical shots, but the effectiveness of predicting future experiments is compromised where there are changes in the state of the main amplifier.

2.2. Influencing factors

The laser pulses are spatiotemporally three-dimensional. The near-field distribution is non-uniform in space and inconsistent across shots, which does not exactly match the F-N equation assumptions. Although the temporal waveform is the integration of the laser pulse in space over time, the spatial and temporal distribution interact and are coupled. The near-field area influences the power density of the laser, which determines the gain saturation. Moreover, the inhomogeneous near-field distribution exacerbates the self-focusing phenomenon caused by the nonlinearity in laser transmission and affects amplification performance. Thus, we defined four features to characterize the spatial distribution of the laser pulses, which are near-field area (S), contrast (C), modulation (M), and inhomogeneity (A).

$$S = \frac{\sum_{x=1}^N \sum_{y=1}^M [u(x, y) - \bar{u}_{bg}] dx dy}{\bar{u}_{sig}}, \quad (9)$$

$$C = \frac{w_{sig}}{\sqrt{2}\bar{u}_{sig}}, \quad (10)$$

$$M = \frac{\max[u(x, y) - \bar{u}_{bg}]}{\bar{u}_{sig}}, \quad (11)$$

$$A = \frac{\sqrt{(\Delta u)^2}}{a_2}, \quad (12)$$

where $u(x, y)$ is the grayscale distribution of the near-field image, \bar{u}_{bg} and \bar{u}_{sig} is the average grayscale value of the background part and the signal part, w_{sig} is the half-height width of the signal peak in the grayscale histogram, Δu is the deviation between the measured distribution and the fitted Gaussian distribution of the signal part in the grayscale histogram, and a_2 is the amplitude obtained from the Gaussian fit.

Some other factors also affect the amplification. In terms of each trial shot, the pulse duration affects the gain recovery effect, which in turn determines the amplification capability [11]. Between continual trial shots, the interval time affects the recovery of components, which influences the amplifying performance of the latter experiment. As time-in-service grows, the aging and degradation of components cause decay of the gain performance. In addition, the ambient temperature and humidity in specific locations impact the performance of vital components, such as fibers in the front end, capacitors and cables in the energy storage, etc.

All of the parameters analyzed above may have different degrees of influence on the temporal shapes. We performed a correlation analysis of all the possible influencing factors and selected 16 factors with weak intercorrelation, shown in Table 1.

Table 1. Selected influence factors for prediction.

No.	Name	Influence on the temporal shapes
1	Pulse width	Gain recovery effect
2	Service time	Aging and degradation of components
3	Time interval	Recovery of components
4	Area	Power density and gain saturation
5	Contrast	
6	Modulation	Nonlinear effect and gain saturation
7	Uniformity	
8	Front end temperature 01	Components in front end, such as fibers, affect the spectrum
9	Front end temperature 02	
10	South energy storage temperature 01	Components in south energy storage, such as capacitors
11	North energy storage temperature 01	Components in north energy storage, such as capacitors
12	Front end humidity 01	Components in front end, such as fibers, affect the spectrum
13	Front end humidity 02	
14	South energy storage humidity 02	Components in south energy storage, such as capacitors
15	North energy storage humidity 01	Components in north energy storage, such as capacitors and cables
16	North energy storage humidity 05	

Although these parameters play roles in the transmission and amplification of the laser, it is hard to express them in an analytics-based physical model. Firstly, some factors are not standard physical quantities, such as device service time. It qualitatively represents the degradation state of components instead of quantitative representation with physical models. Secondly, several parameters are coupled to influence the gain process. For example, the ambient parameters and the state factors affect the performance of some vital components together. Moreover, even if several complex analytical models are built, they are hard to be solved. Thus, since the analytical models have inherent shortcomings and are difficult to be corrected, it is necessary to introduce other methods to achieve more accurate temporal pulse shape prediction.

2.3. Neural network model

A novel neural network is designed to consider the effect of the above parameters on the output waveform for a better description of the amplifying process. The two types of input data

(influence factors and temporal input shapes) are fed to three branches. Branch 1 and 3 extract data characteristics. Branch 2 converges these characteristics, represents the transmission and gain process and outputs the amplified temporal shapes. The specific structure is shown in Fig. 2.

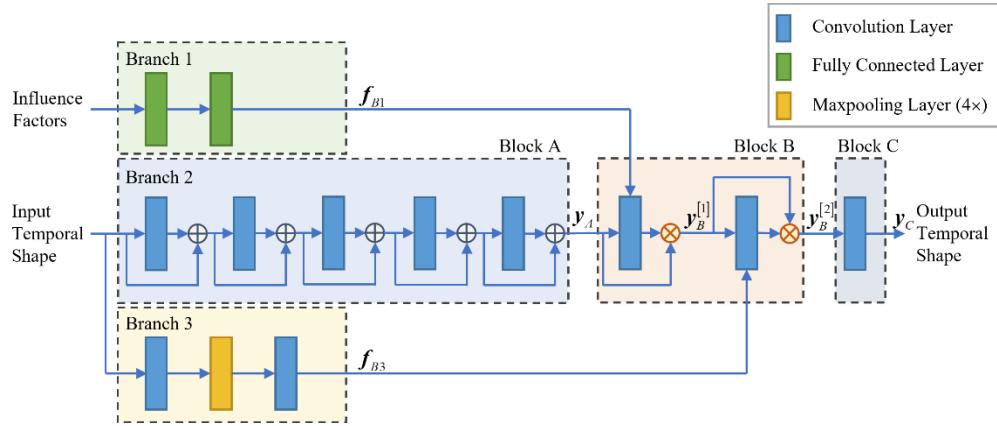


Fig. 2. The architecture of the proposed temporal shape prediction model based on a convolutional neural network.

