Convolutional neural network for transient grating frequency-resolved optical gating trace retrieval and its algorithm optimization

Siyuan Xu(许思源), Xiaoxian Zhu(朱孝先), Ji Wang(王佶), Yuanfeng Li(李远锋), Yitan Gao(高亦谈), Kun Zhao(赵昆)†, Jiangfeng Zhu(朱江峰)‡, Dacheng Zhang(张大成), Yunlin Chen(陈云琳), and Zhiyi Wei(魏志义)

What follows is a list of articles you may be interested in

Optoelectronic memristor for neuromorphic computing

Wuhong Xue(薛武红), Wenjuan Ci(次文娟), Xiao-Hong Xu(许小红), Gang Liu(刘钢)

Chin. Phys. B, 2020, 29 (**4**): 048401.   DOI: 10.1088/1674-1056/ab75da

Effects of Shannon entropy and electric field on polaron in RbCl triangular quantum dot

M Tiotsop, A J Fotue, S C Kenfack, N Issofa, H Fotsin, L C Fai

Chin. Phys. B, 2016, 25 (**4**): 048401.   DOI: 10.1088/1674-1056/25/4/048401

Analysis and modeling of resistive switching mechanism oriented to fault tolerance of resistive memory based on memristor

Huang Da, Wu Jun-Jie, Tang Yu-Hua

Chin. Phys. B, 2014, 23 (**3**): 038404.   DOI: 10.1088/1674-1056/23/3/038404

Comparative study of electrical properties of Au/n-Si (MS) and Au/Si$_3$N$_4$/n-Si (MIS) Schottky diodes

Adem Tataroğlu

Chin. Phys. B, 2013, 22 (**6**): 068402.   DOI: 10.1088/1674-1056/22/6/068402

Analysis and modeling of resistive switching mechanisms oriented to resistive random-access memory

Huang Da, Wu Jun-Jie, Tang Yu-Hua

Chin. Phys. B, 2013, 22 (**3**): 038401.   DOI: 10.1088/1674-1056/22/3/038401

Convolutional neural network for transient grating frequency-resolved optical gating trace retrieval and its algorithm optimization[*]

Siyuan Xu(许思源)[1,2], Xiaoxian Zhu(朱孝先)[2,3], Ji Wang(王佶)[2,4], Yuanfeng Li(李远锋)[1,2], Yitan Gao(高亦谈)[2,3],
Kun Zhao(赵昆)[2,5,†], Jiangfeng Zhu(朱江峰)[1,‡], Dacheng Zhang(张大成)[1],
Yunlin Chen(陈云琳)[4], and Zhiyi Wei(魏志义)[2,3,5]

[1] *School of Physics and Optoelectronic Engineering, Xidian University, Xi'an 710071 China*
[2] *Beijing National Laboratory for Condensed Matter Physics, Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China*
[3] *University of Chinese Academy of Sciences, Beijing 100049, China*
[4] *Institute of Applied Micro-Nano Materials, School of Science, Beijing Jiaotong University, Beijing 100044, China*
[5] *Songshan Lake Material Laboratory, Dongguan 523808, China*

A convolutional neural network is employed to retrieve the time-domain envelop and phase of few-cycle femtosecond pulses from transient-grating frequency-resolved optical gating (TG-FROG) traces. We use theoretically generated TG-FROG traces to complete supervised trainings of the convolutional neural networks, then use similarly generated traces not included in the training dataset to test how well the networks are trained. Accurate retrieval of such traces by the neural network is realized. In our case, we find that networks with exponential linear unit (ELU) activation function perform better than those with leaky rectified linear unit (LRELU) and scaled exponential linear unit (SELU). Finally, the issues that need to be addressed for the retrieval of experimental data by this method are discussed.

