https://doi.org/10.15407/ujpe70.8.550

YU.V. PUSTOVIT, YE.P. LYTVENIUK, YE.D. LIMAREV

Taras Shevchenko National University of Kyiv
(64/13, Volodymyrska Str., Kyiv 01601, Ukraine; e-mail: jura.pustvit@gmail.com)

U-NET BASED METHOD FOR ARPES SPECTRA PROCESSING

Angle-resolved photoemission spectroscopy (ARPES) is a powerful tool for investigating the electronic structure of materials. However, resolving the electronic dispersion from ARPES spectra can be challenging due to the broadening effects, presence of different types of noise, resolution limitations, etc. This paper proposes a new approach for determining dispersion from ARPES spectra based on the U-Net neural network. The energy band extraction problem is regarded as the semantic segmentation task. We will show that the U-Net trained only with generated data can determine band structure from the experimentally obtained spectra, without prior denoising.

K e y w o r d s: ARPES, neural networks, electronic structure, U-Net.

1. Introduction

The importance of machine learning and neural networks in physics is rapidly increasing. These technologies prove their ability to tackle complex problems that were previously considered intractable. The areas, where they impact significantly are quite vast: from astronomy [1] to high-energy physics [2]. The problems solved using these approaches also differ dramatically: from simple image denoising to identification of the Higgs boson [1, 2].

Angle-resolved photoemission spectroscopy is one of the experimental techniques that, in recent years, has developed many upgrades such as time-resolved ARPES, spin-resolved APRES, spatially resolved ARPES *etc.* [3, 4]. However, the classical methods, such as EDC-, MDC-analysis [5, 6], or different image processing methods (the second-derivative [7], curvature [8], minimum gradient [9]) are not feasible for processing large amounts of data with complicated dimensionality. So, nowadays, to cope with more complex experimental data different machine learning methods have been proposed. Such methods give the ability to denoise images using various approaches (based on the convolutional neural network [10], training set-free methods [11, 12]), extract (visualize) spectra features [13] or extract the features and denoise simultaneously using the autoencoder [14].

In this paper, we propose a novel approach for electronic band dispersion determination from the experimentally obtained ARPES spectra. The problem of the electronic dispersion determination is regarded as the semantic segmentation task. The goal is to categorize each image pixel into a particular class. The proposed method utilizes a U-Net neural network with an attention block to discriminate between 'background' and electronic dispersion pixels. As has been shown, such a network can handle experimentally obtained spectra without preprocessing, even though only generated data are used for the training and validation.

2. Method

ARPES enables the direct observation of the electronic band structure [4]. However, the band structure can be distorted and broadened due to various factors, including electron-electron, electron-phonon interactions, and crystal impurities. These factors can cause energy shifts and peak broadening. In addition, the resulting spectra are affected by noises (shot, Shirley background) and resolution limitations. To extract the pure electronic dispersion, several methods are used. They include analytic methods such as the MDC-,EDC- methods and different image pro-

C i t a t i o n: Pustovit Yu.V., Lytveniuk Ye.P., Limarev Ye.D. U-Net based method for ARPES spectra processing. *Ukr. J. Phys.* **70**, No. 8, 550 (2025). https://doi.org/10.15407/ujpe70.8.550.

cessing techniques such as second derivative, curvature, and minimum gradient methods.

Recently, there were proposed different neural network-based methods for electronic dispersion determination [13, 14].

For example, the determination of the electronic band dispersion from noisy and broadened spectra is approached by solving an inverse transformation problem [13]. The idea is to use a convolutional neural network to find the appropriate inverse transformation for the spectrum. However, despite the tendency of the proposed method to outperform the aforementioned traditional methods, this approach has some disadvantages. First, the image should be preprocessed (denoised) to determine the dispersion, so, certain details can be lost. Second, CNN considers the connections between the neighboring pixels but does not consider long-range dependencies. The limited receptive field of CNN (each output pixel is only influenced by its 13-by-13 neighborhood) may not capture all the crucial relationships between distant pixels that influence electronic dispersion [13].

Our approach treats the task of electronic dispersion determination as a segmentation task. The spectrum is regarded as a segmentation map, where each pixel is classified into one of two classes. The first class represents pixels belonging to electronic dispersion. The second class encompasses background pixels associated with broadening, noise and other artifacts. This task is similar to the background removal task. In our case, the "background" includes all pixels that are not directly related to the electronic dispersion.