Branch 2 is the main body part of the proposed neural network, subdivided into three blocks. Block A uses five 1-D convolutional layers with additive skip connections to characterize the commonality of laser transmission and the gain process.

With an additive skip connection, the output of the i th layer in Block A can be expressed as

$$y^{[i]} = a[f^{[i]} * y_{\text{pad}}^{[i-1]}] + y^{[i-1]}, \quad (13)$$

where $y^{[i]}$ and $y^{[i-1]}$ is the output of the i th and $(i-1)$ th layer with the same size (2000 elements), and $y_{\text{pad}}^{[i-1]}$ is the result of padding $y^{[i-1]}$. $f^{[i]}$ is the convolution filter of the i th layer, $*$ indicates convolution operation and the result can be represented as $z^{[i]}$. a is the activation function.

Circular padding is used to expand the output of the previous layer, which pads $(K-1)/2$ values before and after the vector respectively to ensure that the output size after convolution operation is the same as the input size. The convolution operation can be specifically expressed as

$$z^{[i]}(t) = \sum_{k=1}^K f^{[i]}(k) \times y_{\text{pad}}^{[i-1]}(t+k), \quad (14)$$

where K is the size of convolution filter $f^{[i]}$.

We design these convolution filters with different sizes K (121, 101, 51, 31, 21) to obtain frequency response curves with various resolutions according to the convolution theorem. The successive convolutions are equivalent to the product of the response curves in the frequency domain, and thus the response is refined.

$$\tilde{F}(\omega) = \prod_{i=1}^n F^{[i]}(\omega), \quad (15)$$

where $F^{[i]}(\omega)$ is the frequency domain distribution of the i th layer convolution filter and $\tilde{F}(\omega)$ is the integrated frequency response of each layer.

However, continual products cause the disappearance of tiny responses, resulting in the loss of information. Additive skip connections enrich the details after multi-layer convolution. In the

frequency domain, it extends the continuous products of high-order filters into the polynomials of each ordered product, which enhances the representation of the inherent properties of the gain process.

$$\tilde{F}(\omega) = \sum_{j=1}^n \left(\prod_{i=j}^n F^{[i]}(\omega) \right) + 1. \quad (16)$$

It is worth noting that $F^{[i]}(\omega)$ here does not depend only on the convolution filters but is also affected by the activation function. The activation function here is demanded not to affect the calculation of the frequency domain response, so a segmented linear activation function is required. The most common segmented linear rectification function (ReLU) is suppressed when the input x is less than 0. Although this property is very useful in classification problems, it leads to dead training for the regression problem of this study. Therefore, Leaky ReLU is chosen to ensure efficient training and no interference with the spectrum [40].

$$\text{LeakyReLU}(x) = \max(x, 0.01x). \quad (17)$$

Then, in Block B, we design convolutional layers with dynamic filters composed of features instead of constant filters composed of network parameters as classical CNN does. The features extracted from Branch 1 and 3 are set as convolutional filters to characterize the specificity of the physical process. The convolution operations here are the same as Eq. (14).

With influence factors input, which is scalar with low correlation, Branch 1 adopts two fully connected layers with 50 and 41 nodes and Tanh activation functions to map them into a 50-dimensional space and then obtains their 41-dimensional feature expressions. These features are fused with the output of Block A y_A more efficiently, as a filter of the convolution layer. Moreover, a multiplicative skip connection is designed to represent the gain, which is expressed as

$$y_B^{[1]} = \{\sigma[f_{B1} * y_{A\text{pad}}] + 1\} \cdot y_A, \quad (18)$$

where $\sigma(\cdot)$ is the sigmoid function, f_{B1} is the convolutional filter which is the output features of Branch 1, y_A is the input of this layer which is the output of Block A, $y_{A\text{pad}}$ is y_A with circular padding, and $y_B^{[1]}$ is the output of this layer. The multiplicative skip connection is still inherently additive but is more suitable for describing gain in the time domain.

In Branch 3 at the bottom, with time-series waveform data input, a tiny convolutional network is adopted to extract its main features as a 51-element vector. It contains a convolutional layer with filter size of 100 and stride of 4, a max-pooling layer with kernel size of 4 and stride of 4, a convolutional layer with filter size of 19 and stride of 2, and Tanh activation functions after each convolutional layer. The output of Branch 3 is then assigned to the filter of the second convolutional layer in Block B.

$$y_B^{[2]} = \{\sigma[f_{B3} * y_{B\text{pad}}^{[1]}] + 1\} \cdot y_B^{[1]}, \quad (19)$$

where f_{B3} is the convolutional filter which is the output features of Branch 3, $y_{B\text{pad}}^{[1]}$ is circular padded $y_B^{[1]}$, and $y_B^{[2]}$ is the output of Block B.

The convolution in the network is more like the cross-correlation function because no flipping is performed before multiplications. In extracting the features, the contribution of the background noise of the waveform is small, so the intercorrelation calculation here increases the signal-to-noise ratio. Due to the pulse shape distortion characteristics of the amplification process, the latter part of the input and output pulses are more similar and contribute more to the cross-correlation results. Thus, this layer enhances the response to the gain recovery effect.