Keywords: transient-grating frequency-resolved optical gating, convolutional neural network, activation function, phase retrieval algorithm

PACS: 84.37.+q, 84.35.+i, 07.05.Pj, 42.65.Hw   **DOI:** 10.1088/1674-1056/abf0ff

1. Introduction

Generation of femtosecond (10^{-15} s) pulses allows an ultrafast revolution in science and technology.[1] One of the most significant applications of femtosecond laser pulses, especially few-cycle femtosecond pulses, is to generate isolated attosecond pulses (1^{-18} s) via high-order harmonic generation in gas targets.[2,3] In experiments, it is necessary and important to determine the pulse duration as well as detailed temporal profile of the electric field of a few-cycle pulse. For nanosecond pulses, the pulse width can be measured directly using a photodiode, while it cannot for picosecond and shorter pulses. Autocorrelation[4] can measure the pulse width of pico- and femtosecond pulses, but does not give details of their temporal profiles. One of the popular techniques employed in ultrafast research labs to characterize femtosecond pulses nowadays is the frequency-resolved optical gating (FROG) which is applicable for pulses over a wide range of wavelengths[5] and close to one optical cycle, which is 2.67 fs for the center wavelength at 800 nm.

FROG measurement requires splitting the pulse to be characterized into two variably delayed replicas. The two pulse replicas are then crossed in an instantaneously responding nonlinear-optical medium. The nonlinear signal produced by the two pulses is usually measured by a spectrometer, and as the delay between the two pulses changes, a series of spectra of the nonlinear signal are recorded. Such spectra compose a two-dimensional spectrogram with delay (time) as one axis and wavelength or frequency as the other. This spectrogram is commonly named FROG trace, which contains all the information necessary to describe the measured pulse. However, the pulse envelope, electric field, or pulse width cannot be calculated or derived from the FROG trace directly. It requires a non-trivial iterative algorithm to retrieve the spectrum and spectral phase, or equivalently, the electric field temporal profile and phase, of the pulse from the recorded trace.[5,6] Using the traditional FROG algorithm to retrieve the temporal profile and phase of the pulse requires considerable computing power and takes a long time. To this end, an alternative method to the phase retrieval from transient-grating FROG (TG-FROG) trace is demonstrated using a convolutional neural network (CNN). Similar approach has been shown to be successful for SHG-FROG trace retrieval[7] and attosecond streaking trace retrieval.[8]

[†]Corresponding author. E-mail: zhaokun@iphy.ac.cn
[‡]Corresponding author. E-mail: jfzhu@xidian.edu.cn

A FROG spectrogram or trace is a two-dimensional (2D) array and each element contains the FROG signal strength at the corresponding frequency and delay time. This array may be considered as an image (Fig. 1), in which each pixel corresponds to an element in the 2D array. The inversion of such traces to acquire the information of the femtosecond pulses may be considered as image recognition because different pulses produce FROG traces with different shapes, structures, sizes, tilts or slopes, and so on (Fig. 1). For image recognition tasks, neural networks[9,10] are able to achieve better learning effects by keeping important parameters and removing a large number of unimportant parameters. We train a convolutional neural network to learn the inverse mapping of the TG-FROG measurement.[11] For phase retrieval of femtosecond pulses, the neural network is used to find the mapping function between the FROG trace and the pulse electric field, which is achieved by training the network with FROG traces of pulses with known electric fields and phases. FROG traces serve as the inputs to the network, and the time domain electric fields and phases of the pulses are the output. We choose CNN because it is one of the most popular choices in image recognition tasks. Such a neural network has a multilayer structure. Typical CNN structure includes the input layer, convolution layer, pooling layer, and fully connected layer. Each layer contains a number of neurons that are connected to the previous layer. A neuron generates an output signal when the input exceeds a certain threshold. During a training process, the CNN constantly updates the weight of each neuron, and try to make the features represented by each neuron match the desired output.[9,10] In CNN, data features are extracted by convolution, and an activation function introduces nonlinear factors. The pooling layer compresses the input feature map and extracts deeper features. Finally, a fully connected layer is used to connect the extracted features.

2. Transient-grating frequency-resolved optical gating

The signal recorded in a TG-FROG trace is produced via four-wave-mixing (FWM) which is a third-order nonlinear process.[11,12] The input beam is divided into three identical beams. The phase matching is achieved by arranging the three input beams and one signal beam (the output) so that their points of intersection with a plane parallel to the focusing lens form the vertices of a rectangle. The three input beams are focused at one point in a piece of fused silica as the nonlinear medium by a concave mirror. Two of the beams are overlapped in space, and the pulses are overlapped in time as well. The TG signal strength versus the delay of the third input beam yields the (third-order) intensity autocorrelation. By recording the spectra of the TG signal as a function of the delay, a