2.1. U-Net with Attention

One of the most frequently used neural-network-based approaches for semantic segmentation is the U-Net neural network, and its variations [15]. U-Net was originally proposed for biomedical image segmentation. To visualize the positions of the electronic energy bands, we employed a modified Attention U-Net network[16]. This network incorporates an additional attention block that captures a sufficiently large receptive field, enabling the consideration of long-range dependencies without significantly complicating the U-Net model [16].

The U-Net architecture is composed of three main parts: the encoder (or contracting path), the symmetric decoder (or expansive path), and a "latent" layer situated between the encoder and decoder. The core building block of this neural network involves the repeated application of two unpadded 3×3 convolutions, each followed by batch normalization and a ReLU activation.

The encoder consists of four blocks, each followed by a max-pooling layer for downsampling. The number of kernels in the convolutional layers follows the pattern of $64 \times n$, where n represents the block number.

The latent layer applies two unpadded convolutional layers with 64×16 kernels, each followed by batch normalization and ReLU activation.

The decoder mirrors the encoder with four blocks. At each step in the decoder, an attention gate processes the corresponding feature map from the encoder. The output of the attention gate is concatenated with the upsampled features (obtained using 2×2 upsampling) and further processed by the decoder block. The number of kernels in the convolutional layers follows the pattern of $64 \times 8/2^{n-1}$, where n represents the decoder's block number.

The final layer 1×1 convolution with 2 kernels followed by the batch normalization and sigmoid activation. It is used to map the decoder's output to the classes segmentation map.

Image pixels associated with electronic band dispersion occupy only a small portion of the image, so there is a significant imbalance in pixels belonging to different classes. To address the issue of class imbalance, focal loss was used as the loss function[17]. Its main advantage is the division of samples into easy and hard. The easy sample is already well-classified (the predicted probability is in a vicinity of 1), so, its contribution can significantly be decreased or downweighted. In this case, easy samples contribution to the total loss function is small even if their number is large. The model focuses on training a subset of hard examples, for which the predicted probability is much lesser than 1. [17] This allows one to avoid situations when the model achieves nearly 100% accuracy by simply assigning everything to the class that occurs most frequently in the image.

2.2. Training process

A major obstacle in the training process is the lack of available training data derived from experimental spectra. This is primarily attributed to the inherent difficulty of accurately determining dispersion, a task

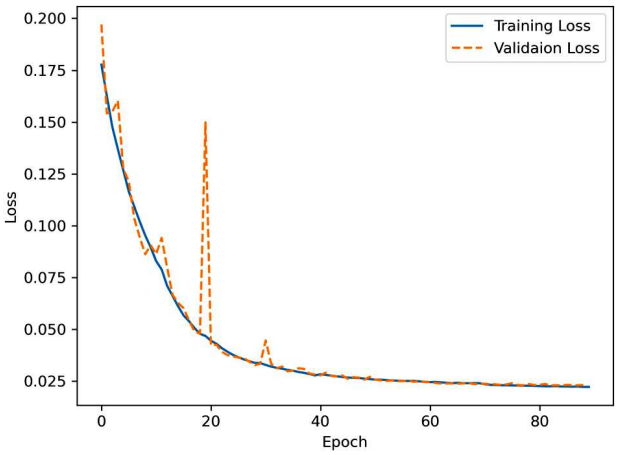

Fig. 1. Loss function during training process

Parameters values

Parameter	Min	Max
α	0	1.5
λ	2.1	8
l	–0.3	1.3
m	1	20
Imp	0	1

that cannot be automated and is not accurate enough if performed by "human guess" [13]. Therefore, the generated dataset has been used, where the position of the electronic dispersion can be precisely determined [10, 13, 14]. Typically, the training set used for neural network training for APRES spectra processing is quite small (less than 10,000 examples) [10, 13, 14].