Finally, in Block C, which is considered a head in this network, a convolutional layer with a filter size of 5 and stride of 2, is used to shape the results and final error compensation.

During network training, the loss function, which evaluates the difference between the temporal shape of the network prediction and the measured data, is a mean squared difference function, calculated as

$$L = \frac{1}{N} \sum_{i=1}^N \sum_{t=1}^M T_s \left| P_i(t) - \hat{P}_i(t) \right|^2, \quad (20)$$

where $P_i(t)$ and $\hat{P}_i(t)$ are the measured and predicted temporal shapes of the output laser pulse i of the main amplifier; N is the mini-batch size, and M is the amount of sample points of temporal shapes. Since the facility state changes over time, the newer data reflects the current state of the amplifier better. We add the service shot time T_s as coefficients in the loss function so that the more recent data has a greater impact on the model training, making the trained network model match the up-to-date state of the facility.

To accelerate the optimization and avoid falling into local optimum in training, the learning rate is set as a piecewise function changing with epochs, shown in Fig. 3. A gradually increasing learning rate warms up the beginning of the training and helps find a suitable optimization path. After 1500 epochs, the learning rate is kept twice the initial value to accelerate and skip the local minima trap. After 2000 epochs, the learning rate decreases dramatically after every 1000 epochs of training to ensure a fine search for suitable parameters near the optimal solution.

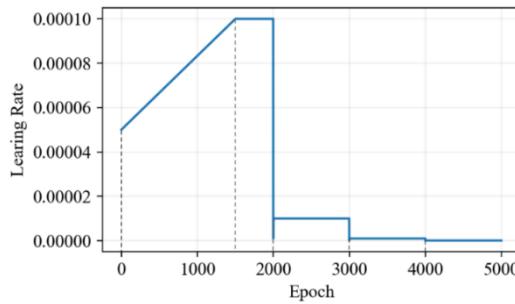


Fig. 3. Variation of learning rate with epochs.

3. Experiments and results

The dataset is obtained from one laser beamline over a four-year operation from 2018 to 2021 of the high-power laser facility discussed in this paper. The measured data are preprocessed to remove redundancy and stored as standard data in the database. The input and output temporal shapes are stretched so that the sampling intervals are 10ps and the edges are aligned. They are then linear normalized for better network training [41]. A typical preprocessed sample is shown as blue curves in Figs. 4(a) and (d). The measured output waveforms in operation are used as the ground truth to train and test the proposed model.

The available sample amount is 690, which is expanded to obtain a more robust trained network. Since the time-series data values have strong mapping relationships, some conventional expansion methods are not available, such as flipping, cropping, deformation, etc. In this paper, two data augmentation methods are adopted, which are signal shifting and random addition of background noise. Specifically, the signal and background parts of each sample are separated. For the signal, we randomly move it forward or backward. For the background, we randomly scale the noise amplitude by 0.5-1.5. Augmented data are shown in Figs. 4(b),(c),(e),(f). All data are arranged in the order of experiment time, and the previous 90% of the data are used as the training set and are expanded to 5 times the number of samples for training the neural network.

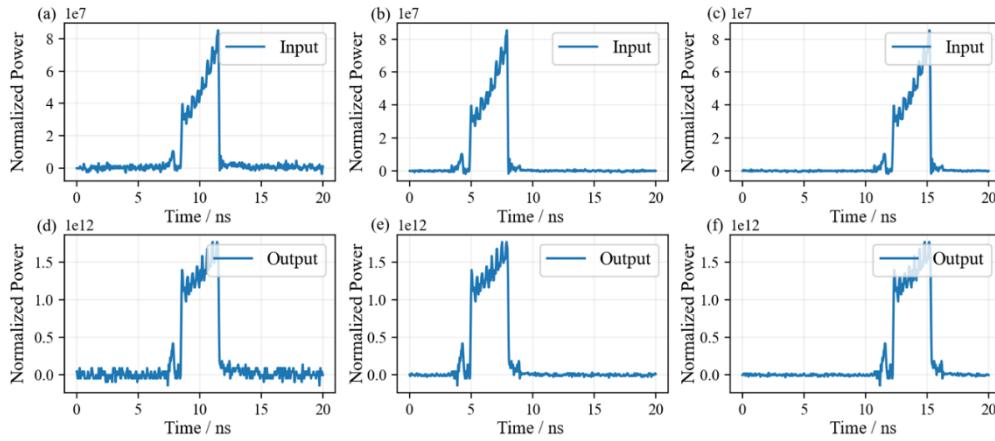


Fig. 4. Samples of training dataset. Preprocessed (a) input and (d) output waveforms measured in a trial. Augmented (b),(c) input and (e),(f) output waveforms in training dataset.

The latter 10% of data with time closer to the present, with 69 sets of samples, are used as the testing set to test the performance of the proposed method in this paper.

The network is implemented with PyTorch 1.1.0 based on Python 3.6.4. Network training and testing are performed on a PC with a Core i7-7800X CPU (3.50 GHz) and 48 GB of RAM using an NVIDIA GeForce GTX 1080Ti GPU. The parameters in the neural network are randomly initialized as tiny values before training. The training process using Adam optimizer takes approximately 10 hours for 3105 pairs of waveform vectors with a length of 2000 elements and a batch size of 20. After training, input wave shapes from the testing dataset are randomly input into the network, and the predicted power waveforms are rapidly produced as the output.