TG-FROG trace is obtained which contains the necessary information for complete and unambiguous retrieval of the input pulse. The expression for the TG-FROG trace is

$$I_{\text{TGFROG}}(\omega, \tau) = \left| \int_{-\infty}^{\infty} E(\text{t})^2 E^*(t - \tau) \exp(-i\omega t) \, dt \right|^2,$$

where I_{TGFROG} is the TG-FROG trace, $E(t)$ is the electric field of the input pulse, τ is the delay between the third pulse and the other two.[6,11] Such an expression represents the well-known phase retrieval problem, which is non-trivial and cannot be solved directly. Iterative trial-and-error algorithms have been employed to retrieve the electric field and phase of the pulse from the trace, but the calculations are usually time-consuming yet not as accurate as one might expect. On the other hand, through a trained neural network, the calculation time is greatly shortened and the retrieval is almost instant.

A Fourier-transform-limited pulse appears as a symmetric and narrow peak in the FROG trace. When dispersion is added to the spectral phase of the pulse so that the pulse becomes chirped, the peak in the trace becomes widened and tilted, as shown in Fig. 1. When the second-order dispersion is added, the pulse is linearly chirped and the peak is tilted. Higher order dispersion makes the pulse nonlinearly chirped. The third-order dispersion produces small ripples on one side of the main pulse and changes the main peak envelope as well. In order to identify the pulse correctly from the trace, the neural network needs to learn the relationship between the phase of the pulse and the structure of the peak in the trace. This learning process may be understood as an analysis of the image characteristics of the trace.

3. Convolutional neural network

To train the neural network, we establish a forward program to produce a sufficient number of traces from pulses with known parameters. The pulses with 750 nm center wavelength have pulse widths ranging from 4 fs to 9 fs. Actual pulses tend to have complex dispersion but higher order dispersion is usually small so we only add dispersions of the 2nd to 5th order for simplicity without losing the generality. Figure 1 shows the traces, electric fields, and temporal phases of pulses with different order dispersions, whose transform-limited pulse widths are all 6 fs. It is shown in Fig. 1 that the 2nd and 4th order dispersions cause a symmetrical change in the pulse envelop in the time domain, and the traces are tilted to one side, while the 3rd and 5th order dispersions introduce small ripples to one side of the main pulse envelop.

The CNN used for FROG phase retrieval consists of four convolution layers with filter sizes of 2×2. The number of filters for each layer of convolution is 128, 64, 32, and 32, respectively. The stride parameter of each filter is set to 1 so the convolutional block outputs a series of feature maps with

dimensions identical to the input one. In each layer of convolution calculation, an exponential linear unit (ELU) is used to introduce nonlinear factors into the neural network. In order to compress the number of data and parameters, reduce overfitting, and improve the fault tolerance of the neural network, max pooling of 2×2 is used in each layer of convolution computation after the ELU. The neural network has two fully connected layers with sizes of 512 and 1024 respectively after four convolutional layers in order to establish contact with the output representing the intensity and phase of the femtosec-