The generated data set was used to train the proposed U-Net neural network. The training data set of the neural network consists of pairs of images: as the input – the generated spectra, which can contain up to 2 zones, as the output corresponding labeled image (segmentation map), where each pixel within an image has a specific label. The dimensionality of the input image is $128 \times 128 \times 1$ and the output is $128 \times 128 \times 2$. Each pixel of the output is labeled using a one-hot encoded vector $\binom{0}{1}$ or $\binom{1}{0}$ (electronic band dispersion pixel). It uses a binary vector, where only one element is set to 1, indicating whether the pixel belongs to the electronic dispersion or background class. A one-particle spectral function, the imaginary part of Green's function for one-electron excitations (quasiparticles), has been used to simulate the detected experimental one-band spectra:

$$A(\omega, k) = \frac{\Sigma''(\omega)}{(\omega - \varepsilon(k) - \Sigma'(\omega))^2 + \Sigma''(\omega)^2},$$

where $\Sigma'(\omega)$, $\Sigma''(\omega)$ – are the real and the imaginary parts of the quasiparticle self-energy, respectively, reflects all the interaction of electrons in the crystal, $\varepsilon(k) = mk^2 + l$ – "bare" electron dispersion(is approximated by simple parabola). $\Sigma'(\omega) = \lambda\omega$ and $\Sigma''(\omega) = \alpha\omega^2 + \text{Imp}$, where constant parameter Imp is added to consider the scattering on impurities. The values of the parameters have been randomly generated from the range of values presented in Table.

Due to the spectrum's resolution exceeding $128 \times \times 128$ pixels, it will be divided into smaller, more manageable sub-regions during analysis. Compared to the entire spectrum, the relatively small size of these sub-regions implies a low probability of any given sub-region containing more than two electronic bands. So, we employed training data that included spectra containing up to 2 zones. The resulting two-band spectra were obtained by adding two one-band spectra.

Since the experiment data are always noisy, two types of noise have been added to the simulated spectra namely:

• Shirley noise to describe the background related to inelastically scattered electrons ($\beta\omega^2$, where β in the range from 0 to 1.5).

• Poisson noise multiplying each generated image pixel by uniformly distributed random value (from 0 to 0.5) and adding to the image.

The generated spectra were multiplied by the Fermi–Dirac distribution with temperatures ranging from 0 to 300 Kelvin to simulate the effects related to the Fermi level.

To make the training process of the neural network less time and memory-consuming we trained the neural network in several phases

1. Initial Training (20 Epochs). The initial training phase utilizes a dataset of 1000 examples (training set 900 examples, validation 100 examples).

2. The second phase introduces an expansion of the dataset. The initial 1000 examples are replaced with a new dataset of 1500 examples (1350 for training and 150 for validation).

3. New dataset consists of another 1500 examples (1350 for training and 150 for validation) and was used during 30 epochs.