We use the three optimized models based on the F-N equation introduced in Section 2.1 to compare with the proposed model in this paper. The parameters in the three compared methods are determined by curve fitting with historical experiments. The key parameters in models are fitted using the data from the previous ten experiments of the current shot to be predicted. For the F-N with updated G_0 model, the key parameters E_s , G_0 , and parameters a and b in function p are fitted according to Eq. (3) and Eq. (4). For the F-N with loss model, E_s , G_0 , and A are fitted according to Eq. (5). For the multi-pass amplification model in PSOPS, E_s , G_0 , and β are fitted according to Eq. (6)–(8).

The prediction results of the proposed prediction model and the other three F-N based optimized models are compared to the measurement data in experiments as shown in Fig. 5 and Fig. 6. Each column is an experimental shot, and each row represents the prediction result of each method, while the last row shows the near-field beam distribution. The prediction results of all four models approximately match the measurement, but there are some differences in the details. The parameters of the three fitting-based models are sensitive and sometimes do not describe the relationship between input and output energy well. It makes some prediction results significantly underpowered. The power density is influenced by the near-field distribution, so that the gain saturation effect changes, but it is difficult to accurately describe the different saturation relationships for each shot by the fitted parameters. In addition, the gain recovery is not well handled in F-N based models, which makes the prediction results appear high on the leading edges or low on the trailing edges.

For the updated G_0 model, it was originally designed for the preamplifier in another Chinese high-power laser facility, which has fewer influencing factors and less complexity of the amplification process. Thus, the updated model for G_0 may need to introduce more nonlinear

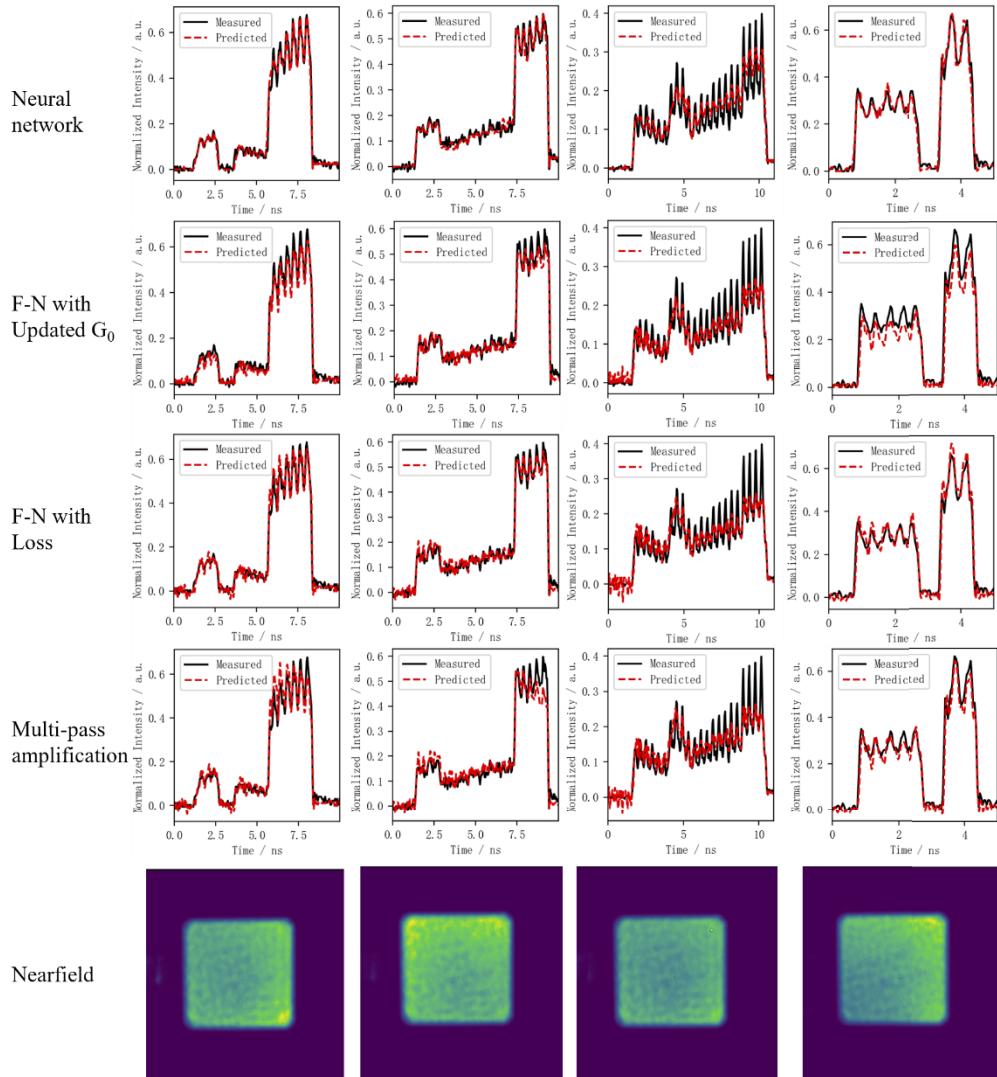


Fig. 5. The prediction results of various methods for different temporal shapes with corresponding near-field beam distributions.