ond pulse corresponding to the FROG trace. The inputs are FROG traces with size of 512×512. The size of the output array is 1024×1 $(512 + 512)$ which contains the spectral intensity with length of 512 and the phase with length of 512. Our CNN structure is shown in Fig. 2. The network is trained using the Tensorflow Python library (version 1.13) running on a graphics card for increased speed. The supervised learning converges to a solution after it optimizes weights for all the 3600 sample traces more than 60000 times, running in a NVIDIA RTX 2060S GPU card.
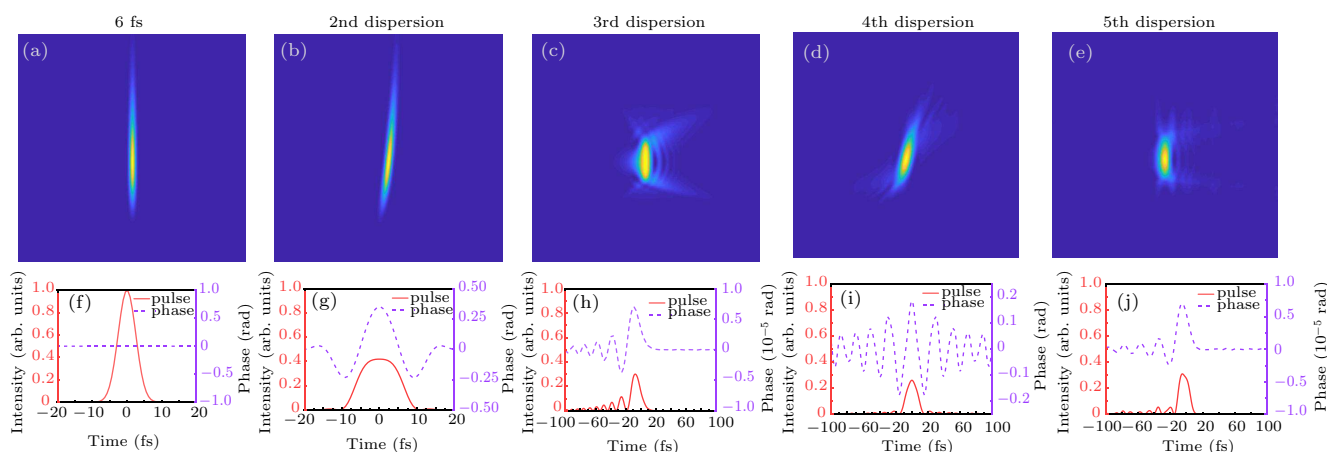

Fig. 1. TG-FROG traces of pulses at (a) Fourier transform limit, with (b) 2nd order dispersion, (c) 3rd order dispersion, (d) 4th order dispersion, and (e) 5th order dispersion, respectively. (f)–(j) The pulse temporal profiles and phases corresponding to (a)–(e).
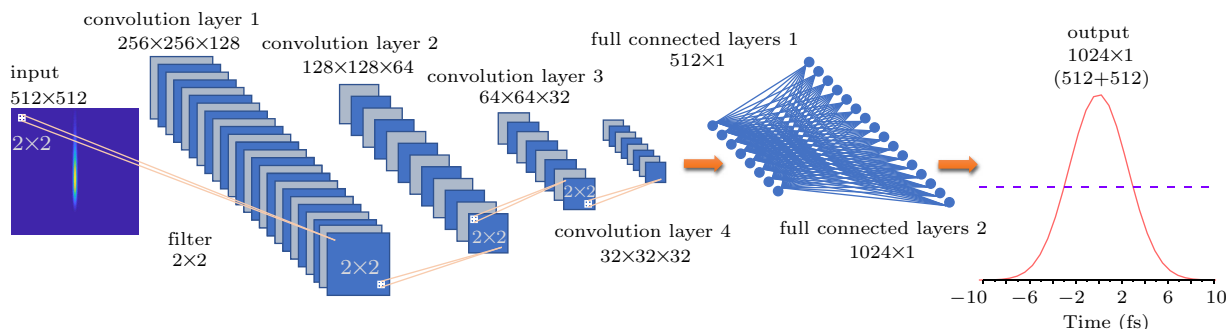

Fig. 2. Schematic of the convolutional neural network.

4. Results

After training the network with a set of 3600 samples, we apply the trained model to FROG-traces that have not been added to the training set, as shown in Fig. 3. The output time domain electric field is compared with the electric field used to produce this particular trace by the forward program. The main peak and the corresponding phase of the pulse restore the real electric field very well, but the small pulses around the main peak are not reproduced well, with a certain amount of noise. Similarly, the retrieved phase near the main peak accurately shows the oscillation of the phase, but the value is different from the real phase. However, the oscillations are not well restored away from the main peak. This is because the model pays more attention to the learning of high-intensity