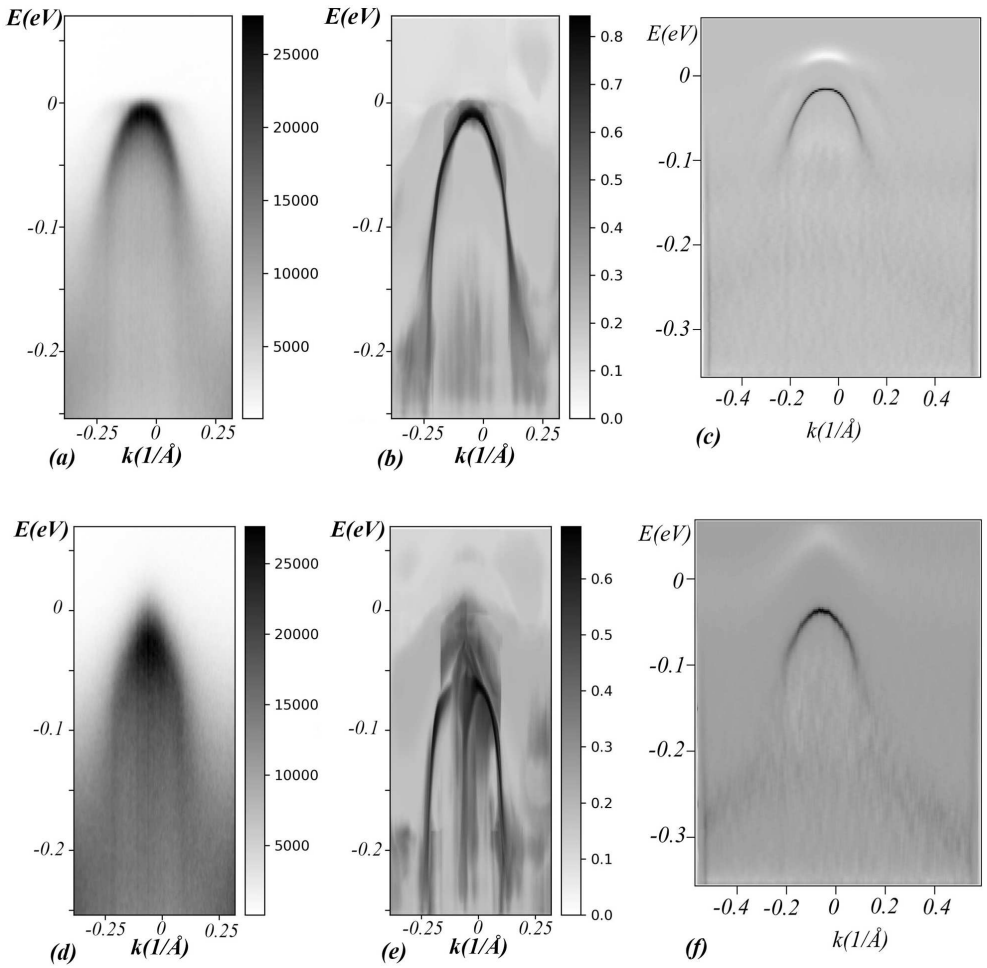

Fig. 2. The ARPES spectra of the Brillouin zone Z point of Fe(Se, Te) obtained using linear vertical polarization for temperature 20 K (*a*) and 220 K (*d*) and the corresponding results of its procession using our U-Net (*b, e*) and the curvature method (*c, f*)

4. Final Dataset Expansion and Training Completion (20 Epochs). The training duration is reduced to 20 epochs and the dataset is extended to 2,000 examples (1,800 for training and 200 for validation).

That dataset expansions and changes expose the network to a wider range of examples, enhancing its generalization ability and preventing the neural network from overfitting. This ensures that the network has sufficient time to learn from the new dataset and maintain its overall performance.

The Adam optimizer is adopted to train the network for 90 epochs. Figure 1 shows that the value of the loss function for training and the validation set decreases monotonically and becomes almost constant in the last 10 epochs, indicating good convergence. In particular, the values of the loss functions for both sets are very similar.

3. Results

To compare the method's performance and reliability, it has been applied to the experimentally obtained spectra of the Brillouin zone Z point of Fe(Se, Te). The spectra were obtained at various temperatures (from 20 K to 220 K), allowing us to track the network's performance under different noise levels. Additionally, the spectra encompass both single and double-zone configurations. Since the image size is larger than 128 by 128 pixels, model predictions were calculated for all possible 128 by 128 image patches. The pixel with the highest value (the high-

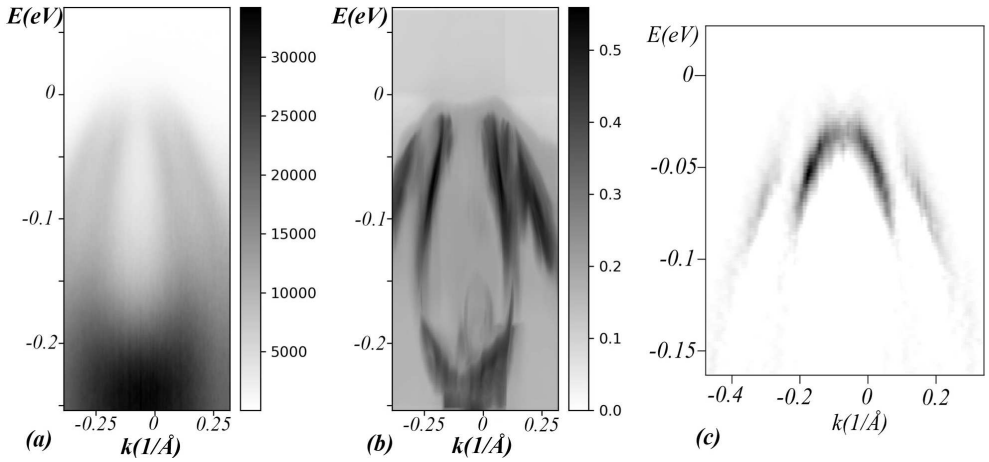

Fig. 3. The ARPES spectrum of the Brillouin zone Z point of Fe(Se, Te) obtained using linear horizontal polarization for temperature 160 K (*a*) and the corresponding results of its procession using our U-Net (*b*) and the curvature method (*c*)