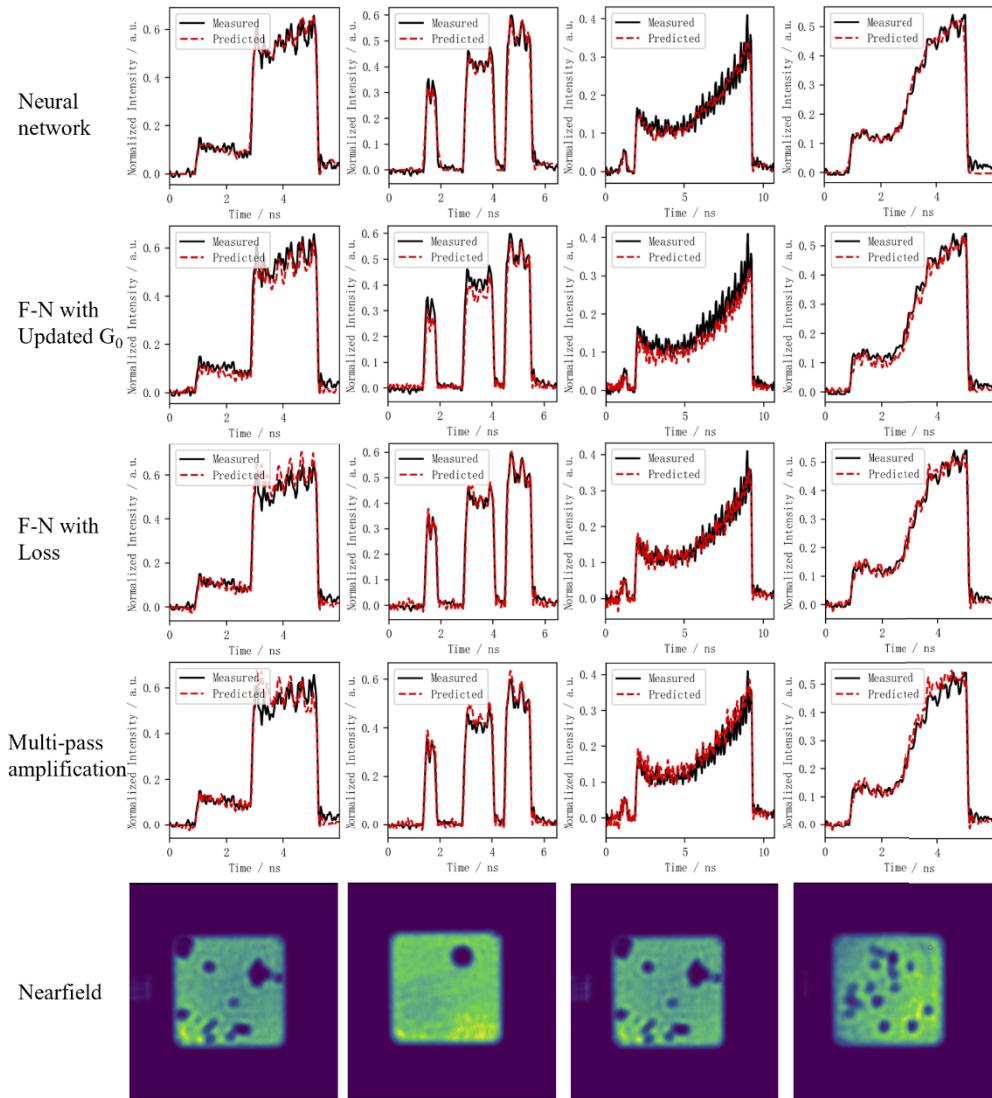


Fig. 6. The prediction results of various methods for different temporal shapes with masked near-field beam distributions.

functions to fit the current state of the device better. The prediction of the F-N with loss model performs better, but when the state of the main amplifier changes abruptly, it cannot describe it accurately with the previous ten shots, shown as the last row in Fig. 5 and the first row in Fig. 6. The multi-pass amplification model is a more accurate model considering more effects. However, due to the different operation strategies between OMEGA EP and the facility in this paper, the parameters that need to be measured in PSOPS model can only be obtained by fitting them with data of the historical experiments in this application. The stability of the devices also differs, specifically the gain capacity of the different shots in one day changes in this paper while the OMEGA EP does not, so the predictions of the PSOPS model are biased in this paper. In comparison, the proposed method has better prediction performance in various waveform and near-field distribution cases.

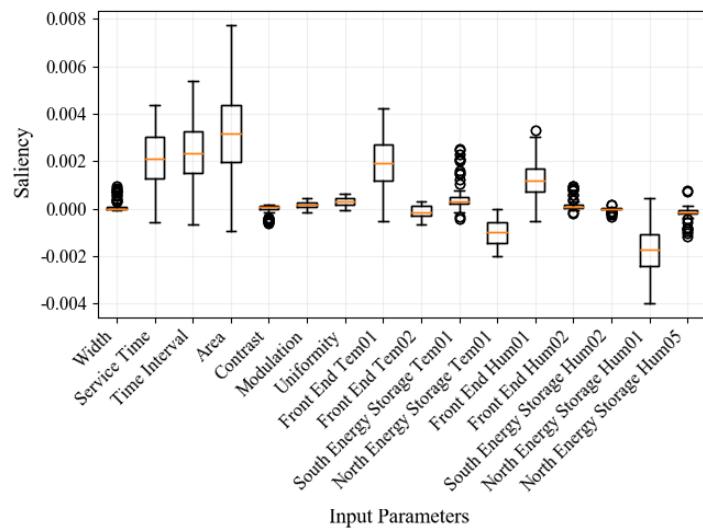


Fig. 7. Saliency results of each input parameters.