values during the training, and treats low-intensity information as noise, which makes the learning of low-intensity information insufficient. This phenomenon does not affect our prediction of the pulse, and we will try to solve this problem in follow-up works. The loss and learning rates are shown in Fig. 4. The loss represents the error between the learning result and the actual value in a supervised learning. The average loss of the training samples is 2×10^{-3} after training in our neural network, which is fairly good. The learning rate determines whether the objective function converges to a local minimum and when it converges to the minimum, the learning rate is stable at 10^{-5} within 50000 steps. We use root mean squared error (RMSE) to verify the error between the predicted result and the actual pulse and phase, and show them in each image.
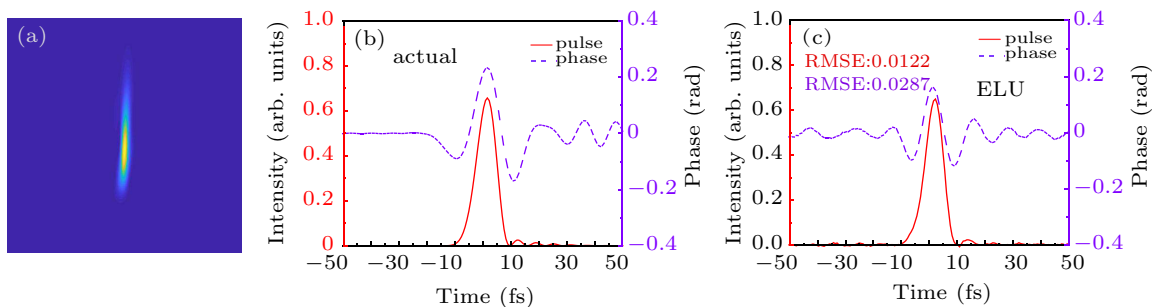
Fig. 3. (a) Input theoretically calculated TG-FROG trace. (b) Actual pulse envelop and phase corresponding to the TG-FROG trace. (c) Pulse envelop and phase retrieved by the neural network with ELU activation function.
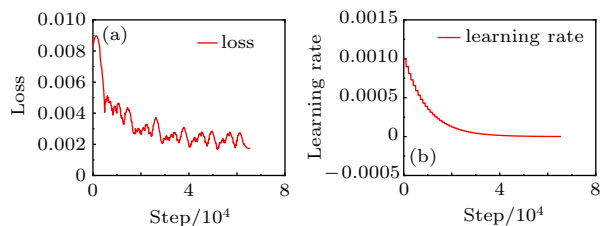

Fig. 4. (a) The loss and (b) the learning rates, stabilize at step 50000.

In our neural network we use ELU as the activation function to introduce nonlinearity, whose function form and curve are shown in Fig. 5. ELU is an activation function based on rectified linear unit (RELU) that has an extra constant α that defines function smoothness when inputs are negative. In our neural network training, we set $\alpha = 1$. Advantages of ELU include tending to converge faster than RELU and better generalization performance than RELU. ELU is fully continuous and differentiable, does not have a vanishing gradients problem, and does not have an exploding gradients problem, and dead RELU problem. ELU is slower to compute, but it compensates this by faster convergence during training. The performance of the ELU function in our neural network training meets our expectations.
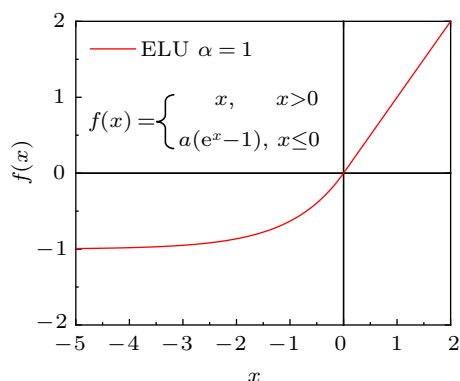

Fig. 5. ELU activation function expression and curve.

We compare the prediction of pulse envelop and phase by neural networks trained under several different activation functions, including leaky rectified linear unit (LRELU) and scaled exponential linear units (SELU). Their training results are shown in Fig. 6. Each network has been trained enough

times to make the learning rate saturated. The result of using the SELU function to predict (Fig. 6(a)) shows that the width of the main pulse is slightly larger, but the slightly sharper trailing edge as shown in the actual pulse (Fig. 3(b)) is not observed in the predicted result, and the oscillation of the phase is not well learned. The result of LRELU (Fig. 6(b)) shows that the main peak is split, which is wrong, and the phase oscillation is reproduced better than SELU, but the value is smaller than the actual pulse. In the application to traces not in the training data set, the neural networks with these two activation functions do not achieve results as good as ELU. At the same time, the RMSE results also show that using the ELU activation function performs best in the retrieval of ultrashort laser pulses.
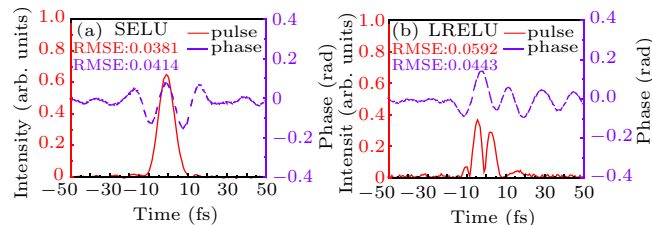

Fig. 6. Pulse envelops and phases retrieved by the neural network with (a) SELU activation function and (b) LRELU activation function.