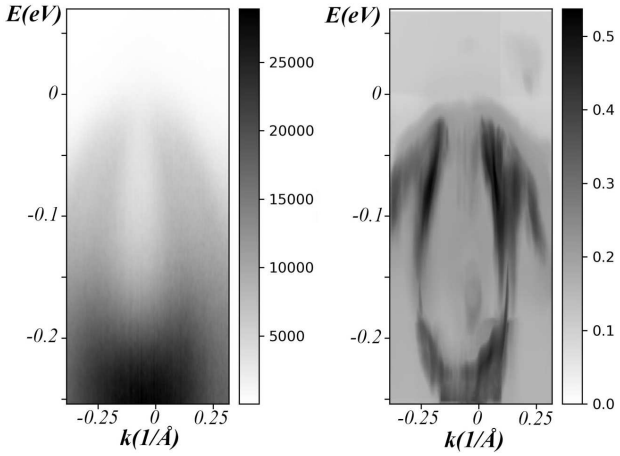

Fig. 4. The ARPES spectrum of the Brillouin zone Z point of Fe(Se, Te) obtained using linear horizontal polarization for temperature 220 K results (left image) and the results of its procession using our U-Net (right)

est of all possible probabilities that the pixel belongs to electronic dispersion) was selected for overlapping patches. This approach helps considering all possible long-range dependencies between image pixels available for such neural network resolution.

We compared our method's results to those obtained using the curvature method from [18] to evaluate its performance. However, in [18] each spectrum was pre-processed by smoothing to reduce noise before applying the curvature method. The curvature method's results were optimized by tuning the parameter C_0 to achieve the best result [8]. As shown in Fig. 2, our method enables the visualization of dispersion across a broader energy range, showing greater resilience to the noise introduced by inelastically scattered electrons. By considering long-range pixel dependencies, the network can capture the band-like nature of electronic dispersion. As illustrated in Fig. 2, *e*, the network extends the dispersion pattern based on previously predicted pixels. Unlike the curvature and most other methods, our neural network approach directly processes the raw spectra without requiring a prior smoothing.

The ability of neural networks to bypass smoothing is essential, when dealing with spectra containing multiple, adjacent zones. In such cases, smoothing often leads to a loss of information, so, it is quite hard to find trade-offs between noise reduction and zone information loss. Especially for high temperatures, the signal becomes weaker compared to the noise, making it more difficult to extract or interpret the desired information from spectra. For example, for spectra obtained at 160 K, the curvature method reveals only a single discernible zone (Fig. 3, *c*), while the other zone is barely detectable. At 220 K, the curvature method fails to extract meaningful information about the zone positions. Unlike the curvature method, our proposed neural network effectively visualizes bands even in complex scenarios. As seen in Figs. 3, and 4, the neural network enables visualization of the bands. High noise levels significantly hinder the model's ability to classify pixels accurately. This is evident in the drop in prediction ac-

curacy, where the probability of a pixel belonging to electronic dispersion barely exceeds 0.5. In contrast, lower temperatures typically yield probabilities closer to 0.8, indicating a much clearer distinction between dispersion and non-dispersion regions.

4. Conclusions

The paper proposes to regard the problem of determining electronic dispersion from ARPES spectra as a semantic segmentation task. We trained the U-Net with an attention mechanism to solve this problem, using a small dataset of generated spectra and corresponding segmentation maps (2,000 examples). Processing experimentally obtained spectra of Fe(Se, Te) (with different radiation polarizations, temperature interval 20–200 K) demonstrated the effectiveness of the proposed network. Compared to the curvature method, the proposed network is more robust to noise and, thus, achieves better results in complicated situations (multiple bands close to each other, high noise levels). Additionally, since no spectra preprocessing is required, the U-Net-based method can be used for the automated processing of large volumes of data.