We quantitatively evaluate the average prediction performance and accuracy of the proposed method with testing data containing 69 shots. Widely adopted metrics are used, namely peak valley value (PV) and root mean square (RMS). PV and RMS with 1ns boxcar smoothing are also evaluated. The specific performances of the three F-N equation based optimized model and the proposed neural network method are shown in Table 2.

Table 2. Prediction performance of different approaches.

Methods	PV / %	PV _{1ns} / %	RMS / %	RMS _{1ns} / %
F-N with updated G0	14.88	8.54	15.79	11.99
F-N with Loss	7.52	3.20	10.98	4.74
Multi-pass amplification	8.82	5.20	12.88	7.39
Neural network	6.65	3.02	7.93	3.43

The results of the proposed model have less PV and RMS than the other models, which shows that the pulse shapes predicted by the proposed method are closer to the true values for the whole test set, and the details are more accurate. In general, the physical models based on the FN equation can describe the trend of pulse shapes, but they are biased in some details. It is more difficult to improve the physical model, difficult to resolve when more influences need to be introduced, and not easily adaptable to a more variable system. Neural networks, on the

other hand, are good at describing the relationships between different kinds of data and have an advantage in this case.

To investigate the reason for the greater accuracy of the proposed method, we analyzed the effect of the additionally introduced input parameters on the results. The saliency of input parameters is defined as the partial derivative of the L1 loss function to each input parameter, denoted as

$$\frac{\partial L}{\partial x_i} = \frac{\partial \left[\sum_{t=1}^M \|\hat{P}(t) - P_{FN}(t)\|_1 \right]}{\partial (x_i - \bar{x}_i)}, \quad (21)$$

where $\hat{P}(t)$ and $P_{FN}(t)$ is the prediction of output temporal shapes with the proposed neural network method and the F-N equation, x_i is the i th input parameters, and \bar{x}_i is the average of them. It characterizes how the variation of each input parameter affects the difference between the network output and the theoretical calculation, which is calculated for each sample and statistically shown in Fig. 7.

Some of the input parameters have a significant impact on the results. Among them, the service time characterizes the slow change of the amplifier state after a long-time operation. The time interval indicates the recovery time, which affects the readiness of the amplifier and is a sudden change for each shot. The nearfield distribution determines the power density of the pulses, which affects whether the operation region of the amplification reaches the zone of saturation. The saliency plot shows that the near-field area has the greatest degree of influence. The temperature and humidity of the front-end 01 monitoring point have obvious saliency, and the front-end affects the laser spectrum. But the specific contribution needs to be observed by increasing the monitoring equipment. North energy storage humidity also impacts the charging and discharging process. However, the detailed role still needs to be analyzed by more measurement equipment. The pulse width has a partially discrete distribution in the figure because the gain recovery effect varies only when the pulse width is large. The saliency analysis is consistent with the theoretical model and the prior knowledge accumulated during operations [42]. Moreover, the results also show that the influence of temperature and humidity parameters on the results cannot be neglected. Operating personnel should maintain them in operation and add auxiliary monitoring systems in the future.

4. Conclusions

In summary, a neural network model is presented to predict the temporal output pulse shape of the main amplifier in a high-power ICF facility. With analysis of the amplification characteristics, a purpose-designed convolutional neural network concerning 16 additional parameters is proposed to improve the representation of the physical process. Different branches extract features that serve as convolutional filters to characterize the states of the amplifier more directly. Additive skip connections prevent frequency domain information from being lost after multi-layer convolution, and multiplicative skip connections fit the gain. To enhance the robustness of the proposed model, we augment the training dataset by shifting the signal and adjusting the background noise. The performances of the conventional analytical methods in LPOSS and PSOPS and the proposed neural network model are investigated based on experimental data. The proposed method achieved better prediction results with an RMS of 7.93% (3.43% with a boxcar smoothing) and a PV of 6.65% (3.02% with a boxcar smoothing), compared to other physical based models. The saliency of each input parameter is further analyzed to explore the reasons for the accuracy improvement. The analysis results are consistent with prior knowledge and point out some influencing factors neglected before to be observed and controlled. The proposed nonanalytical prediction method based on a convolutional neural network is feasible in the main amplifier laser pulse prediction. It has potential in other complex measurable physical processes in the high-power field.

Disclosures. The authors declare no conflicts of interest.