Recently, a new activation function, Gaussian error linear unit (GELU), was proposed, which obtained better results than the ELU function in the learning of MNIST's (Mixed National Institute of Standards and Technology database) handwritten digital database. Since our neural network is based on Google's TensorFlow, the version is 1.13, and the GELU function only supports the latest version 2.0. We cannot currently verify the effect of this activation function in our training. In future works, we will compare more activation functions.

In order to use the neural network trained with theoretical data to retrieve the experimentally measured TG-FROG traces, the following issues still need to be addressed. First, it is related to the central wavelength of the theoretical pulse we generate. In the forward calculation, we set the wavelength of the carrier wave at 750 nm, and the actual carrier wave of a real pulse may not be at this wavelength. Second, when we calculate and generate theoretical traces, we use a Gaussian function

spectrum, while the actual measured spectrum in experiments is not a Gaussian function, which affects the structure of the trace. Third, in the experimental measurement, other nonlinear effects such as self-phase modulation (SPM) in the nonlinear crystal also change the characteristics of the trace. Fourth, different CNN parameters and activation functions still need to be explored for experimental traces.

5. Conclusion and perspectives

In conclusion, we train a CNN to reconstruct the time-domain envelop and phase of ultrashort laser pulses from TG-FROG traces. We use a forward program to generate numerically time domain pulse envelops, phases, and FROG traces of femtosecond pulses to train CNN under supervised training. Through the application to similarly generated FROG traces outside the training set, it is proved that our method based on CNN is feasible. We compare the prediction results of several activation functions, and the results show that the appropriate activation function improves the correctness of the neural network training. In our case, the ELU function performs better than SELU and LRELU. Finally, we point out several issues that need to be addressed for neural networks trained with theoretical data to predict experimental data. Currently this method has been proved to be able to achieve envelop and phase retrieval of theoretically calculated FROG traces. In order to retrieve traces measured in experiments, a large number of experiments and theoretical data close to the experimental conditions are necessary first. Issues of noise processing and further optimization of the neural network also need to be addressed.

References

[1] Jimenez R, Fleming G R, Kumar P V and Maroncelli M 1994 *Nature* **369** 471
[2] Hentschel M, Kienberger R, Spielmann C, Reider G A, Milosevic N, Brabec T, Corkum P, Heinzmann U, Drescher M and Krausz F 2001 *Nature* **414** 509
[3] Li J, Ren X, Yin Y, Zhao K, Chew A, Cheng Y, Cunningham E, Wang Y, Hu S, Wu Y, Chini M and Chang Z 2017 *Nat. Commun.* **8** 186
[4] Reid D T, Padgett M, Mcgowan C, Sleat W E and Sibbett W 1997 *Opt. Lett.* **22** 233
[5] Kane D J and Trebino R 1993 *IEEE J. Quantum Electron* **29** 571
[6] Trebino R, Delong K W, Fittinghoff D N, Sweetser J N, Krumbügel M A, Richman B A and Kane D J 1997 *Rev. Sci. Instrum.* **68** 3277
[7] Zahavy T, Dikopoltsev A, Moss D, Haham G I, Cohen O, Mannor S and Segev M 2018 *Optica* **5** 666
[8] White J and Chang Z 2019 *Opt. Express* **27** 4799
[9] Krizhevsky A, Sutskever I and Hinton G E 2017 *Commun. ACM* **60** 84
[10] Lecun Y, Bengio Y and Hinton G 2015 *Nature* **521** 436
[11] Sweetser J N, Fittinghoff D N and Trebino R 1997 *Opt. Lett.* **22** 519
[12] Eichler H J, Gunther P and Pohl D W 1986 *Laser-Induced Dynamic Gratings* (Berlin: Springer-Verlag)