1. M.J. Smith, J.E. Geach. Astronomia ex machina: A history, primer and outlook on neural networks in astronomy *Royal Soc. Open Sci.* **10** (5), 221454 (2023).
2. C. Adam-Bourdarios, G. Cowan, C. Germain-Renaud, I. Guyon, B. Kégl, D. Rousseau. The higgs machine learning challenge. *J. Phys. Conf. Ser.* **664**, 072015 (2015).
3. A. Damascelli, Z. Hussain, Z.-X. Shen. Angle-resolved photoemission studies of the cuprate superconductors. *Rev. Mod. Phys.* **75**, 473 (2003).
4. J.A. Sobota, Yu. He, Z.-X. Shen. Angle-resolved photoemission studies of quantum materials. *Rev. Mod. Phys.* **93**, 025006 (2021).
5. T. Valla, A.V. Fedorov, P.D. Johnson, B.O. Wells, S.L. Hulbert, Q. Li, G.D. Gu, N. Koshizuka. Evidence for quantum critical behavior in the optimally doped cuprate. *Science* **285**, 2110 (1999).
6. Z.-X. Shen, J.R. Schrieffer. Momentum, temperature, and doping dependence of photoemission lineshape and implications for the nature of the pairing potential in high-T_c superconducting materials. *Phys. Rev. Lett.* **78**, 1771 (1997).
7. R.C. Gonzalez, R.E. Woods. *Digital Image Processing* (Prentice Hall, 2008) [ISBN: 978-0131687288].
8. P. Zhang, P. Richard, T. Qian, Y.-M. Xu, X. Dai, H. Ding. A precise method for visualizing dispersive features in image plots. *Rev. Sci. Instrum.* **82**, 043712 (2011).
9. Yu. He, Y. Wang, Z.-X. Shen. Visualizing dispersive features in 2D image via minimum gradient method. *Rev. Sci. Instrum.* **88**, 073903 (2017).
10. Yo. Kim, D. Oh, S. Huh, D. Song, S. Jeong, Ju. Kwon, M. Kim, D. Kim, H. Ryu, J. Jung, W.Kyung, B. Sohn, S. Lee, J. Hyun, Ye. Lee *et al.* Deep learning-based statistical noise reduction for multidimensional spectral data. *Rev. Sci. Instrum.* **92**, 073901 (2021).
11. D. Huang, J. Liu, T. Qian, Y.F. Yang. Spectroscopic data de-noising via training-set-free deep learning method. *Sci. China: Phys. Mech. Astron.* **66**, 267011 (2023).
12. J. Liu, D. Huang, Y.F. Yang, T. Qian. Removing grid structure in angle-resolved photoemission spectra via deep learning method. *Phys. Rev. B* **107**, 165106 (2023).
13. H. Peng, X. Gao, Yu He, Y. Li, Y. Ji, Ch. Liu, S.A. Ekahana, D. Pei, Z. Liu, Z. Shen, Yu. Chen. Super resolution convolutional neural network for feature extraction in spectroscopic data. *Rev. Sci. Instrum.* **91**, 033905 (2020).
14. F. Restrepo, Ju. Zhao, U. Chatterjee. Denoising and feature extraction in photoemission spectra with variational auto-encoder neural networks. *Rev. Sci. Instrum.* **93**, 065106 (2022).
15. O. Ronneberger, P. Fischer, T. Brox. U-Net: Convolutional networks for biomedical image segmentation. arXiv: 1505.04597[cs.CV] (2015).
16. O. Oktay, J. Schlemper, L. Le Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y Hammerla, B. Kainz, B. Glocker, D. Rueckert. Attention U-Net: Learning where to look for the pancreas. arXiv: 1804.03999v3[cs.CV] (2018).
17. T. Lin, P. Goyal, R. Girshick, K. He, P. Dollar. Focal loss for dense object detection. *2017 IEEE International Conference on Computer Vision (ICCV)* (2017), p. 2999.
18. Yu. V. Pustovit, A.A. Kordyuk. Temperature induced shift of electronic band structure in Fe(Se, Te). *Low Temp. Phys.* **45**, 1172 (2019).

Received 05.06.24

Ю.В. Пустовіт, Є.П. Литвенюк, Є.Д. Лімарєв

МЕТОД ОБРОБКИ ARPES
СПЕКТРІВ НА ОСНОВІ U-NET

Фотоемісійна спектроскопія з розділенням по куту (ARPES) – це потужний інструмент для дослідження електронної структури матеріалів. Однак визначення електронної дисперсії з цих спектрів може бути ускладненим через ефекти уширення, наявність різних типів шуму, обмеження роздільної здатності тощо. У цій статті пропонується новий підхід до визначення дисперсії зі спектрів ARPES на основі нейронної мережі U-Net. Задача виділення енергетичних зон розглядається як задача семантичної сегментації. Ми показуємо, що U-Net, навчена лише на основі згенерованих даних, може визначити зонну структуру з експериментально отриманих спектрів без попереднього знешумлення.

Ключові слова: ARPES, нейронні мережі, електронна структура, U-Net.