Data availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. C. C. Kuranz, H. S. Park, C. M. Huntington, A. R. Miles, B. A. Remington, T. Plewa, M. R. Trantham, H. F. Robey, D. Shvarts, and A. Shimony, "How high energy fluxes may affect Rayleigh–Taylor instability growth in young supernova remnants," *Nat. Commun.* **9**(1), 1564 (2018).
2. B. Canaud, F. Garaude, P. Ballereau, J. Bourgade, C. Clique, D. Dureau, M. Houry, S. Jaouen, H. Jourdren, and N. Lecler, "High-gain direct-drive inertial confinement fusion for the Laser Mégajoule: recent progress," *Plasma Phys. Contr. F.* **49**(12B), B601–B610 (2007).
3. N. Lemos, L. Cardoso, J. Geada, G. Figueira, F. Albert, and J. M. Dias, "Guiding of laser pulses in plasma waveguides created by linearly-polarized femtosecond laser pulses," *Sci. Rep.* **8**(1), 3165 (2018).
4. T. A. Laurence, S. Ly, J. D. Bude, S. H. Baxamusa, and P. Ehrmann, "Energy transfer networks: Quasicontinuum photoluminescence linked to high densities of defects," *Phys. Rev. Mater.* **1**(6), 065201 (2017).
5. B. M. Van Wonterghem, S. J. Brereton, R. F. Burr, P. Folta, D. L. Hardy, N. N. Jize, T. R. Kohut, T. A. Land, and B. T. Merritt, "Operations on the National Ignition Facility," *Fusion Sci. Technol.* **69**(1), 452–469 (2016).
6. M. L. Spaeth, K. Manes, D. Kalantar, P. Miller, J. Heebner, E. Bliss, D. Spec, T. Parham, P. Whitman, and P. Wegner, "Description of the NIF laser," *Fusion Sci. Technol.* **69**(1), 25–145 (2016).
7. E. I. Moses, J. D. Lindl, M. L. Spaeth, R. W. Patterson, R. H. Sawicki, L. J. Atherton, P. A. Baisden, L. J. Lagin, D. W. Larson, and B. J. MacGowan, "Overview: Development of the National Ignition Facility and the Transition to a User Facility for the Ignition Campaign and High Energy Density Scientific Research," *Fusion Sci. Technol.* **69**(1), 1–24 (2016).
8. M. L. Spaeth, K. R. Manes, M. Bowers, P. Celliers, J.-M. D. Nicola, P. D. Nicola, S. Dixit, G. Erbert, J. Heebner, and D. Kalantar, "National ignition facility laser system performance," *Fusion Sci. Technol.* **69**(1), 366–394 (2016).
9. E. I. Moses, R. E. Bonanno, C. A. Haynam, R. L. Kauffman, B. J. MacGowan, R. W. Patterson, R. H. Sawicki, and B. M. Van Wonterghem, "The national ignition facility: path to ignition in the laboratory," *J. Phys. IV* **133**(2), 57 (2006).
10. G. Brunton, A. Casey, M. Christensen, R. Demaret, M. Fedorov, M. Flegel, P. Folta, T. Frazier, M. Hutton, and L. Kegelmeyer, "Control and Information Systems for the National Ignition Facility," *Fusion Sci. Technol.* **69**(1), 352–365 (2016).
11. M. Shaw, R. House, W. Williams, C. Haynam, R. White, C. Orth, and R. Sacks, "Laser Performance Operations Model (LPOM): A computational system that automates the setup and performance analysis of the National Ignition Facility," *J. Phys.: Conf. Ser.* **112**(3), 032022 (2008).
12. W. Koechner, *Solid-State Laser Engineering* (Springer, 2006).
13. M. S. Hutton, S. Azevedo, R. Beeler, R. Bettenhausen, E. Bond, A. Casey, J. Liebman, A. Marsh, T. Pannell, and A. Warrick, "Experiment archive, analysis, and visualization at the National Ignition Facility," *Fusion Eng. Des.* **87**(12), 2087–2091 (2012).
14. J. M. Di Nicola, T. Bond, M. Bowers, L. Chang, M. Hermann, R. House, T. Lewis, K. Manes, G. Mennerat, B. MacGowan, R. Negres, B. Olejniczak, C. Orth, T. Parham, S. Rana, B. Raymond, M. Rever, S. Schrauth, M. Shaw, M. Spaeth, B. Van Wonterghem, W. Williams, C. Widmayer, S. Yang, P. Whitman, and P. Wegner, "The national ignition facility: laser performance status and performance quad results at elevated energy," *Nucl. Fusion* **59**(3), 032004 (2019).
15. M. J. Guardalben, M. Barczys, B. E. Kruschwitz, M. Spilatro, L. J. Waxer, and E. M. Hill, "Laser-system model for enhanced operational performance and flexibility on OMEGA EP," *High Power Laser Sci. Eng.* **8**, e8 (2020).
16. W. Zheng, X. Wei, Q. Zhu, F. Jing, D. Hu, J. Su, K. Zheng, X. Yuan, H. Zhou, W. Dai, W. Zhou, F. Wang, D. Xu, X. Xie, B. Feng, Z. Peng, L. Guo, Y. Chen, X. Zhang, L. Liu, D. Lin, Z. Dang, Y. Xiang, and X. Deng, "Laser performance of the SG-III laser facility," *High Power Laser Sci. Eng.* **4**, e21 (2016).
17. J. Su, W. Wang, J. Feng, Z. Peng, and X. Zhang, "The code SG99 for high-power laser propagation and its applications," *High-Power Lasers and Applications III* **5627**, 527–531 (2005).
18. W. Zheng, X. Wei, Q. Zhu, F. Jing, D. Hu, X. Yuan, W. Dai, W. Zhou, F. Wang, and D. Xu, "Laser performance upgrade for precise ICF experiment in SG-III laser facility," *Matter Radiat. Extremes* **2**(5), 243–255 (2017).
19. Z. Zhaoyu, Z. Junpu, L. Sen, L. Yue, Y. Ke, T. Xiaocheng, H. Xiaoxia, C. Bo, and Z. Wanguo, "Precise laser pulse shaping technology and application with high energy stability," *High Power Laser and Particle Beams* **34**(3), 031011 (2022).
20. E. S. Bliss, D. R. Speck, and W. W. Simmons, "Direct interferometric measurements of the nonlinear refractive index coefficient n_2 in laser materials," *Appl. Phys. Lett.* **25**(12), 728–730 (1974).
21. B. K. Spears, J. Brase, P.-T. Bremer, B. Chen, J. Field, J. Gaffney, M. Kruse, S. Langer, K. Lewis, R. Nora, J. L. Peterson, J. J. Thiagarajan, B. V. Essen, and K. Humbird, "Deep learning: A guide for practitioners in the physical sciences," *Phys. Plasmas* **25**(8), 080901 (2018).
22. G. E. Hinton and R. R. Salakhutdinov, "Reducing the Dimensionality of Data with Neural Networks," *Science* **313**(5786), 504–507 (2006).

23. C. Amorin, L. M. Kegelmeyer, and W. P. Kegelmeyer, "A hybrid deep learning architecture for classification of microscopic damage on National Ignition Facility laser optics," *Stat. Anal. Data Min: The ASA Data Sci. J.* **12**(6), 505–513 (2019).

24. T. N. Mundhenk, L. M. Kegelmeyer, and S. K. Trummer, "Deep learning for evaluating difficult-to-detect incomplete repairs of high fluence laser optics at the National Ignition Facility," in *Thirteenth International Conference on Quality Control by Artificial Vision 2017*, vol.10338, (2017), 109–116.

25. X. Chu, H. Zhang, Z. Tian, Q. Zhang, F. W. ng, J. Chen, and Y. Geng, "Detection of laser-induced optical defects based on image segmentation," *High Power Laser Sci. Eng.* **7**(4), e66 (2019).

26. K.-J. Boehm, Y. Ayzman, R. Blake, A. Garcia, K. Sequoia, S. Sundram, and W. Sweet, "Machine learning algorithms for automated NIF capsule mandrel selection," *Fusion Sci. Technol.* **76**(6), 749–757 (2020).

27. K. D. Humbird, J. L. Peterson, and R. G. McClarren, "Predicting the time-evolution of multi-physics systems with sequence-to-sequence models," arXiv:1811.05852 (2018).

28. B. Dya, B. Bla, B. Hla, B. Jya, and C. Ljb, "Remaining Useful Life Prediction of Roller Bearings based on Improved 1D-CNN and Simple Recurrent Unit," *Measurement* **175**, 109166 (2021).

29. J. Li, X. Li, and D. He, "A Directed Acyclic Graph Network Combined with CNN and LSTM for Remaining Useful Life Prediction," *IEEE Access* **7**, 75464–75475 (2019).

30. S. Bai, J. Z. Kolter, and V. Koltun, "An empirical evaluation of generic convolutional and recurrent networks for sequence modeling," arXiv:1803.01271 (2018).

31. P. Hewage, A. Behera, M. Trovati, E. Pereira, M. Ghahremani, F. Palmieri, and Y. Liu, "Temporal convolutional neural (TCN) network for an effective weather forecasting using time-series data from the local weather station," *Soft Comput.* **24**(21), 16453–16482 (2020).

32. J. Fan, K. Zhang, Y. Huang, Y. Zhu, and B. Chen, "Parallel spatio-temporal attention-based TCN for multivariate time series prediction," *Neural Comput. Appl.* 1–10 (2021).

33. K. Xu, M. Qin, F. Sun, Y. Wang, Y.-K. Chen, and F. Ren, "Learning in the frequency domain," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2020), 1740–1749.

34. Y. Wang, C. Xu, C. Xu, and D. Tao, "Packing convolutional neural networks in the frequency domain," *IEEE T. Pattern Anal.* **41**(10), 2495–2510 (2019).

35. H. Wang, X. Wu, Z. Huang, and E. P. Xing, "High-frequency component helps explain the generalization of convolutional neural networks," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2020), 8684–8694.

36. G. Chen, P. Peng, L. Ma, J. Li, L. Du, and Y. Tian, "Amplitude-phase recombination: Rethinking robustness of convolutional neural networks in frequency domain," in *Proceedings of the IEEE/CVF International Conference on Computer Vision* (2021), 458–467.

37. H. Chen, J. Jia, W. Niu, Y. Zhao, and N. Chi, "Hybrid frequency domain aided temporal convolutional neural network with low network complexity utilized in UVLC system," *Opt. Express* **29**(3), 3296–3308 (2021).

38. L. Ping, W. Wei, J. Sai, H. Wanqing, W. Wenyi, S. Jingqin, and Z. Runchang, "The shaped pulses control and operation on the SG-III prototype facility," *Laser Phys.* **28**(4), 045004 (2018).

39. C. Bibeau, J. B. Trenholme, and S. A. Payne, "Pulse length and terminal-level lifetime dependence of energy extraction for neodymium-doped phosphate amplifier glass," *IEEE J. Quantum Electron.* **32**(8), 1487–1496 (1996).

40. A. K. Dubey and V. Jain, "Comparative Study of Convolution Neural Network's Relu and Leaky-Relu Activation Functions," in *Applications of Computing, Automation and Wireless Systems in Electrical Engineering* (Springer, 2019), 873–880.

41. J. Sola and J. Sevilla, "Importance of input data normalization for the application of neural networks to complex industrial problems," *IEEE Trans. Nucl. Sci.* **44**(3), 1464–1468 (1997).

42. Z. Lu, G. Yuanchao, L. Guodong, L. Lanqin, C. Fengdong, L. Bingguo, H. Dongxia, Z. Wei, and P. Zhitao, "Laser energy prediction with ensemble neural networks for high-power laser facility," *Opt. Express* **30**(3), 4046–4057 (2022).